Research Article Vol. 41, No. 3 / March 2024 / Journal of the Optical Society of America A A1

Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

ADMM-inspired image reconstruction for terahertz off-axis digital holography

Murielle Kirkove,[1,*] Yuchen Zhao,[1] Olivier Leblanc,[2] Laurent Jacques,[2] and Marc Georges[1]

[1]Centre Spatial de Liège, STAR Research Unit, Université de Liège, Liege Science Park, B-4031 Angleur, Belgium
[2]ISPGroup, INMA, ICTEAM Institute, Université Catholique de Louvain, 1348 Louvain-La-Neuve, Belgium
*M.Kirkove@uliege.be

Received 25 August 2023; revised 24 October 2023; accepted 20 November 2023; posted 20 November 2023; published 8 December 2023

Image reconstruction in off-axis terahertz digital holography is complicated due to the harsh recording conditions and the non-convexity form of the problem. In this paper, we propose an inverse problem-based reconstruction technique that jointly reconstructs the object field and the amplitude of the reference field. Regularization in the wavelet domain promotes a sparse object solution. A single objective function combining the data-fidelity and regularization terms is optimized with a dedicated algorithm based on an alternating direction method of multipliers framework. Each iteration alternates between two consecutive optimizations using projections operating on each solution and one soft thresholding operator applying to the object solution. The method is preceded by a windowing process to alleviate artifacts due to the mismatch between camera frame truncation and periodic boundary conditions assumed to implement convolution operators. Experiments demonstrate the effectiveness of the proposed method, in particular, improvements of reconstruction quality, compared to two other methods.

https://doi.org/10.1364/JOSAA.504126

1. INTRODUCTION

Terahertz (THz) waves refer to electromagnetic radiations with a frequency of 0.1–10 THz, located between the infrared and microwave frequencies. The THz band owns the unique property of penetrating non-polar materials such as polymers, fabrics, biological tissues, and semiconductors without ionization damage [1–3], making THz imaging an emerging technique of high potentials in security, biomedical, and industrial non-destructive testing applications [4,5]. Terahertz (THz) imaging techniques have been growing for the last decades. Various imaging techniques have been successfully migrated from the visible to the THz band [5]. Digital holography (DH) was one of the first techniques in the coherent lensless imaging family implemented in the THz band [6]. It permits to reconstruct the complex wavefront ψ of an object from one interference pattern recorded by a digital image sensor. Various configurations for holographic setups exist [7] including the two complementary techniques of Gabor inline and off-axis holography. For the THz domain, the off-axis configuration is an appropriate choice when the investigated object is non-sparse and complex, as discussed in [8]. This configuration uses off-axis illumination with a separate reference beam (see Fig. 1).

Image reconstruction in off-axis DH consists in recovering the object field ψ from intensity measurements. Compared to the visible configuration, the off-axis THz DH suffers from a few particular problems that make the reconstruction task challenging [9]. First, the recording distance should be minimized to several centimeters to get a reasonable lateral resolution. Such a limited configuration makes the injection of reference wave more difficult, possibly causing excessive diffraction fringes that degrade the reference wave uniformity. Second, due to the large wavelength range, the diffraction effect is more significant in the THz range. The excessive diffraction caused by the optic mounts or the camera housing contributes to bright and dark fringes deteriorating the reference wave uniformity. Finally, the pixel resolutions of the THz cameras are lower than the visible cameras. The above-mentioned problems inevitably deteriorate the reconstruction quality.

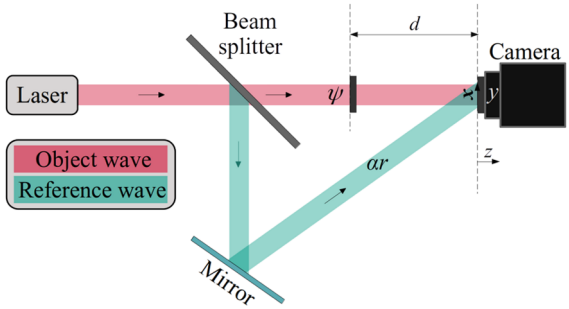

Fig. 1. Off-axis DH setup.

Classical reconstruction approaches rely on filtering in the Fourier domain [10]. In particular, in the direct Fourier method, the reference beam is set to an oscillating function that shifts the spectral content of the input image and allows recovery of the object field by removing the zero- and minus one-order terms of the spectral information [11,12]. However, classical Fourier approaches are corrupted by ringing artifacts due to the loss of high-frequency components.

Inverse problem (IP)-based methods are alternative approaches for solving reconstruction problems in various fields of imaging. They consist of estimating images by minimizing an objective/cost function promoting both a good fidelity to the available observations and specific image structures. In off-axis DH, the reconstruction problem is complicated due to its non-convexity form and the difficulty of estimating the amplitude of the reference field that displays slow spatial variations due to the non-ideal holographic devices. In off-axis DH, the first IP-based approaches come from the work of Sotthivirat and Fessler [13–15] using a penalized-likelihood objective function and a regularization based on an edge-preserving filter. The objective function is minimized by solving a simple surrogate problem with an iterative algorithm.

To cope with the non-uniformity of the reference field amplitude, Bourquard *et al.* proposed an IP-based technique in which an additive unknown map $\boldsymbol{\alpha}$ is introduced to represent the relative intensity of the reference beam [16]. The optimization is obtained in a dual-step way: the data-fidelity term is minimized up to a given precision, and then the object solution is regularized by a total variation (TV) regularization (both minimizations being solved with gradient-descent algorithms). This method demonstrated that IP-based methods can greatly reduce ringing artifacts and slightly increase resolution recovery compared to classical Fourier methods. Furthermore, it is robust to noise and data sub-sampling [9,17,18]. Nevertheless, this method suffers from weaknesses. The TV regularization, promoting sparsity in the gradient domain, is adapted to cartoon-shape, piecewise constant models [19]; this is, however, not a realistic hypothesis for all images depending on the shape of the object being observed. The interleaving of two separate objective functions in the data and solution spaces leads to slow convergence [17] and can miss points that jointly minimize the data-fidelity and regularization terms. More recently, Schretter *et al*. presented in [17] another IP-based technique in which the non-convex problem is replaced by a sequence of surrogate convex problems using a simple projection operator on a subspace of solutions in the data domain. The convex problems are solved with a Nesterov acceleration of the simultaneous Kaczmarz method. No explicit regularization is employed, but instead the solution $\boldsymbol{\psi}$ with minimum ℓ_2-norm in the wavelet domain is selected. The amplitude of the reference field is not introduced as a variable of the problem, but instead a global gain calibration factor is estimated at each iteration and integrated into the projection operator.

The works of [16,17] considered off-axis DH reconstruction in the visible domain. In this work, we propose an IP-based reconstruction technique for off-axis DH in the THz domain. As the technique [16], it *jointly* reconstructs the object field $\boldsymbol{\psi}$ and the amplitude of the reference field $\boldsymbol{\alpha}$. In the visible domain, the modeling of $\boldsymbol{\alpha}$ by a function showing slow spatial variations is compliant with the reality. In THz, the homogeneity of the reference field amplitude is disturbed with fringes. In [16], the $\boldsymbol{\alpha}$ map is defined on a sampled 2D pixel grid that can be as fine as the one of $\boldsymbol{\psi}$, and the only regularization applied to it is the possibility to down-sample it. This is incompatible with the smoothing property of $\boldsymbol{\alpha}$ and, due to the increase of unknowns making the reconstruction problem more ill-posed, leads to slow convergence and often divergence when the down-sampling is not chosen adequately. To avoid such issues, in this work, we consider a simpler model for the $\boldsymbol{\alpha}$ map. Indeed, if we do not take the fringe disturbances into account, we can consider $\boldsymbol{\alpha}$ as a smooth function that is, in first approximation, spatially constant. Such assumption reduces the model of $\boldsymbol{\alpha}$ to one single parameter. Contrary to what has been suggested in [16,17], in this work, a single objective function combining both data-fidelity and regularization terms is optimized. Our scheme thus targets the finding of critical points (local minima) of this function, which can be missed by the alternate formulation. A ℓ_1-norm regularizer expressed in the wavelet domain provides a sparse solution $\boldsymbol{\psi}$.

The optimization of our reconstruction method relies on a dedicated algorithm involving an alternating direction method of multipliers (ADMM)-based framework. Examples of its application for solving other phase retrieval problems can be found in [20,21]. However, as far as we know, this has never been applied to off-axis DH reconstruction. In addition, we developed an extension of the original ADMM algorithm to consider $\boldsymbol{\alpha}$ as a variable of the problem. Each iteration of our optimization algorithm alternates between two consecutive optimizations: one with respect to variable $\boldsymbol{\psi}$ and one with respect to variable $\boldsymbol{\alpha}$. The optimizations use projection operators derived from the one defined in [17] to minimize the data-fidelity terms and one soft thresholding operator to minimize the regularization term. The projections operate not only on solution $\boldsymbol{\psi}$ but also on solution $\boldsymbol{\alpha}$.

Our method, as many other reconstruction methods, adopts a periodic boundary condition (BC). Such assumption is unnatural but merely motivated by computational convenience. Nevertheless, in the case of off-axis DH in the THz domain where the frame is truncated by the camera, a periodic BC assumption creates discontinuities in the input image by periodic reproduction. This mismatch produces oscillation artifacts emanating from the boundaries of the solution throughout the whole image [22–25]. In the proposed method, a simple pre-processing by apodization is adopted, as in [8], to alleviate these artifacts.

2. OFF-AXIS DIGITAL HOLOGRAPHY

A. Geometry

The setup of off-axis DH is depicted in Fig. 1. A laser beam is separated by a beam splitter into two parts with equivalent intensity: the reference and the object beam, traveling in different directions. The camera then records the interference pattern, i.e., a hologram.

B. Forward Model

The general continuous DH model that we consider is standard in the literature [16]. In the case of a thin object, located at $z=-d$ from the detector plane, the observed intensity $y(\boldsymbol{x})$ at spatial position $\boldsymbol{x}$ is then given by this forward model:

$$y(\boldsymbol{x}) = |o(\boldsymbol{x}, 0) + r(\boldsymbol{x}, 0)|^2, \tag{1}$$

where $o(\boldsymbol{x}, z)$ and $r(\boldsymbol{x}, z)$ denote the (complex) object and reference fields at position $\boldsymbol{x}$ and depth $z \leq 0$, respectively. The object field measured at $z=0$ amounts to a blurred observation of the object $\psi(\boldsymbol{x}) = o(\boldsymbol{x}, -d)$ that we want to estimate,

$$o(\boldsymbol{x}, 0) = (A_d \psi)(\boldsymbol{x}, 0),$$

where A_d is the instrumental point spread function. Its expression in the Fourier domain results from the successive application of the spatial-frequency cutoff function associated to the microscope objective and the digital propagation function with distance d (see, e.g., [16]).

However, in addition to ψ, the amplitude function $\alpha(\boldsymbol{x}) = |r(\boldsymbol{x}, 0)|$ of the reference field r is also unknown and must be estimated. Introducing $\Phi(\boldsymbol{x}) = \arg(r(\boldsymbol{x}, 0))$ as the phase of r, and $\bar{r} = e^{-\mathrm{i}\Phi}$, Eq. (1) becomes

$$y(\boldsymbol{x}) = |(A_d \psi)(\boldsymbol{x}, 0) + (\alpha \bar{r})(\boldsymbol{x}, 0)|^2.$$

The sampling induced by the N-pixel camera followed by the vectorization (transforming images into vectors) results in the following discrete forward model:

$$\bar{\boldsymbol{y}} = |\boldsymbol{o} + \boldsymbol{\alpha} \odot \boldsymbol{r}|^2 = |\boldsymbol{A}_d \boldsymbol{\psi} + \boldsymbol{\alpha} \odot \boldsymbol{r}|^2$$

where the square modulus is applied component-wise onto vectors, the vectors $\bar{\boldsymbol{y}} \in \mathbb{R}_+^N$, $\boldsymbol{o} \in \mathbb{C}^N$, $\boldsymbol{\psi} \in \mathbb{C}^N$, $\boldsymbol{\alpha} \in \mathbb{R}_+^N$, and $\boldsymbol{r} \in \mathbb{C}^N$ are obtained by sampling and vectorizing the functions y, o, ψ, α, and $\bar{r}$, respectively, and $\boldsymbol{A}_d \in \mathbb{C}^{N \times N}$ is the doubly block circulant matrix associated with the filter A_d. In this work, we assume that α is spatially constant (see introduction), i.e., $\boldsymbol{\alpha} = (\alpha, \ldots, \alpha)^\intercal$; therefore, the previous expression becomes

$$\bar{\boldsymbol{y}} = |\boldsymbol{o} + \alpha \boldsymbol{r}|^2 = |\boldsymbol{A}_d \boldsymbol{\psi} + \alpha \boldsymbol{r}|^2. \tag{2}$$

The discrete forward model also takes into account the external disturbances (noise) intrinsic to the physical measurement process. Due to its simplicity, an additive white Gaussian noise (AWGN) model is typically adopted in imaging systems as DH where the true noise model is unknown [26]. This consideration yields

$$\boldsymbol{y} = \bar{\boldsymbol{y}} + \boldsymbol{n} \tag{3}$$

where $\boldsymbol{n} \in \mathbb{R}^N$ represents an AWGN noise of variance σ^2.

3. RECONSTRUCTION METHOD

A. Reconstruction Problem

The objective of off-axis DH reconstruction is to deduce from the recorded measurements $\boldsymbol{y}$ the unknown complex image $\boldsymbol{\psi}$ and the (unknown) amplitude α, in other words, to solve the inverse problem related to the forward relation in Eqs. (2) and (3). This is an instance of the phase retrieval problem [27], that is, the recovery of a complex signal from the magnitude of a transform of that signal. With an IP-based approach, the objective of off-axis DH reconstruction is to retrieve two estimates $\tilde{\boldsymbol{\psi}}$ and $\tilde{\alpha}$ by minimizing the following data-fidelity term, or cost function, which measures the distance between the actual measurements and the ones given by two candidates and the forward model defined as Eqs. (2) and (3):

$$D(\boldsymbol{\psi}, \alpha) = \| \boldsymbol{y} - |\boldsymbol{A}_d \boldsymbol{\psi} + \alpha \boldsymbol{r}|^2 \|_2^2 . \tag{4}$$

Solving this problem is challenging. First, it is non-convex relatively to the unknowns $\boldsymbol{\psi}$ and α. This can be seen by the cross-term obtained when developing the square modulus in Eq. (4). Therefore, D can have many local minima. Second, the problem is ill-posed: it is undetermined since the number of unknowns (the dimension of $\boldsymbol{\psi}$ and α) is larger than the number of measurements (the dimension of $\boldsymbol{y}$), which makes it difficult to separate solutions $\boldsymbol{\psi}$ and α. Moreover, it is unstable and highly sensitive to noise due to the frequency distribution of the operator $\boldsymbol{A}_d$. Therefore, regularization is mandatory for selecting a meaningful solution among all compatible solutions.

We regularize the solution $\boldsymbol{\psi}$ in the wavelet domain by minimizing the ℓ_1-norm of its wavelet coefficients for promoting sparse solutions [17,28]. We consider here the frame synthesis (FS) formulation of the wavelet-based regularization [22,29] where the regularizer is applied to the frame coefficients and the image to reconstruct is synthesized from its frame coefficients. As in [17], the wavelet representation is obtained by the discrete wavelet transform (DWT) [30] with the biorthogonal Cohen-Daubechies-Feauveau (CDF) 9/7 wavelet and is implemented with the lifting scheme representation [31]. This wavelet, used in the JPEG 2000 compression standard for lossy compression [32], was recommended by Bettens *et al.* [28] for yielding sparse representations and robust reconstructions for compressive DH.

We rewrite the noise-free model of Eq. (2) and the data-fidelity term of Eq. (4) with respect to the wavelet coefficients $\boldsymbol{c}_\psi$ of $\boldsymbol{\psi}$ as

$$\bar{\boldsymbol{y}}(\boldsymbol{c}_\psi, \alpha) = \left| \boldsymbol{A}_d \boldsymbol{W}^{-1} \boldsymbol{c}_\psi + \alpha \boldsymbol{r} \right|^2, \tag{5}$$

$$D(\boldsymbol{c}_\psi, \alpha) = \| \boldsymbol{y} - \bar{\boldsymbol{y}}(\boldsymbol{c}_\psi, \alpha) \|_2^2, \tag{6}$$

where $\boldsymbol{W}$ is the matrix of fast DWT. Taking into consideration these equations, the wavelet-based regularization of the off-axis DH reconstruction leads to the following minimization problem:

$$(\tilde{\boldsymbol{c}_\psi}, \tilde{\alpha}) = \underset{\boldsymbol{c}_\psi, \alpha}{\operatorname{argmin}} \; D(\boldsymbol{c}_\psi, \alpha) + \lambda \, \| \boldsymbol{c}_\psi \|_1, \tag{7}$$

where $\lambda > 0$ controls the regularization strength.

B. Reconstruction Algorithm

The minimization problem Eq. (7) is non-convex. Despite the lack of guarantees of convergence toward an optimal solution, we still use a convex optimization algorithm to solve it, in this case, an algorithm inspired by the ADMM approach. The

ADMM is a particular variant of the augmented Lagrangian methods (ALM) that are part of a certain class of algorithms for solving constrained-optimization problems (COP) and were originally known as the method of multipliers [33]. The ALMs replace a COP by a series of unconstrained-optimization problems and add a penalty term to the objective function. The ADMM, originally proposed in [34], has been extensively explored in recent years in wide variety of problems such as image processing [35], applied machine learning and statistics [36], sparse optimizations, and other relevant fields [34]. Most of the works on the ADMM algorithm focused on the real domain. The application to off-axis DH problem is made possible by the generalization of the ADMM algorithm to the complex domain presented by Li *et al.* in [33], called here the ADMMCP algorithm. We developed an extension of this algorithm to integrate the dependency of the optimization on the variable α, which is presented in Appendix B. The application of the ADMMCP algorithm to off-axis DH reconstruction problem requires the computation of proximity operators of the data-fidelity function, which is detailed in Appendix C.

Our extended ADMMCP algorithm, described in Appendix B, is intended to solve problems in the form of Eq. (B3). We adapted it for solving problem Eqs. (5)–(7). First, we express Eq. (7) in the form of Eq. (B3), where $\boldsymbol{x}_1 = \boldsymbol{c}_\psi \in C_{x_1} \subset \mathbb{C}^N$, $x_2 = \alpha \in C_{x_2} \subset \mathbb{R}$, $\boldsymbol{z}_1 \in C_{z_1} \subset \mathbb{C}^N$, $z_2 \in C_{z_2} \subset \mathbb{R}$, $f(\boldsymbol{x}_1, x_2) = D(\boldsymbol{x}_1, x_2)$, $g_1(\boldsymbol{z}_1) = \lambda \|\boldsymbol{z}_1\|_1$, $g_2(z_2) = 0$, $\boldsymbol{A}_1 = \boldsymbol{A}_2 = \boldsymbol{I}$, $\boldsymbol{B}_1 = \boldsymbol{B}_2 = -\boldsymbol{I}$, $\boldsymbol{c}_1 = 0$, and $c_2 = 0$. Since the matrices $\boldsymbol{A}_l$ and $\boldsymbol{B}_l$ are not singular, we use the expression Eq. (B6) of the extended ADMMCP iterations. In addition, for the variable $x_2 = \alpha$, the equality $\text{prox}_0(\boldsymbol{v}) = \boldsymbol{v}$ permits the reduction of the three related iterations to the single iteration $\alpha^{k+1} = \text{prox}_{\frac{1}{2\rho}D}(\alpha^k)$. Therefore, the iterations of the extended ADMMCP become

$$\begin{cases} \boldsymbol{c}_\psi^{k+1} &= \text{prox}_{\frac{1}{2\rho}D}(\boldsymbol{z}^k - \boldsymbol{u}^k) \\ \boldsymbol{z}^{k+1} &= \text{prox}_{\frac{\lambda}{2\rho}\|\cdot\|_1}(\boldsymbol{c}_\psi^{k+1} + \boldsymbol{u}^k) \\ \boldsymbol{u}^{k+1} &= \boldsymbol{u}^k + \boldsymbol{c}_\psi^{k+1} - \boldsymbol{z}^{k+1} \\ \alpha^{k+1} &= \text{prox}_{\frac{1}{2\rho}D}(\alpha^k) \end{cases}.$$

The $\boldsymbol{c}_\psi$- and the α-updates use the proximity operator of the same function D implicitly considered as a single-variable function of $\boldsymbol{c}_\psi$ and $\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$ at the first and fourth line, respectively; they are respectively obtained by replacing in expressions Eqs. (C9) and (C10) of the proximity operators $\boldsymbol{v}$ by $\boldsymbol{z}^k - \boldsymbol{u}^k$ and by α^k for γ set at $\frac{1}{2\rho}$. The $\boldsymbol{z}$-update is given by the proximity operator of the ℓ_1-norm, which is the soft-thresholding operator extended to the complex domain [33]. If $\mathcal{S}_\gamma$ denotes this operator for the threshold γ, the $\boldsymbol{z}$-update is $\boldsymbol{z}^{k+1} = \mathcal{S}_{\frac{\lambda}{2\rho}}(\boldsymbol{c}_\psi^{k+1} + \boldsymbol{u}^k)$. Since Eqs. (C9) and (C10) do not depend on parameter γ, the expressions of $\boldsymbol{c}_\psi$- and the α- updates do not depend on parameter ρ, the dependency to this parameter in the $\boldsymbol{z}$-update can be removed, and the previous iterations become

$$\begin{cases} \boldsymbol{c}_\psi^{k+1} &= \boldsymbol{W}\boldsymbol{A}_d^\dagger \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}_d \boldsymbol{W}^{-1}(\boldsymbol{z}^k - \boldsymbol{u}^k)),\ \boldsymbol{b} = \alpha^k \boldsymbol{r} \\ \boldsymbol{z}^{k+1} &= \mathcal{S}_\lambda(\boldsymbol{c}_\psi^{k+1} + \boldsymbol{u}^k) \\ \boldsymbol{u}^{k+1} &= \boldsymbol{u}^k + \boldsymbol{c}_\psi^{k+1} - \boldsymbol{z}^{k+1} \\ \alpha^{k+1} &= \langle |\boldsymbol{P}_{\boldsymbol{b},\delta}(\alpha^k \boldsymbol{r})| \rangle,\ \boldsymbol{b} = \boldsymbol{A}_d \boldsymbol{W}^{-1} \boldsymbol{c}_\psi \end{cases}, \quad (8)$$

Algorithm 1. ADMMCP-based algorithm

Initialization: $\boldsymbol{c}_\psi^0 = 0$, $\alpha^0 = 0$, $\boldsymbol{z}^0 = 0$, $\boldsymbol{u}^0 = 0$
while stopping condition not reached **do**
  $\boldsymbol{b} \leftarrow \alpha^k \boldsymbol{r}$
  $\boldsymbol{c}_\psi^{k+1} \leftarrow \boldsymbol{W}\boldsymbol{A}_d^\dagger \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}_d \boldsymbol{W}^{-1}(\boldsymbol{z}^k - \boldsymbol{u}^k))$
  $\boldsymbol{z}^{k+1} \leftarrow \mathcal{S}_\lambda(\boldsymbol{c}_\psi^{k+1} + \boldsymbol{u}^k)$
  $\boldsymbol{u}^{k+1} \leftarrow \boldsymbol{u}^k + \boldsymbol{c}_\psi^{k+1} - \boldsymbol{z}^{k+1}$
  $\boldsymbol{b} \leftarrow \boldsymbol{A}_d \boldsymbol{W}^{-1} \boldsymbol{c}_\psi$
  $\alpha^{k+1} \leftarrow \langle |\boldsymbol{P}_{\boldsymbol{b},\delta}(\alpha^k \boldsymbol{r})| \rangle$

which leads to the ADMMCP-based algorithm for the problem of Eqs. (5)–(7) (see Algorithm 1).

This algorithm alternates between two consecutive optimizations: one with respect to $\boldsymbol{\psi}$ and one with respect to α. Each optimization uses one projection operator $\boldsymbol{P}_{\boldsymbol{b},\delta}$, defined by Eq. (C7) and based on the projector defined in [17] to minimize the data-fidelity term, and the optimization with respect to $\boldsymbol{\psi}$ uses one soft thresholding operator to minimize the regularization term. Contrary to what has been proposed in [17], the projections operate not only on $\boldsymbol{\psi}$ but also on α.

Algorithms dedicated to inverse problem approaches involve a set of parameters controlling the optimization process. For clarity, we use the term hyperparameters, borrowed from the machine learning field, to name such parameters and distinguish them from the parameters estimated by the learning process. For coherence, we also use this term for the parameters controlling the quality of the reconstruction of the classical Fourier methods. In our experiments, we fix the number of wavelet decomposition levels at the largest possible value (depending on the dimensions of the solution $\boldsymbol{\psi}$) and the hyperparameter μ involved in the definition of the projection operators $\boldsymbol{P}_{\boldsymbol{b},\delta}$ at $\mu = 3$. The constant σ that quantifies the noise level and is also used in the definition of $\boldsymbol{P}_{\boldsymbol{b},\delta}$ can be seen as another regularization hyperparameter due to its difficult estimation and its effect on the regularization [see Eq. (C7)]. In addition, to use the advantages of the wavelet-based regularization for adjusting the regularization of the components of $\boldsymbol{\psi}$ according to their resolution, we introduced a dependency on the levels l in the regularization strength by defining $\lambda(l) = \lambda f^{-(l-1)}(f \geq 1)$. A value of f larger than 1 has the effect of decreasing the regularization strength as l increases and, thus, as the resolution decreases. Using such setting for f has, thus, the effect to more regularize the high-resolved structures. Our algorithm depends, thus, on only four processing hyperparameters: the regularization strength λ, its modulation factor f, another regularization hyperparameter σ, and the number of iterations n_i.

Since the cost function of problem Eq. (7) is non-convex, the solution obtained with a convex optimization algorithm should depend on the initialization. Nevertheless, in practice, we observed that our algorithm converges to the same solution regardless of the initial estimates. We set the initial estimate of $\boldsymbol{\psi}$ as the solution of the direct Fourier method since it leads to a faster convergence.

C. Protocol for Adjusting Hyperparameters

The adjustment of the hyperparameters is a complex work due to the high dimension of the hyperparameters space. To provide a strategy, we established a protocol based on the analysis

of the influences of the hyperparameters on the quality of the reconstruction with the help of following criteria: visual characteristics as the proportion of artifacts and details in the solution $\boldsymbol{\psi}$ and performance measures computed between the solution $\boldsymbol{\psi}(\boldsymbol{\alpha})$ and its true distribution $\boldsymbol{\psi}_0(\boldsymbol{\alpha}_0)$. The following complementary similarity measures were used: the signal-to-noise ratio (SNR) and the structural similarity index measure (SSIM). The SNR is inversely proportional to the magnitude of the reconstruction error. It ranges between $[0, +\infty]$, and a higher SNR value indicates less difference between the reconstructed image and its original distribution. The SNR measure is relevant from a mathematical point of view, but it is not always informative since it does not correlate well with visual evaluation. The SSIM was developed to circumvent this problem [37]. It ranges between [0,1], and a higher SSIM value indicates better similarity between the reconstructed image and its original distribution.

It is instructive to note that, even if the protocol has been proven on synthetic data, it does not depend on the performance measures for its application to real data. The protocol is as follows. In a first step, the hyperparameters λ and f are tuned for adjusting the regularization of the components of $\boldsymbol{\psi}$ according to their resolution. To that aim, the hyperparameter λ is first decreased as much as possible to better reconstruct the details of $\boldsymbol{\psi}$ while not adding artifacts and residual noise; then, the hyperparameter f is increased to get a sufficient level of details in the components of low resolution without adding artifacts, and finally the hyperparameter λ is increased for reducing the level of residual noise and correcting a possible lack of uniformity in the areas of low resolution. In a second step, the hyperparameter σ is increased to decrease the possible residual noise and artifacts. In a final step, the hyperparameter n_i is increased as much as possible until there is no significant improvement of the solution $\boldsymbol{\psi}$. This adjustment is justified because, in our experiments, the method is convergent in the sense that all criteria used in the protocol are optimized at a given number of iterations $n_i = n_{i0}$ and remain stable at larger number of iterations, thus yielding a solution independent of the choice of n_i provided that it is larger than n_{i0}.

D. Pre-processing

Periodic boundary condition (BC) is unnatural but merely motivated by computational convenience. Our method involves the acquisition system operator and its pseudo-inverse ($\boldsymbol{A}_d$ and $\boldsymbol{A}_d^{\dagger}$) that are convolution operators. The convolution effect implies that the pixels located near the boundary of the output depend on pixels of the input outside of its domain. The typical way to formalize this issue is to adopt a periodic BC that allows very fast implementation of the convolution using point-wise multiplication in the Fourier domain, efficiently realized using the FFT. In the case of off-axis DH in the THz domain where the frame is truncated by the camera, periodic reproduction of the input image creates discontinuities. This mismatch produces oscillation artifacts emanating from the boundaries of the solution throughout the whole image [22–25]. Such artifacts, which we call artifacts due to camera frame truncation due to their origin, are produced by any reconstruction method assuming periodic BC assumption. The simplest way to alleviate these artifacts is, as proposed for direct methods, to apply a pre-processing by apodization [8]. This solution has the defect of producing a decrease in intensity in the border area of the reconstructions. For a solution that does not have this limitation, as proposed in [22,38], a more realistic BC—an unknown BC—instead of a periodic BC should be considered for both solutions $\boldsymbol{\psi}$ and α. Nevertheless, the proposed pre-processing technique has the advantage of being simple to implement. As in [8], apodization is performed with a 2D Tukey window.

4. EXPERIMENTS

We implemented the described reconstruction method in MATLAB. Executable files of the software are available upon request from the authors.

The experimental setup is described in Section 4.A. It involves a laser (2.52 THz, wavelength $\lambda = 118.83$ μm) and a camera (480×640 pixels with pitch of 17 μm).

We have reconstructed synthetic and real data with the proposed method. For a comparison purpose, we also applied reconstruction of the direct Fourier method and the method of Bourquard *et al.* [16]. The direct Fourier method uses one single hyperparameter: the cutoff frequency of the low-pass filtering in the Fourier domain. The main hyperparameters of the method of Bourquard *et al.* are ϵ_{SNR}, the precision up to which the data-fidelity term is minimized converted into SNR, γ, the value balancing the respective influence of the phase and amplitude regularizations, μ, the step size of the gradient-descent for TV regularization, d_α, the down-sampling scale for the $\boldsymbol{\alpha}$ map, and n_i the number of iterations. To optimize the reconstruction quality of the proposed method, we use the protocol described in Section 3.C. For experiments on the synthetic data, the optimization of the quality reconstruction required us to set f at 1.5, which implies an adaptation of the regularization with respect to the resolution. For experiments on the real data, no such adaptation was made since the protocol led to the value 1 for f. There is no such protocol for the two other methods. For the direct Fourier method, we adjust the cutoff frequency to keep most of the signal of interest while removing most of the signal of the reference field. For a fair comparison, we developed for the method of Bourquard *et al.* an ad-hoc protocol similar to ours. It is worth noting that the adjustment of the hyperparameters is more complicated for the method of Bourquard *et al.* than for the proposed method due to non-convergence and interdependence of hyperparameters ϵ_{SNR} and μ. In addition, to compare the methods in term of the frequency recovery, we compute the power spectral density (PSD) of the measures $\boldsymbol{y}$, their reconstructions $\boldsymbol{\psi}$, and their true distributions $\boldsymbol{\psi}_0$, in case of synthetic data.

A. Experiments on Synthetic Data

1. Data Description

The off-axis angle is set to 45° in vertical and horizontal directions, and the distance between the object and the recording plane is 8.5 mm. We performed simulations of a highly transparent phase object. The amplitude and the phase distributions of the simulated object are both a USAF resolution target. The values of the amplitude and the phase are, respectively,

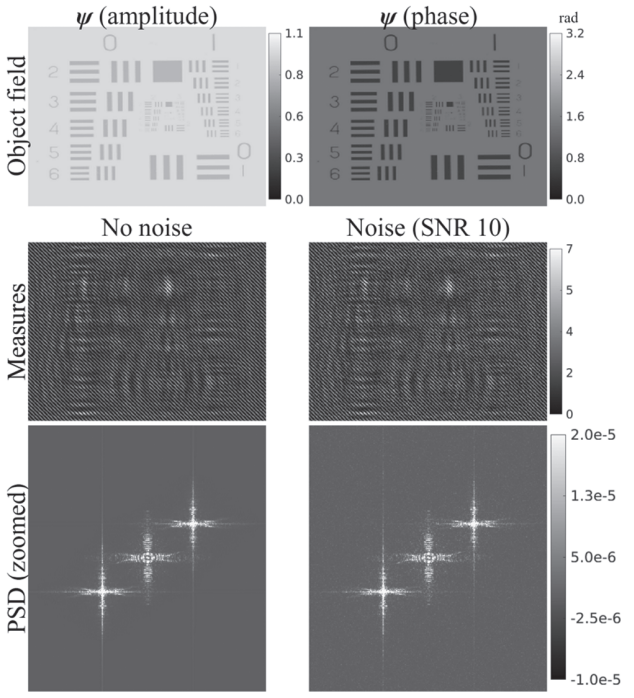

Fig. 2. Synthetic data: cases without noise and with noise of SNR 10.

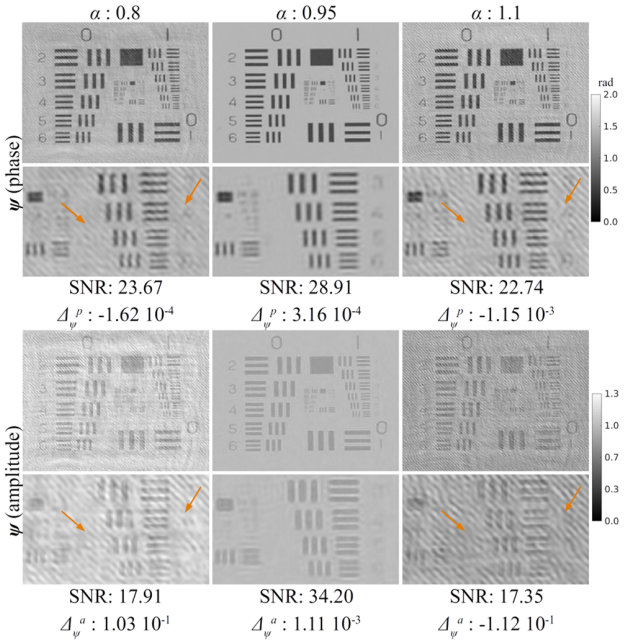

Fig. 3. Reconstruction of synthetic data without noise by the proposed method: effect of incorrect estimation of α (orange arrows indicate oblique fringe artifacts).

in the range [0.8, 1] and $[0, \pi]$ after normalization. We simulated the amplitude of the reference field with a constant. We generated several holograms, one without noise and others with addition of AWGN noise. To evaluate the limitations of the reconstruction methods with respect to the noise level, we generated three sets of holograms with different noise levels corresponding to SNRs in {10,20,30} dB, each of these sets counting 10 holograms. Figure 2 shows simulation results without noise and with noise of SNR 10. For all the methods, the hyperparameters are adjusted on one image of each data set and then are used for each other image in the data set. In the method of Bourquard *et al.*, the $\boldsymbol{\alpha}$ map can be spatially variable or not. For a fair comparison, since the true $\boldsymbol{\alpha}$ map is spatially constant, we consider the second option.

2. Results

To better analyze the results on synthetic data, we introduce performance measures to quantify the average deviation between the estimated and true distributions of all solutions: for α, the relative difference Δ_α between α and its true value α_0; for $\boldsymbol{\psi}$, the relative difference $\Delta^a_{\boldsymbol{\psi}}$ ($\Delta^p_{\boldsymbol{\psi}}$) between the means of the amplitude (phase) of $\boldsymbol{\psi}$ and of its true distribution $\boldsymbol{\psi}_0$.

The off-axis DH reconstruction is a difficult problem to solve due to the mixing between object field $\boldsymbol{\psi}$ and reference field $\boldsymbol{\alpha r}$. We first show the influence of an incorrect estimation of $\boldsymbol{\alpha r}$ on the solution $\boldsymbol{\psi}$ for the proposed method. The conclusions drawn from this study are also valid for all methods based on inverse problem approaches.

To estimate the effect of a mis-estimation of the amplitude of the reference field, we reconstructed the synthetic data by fixing α at different values around its true value $\alpha_0 \sim 0.95$ and considering the hyperparameters selected for the normal reconstruction (with variability of α). Figure 3 shows the results for the synthetic data without noise for $\alpha = 0.95$ and α-values equidistant from 0.95 at which the reconstruction quality starts to decrease. As it is shown, a relatively small mis-estimation of α is enough to reduce the reconstruction quality and the performance results. More specifically, it produces oblique fringe artifacts in $\boldsymbol{\psi}$ (see orange arrows) in the direction of the fringes in the acquired reference field and hologram (see Fig. 2) and an over-estimation (under-estimation) of α produces an under-estimation (over-estimation) of the average amplitude of $\boldsymbol{\psi}$. The conclusions drawn from this study are valid for all the selected noise levels, and the only difference is that the quality of the reconstruction is more sensitive to an incorrect estimation of α when the noise level is lower. Indeed, the introduction of high noise levels in the algorithm increases the risk of various kinds of artifacts and remaining noise that dominates the fringe artifacts. In addition, for transparent phase objects for which the amplitude is of low contrast, the deterioration of $\boldsymbol{\psi}$ by all the artifacts is more visible on its amplitude than on its phase.

To estimate the effect of a mis-estimation of the phase of the reference field, we reconstructed the synthetic data without noise by fixing α at its true value α_0 and by performing the reconstruction with incorrect phases of $\boldsymbol{r}$ associated to a same off-axis angle θ in both direction close to its true value θ_0 (45°). Figure 4 shows the results for $\theta = \theta_0$ and for values of θ equidistant from θ_0 at which the reconstruction quality starts to decrease. As it is shown, a very small mis-estimation of the phase of $\boldsymbol{r}$ is enough to reduce the reconstruction quality and the performance results; more specifically, it produces line artifacts in both components of $\boldsymbol{\psi}$. Therefore, since the off-axis angles of $\boldsymbol{r}$ are not unknowns of the reconstruction problem, it is crucial to correctly calculate them. Their estimation, based on the analysis of the interference patterns of $\boldsymbol{y}$ in the frequency

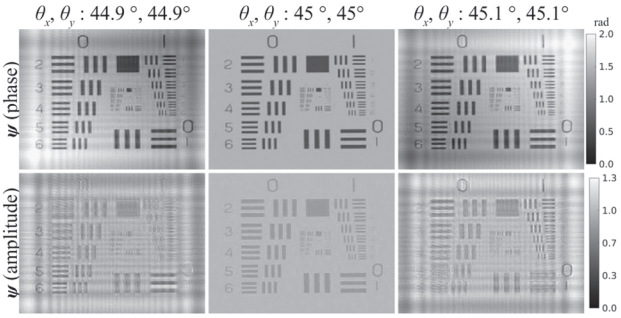

Fig. 4. Reconstruction of synthetic data without noise by the proposed method: effect of incorrect estimation of phase of $\boldsymbol{r}$.

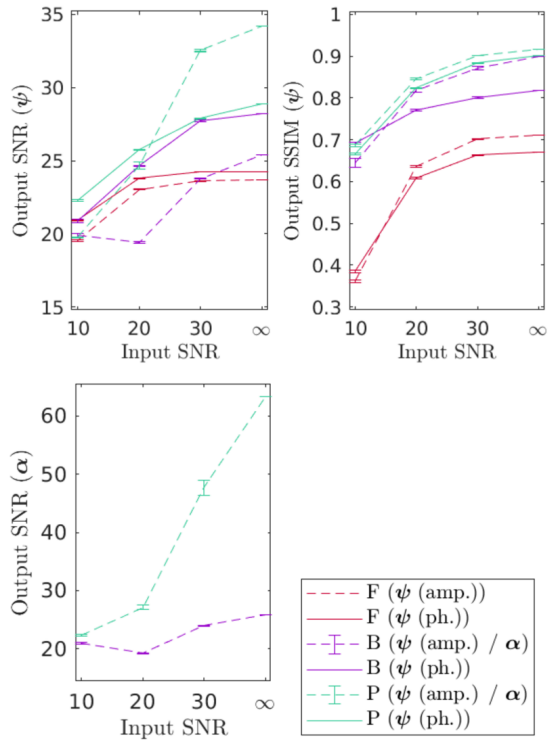

Fig. 5. Reconstruction of synthetic data: performance measures with respect to noise level. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B), and the proposed method (P).

spectrum, provides an error much lower than 1°, which does not cause any reconstruction problems.

Figure 5 shows the performance measures of the comparative study of the three methods with respect to the noise level, i.e., their values for the data without noise and the mean and standard deviation over the 10 trials for the data with noise. We can see that the proposed method (P) provides at all noise levels, the best SSIMs on amplitude of $\boldsymbol{\psi}$, the best SNRs on phase of $\boldsymbol{\psi}$, and the best SNRs on $\boldsymbol{\alpha}$. It also provides the best SSIMs on phase of $\boldsymbol{\psi}$ and the best SNRs on amplitude of $\boldsymbol{\psi}$ at all noise levels except at the lowest one where the best values are given by the method of Bourquard *et al.* (B). The latter produces the highest standard deviations except for the SSIM on phase of $\boldsymbol{\psi}$ at the highest noise level. This shows a strong variation of the results qualities with respect to the noise realizations.

Figure 6 illustrates the best performance of the proposed method at the highest noise levels. Since the three compared methods behave similarly on noise-free data and on noisy data of highest SNRs (20 and 30), we show in this figure the results on the noise-free data. The direct Fourier method (F) shows ringing artifacts on the whole image (see oscillations in the flat areas) resulting from the loss of high-frequency components. As shown by comparing the PSDs of all reconstructions with the one of the simulation, the methods based on inverse problem approaches permit better recovery of high-frequency components. The average amplitude of $\boldsymbol{\psi}$ is mis-estimated by the two other methods due to a mis-estimation of α (which is underestimated by the method of Bourquard *et al.* (B) as indicated by Δ_α). On the other hand, the proposed method (P) yields a correct estimation of α and the average amplitude of $\boldsymbol{\psi}$. In addition, as shown in the zoomed images and the cut profile, in comparison to the method of Bourquard *et al.* (B), it provides slightly better resolution and weaker signal fluctuations in the flat areas. The only weakness of the proposed method (P) is the production of blocky artifacts that are faintly visible. Such artifacts are due to Gibbs problems; they are unavoidable for any kind of regularization, and the effects are well-localized for wavelet-based regularization.

Figure 7 illustrates the poorer performance of the proposed method at the lowest noise level. In this case, the average amplitude of $\boldsymbol{\psi}$ is underestimated due to an overestimation of α as indicated by Δ_α. As such value of Δ_α, at lower noise level, the reconstructed $\boldsymbol{\psi}$ would have been altered with oblique fringe artifacts, but, as mentioned here-above, for this noise level, the fringe artifacts are dominated by other artifacts, here blocky artifacts, and remaining noise. In addition, for transparent phase objects, for which the amplitude of $\boldsymbol{\psi}$ is of low contrast, the deterioration of $\boldsymbol{\psi}$ by these artifacts is more visible on its amplitude than on its phase, and, as shown in the zoomed images and the cut profile, compared to the two other methods, the proposed method (P) provides on the phase more resolution and a signal closer to the original one.

We conclude from synthetic experiments that the proposed method outperforms the two other methods for low to medium noise levels. For a high noise level, it outperforms the two other methods only for the reconstruction of the phase. This conclusion must be nuanced since, for a highly transparent phase object, the main objective of the reconstruction is to retrieve its phase.

B. Experiments on Real Data

1. Data Description

The off-axis angle is set to 41°/21° in the horizontal/vertical directions, and the distance between the object and the recording plane is 9.5 mm. The investigated sample (Fig. 8) is a visibly opaque 550-μm-thick polypropylene (PPP) slab with different engraved patterns. Polypropylene is highly transparent at 2.52 THz. Patterns 1–3 are three squares of surface area $1 \times 1\ \text{mm}^2$ and depth 44 μm, 72 μm, and 105 μm, respectively. The line widths of patterns 4–6 are 245 μm, 155 μm, and 140 μm to evaluate lateral resolution performance. Figure 8

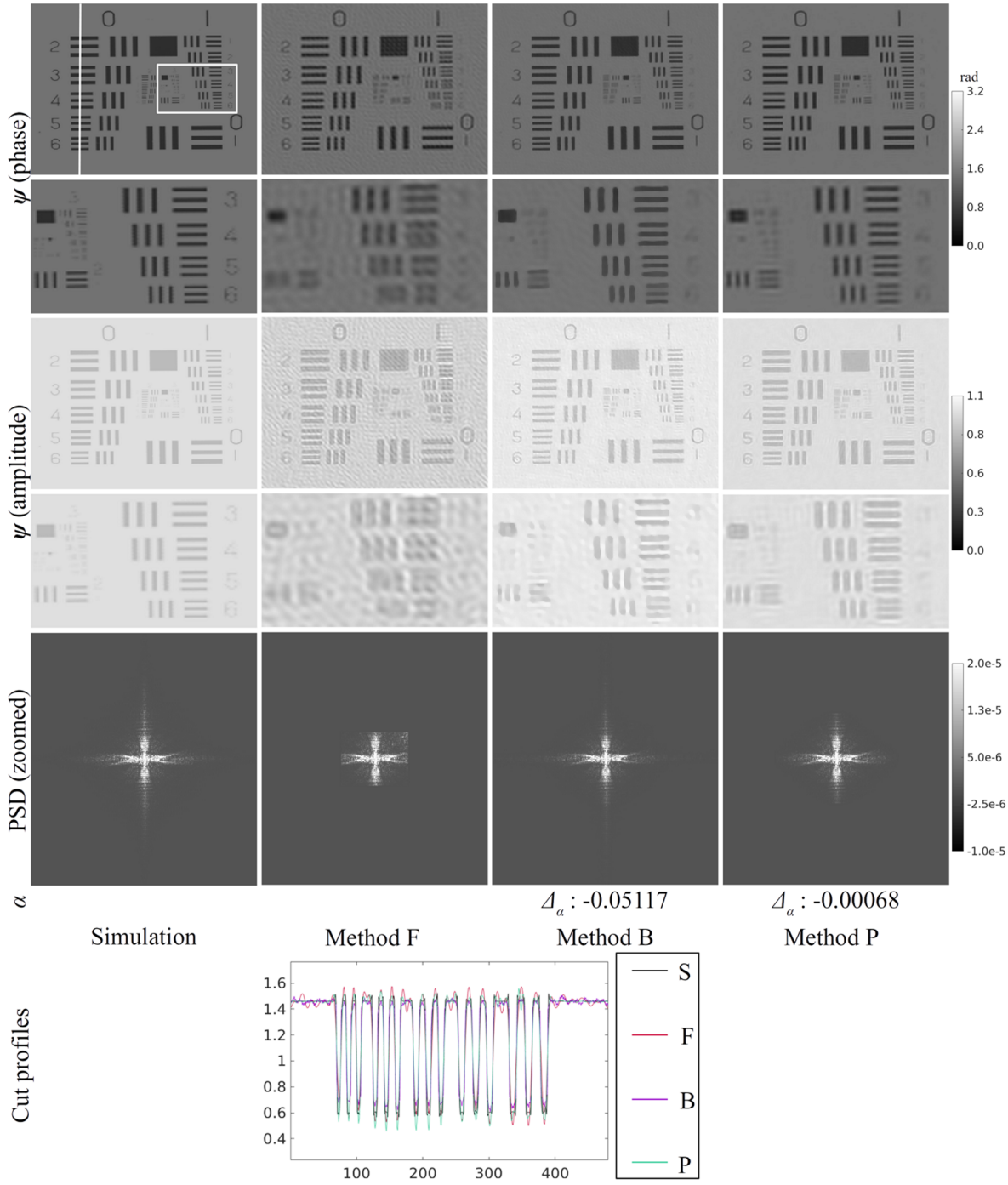

Fig. 6. Reconstruction of synthetic data without noise. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B), and the proposed method (P).

(Measure) shows the recorded hologram. An acquisition obtained in the same condition without any object provides an acquisition of the amplitude of the reference field (Fig. 8, Reference). This image illustrates a typical reference wavefront with diffraction patterns caused by the camera truncation.

2. Results

We performed the reconstructions of the real data with all the considered methods. Since the true $\boldsymbol{\alpha}$ map is certainly not spatially constant, for a fair comparison, we select in the method of Bourquard *et al.* the option for spatially adaptive estimation of $\boldsymbol{\alpha}$, and we set d_α at 8 as suggested in [16]. In Fig. 9, we observed, in the results of the three methods, artifacts emanating from the boundaries of the images throughout the whole image (see red arrows). Such artifacts, named here artifacts due to camera frame truncation, are produced by any reconstruction method assuming periodic BC due to the mismatch between such assumption and frame truncation by the camera (see Section 3.D). The implementation of the two other methods also use periodic BC (for Fourier filtering in the direct Fourier method and for the convolution operators in the method of

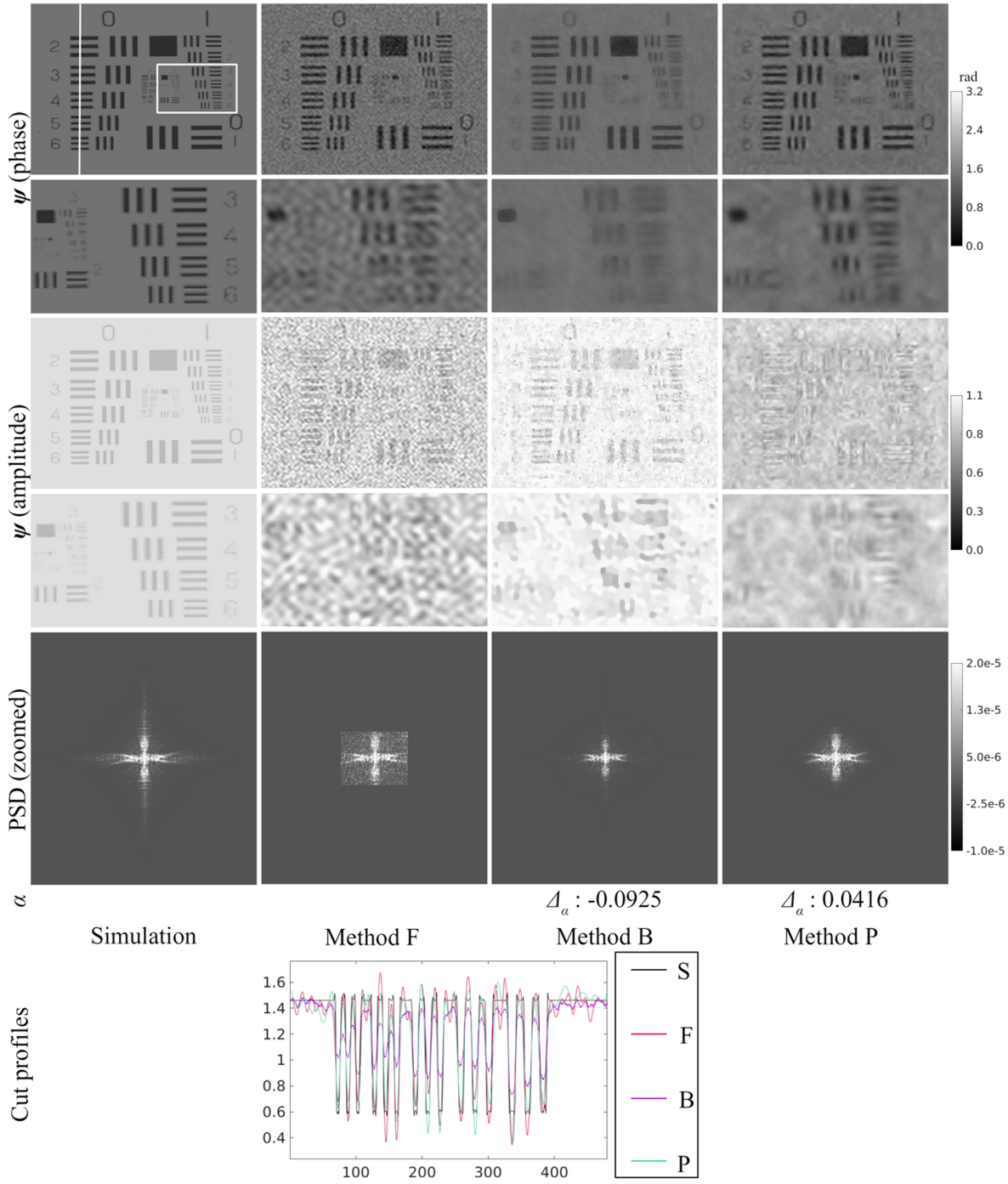

Fig. 7. Reconstruction of synthetic data with noise of SNR 10: comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B), and the proposed method (P).

Bourquard *et al.*). The pre-processing method by apodization described in Section 3.D is therefore applied before all three compared methods.

Figure 10 shows the results obtained with the proposed pre-processing. As expected, it permits the reduction of the artifacts due to camera frame truncation for all the considered methods. Compared to the two other methods, the proposed method (P) removes the zero- and minus one-order terms of the acquired spectrum more efficiently, providing gain in resolution. As for synthetic data, the direct Fourier method (F) yields ringing artifacts on the whole image (see oscillations in the flat areas). In addition, it produces, as is also the case for method of Bourquard *et al.* (B), oblique fringe artifacts (see orange arrows). Although the method of Bourquard *et al.* (B) uses spatially adaptive estimation of the $\boldsymbol{\alpha}$ map, it provides, as the direct Fourier method (F), lower quality reconstruction. As can be seen by comparison with Fig. 8, the $\boldsymbol{\alpha}$ map provided by method of Bourquard *et al.* (B) contains information not coming only from amplitude of the reference field but also from the observed measures including a large part of solution $\boldsymbol{\psi}$. It therefore does not correctly

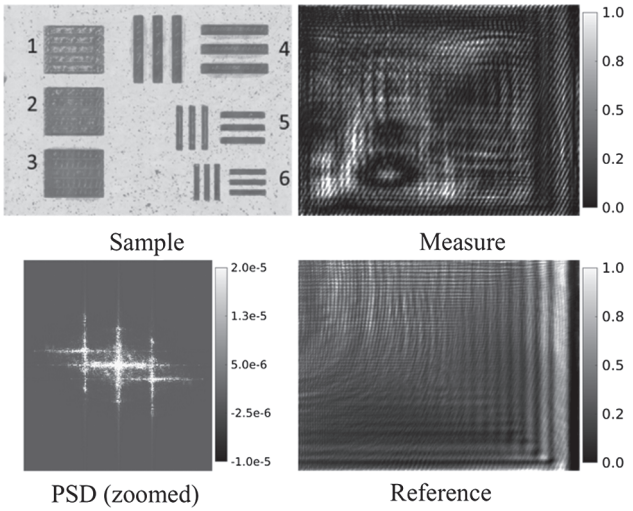

Fig. 8. Real data: white light image of the sample, recorded holograms with or without the presence of the object and PSD.

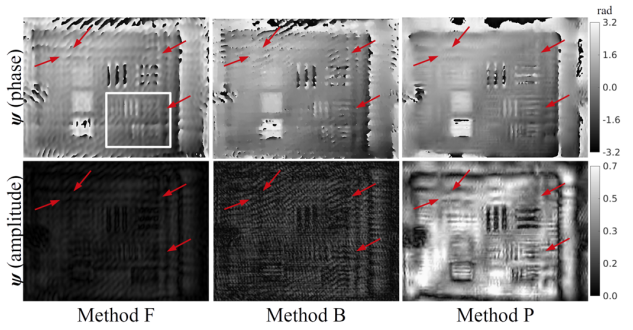

Fig. 9. Reconstruction of real data without pre-processing: comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B), and the proposed method (P) (red arrows indicate artifacts due to camera frame truncation).

separate solutions $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$. It is understandable since, by assuming spatial variability of $\boldsymbol{\alpha}$, the number of unknowns of the reconstruction problem is increased making the inverse problem even more ill-posed and therefore more difficult to solve. In addition, it over-estimates the mean of the $\boldsymbol{\alpha}$ map that produces, as explained in Section A.2, an underestimation of the average amplitude of $\boldsymbol{\psi}$ and oblique fringe artifacts in $\boldsymbol{\psi}$ in the direction of the fringes in the acquired hologram. The proposed method (P) generates an α-value certainly closer to the true value α_0 since it does not cause any oblique fringe artifacts, which results in better visibility of the patterns. Nevertheless, the scalar hypothesis for $\boldsymbol{\alpha}$ could be questioned since it does not permit modeling of the fringes in the reference field caused by the camera truncation. To investigate this question, we used one acquisition of the reference field to model $\boldsymbol{\alpha}$ as the square root of this acquisition modulated by a constant. Integrating this model in the proposed method delivers very similar results for the real experiments. This proves that such modeling does not bring any improvement compared to the constant modeling. In addition, the difficulties of the method of Bourquard *et al.* to deal with the complete spatial variability for $\boldsymbol{\alpha}$ suggest that a simpler model with few parameters, in particular, one single scalar, is a good approximation for modeling $\boldsymbol{\alpha}$.

We conclude from real experiments that the proposed method outperforms the two other methods.

5. CONCLUSION AND DISCUSSION

This work presented a convergent iterative method based on an inverse problem approach to image reconstruction of off-axis holograms. A key advantage of IP approaches is their resolution gain with flexible regularization priors, as well as their ability to model noise and to consider sub-sampled data. The method jointly reconstructs the object field as well as the amplitude of the reference field. The tuning of the hyperparameters is reliable due to their intuitive meaning. In addition, a protocol has been proposed for adjusting the hyperparameters in order to optimize the quality of the reconstruction. The method shows benefits due to the wavelet-based regularization: it can apply adaptive regularization with respect to the awaited resolution. For application in real situations, we proposed a pre-processing by apodization to reduce the artifacts due to the mismatch between the camera frame truncation and periodic boundary conditions assumed for our implementation. Experiments demonstrate improvements in terms of image quality such as gain in resolution and absence of fringe and ringing effects, compared to two other reconstruction methods, i.e., the direct Fourier method and the method of Bourquard *et al.* [16].

The protocol proposed for tuning the hyperparameters could be improved. Ideally, adjusting these hyperparameters could be done by optimizing additional metrics such as the sharpness of the object edges. Another possibility is to study the spectral characteristics of the residual noise in the measurement domain, as commonly done in image deconvolution methods [39,40] in order, e.g., to optimize the whiteness of the residual noise.

The experiments on real data suggested that a simple model for $\boldsymbol{\alpha}$ is a good approximation. Nevertheless, the difficulties with the separation of the solutions $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$ due to the ill-posedness of the reconstruction problem can certainly be reduced with other choices for the parameterization of $\boldsymbol{\alpha}$. We should analyze other modeling for $\boldsymbol{\alpha}$ requiring only few parameters to not increase the space of unknowns too much and thus not make the reconstruction problem more ill-posed and more difficult to solve. This should provide a finer comparison with the method of Bourquard *et al*. in terms of the influence of the spatially invariant or spatially variant hypothesis for $\boldsymbol{\alpha}$. As many other algorithms not intended for non-convex optimization, our algorithm has difficulties dealing with many non-optimal minima, which prevents the handling of complex models for $\boldsymbol{\alpha}$. Variants of the proximal gradient approach, initially developed for convex optimization, have been studied for solving non-convex problems [41,42]. We should examine if such optimization approaches can better deal with more complex models of $\boldsymbol{\alpha}$. An improved separation between solutions $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$ should be provided with an increase of information about $\boldsymbol{\alpha}$. Therefore, we should also study how to integrate extra measurements in the data-fidelity term to improve the quality of the reconstruction.

The pre-processing by apodization produces a decrease in intensity in the border area of the reconstructions. A solution to avoid this limitation would be to consider, instead of a periodic BC, an unknown BC for both solutions $\boldsymbol{\psi}$ and α, as proposed in [22,38].

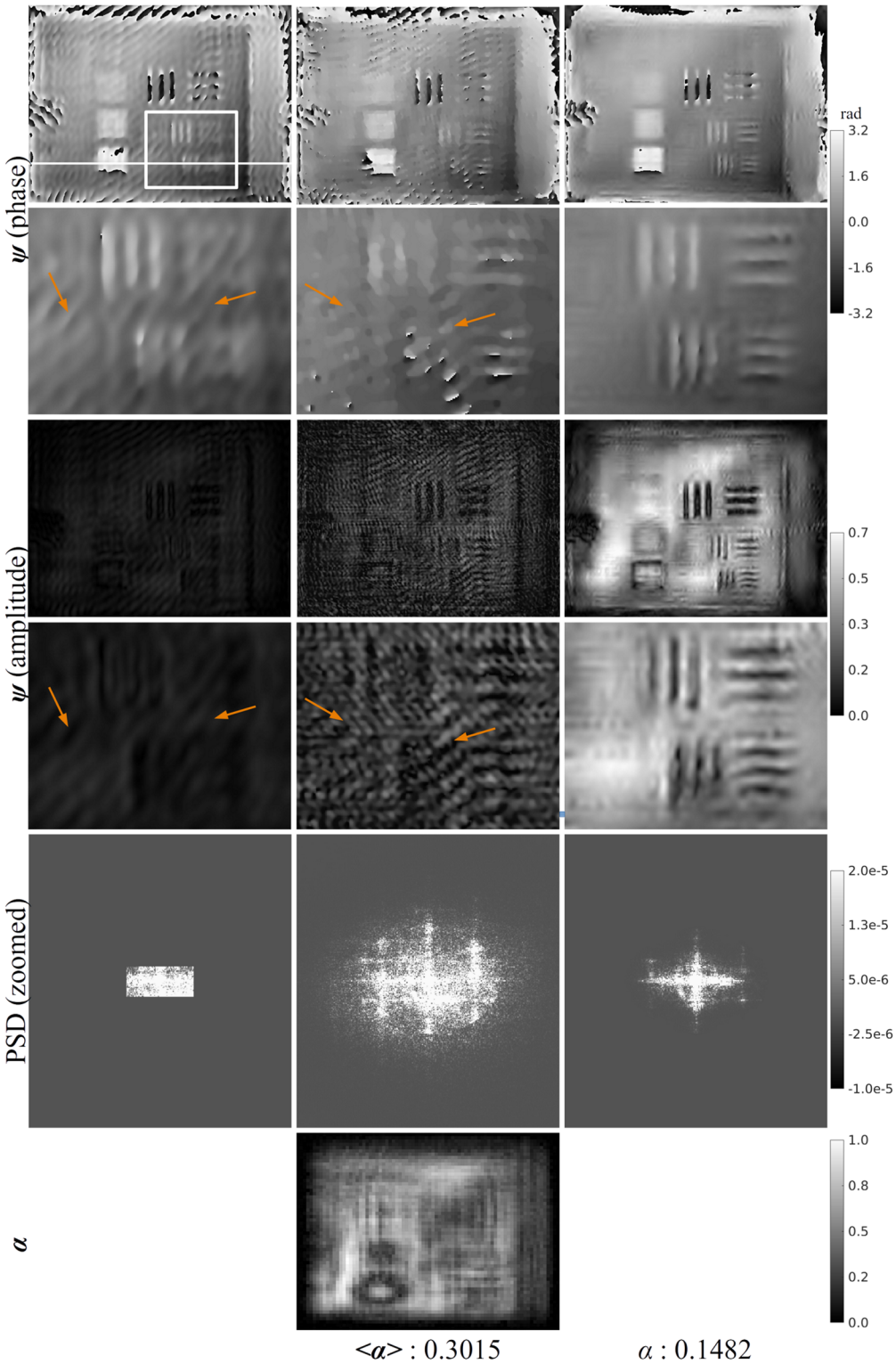

Fig. 10. Reconstruction of real data with pre-processing: comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B), and the proposed method (P) (orange arrows indicate oblique fringe artifacts).

The development of the proposed algorithm is inspired by the difficulties encountered in THz DH, and it is independent of the recording wavelengths. The potential of proposed algorithm should be studied in future works.

APPENDIX A: MATHEMATICAL NOTATIONS

In this paper, light symbols denote scalars (or scalar functions), and bold symbols refer to vectors and matrices (e.g., $\eta \in \mathbb{R}$, $g \in L_2(\mathbb{R})$, $\boldsymbol{f} \in \mathbb{R}^N$, $\boldsymbol{G} \in \mathbb{C}^{N\times N}$). Expression $f(x;\theta)$ denotes

a function f of variable x parameterized by θ. The convolution between two functions f and g is written $f \circledast g$. The kth component (iterate) of a vector $\boldsymbol{x}$ is denoted by x_k (x^k). The mean, the absolute value (taken element-wise), the ℓ_1-norms, and the ℓ_2- norms of a vector $\boldsymbol{x}$ are denoted by $\langle \boldsymbol{x} \rangle$, $|\boldsymbol{x}|$, $\| \boldsymbol{x} \|_1$, and $\| \boldsymbol{x} \|_2$, respectively. The element-wise product of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are denoted by $\boldsymbol{x} \odot \boldsymbol{y}$. The superscripts $.^\intercal$, $.^A$, $.^{-1}$, and $.^\dagger$ are used for the transpose, the adjoint, the inverse, and the Moore-Penrose pseudo-inverse of a matrix or an operator. All images are defined on a rectangular grid of $N = N_x \times N_y$ pixels and are represented by vectors of size N after vectorization in lexicographical order.

APPENDIX B: ADMMCP ALGORITHM AND EXTENSION

The ADMM is devoted to the solving of specific convex separable optimization problems, expressed, in the complex domain, as

$$\min\{f(\boldsymbol{x}) + g(\boldsymbol{z}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{z} = \boldsymbol{c}\}, \tag{B1}$$

where $\boldsymbol{x} \in C_x \subset \mathbb{C}^n$, $\boldsymbol{z} \in C_z \subset \mathbb{C}^m$, and C_x, C_z are convex sets, $\boldsymbol{A}$ and $\boldsymbol{B}$ are arbitrary matrices of $\mathbb{C}^{p\times n}$ and of $\mathbb{C}^{p\times m}$, respectively, $\boldsymbol{c}$ is a vector in $\mathbb{C}^p$, and f, g are proper, closed, convex real-valued functions. The ADMMCP is composed of the following iterations:

$$\begin{cases} \boldsymbol{x}^{k+1} = \operatorname{argmin}_{\boldsymbol{x}} f(\boldsymbol{x}) + \rho \left\| \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{z}^k - \boldsymbol{c} + \boldsymbol{u}^k \right\|_2^2 \\ \boldsymbol{z}^{k+1} = \operatorname{argmin}_{\boldsymbol{z}} g(\boldsymbol{z}) + \rho \left\| \boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{z} - \boldsymbol{c} + \boldsymbol{u}^k \right\|_2^2 \\ \boldsymbol{u}^{k+1} = \boldsymbol{u}^k + \boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{z}^{k+1} - \boldsymbol{c} \end{cases}, \tag{B2}$$

where $\boldsymbol{u}$ is the so-called scaled dual variable and ρ is a penalty parameter. Li *et al.* proved that under the assumptions that the extended-real-valued functions f, g are proper, closed, and convex functions and the Lagrangian function of the problem of Eq. (B1) has a saddle point, the ADMMCP converges.

We consider an extension of the original ADMMCP algorithm for adding variables in the problem of Eq. (B1). More precisely, we deal with the following generalization of Eq. (B1):

$$\min\{f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L) + \sum_{l=1}^{L} g_l(\boldsymbol{z}_l) : \forall 1 \le l \le L,\ \boldsymbol{A}_l\boldsymbol{x}_l + \boldsymbol{B}_l\boldsymbol{z}_l = \boldsymbol{c}_l\}, \tag{B3}$$

where $\boldsymbol{x}_l \in C_{x_l} \subset \mathbb{C}^{n_l}$, $\boldsymbol{z}_l \in C_{z_l} \subset \mathbb{C}^{m_l}$, $L \ge 1$, f is a proper, closed, convex function, C_{x_l} is a convex set and for all $1 \le l \le L$, C_{z_l} are convex sets, $\boldsymbol{A}_l$ and $\boldsymbol{B}_l$ are arbitrary matrices of $\mathbb{C}^{p_l\times n_l}$ and of $\mathbb{C}^{p_l\times m_l}$, respectively, c_l is a vector of $\mathbb{C}^{p_l}$, and g_l represents proper, closed, convex real-valued functions. We map the problem of Eq. (B3) into the form of Eq. (B1) as follows. First, we define matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as $\boldsymbol{A} = diag(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_L) \in \mathbb{C}^{p\times n}$, $\boldsymbol{B} = diag(\boldsymbol{B}_1, \ldots, \boldsymbol{B}_L) \in \mathbb{C}^{p\times m}$, where $p = p_1 + \ldots + p_L$, $n = n_1 + \ldots + n_L$, and $m = m_1 + \ldots + m_L$, we define the vectors $\boldsymbol{x}$, $\boldsymbol{z}$, $\boldsymbol{c}$, and the iterate $\boldsymbol{u}^k$ as

$$\begin{aligned} &\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L)^\intercal \in \mathbb{C}^n,\ \boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_L)^\intercal \in \mathbb{C}^m, \\ &\boldsymbol{c} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_L)^\intercal \in \mathbb{C}^p,\ \boldsymbol{u}^k = (\boldsymbol{u}_1^k, \ldots, \boldsymbol{u}_L^k)^\intercal \in \mathbb{C}^p, \end{aligned} \tag{B4}$$

and we define the function g on $\mathbb{C}^m$ as $g(\boldsymbol{z}) = \sum_{l=1}^{L} g_l(\boldsymbol{z}_l)$. The separable structure of g, the block structure of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, and the decomposition of Eq. (B4) allow us to rewrite the expression of iterations of Eq. (B2) in terms of the initial matrices, vectors, and variables of the problem of Eq. (B3), i.e., the matrices $\boldsymbol{A}_l$ and $\boldsymbol{B}_l$, the vectors $\boldsymbol{c}_l$, and the variables $\boldsymbol{z}_l$.

The $\boldsymbol{z}$-update in Eq. (B2) can be decoupled into L independent minimizations, each of the form $\boldsymbol{z}_l^{k+1} = \operatorname{argmin}_{\boldsymbol{z}_l} g(\boldsymbol{z}_l) + \rho \| \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l - \boldsymbol{c}_l + \boldsymbol{u}_l^k \|_2^2$, and, although function f is not separable, the $\boldsymbol{x}$-update in Eq. (B2) can be decoupled into L independent minimizations, each of the form $\boldsymbol{x}_l^{k+1} = \operatorname{argmin}_{\boldsymbol{x}_l} f_l(\boldsymbol{x}_l) + \rho \| \boldsymbol{A}_l\boldsymbol{x}_l + \boldsymbol{B}_l\boldsymbol{z}_l^k - \boldsymbol{c}_l + \boldsymbol{u}_l^k \|_2^2$, if we define f_l as the function f considered on the variable $\boldsymbol{x}_l$, the other variables being constant. The $\boldsymbol{u}$-update in Eq. (B2) can also be decoupled into L updates, each of the form $\boldsymbol{u}_l^{k+1} = \boldsymbol{u}_l^k + \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l^{k+1} - \boldsymbol{c}_l$. Consequently, the extended ADMMCP iterations are made of $3 \times L$ independent iterations corresponding to $3 \times L$ optimization problems:

$$\begin{cases} \boldsymbol{x}_l^{k+1} = \operatorname{argmin}_{\boldsymbol{x}_l} f_l(\boldsymbol{x}_l) + \rho \left\| \boldsymbol{A}_l\boldsymbol{x}_l + \boldsymbol{B}_l\boldsymbol{z}_l^k - \boldsymbol{c}_l + \boldsymbol{u}_l^k \right\|_2^2 \\ \boldsymbol{z}_l^{k+1} = \operatorname{argmin}_{\boldsymbol{z}_l} g_l(\boldsymbol{z}_l) + \rho \left\| \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l - \boldsymbol{c}_l + \boldsymbol{u}_l^k \right\|_2^2 \\ \boldsymbol{u}_l^{k+1} = \boldsymbol{u}_l^k + \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l^{k+1} - \boldsymbol{c}_l \end{cases}. \tag{B5}$$

This algorithm converges under the assumptions that the extended-real-valued functions f, g_l are proper, closed, and convex functions and the Lagrangian function of the problem of Eq. (B3) has a saddle point. The proof is direct: we easily show that, under these assumptions, the functions f, g defined for the mapping of the problem of Eq. (B3) into the form of Eq. (B1) are proper, closed, and convex functions and that the Lagrangian function of the problem of Eq. (B1) for this mapping has a saddle point too. Finally, if the matrices $\boldsymbol{A}_l$ and $\boldsymbol{B}_l$ are all not singular, we can rewrite the two first lines of iterations of Eq. (B5) by using proximity operators (see definition in Appendix C),

$$\begin{cases} \boldsymbol{x}_l^{k+1} = \operatorname{prox}_{\frac{1}{2\rho} f_l}(\boldsymbol{A}_l^{-1}(\boldsymbol{c}_l - \boldsymbol{B}_l\boldsymbol{z}_l^k - \boldsymbol{u}_l^k)) \\ \boldsymbol{z}_l^{k+1} = \operatorname{prox}_{\frac{1}{2\rho} g_l}(\boldsymbol{B}_l^{-1}(\boldsymbol{c}_l - \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} - \boldsymbol{u}_l^k)) \\ \boldsymbol{u}_l^{k+1} = \boldsymbol{u}_l^k + \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l^{k+1} - \boldsymbol{c}_l \end{cases}. \tag{B6}$$

APPENDIX C: PROXIMITY OPERATORS OF DATA-FIDELITY FUNCTION

Proximity operators, originally introduced in the real domain in [43], have been broadly used in signal processing where convex optimization in the real domain is required (see for example [22] and references therein). The proximity operator of any proper, closed, and convex function g with penalty γ can be generalized on the complex domain as $\operatorname{prox}_{\gamma g}(\boldsymbol{v}) = \operatorname{argmin}_{\boldsymbol{x}\in\mathbb{C}^n} \frac{1}{2} \| \boldsymbol{x} - \boldsymbol{v} \|_2^2 + \gamma g(\boldsymbol{x})$. In this appendix, we blindly assume that the data-fidelity function is convex, and we compute its proximity operators with respect to variables $\boldsymbol{c}_\psi$ and α.

We successively compute the proximity operator of the data-fidelity function D [see Eqs. (5) and (6)] for the minimization variables $\boldsymbol{c}_\psi$ and α. For α, we must take into account that it is a strictly positive real number, i.e., the complex vector $\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$ is in the direction of the unit vector $\boldsymbol{r}$. For this purpose, we take $\boldsymbol{\alpha}'$

as the minimization variable while α will be computed as $\langle|\boldsymbol{\alpha}'|\rangle$. According to that, we consider the following generic form of the data-fidelity function of the variable $\boldsymbol{x} \in \mathbb{C}^N$:

$$D(\boldsymbol{x}) = \| \boldsymbol{y} - |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|^2 \|_2^2 \tag{C1}$$

where

$$\boldsymbol{x} = \boldsymbol{c}_\psi,\ \boldsymbol{A} = \boldsymbol{A}_d \boldsymbol{W}^{-1},\ \boldsymbol{b} = \alpha \boldsymbol{r} \tag{C2}$$

for variable $\boldsymbol{c}_\psi$ and

$$\boldsymbol{x} = \boldsymbol{\alpha}' = \alpha \boldsymbol{r},\ \boldsymbol{A} = \boldsymbol{I},\ \boldsymbol{b} = \boldsymbol{o} = \boldsymbol{A}_d \boldsymbol{\psi} \tag{C3}$$

for variable α.

The proximity operator of the function of Eq. (C1) is given by

$$\text{prox}_{\gamma D}(\boldsymbol{v}) = \underset{\boldsymbol{x}}{\text{argmin}} \frac{1}{2} \| \boldsymbol{x} - \boldsymbol{v} \|_2^2 + \gamma \| \boldsymbol{y} - |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|^2 \|_2^2, \tag{C4}$$

which can be solved from the work of Brandwood [44] and requires canceling the conjugate gradient of the function to minimize. With the resulting equations being difficult to resolve, we proceed in another way. If we take the approximated assumption of a convex optimization context, the optimization of Eq. (C4) is equivalent to

$$\underset{\boldsymbol{x}}{\text{argmin}} \| \boldsymbol{x} - \boldsymbol{v} \|_2^2 : \left| \boldsymbol{y} - |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|^2 \right| \leq \boldsymbol{\delta},\ \sum_{j=1}^{N} \delta_j^2 = \epsilon,\ \epsilon > 0,$$

which, by assuming that all the components of the vector $\boldsymbol{\delta}$ are the same constant δ, can be solved by the following minimization:

$$\underset{\boldsymbol{x}}{\text{argmin}} \| \boldsymbol{x} - \boldsymbol{v} \|_2^2 : \left| y_j - \left| (\boldsymbol{A}\boldsymbol{x})_j + b_j \right|^2 \right| \leq \delta,\ \delta > 0,\ \forall 1 \leq j \leq N. \tag{C5}$$

We implement an approximation of the solution of problem Eq. (C5). First, we split it in two optimization problems: first we obtain $\boldsymbol{A}\boldsymbol{x}$ through

$$\underset{\boldsymbol{a}}{\text{argmin}} \| \boldsymbol{a} - \boldsymbol{A}\boldsymbol{v} \|_2^2 \text{ s.t. } \left| \boldsymbol{y} - |\boldsymbol{a} + \boldsymbol{b}|^2 \right| \leq \delta, \tag{C6}$$

and then we estimate $\boldsymbol{x}$ solving the well-known least-squares problem $\text{argmin}_{\boldsymbol{x}} \| \boldsymbol{x} \|_2^2$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{a}$, whose solution is $\boldsymbol{x} = \boldsymbol{A}^\dagger \boldsymbol{a}$. Second, we give an approximate solution of Eq. (C6) with the use of the projection operator initially designed by Schretter *et al.* [17] for projecting any vector $\boldsymbol{x}$ of the data domain on its sub-space of solutions $\boldsymbol{o}$ satisfying Eqs. (2) and (3) of the forward model. More precisely, since in Eq. (3) the noise $\boldsymbol{n}$ is unknown, Schretter *et al.* defined the projection operator on the sub-set $\boldsymbol{o} \in \mathbb{C}^N : |\boldsymbol{y} - |\boldsymbol{o} + \boldsymbol{r}|^2| \leq \epsilon^2$, where ϵ^2 is the variance of the noise on signal $\sqrt{\boldsymbol{y}}$. We consider here an extension in order to project on the sub-space of solutions $\boldsymbol{o}$ or of solutions $\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$ of Eqs. (2) and (3) of the forward model. In addition, for taking the noise into account, we rather define the tolerance margin as δ, where $\delta = \mu\sigma, \sigma$ is the standard deviation of the noise of the signal $\boldsymbol{y}$, and μ is a positive factor. Therefore, for application to problem Eq. (C6), we define the projection operator $\boldsymbol{P}_{\boldsymbol{b},\delta}$ on the set $\{\boldsymbol{a} \in \mathbb{C}^N : |\boldsymbol{y} - |\boldsymbol{a} + \boldsymbol{b}|^2| \leq \delta\}$ by

$$\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{a}) = \sqrt{\boldsymbol{y}_\delta} \frac{\boldsymbol{a} + \boldsymbol{b}}{|\boldsymbol{a} + \boldsymbol{b}|} - \boldsymbol{b},$$
$$\boldsymbol{y}_\delta = \begin{cases} \max(\boldsymbol{y} - \delta, \boldsymbol{0}) & \text{if } |\boldsymbol{a} + \boldsymbol{b}|^2 < \max(\boldsymbol{y} - \delta, \boldsymbol{0}) \\ \boldsymbol{y} + \delta & \text{if } |\boldsymbol{a} + \boldsymbol{b}|^2 > \boldsymbol{y} + \delta \\ |\boldsymbol{a} + \boldsymbol{b}|^2 & \text{otherwise} \end{cases}, \tag{C7}$$

where all operations are considered component-wise and the variable $\boldsymbol{b}$ can be $\boldsymbol{o}$ or $\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$. This provides an approximation of solution of Eq. (C6) as $\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}\boldsymbol{v})$, which yields the following approximated solution of Eq. (C4):

$$\boldsymbol{x} = \boldsymbol{A}^\dagger \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}\boldsymbol{v}). \tag{C8}$$

The approximate solution of Eq. (C4) is obtained in replacing in Eqs. (C7) and (C8) the variables $\boldsymbol{A}$ and $\boldsymbol{b}$ by their definitions [Eq. (C2) for variable $\boldsymbol{c}_\psi$ and Eq. (C3) for variable $\boldsymbol{\alpha}'$], which gives

$$\text{prox}_{\gamma D}(\boldsymbol{v}) \approx \boldsymbol{W}\boldsymbol{A}_d^\dagger \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}_d \boldsymbol{W}^{-1}\boldsymbol{v}),\ \boldsymbol{b} = \alpha \boldsymbol{r} \tag{C9}$$

for variable $\boldsymbol{c}_\psi$,

$$\text{prox}_{\gamma D}(\boldsymbol{v}) \approx \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{v}\boldsymbol{r}),\ \boldsymbol{b} = \boldsymbol{o} = \boldsymbol{A}_d \boldsymbol{\psi}$$

for variable $\boldsymbol{\alpha}'$, and finally

$$\text{prox}_{\gamma D}(\boldsymbol{v}) \approx \langle \left| \boldsymbol{P}_{\boldsymbol{b},\delta}(\upsilon \boldsymbol{r}) \right| \rangle,\ \boldsymbol{b} = \boldsymbol{o} = \boldsymbol{A}_d \boldsymbol{\psi} \tag{C10}$$

for variable α. Expressions of Eqs. (C9) and (C10) show that the proximity operators of the data-fidelity function are obtained by projecting each solution ($\boldsymbol{o}$ or $\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$) so that its combination with the other solution ($\boldsymbol{\alpha}' = \alpha \boldsymbol{r}$ or $\boldsymbol{o}$) is compatible with the measurements $\boldsymbol{y}$.

Funding. European Regional Development Fund/Wallonia region (2014-2020 "En Mieux" program, project TERA4ALL), "Fonds de la Recherche Scientifique – FNRS" (grant T.0136.20, project Learn2Sense).

Acknowledgment. This study could not have been accomplished without the work provided by two students, H. Blaschyk and L. Barriol, for their master internships. We also thank Juriy Hastanin (CSL) for helpful comments about the manuscript.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. Y. S. Lee, *Principles of Terahertz Science and Technology* (Springer, 2009), Vol. **170**.
2. A. Redo-Sanchez, N. Laman, B. Schulkin, and T. Tongue, "Review of terahertz technology readiness assessment and applications," J. Infrared Millim. Terahertz Waves **34**, 500–518 (2013).
3. M. Tonouchi, "Cutting-edge terahertz technology," Nat. Photonics **1**, 97–105 (2007).
4. H. Guerboukha, K. Nallappan, and M. Skorobogatiy, "Toward real-time terahertz imaging," Adv. Opt. Photonics **10**, 843–938 (2018).
5. D. M. Mittleman, "Twenty years of terahertz imaging [Invited]," Opt. Express **26**, 9417–9431 (2018).
6. L. Valzania, Y. Zhao, L. Rong, D. Wang, M. Georges, E. Hack, and P. Zolliker, "THz coherent lensless imaging," Appl. Opt. **58**, G256–G275 (2019).

7. M. Kim, "Principles and techniques of digital holographic microscopy," SPIE Rev. **1**, 018005 (2010).
8. Y. Zhao, J.-F. Vandenrijt, M. Kirkove, and M. Georges, "Iterative phase-retrieval-assisted off-axis terahertz digital holography," Appl. Opt. **58**, 9208–9216 (2019).
9. M. Kirkove, Y. Zhao, and M. P. Georges, "Inverse-problem-based algorithm for sparse reconstruction of Terahertz off-axis holograms," in *OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP)* (Optica Publishing Group, 2021), paper DM1B.4.
10. E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," Appl. Opt. **38**, 6994–7001 (1999).
11. N. Pavillon, C. S. Seelamantula, J. Kühn, M. Unser, and C. Depeursinge, "Suppression of the zero-order term in off-axis digital holography through nonlinear filtering," Appl. Opt. **48**, H186–H195 (2009).
12. C. S. Seelamantula, N. Pavillon, C. Depeursinge, and M. Unser, "Exact complex-wave reconstruction in digital holography," J. Opt. Soc. Am. A **28**, 983–992 (2011).
13. J. A. Fessler and S. Sotthivirat, "Simplified digital holographic reconstruction using statistical methods," in *International Conference on Image Processing-ICIP* (2004), Vol. **4**, pp. 2435–2438.
14. S. Sotthivirat and J. A. Fessler, "Reconstruction from digital holograms by statistical methods," in *37th Asilomar Conference on Signals, Systems and Computers* (2003), Vol. **2**, pp. 1928–1932.
15. S. Sotthivirat and J. A. Fessler, "Penalized-likelihood image reconstruction for digital holography," J. Opt. Soc. Am. A **21**, 737–750 (2004).
16. A. A. Bourquard, N. Pavillon, E. Bostan, C. Depeursinge, and M. Unser, "A practical inverse-problem approach to digital holographic reconstruction," Opt. Express **21**, 3417–3433 (2013).
17. C. Schretter, D. Blinder, S. Bettens, H. Ottevaere, and P. Schelkens, "Regularized non-convex image reconstruction in digital holographic microscopy," Opt. Express **25**, 16491–16508 (2017).
18. Y. Zhao, M. Kirkove, and M. P. Georges, "Inverse-problem based algorithm for THz off-axis digital holography reconstruction," in *Imaging and Applied Optics Congress* (Optica Publishing Group, 2020), paper HF4G.6.
19. C. A. González González, "Solving inverse problems in imaging using robust and regularized optimization," Ph.D. thesis (Université libre de Bruxelles, 2016).
20. F. Soulez, É. Thiébaut, A. Schutz, A. Ferrari, F. Courbin, and M. Unser, "Proximity operators for phase retrieval," Appl. Opt. **55**, 7412–7421 (2016).
21. D. S. Weller, A. Pnueli, G. Divon, O. Radzyner, Y. C. Eldar, and J. A. Fessler, "Undersampled phase retrieval with outliers," IEEE Trans. Comput. Imaging **1**, 247–258 (2015).
22. M. S. C. Almeida and M. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," IEEE Trans. Image Process. **22**, 3074–3086 (2013).
23. M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," IEEE Trans. Image Process. **12**, 906–916 (2003).
24. E. Cuche, P. Marquet, and C. Depeursinge, "Aperture apodization using cubic spline interpolation: application in digital holographic microscopy," Opt. Commun. **182**, 59–69 (2000).
25. F. Dubois, O. Monnom, C. Yourassowsky, and J.-C. Legros, "Border processing in digital holography by extension of the digital hologram and reduction of the higher spatial frequencies," Appl. Opt. **41**, 2621–2626 (2002).
26. D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inf. Theory **41**, 613–627 (1995).
27. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
28. S. Bettens, H. Yan, D. Blinder, H. Ottevaere, C. Schretter, and P. Schelkens, "Studies on the sparsifying operator in compressive digital holography," Opt. Express **25**, 18656–18676 (2017).
29. M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," Inverse Probl. **23**, 947 (2007).
30. S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. (Academic, 2008).
31. W. Sweldens, "The lifting scheme: a custom-design construction of biorthogonal wavelets," Appl. Comput. Harmon. Anal. **3**, 186–200 (1996).
32. P. Schelkens, A. Skodras, and T. Ebrahimi, *The JPEG 2000 Suite* (Wiley, 2009).
33. L. Li, X. Wang, and G. Wang, "Alternating direction method of multipliers for separable convex optimization of real functions in complex variables," Math. Probl. Eng. **2015**, 1–14 (2015).
34. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Found. Trends Mach. Learn. **3**, 1–122 (2011).
35. M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," SIAM J. Sci. Comput. **32**, 2710–2736 (2010).
36. K. Scheinberg, S. Ma, and D. Goldfarb, "Sparse inverse covariance selection via alternating linearization methods," in *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 2-NIPS* (Curran Associates, 2010), Vol. **23**, pp. 2101–2109.
37. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).
38. M. Kirkove, Y. Zhao, P. Blain, J.-F. Vandenrijt, and M. Georges, "Thermography-inspired processing strategy applied on shearography towards nondestructive inspection of composites," Proc. SPIE **11056**, 110560G (2019).
39. M. S. C. Almeida and M. A. T. Figueiredo, "Parameter estimation for blind and non-blind deblurring using residual whiteness measures," IEEE Trans. Image Process. **22**, 2751–2763 (2013).
40. A. Gonzalez, V. Delouille, and L. Jacques, "Non-parametric PSF estimation from celestial transit solar images using blind deconvolution," J. Space Weather Space Clim. **6**, A1 (2016).
41. N. Parikh and S. Boyd, "Proximal algorithms," Found. Trends Optim. **1**, 127–239 (2014).
42. H. Attouch, J. Bolte, and B. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods," Math. Program. **137**, 91–129 (2013).
43. J.-J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," C.R. Acad. Sci. Paris Ser. A Math. **255**, 2897–2899 (1962).
44. D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," IEEE Proc. F: Commun. Radar Signal Process. **130**, 11–16 (1983).