Self-calibration approach for optical long-baseline interferometry imaging

Serge Meimon,[1,*] Laurent M. Mugnier,[1] and Guy Le Besnerais[2]

[1]*Département d'Optique Théorique et Appliquée, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon cedex, France*

[2]*Département Traitement de l'Information et Modélisation, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon cedex, France*

**Corresponding author: meimon@onera.fr*

Received May 28, 2008; revised October 14, 2008; accepted October 17, 2008; posted October 23, 2008 (Doc. ID 96631); published December 19, 2008

Current optical interferometers are affected by unknown turbulent phases on each telescope. In the field of radio interferometry, the self-calibration technique is a powerful tool to process interferometric data with missing phase information. This paper intends to revisit the application of self-calibration to optical long-baseline interferometry (OLBI). We cast rigorously the OLBI data processing problem into the self-calibration framework and demonstrate the efficiency of the method on a real astronomical OLBI data set. © 2008 Optical Society of America

OCIS codes: 120.3180, 100.3020, 100.3190.

1. INTRODUCTION

Optical long-baseline interferometry (OLBI) aims to combine light collected by widely separated telescopes to access angular resolutions beyond the diffration limit of each individual aperture. Long-baseline interferometers measure a discrete set of spatial frequencies of the observed object, or Fourier data. Due to instrumental complexity, current interferometers recombine only a few telescopes, and even several nights of observation lead to a very limited number of Fourier data; moreover, due to the atmospheric turbulence, it is very difficult to get reliable phase information from ground-based interferometry [1]. Hence OLBI has to deal with severe underdetermination and missing phase information.

The classical answer to underdetermination is to use a parametric approach, i.e., to search for an object entirely described by a small set of parameters (for instance, a circular object with a parametric attenuation profile). With a "good model," such an approach allows a reliable and precise estimation of astrophysical parameters. A good model should limit as much as possible the number of free parameters, while allowing a description of all the object's features, because parametric inversion cannot reveal unguessed features. The χ^2 fit is often used as a model quality diagnosis, since an inadequate model will often result in a poor fit to the data, thus revealing that a new model (with more parameters or different parameters) is needed. However, it does not reveal which new model must be adopted.

As progress in instrumental issues gives access to better frequency coverage, i.e., to potentially finer descriptions of the object, the choice of the model becomes more difficult. An alternate and complementary approach is then nonparametric reconstruction, which we will call "optical long-baseline interferometric imaging" (OLBII). Imaging means that the object is described by a large set of parameters, such as coefficients of the object's decomposition in some spatial functional basis, while underdetermination is tackled by regularization tools. Imaging is useful to understand the structure of a complex object when prior information is limited.

From the beginning, OLBII has been influenced by the remarkable techniques developed in radio interferometry with very large baselines (VLBI) [2]. For instance, the "WIPE" OLBII technique of Lannes *et al.* [3] is inspired by the well-known CLEAN method [4]. As regards the missing phase problem, the self-calibration technique proposed in radio interferometry by Cornwell and Wilkinson [5] underlies recent work in OLBII [6].

This paper intends to revisit the application of self-calibration to OLBI. Our contribution is threefold:

1. We cast rigorously the OLBI data processing problem into the self-calibration framework, with consideration of the second-order statistics of the noise.

2. We propose WISARD (for Weak-phase Interferometric Sample Alternating Reconstruction Device), a self-calibration algorithm dedicated to OLBII, which uses the proposed data model within a Bayesian regularization approach.

3. We demonstrate the efficiency of WISARD on a real astronomical OLBI data set.

The paper is organized as follows: Section 2 describes the observation model of OLBI, briefly presents a Bayesian approach, and discusses the main problems that are encountered because of the incomplete OLBI data. Section 3 is devoted to the derivation of a specific myopic model, which achieves a good approximation of the data model and leads to self-calibration techniques. One such technique, WISARD, is proposed in Section 4. Results of WISARD on simulated and real astronomical data sets are presented in Section 5. Our conclusions are given in Section 6. Most mathematical derivations are gathered in the appendixes.

2. REALISTIC OBSERVABLES IN OPTICAL LONG-BASELINE INTERFEROMETRY

A. Ideal Interferometric Data

Here we describe the ideal data, i.e., without aberrations, noise, or turbulence effects, produced by an N_{t}-telescope interferometer observing a monochromatic source with wavelength λ. The brightness distribution of the source is denoted $x(\boldsymbol{\xi})$, $\boldsymbol{\xi}$ being angular coordinates on the sky. Individual telescopes T_k of the interferometer are located at three-space positions $\vec{OT_k}$, and we denote $\boldsymbol{r}_k(t)$ the projection of $\vec{OT_k}$ onto $\mathcal{P}$, the plane normal to the pointing direction. Because of the Earth's rotation, the pointing direction changes during an observing night, so these projected vectors are time dependent.

Each pair (T_k,T_l) of telescopes yields a fringe pattern with a 2D spatial frequency $\boldsymbol{\nu}_{kl}(t)\triangleq\frac{\boldsymbol{u}_{kl}(t)}{\lambda}$, where $\boldsymbol{u}_{kl}(t)$ is the baseline

$$\boldsymbol{u}_{kl}(t)\triangleq\boldsymbol{r}_l(t)-\boldsymbol{r}_k(t), \tag{1}$$

that is, the projection of the vector $\vec{T_kT_l}$ onto $\mathcal{P}$.

Measuring the position and contrast of these fringes yields a phase $\phi_{\mathrm{data}}^{kl}(t)$ and an amplitude $a_{\mathrm{data}}^{kl}(t)$, which can be grouped together in a complex visibility:

$$y_{kl}^{\mathrm{data}}(t)\triangleq a_{kl}^{\mathrm{data}}(t)\mathrm{e}^{\mathrm{i}\phi_{kl}^{\mathrm{data}}(t)}. \tag{2}$$

According to the Van Cittert–Zernike theorem [7], complex visibilities are *ideally* linked to the normalized Fourier transform (FT) of $x(\boldsymbol{\xi})$ at the 2D spatial frequency $\boldsymbol{\nu}_{kl}(t)$ through

$$y_{kl}^{\mathrm{data}}(t)=\eta_{kl}(t)\frac{\mathrm{FT}[x(\boldsymbol{\xi})](\boldsymbol{\nu}_{kl}(t))}{\mathrm{FT}[x(\boldsymbol{\xi})](\boldsymbol{0})}. \tag{3}$$

The instrumental visibility $\eta_{kl}(t)$ accounts from the many potential sources of visibility loss: residual perturbations of the wavefront at each telescope, differential tilts between telescopes, differential polarization effects, nonzero spectral width, etc. In practice, the instrumental visibility is calibrated on a star reputed to be unresolved by the interferometer before the object of interest is observed and is compensated for in the preprocessing of the raw data. Thanks to this calibration step, we replace $\eta_{kl}(t)$ with 1 in Eq. (3).

For the sake of clarity, we consider a *complete* N_{t}-telescope array in what follows, i.e., one in which all the possible two-telescope baselines can be formed simultaneously, and a nonredundant interferometer configuration, where each baseline provides a different spatial frequency. Extension to incomplete and redundant settings is straightforward. Thus, at each time t, there are

$$N_b=\binom{N_{\mathrm{t}}}{2}=\frac{N_{\mathrm{t}}(N_{\mathrm{t}}-1)}{2} \tag{4}$$

complex observation equations such as Eq. (3).

Let us briefly introduce the discretized observation model. The sought brightness distribution x is represented by the coefficients $\boldsymbol{x}$ of its projection onto some convenient spatial basis (box functions, sinc's, wavelets, prolate spheroidal functions, etc). The normalized discrete-continuous Fourier matrix $\boldsymbol{H}(t)$ maps the chosen discrete spatial representation into the real-valued instantaneous frequency coverage $\{\boldsymbol{\nu}_{kl}(t)\}_{1\leq k<l\leq N_{\mathrm{t}}}$, and we further define

$$\begin{cases}\boldsymbol{a}(\boldsymbol{x},t)\triangleq|\boldsymbol{H}(t)\boldsymbol{x}|,\\ \boldsymbol{\phi}(\boldsymbol{x},t)\triangleq\arg\{\boldsymbol{H}(t)\boldsymbol{x}\}.\end{cases} \tag{5}$$

B. Effect of Atmospheric Turbulence on Short-Exposure Measurements

At optical wavelengths, atmospheric turbulence affects phase measurements through path-length fluctuations. The statistics of these fluctuations can be described by a time-scale parameter, the coherence time τ_0, typically around 10 ms, and by a space-scale parameter, the Fried parameter r_0 [[8]]. We assume that the diameter of the elementary apertures is small relative to the Fried parameter or that each telescope is corrected from the effects of turbulence by adaptive optics. The remaining turbulent effects on the interferometric measurements can be seen as a delay line between the two telescopes T_k and T_l, which affects short-exposure phase measurements through an additive differential piston $\varphi_l(t)-\varphi_k(t)$:

$$\phi_{kl}^{\mathrm{data}}(t)=\phi_{kl}(\boldsymbol{x},t)+\varphi_l(t)-\varphi_k(t)+\mathrm{noise}[2\pi] \tag{6}$$

or, in a matrix formulation:

$$\boldsymbol{\phi}^{\mathrm{data}}(t)=\boldsymbol{\phi}(\boldsymbol{x},t)+\boldsymbol{B}\boldsymbol{\varphi}(t)+\mathrm{noise}[2\pi], \tag{7}$$

where $N_b\times N_{\mathrm{t}}$ operator $\boldsymbol{B}$, called the baseline operator, is defined in Appendix A.

Because the differential pistons are zero mean, one might think that the object phase $\boldsymbol{\phi}(\boldsymbol{x},t)$ could be recovered from Eq. (7) by averaging over many realizations of the atmosphere. However, for a long baseline relative to the Fried parameter, the optical path difference between apertures introduced by turbulence may be very much greater than the observation wavelength and thus lead to random pistons much larger than 2π. The 2π-wrapped perturbation that affects phase (7) is then practically uniformly distributed in $[0,2\pi]$. In consequence, averaging the short-exposure phase measurements (7) does not improve the signal-to-noise ratio (SNR).

In phase referencing techniques (see [9]), the turbulent pistons are measured in order to subtract them in Eq. (7). However powerful and promising, these methods require specific hardware and are not feasible for all sources. The only other way to obtain exploitable long-exposure data then is to form piston-free short-exposure observables *before* the averaging.

C. Piston-Free Short-Exposure Observables

Piston-free short-exposure phase observables are quantities $f(\boldsymbol{\phi}^{\mathrm{data}}(t))$ in which the turbulent term $\boldsymbol{B}\boldsymbol{\varphi}(t)$ cancels out:

$$f(\phi^{\text{data}}(t)) = f(\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{B}\boldsymbol{\varphi}(t)) = f(\boldsymbol{\phi}(\boldsymbol{x},t)). \tag{8}$$

For an interferometric array of three telescopes or more, the closure phases [10] are one famous example, in which f is a linear operator performing triplewise summation of the phases. For any set of three telescopes (T_k, T_l, T_m), the short-exposure visibility phase data are

$$\begin{cases} \phi_{kl}^{\text{data}}(t) = \phi_{kl}(\boldsymbol{x},t) + \varphi_l(t) - \varphi_k(t) + \text{noise}[2\pi], \\ \phi_{lm}^{\text{data}}(t) = \phi_{lm}(\boldsymbol{x},t) + \varphi_m(t) - \varphi_l(t) + \text{noise}[2\pi], \\ \phi_{mk}^{\text{data}}(t) = \phi_{mk}(\boldsymbol{x},t) + \varphi_k(t) - \varphi_m(t) + \text{noise}[2\pi], \end{cases} \tag{9}$$

and the turbulent pistons cancel out in the closure phase defined by

$$\begin{aligned} \beta_{klm}^{\text{data}}(t) &\triangleq \phi_{kl}^{\text{data}}(t) + \phi_{lm}^{\text{data}}(t) + \phi_{mk}^{\text{data}}(t) + \text{noise}[2\pi] \\ &= \phi_{kl}(\boldsymbol{x},t) + \phi_{lm}(\boldsymbol{x},t) + \phi_{mk}(\boldsymbol{x},t) + \text{noise}[2\pi] \\ &\triangleq \beta_{klm}(\boldsymbol{x},t) + \text{noise}[2\pi]. \end{aligned} \tag{10}$$

We have the following properties:

- The set of all three-telescope closure phases that can be formed using a complete array is generated by the $(N_{\text{t}}-1)(N_{\text{t}}-2)/2$ closure phases $\beta_{1kl}^{\text{data}}(t)$, $k<l$, i.e., the closure phase that includes telescope T_1 (indeed, $\beta_{klm}^{\text{data}} = \beta_{1kl}^{\text{data}} + \beta_{1lm}^{\text{data}} - \beta_{1km}^{\text{data}}$). In what follows, these canonical closure phases are grouped together in a vector $\boldsymbol{\beta}^{\text{data}}$, and $\boldsymbol{C}$ denotes the linear closure operator such that $\boldsymbol{C}\boldsymbol{\phi}^{\text{data}} = \boldsymbol{\beta}^{\text{data}}$ (see Appendix A).
- If f is a continuous differentiable function verifying property (8), then

$$f(\boldsymbol{\phi}) = g(\boldsymbol{C}\boldsymbol{\phi}), \tag{11}$$

where g is some continuous differentiable function. In other terms, there is essentially *no operator other than the closure operator* that cancels out the effect of turbulence on short-exposure visibility phases (this property holds only in the monochromatic case).

The proof of the second property is given in Appendix B.

D. Long-Exposure Observables Data Model

To minimize the effect of noise, one is led to average short-exposure measurements into long-exposure observables, chosen so that they are asymptotically unbiased. The averaging time must be short enough with respect to the Earth's rotation so that the baseline does not change, and long enough to reach an acceptable SNR. The averaged quantities are generally these:

- averaged *squared amplitudes* $\boldsymbol{s}^{\text{data}}(t) = \langle \boldsymbol{a}^{\text{data}}(t+\tau)^2 \rangle_\tau$,
- averaged *bispectra* $\boldsymbol{V}_{1kl}^{\text{data}}(t) = \langle y_{1k}^{\text{data}}(t+\tau) \cdot y_{kl}^{\text{data}}(t+\tau) \cdot y_{l1}^{\text{data}}(t+\tau) \rangle_\tau$, $k<l$. Squared amplitudes are preferred to amplitudes because their bias can be estimated and subtracted from the data. Short-exposure bispectra are continuous differentiable functions verifying property (8) and so correspond to a particular choice of g in Eq. (11). In the absence of noise, the averaged bispectrum amplitudes are redundant with the averaged squared amplitudes. Although they should be useful in low-SNR conditions, averaged bispectrum amplitudes are not considered in what follows. The averaged bispectrum phases $\beta_{1kl}^{\text{data}}(t)$, $k<l$ constitute unbiased long-exposure closure phase estimators. As such, they are linked to the object phases $\boldsymbol{\phi}(\boldsymbol{x},t)$ through

$$\boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \text{noise}[2\pi]. \tag{12}$$

It is shown in Appendix A that the kernel of the closure operator $\boldsymbol{C}$ is of dimension $(N_{\text{t}}-1)$. Hence Eq. (12) implies that optical interferometry through turbulence has to deal with partial phase information. This result can also be obtained by counting up phase unknowns for each instant of measurement t: there are $N_{\text{t}}(N_{\text{t}}-1)/2$ unknown object visibility phases and $(N_{\text{t}}-1)(N_{\text{t}}-2)/2$ observable independent closure phases, which results in $(N_{\text{t}}-1)$ missing phase data. As is well known in the radio interferometric community, the greater the number of apertures in the array, the smaller the proportion of missing phase information.

The long-exposure observables considered in this paper are noisy squared amplitudes $\boldsymbol{s}^{\text{data}}(t)$ and closure phases $\boldsymbol{\beta}^{\text{data}}(t)$. The only statistics usually available are the variances for each observable (as, for instance, in the OIFITS data exchange format [11]). The assumed noise distribution is consequently zero-mean white Gaussian:

$$\begin{cases} \boldsymbol{s}^{\text{data}}(t) = \boldsymbol{a}^2(\boldsymbol{x},t) + \boldsymbol{s}^{\text{noise}}(t), & \boldsymbol{s}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{s}(t)}), \\ \boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}^{\text{noise}}(t)[2\pi], & \boldsymbol{\beta}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{\beta}(t)}). \end{cases} \tag{13}$$

The matrices $\boldsymbol{R}_{\boldsymbol{s}(t)}$ and $\boldsymbol{R}_{\boldsymbol{\beta}(t)}$ are diagonal, with variances related to the integration time, although correlations may be produced by the use of the same reference stars in the calibration process [12].

E. Bayesian Reconstruction Methods

This approach first forms the anti-log-likelihood according to model (13):

$$J^{\text{data}}(\boldsymbol{x}) = \sum_t J^{\text{data}}(\boldsymbol{x},t) = \sum_t \chi^2_{\boldsymbol{s}(t)}(\boldsymbol{x}) + \chi^2_{\boldsymbol{\beta}(t)}(\boldsymbol{x}), \tag{14}$$

where $\chi^2_{\boldsymbol{s}(t)}(\boldsymbol{x}) = \boldsymbol{s}^{\text{data}}(t)$ denotes the classical χ^2 statistic $(\boldsymbol{s}^{\text{data}}(t) - \boldsymbol{\alpha}^2(\boldsymbol{x},t))^{\text{T}} \boldsymbol{R}_{\boldsymbol{s}(t)}^{-1} (\boldsymbol{s}^{\text{data}}(t) - \boldsymbol{\alpha}^2(\boldsymbol{x},t))$. Closure terms $\chi^2_{\boldsymbol{\beta}(t)}(\boldsymbol{x})$ are a weighted quadratic distance between complex phasors [13] instead of a χ^2 statistic over closure phase residuals. One then associates $\boldsymbol{J}^{\text{data}}$ with a regularization term to account for the incompleteness of the data in such inverse problems and minimizes the composite criterion

$$J(\boldsymbol{x}) = J^{\text{data}}(\boldsymbol{x}) + J^{\text{prior}}(\boldsymbol{x}) \tag{15}$$

under the following constraints:

$$\forall (p,q), \qquad x(p,q) \geqslant 0,$$

$$\sum_{p,q} x(p,q) = 1. \tag{16}$$

The first requires positivity of the sought object, and the second is a constraint of unit flux. Indeed, fringe visibilities are by definition flux-normalized quantities [i.e., normalized by the FT of the object at the null frequency; see Eq. (3)], so the data are independent of the total flux of the sought object (of course an interferometer is sensitive to the total flux of the source, but this last value is not contained in the fringe visibility itself).

The regularization term J^{prior} is chosen to enforce some properties of the object that are known *a priori* (smoothness, spiky behavior, positivity, etc.) and should also ease the minimization. Simple and popular regularization terms are convex separable penalizations of the object pixels (i.e., white priors) or of the object spatial derivatives (for instance, first-order derivative or gradient). In what follows, we quickly describe the prior terms used in this paper. These priors are more extensively described and compared in [14]. For a general review on regularization, see [15].

Entropic priors belong to the family of white priors and often allow one to obtain a clean image while preserving its sharp spiky features, whereas quadratic penalization tends to soften the reconstructed map. The white quadratic-linear (or $L_2L_1^w$) penalization given by

$$L_2L_1^w(\boldsymbol{x}) = \delta^2 \sum_{p,q} \frac{\boldsymbol{x}(p,q)}{s\delta} - \ln\left(1 + \frac{\boldsymbol{x}(p,q)}{s\delta}\right) \tag{17}$$

that we use in Section 5 leads to a kind of entropic regularization, in the sense of [16]. We propose a nominal setting of the two parameters δ and s:

$$s = 1/N_{\text{pix}}; \qquad \delta = 1. \tag{18}$$

As regards regularization based on the object's spatial derivatives, we shall consider here only quadratic penalization, but convex quadratic-linear L_2L_1 penalization functions could also be invoked.

Reference [17] is one of the works that adopts such a Bayesian approach for processing OLBI using a constrained local descent method to minimize Eq. (15). A convex data criterion J, i.e., such that $J(k\cdot x_1+(1-k)\cdot x_2) \leqslant k\cdot J(x_1)+(1-k)\cdot J(x_2)$, $\forall x_1,x_2$, $\forall k \in [0,1]$, has no local minima, which makes the minimization much easier. Unfortunately, the criterion J is nonconvex. To be more precise, the difficulty of the problem can be summed up as follows:

(i) The small number of Fourier coefficients makes the problem underdetermined. Here the regularization term and the positivity constraint can help by limiting the high frequencies of the reconstructed object [6].

(ii) Closure phase measurements imply missing phase information and make the Fourier synthesis problem nonconvex. Adding a regularization term does not generally correct the problem [18].

(iii) Phase and modulus measurements with additive Gaussian noise lead to a non-Gaussian likelihood and a nonconvex log-likelihood with respect to $\boldsymbol{x}$. As a consequence, even with no missing phases, some approximation of the real observable statistics is necessary to get a convex data fidelity term. This data conversion from polar to Cartesian coordinates, which is commonly used in the field of radar processing [19], has been studied only recently in OLBI [20]; see Subsection 3.C.

These characteristics imply that optimizing J by a local descent algorithm can work only if the initialization selects the "right" valley of the criterion. The design of a good initial position is very case dependent and will not be extensively addressed here. The other key aspects are then the followed path, i.e., the minimization method, and the shape of the function to minimize, i.e., the behavior of the criterion $\boldsymbol{x} \mapsto J(\boldsymbol{x})$. This paper addresses both aspects:

- We design a specific OLBI criterion $\mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha})$ where two sets of variables appear explicitly, one in the spatial domain $\boldsymbol{x}$, describing the sought object, and another in the Fourier phase domain $\boldsymbol{\alpha}$, which accounts for the missing phase information. This specific criterion is designed to solve (iii), i.e., so that for a known $\boldsymbol{\alpha}$, the criterion is convex with respect to $\boldsymbol{x}$. In other words, if we had all the complex visibility phase measurements instead of just the closure phases, our criterion $\boldsymbol{x} \mapsto \mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha})$ *would be convex*;
- We adopt an alternate minimization method, working on the two sets of variables. This approach can be related to "myopic" approaches of some inverse problems, where missing data concerning the instrumental response are modeled and sought for during the inversion [21]. Alternate minimization methods are inspired by self-calibration methods in radio interferometry and have been used in optical interferometry by Lannes *et al.* [6]. However, the criterion used in [6] was essentially imported from radio interferometry and does not match OLBI data model [13]. Our main contribution is to derive a criterion that accounts for data model (13), while allowing an efficient alternate minimization. This construction is the subject of the next section.

3. EQUIVALENT MYOPIC MODEL FOR SELF-CALIBRATION

The aim of this section is to approximate the data model of Eq. (13):

$$\boldsymbol{s}^{\text{data}}(t) = \boldsymbol{a}^2(\boldsymbol{x},t) + \boldsymbol{s}^{\text{noise}}(t), \qquad \boldsymbol{s}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{s}(t)}), \tag{19}$$

$$\boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}^{\text{noise}}(t)[2\pi],$$

$$\boldsymbol{\beta}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{\beta}(t)}) \tag{20}$$

by a myopic linear model with additive complex Gaussian noise of the following form:

$$\boldsymbol{y}^{\text{data}}(t) = \mathcal{F}_{\boldsymbol{\alpha}(t)} \cdot \boldsymbol{H}(t)\boldsymbol{x} + \boldsymbol{y}^{\text{noise}}(t), \tag{21}$$

where the operator $\cdot$ denotes componentwise multiplication and $\mathcal{F}_{\boldsymbol{\alpha}(t)}$ is a vector of phasors depending on phase aberration parameters $\boldsymbol{\alpha}(t)$, which are defined in Subsection 3.B. This will be done in three steps:

- Subsection 3.A is devoted to the derivation of the observation model for the pseudo amplitude term $\boldsymbol{a}^{\text{data}}(t)$ from Eq. (19).
- Subsection 3.B is devoted to the derivation of the observation model for the pseudo phase term $\boldsymbol{\phi}^{\text{data}}(t)$ from Eq. (20).
- Subsection 3.C shows how to combine pseudo phase and pseudo amplitude models in a complex model such as Eq. (21) while solving problem (iii) of Subsection 2.E.

A. Pseudo Amplitude Data Model

In Eq. (19), we have assumed a Gaussian distribution for $\boldsymbol{s}^{\text{data}}(t)$ around $\boldsymbol{s}(\boldsymbol{x},t)$, which is questionable, since squared amplitudes should be nonnegative. However, such a statistic model is acceptable provided that the probability of a negative component of $\boldsymbol{s}^{\text{data}}(t)$ is very weak. For uncorrelated measurements, this assumption corresponds to mean values much greater than the corresponding standard deviation. Appendix D shows how to build the mean and covariance matrix of the square root of such a distribution. The mean vector is taken as the pseudo amplitude data $\boldsymbol{a}^{\text{data}}(t)$ and the covariance matrix called $\boldsymbol{R}_{\boldsymbol{a}(t)}$.

Observation model (19) can then be approximated by the following pseudo amplitude data model:

$$\boldsymbol{a}^{\text{data}}(t) = \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}^{\text{noise}}(t), \qquad \boldsymbol{a}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{a}(t)}). \tag{22}$$

B. Pseudo Phase Data Model

We start from a generalized inverse solution to the phase closure equation of Eq. (20). The generalized inverse $\boldsymbol{C}^{\dagger}$ of $\boldsymbol{C}$, defined by $\boldsymbol{C}^{\dagger} \triangleq \boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1}$, is such that $\boldsymbol{C}\boldsymbol{C}^{\dagger}=\mathbf{I}\mathrm{d}$. By applying it on all the terms of Eq. (20), we obtain

$$\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t) + 2\pi\boldsymbol{C}^{\dagger}\boldsymbol{\kappa}, \tag{23}$$

where $\boldsymbol{\kappa}$ is a vector of integers to account for the fact that each phase component is measured modulo 2π. We define

$$\boldsymbol{\phi}^{\text{data}}(t) \triangleq \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{data}}(t), \tag{24}$$

$$\boldsymbol{\phi}^{\text{ker}}(t) \triangleq (\boldsymbol{C}^{\dagger}\boldsymbol{C} - \mathbf{I}\mathrm{d})\boldsymbol{\phi}(\boldsymbol{x},t) + 2\pi\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} \tag{25}$$

and obtain

$$\boldsymbol{\phi}^{\text{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\phi}^{\text{ker}}(t) + \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t). \tag{26}$$

Vector $\boldsymbol{\phi}^{\text{ker}}(t)$ belongs to the 2π-wrapped kernel of operator $\boldsymbol{C}$:

$$\boldsymbol{C}\boldsymbol{\phi}^{\text{ker}}(t) = (\underbrace{\boldsymbol{C}\boldsymbol{C}^{\dagger}}_{=\mathbf{I}\mathrm{d}}\boldsymbol{C} - \boldsymbol{C})\boldsymbol{\phi}(\boldsymbol{x},t) + 2\pi\underbrace{\boldsymbol{C}\boldsymbol{C}^{\dagger}}_{=\mathbf{I}\mathrm{d}}\boldsymbol{\kappa} = 2\boldsymbol{\kappa}\pi = \mathbf{0}[2\pi].$$

As shown in Appendix C, if $\boldsymbol{\phi}^{\text{ker}}=0[2\pi]$, there exists a real vector $\boldsymbol{\alpha}(t)$ of dimension $N_{\mathrm{t}}-1$ such that $\boldsymbol{\phi}^{\text{ker}}(t) =\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)[2\pi]$, where $\bar{\boldsymbol{B}}$ is obtained by removing the first column of operator $\boldsymbol{B}$. So we have

$$\boldsymbol{\phi}^{\text{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x},t) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t)[2\pi]. \tag{27}$$

Now the problem is that $\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t)$ is a zero-mean random vector with a *singular covariance matrix*:

$$\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)} \triangleq \boldsymbol{C}^{\dagger}\boldsymbol{R}_{\boldsymbol{\beta}(t)}\boldsymbol{C}^{\dagger\mathrm{T}}.$$

To obtain a strictly convex log-likelihood, we have to approximate this term by a proper Gaussian vector $\boldsymbol{\phi}^{\text{noise}}(t)$, with an invertible covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$ chosen so as to correctly fit the second-order statistics of the noise in phase closure measurement equation (20). This last requirement can be written as the following equation:

$$\boldsymbol{C}\boldsymbol{R}_{\boldsymbol{\phi}(t)}\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{R}_{\boldsymbol{\beta}(t)}. \tag{28}$$

In other words, we are led to choose an invertible covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$ so as to mimic the statistical behavior of the closures, which is expressed by Eq. (28).

We propose to modify matrix $\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}$ by setting its nondiagonal components to 0, *i.e.*, to use the following diagonal matrix:

$$\{\boldsymbol{R}_{\boldsymbol{\phi}(t)}\}_{ij} = \begin{cases} 3\cdot\{\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}\}_{ij} & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}. \tag{29}$$

The factor 3 allows us to preserve the total weight of the phase term in the log-likelihood by satisfying the condition

$$\sum_{i,j}|\{\boldsymbol{R}_{\boldsymbol{\phi}(t)}\}_{ij}| = \sum_{i,j}|\{\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}\}_{ij}|.$$

There are several ways of choosing $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$, and we propose this particular choice without claiming it is optimal. Note that the myopic model derived in what follows can accommodate to any choice of a proper (*i.e.*, invertible) covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$.

With Eqs. (24), (27), and (29), we obtain the visibility phase pseudo data model:

$$\boldsymbol{\phi}^{\text{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x},t) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}^{\text{noise}}(t)[2\pi],$$

$$\boldsymbol{\phi}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}). \tag{30}$$

C. Pseudo Complex Visibility Data Model

Gathering Eqs. (22) and (30), we have finally approximated the data model [Eqs. (19) and (20)] by

$$\begin{cases} \boldsymbol{a}^{\text{data}}(t) = \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}^{\text{noise}}(t), \\ \boldsymbol{\phi}^{\text{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x},t) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}^{\text{noise}}(t)[2\pi], \\ \text{with } \boldsymbol{a}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{a}(t)}), \quad \boldsymbol{\phi}^{\text{noise}}(t) \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}). \end{cases} \tag{31}$$

We form pseudo complex visibility measurements $\boldsymbol{y}^{\text{data}}(t)$ defined by

$$\boldsymbol{y}^{\text{data}}(t) \triangleq \boldsymbol{a}^{\text{data}}(t)\cdot \mathrm{e}^{\mathrm{i}\boldsymbol{\phi}^{\text{data}}(t)}. \tag{32}$$

The approach proposed in [20], which we recall and generalize in Appendix E, is based on an approximated complex visibility data model:

$$\boldsymbol{y}^{\text{data}}(t) = \boldsymbol{H}(t)\boldsymbol{x}\cdot \mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)} + \boldsymbol{y}^{\text{noise}}(t). \tag{33}$$

This is exactly the sought model stated at the beginning of this section in Eq. (21), with $\mathcal{F}_{\boldsymbol{\alpha}(t)}=\mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}$. We now define the myopic observation model as follows:

$$\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t)) \triangleq \boldsymbol{H}(t)\boldsymbol{x}\cdot \mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}. \tag{34}$$

As shown in Appendix E, the mean value $\bar{\boldsymbol{y}}^{\text{noise}}(t)$ and covariance matrix $\boldsymbol{R}_{\boldsymbol{y}^{\text{noise}}(t)}$ of the additive complex noise term $\boldsymbol{y}^{\text{noise}}(t)$ are carefully designed so that the corresponding data likelihood criterion is convex quadratic with respect to the complex $\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))$ while remaining close to the real nonconvex model. To illustrate these properties, we consider one complex visibility and plot in the complex plane the distribution of $\boldsymbol{y}^{\text{data}}(t)$ around $\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))$ for the true noise distribution—i.e., a polar Gaussian noise in phase and modulus—and our Cartesian Gaussian approximation (see Fig. 1) In particular, the "elliptic" covariance matrix we propose (which yields elliptic contour plots in Fig. 1) is preferable to the more classical

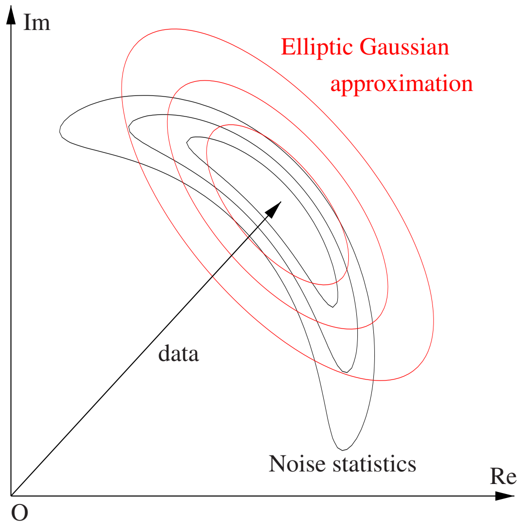

Fig. 1. (Color online) Contour plots of a polar Gaussian distribution and of its Cartesian Gaussian approximation.

"circular" approximation that appears in previous contributions on OLBI [22]. The latter can be described by half as many parameters as needed for the elliptic one (one radius for a circle, instead of a short axis and a long axis for an ellipsis), but it is clearly less accurate [20] (such a noise statistics description has also been investigated for the complex bispectra in the OIFITS data exchange format [11]).

From Eq. (33), we build Chi-2 statistics over real and imaginary parts of the observation equation

$$\chi^2_{\boldsymbol{y}(t)}(\boldsymbol{x},\boldsymbol{\alpha}(t)) \triangleq \begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}^{\mathrm{data}}(t) - \boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t)) - \bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \\ \mathrm{Im}\{\boldsymbol{y}^{\mathrm{data}}(t) - \boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t)) - \bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \end{bmatrix}^{\mathrm{T}} \times \boldsymbol{R}_{\boldsymbol{y}^{\mathrm{noise}}(t)}^{-1} \begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}^{\mathrm{data}}(t) - \boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t)) - \bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \\ \mathrm{Im}\{\boldsymbol{y}^{\mathrm{data}}(t) - \boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t)) - \bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \end{bmatrix}.$$

And we finally propose the myopic goodness-of-fit criterion:

$$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_t \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}(t),t) = \sum_t \chi^2_{\boldsymbol{y}(t)}(\boldsymbol{x},\boldsymbol{\alpha}(t)). \quad (35)$$

We can now design a myopic Bayesian approach to the reconstruction problem by combining the data term with a regularization term along the lines of Subsection 2.E:

$$\mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha}) = \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) + \mathcal{J}^{\mathrm{prior}}(\boldsymbol{x}). \quad (36)$$

The next section describes an alternate minimization technique applied to regularized criterion (36).

4. WISARD

In this section, we describe WISARD, standing for Weak-phase Interferometric Sample Alternating Reconstruction Device, a self-calibration method for OLBII.

A. Global Structure of WISARD

WISARD is made of four major blocks:

- A first block recasts the raw data (i.e., closure phases and squared visibilities) in myopic data (i.e., phases and moduli) as described in Subsections 3.A and 3.B.
- A second "convexification block" computes a Gaussian approximation of the pseudo visibility data model as described in Subsection 3.C.
- A third block builds a guess for the object $\boldsymbol{x}$ and aberrations $\boldsymbol{\alpha}$ (i.e., a good starting point).
- Finally, the self-calibration block performs the minimization of regularized criterion (36), under constraints (16). It alternates optimization of the object for given aberrations and optimization of the aberrations for the current object.

The structure of WISARD is sketched in Fig. 2. The principles that underline the three first blocks of WISARD have been described in previous sections, while details on the self-calibration minimization are gathered in the next one.

B. Self-Calibration Block

In the following, we describe the three key components of the self-calibration block.

Minimization with respect to $\boldsymbol{x}$. The criterion $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha})$ we have derived is quadratic and hence convex with respect to the object $\boldsymbol{x}$. Hence the minimization versus $\boldsymbol{x}$ does not raise special difficulties.

Minimization with respect to $\boldsymbol{\alpha}$. $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha})$ is the sum of terms involving only measurements obtained at one time instant t [Eq. (35)]:

$$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_t \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}(t),t).$$

Because the time between two measurements is much greater than the turbulence coherent time (around

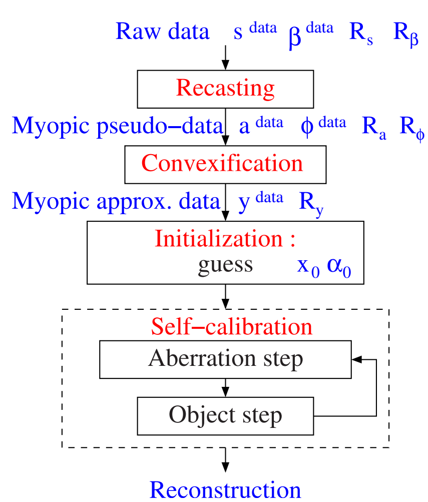

Fig. 2. (Color online) WISARD algorithm loop.

10 ms), aberrations $\boldsymbol{\alpha}(t)$ at two different instants are statistically independent. We can then solve separately for each set of $\boldsymbol{\alpha}(t)$, which dramatically reduces the complexity of the minimization. The number of $\boldsymbol{\alpha}(t)$ components to solve for is $(N_t - 1)$ and the minimization is delicate, as the criterion exhibits periodic structures that have been studied in [22].

However, exact minimization is affordable for a three-telescope interferometric array. In this case we have to perform several two-parameter minimizations, and each one can be efficiently initialized by an exhaustive search on a 2D grid, which ensures we avoid local minima. On the other hand, when N_t gets high enough, e.g., 6, then the number of $\boldsymbol{\alpha}(t)$ to solve for, e.g., 5, gets small compared to the number of closure phases available, e.g., 15. With a three-telescope array, 2/3 of the phase information is missing, whereas with a six-telescope array, only 1/3 of the phase information is missing. In this last case, which corresponds to the processing of synthetic data presented in Subsection 5.A, the reconstructions were straightforward, and no effects of the local minima in $\boldsymbol{\alpha}$ were witnessed.

In other words, coping with the ambiguities in $\boldsymbol{\alpha}$, for instance, with the specific criterion proposed in [22], may be necessary only for $N_t=4$ or $N_t=5$. For $N_t=3$, an exhaustive search is possible, and for $N_t \geqslant 6$, ambiguities in $\boldsymbol{\alpha}$ do not have, according to our experience, a major impact on reconstruction.

Starting point: object and aberration guess $\boldsymbol{x}_0$ *and* $\boldsymbol{\alpha}_0$. If a parametric model of the observed stellar source is not available, the object starting point is a mean square solution, from which we extract the positive part. The first step in the self-calibration block is a minimization with respect to $\boldsymbol{\alpha}$ for $\boldsymbol{x}=\boldsymbol{x}_0$.

5. RESULTS

This section presents some results of processing by the WISARD algorithm, with both synthetic and experimental data.

A. Processing of Synthetic Data

The first example takes synthetic interferometric data that were used in the international Imaging Beauty Contest organized by P. Lawson for the International Astronomical Union (IAU) [23]. These data simulate the observation of the synthetic object shown in Fig. 3 with the Navy Prototype Optical Interferometer (NPOI) [24] six-telescope interferometer. The corresponding frequency coverage, shown in Fig. 3, has a structure in arcs of circles typical of the supersynthesis technique, which consists in repeating the measurements over several nights of observation so that the same baselines access different measurement spatial frequencies because of the Earth's rotation. In total, there are 195 square visibility modules and 130 closure phases, together with the associated variances.

Six reconstructions obtained with WISARD are shown in Fig. 4. On the upper row is a reconstruction using a quadratic regularization based on a power spectral density model in $1/|u|^3$ for a weak, a strong, and a correct regularization parameter. The latter gives a satisfactory level of smoothing but does not restore the peak in the center of the object. The peak is visible in the under-regularized reconstruction on the left but at the cost of too high a residual variance.

The reconstruction presented on the lower row is a good trade-off between smoothing and restoration of the central peak thanks to the use of the white $L_2L_1^w$ prior term introduced in Subsection 2.E. The automatically set parameters [Eq. (18)] are very satisfactory (left), and a light tuning (center and right) allows an even better reconstruction. The goodness of fit of the $L_2L_1^w$ reconstruction can be appreciated in Fig. 5. The crosses (red online) show the reconstructed visibility moduli (i.e., of the FT of the reconstructed object at the measurement frequencies), and the squares (blue online) are the moduli of the measured visibilities. The difference between the two, weighted by 10 times the standard deviation of the moduli, is shown as the dotted curve. The mean value of this difference is 0.1, which shows a good fit (to within 1 σ).

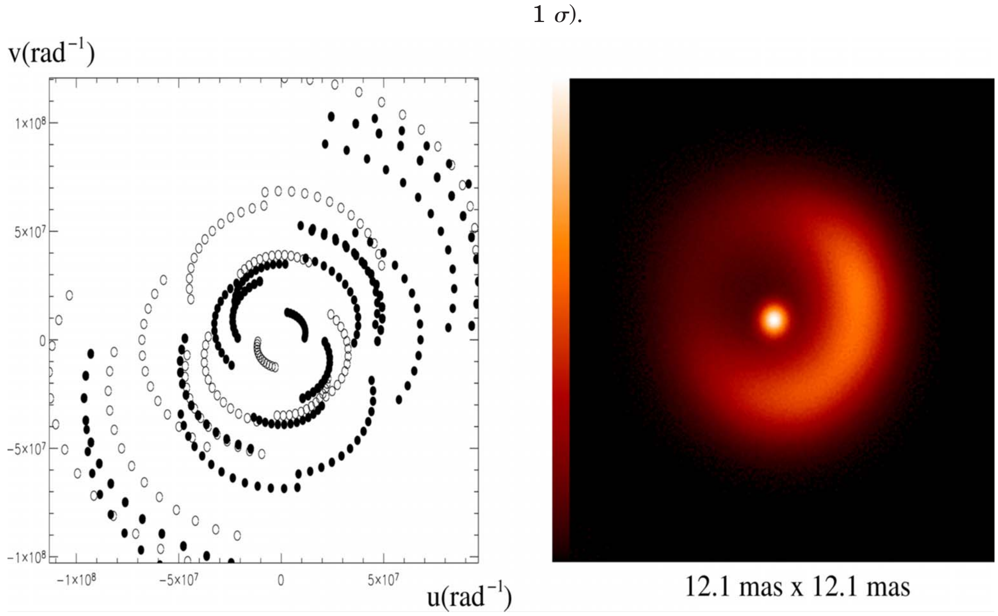

Fig. 3. (Color online) Synthetic object (right) and frequency coverage (left) from the Imaging Beauty Contest 2004.

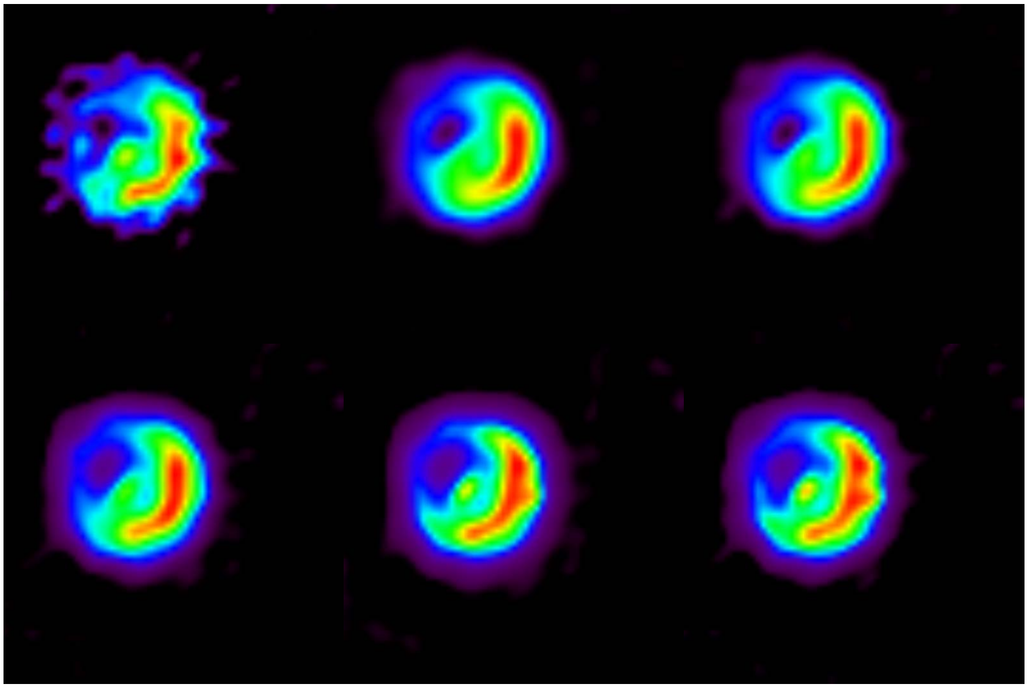

Fig. 4. (Color online) Reconstructions with WISARD. Upper row, under-regularized quadratic model (left), over-regularized quadratic model (center), quadratic model with correct regularization parameter (right). Lower row, white $L_2L_1^{uu}$ model with automatically set scale and delta parameters (left), white $L_2L_1^{uu}$ model with half-scale (center), white $L_2L_1^{uu}$ model with half-delta (right). Each image field is 12.1×12.1 mas.

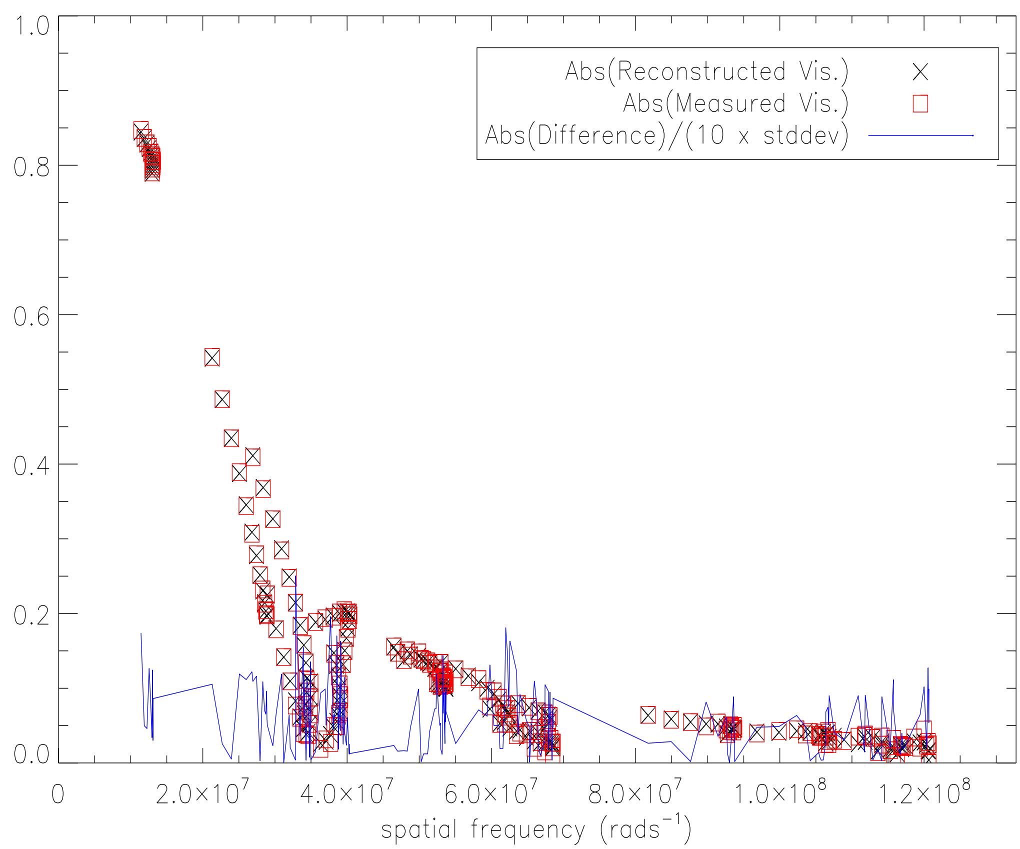

Fig. 5. (Color online) Goodness of fit at WISARD convergence.

B. Processing of Experimental Data

Here we present the reconstruction of the star χ Cygni from experimental data using the WISARD algorithm. The data were obtained by S. Lacour and S. Meimon under the leadership of G. Perrin during a measuring campaign on the IR/Optical Telescope Array (IOTA) interferometer [25] in May 2005. As already mentioned, each measurement has to be calibrated by observation of an object that acts as a point source at the instrument's resolving power. The calibrators chosen were HD 180450 and HD 176670.

The star χ Cygni is a Mira-type star, Mira itself being an example of such stars. Perrin *et al.* [26] propose a model of Mira-type stars, composed of a photosphere, an empty layer, and a thin molecular layer. The aim of the mission was to obtain images of χ Cygni in the H band (1.65 μm ± 175 nm) and, in particular, to highlight possible assymmetric features in the structure of the molecular layer.

Figure 6 shows, on the left, the $u-v$ coverage obtained, i.e., the set of spatial frequencies measured, multiplied by the observation wavelength. Because the sky is habitually represented with the west on the right, the coordinates used are, in fact, $-u,v$. The domain of the accessible $u-v$ plane is constrained by the geometry of the interferometer and the position of the star in the sky. The "hourglass" shape is characteristic of the IOTA interferometer, and entails nonuniform resolution that affects the image reconstruction, shown on the right. The reconstructed angular field has sides of 60 mas. In addition to the positivity constraint, the regularization term used is the $L_2L_1^w$ term described in Subsection 2.E. The interested reader will find an astrophysical interpretation of this result in [27].

6. CONCLUDING COMMENTS

We have proposed a complete and precise self-calibration approach to optical interferometry image reconstruction. After pointing out the data model specificities in the OLBI context, we have emphasized the sources of underdeterminations, which make a classical Bayesian criterion descent method critical. Namely, the main problems are the phase underdetermination caused by turbulence effects, and, as noted only recently, the polar coordinate structure of the data model.

We have built a specially designed approximate myopic data model in order to derive a self-calibration method. Special care was given to the design of the second-order statistics of the myopic model, an aspect that was ignored in previous related works.

We have extended our previous work on polar data conversion [20] and proposed a convex approximation of the noise model that reduces the number of local minima of the criterion to minimize.

We also addressed integer ambiguities induced by closure phase wrapping, which are classical when dealing with phase data, and have discussed their impact on the image reconstruction quality: for three-telescope data, we have proposed an exhaustive search method, and we have witnessed that these ambiguities do not raise any particular problem when processing the interferometer data of six or more telescopes. Concerning the remaining case of four to five telescopes, the work by Lannes [22] should be worth investigating. On the other hand, global minimization methods were left aside because of their intensive computation needs. As computer performance increases, these methods might be, in the years to come, an appropriate way to deal with local minima.

All these developments allowed us to propose WISARD, a self-calibration method for OLBII reconstruction and to demonstrate its efficiency on simulated data.

Finally, WISARD was also used to successfully process real astronomical OLBI data sets. These results were made possible thanks to a close partnership with the astronomers Sylvestre Lacour and Guy Perrin of the Observatoire de Paris Meudon, within the PHASE partnership (Partenariat Haute résolution Angulaire Sol-Espace). Indeed, an accurate astronomical model of the observed stellar object is a precious guideline for reconstructing a complex image from OLBI data. From the author's point

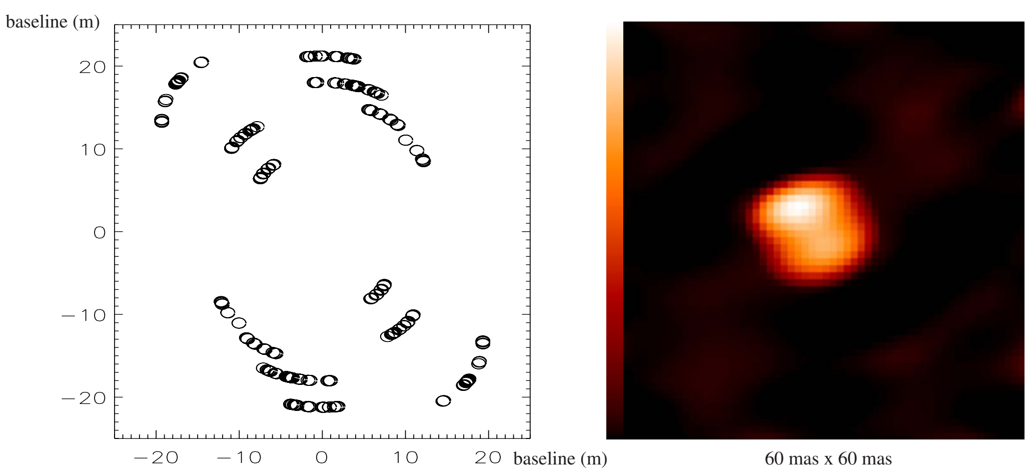

Fig. 6. (Color online) Frequency coverage (left) and reconstruction of the star χ Cygni (right).

of view, such a collaboration is essential to the success of OLBII techniques.

APPENDIX A: BASELINE AND CLOSURE OPERATORS C AND B

Let N_t be the number of telescopes of the interferometric array. We have the following definitions:

$$\boldsymbol{B}_2 \triangleq [-1 \quad 1], \tag{A1}$$

$$\boldsymbol{B}_{N_t} \triangleq \left[\begin{array}{c|c} -\mathbf{1}_{N_t-1} & \mathbf{Id}_{N_t-1} \\ \hline \boldsymbol{O} & \boldsymbol{B}_{N_t-1} \end{array}\right], \tag{A2}$$

$$\boldsymbol{C}_{N_t} \triangleq \left[-\boldsymbol{B}_{N_t-1} \quad \mathbf{Id}_{[(N_t-1)(N_t-2)]/2}\right], \tag{A3}$$

for $N_t \geqslant 3$.

In what follows, we prove that $\ker \boldsymbol{C} = \operatorname{im} \boldsymbol{B}$.

We have $\boldsymbol{C}_{N_t}\boldsymbol{B}_{N_t} = 0$, so

$$\operatorname{im} \boldsymbol{B} \subset \ker \boldsymbol{C}. \tag{A4}$$

It is straightforward to prove by recurrence that $\boldsymbol{B}_{N_t} \cdot \mathbf{1}_{N_t} = 0$, which yields rank $\boldsymbol{B}_{N_t} \leqslant N_t - 1$. Because $\boldsymbol{B}_{N_t}$ contains $\mathbf{Id}_{N_t-1}$, we gather

$$\dim \operatorname{im} \boldsymbol{B} \triangleq \operatorname{rank} \boldsymbol{B} = N_t - 1. \tag{A5}$$

Here $\boldsymbol{C}_{N_t}$ contains $\mathbf{Id}_{(N_t-1)(N_t-2)/2}$, which yields rank $\boldsymbol{C}_{N_t} \geqslant (N_t-1)(N_t-2)/2$, or

$$\dim \ker \boldsymbol{C}_{N_t} \leqslant N_t - 1. \tag{A6}$$

With Eqs. (A4)–(A6), we gather

$$\ker \boldsymbol{C} = \operatorname{im} \boldsymbol{B}. \tag{A7}$$

APPENDIX B: CHARACTERIZATION OF THE BASELINE PHASE-INDEPENDENT OPERATORS

Here we prove that any continuous differentiable function f verifying property (8)

$$f(\boldsymbol{\phi} + \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{\phi}), \quad \forall\, (\boldsymbol{\phi}, \boldsymbol{\varphi})$$

is such that $f(\boldsymbol{\phi}) = g(\boldsymbol{C}\phi)$, where $\boldsymbol{C}$ has more columns than rows, so its pseudo inverse is defined by $\boldsymbol{C}^{\dagger} \triangleq \boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1}$ and verifies

$$\boldsymbol{C}\boldsymbol{C}^{\dagger} = \mathbf{Id} \tag{B1}$$

and thus

$$\boldsymbol{C}\boldsymbol{C}^{\dagger}\boldsymbol{C} - \boldsymbol{C} = 0 \Rightarrow \boldsymbol{C}(\boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{\phi}) = 0,$$

$$\forall\, \boldsymbol{\phi} \overset{\text{A7}}{\Rightarrow} \exists\, \boldsymbol{\varphi}, (\boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{\phi}) = \boldsymbol{B}\boldsymbol{\varphi},$$

$$\forall\, \boldsymbol{\phi} \Rightarrow \exists\, \boldsymbol{\varphi}, \boldsymbol{\phi} = \boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{B}\boldsymbol{\varphi}, \quad \forall\, \boldsymbol{\phi}.$$

With this we obtain that any f verifying property (8) is such that

$$f(\boldsymbol{\phi}) = f(\boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi}) = g(\boldsymbol{C}\boldsymbol{\phi}).$$

APPENDIX C: WRAPPED KERNEL OF OPERATOR C

The kernel of operator $\boldsymbol{C}$ is given by $\ker \boldsymbol{C} = \operatorname{im} \boldsymbol{B}$ [Eq. (A7)]. With dimensional arguments, it is easy to see that

$$\operatorname{im} \boldsymbol{B} = \operatorname{im} \bar{\boldsymbol{B}},$$

where $\bar{\boldsymbol{B}}$ is obtained by removing the first column of operator $\boldsymbol{B}$, so we have

$$\ker \boldsymbol{C} = \operatorname{im} \bar{\boldsymbol{B}}. \tag{C1}$$

Let us now characterize the set of $\boldsymbol{\phi}^{\ker}$ such that

$$\boldsymbol{C}\boldsymbol{\phi}^{\ker} \equiv 0[2\pi].$$

Because $\boldsymbol{C}$ has integer components, $\boldsymbol{\phi}^{\ker}$ can be considered modulo 2π. With Eq. (C1), we obtain

$$\exists\, \boldsymbol{\alpha}_1,\ \boldsymbol{\phi}^{\ker} \equiv \boldsymbol{C}^{\dagger}(0[2\pi]) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}_1[2\pi]. \tag{C2}$$

Because $\bar{\boldsymbol{B}}$ has integer components, $\boldsymbol{\alpha}_1$ can be considered modulo 2π. The issue here is to evaluate the $\boldsymbol{C}^{\dagger}(0[2\pi])$ term, i.e., the value of $\boldsymbol{C}^{\dagger}(2\pi\boldsymbol{\kappa})$, with $\boldsymbol{\kappa}$ any integer vector.

Equations (A1) show that $\boldsymbol{C} = [\boldsymbol{M}|\mathbf{Id}]$. The integer vector

$$\boldsymbol{\mu} \triangleq \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix}$$

is then such that

$$\boldsymbol{C}\boldsymbol{\mu} = [*|\mathbf{Id}]\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix} = \boldsymbol{\kappa}.$$

Then we have

$$\boldsymbol{C}\boldsymbol{\mu} = \boldsymbol{\kappa} \Rightarrow \boldsymbol{C}\boldsymbol{\mu}' = \boldsymbol{C}\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} \Rightarrow \boldsymbol{C}(\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} - \boldsymbol{\mu}) = 0 \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} - \boldsymbol{\mu} = \boldsymbol{B}\boldsymbol{\alpha}_2 \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$2\pi\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} = 2\pi\boldsymbol{\mu} + \boldsymbol{B}(2\pi\boldsymbol{\alpha}_2) \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$\boldsymbol{C}^{\dagger}(0[2\pi]) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}_1 \equiv \boldsymbol{B}\underbrace{(2\pi\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_1)}_{\boldsymbol{\alpha}}[2\pi].$$

So Eq. (C2) yields

$$\exists\, \boldsymbol{\alpha}, \boldsymbol{\phi}^{\ker} \equiv \bar{\boldsymbol{B}}\boldsymbol{\alpha}[2\pi]. \tag{C3}$$

APPENDIX D: SQUARE ROOT OF A GAUSSIAN DISTRIBUTION

Let us assume we measure the squared value s of a positive value a, with an additive Gaussian noise:

$$s^{\text{data}} = a^2 + s^{\text{noise}}, \tag{D1}$$

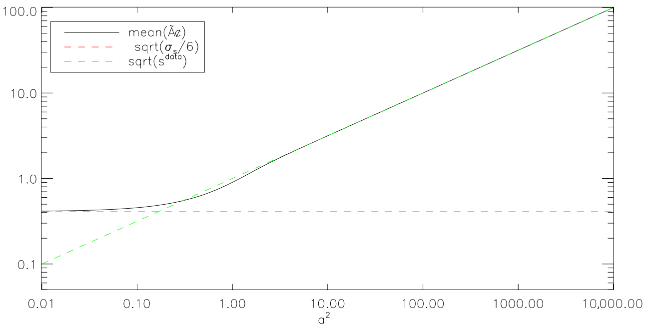

Fig. 7. (Color online) Behavior of $\langle \hat{a} \rangle$ in function of a^2 with a unit σ_s.

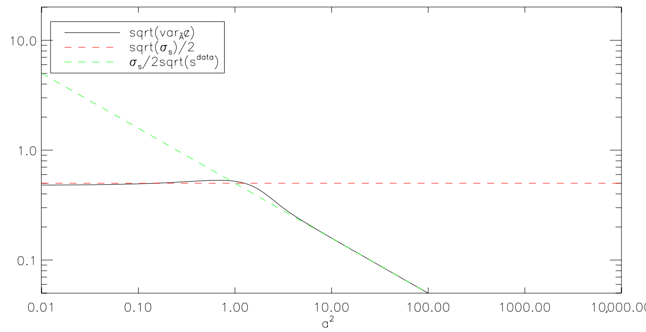

Fig. 8. (Color online) Behavior of $\sqrt{\mathrm{Var}(\hat{a})}$ in function of a^2 with a unit σ_s.

with s^{noise} being zero-mean Gaussian with the variance σ_s^2. Let $\hat{a}$ be the estimator of a from s^{data} defined by

$$\hat{a} = \begin{cases} \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} > 0 \\ 0 & \text{else} \end{cases},$$

where $\hat{a}$ can be seen as pseudo data. The data model of $\hat{a}$ derived from Eq. (D1) is not additive Gaussian. As will be shown in Appendix E, an optimal Gaussian approximation of the data model of $\hat{a}$ would be

$$\hat{a} = a + a^{\text{noise}}, \tag{D2}$$

with a^{noise} a Gaussian noise with a mean equal to $\langle \hat{a} \rangle$ and a standard deviation $\sqrt{\mathrm{Var}(\hat{a})}$.

We have studied the behavior of the mean $\langle \hat{a} \rangle$ and standard deviation $\sqrt{\mathrm{Var}(\hat{a})}$ of this estimator for various values of a^2, with a unit σ_s (see Figs. 7 and 8). We can distinguish two regimes for $\langle \hat{a} \rangle$:

- A low-mean regime, where $a^2 \leqslant \sigma_s/6$: a nonnegligible part of the distribution of s^{data} around a^2 is in the negative domain. Because $\hat{a}$ estimates a null value for a when s^{data} is negative, its mean will depend mainly on the width of the Gaussian wings. A good approximation of $\langle \hat{a} \rangle$ is $\sqrt{\sigma_s/6}$.
- A high-mean regime, where $a^2 \geqslant \sigma_s/6$: most of the distribution of s^{data} around a^2 is in the positive domain. The fact that $\hat{a}$ estimates a null value for a when $s^{\text{data}} < 0$ does not affect its mean $\langle \hat{a} \rangle$, which is close to a. Because a is not known, we choose $\langle \hat{a} \rangle = \sqrt{s^{\text{data}}}$. We can distinguish the same two regimes for $\sqrt{\mathrm{Var}(\hat{a})}$. However, the transition is around σ_s:
- When $a^2 \leqslant \sigma_s$, the fact that $\hat{a}$ estimates a null value for a when s^{data} is negative tends to diminish its standard deviation, which we approximate by $\sqrt{\mathrm{Var}(\hat{a})} \simeq \sqrt{\sigma_s}/2$.
- In the high-mean regime, where $a^2 \geqslant \sigma_s$, most of the distribution of s^{data} around a^2 is in the positive domain, and $\sqrt{\mathrm{Var}(\hat{a})}$ is close to the classical expression. This expression corresponds to a first-order expansion in σ_a:

$$(a + \sigma_a)^2 = a^2 + \sigma_s \Rightarrow 2a\sigma_a \simeq \sigma_s,$$

where $\sigma_s/2a$. Because a is not known, we choose $\sqrt{\mathrm{Var}(\hat{a})} = \sigma_s/2\sqrt{s^{\text{data}}}$. We then propose the pseudo data model

$$a^{\text{data}} = a + a^{\text{noise}},$$

with

$$a^{\text{data}} = \begin{cases} \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} > 0 \\ 0 & \text{else} \end{cases}$$

and a^{noise} a Gaussian noise with mean and standard deviation defined by

$$\bar{a} = \begin{cases} \sqrt{\sigma_s/6} & \text{if } s^{\text{data}} \leqslant \sigma_s/6, \\ \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} \geqslant \sigma_s/6 \end{cases},$$

$$\sigma_a = \begin{cases} \sqrt{\sigma_s}/2 & \text{if } s^{\text{data}} \leqslant \sigma_s \\ \dfrac{\sigma_s}{2\sqrt{s^{\text{data}}}} & \text{if } s^{\text{data}} \geqslant \sigma_s \end{cases}.$$

We also decide to discard the data such that $s^{\text{data}} \leqslant -\sigma_s$.

APPENDIX E: CARTESIAN GAUSSIAN APPROXIMATION TO A POLAR GAUSSIAN DISTRIBUTION

If we define

$$\boldsymbol{y}_{\boldsymbol{\alpha}(t)}(\boldsymbol{x},t) \triangleq \boldsymbol{H}(t)\boldsymbol{x} \cdot \mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}, \tag{E1}$$

Eq. (31) reads

$$\begin{cases} \boldsymbol{a}^{\text{data}}(t) = |\boldsymbol{y}_{\boldsymbol{\alpha}(t)}|(\boldsymbol{x},t) + \boldsymbol{a}^{\text{noise}}(t), & \boldsymbol{a}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{a}(t)}), \\ \boldsymbol{\phi}^{\text{data}}(t) \overset{2\pi}{\equiv} \arg \boldsymbol{y}_{\boldsymbol{\alpha}(t)}(\boldsymbol{x},t) + \boldsymbol{\phi}^{\text{noise}}(t), & \boldsymbol{\phi}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}). \end{cases} \tag{E2}$$

1. General Expression

We consider a polar distribution of a Gaussian vector $\boldsymbol{y}$ of modulus $\boldsymbol{a}$ and phase $\boldsymbol{\phi}$:

$$\boldsymbol{\phi}^{\text{data}} = \bar{\boldsymbol{\phi}} + \boldsymbol{\phi}^{\text{noise}}, \tag{E3}$$

$$\boldsymbol{a}^{\text{data}} = \bar{\boldsymbol{a}} + \boldsymbol{a}^{\text{noise}}, \tag{E4}$$

where $\boldsymbol{\phi}^{\text{noise}}$ and $\boldsymbol{a}^{\text{noise}}$ are zero-mean real Gaussian vectors of covariance matrices $\boldsymbol{R}_{\boldsymbol{a}}$ and $\boldsymbol{R}_{\boldsymbol{\phi}}$ (the vectors $\boldsymbol{\phi}^{\text{noise}}$ and $\boldsymbol{a}^{\text{noise}}$ are assumed uncorrelated).

With the definitions

$$\begin{cases} \bar{\boldsymbol{y}} \triangleq \bar{\boldsymbol{a}} \exp \mathrm{i}\bar{\boldsymbol{\phi}}, \\ \boldsymbol{y}^{\text{noise}} \triangleq \boldsymbol{y}^{\text{data}} - \bar{\boldsymbol{y}}, \\ \boldsymbol{y}^{\text{n}}_{\text{rad}} \triangleq \mathrm{Re}\{\boldsymbol{y}^{\text{noise}} \mathrm{e}^{-\mathrm{i}\bar{\boldsymbol{\phi}}}\}, \\ \boldsymbol{y}^{\text{n}}_{\text{tan}} \triangleq \mathrm{Im}\{\boldsymbol{y}^{\text{noise}} \mathrm{e}^{-\mathrm{i}\bar{\boldsymbol{\phi}}}\}, \\ \bar{\bar{\boldsymbol{y}}}^{\text{noise}} \triangleq \begin{bmatrix} \boldsymbol{y}^{\text{n}}_{\text{rad}} \\ \boldsymbol{y}^{\text{n}}_{\text{tan}} \end{bmatrix}, \end{cases} \tag{E5}$$

we gather

$$\begin{cases} \boldsymbol{y}^{\text{n}}_{\text{rad}} = [\bar{\boldsymbol{a}} + \boldsymbol{a}^{\text{noise}}] \cos \boldsymbol{\phi}^{\text{noise}} - \bar{\boldsymbol{a}}, \\ \boldsymbol{y}^{\text{n}}_{\text{tan}} = [\bar{\boldsymbol{a}} + \boldsymbol{a}^{\text{noise}}] \sin \boldsymbol{\phi}^{\text{noise}}. \end{cases} \tag{E6}$$

A complex vector is Gaussian if and only if each of its components is Gaussian. A complex is Gaussian if and only if, in any Cartesian basis, its two components are Gaussian. So $\boldsymbol{y}$ is Gaussian if and only if $\bar{\bar{\boldsymbol{y}}}^{\text{noise}}$ is Gaussian, which is not the case [20]. In what follows, we show how to optimally approximate the distribution of $\bar{\bar{\boldsymbol{y}}}^{\text{noise}}$ by a Gaussian distribution.

2. Gaussian Approximation

We characterize our Cartesian additive Gaussian approximation, *i.e.,* its mean $\langle \bar{\bar{\boldsymbol{y}}}^{\text{noise}} \rangle$ and covariance $\mathbf{R}_{\bar{\bar{\boldsymbol{y}}}^{\text{noise}}}$, by minimizing the Kullback–Leibler distance between the two noise distributions, which gives [20]

$$\begin{cases} \langle \bar{\bar{\boldsymbol{y}}}^{\text{noise}} \rangle = E\left\{ \begin{bmatrix} \boldsymbol{y}^{\text{n}}_{\text{rad}} \\ \boldsymbol{y}^{\text{n}}_{\text{tan}} \end{bmatrix} \right\} = \begin{bmatrix} \bar{\boldsymbol{y}}^{\text{n}}_{\text{rad}} \\ \bar{\boldsymbol{y}}^{\text{n}}_{\text{tan}} \end{bmatrix}, \\ \boldsymbol{R}_{\bar{\bar{\boldsymbol{y}}}^{\text{noise}}} = E\left\{ \begin{bmatrix} \bar{\boldsymbol{y}}^{\text{n}}_{\text{rad}} - \boldsymbol{y}^{\text{n}}_{\text{rad}} \\ \bar{\boldsymbol{y}}^{\text{n}}_{\text{tan}} - \boldsymbol{y}^{\text{n}}_{\text{tan}} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{y}}^{\text{n}}_{\text{rad}} - \boldsymbol{y}^{\text{n}}_{\text{rad}} \\ \bar{\boldsymbol{y}}^{\text{n}}_{\text{tan}} - \boldsymbol{y}^{\text{n}}_{\text{tan}} \end{bmatrix}^{\mathrm{T}} \right\}, \end{cases} \tag{E7}$$

and we define

$$\boldsymbol{R}_{\bar{\bar{\boldsymbol{y}}}^{\text{noise}}} \triangleq \begin{bmatrix} \boldsymbol{R}_{\text{rad,rad}} & \boldsymbol{R}_{\text{rad,tan}} \\ \boldsymbol{R}^{\mathrm{T}}_{\text{rad,tan}} & \boldsymbol{R}_{\text{tan,tan}} \end{bmatrix}.$$

For a zero-mean Gaussian vector $\boldsymbol{\phi}^{\text{noise}}$ of covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}}$,

$$E\{\sin \phi_i^{\text{noise}}\} = 0,$$

$$E\{\cos \phi_i^{\text{noise}}\} = \exp\left(- \frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}}}{2} \right),$$

$$E\{\sin \phi_i^{\text{noise}} \sin \phi_j^{\text{noise}}\} = \sinh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \exp\left(- \frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}}}{2} \right),$$

$$E\{\cos \phi_i^{\text{noise}} \cos \phi_j^{\text{noise}}\} = \cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \exp\left(- \frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}}}{2} \right),$$

$$E\{\cos \phi_i^{\text{noise}} \sin \phi_j^{\text{noise}}\} = 0. \tag{E8}$$

By combining Eq. (E7), (E5), (E6), and (E8), we obtain

$$E\{\boldsymbol{y}^{\text{n}}_{\text{rad}_i}\} = \bar{a}_i[\mathrm{e}^{-\boldsymbol{R}_{\boldsymbol{\phi}_{ii}}/2} - 1],$$

$$E\{\boldsymbol{y}^{\text{n}}_{\text{tan}_i}\} = 0,$$

$$[\boldsymbol{R}_{\text{rad,rad}}]_{ij} = [\bar{a}_i \bar{a}_j (\cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} - 1) + \boldsymbol{R}_{a_{ij}} \cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}}] \cdot \mathrm{e}^{-[(\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}})/2]},$$

$$[\boldsymbol{R}_{\text{rad,tan}}]_{ij} = 0,$$

$$[\boldsymbol{R}_{\text{tan,tan}}]_{ij} = (\bar{a}_i \bar{a}_j + \boldsymbol{R}_{a_{ij}}) \sinh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \mathrm{e}^{-[(\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}})/2]}. \tag{E9}$$

3. Scalar Case

Now we make the additional assumption that both $\boldsymbol{\phi}^{\text{noise}}$ and $\boldsymbol{a}^{\text{noise}}$ are decorrelated, i.e.,

$$\begin{cases} \boldsymbol{R}_{\boldsymbol{a}} = \mathrm{Diag}\{\sigma^2_{a,i}\}, \\ \boldsymbol{R}_{\boldsymbol{\phi}} = \mathrm{Diag}\{\sigma^2_{\phi,i}\}. \end{cases}$$

We obtain

$$\begin{cases} \boldsymbol{R}_{\text{rad,rad}} = \mathrm{Diag}\{\sigma^2_{rad,i}\}, \\ \boldsymbol{R}_{\text{tan,tan}} = \mathrm{Diag}\{\sigma^2_{tan,i}\}, \\ \boldsymbol{R}_{\text{rad,tan}} = 0, \end{cases}$$

with

$$\sigma^2_{rad,i} = \frac{\bar{a}_i^2}{2}(1 - \mathrm{e}^{-\sigma^2_{\phi,i}})^2 + \frac{\sigma^2_{a,i}}{2}(1 + \mathrm{e}^{-2\sigma^2_{\phi,i}}),$$

$$\sigma^2_{tan,i} = \frac{\bar{a}_i^2}{2}(1 - \mathrm{e}^{-2\sigma^2_{\phi,i}}) + \frac{\sigma^2_{a,i}}{2}(1 - \mathrm{e}^{-2\sigma^2_{\phi,i}}). \tag{E10}$$

In this case, we can plot for one complex visibility the true noise distribution—i.e., a Gaussian noise in phase and modulus—and our Gaussian approximation (see Fig. 1).

ACKNOWLEDGMENTS

The authors want to express their special thanks to Eric Thiébaut for his support and for letting them use his minimization software. Serge Meimon is very grateful to Guy Perrin and Sylvestre Lacour, who allowed him to participate in two IOTA observing campaigns. We also thank all the people who contributed to the existence and success of the IOTA interferometer, in particular John Monnier, Wes Traub, Jean-Philippe Berger, and Marc Lacasse. Serge Meimon also thanks Vincent Bix Josso for his help with Appendix B. Serge Meimon and Laurent Mugnier acknowledge support from PHASE, the space- and ground-based high-angular-resolution partnership among ONERA, Observatoire de Paris, CNRS, and University Denis Diderot Paris 7.

REFERENCES

1. J. D. Monnier, "An introduction to closure phases," in *Principles of Long Baseline Stellar Interferometry*, P. R. Lawson, ed. (Jet Propulsion Laboratory, 1999), Chap. 13, pp. 203–239.
2. A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio-Astronomy* (Wiley Interscience, 1986).
3. A. Lannes, E. Anterrieu, and P. Maréchal, "Clean and wipe," Astron. Astrophys. Suppl. Ser. **123**, 183–198 (1997).
4. J. Hogbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," Astron. Astrophys. Suppl. Ser. **15**, 417–426 (1974).
5. T. J. Cornwell and P. N. Wilkinson, "A new method for making maps with unstable radio interferometers," Mon. Not. R. Astron. Soc. **196**, 1067–1086 (1981).
6. A. Lannes, "Weak-phase imaging in optical interferometry," J. Opt. Soc. Am. A **15**, 811–824 (1998).
7. J. W. Goodman, *Statistical Optics* (Wiley-Interscience, 1985).
8. D. L. Fried, "Statistics of a geometric representation of wavefront distortion," J. Opt. Soc. Am. **55**, 1427–1435 (1965).
9. A. Quirrenbach, "Phase referencing," in *Principles of Long Baseline Stellar Interferometry*, P. R. Lawson, ed. (Jet Propulsion Laboratory, 1999), Chap. 9, pp. 143–160.
10. R. C. Jennison, "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distribution of small angular extent," Mon. Not. R. Astron. Soc. **118**, 276–284 (1958).
11. T. A. Pauls, J. S. Young, W. D. Cotton, and J. D. Monnier, "A data exchange standard for optical (visible/IR) interferometry," Publ. Astron. Soc. Pac. **117**, 1255–1262 (2005).
12. G. Perrin, "The calibration of interferometric visibilities obtained with single-mode optical interferometers. Computation of error bars and correlations," Astron. Astrophys. **400**, 1173–1181 (2003).
13. C. Haniff, "Least-squares Fourier phase estimation from the modulo 2π bispectrum phase," J. Opt. Soc. Am. A **8**, 134–140 (1991).
14. G. Le Besnerais, S. Lacour, L. M. Mugnier, E. Thiébaut, G. Perrin, and S. Meimon, "Advanced imaging methods for long-baseline optical interferometry," IEEE J. Sel. Top. Signal Process. (to be published).
15. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024–2036 (1989).
16. R. Nityananda and R. Narayan, "Maximum entropy image reconstruction—a practical non-information-theoretic approach," J. Astrophys. Astron. **3**, 419–450 (1982).
17. E. Thiébaut, P. J. V. Garcia, and R. Foy, "Imaging with Amber/VLTI: the case of microjets," Astrophys. Space Sci. **286**, 171–176 (2003).
18. J. Navaza, "Accurate solutions of the maximum entropy equations. Their impact on the foundations of direct methods," in *Crystallographic Computing 5: From Chemistry to Biology*, Oxford Internatiural Union of Crystallography Book Series (Oxford University Press, 1991), pp. 317–323.
19. Y. Bar-Shalom and X.-R. Li, *Multitarget–Multisensor Tracking: Principles and Techniques* (YBS Publishing, 1995).
20. S. Meimon, L. M. Mugnier, and G. Le Besnerais, "Convex approximation of the likelihood in optical interferometry," J. Opt. Soc. Am. A **22**, 2348–2356 (2005).
21. L. M. Mugnier, G. Le Besnerais, and S. Meimon, "Inversion in optical imaging through atmospheric turbulence," in *Bayesian Approach to Inverse Problems*, J. Idier, ed., Digital Signal and Image Processing Series (ISTE/Wiley, 2008), Chap. 10, pp. 243–283.
22. A. Lannes, "Integer ambiguity resolution in phase closure imaging," J. Opt. Soc. Am. A **18**, 1046–1055 (2001).
23. P. R. Lawson, W. D. Cotton, C. A. Hummel, J. D. Monnier, M. Zhao, J. S. Young, H. Thorsteinsson, S. C. Meimon, L. Mugnier, G. Le Besnerais, E. Thiébaut, and P. G. Tuthill, "An interferometry imaging beauty contest," Proc. SPIE **5491**, 886–899 (2004).
24. J. A. Benson, C. A. Hummel, and D. Mozurkewich, "Simultaneous 6-station observations with the NPOI," Proc. SPIE **4838**, 358–368 (2003).
25. F. P. Schloerb, J.-P. Berger, N. P. Carleton, P. Hagenauer, P. Y. Kern, P. R. Labeye, M. G. Lacasse, F. Malbet, R. Millan-Gabet, J. D. Monnier, M. R. Pearlman, E. Pedretti, K. Rousselet-Perraut, S. D. Ragland, P. A. Schuller, W. A. Traub, and G. Wallace, "IOTA: recent science and technology," Proc. SPIE **6268**, 62680I (2006).
26. G. Perrin, S. Ridgway, B. Mennesson, W. Cotton, J. Woillez, T. Verhoelst, P. Schuller, V. Coudé du Foresto, W. Traub, R. Millan-Galbet, and M. Lacasse, "Unveiling Mira stars behind the molecules. Confirmation of the molecular layer model with narrow band near-infrared interferometry," Astron. Astrophys. **426**, 279–296 (2004).
27. S. Lacour, "Imagerie des étoiles évoluées par interférométrie. Réarrangement de pupille," Ph.D. thesis (University of Paris VI, 2007).