# 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*

**Abstract:** 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.

## 1. Introduction

Terahertz (THz) waves refer to electromagnetic radiations with a frequency of 0.1-10 THz, locating 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 $\psi$ 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 $\psi$ 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 which degrade the reference wave uniformity. Secondly, the light field truncation by the cameras and the diffraction effect are severe in THz. Finally, the pixel resolutions of the THz cameras are lower than the visible cameras. The above-mentioned problems inevitably deteriorate the reconstruction quality.

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 to recover 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 in estimating images by minimizing an objective/cost function promoting both a good fidelity to the available observations and specific image structures (such as sparsity or low-rankness). In off-axis DH, the reconstruction problem is complicated due to its non-convexity form and the difficulty to estimate 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 simpler 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 sub-space 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 $\ell_2$-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 in 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 truncation by the camera aperture is more important, what disturbs the homogeneity of the reference field amplitude with fringes. In [16], the $\boldsymbol{\alpha}$ map is defined on a sampled 2-D 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 lead 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 fringes 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 $\ell_1$-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

### 2.1. 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.*, an hologram.

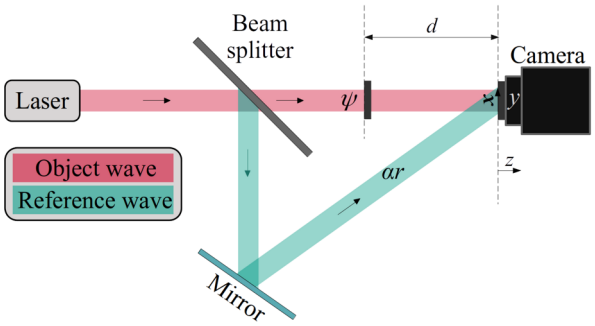



Fig. 1. Off-axis DH setup

### 2.2. Forward model

The general continuous DH model that we consider is standard in the literature [16]. In the case of a thin object, localized 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 instrumental point-spread-function (PSF). The expression of the PSF in the Fourier domain results from the successive application of the spatial-frequency cut-off function associated to the microscope objective and the digital propagation function with distance $d$ (see, *e.g.*, [16]).

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

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

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 $y$, $o$, $\psi$, $\alpha$ 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 $\alpha$ is spatially constant (see introduction), *i.e.*, $\boldsymbol{\alpha} = (\alpha, \ldots, \alpha)^\top$, 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 distortions 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 $\sigma^2$.

## 3. Reconstruction method

### 3.1. 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 $\alpha$; in other words, to solve the inverse problem related to the forward relation in equations (2)-(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 $\widetilde{\boldsymbol{\psi}}$ and $\widetilde{\alpha}$ 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 (2)-(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 $\alpha$. This can be seen by the cross-term obtained when developing the square modulus in (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 $\alpha$) is larger than the number of measurements (the dimension of $\boldsymbol{y}$) and it is unstable and highly sensitive to noise due to the frequency distribution

of the operator $\boldsymbol{A}_d$. Therefore, regularization is mandatory to select a meaningful solution among all compatible solutions.

We regularize the solution $\boldsymbol{\psi}$ in the wavelet domain by minimizing the $\ell_1$-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 (2) and the data-fidelity term (4) with respect to the wavelet coefficients $\boldsymbol{c}_\psi$ of $\boldsymbol{\psi}$ as

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

and

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

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

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

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

### 3.2. Reconstruction algorithm

The minimization problem (7) is non-convex. Despite the lack of guarantees of convergence towards 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)s that are part of a certain class of algorithms for solving constrained-optimization problems (COP)s 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 ADMMCP algorithm. We developed an extension of this algorithm to integrate the dependency of the optimization on the variable $\alpha$, which is presented in Appendix B. The application of the ADMMCP algorithm to off-axis DH reconstruction problem requires to compute 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 (11). We adapted it for solving problem (5)-(7). First, we express (7) in the form of (11) 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 (14) of the extended ADMMCP iterations. In addition, for the variable $x_2 = \alpha$, the equality $\operatorname{prox}_0(\boldsymbol{v}) = \boldsymbol{v}$ permits to reduce the three related iterations to the single iteration $\alpha^{k+1} = \operatorname{prox}_{\frac{1}{2\rho}D}(\alpha^k)$. Therefore, the

iterations of the extended ADMMCP become

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

The $\boldsymbol{c}_{\psi}$- and the $\alpha$-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 obtained by replacing in eqs. (23) and (24) $\boldsymbol{v}$ by $\boldsymbol{z}^k - \boldsymbol{u}^k$ and by $\alpha^k$ for $\gamma$ set at $\frac{1}{2\rho}$. The $z$-update is given by the proximity operator of the $\ell_1$-norm, which is the soft-thresholding operator extended to the complex domain [33]. If $\mathcal{S}_{\gamma}$ denotes this operator for the threshold $\gamma$, the $z$-update is $\boldsymbol{z}^{k+1} = \mathcal{S}_{\frac{\lambda}{2\rho}}(\boldsymbol{c}_{\psi}^{k+1} + \boldsymbol{u}^k)$. Since the $\boldsymbol{c}_{\psi}$- and the $\alpha$- updates do not depend on parameter $\rho$, the dependency to this parameter in the $z$-update can be removed and the previous iterations become

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

which leads to the following ADMMCP-based algorithm for problem (5)- (7).

**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**
  $\boldsymbol{b} \leftarrow \alpha^k\boldsymbol{r}$
  $\boldsymbol{c}_{\psi}^{k+1} \leftarrow \boldsymbol{W}\boldsymbol{A}_d^{\dagger}\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}_d\boldsymbol{W}^{-1}(\boldsymbol{z}^k - \boldsymbol{u}^k))$
  $\boldsymbol{z}^{k+1} \leftarrow \mathcal{S}_{\lambda}(\boldsymbol{c}_{\psi}^{k+1} + \boldsymbol{u}^k)$
  $\boldsymbol{u}^{k+1} \leftarrow \boldsymbol{u}^k + \boldsymbol{c}_{\psi}^{k+1} - \boldsymbol{z}^{k+1}$
  $\boldsymbol{b} \leftarrow \boldsymbol{A}_d\boldsymbol{W}^{-1}\boldsymbol{c}_{\psi}$
  $\alpha^{k+1} \leftarrow \langle\left|\boldsymbol{P}_{\boldsymbol{b},\delta}(\alpha^k\boldsymbol{r})\right|\rangle$
**end while**

This algorithm alternates between two consecutive optimizations, one with respect to $\boldsymbol{\psi}$ and one with respect to $\alpha$. Each optimization uses one projection operator ($\boldsymbol{P}_{\boldsymbol{b},\delta}$) based on the projector defined in [17] (see eq. (21)) 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 $\alpha$.

Since the cost function of problem (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 estimate. We set it as the solution of the direct Fourier method since it leads to a faster convergence. 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 parameter $\mu$ involved in the definition of the projection operators $\boldsymbol{P}_{\boldsymbol{b},\delta}$ at $\mu = 3$. The constant $\sigma$ that quantifies the noise level and

is also used in the definition of $\boldsymbol{P}_{\boldsymbol{b},\delta}$ can be seen as another regularization parameter due to its difficult estimation and its effect on the regularization (see eq. (21)). 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 $\lambda(l) = \lambda f^{-(l-1)} (f > 0)$. Our algorithm depends thus on only four processing parameters: the regularization strength $\lambda$, its modulation factor $f$, another regularization parameter $\sigma$ and the number of iterations $n_i$.

### 3.3. Protocol for adjusting reconstruction parameters

The adjustment of the reconstruction parameters is a complex work due to the high dimension of the parameters space. To provide a strategy, we established a protocol based on the analysis of the influences of the reconstruction parameters 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 $\boldsymbol{\psi}$ and its true distribution $\boldsymbol{\psi}_0$. 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 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 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 parameters $\lambda$ and $f$ are tuned for adjusting the regularization of the components of $\boldsymbol{\psi}$ according to their resolution. To that aim, the parameter $\lambda$ is first decreased as much as possible to better reconstruct the details of $\boldsymbol{\psi}$ while not adding artifacts and residual noise; then, the parameter $f$ is increased to get a sufficient level of details in the components of low resolution without adding artifacts, and finally the parameter $\lambda$ 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 parameter $\sigma$ is increased to decrease the possible residual noise and artifacts. In a final step, the parameter $n_i$ 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 yielding thus a solution independent of the choice of $n_i$ provided that it is larger than $n_{i0}$.

### 3.4. 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, that 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 decrease in intensity in the border area of the reconstructions. A solution that has not this limitation would be to consider, as proposed in [22, 38], for both solutions $\boldsymbol{\psi}$ and $\alpha$, a more realistic BC - an unknown BC- instead of a periodic BC. 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.

The experimental setup is described in fig. 1 section 2.1. It involves a laser (2.52 THz, wavelength $\lambda$ = 118.83 $\mu$m) and a camera (480 × 640 pixels with pitch of 17 $\mu$m).

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 parameter: the cutoff frequency of the low-pass filtering in the Fourier domain. The main processing parameters of the method of Bourquard *et al.* are $\epsilon_{SNR}$, the precision up to which the data-fidelity term is minimized converted into SNR, $\gamma$, the value balancing the respective influence of the phase and amplitude regularizations, $\mu$, the step size of the gradient-descent for TV regularization, $d_{\boldsymbol{\alpha}}$, 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 3.3. 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 to note that the adjustment of the parameters is more complicated for the method of Bourquard *et al.* than for the proposed method, due to non-convergence and interdependence of parameters $\epsilon_{SNR}$ and $\mu$.

### *4.1. Experiments on synthetic data*

#### 4.1.1. Data description

Tthe 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 simulation results are shown in the first column of Fig. 3 and Fig. 4. 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 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. For all the methods, the processing parameters 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 vary spatially or not. For a fair comparison, since the true $\boldsymbol{\alpha}$ map is spatially constant, we consider the second option.

#### 4.1.2. Results

Fig. 2 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 the best SSIMs on amplitude and the best SNRs on phase at all noise levels, the best SSIMs on phase and the best SNRs on amplitude at all noise levels except at the lowest one where

their values are close to the best ones given by the method of Bourquard *et al.* (B). This latter produces the highest standard deviations except for the SSIM on phase at the highest noise level. This shows a strong variation of the results qualities with respect to the noise realizations.

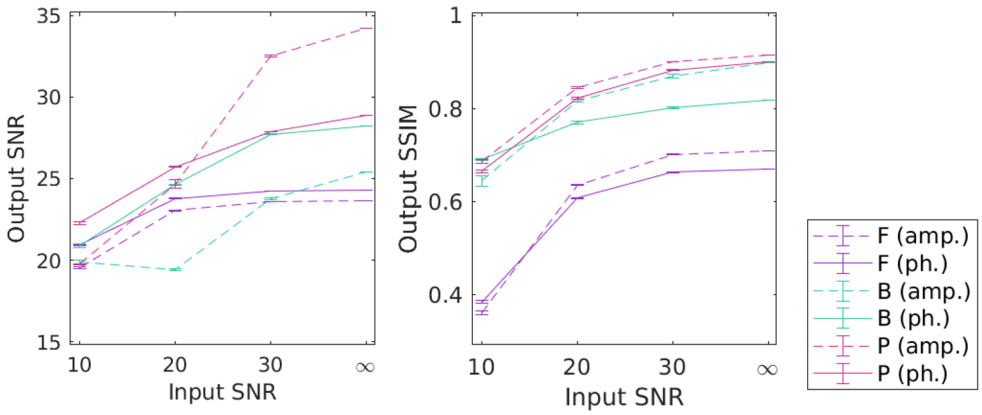



Fig. 2. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B) and the proposed method (P): reconstruction of synthetic data: performance measures with respect to noise level.

Fig. 3 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 resulting from the loss of high-frequency components. The magnitude of the amplitude of $\boldsymbol{\psi}$ is mis-estimated by the two other methods due to a mis-estimation of the $\boldsymbol{\alpha}$ mean. In comparison, the proposed method (P) yields a correct estimation of the $\boldsymbol{\alpha}$ mean resulting in a correct estimation of the magnitude of the amplitude of $\boldsymbol{\psi}$. In addition, it provides more signal details (slightly more for the phase of $\boldsymbol{\psi}$, in comparison to the method of Bourquard *et al.* (B)) resulting in better visibility of the patterns and better performance measures. The only weakness of the proposed method 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.

Fig. 4 illustrates the poorer performance of the proposed method at the lowest noise level. In this case, blocky artifacts strongly disturb the background of the amplitude of $\boldsymbol{\psi}$, preventing a correct identification of the patterns. In addition, the magnitude of the amplitude of $\boldsymbol{\psi}$ is underestimated. These defects explain the lower SNR on the amplitude for the proposed method. This conclusion must be nuanced since, for a highly transparent phase object, the main objective of the reconstruction is to retrieve its phase; and, compared to the two other methods, the proposed method (P) provides more signal details (slightly more compared to the method of Bourquard *et al.* (B)) on the phase.

We conclude from synthetic experiments that the proposed method outperforms the two other methods for low to medium noise levels. For high noise level, it outperforms the two other methods only for the reconstruction of the phase.

### *4.2. Experiments on real data*

#### 4.2.1. 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. 5(a)) is a visibly opaque 550 $\mu$m-thick polypropylene (PPP) slab with different engraved patterns. Polypropylene

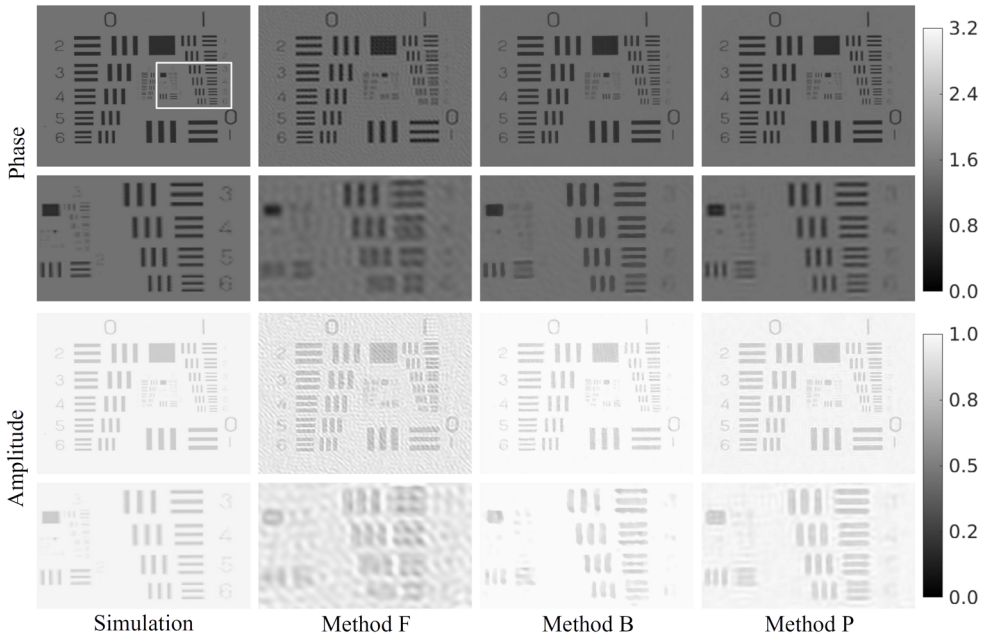



Fig. 3. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B) and the proposed method (P): reconstruction of synthetic data without noise

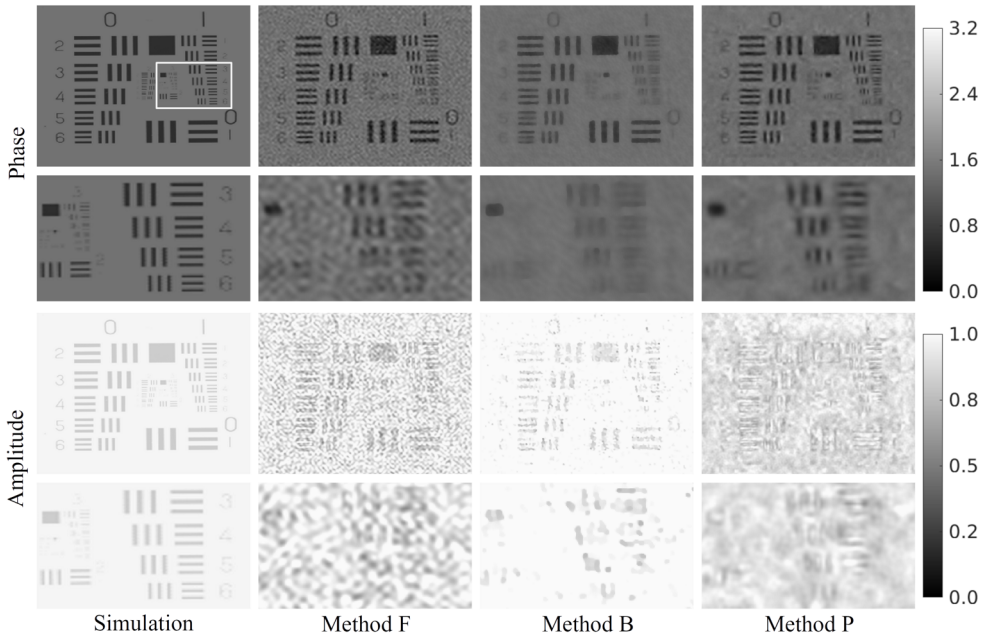
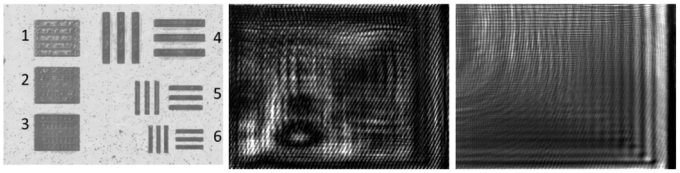



Fig. 4. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B) and the proposed method (P): reconstruction of synthetic data with noise of SNR 10

is highly transparent at 2.52 THz. Fig. 5(b) 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. 5(c)). In this case, 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 $\boldsymbol{\alpha}$.




(a) (b) (c)

Fig. 5. Real data: (a) White light image of the sample, (b-c) recorded holograms with (b) or without (c) the presence of the object.

#### 4.2.2. Results

We performed the reconstructions of the real data with all the considered methods. In the first part (the three first columns) of Fig. 6, we observed, in the results of the three methods, artifacts emanating from the boundaries of the images throughout the whole image. 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 3.4). 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 Bourquard *et al.* ). The pre-processing method by apodization described in section 3.4 is therefore applied before application of all three compared methods to reduce these artifacts.

The second part (the three last columns) of Fig. 6 shows the results obtained with the proposed pre-processing. As expected, it permits to reduce the artifacts due to camera frame truncation for all the considered methods. As for synthetic data, the direct Fourier method (F) yields ringing artifacts on the whole image. In addition, it produces, as is also the case for method of Bourquard *et al.* (B), fringe artifacts. Although the method of Bourquard *et al.* (B) uses spatially adaptive estimation of the $\boldsymbol{\alpha}$ map, it provides lower quality reconstruction. In comparison, the proposed method (P) produces improved image quality with better visibility of the patterns. It yields blocky artifacts but they are here faintly visible. As shown here and worth noted in section 3.4, the pre-processing by apodization produces decrease in intensity in the border area of the reconstructions. A solution that has not this limitation would be to consider, instead of a periodic BC, an unknown BC for both solutions $\boldsymbol{\psi}$ and $\alpha$, as proposed in [22, 38].

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

In THz, the amplitude of the reference field is disturbed by addition of fringes due to camera frame truncation. In this work, we assumed that the $\boldsymbol{\alpha}$ map is spatially constant by not modeling the fringes. This model is limited to provide a correct separation of the solutions $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$. To investigate the effect of modeling the fringes, we used one acquisition $\boldsymbol{y}_r$ of the reference field to approximate the fringes of $\boldsymbol{\alpha}$ and consider the model $\alpha\boldsymbol{\alpha}_n$ where $\alpha$ is a constant and $\boldsymbol{\alpha}_n$ is the normalized version of $\sqrt{\boldsymbol{y}_r}$. Integrating this model in the proposed method delivers very similar results. This proves that, a non-constant modeling of $\boldsymbol{\alpha}$ does not bring any improvement and thus validates our assumption for the model of $\boldsymbol{\alpha}$. Our algorithm, as many other algorithms not intended for non-convex optimization, has difficulties to deal with many non-optimal minima, these preventing the handling of a more complex model for $\boldsymbol{\alpha}$ requiring additional parameters.

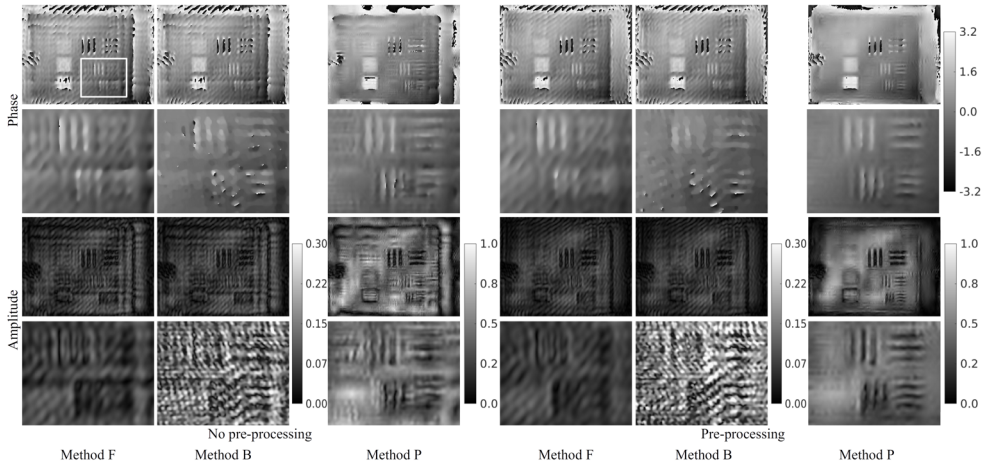



Fig. 6. Comparative study of three methods: the direct Fourier method (F), the method of Bourquard *et al.* (B) and the proposed method (P): reconstruction of real data

Variants of the proximal gradient approach, initially developed for convex optimization, have been studied for solving non-convex problems [39, 40]. It could be interesting to examine if such optimization approaches can better deal with a more complex model of $\alpha$ and thus can help to better separate solutions $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$.

The lack of information about $\boldsymbol{\alpha}$ can also be a cause of this incorrect separation. Therefore, in the future, we should also examine how to integrate additional prior information on $\boldsymbol{\alpha}$ by considering extra measurements in the data-fidelity term to improve the quality of the reconstruction.

## 5. Conclusions

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 its resolution gain with flexible regularization priors, its 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 reconstruction parameters is reliable due to their intuitive meaning. In addition, a protocol has been proposed for adjusting the parameters 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].

## A. Mathematical notations

In this paper, light symbols denote scalars (or scalar functions), and bold symbols refer to vectors and matrices (*e.g.*, $\eta \in \mathbb{R}$, $g \in L_2(\mathbb{R})$, $\boldsymbol{f} \in \mathbb{R}^N$, $\boldsymbol{G} \in \mathbb{C}^{N\times N}$). $f(x;\theta)$ denotes a function $f$ of variable $x$ parameterized by $\theta$. The convolution between two functions $f$ and $g$ is written $f \circledast g$. The $k$-th component (iterate) of a vector $\boldsymbol{x}$ is denoted by $x_k$ ($x^k$). The mean, the absolute value (taken element-wise), the $\ell_1$-, the $\ell_2$- 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 $.^\top$, $.^A$, $\cdot^{-1}$ and $\cdot^\dagger$ 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.

## B. ADMMCP algorithm and extension

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

$$\min \{f(\boldsymbol{x}) + g(\boldsymbol{z}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{z} = \boldsymbol{c}\} \tag{9}$$

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 and differentiable real-valued functions. The ADMMCP is composed of the following iterations:

$$\begin{cases} \boldsymbol{x}^{k+1} &= \operatorname{argmin}_{\boldsymbol{x}} f(\boldsymbol{x}) + \rho\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{z}^k - \boldsymbol{c} + \boldsymbol{u}^k\|_2^2 \\ \boldsymbol{z}^{k+1} &= \operatorname{argmin}_{\boldsymbol{z}} g(\boldsymbol{z}) + \rho\|\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{z} - \boldsymbol{c} + \boldsymbol{u}^k\|_2^2 \\ \boldsymbol{u}^{k+1} &= \boldsymbol{u}^k + \boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{z}^{k+1} - \boldsymbol{c} \end{cases}, \tag{10}$$

where $\boldsymbol{u}$ is the so-called scaled dual variable and $\rho$ 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 problem (9) has a saddle point, the ADMMCP converges.

We consider an extension of the original ADMMCP algorithm for adding variables in problem (9). More precisely, we deal with the following generalization of (9):

$$\min\{f(\boldsymbol{x}_1, ..., \boldsymbol{x}_L) + \sum_{l=1}^{L} g_l(\boldsymbol{z}_l) : \begin{array}{c} \forall 1 \le l \le L, \\ \boldsymbol{A}_l\boldsymbol{x}_l + \boldsymbol{B}_l\boldsymbol{z}_l = \boldsymbol{c}_l \end{array}\}, \tag{11}$$

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 and differentiable 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 $g_l$ are proper, closed, convex and differentiable real-valued functions. We map problem (11) into form (9) as follows. First, we define matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as $\boldsymbol{A} = diag(\boldsymbol{A}_1, ..., \boldsymbol{A}_L) \in \mathbb{C}^{p\times n}$, $\boldsymbol{B} = diag(\boldsymbol{B}_1, ..., \boldsymbol{B}_L) \in \mathbb{C}^{p\times m}$ where $p = p_1 + ... + p_L$, $n = n_1 + ... + n_L$, and $m = m_1 + ... + m_L$, we define the vectors $\boldsymbol{x}$, $z$, $\boldsymbol{c}$ and the iterate $\boldsymbol{u}^k$ as

$$\begin{aligned} &\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_L)^\top \in \mathbb{C}^n \quad , z = (z_1, ..., z_L)^\top \in \mathbb{C}^m \\ &\boldsymbol{c} = (\boldsymbol{c}_1, ..., \boldsymbol{c}_L)^\top \in \mathbb{C}^p \quad , \boldsymbol{u}^k = (\boldsymbol{u}_1^k, ..., \boldsymbol{u}_L^k)^\top \in \mathbb{C}^p, \end{aligned} \tag{12}$$

and we define the function $g$ on $\mathbb{C}^m$ as $g(z) = \sum_{l=1}^{L} g_l(z_l)$. The separable structure of $g$, the block structure of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and the decomposition (12) allow to rewrite the expression of iterations (10) in terms of the initial matrices, vectors and variables of the problem (11), i.e. the matrices $\boldsymbol{A}_l$ and $\boldsymbol{B}_l$, the vectors $\boldsymbol{c}_l$ and the variables $z_l$.

The $z$-update in eq. (10) can be decoupled into $L$ independent minimizations, each of the form $\boldsymbol{z}_l^{k+1} = \operatorname{argmin}_{z_l} g(z_l) + \rho\|\boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l z_l - \boldsymbol{c}_l + \boldsymbol{u}_l^k\|_2^2$ and, although function $f$ is not separable, the $\boldsymbol{x}$-update in eq. (10) 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. (10) 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:

$$\left\{\begin{array}{lcl} \boldsymbol{x}_l^{k+1} & = & \text{argmin}_{\,\boldsymbol{x}_l} f_l(\boldsymbol{x}_l) \\ & & +\rho\|\boldsymbol{A}_l\boldsymbol{x}_l + \boldsymbol{B}_l\boldsymbol{z}_l^k - \boldsymbol{c}_l + \boldsymbol{u}_l^k\|_2^2 \\ \boldsymbol{z}_l^{k+1} & = & \text{argmin}_{\,\boldsymbol{z}_l} g_l(\boldsymbol{z}_l) \\ & & +\rho\|\boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l - \boldsymbol{c}_l + \boldsymbol{u}_l^k\|_2^2 \\ \boldsymbol{u}_l^{k+1} & = & \boldsymbol{u}_l^k + \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l^{k+1} - \boldsymbol{c}_l \end{array}\right. . \tag{13}$$

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 problem (11) has a saddle point. The proof is direct: we easily show that, under these assumptions, the functions $f$, $g$ defined for the mapping of problem (11) into form (9) are proper, closed and convex functions and that the Lagrangian function of problem (9) 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 (13) by using proximity operators (see definition in Appendix C):

$$\left\{\begin{array}{lcl} \boldsymbol{x}_l^{k+1} & = & \text{prox}\,\frac{1}{2\rho} f_l(\boldsymbol{A}_l^{-1}(\boldsymbol{c}_l - \boldsymbol{B}_l\boldsymbol{z}_l^k - \boldsymbol{u}_l^k)) \\ \boldsymbol{z}_l^{k+1} & = & \text{prox}\,\frac{1}{2\rho} g_l(\boldsymbol{B}_l^{-1}(\boldsymbol{c}_l - \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} - \boldsymbol{u}_l^k)) \\ \boldsymbol{u}_l^{k+1} & = & \boldsymbol{u}_l^k + \boldsymbol{A}_l\boldsymbol{x}_l^{k+1} + \boldsymbol{B}_l\boldsymbol{z}_l^{k+1} - \boldsymbol{c}_l \end{array}\right. . \tag{14}$$

## C. Proximity operators of data-fidelity function

Proximity operators, originally introduced in the real domain in [41], 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 $\gamma$ can be generalized on the complex domain as $\text{prox}_{\gamma g}(\boldsymbol{v}) = \text{argmin}_{\,\boldsymbol{x}\in\mathbb{C}^n} \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{v}\|_2^2 + \gamma g(\boldsymbol{x})$. 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 $\alpha$.

We successively compute the proximity operator of the data-fidelity function $D$ (see eqs. (5)-(6)) for the minimization variables $\boldsymbol{c}_\psi$ and $\alpha$. For $\alpha$, 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 $\alpha$ 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 $\boldsymbol{x} \in \mathbb{C}^N$:

$$D(\boldsymbol{x}) = \|\boldsymbol{y} - |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|^2\|_2^2 \tag{15}$$

where

$$\boldsymbol{x} = \boldsymbol{c}_\psi, \boldsymbol{A} = \boldsymbol{A}_d\boldsymbol{W}^{-1}, \boldsymbol{b} = \alpha\boldsymbol{r} \tag{16}$$

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

$$\boldsymbol{x} = \boldsymbol{\alpha}' = \alpha\boldsymbol{r}, \boldsymbol{A} = \boldsymbol{I}, \boldsymbol{b} = \boldsymbol{o} = \boldsymbol{A}_d\boldsymbol{\psi} \tag{17}$$

for variable $\alpha$.

The proximity operator of function (15) is given by

$$\text{prox}_{\gamma D}(\boldsymbol{v}) = \underset{\boldsymbol{x}}{\text{argmin}}\ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{v}\|_2^2 + \gamma\|\boldsymbol{y} - |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|^2\|_2^2, \tag{18}$$

which can be solved from the work of Brandwood [42] and requires canceling the conjugate gradient of the function to minimize. 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 (18) is equivalent to

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

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

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{x}-\boldsymbol{v}\|_2^2 : \left|y_j - \left|(\boldsymbol{A}\boldsymbol{x})_j + b_j\right|^2\right| \leq \delta,\ \delta > 0, \forall 1 \leq j \leq N. \tag{19}$$

We implement an approximation of the solution of problem (19). First, we split it in two optimization problems: first we obtain $\boldsymbol{A}\boldsymbol{x}$ through

$$\underset{\boldsymbol{a}}{\operatorname{argmin}} \|\boldsymbol{a}-\boldsymbol{A}\boldsymbol{v}\|_2^2 \text{ s.t. } \left|\boldsymbol{y}-|\boldsymbol{a}+\boldsymbol{b}|^2\right| \leq \delta, \tag{20}$$

then we estimate $\boldsymbol{x}$ solving the well-known least-squares (LS) problem $\underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{x}\|_2^2$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{a}$ whose solution is $\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{a}$. Secondly, we give an approximate solution of (20) 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 equations (2) - (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 $\boldsymbol{o} \in \mathbb{C}^N : \left|\boldsymbol{y}-|\boldsymbol{o}+\boldsymbol{r}|^2\right| \leq \epsilon^2$ where $\epsilon^2$ 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 equations (2) - (3) of the forward model. In addition, for taking the noise into account, we rather define the tolerance margin as $\delta$ where $\delta = \mu\sigma$ and $\sigma$ is the standard deviation of the noise of the signal $\boldsymbol{y}$ and $\mu$ is a positive factor. Therefore, for application to problem (20), we define the projection operator $\boldsymbol{P}_{\boldsymbol{b},\delta}$ on the set $\left\{\boldsymbol{a} \in \mathbb{C}^N : \left|\boldsymbol{y}-|\boldsymbol{a}+\boldsymbol{b}|^2\right| \leq \delta\right\}$ by

$$\begin{aligned} \boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{a}) &= \sqrt{\boldsymbol{y}_\delta}\frac{\boldsymbol{a}+\boldsymbol{b}}{|\boldsymbol{a}+\boldsymbol{b}|} - \boldsymbol{b}, \\ \boldsymbol{y}_\delta &= \begin{cases} \max(\boldsymbol{y}-\delta, \boldsymbol{0}) & \text{if } |\boldsymbol{a}+\boldsymbol{b}|^2 < \max(\boldsymbol{y}-\delta, \boldsymbol{0}) \\ \boldsymbol{y}+\delta & \text{if } |\boldsymbol{a}+\boldsymbol{b}|^2 > \boldsymbol{y}+\delta \\ |\boldsymbol{a}+\boldsymbol{b}|^2 & \text{otherwise.} \end{cases} \end{aligned} \tag{21}$$

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 (20) as $\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}\boldsymbol{v})$, which yields the following approximated solution of (18):

$$\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}\boldsymbol{v}). \tag{22}$$

The approximate solution of (18) is obtained in replacing in eqs. (21)-(22) the variables $\boldsymbol{A}$ and $\boldsymbol{b}$ by their definitions (eq. (16) for variable $\boldsymbol{c}_\psi$ and eq. (17) for variable $\boldsymbol{\alpha}'$), which gives

$$\operatorname{prox}_{\gamma D}(\boldsymbol{v}) \approx \boldsymbol{W}\boldsymbol{A}_d^{\dagger}\boldsymbol{P}_{\boldsymbol{b},\delta}(\boldsymbol{A}_d\boldsymbol{W}^{-1}\boldsymbol{v}),\ \boldsymbol{b} = \boldsymbol{\alpha}\boldsymbol{r} \tag{23}$$

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

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

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

$$\operatorname{prox}_{\gamma D}(\boldsymbol{v}) \approx \langle\left|\boldsymbol{P}_{\boldsymbol{b},\delta}(v\boldsymbol{r})\right|\rangle,\ \boldsymbol{b} = \boldsymbol{o} = \boldsymbol{A}_d\boldsymbol{\psi} \tag{24}$$

for variable $\alpha$. Expressions (23) and (24) 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

The research presented in this paper was funded by project TERA4ALL in the frame of 2014-2020 ERDF/Wallonia region "En Mieux" program and supported by the "Fonds de la Recherche Scientifique – FNRS" under 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.

## References

1. Y. S. Lee, *Principles of Terahertz Science and Technology*, vol. 170 (Springer Science & Business Media, 2009).
2. A. Redo-Sanchez, N. Laman, B. Schulkin, and T. Tongue, "Review of Terahertz Technology Readiness Assessment and Applications," J. Infrared, Millimeter, 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. Photon. **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. Vol. 1, id. 018005 (2010). **1**, 8005 (2009).
8. Y. Zhao, J.-F. Vandenrijt, M. Kirkove, and M. Georges, "Iterative phase-retrieval-assisted off-axis terahertz digital holography," Appl. Opt. **58**, 9208 (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), 4, p. 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 (2011).
13. J. A. Fessler and S. Sotthivirat, "Simplified digital holographic reconstruction using statistical methods," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, vol. 4 (2004), pp. 2435–2438 Vol. 4.
14. S. Sotthivirat and J. A. Fessler, "Reconstruction from digital holograms by statistical methods," in *The Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2003,* vol. 2 (2003), pp. 1928–1932 Vol.2.
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 (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 (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), p. HF4G.6.
19. C. A. González González, "Solving Inverse Problems in Imaging using Robust and Regularized Optimization," Ph.D. thesis, Université libre de Bruxelles, Belgium (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. on 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. on 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. on 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. on 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, Third Edition: The Sparse Way* (Academic Press, Inc., USA, 2008), 3rd ed.
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. P. Boyd, N. Parikh, E. K. wah 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. on 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,* vol. 23 of *NIPS'10* (Curran Associates Inc., Red Hook, NY, USA, 2010), 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. on Image Process. **13**, 600–612 (2004).
38. M. Kirkove, Y. Zhao, P. Blain, J.-F. F. Vandenrijt, and M. Georges, "Thermography-inspired processing strategy applied on shearography towards nondestructive inspection of composites," in *Optical Measurement Systems for Industrial Inspection XI,* vol. 11056 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*International Society for Optics and Photonics (SPIE, 2019), p. 110560G.
39. N. Parikh and S. Boyd, "Proximal algorithms," Found. Trends Optim. **1**, 127–239 (2014).
40. 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** (2011).
41. 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).
42. 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).