Statistical Behavior of Joint Least-Square Estimation in the Phase Diversity Context

Jérôme Idier, Laurent Mugnier, and Amandine Blanc

Abstract—The images recorded by optical telescopes are often degraded by aberrations that induce phase variations in the pupil plane. Several wavefront sensing techniques have been proposed to estimate aberrated phases. One of them is *phase diversity*, for which the joint least-square approach introduced by Gonsalves *et al.* is a reference method to estimate phase coefficients from the recorded images. In this paper, we rely on the asymptotic theory of Toeplitz matrices to show that Gonsalves' technique provides a consistent phase estimator as the size of the images grows. No comparable result is yielded by the classical *joint maximum likelihood* interpretation (*e.g.*, as found in the work by Paxman *et al.*). Finally, our theoretical analysis is illustrated through simulated problems.

Index Terms—Error analysis, least-squares methods, optical image processing, parameter estimation, phase diversity, statistics, Toeplitz matrices.

I. INTRODUCTION

THE images recorded by optical telescopes are often degraded by aberrations that induce phase variations in the pupil plane. In the case of ground telescopes, atmospheric turbulence is typically responsible for such phase aberrations. Imperfections of the optical system are another important source of errors, most of the latter being static while the former evolves with atmospheric turbulence.

Phase aberration is an ackowledged cause of degradation of the optical transfer function (OTF). The situation becomes far more favorable if the aberrated phases can be inferred and compensated. Several wavefront sensing techniques have been proposed to allow phase estimation. One of them is Gonsalves' *phase diversity* technique [1], [2]. It consists in the simultaneous acquisition of the usual focal plane image and of (at least) one additional image with a known defocus. Then the aberrations are numerically estimated using the information brought by the set of measured images.

Joint least-square (JLS) estimation of the aberrations and the observed object has been proposed by Gonsalves [1], [2], and it has since become the reference phase diversity technique. In

[3], a statistical interpretation is given: The JLS estimate can be viewed as a joint maximum likelihood (ML) solution under the assumption of additive white Gaussian noise. In the present paper, our main objective is to examine its asymptotical statistical properties w.r.t. aberration estimation. By "asymptotical," we refer to a situation where the number of data points grows to infinity, and preferentially to the case where the size of the acquired images is arbitrarily large. Such a situation is clearly formal, *i.e.*, it is not aimed to be reproduced in practice.[1] However, the asymptotical behavior of the solution may bring meaningful information about its behavior in realistic situations.

As the number of data points grows to infinity, the optimality of ML estimation is granted in a wide theoretical framework. Unfortunately, the usual JLS solution to the phase estimation problem does not pertain to this framework, since the number of unknowns (*i.e.*, both the aberrated phase parameters *and the object*) increases with the number of observations. It rather corresponds to an approach studied by Little and Rubin [5]. According to their conclusions, this approach is not generally reliable from the statistical viewpoint, especially when the relative proportion of unknowns does not go to zero as the size of the data set increases.

In [6] and [7], a true ML estimate in the sense of [5] is proposed for the aberration parameters in the context of phase diversity: The unknown object is treated as a nuisance parameter, which means that it is integrated out to form the likelihood with respect to phase parameters. In contrast with the JLS solution, the theoretical asymptotical optimality of such a solution is granted.

Yet, it has been established by practical evidence that the behavior of the JLS solution to the phase parameter estimation is globally satisfactory. It is the aim of the present paper to examine the statistical properties of JLS type solutions more specifically. To our best knowledge, this is a fully open question, since the only few contributions devoted to statistical analysis of phase diversity imaging assume that the source object is known [8], [9].

Our main result is that the JLS solution possesses the essential features of a *minimum contrast estimator* [10, Section 3.2]. As such, it is a consistent estimator (*i.e.*, it converges toward the true value as the size of the data set increases).

II. DATA MODEL

Let $\boldsymbol{i}_N = \{i(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\}$ correspond to a focused image measured on a square grid $\Omega_N = \{-N, \ldots, N\} \times$

---

[1]Physical phenomena should then be taken into account, such as anisoplanatism in the case of extended objects observed through turbulence [4].

$\{-N, \ldots, N\}$ of size $Q_N \times Q_N$, with $Q_N = 2N + 1$. In the isoplanatic patch of the telescope, it is obtained by noisy convolution of the object o with the focused point-spread function (PSF) h: $\forall \boldsymbol{x} \in \Omega_N$

$$i(\boldsymbol{x}) = (h \star o)(\boldsymbol{x}) + n(\boldsymbol{x}) = \int_{\mathbb{R}^2} h(\boldsymbol{y})o(\boldsymbol{x} - \boldsymbol{y})d\boldsymbol{y} + n(\boldsymbol{x}) \quad (1)$$

where $\star$ denotes convolution between functions of $\mathbb{R}^2$ and n corresponds to observation noise. The PSF is given by $h = |z|^2$, with

$$\forall \boldsymbol{x} \in \mathbb{R}^2, \quad z(\boldsymbol{x}) = \int_{\mathbb{R}^2} Z(\boldsymbol{\nu})e^{2i\pi\langle x, \boldsymbol{\nu}\rangle}d\boldsymbol{\nu} \quad (2)$$

where $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^2$ and $Z = Ae^{i\phi}$. The aperture function $A = |Z|$ is known and of limited spatial extent, and ϕ is the unknown aberrated phase function. Following [3] and others, we shall consider a finite linear decomposition for ϕ: $\phi = \sum_{j=1}^{J} \alpha_j \psi_j$. Typically, $\{\psi_j\}$ is a set of Zernike polynomials [11].

Let us remark that the OTF (*i.e.*, the Fourier transform of the PSF) has a finite support since A is of limited extent. In the sequel, it will be assumed that the OTF vanishes outside the square $\Gamma = [-1/2, 1/2] \times [-1/2, 1/2]$, so that no aliasing effect occurs. Then, it is possible to cast the observation model within a fully discrete framework

$$\forall \boldsymbol{x} \in \Omega_N, \quad i(\boldsymbol{x}) = (h * o)(\boldsymbol{x}) + n(\boldsymbol{x}) \quad (3)$$

where $*$ denotes convolution between functions of $\mathbb{Z}^2$, *i.e.*, $(h * o)(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \mathbb{Z}^2} h(\boldsymbol{y})o(\boldsymbol{x} - \boldsymbol{y})$.

In order to allow practical computations, the convolution product in (3) must be restricted to finite arrays. Depending on the assumption made at the boundaries, several alternatives are possible, none of which being exact unless the object has a known finite support. In particular, the *periodic boundary* condition corresponds to *cyclic convolution*. For the sake of computational simplicity, it is the most commonly adopted approximation.

In all cases, the approximate observation model can then be described in a vector-matrix formulation using lexicographical orderings of the image and of the object [12]. Let $\mathrm{col}\{\boldsymbol{v}_N\}$ denote a column vector of length Q_N^2 corresponding to an array $\boldsymbol{v}_N = \{v(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\}$ scanned in lexicographical order. Then, all approximate observation models read

$$\mathrm{col}\{\boldsymbol{i}_N\} \simeq \mathbf{H}_{NM}\mathrm{col}\{\boldsymbol{o}_M\} + \mathrm{col}\{\boldsymbol{n}_N\} \quad (4)$$

where $\mathbf{H}_{NM}$ is a $Q_N^2 \times Q_M^2$ *Toeplitz-block-Toeplitz* (TBT) matrix with blocks of size $Q_N \times Q_M$, where M depends on the adopted approximation. In the case of cyclic convolution, $M = N$ and $\mathbf{H}_{NM} = \mathbf{H}_N$ is a square matrix with a *circulant-block-circulant* (CBC) structure: $\mathbf{H}_N \mathrm{col}\{\boldsymbol{o}_N\} = \mathrm{col}\{\boldsymbol{h}_N \circledast \boldsymbol{o}_N\}$, where the symbol $\circledast$ denotes cyclic convolution between two-dimensional (2-D) finite arrays. The most usual approximation corresponds to $\boldsymbol{h}_N = |\boldsymbol{z}_N|^2$ where $\boldsymbol{z}_N$ is obtained from Z by inverse 2-D discrete Fourier transform (DFT)

$$\boldsymbol{z}_N(\boldsymbol{x}) = Q_N^{-2} \sum_{y \in \Omega_N} Z\left(\frac{\boldsymbol{y}}{Q_N}\right) e^{2i\pi\langle x, y\rangle/Q_N}$$

which is practically implemented using fast Fourier transform techniques.

In defocused planes, the observation model (1)–(3) generalizes under the following form: $\forall k = 2, \ldots, K$

$$\forall \boldsymbol{x} \in \Omega_N, \quad i_k(\boldsymbol{x}) = (h_k \star o)(\boldsymbol{x}) + n_k(\boldsymbol{x}),$$
$$= (h_k * o)(\boldsymbol{x}) + n_k(\boldsymbol{x}) \quad (5)$$

$$\forall \boldsymbol{x} \in \mathbb{R}^2, \quad h_k(\boldsymbol{x}) = |z_k(\boldsymbol{x})|^2 \quad (6)$$

$$z_k(\boldsymbol{x}) = \int_\Gamma Z_k(\boldsymbol{\nu})e^{2i\pi\langle x, \boldsymbol{\nu}\rangle}d\boldsymbol{\nu} \quad (7)$$

with $Z_k = Ae^{i\phi_k}$, where $\phi_k - \phi = \theta_k$ are known phase increments. The usual approximation of cyclic convolution corresponds to[2]

$$\boldsymbol{i}_{kN} \simeq \mathbf{H}_{kN}\boldsymbol{o}_N + \boldsymbol{n}_{kN} = \boldsymbol{h}_{kN} \circledast \boldsymbol{o}_N + \boldsymbol{n}_{kN} \quad (8)$$

where $\forall \boldsymbol{x} \in \Omega_N$

$$\boldsymbol{h}_{kN}(\boldsymbol{x}) = |\boldsymbol{z}_{kN}(\boldsymbol{x})|^2 \quad (9)$$

with

$$\boldsymbol{z}_{kN}(\boldsymbol{x}) = Q_N^{-2} \sum_{\boldsymbol{y} \in \Omega_N} Z_k\left(\frac{\boldsymbol{y}}{Q_N}\right) e^{2i\pi\langle \boldsymbol{x}, \boldsymbol{y}\rangle/Q_N}. \quad (10)$$

III. JOINT LEAST-SQUARE APPROACH

Let us introduce the following penalized least-square criterion:

$$\mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}_N) = \sum_{k=1}^{K} \|\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N\|^2 + \mu_N\|\mathbf{F}_N\boldsymbol{o}_N\|^2 \quad (11)$$

where conventionaly, the index value $k = 1$ refers to focused quantities (*e.g.*, $\boldsymbol{i}_1 \equiv \boldsymbol{i}$). In what follows, the default range of summation on image indices is $1, \ldots, K$.

Within the probabilistic framework, choosing the above penalizing term corresponds to assuming that the object $\boldsymbol{o}_N$ is a centered random vector with a covariance matrix proportional to $(\mathbf{F}_N^t\mathbf{F}_N)^{-1}$ (provided that $\mathbf{F}_N^t\mathbf{F}_N$ is actually invertible). Let us also remark that the original approach introduced by Gonsalves corresponds to $K = 2$ and $\mu_N = 0$. Choosing $\mathbf{F}_N$ as filtering low frequencies out and setting μ_N to some strictly positive value has a favorable regularizing effect on the quality of the restored object [7, Fig. 4(b)]. At low signal-to-noise ratio, it is also favorable with respect to phase estimation [7, Fig. 4(a)].

Finding a joint minimizer $(\hat{\boldsymbol{\alpha}}_N, \hat{\boldsymbol{o}}_N)$ of $\mathcal{J}_N$ can be partially simplified as follows [2], [3], [13]. For any $\boldsymbol{\alpha}$, minimizing $\mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}_N)$ as a function of $\boldsymbol{o}_N$ amounts to solving a quadratic programming problem. The set of solutions is characterized by the normal equation

$$\mathbf{R}_N\boldsymbol{o}_N = \sum_k \mathbf{H}_{kN}^t\boldsymbol{i}_{kN} \quad (12)$$

where

$$\mathbf{R}_N = \sum_k \mathbf{H}_{kN}^t\mathbf{H}_{kN} + \mu_N\mathbf{F}_N^t\mathbf{F}_N$$

---

[2]Thereafter, the operation of lexicographical reordering $\mathrm{col}\{\cdot\}$ is understood whenever unambiguous.

is the associated normal matrix. In particular, $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha}) = \mathbf{R}_N^+ \sum_k \mathbf{H}_{kN}^t \boldsymbol{i}_{kN}$ is the *generalized inverse* solution (matrix $\mathbf{R}_N^+$ denotes the pseudo inverse of $\mathbf{R}_N$) [14]. If $\mathbf{R}_N$ is full rank, then $\mathbf{R}_N^+ = \mathbf{R}_N^{-1}$ and $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})$ is the unique solution of (12). Let

$$\mathcal{K}_N(\boldsymbol{\alpha}) = \min_o \mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}) = \mathcal{J}_N(\boldsymbol{\alpha}, \hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})).$$

Basic algebraic manipulations yield

$$\mathcal{K}_N(\boldsymbol{\alpha}) = \mathcal{K}_N' - \mathcal{K}_N''(\boldsymbol{\alpha}) \tag{13}$$

$$\mathcal{K}_N' = \sum_k \|\boldsymbol{i}_{kN}\|^2 \tag{14}$$

$$\mathcal{K}_N''(\boldsymbol{\alpha}) = \left(\sum_k \boldsymbol{i}_{kN}^t \mathbf{H}_{kN} \right) \mathbf{R}_N^+ \left(\sum_k \mathbf{H}_{kN}^t \boldsymbol{i}_{kN} \right). \tag{15}$$

Hence, in order to obtain $\hat{\boldsymbol{\alpha}}_N$, it suffices to maximize $\mathcal{K}_N''$, which only depends on the J unknown phase parameters $\boldsymbol{\alpha}$ through matrices $\mathbf{H}_{kN}$. Since no closed-form expression of the maximizer of $\mathcal{K}_N''$ is available, one must resort to some iterative optimization algorithm [3], [7], [13].

The case where matrices $\mathbf{H}_{kN}$ and $\mathbf{F}_N$ are CBC introduces further simplifications, since the eigenvalues of a CBC matrix correspond to the 2-D DFT of its first row [15]. When such approximations are adopted, it is numerically preferable (and formally equivalent according to Parseval identity), to maximize $\mathcal{K}_N''$ using quantities expressed in the Fourier domain [3], [7], [13].

IV. ASYMPTOTIC BEHAVIOR OF THE JLS SOLUTION

In this section, "true" quantities are denoted using a tilde, *e.g.*, $\tilde{o}$ and $\tilde{h}_k$ denote the true object and the true kth PSF, respectively.

An asymptotical study of the behavior of the JLS solution $\hat{\boldsymbol{\alpha}}_N$ needs to refer to a statistical framework. Here, we shall assume that the noise signals $\left\{ n_k(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{Z}^2 \right\}$ are white, centered, identically distributed, of same finite variance σ^2, and uncorrelated: $\forall \, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}^2, \forall \, k, \ell$

$$E\left[n_k(\boldsymbol{x}) \right] = 0 \tag{16}$$

$$E\left[n_k(\boldsymbol{x}) n_\ell(\boldsymbol{y}) \right] = \sigma^2 \delta_{xy} \delta_{k\ell} \tag{17}$$

where δ is the Kronecker delta symbol: $\delta_{xy} = 1$ if $x = y$, $\delta_{xy} = 0$ otherwise.

By "asymptotical," one could refer to at least three limiting situations

- high-SNR case : $\sigma \to 0$
- highly diverse case : $K \to +\infty$
- large size case : $N \to +\infty$.

This paper focuses on the last case, because it corresponds to a realistic situation (in the sense that N is usually much greater than J). Moreover, the other two cases can be studied in the usual ML framework, since they correspond to situations where the number of unknowns $Q_N^2 + J$ remains constant.

In the framework of *minimum contrast* estimation, one minimizes an objective function $\mathcal{L}_N(\boldsymbol{\alpha})$ that holds the following properties [10]:

C1) as $N \to +\infty$, $\mathcal{L}_N(\boldsymbol{\alpha})$ uniformly converges in probability toward a limiting function $\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$;

C2) $\mathcal{L}$ is a *contrast function relative to* $\tilde{\boldsymbol{\alpha}}$, *i.e.*, its minimum value as a function of $\boldsymbol{\alpha}$ is uniquely attained at $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$.

Under quite general regularity conditions, minimum contrast estimators are *weakly consistent* [10, Section 3.2.3], *i.e.*, the minimizer $\hat{\boldsymbol{\alpha}}_N$ of $\mathcal{L}_N(\boldsymbol{\alpha})$ converges in probability toward $\tilde{\boldsymbol{\alpha}}$

$$\forall \, \varepsilon > 0, \quad \lim_{N \to \infty} \Pr(\|\hat{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\| > \varepsilon) = 0$$

which will be noted $\text{plim}_{N \to \infty} \hat{\boldsymbol{\alpha}}_N = \tilde{\boldsymbol{\alpha}}$. Under additional conditions, one can also establish that $\hat{\boldsymbol{\alpha}}_N$ is asymptotically normally distributed around $\tilde{\boldsymbol{\alpha}}$ with a standard deviation proportional to $1/Q_N$ [10, Section 3.3.4].

Least-square estimation constitutes a fundamental case of contrast estimation. Minimizing $\mathcal{K}_N$ actually falls within the *nonlinear generalized least-square* (NLGLS) approach: $\mathcal{K}_N$ is a quadratic objective function of the data $\boldsymbol{i}_{kN}$, which are nonlinear functions of the unknowns $\boldsymbol{\alpha}$. Moreover, $\mathcal{K}_N$ is not merely a sum of squared residuals, hence the mention *"generalized"*.

Both theory and practice of least-square estimation are well documented, particularly in the field of econometrics. For instance, [16] provides a detailed review of asymptotical statistical properties of least-square estimation. Some contributions address problems (such as estimation in an *errors in variables* model [17]) that are structurally close to phase diversity estimation using the JLS approach. Yet, we have been unable to find directly applicable results to the phase diversity problem. Nonetheless, a tailormade statistical study does seem achievable within the NLGLS framework. In the present paper, we only outline the main conditions that lead to establish consistency.

The most important step is to check that the limiting behavior of $\mathcal{L}_N = \mathcal{K}_N / Q_N^2$ meets Conditions C1 and C2 related to minimum contrast estimation. As the image size increases, two phenomena must be taken into account to establish the limiting expression of $\mathcal{L}_N$.

On the one hand, the effect of approximating the convolution on finite arrays vanishes. This phenomenon can be mathematically studied using Gray's theory of *asymptotically equivalent* matrices [18].

Definition 1: Two series of square matrices $\mathbf{A}_n, \mathbf{B}_n$ of size $n \times n$ are said asymptotically equivalent (which is denoted $\mathbf{A}_n \sim \mathbf{B}_n$) if

- $\mathbf{A}_n, \mathbf{B}_n$ are uniformly bounded in strong norm (*i.e.*, their maximal singular value is uniformly bounded);
- $\lim_{n \to \infty} \|\mathbf{A}_n - \mathbf{B}_n\|_F^2 / n = 0$, where $\|\cdot\|_F$ is the Frobenius norm: $\|\mathbf{A}\|_F^2 = \text{Tr}\{\mathbf{A}^t \mathbf{A}\}$ ($\text{Tr}\{\cdot\}$ is the trace of a square matrix, *i.e.*, the sum of its diagonal elements, which is also the sum of its eigenvalues).

Specifically, important results establish the asymptotical equivalence between Toeplitz and circulant matrices [18], and between TBT and CBC matrices [15], [19].

On the other hand, the random behavior of noise signals is averaged, according to a large numbers effect. Actually, we will also have to consider the true object $\tilde{o}$ from a statistical viewpoint, the latter being considered as a second-order stationary random process.

To simplify the derivations, we only establish the expression of $\mathcal{L} = \lim_{N\to\infty} E[\mathcal{L}_N]$, implicitly assuming that $\mathcal{L}_N$ uniformly converges in probability toward $\mathcal{L}$ under appropriate additional hypotheses (at least, $\tilde{o}$ should be a *correlation-ergodic* stationary random process). In Section V, it is checked by simulations that this conjectured behavior is in good agreement with practice.

Theorem 1: Let us assume the following.

1) The true object $\tilde{o}$ is a second-order stationary random process, centered, with a stable[3] correlation function $\tilde{r}$. Let $\tilde{P}$ denote the power spectral density function of $\tilde{o}$: $\tilde{P}(\boldsymbol{\nu}) = \sum_{x \in Z^2} \tilde{r}(\boldsymbol{x}) e^{-2i\pi\langle\boldsymbol{x},\boldsymbol{\nu}\rangle}$.
2) The noise signals n_k fulfill (16) and (17) and they are uncorrelated with $\tilde{o}$.
3) The cyclic convolution approximation (8) has been adopted: $\mathbf{H}_{kN}\boldsymbol{o}_N = \boldsymbol{h}_{kN} \circledast \boldsymbol{o}_N$.
4) $\forall\, N$, $\mathbf{F}_N$ is CBC and there exists an impulse response $f = \left\{ f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{Z}^2 \right\}$ such that $\mathbf{F}_N\boldsymbol{o}_N = \boldsymbol{f}_N \circledast \boldsymbol{o}_N$ with $\boldsymbol{f}_N = \{ f(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N \}$.
5) $\lim_{N\to\infty} \mu_N = \mu$ and $\forall\, \boldsymbol{\nu} \in \Gamma$, $\boldsymbol{\alpha}$

$$D(\boldsymbol{\nu}) = \mu|\rho(\boldsymbol{\nu})|^2 + \sum_k |\gamma_k(\boldsymbol{\nu})|^2 > d > 0$$

where γ_k and ρ are the 2-D discrete time Fourier transforms (DTFT) of h_k and f, respectively

$$\gamma_k(\boldsymbol{\nu}) = \sum_{x \in Z^2} h_k(\boldsymbol{x}) e^{-2i\pi\langle\boldsymbol{x},\boldsymbol{\nu}\rangle}$$

$$\rho(\boldsymbol{\nu}) = \sum_{x \in Z^2} f(\boldsymbol{x}) e^{-2i\pi\langle\boldsymbol{x},\boldsymbol{\nu}\rangle}.$$

Then, $E[\mathcal{L}_N(\boldsymbol{\alpha})] = E[\mathcal{K}_N(\boldsymbol{\alpha})]/Q_N^2$ converges toward $\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = \mathcal{L}'(\tilde{\boldsymbol{\alpha}}) - \mathcal{L}''(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$, with

$$\mathcal{L}'(\tilde{\boldsymbol{\alpha}}) = K\sigma^2 + \int_\Gamma \tilde{P}(\boldsymbol{\nu}) \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu} \qquad (18)$$

$$\mathcal{L}''(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = \int_\Gamma \frac{\tilde{P}(\boldsymbol{\nu})}{D(\boldsymbol{\nu})} \left| \sum_k \gamma_k^*(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu}) \right|^2 d\boldsymbol{\nu}$$
$$+ \int_\Gamma \frac{\sigma^2}{D(\boldsymbol{\nu})} \sum_k |\gamma_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}. \qquad (19)$$

Proof: See Appendix A.

Remark 1: In the first assumption of Theorem 1, it would be more realistic to suppose that the object has a strictly positive mean value $\tilde{m}$. Such a modification introduces the following additional term to $\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$

$$\mathcal{L}''' = \tilde{m}^2 \sum_k \tilde{\gamma}_k^2(0) + \frac{\tilde{m}^2}{D(0)} \left(\sum_k \gamma_k(0) \tilde{\gamma}_k(0) \right)^2$$

which is a constant term since

$$\gamma_k(0) = \tilde{\gamma}_k(0) = \int_\Gamma A^2(\nu) d\boldsymbol{\nu}$$

does not depend on $\boldsymbol{\alpha}$. For this reason, we maintain $\tilde{m} = 0$ in the rest of the paper.

[3] For example, absolutely summable: $\sum_{\boldsymbol{x} \in Z^2} |\tilde{r}(\boldsymbol{x})| < \infty$.

Remark 2: According to Assumption 3, the range of Theorem 1 is restricted to cyclic convolution, although generalization to alternate boundary conditions seems possible.

Theorem 2: If $\mu = 0$, then $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ minimizes $\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = \mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}; \mu = 0)$ and the minimum value is $\mathcal{L}^0(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = (K - 1)\sigma^2$.

Proof: See Appendix B. ∎

Several remarks can be made concerning Theorem 2.

Remark 3: To benefit from the statistical property of Theorem 2, it is required that the regularization term in (11) asymptotically vanish. Alternately, in strict conformity with a Bayesian approach, one could rather choose μ and ρ according to $\forall\, \boldsymbol{\nu}$, $\mu|\rho(\boldsymbol{\nu})|^2 = \sigma^2/\tilde{P}(\boldsymbol{\nu})$. Then, $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})$ asymptotically corresponds to the Wiener solution. In this case, the identities of Appendix B yield $\mathcal{L}^W(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = K\sigma^2$ and

$$\mathcal{L}^W(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) \geq K\sigma^2 + \int_\Gamma \frac{\sigma^2}{D(\boldsymbol{\nu})} \sum_k \left(|\tilde{\gamma}_k(\boldsymbol{\nu})|^2 - |\gamma_k(\boldsymbol{\nu})|^2 \right) d\boldsymbol{\nu}.$$

Unfortunately, this does not allow to conclude that $\mathcal{L}^W(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}})$ is the minimum value of $\mathcal{L}^W(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$.

Remark 4: According to Theorem 2, it seems preferable not to regularize the criterion when the dataset is large enough. This theoretical result meets the conclusions drawn from simulated experiments: In the most favorable situations (such as [7, Fig. 4(b)]), the empirical mean squared error (MSE) $\|\hat{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\|^2$ is an increasing function of μ_N. In such favorable cases, the estimation variance is small, so the MSE is mainly due to bias. In less favorable situations (such as [7, Fig. 4(a)]), penalization also creates bias, but, at the same time, it has a favorable effect on variance. This is a classical situation of bias/variance compromise. In Section V, the same phenomenon is reexamined as a function of the size of the dataset.

Remark 5: If (5) holds without aliasing, then the OTF $\gamma_k(\boldsymbol{\nu})$ necessarily vanishes on the boundaries of the square Γ. As a consequence, Assumption 5 of Theorem 1 holds only if $\mu > 0$, which contradicts the assumption of Theorem 2. Choosing strictly positive, possibly very small, values of μ is a satisfying option in practice. From a more theoretical viewpoint, a possibility to alleviate Assumption 5 is to modify the original least-square criterion. More precisely, let us replace each fidelity-to-data term $\|\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N\|^2$ by a generalized least-square term $\|\mathbf{G}_N(\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N)\|^2$, where $\mathbf{G}_N$ is a CBC matrix defined from an impulse response g that cancels high frequencies out (at least those that violate the condition $\sum_k |\gamma_k(\boldsymbol{\nu})|^2 > d > 0$). Incorporating $\mathbf{G}_N$ into our previous calculations, we are led to the following conclusions.

- Expression (13) of criterion $\mathcal{K}$ is still available provided that $\boldsymbol{i}_{kN}$ be replaced by $\mathbf{G}_N\boldsymbol{i}_{kN}$ in (14) and (15). As a particular case, $\mu_N = 0$ provides Löfdahl and Scharmer's solution based on *noise filtering* [20, Section 2.3]. A comparable data filtering procedure is proposed in [21].
- Theorem 2 still holds, provided that

$$\mathcal{L}^0(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = (K - 1)\sigma^2 \|g\|^2.$$

- Let us assume $\mu = 0$. Then Assumption 5 in Theorem 1 can be replaced by

5'). Let β be the DTFT of g. There exists $d > 0$ such that $\forall \, \boldsymbol{\nu} \in \Gamma, \forall \, \boldsymbol{\alpha}$

$$\sum_k |\gamma_k(\boldsymbol{\nu})|^2 \le d \Longrightarrow \beta(\boldsymbol{\nu}) = 0.$$

Finally, let us seek conditions under which $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ is actually the *unique* minimizer of

$$\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = (K-1)\sigma^2$$

$$+ \int_\Gamma \tilde{P}(\boldsymbol{\nu}) \left(\sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 - \frac{\left| \sum_k \gamma_k^*(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu}) \right|^2}{\sum_k |\gamma_k(\boldsymbol{\nu})|^2} \right) d\boldsymbol{\nu}.$$

Any value of $\boldsymbol{\alpha}$ that cancels the integral part of $\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$ is obviously a minimizer, by necessary and sufficient condition. Then, equivalently, $(\gamma_1(\boldsymbol{\nu}), \ldots, \gamma_K(\boldsymbol{\nu}))$ and $(\tilde{\gamma}_1(\boldsymbol{\nu}), \ldots, \tilde{\gamma}_K(\boldsymbol{\nu}))$ are colinear for all $\boldsymbol{\nu}$, *i.e.*,

$$\exists \lambda \text{ such that } \forall \, k, \boldsymbol{\nu}, \quad \gamma_k(\boldsymbol{\nu}) = \lambda(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu})$$

$$\Longleftrightarrow \exists g \text{ such that } \forall \, k, \quad h_k = g * \tilde{h}_k. \quad (20)$$

Let us assume that there exists a filter g such that (20) holds for some $\boldsymbol{\alpha} \ne \tilde{\boldsymbol{\alpha}}$. This means that we are facing a strong identifiability problem: The two solutions $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ and $(\boldsymbol{\alpha}, g^{-1} * \tilde{o})$ are not distinguishable from each other on the basis of the data, whatever the size of measured images and whatever the adopted method of estimation. Such a situation happens if the phase diversity functions θ_k are not appropriately chosen, *e.g.*, [22, Appendix B].

- From only one measured image $(K = 1)$, the sign of the symmetric component of ϕ (*i.e.*, ϕ_s with $\phi_s(\boldsymbol{\nu}) = (\phi(\boldsymbol{\nu}) + \phi(-\boldsymbol{\nu}))/2$) is not identifiable.
- The same undeterminacy holds if the phase diversity functions θ_k are chosen antisymmetric ($\theta_k(-\boldsymbol{\nu}) = -\theta_k(\boldsymbol{\nu})$, $\forall \, \boldsymbol{\nu}$). This does not occur in practice since defocus corresponds to $\theta_k(\boldsymbol{\nu}) \propto d_k \|\boldsymbol{\nu}\|^2$, where d_k is the kth defocus distance.
- In any case, the couple $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ is only identifiable up to an arbitrary spatial shift $g(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$ (here δ is the Dirac delta function), *i.e., tilt coefficients* are not identifiable [7].
- From only one measured image $(K = 1)$, the sign of the symmetric component of ϕ (*i.e.*, ϕ_s with $\phi_s(\boldsymbol{\nu}) = (\phi(\boldsymbol{\nu}) + \phi(-\boldsymbol{\nu}))/2$) is not identifiable.
- The same undeterminacy holds if the phase diversity functions θ_k are chosen antisymmetric ($\theta_k(-\boldsymbol{\nu}) = -\theta_k(\boldsymbol{\nu})$, $\forall \, \boldsymbol{\nu}$). This does not occur in practice since defocus corresponds to $\theta_k(\boldsymbol{\nu}) \propto d_k \|\boldsymbol{\nu}\|^2$, where d_k is the kth defocus distance.
- In any case, the couple $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ is only identifiable up to an arbitrary spatial shift $g(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$ (here δ is the Dirac delta function), *i.e., tilt coefficients* are not identifiable [7].

V. SIMULATION STUDY

A. Conditions of Simulation

This section proposes an empirical study of the statistical behavior of estimated phase coefficients as a function of the size of the observed images.

TABLE I
VALUES OF ZERNIKE COEFFICIENTS (EXPRESSED IN RADIANS) USED TO SIMULATE THE FOCUSED PSF

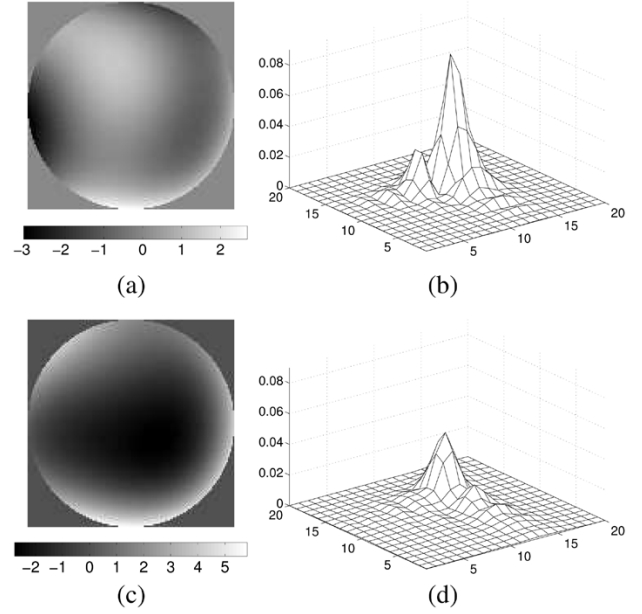| j	1	2	3	4	5	6	7
α_j	0	0	0	-0.2	0.3	-0.45	0.4
j	8	9	10	11	12	13	14
α_j	0.3	-0.25	0.35	0.2	0.1	0.05	-0.05
j	15	16	17	18	19	20	21
α_j	0.05	0.02	0.01	-0.01	-0.02	0.01	0.01

Fig. 1. (a) Aberrated phase $\phi_1 = \phi = \sum_{j=1}^J \alpha_j \psi_j$ in radians, where α_j are given in Table I. (b) Central 20×20 part of the resulting PSF h_1 on $\mathbb{Z}^2$. (c) Aberrated phase in radians after defocus $\phi_2 = \phi + \theta_2$. (d) central 20×20 part of the resulting PSF h_2. Both h_1 and h_2 almost vanish outside a central square of 20×20 pixels.

Following [7], we have simulated a focused PSF h_1 using the first 21 Zernike polynomials with coefficients given by Table I [see Fig. 1(a) and (b)]. We have also simulated one defocused PSF h_2 using $\theta_2 = \delta \alpha_4 \psi_4$, with $\delta \alpha_4 = 1.814$ radians [see Fig. 1(c) and (d)].

On the other hand, we have selected two different objects o.

- "`Noise`": The object is a Gaussian white noise sampled on a 512×512 grid.
- "`Earth`": The object is an Earth view sampled on the same 512×512 grid depicted on Fig. 2(a).

Couples of observed images of size 512×512 have been obtained using the approximate model (8), only the central 256×256 part of them being considered afterwards in order to get rid of the effect of cyclic convolution [see Fig. 2(b) and (c), respectively]. Finally, images have been corrupted by realizations of white Gaussian noise with a realistic signal to noise ratio of 100 dB.

In the sequel, $\boldsymbol{\alpha} = [\alpha_4, \ldots, \alpha_{21}]$, *i.e.*, the first three coefficients have not been estimated.

- The *piston* coefficient α_1 is a constant added to the phase and has no influence on the shape of the PSF.
- The *tilt* coefficients α_2, α_3 introduce a shift in the image that is of no importance for extended object.
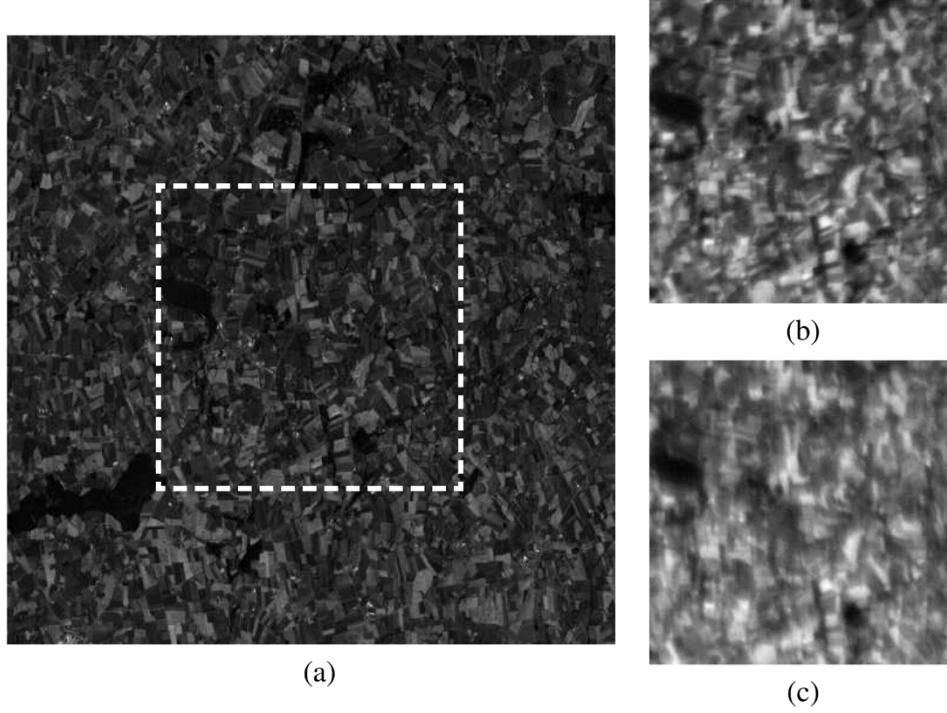
Fig. 2.   (a) The 512×512 Earth view used to simulate the extended object called "`Earth`." The original source is an image taken by satellite SPOT3, which has been downsampled by a factor two in both directions—Copyright CNES/Distribution SPOT IMAGE. (b) Central 256×256 part of noiseless focused image. (c) Central 256×256 part of noiseless defocused image.

By mean of Monte-Carlo simulations using several independent couples of realizations of noise, it is possible to evaluate the statistical performance of $\hat{\boldsymbol{\alpha}}_N$ for different values of N. Here, we have evaluated and displayed the following quantities for $R = 50$ independent realizations of noise.

- The Euclidian norm of the empirical bias vector (henceforth referred as the bias of $\hat{\boldsymbol{\alpha}}_N$, for sake of brevity): $b_N \triangleq \|\check{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\|$ where $\check{\boldsymbol{\alpha}}_N = \sum_{r=1}^{R} \hat{\boldsymbol{\alpha}}_N^{(r)}/R$.
- The Euclidian norm of the vector of empirical standard deviations (henceforth referred as the standard deviation of $\hat{\boldsymbol{\alpha}}_N$)

$$\sigma_N \triangleq \left(\frac{1}{R} \sum_{r=1}^{R} \|\hat{\boldsymbol{\alpha}}_N^{(r)} - \check{\boldsymbol{\alpha}}_N\|^2 \right)^{(1/2)} .$$

- The square root of the empirical MSE (henceforth referred as RMSE) $r_N \triangleq \sqrt{b_N^2 + \sigma_N^2}$.

B. "`Noise`" Object

Fig. 3(a)–(c) deal with the "`Noise`" object. They, respectively, depict the bias b_N, standard deviation σ_N and RMSE r_N of $\hat{\boldsymbol{\alpha}}_N$ as functions of the regularization parameter μ_N. In the "`Noise`" case, maximization of (15) has been considered with $\mathbf{F}_N$ equal to the identity matrix and under the usual CBC approximation for $\mathbf{H}_{1N}$ and $\mathbf{H}_{2N}$. Three nested images have been tested, of size $Q_N \times Q_N$, with $Q_N \in \{64, 128, 256\}$.

On the one hand, processing a larger image appears favorable in terms of bias [Fig. 3(a)]. The reason is that the effect of the CBC approximation becomes negligible for large size images.

However, the relative improvement is more substantial for small values of μ_N. This empirical observation fully meets the conclusions of Section IV: As Q_N grows, a vanishing series of μ_N is required to get an asymptotically unbiased estimator $\hat{\boldsymbol{\alpha}}_N$.

On the other hand, the standard deviation of $\hat{\boldsymbol{\alpha}}_N$ is a decreasing function of μ_N [Fig. 3(b)]. It also decreases with Q_N, and it is important to notice that the corresponding decreasing rate is rather independent of μ_N. This is not surprising since random fluctuations are averaged whatever the value of μ_N.

As a global consequence in terms of bias/variance compromise, the minimizer of the RMSE r_N shifts leftward as Q_N grows [Fig. 3(c)]. The minimum value is $\min_{\mu_N} r_N = 0.2116$, 0.1052, and 0.0476 for $Q_N = 64$, 128, and 256, respectively: It roughly decreases proportionaly to $1/Q_N$, although other values of Q_N should be tested to assess the actual decrease rate.

C. "`Earth`" Object

Fig. 4(a)–(c) deal with the "`Earth`" object. Maximization of (15) has been considered in the same conditions as in Section V-B, provided that $\mathbf{F}_N$ has been deduced from a power spectral density model with parameters fitted using the true object (see [7, Eq. (13)]).

Results depicted on Figs. 3 and 4 are comparable, except that the bias reaches much larger values in the present case, even for the largest size of image. This is a consequence of *edge effects* due to the adopted cyclic convolution approximation in the presence of extended, structured objects: Nonrealistic sizes of images should be processed to get statistically meaningfull estimates of $\boldsymbol{\alpha}$.
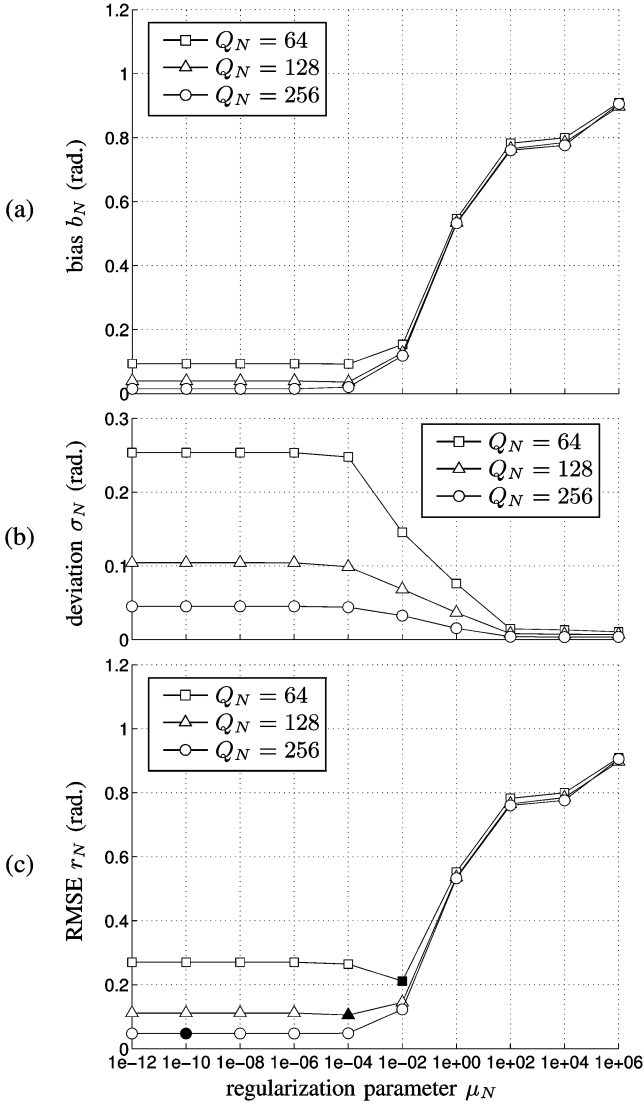
Fig. 3. (a) Bias b_N. (b) Standard deviation σ_N. (c) RMSE r_N of $\hat{\alpha}_N$ in the "`Noise`" case, *i.e.*, the true object is the realization of a 2-D Gaussian white noise, of size $Q_N \times Q_N$. For the RMSE, the minimum value of each curve is indicated by a blackened symbol.

To overcome this difficulty, Löfdahl and Scharmer introduced a *tapering* technique [20],[4] where:

- the observed images are windowed in order to apodize the edges;
- in the least-square criterion (11), the fidelity-to-data term is modified: The norm is only considered over a central, nonapodized part of the images.

We have applied this technique to the data simulated in the "`Earth`" case, using a 2-D modified Hamming apodization window with a central plateau of size $Q_N/2 \times Q_N/2$. Fig. 5(a) [compared to Fig. 4(a)] shows an impressive effect on bias. Even if the variance slightly increases, at least for large size images [Fig. 5(b) compared to Fig. 4(b)], the overall effect is largely favorable in term of RMSE [Fig. 5(c) compared to Fig. 4(c)].

Our theoretical study has been derived in the case where the images are not apodized by Löfdahl and Scharmer's tapering technique. However, given the results depicted on Fig. 5, it

<hr>

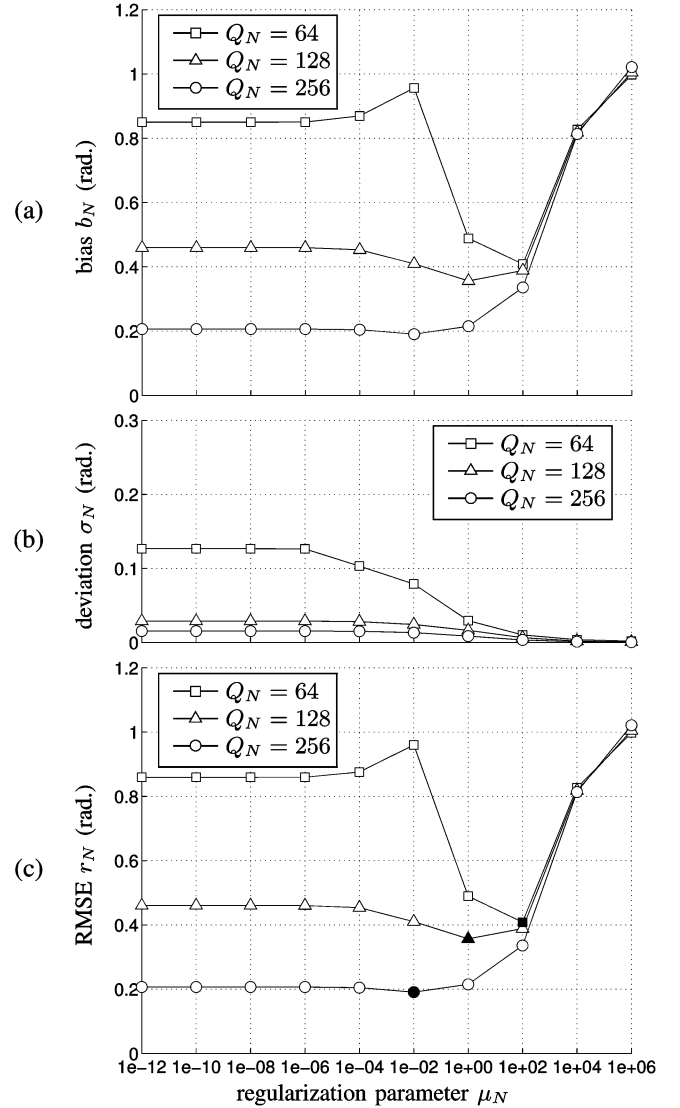[4]See also [23] for an alternative technique based on the use of a guard-band.

Fig. 4. Same as Fig. 3, except that the true object is the Earth view depicted at Fig. 2(a).

seems reasonable to expect that consistency results still hold, provided that the size of the apodized edges vanishes as Q_N grows.

VI. CONCLUSION

In this paper, we have studied some important statistical properties of the phase diversity technique introduced by Gonsalves [1], [2]. In particular, it has been shown that Gonsalves' technique is a *minimum contrast method*, with respect to phase estimation. As a consequence, it provided a consistent phase estimator as the size of the processed images grows (putting aside practical and physical limitations). No comparable result is yielded by the classical *joint ML* interpretation (*e.g.*, as found in [3]). In particular, the Gaussian character of the noise is not a prerequisite in our convergence study.

By simulation, we have checked that the JLS method behaves as predicted by theory in the case of extended objects. We have also observed that the edge effects due to cyclic convolution introduce a strong bias on phase estimation, that only slowly diminishes as the image size grows. Modified versions of the
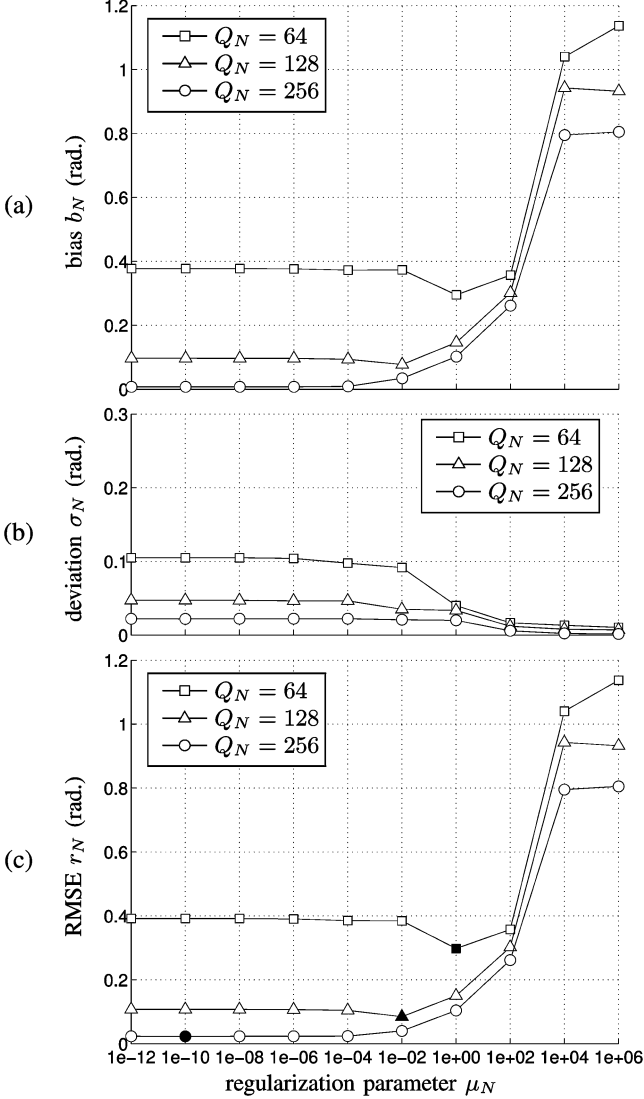
Fig. 5.　Same as Fig. 4, except that observed images have been apodized using the tapering technique introduced in [20].

JLS method are then required to recover meaningful estimates. We have more specifically considered the tapering technique proposed in [20], and we have empirically verified that the latter technique is still statistically convergent.

Finally, some alternative error metrics have been introduced to replace the criterion $\mathcal{K}_N$ induced by the JLS approach [24], [25], for the sake of faster computations. An interesting perspective would be to study if such alternative error metrics are still minimum contrast functions.

APPENDIX A

Proof of Theorem 1

Let us decompose the observed images according to $\boldsymbol{i}_{kN} = \boldsymbol{y}_{kN} + \boldsymbol{n}_{kN}$, where $\boldsymbol{y}_{kN} = \left\{ (\tilde{h}_k * \tilde{o})(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N \right\}$. Then, we have

$$\mathcal{K}'_N = \sum_k \|\boldsymbol{i}_{kN}\|^2$$
$$= \sum_k \left(\|\boldsymbol{n}_{kN}\|^2 + \|\boldsymbol{y}_{kN}\|^2 + 2\boldsymbol{y}_{kN}^t \boldsymbol{n}_{kN} \right).$$

Hence

$$E\left[\mathcal{L}'_N\right] = Q_N^{-2} E\left[\mathcal{K}'_N\right] = K\sigma^2 + Q_N^{-2} \sum_k E\left[\|\boldsymbol{y}_{kN}\|^2\right]$$

according to Assumptions 1 and 2. The second term converges toward the second term of (18), according to Parseval identity. As a consequence, we have $\lim_{N\to\infty} E\left[\mathcal{L}'_N\right] = \mathcal{L}'$.

The main part of the proof is to express $\lim_{N\to\infty} E\left[\mathcal{L}''_N\right]$, where $\mathcal{L}''_N = Q_N^{-2} \mathcal{K}''_N$. In particular, we need to examine the asymptotical behavior of matrices $\mathbf{H}_{kN}$. This is done in the following technical lemma.

Lemma 1: For all k, we have $\mathbf{H}_{kN} \sim \bar{\mathbf{H}}_{kN}$, where $\bar{\mathbf{H}}_{kN}$ is the $Q_N^2 \times Q_N^2$ CBC matrix whose eigenvalues are equally distributed on γ_k, i.e., $\{\gamma_k(\boldsymbol{x}/Q_N), \boldsymbol{x} \in \Omega_N\}$.

Proof: Let us introduce the DTFT of h_{kN}: $\forall \boldsymbol{\nu} \in \mathbb{R}^2$, $k = 1, \ldots, K$

$$\gamma_{kN}(\boldsymbol{\nu}) = \sum_{\boldsymbol{x} \in \Omega_N} h_{kN}(\boldsymbol{x}) e^{-2i\pi\langle\boldsymbol{\nu},\boldsymbol{x}\rangle}$$

so that $\mathbf{H}_{kN}$ is the CBC matrix whose eigenvalues are equally distributed on γ_{kN}. According to (9) and (10), we have

$$\gamma_{kN}(\boldsymbol{\nu}) = Q_N^{-2} \sum_{\boldsymbol{y} \in \Omega_N} Z_k^*\left(\boldsymbol{\nu} + \frac{\boldsymbol{y}}{Q_N}\right) Z_k(\boldsymbol{\nu}) \qquad (21)$$

where Z_k is extended over $\mathbb{R}^2$ in a periodic manner: $\forall \boldsymbol{\nu} \in \Gamma$, $\boldsymbol{x} \in \mathbb{Z}^2$, $Z_k(\boldsymbol{\nu} + \boldsymbol{x}) = Z_k(\boldsymbol{\nu})$. On the other hand, (6) and (7) also yield

$$\gamma_k(\boldsymbol{\nu}) = \int_\Gamma Z_k^*(\boldsymbol{\nu} + \boldsymbol{\eta}) Z_k(\boldsymbol{\eta}) d\boldsymbol{\eta}. \qquad (22)$$

It is clear from (21) and (22) that $\mathbf{H}_{kN}$ and $\bar{\mathbf{H}}_{kN}$ are uniformly bounded by $\max_{\boldsymbol{\nu} \in \Gamma} A^2(\boldsymbol{\nu})$. It is also clear from the same equations that $\gamma_{kN}(\boldsymbol{\nu})$ is a Rieman sum that uniformly converges toward $\gamma_k(\boldsymbol{\nu})$ when N grows to infinity

$$\forall \varepsilon > 0, \exists N_0 \text{ such that } \forall \boldsymbol{\nu}, N \geq N_0, \quad |\gamma_{kN}(\boldsymbol{\nu}) - \gamma_k(\boldsymbol{\nu})| \leq \varepsilon.$$

As a consequence, we have

$$\|\mathbf{H}_{kN} - \bar{\mathbf{H}}_{kN}\|_F^2 = \sum_{\boldsymbol{x} \in \Omega_N} \left| \gamma_{kN}\left(\frac{\boldsymbol{x}}{Q_N}\right) - \gamma_k\left(\frac{\boldsymbol{x}}{Q_N}\right) \right|^2$$
$$\leq Q_N^2 \varepsilon^2$$

for any ε, provided that N is large enough. Hence, $\lim_{N\to\infty} \|\mathbf{H}_{kN} - \bar{\mathbf{H}}_{kN}\|_F^2 / Q_N^2 = 0$. ∎

Let $\mathbf{U}_{k\ell N} = \mathbf{H}_{kN} \mathbf{R}_N^+ \mathbf{H}_{\ell N}^t$ and $\mathbf{S}_{k\ell N} = E\left[\boldsymbol{i}_{kN} \boldsymbol{i}_{\ell N}^t\right] = E\left[\boldsymbol{y}_{kN} \boldsymbol{y}_{\ell N}^t\right] + \sigma^2 \mathbf{I}_N$ (where $\mathbf{I}_N$ is the $Q_N^2 \times Q_N^2$ identity matrix). Given (15), we have[5]

$$E\left[\mathcal{K}''_N\right] = \sum_{k,\ell} E\left[\boldsymbol{i}_{kN}^t \mathbf{U}_{k\ell N} \boldsymbol{i}_{\ell N}\right] = \sum_{k,\ell} \text{Tr}\left\{\mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N}\right\}.$$

According to Lemma 1 and to [18, Theorem 2.1], each matrix $\mathbf{U}_{k\ell N}$ is asymptotically equivalent to another CBC matrix, whose eigenvalues are equally distributed on

$$U_{k\ell}(\boldsymbol{\nu}) = \frac{\gamma_k^*(\boldsymbol{\nu})\gamma_\ell(\boldsymbol{\nu})}{D(\boldsymbol{\nu})}, \quad \boldsymbol{\nu} \in \Gamma.$$

Let us remark that Assumption 5 is needed here to ensure that $\mathbf{R}_N^+$ is uniformly bounded in strong norm.

---

[5]Recall that $\text{Tr}\left\{\mathbf{A}^t \mathbf{B}\right\} = \text{Tr}\left\{\mathbf{B}\mathbf{A}^t\right\}$ for any two matrices $\mathbf{A}$, $\mathbf{B}$ of same dimensions.

On the other hand, $\mathbf{S}_{k\ell N}$ is an intercorrelation matrix, which is TBT since $\tilde{o}$ is stationary. Moreover, $\mathbf{S}_{k\ell N}$ is generated by the crosscorrelation sequence $\tilde{s}_{kl}$ defined by

$$\tilde{s}_{kl}(\boldsymbol{x}) = \left(\tilde{h}_k^{(-)} * \tilde{h}_\ell * \tilde{r}\right)(\boldsymbol{x}) + \sigma^2 \delta_{k\ell} \delta_{x\mathbf{0}}$$

where $\tilde{h}_k^{(-)}$ is the reversed version of $\tilde{h}_k$: $\tilde{h}_k^{(-)}(\boldsymbol{x}) = \tilde{h}_k(-\boldsymbol{x})$. The sequence $\tilde{s}_{kl}$ is stable for all k, ℓ since

- $\tilde{r}$ is stable by assumption;
- $\forall k$, $\tilde{h}_k$ is stable since given (6) and (7), we have the following Parseval identity

$$\sum_{x \in Z^2} |\tilde{h}_k(\boldsymbol{x})| = \sum_{x \in Z^2} |\tilde{z}_k(\boldsymbol{x})|^2 = \int_\Gamma A^2(\boldsymbol{\nu}) d\boldsymbol{\nu}$$

- convolution products and sums of stable functions are stable.

Thus, according to [18, Lemma 4.6][6], $\mathbf{S}_{k\ell N}$ is asymptotically equivalent to a CBC matrix whose eigenvalues are equally distributed on $S_{k\ell}(\boldsymbol{\nu}), \boldsymbol{\nu} \in \Gamma$, where

$$S_{k\ell}(\boldsymbol{\nu}) = \tilde{\gamma}_k^*(\boldsymbol{\nu}) \tilde{\gamma}_\ell(\boldsymbol{\nu}) \tilde{P}(\boldsymbol{\nu}) + \sigma^2 \delta_{k\ell}.$$

Finally, according to [18, Theorem 2.1], the matrix product $\mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N}$ is also asymptotically CBC, with eigenvalues equally distributed on $S_{k\ell}^*(\boldsymbol{\nu}) U_{k\ell}(\boldsymbol{\nu})$, so we get the following converging Riemann sum

$$\lim_{N \to \infty} Q_N^{-2} \sum_{k,\ell} \text{Tr} \left\{ \mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N} \right\} = \int_\Gamma \sum_{k,\ell} S_{k\ell}^*(\boldsymbol{\nu}) U_{k\ell}(\boldsymbol{\nu}) d\boldsymbol{\nu}$$
$$= \mathcal{L}''(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}).$$

APPENDIX B

Proof of Theorem 2

Evaluation of $\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}})$ is straightforward

$$\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = K\sigma^2 + \int_\Gamma \frac{\mu |\rho(\boldsymbol{\nu})|^2 \tilde{P}(\boldsymbol{\nu}) - \sigma^2}{\tilde{D}(\boldsymbol{\nu})} \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}$$

which takes an extremely simple form when $\mu = 0$

$$\mathcal{L}^0(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = (K - 1)\sigma^2.$$

Furthermore, let us show that $(K - 1)\sigma^2$ is actually the minimum value of $\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$. By Cauchy–Schwartz inequality

$$\forall \boldsymbol{\nu} \in \Gamma, \quad \left| \sum_k \gamma_k^*(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu}) \right|^2 \leq \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 \sum_k |\gamma_k(\boldsymbol{\nu})|^2$$

[6] For sake of correctness, Gray's asymptotical result only applies to Toeplitz, not necessarily Hermitian, matrices. In extensions to TBT matrices found in [15], [19], only the Hermitian case is considered. Here, we shall admit that Gray's result extends to TBT matrices, not necessarily Hermitian.

which allows to deduce from (18) and (19) that $\forall \boldsymbol{\alpha}$

$$\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) \geq (K - 1)\sigma^2$$
$$+ \mu \int_\Gamma \frac{|\rho(\boldsymbol{\nu})|^2}{D(\boldsymbol{\nu})} \left(\tilde{P}(\boldsymbol{\nu}) \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 + \sigma^2 \right) d\boldsymbol{\nu}.$$

When $\mu = 0$, the latter inequality drastically simplifies according to

$$\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) \geq (K - 1)\sigma^2$$

which proves the assertion.

REFERENCES

[1] R. A. Gonsalves and R. Chidlaw, "Wavefront sensing by phase retrieval," in *Proc. SPIE Applications of Digital Image Processing III*, vol. 207, A. Tescher, Ed., 1979, pp. 32–39.

[2] R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.*, vol. 21, no. 5, pp. 829–832, 1982.

[3] R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations using phase diversity," *J. Opt. Soc. Amer. A*, vol. 9, no. 7, pp. 1072–1085, 1992.

[4] R. G. Paxman, B. J. Thelen, and J. H. Seldin, "Phase-diversity correction of turbulence-induced space-variant blur," *Opt. Lett.*, vol. 19, no. 16, pp. 1231–1233, Aug. 1994.

[5] R. J. A. Little and D. B. Rubin, "On jointly estimating parameters and missing data by maximizing the complete-data likelihood," *Amer. Stat.*, vol. 37, pp. 218–220, Aug. 1983.

[6] A. Blanc, J. Idier, and L. Mugnier, "Novel estimator for the aberrations of a space telescope by phase diversity," in *Proc. SPIE Int. Symp. Astronomical Telescopes and Instrumentation*, Munich, Germany, Mar. 2000, pp. 728–736.

[7] A. Blanc, L. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," *J. Opt. Soc. Amer. A*, vol. 20, no. 6, pp. 1035–1045, Jun. 2003.

[8] J. R. Fienup, J. Marron, T. J. Schulz, and J. H. Seldin, "Hubble space telescope characterized by using phase-retrieval algorithms," *Appl. Opt.*, vol. 32, pp. 1747–1767, 1993.

[9] D. J. Lee, M. C. Roggemann, and B. M. Welsh, "Cramer-Rao analysis of phase-diverse wave-front sensing," *J. Opt. Soc. Amer.*, vol. 16, no. 5, pp. 1005–1015, 1999.

[10] D. Dacunha-Castelle and M. Duflo, *Probabilités et Statistiques, 2. Problèmes à Temps Mobile*, 1st ed. Paris, France: Masson, 1983.

[11] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Amer.*, vol. 66, no. 3, pp. 207–211, 1976.

[12] B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," *IEEE Trans. Comput.*, vol. C-22, no. 8, pp. 805–812, Aug. 1973.

[13] C. R. Vogel, T. Chan, and R. Plemmons, "Fast algorithms for phase-diversity-based blind deconvolution," in *Proc. SPIE Adaptive Optical System Technologies*, vol. 3353, D. Bonaccini and R. K. Tyson, Eds., 1998, pp. 994–1005.

[14] M. Z. Nashed, *Generalized Inverses and Applications*. New York: Academic, 1976.

[15] P. A. Voois, "A theorem on the asymptotic eigenvalue distribution of Toeplitz-block-Toeplitz matrices," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 1837–1841, Jul. 1996.

[16] T. Amemiya, "Non-linear regression models," in *Handbook of Econometrics*, Z. Griliches and M. D. Intriligator, Eds. Amsterdam, The Netherlands, 1983, vol. 1, ch. 6, pp. 333–389.

[17] L. J. Gleser, "Estimation in a multivariate errors in variables regression model: Large sample results," *Ann. Stat.*, vol. 9, no. 1, pp. 24–44, 1981.

[18] R. M. Gray. (2002) Toeplitz and circulant matrices: A review. Inf. Theory Lab., Stanford Univ., Stanford, CA. [Online]. Available: http://www-ee.stanford.edu/~gray/toeplitz.html
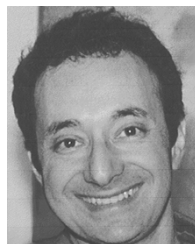
[19] N. K. Bose and K. J. Boo, "Asymptotic eigenvalue distribution of block-Toeplitz matrices," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 858–961, Mar. 1998.

[20] M. G. Löfdahl and G. B. Scharmer, "Wavefront sensing and image restoration from focused and defocused solar images," *Astron. Astrophys.*, vol. 107, pp. 243–264, 1994.

[21] D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wavefront sensing," *Appl. Opt.*, pp. 9186–9197, Dec. 1997.

[22] A. Blanc, "Identification de Réponse Impulsionnelle et Restauration D'images: Apports de la Diversité de Phase," Ph.D. dissertation, Université de Paris-Sud, Orsay, France, 2002.

[23] J. H. Seldin and R. G. Paxman, "Phase-diverse speckle reconstruction of solar data," in *Proc. SPIE Image Reconstruction and Restoration*, vol. 2302, T. Schulz and D. Snyder, Eds., 1994, pp. 268–280.

[24] R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," *Appl. Opt.*, vol. 33, no. 27, pp. 6533–6546, 1994.

[25] M. G. Löfdahl and G. B. Scharmer, "A predictor approach to closed-loop phase-diversity wavefront sensing," in *Proc. SPIE 4013, UV, Optical, and IR Telescopes and Instrumentation VI*, Munich, Germany, 2000, pp. 737–748.

Laurent Mugnier graduated from the Ecole Polytechnique, France, in 1988, and received the Ph.D. degree from the Ecole Nationale Supérieure des Télécommunications (ENST), France, in 1992, for his work on the digital reconstruction of incoherent-light holograms.

In 1994, he joined ONERA, where he is currently a Senior Research Scientist in the field of inverse problems and high-resolution optical imaging. His current research interests include wavefront sensing, image restoration, and image reconstruction, in particular, for adaptive-optics corrected imaging through turbulence, for Earth observation, and for optical interferometry. His publications include a chapter in a reference book on inverse problems, 20 papers in peer-reviewed international journals, and more than 40 papers in conference proceedings.

Jérôme Idier was born in France in 1966. He received the diploma degree in electrical engineering from the École Supérieure d'Électricité, Gif-sur-Yvette, France, in 1988, and the Ph.D. degree in physics from the University of Paris-Sud, Orsay, France, in 1991.

In 1991, he joined the Centre National de la Recherche Scientifique. He is currently with the Institut de Recherche en Communications et Cybernétique, Nantes, France. His major scientific interests are in probabilistic approaches to inverse problems for signal and image processing.

Amandine Blanc received the Ph.D. degree in image processing from the University of Paris-Sud, Orsay, France, in 2002.

She is currently with the Laboratoire d'Astrophysique de l'Observatoire de Grenoble, Grenoble, France. Her main subjects of interest are image restoration, wavefront sensing, and adaptive optics.