Research Article | Vol. 25, No. 18 | 4 Sep 2017 | OPTICS EXPRESS 21786
Optics EXPRESS

Efficient inversion of multiple-scattering model for optical diffraction tomography

EMMANUEL SOUBIES,[1,2,*] THANH-AN PHAM,[1,2] AND MICHAEL UNSER[1]

[1] *Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[2] *These authors contributed equally to this paper*
[*] *emmanuel.soubies@epfl.ch*

Abstract: Optical diffraction tomography relies on solving an inverse scattering problem governed by the wave equation. Classical reconstruction algorithms are based on linear approximations of the forward model (Born or Rytov), which limits their applicability to thin samples with low refractive-index contrasts. More recent works have shown the benefit of adopting nonlinear models. They account for multiple scattering and reflections, improving the quality of reconstruction. To reduce the complexity and memory requirements of these methods, we derive an explicit formula for the Jacobian matrix of the nonlinear Lippmann-Schwinger model which lends itself to an efficient evaluation of the gradient of the data-fidelity term. This allows us to deploy efficient methods to solve the corresponding inverse problem subject to sparsity constraints.

© 2017 Optical Society of America

OCIS codes: (180.1655) Coherence tomography; (180.3170) Interference microscopy; (100.3010) Image reconstruction techniques.

References and links

1. E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," Opt. Commun. **1**, 153–156 (1969).
2. D. Jin, R. Zhou, Z. Yaqoob, and P. T. So, "Tomographic phase microscopy: Principles and applications in bioimaging," J. Opt. Soc. Am. B **34**, B64–B77 (2017).
3. A. Devaney, "Inverse-scattering theory within the Rytov approximation," Opt. Lett. **6**, 374–376 (1981).
4. E. Mudry, P. C. Chaumet, K. Belkebir, and A. Sentenac, "Electromagnetic wave imaging of three-dimensional targets using a hybrid iterative inversion method," Inverse Probl. **28**, 065007 (2012).
5. U. S. Kamilov, D. Liu, H. Mansour, and P. T. Boufounos, "A recursive Born approach to nonlinear inverse scattering," IEEE Signal Process. Lett. **23**, 1052–1056 (2016).
6. H.-Y. Liu, U. S. Kamilov, D. Liu, H. Mansour, and P. T. Boufounos, "Compressive imaging with iterative forward models," in "IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)," (IEEE, 2017), pp. 6025–6029.
7. H.-Y. Liu, D. Liu, H. Mansour, P. T. Boufounos, L. Waller, and U. S. Kamilov, "SEAGLE: Sparsity-driven image reconstruction under multiple scattering," arXiv preprint arXiv:1705.04281 (2017).
8. D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, vol. 93 (Springer Science & Business Media, 2012).
9. J. A. Schmalz, G. Schmalz, T. E. Gureyev, and K. M. Pavlov, "On the derivation of the Green's function for the Helmholtz equation using generalized functions," Am. J. Phys. **78**, 181–186 (2010).
10. A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (SIAM, 2001).
11. W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, and M. S. Feld, "Tomographic phase microscopy," Nat. Methods **4**, 717 (2007).
12. B. Chen and J. J. Stamnes, "Validity of diffraction tomography based on the first Born and the first Rytov approximations," Appl. Opt. **37**, 2996–3006 (1998).
13. Y. Sung, W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, and M. S. Feld, "Optical diffraction tomography for high resolution live cell imaging," Opt. Express **17**, 266–277 (2009).
14. Y. Sung and R. R. Dasari, "Deterministic regularization of three-dimensional optical diffraction tomography," J. Opt. Soc. Am. A **28**, 1554–1561 (2011).
15. J. Lim, K. Lee, K. H. Jin, S. Shin, S. Lee, Y. Park, and J. C. Ye, "Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography," Opt. Express **23**, 16933–16948 (2015).
16. P. C. Chaumet and K. Belkebir, "Three-dimensional reconstruction from real data using a conjugate gradient-coupled dipole method," Inverse Probl. **25**, 024003 (2009).

#302238 https://doi.org/10.1364/OE.25.021786
Journal © 2017 Received 11 Jul 2017; revised 23 Aug 2017; accepted 23 Aug 2017; published 29 Aug 2017

Optics EXPRESS

17. K. Belkebir, P. C. Chaumet, and A. Sentenac, "Superresolution in total internal reflection tomography," J. Opt. Soc. Am. A **22**, 1889–1897 (2005).
18. A. Abubakar and P. M. van den Berg, "The contrast source inversion method for location and shape reconstructions," Inverse Probl. **18**, 495 (2002).
19. U. Kamilov, I. Papadopoulos, M. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," Optica **2**, 517–522 (2015).
20. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imaging Sci. **2**, 183–202 (2009).
21. Y. Nesterov, "A method of solving a convex programming problem with convergence rate $O(1/k^2)$," Soviet Math. Dokl. **27**, 372–376 (1983).
22. P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," Multiscale Model Simul. **4**, 1168–1200 (2006).
23. Y. Nesterov, "Gradient methods for minimizing composite functions," Math. Prog. **140**, 125–161 (2013).
24. H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods," Math. Prog. **137**, 91–129 (2013).
25. A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," IEEE Trans. Image Process. **18**, 2419–2434 (2009).
26. A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," J. Math. Imaging Vision **40**, 120–145 (2011).
27. U. S. Kamilov, "A parallel proximal algorithm for anisotropic total variation minimization," IEEE Trans. Image Process. **26**, 539–548 (2017).
28. 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).
29. S. Lefkimmiatis, J. Ward, and M. Unser, "Hessian Schatten-norm regularization for linear inverse problems," IEEE Trans. Image Process. **22**, 1873–1888 (2013).
30. L. Bottou, "Large-scale machine learning with stochastic gradient descent," in "Proceedings of COMPSTAT'2010: 19th International Conference on Computational StatisticsParis France, August 22-27, 2010 Keynote, Invited and Contributed Papers," Y. Lechevallier and G. Saporta, eds. (Physica-Verlag HD, Heidelberg, 2010), pp. 177–186.
31. A. J. Devaney, *Mathematical Foundations of Imaging, Tomography and Wavefield Inversion* (Cambridge University Press, 2012).
32. J. A. Stratton, *Electromagnetic Theory* (John Wiley & Sons, 2007).
33. M. Unser, E. Soubies, F. Soulez, M. McCann, and L. Donati, "GlobalBioIm: A unifying computational framework for solving inverse problems," in "Proceedings of the OSA Imaging and Applied Optics Congress on Computational Optical Sensing and Imaging (COSI'17)," (San Francisco CA, USA, 2017). Paper no. CTu1B.
34. Jean-Michel Geffrin, Pierre Sabouroux and Christelle Eyraud, "Free space experimental scattering database continuation: experimental set-up and measurement precision," Inverse Probl. **21**, 6 (2005).
35. J.-J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," C. R. Acad. Sci. Ser. A Math. **255**, 2897–2899 (1962).

1. Introduction

Optical diffraction tomography (ODT) was introduced in [1] by E. Wolf in the late '60s. It is a microscopic technique that retrieves the distribution of refractive indices in biological samples out of holographic measurements of the scattered complex field produced when the sample is illuminated by an incident wave. This method is of particular interest in biology because, contrarily to fluorescence imaging, it does not require any staining of the sample [2]. It proceeds by solving an inverse scattering problem, where the scattering phenomenon is governed by the wave equation. There is a vast literature on inversion methods going from linearized models (Born, Rytov) [1, 3] to nonlinear ones [4–7]. It is worth noting that the scattering model, along with its associated inverse problem, is generic and not limited to optical diffraction tomography. In particular, it is encountered in many other fields such as acoustics, microwave imaging, or radar applications [8].

1.1. From the wave equation to the Lippmann-Schwinger integral equation

Let us consider an unknown object of refractive index $n(\mathbf{x})$ lying in the region $\Omega \subseteq \mathbb{R}^D$ ($D \in \{2, 3\}$) and being immersed in a medium of refractive index n_b, as depicted in Fig. 1. This sample is

Optics EXPRESS

illuminated by the incident plane wave

$$u^{\text{in}}(\mathbf{x},t) = \text{Re}\left(u_0 \text{e}^{\text{i}\mathbf{k}\cdot\mathbf{x}-\text{i}\omega t}\right), \tag{1}$$

where the wave vector $\mathbf{k} \in \mathbb{R}^D$ specifies the direction of the wave propagation, $\omega \in \mathbb{R}$ denotes its angular frequency, and $u_0 \in \mathbb{C}$ defines its complex envelope (amplitude). The resulting total electric field $u(\mathbf{x},t)$ satisfies the wave equation

$$\nabla^2 u(\mathbf{x},t) - \frac{n^2(\mathbf{x})}{\text{c}^2}\frac{\partial^2 u}{\partial t^2}(\mathbf{x},t) = 0, \tag{2}$$

where $\text{c} \simeq 3\times10^8\text{m/s}$ is the speed of light in free space. Denoting by $u(\mathbf{x})$ the complex amplitude of $u(\mathbf{x},t) = \text{Re}\left(u(\mathbf{x})\text{e}^{-\text{i}\omega t}\right)$ and substituting it into Eq. (2), we obtain the inhomogeneous Helmholtz equation

$$\nabla^2 u(\mathbf{x}) + k_0^2 n^2(\mathbf{x}) u(\mathbf{x}) = 0, \tag{3}$$

with the propagating constant in free space $k_0 = \omega/\text{c}$. The total field $u(\mathbf{x})$ is the sum of the scattered field $u^{\text{sc}}(\mathbf{x})$ and of the incident field $u^{\text{in}}(\mathbf{x})$, which is itself a solution of the homogeneous Helmholtz equation $\nabla u^{\text{in}}(\mathbf{x}) + k_0^2 n_{\text{b}}^2 u^{\text{in}}(\mathbf{x}) = 0$. Accordingly, Eq. (3) can be rewritten as (see [1])

$$\nabla^2 u^{\text{sc}}(\mathbf{x}) + k_0^2 n_{\text{b}}^2 u^{\text{sc}}(\mathbf{x}) = -f(\mathbf{x})u(\mathbf{x}), \tag{4}$$

where $f(\mathbf{x}) = k_0^2(n^2(\mathbf{x}) - n_{\text{b}}^2)$ defines the scattering potential function. It follows that

$$u^{\text{sc}}(\mathbf{x}) = \int_\Omega g(\mathbf{x}-\mathbf{x}')f(\mathbf{x}')u(\mathbf{x}')\,\text{d}\mathbf{x}', \tag{5}$$

where $g(\mathbf{x})$ is the Green's function of the shift-invariant differential operator $(\nabla^2 + k_0^2 n_{\text{b}}^2\mathbf{I})$. Specifically, g verifies $\nabla^2 g(\mathbf{x}) + k_0^2 n_{\text{b}}^2 g(\mathbf{x}) = -\delta(\mathbf{x})$, where δ is the Dirac distribution and the minus sign is a convention used in physics. Under Sommerfeld's radiation condition, $g(\mathbf{x})$ is given by [9, and references therein]

$$g(\mathbf{x}) = \begin{cases} \frac{1}{4}H_0^{(1)}(k_0 n_{\text{b}}\|\mathbf{x}\|), & D = 2, \\ \frac{1}{4\pi}\frac{\text{e}^{\text{i}k_0 n_{\text{b}}\|\mathbf{x}\|}}{\|\mathbf{x}\|}, & D = 3. \end{cases} \tag{6}$$

There, $H_0^{(1)}$ is the Hankel function of the first kind. Finally, the total field $u(\mathbf{x})$ is governed by the Lippmann-Schwinger equation

$$u(\mathbf{x}) = u^{\text{in}}(\mathbf{x}) + \int_\Omega g(\mathbf{x}-\mathbf{x}')f(\mathbf{x}')u(\mathbf{x}')\,\text{d}\mathbf{x}'. \tag{7}$$

1.2. Inverse ODT problem: prior work

Let the object be illuminated by a series of incident fields $u_p^{\text{in}}(\mathbf{x})$, $p \in [1\ldots P]$. Records of the resulting total fields $u_p(\mathbf{x})$ at positions $\mathbf{x}_m$ ($m \in [1\ldots M]$) in the detector plane Γ are denoted $\mathbf{y}_p \in \mathbb{C}^M$ (see Fig. 1). The objective is then to retrieve the scattering potential function $f(\mathbf{x})$ (*i.e.,* the refractive index $n(\mathbf{x})$) from the data $\mathbf{y}_p$. Pioneering methods were using linear approximations of the model. For instance, assuming that the scattering field is weak compared to the incident one (*i.e.,* $u^{\text{sc}} \ll u^{\text{in}}$), one can interpret the phase of the transmitted wave as the Radon transform of the refractive index and then reconstruct f using the filtered-back-projection algorithm [10, 11]. This method ignores the effect of diffraction. The first Born approximation [1] has then been proposed as a refined model. Its validity is however limited to thin samples with weak variations

Optics EXPRESS

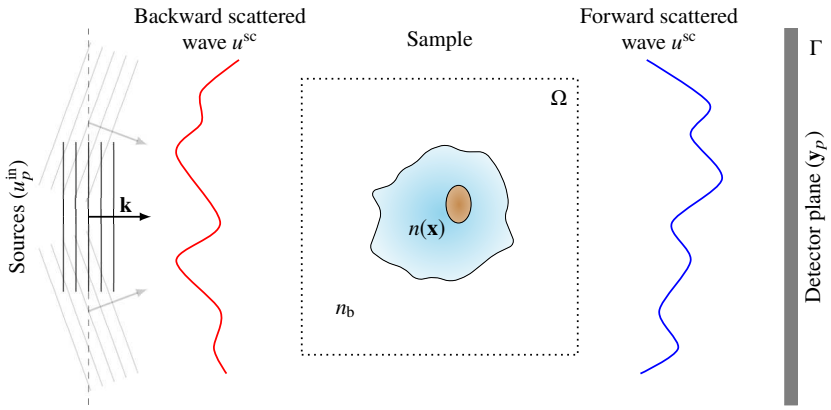

Fig. 1. Optical diffraction tomography. A sample of refractive index $n(\mathbf{x})$ is immersed in a medium of index n_b and illuminated by an incident plane wave (wave vector $\mathbf{k}$). The interaction of the wave with the object produces forward and backward scattered waves. The forward scattered wave is recorded in the detector plane. Optionally, a second detector plane may record the backward scattered wave (see Section 5).

of their refractive index (RI) [12]. A more accurate linearization, less sensitive to the thickness of the sample but still limited to weak RI contrasts, is given by the Rytov approximation [3, 13]. It is derived by assuming that the total field has the form $u(\mathbf{x}) = u^{\text{in}}(\mathbf{x})e^{\phi(\mathbf{x})}$, where $\phi(\mathbf{x})$ is a complex phase function. Both Born and Rytov approximations have been originally used to derive direct inversion methods. They were later used within regularized variational approaches to improve the quality of reconstructed images [14, 15].

Inversion methods that use a nonlinear model have been shown to significantly improve the accuracy of reconstruction. These include the conjugate-gradient method (CGM) [16, 17], the contrast source-inversion method (CSI) [18], the beam-propagation method (BPM) [19], the recursive Born approximation [5], or the hybrid method proposed in [4]. Although still approximate (for instance, they do not properly take reflections into account), they more closely adhere to the model of the physical phenomenon than the linear models, at the price of a higher computational cost. We refer the reader to [2] for additional details concerning existing approximations, regularizations, algorithms, and comparisons.

To address applications with thick samples and large RI contrasts, a better solution is to rely on the exact Lippmann-Schwinger model which accounts for mutiple scattering and reflections. Such an approach has been recently proposed in [6, 7] (SEAGLE algorithm). There, the authors tackle the problem from a variational perspective. They minimize a nonconvex objective using the well known fast iterative shrinkage-thresholding algorithm (FISTA) [20]. Their main contribution is to compute the forward model (which itself requires the inversion of an operator) using Nesterov's accelerated gradient-descent (NAGD) method [21] and, more interestingly, to explicitly compute the gradient of the quadratic data-fidelity term as an error-backpropagation of the forward algorithm. However, the bottleneck of their method is its high memory consumption. Indeed, the error-backpropagation strategy requires one to store all the iterates produced during the computation of the iterative forward model. This can be limiting for large 3D volumes.

1.3. Contributions

To improve the computational efficiency of solvers such as SEAGLE, we provide an explicit expression for the Jacobian of the nonlinear Lippmann-Schwinger operator. This results in an

Optics EXPRESS

efficient method to compute the gradient of the data-fidelity term and avoid recoursing to the memory-consuming error-backpropagation strategy. Another advantage is that the computation of the forward model and of the gradient are now decoupled. They can thus be solved using any numerical scheme. Then, considering simulated data, we show that the proposed method results in a significant reduction of both computational time and memory requirements with respect to SEAGLE, at no loss in quality.

In Section 2.1, we formulate the discrete forward model proposed in [7]. Then, the common approach used to solve the inverse problem subject to sparsity constraints is presented in Section 2.2. There, we highlight our main innovation with respect to SEAGLE, which is a new computation of the gradient of the data-fidelity term. It relies on the derivation of the Jacobian of the forward model, which is given by Proposition 3.1. Finally, Sections 4 and 5 are dedicated to numerical comparisons.

2. Solving the inverse problem

2.1. Formulation of the forward model

In this section, we review the formulation of the forward model that was proposed by Liu *et al.* in [7]. Let the region of interest Ω be divided into $N \in \mathbb{N}$ "pixels". Then, over Ω, we define the discrete version of Eq. (7) as

$$\mathbf{u}_p = \mathbf{u}_p^{\text{in}} + \mathbf{G}\,\mathbf{diag}(\mathbf{f})\mathbf{u}_p, \tag{8}$$

where $\mathbf{u}_p \in \mathbb{C}^N$, $\mathbf{u}_p^{\text{in}} \in \mathbb{C}^N$, $\mathbf{f} \in \mathbb{R}^N$ are the discrete representations of u_p, u_p^{in}, and f, respectively. The diagonal matrix $\mathbf{diag}(\mathbf{f}) \in \mathbb{R}^{N\times N}$ is formed out of the entries of $\mathbf{f}$, while $\mathbf{G} \in \mathbb{C}^{N\times N}$ stands for the matrix of the convolution operator on Ω (convolution with g). One can notice that Eq. (8) is nonlinear with respect to $\mathbf{f}$. On the other hand, given $\mathbf{u}_p^{\text{in}}$ and $\mathbf{f}$, the computation of $\mathbf{u}_p$ amounts to inverting the operator $(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))$. Instead of attempting to compute this inverse directly, the ODT forward model on Ω, for a given $\mathbf{f}$, is defined as

$$\mathbf{u}_p(\mathbf{f}) = \arg\min_{\mathbf{u}\in\mathbb{C}^N} \frac{1}{2}\|(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))\mathbf{u} - \mathbf{u}_p^{\text{in}}\|_2^2. \tag{9}$$

This classical quadratic-minimization problem can be solved iteratively using numerous state-of-the-art algorithms (see Section 4.2). Then, from the total field $\mathbf{u}_p(\mathbf{f})$ (inside Ω), we get measurements $\mathbf{y}_p$ on Γ using a different discretization $\tilde{\mathbf{G}} \in \mathbb{C}^{M\times N}$ of the Green's function (see [7])

$$\mathbf{y}_p = \tilde{\mathbf{G}}\,\mathbf{diag}(\mathbf{f})\mathbf{u}_p(\mathbf{f}) + \mathbf{u}_p^{\text{in}}|_\Gamma, \tag{10}$$

where $\mathbf{u}_p^{\text{in}}|_\Gamma$ denotes the restriction of the field $\mathbf{u}_p^{\text{in}}$ to the area Γ.

2.2. Common optimization strategy

Following the classical variational approach, the estimation of $\mathbf{f} \in \mathbb{R}^N$ from the measurements $\{\mathbf{y}_p \in \mathbb{C}^M\}_{p\in[1\ldots P]}$ is formulated as the optimization problem

$$\widehat{\mathbf{f}} \in \left\{\arg\min_{\mathbf{f}\in\mathbb{R}^N} \left(\mathcal{D}(\mathbf{f}) + \mu\mathcal{R}(\mathbf{f})\right)\right\}, \tag{11}$$

where $\mathcal{D} : \mathbb{R}^N \to \mathbb{R}$ measures the fidelity to data, $\mathcal{R} : \mathbb{R}^N \to \mathbb{R}$ imposes some prior to the solution (regularization), and $\mu > 0$ balances between these two terms. It is customary to consider the data term

$$\mathcal{D}(\mathbf{f}) = \sum_{p=1}^{P} \mathcal{D}_p(\mathbf{f}), \tag{12}$$

Optics EXPRESS

Algorithm 1 Accelerated forward-backward splitting.

Require: $\mathbf{f}^0 \in \mathbb{R}^N$, $(\alpha^k)_{k\in\mathbb{N}}$, $\gamma \in (0, 1/\mathrm{Lip}(\boldsymbol{\nabla}\mathcal{D}))$
1: $\mathbf{v}^1 = \mathbf{f}^0$
2: $k = 1$
3: **while** (not converged) **do**
4: $\quad \mathbf{u}_p^k \leftarrow \mathbf{u}_p(\mathbf{f}^k), \forall p \in [1 \dots P]$ (forward model Eq. (9))
5: $\quad \mathbf{d}^k = \sum_{p=1}^{P} \mathrm{Re}\left(\mathbf{J}_{h_p}^H(\mathbf{f}^k)\tilde{\mathbf{G}}^H(\tilde{\mathbf{G}}\,\mathbf{diag}(\mathbf{f}^k)\mathbf{u}_p^k - \mathbf{y}_p^{\mathrm{sc}})\right)$
6: $\quad \mathbf{f}^k = \mathrm{prox}_{\gamma\lambda\mathcal{R}}\left(\mathbf{v}^k - \gamma\mathbf{d}^k\right)$
7: $\quad \mathbf{v}^{k+1} = \mathbf{f}^k + \alpha^k(\mathbf{f}^k - \mathbf{f}^{k-1})$
8: $\quad k = k + 1$
9: **end while**

where $\forall p \in [1 \dots P]$

$$\mathcal{D}_p(\mathbf{f}) = \frac{1}{2}\|\tilde{\mathbf{G}}\,\mathbf{diag}(\mathbf{f})\mathbf{u}_p(\mathbf{f}) - \mathbf{y}_p^{\mathrm{sc}}\|_2^2, \tag{13}$$

which is well suited for Gaussian noise. Here, $\mathbf{y}_p^{\mathrm{sc}} = (\mathbf{y}_p - \mathbf{u}_p^{\mathrm{in}}|_\Gamma)$ is the scattered measured field at the detector plane Γ and $\mathbf{u}_p(\mathbf{f})$ is given by Eq. (9). As regularizer $\mathcal{R}$, the combination

$$\mathcal{R}(\mathbf{f}) = i_{\geqslant 0}(\mathbf{f}) + \|\boldsymbol{\nabla}\mathbf{f}\|_{2,1} = i_{\geqslant 0}(\mathbf{f}) + \sum_{n=1}^{N}\sqrt{\sum_{d=1}^{D}(\partial_d\mathbf{f})_n^2} \tag{14}$$

of total variation (TV) penalty and nonnegativity constraint is used, where $i_{\geqslant 0}(\mathbf{f}) = \{0, \text{ if } \mathbf{f}_n \geq 0\,\forall n; +\infty, \text{ otherwise}\}$ and ∂_d denotes the gradient operator along the dth direction. This choice is supported by the facts that we consider situations where $n_{\mathrm{b}} \leq n(\mathbf{x}) \Rightarrow f(\mathbf{x}) \geq 0$ and that n and, thus, f can be assumed to be piecewise-constant. It is worth noting that Eq. (11) is nonconvex due to the nonlinearity of the forward operator in Eq. (8). However, since $\mathcal{D}$ is smooth with respect to $\mathbf{f}$, Eq. (11) can be solved by deploying a forward-backward splitting (FBS) method [22] or some accelerated variants [20, 23], as presented in Algorithm 1. The gradient of the data-fidelity term $\mathcal{D}$ is given by

$$\boldsymbol{\nabla}\mathcal{D}(\mathbf{f}) = \sum_{p=1}^{P}\boldsymbol{\nabla}\mathcal{D}_p(\mathbf{f}), \tag{15}$$

with

$$\boldsymbol{\nabla}\mathcal{D}_p(\mathbf{f}) = \mathrm{Re}\left(\mathbf{J}_{h_p}^H(\mathbf{f})\tilde{\mathbf{G}}^H(\tilde{\mathbf{G}}\,\mathbf{diag}(\mathbf{f})\mathbf{u}_p(\mathbf{f}) - \mathbf{y}_p^{\mathrm{sc}})\right), \tag{16}$$

where $\mathbf{J}_{h_p}(\mathbf{f})$ denotes the Jacobian matrix of

$$h_p : \mathbf{f} \mapsto \mathbf{diag}(\mathbf{f})\mathbf{u}_p(\mathbf{f}). \tag{17}$$

Algorithm 1 encompasses FISTA [20] for a specific choice of the sequence $(\alpha^k)_{k\in\mathbb{N}}$. Its convergence is guaranteed in the convex case when $\gamma < 1/\mathrm{Lip}(\boldsymbol{\nabla}\mathcal{D})$, where $\mathrm{Lip}(\boldsymbol{\nabla}\mathcal{D})$ is the Lipschitz constant of $\boldsymbol{\nabla}\mathcal{D}$. In the nonconvex case, a local convergence of the classical FBS algorithm can be shown [24]. Although, to the best of our knowledge, there exists no theoretical proof of convergence of accelerated versions for nonconvex function, Algorithm 1 always converged in our experiments.

2.2.1. Computation of $\mathbf{J}_{h_p}^H(\mathbf{f})$

The computation of $\mathbf{J}_{h_p}^H(\mathbf{f})$, required at line 5 of Algorithm 1, is challenging. The existence of a closed-form solution is made unlikely by the fact that the forward model in Eq. (8) itself requires one to invert an operator. We distinguish two distinct strategies.

Optics EXPRESS

1. SEAGLE: Build an error-backpropagation rule from the NAGD algorithm used to compute the forward model Eq. (9).

2. Ours: Derive an explicit expression of $\mathbf{J}_{h_p}(\mathbf{f})$, as given in Section 3 (Proposition 3.1).

2.2.2. Computation of $\text{prox}_{\gamma\lambda\mathcal{R}}$

Numerous methods have been proposed to compute the proximity operator of $\mathcal{R}$, [25–27]. In SEAGLE, Liu *et al.* use the algorithm proposed by Beck and Teboulle [25]. Here, we compute it using the popular alternating-direction method of multipliers (ADMM) [28], which is well suited to the minimization of the sum of three convex functions. Moreover, it provides a high modularity for spatial regularization since one can easily change from one regularizer (*e.g.,* TV) to another (*e.g.,* Hessian Shatten-norm [29]). Details about the computation of $\mathcal{R}$ are provided in Appendix A.

2.2.3. Speedup strategies

The cost of evaluating the forward model with Eq. (9) and the gradient $\boldsymbol{\nabla}\mathcal{D}$ is proportional to the number P of illuminations $\mathbf{u}_p^{in}$. However, these computations can easily be parallelized by performing the computation for each illumination (or each element of the sum in Eq. (15)) on a separate thread. Moreover, in the spirit of the stochastic gradient-descent algorithm [30], we approximate $\boldsymbol{\nabla}\mathcal{D}$ as

$$\boldsymbol{\nabla}\mathcal{D}(\mathbf{f}) \simeq \sum_{p\in\omega} \boldsymbol{\nabla}\mathcal{D}_p(\mathbf{f}), \tag{18}$$

where ω is a subset of $[1 \dots P]$. We change ω at each iteration. Such a method is known to spare many computations when $\boldsymbol{\nabla}\mathcal{D}_p$ does not admit a simple-form expression.

3. Efficient computation of the gradient $\boldsymbol{\nabla}\mathcal{D}$

The error-backpropagation strategy used in SEAGLE to compute $\mathbf{J}_{h_p}^H(\mathbf{f})$ implies that one must store all the forward iterates. This consumes memory resources and compromises the deployment of the method for large 3D data. Instead, Proposition 3.1 reveals that its computation requires one to invert the operator $(\mathbf{I} - \mathbf{diag}(\mathbf{f})\mathbf{G}^H)$. This operator has the same form (and size) that the operator we invert within the forward computation in Eq. (9) and both can be computed in a similar way, using an iterative algorithm. Moreover, it allows us to decouple the forward and gradient computation in Algorithm 1, which has the two following advantages:

- choice of any iterative algorithm for computing Eq. (9) at line 4 of Algorithm 1, and computing $\mathbf{J}_{h_p}^H(\mathbf{f})$ at line 5 of Algorithm 1 (see Section 4.2);
- reduction of the memory consumption (no needs for storing forward iterates).

Proposition 3.1. *The Jacobian matrix of the function h_p in Eq. (17) is given by*

$$\mathbf{J}_{h_p}(\mathbf{f}) = \Big(\mathbf{I} + \mathbf{diag}(\mathbf{f})(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))^{-1}\mathbf{G}\Big)\,\mathbf{diag}(\mathbf{u}_p(\mathbf{f})). \tag{19}$$

Proof. We use the Gâteaux derivative in the direction $\mathbf{v} \in \mathbb{R}^N$ given by

$$\begin{aligned} \mathrm{d}h_p(\mathbf{f};\mathbf{v}) &= \lim_{\varepsilon\to 0} \frac{\mathbf{diag}(\mathbf{f}+\varepsilon\mathbf{v})\mathbf{u}_p(\mathbf{f}+\varepsilon\mathbf{v}) - \mathbf{diag}(\mathbf{f})\mathbf{u}_p(\mathbf{f})}{\varepsilon} \\ &= \mathbf{diag}(\mathbf{u}_p(\mathbf{f}))\mathbf{v} + \lim_{\varepsilon\to 0} \mathbf{diag}(\mathbf{f})\frac{\mathbf{u}_p(\mathbf{f}+\varepsilon\mathbf{v}) - \mathbf{u}_p(\mathbf{f})}{\varepsilon}. \end{aligned} \tag{20}$$

Optics EXPRESS

Then, from Eq. (8), we get that

$$\begin{aligned}\mathbf{u}_p^{\text{in}} =&(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f} + \varepsilon\mathbf{v}))\mathbf{u}_p(\mathbf{f} + \varepsilon\mathbf{v}) \\ =&(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))\mathbf{u}_p(\mathbf{f} + \varepsilon\mathbf{v}) - \varepsilon\mathbf{G}\,\mathbf{diag}(\mathbf{v})\mathbf{u}_p(\mathbf{f} + \varepsilon\mathbf{v})\end{aligned} \tag{21}$$

and

$$(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))\mathbf{u}_p(\mathbf{f}) = \mathbf{u}_p^{\text{in}}. \tag{22}$$

Combining Eq. (21) and Eq. (22), we obtain that

$$(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))(\mathbf{u}_p(\mathbf{f} + \varepsilon\mathbf{v}) - \mathbf{u}_p(\mathbf{f})) = \varepsilon\mathbf{G}\,\mathbf{diag}(\mathbf{v})\mathbf{u}_p(\mathbf{f} + \varepsilon\mathbf{v}). \tag{23}$$

Finally, we get that

$$\mathrm{d}h_p(\mathbf{f};\mathbf{v}) = \left(\mathbf{I} + \mathbf{diag}(\mathbf{f})(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))^{-1}\mathbf{G}\right)\mathbf{diag}(\mathbf{u}_p(\mathbf{f}))\mathbf{v} \tag{24}$$

and, thus, that

$$\mathbf{J}_{h_p}(\mathbf{f}) = \left(\mathbf{I} + \mathbf{diag}(\mathbf{f})(\mathbf{I} - \mathbf{G}\,\mathbf{diag}(\mathbf{f}))^{-1}\mathbf{G}\right)\mathbf{diag}(\mathbf{u}_p(\mathbf{f})), \tag{25}$$

which completes the proof. □

4. Algorithm analysis

4.1. Memory requirement

In this section, we elaborate on the memory consumption of the proposed method in comparison with SEAGLE. First, let us state that gradient based methods, such as NAGD or CG, have similar memory requirements. It corresponds roughly to three times the size of the optimization variable which is the part that is common to both algorithms. The additional memory requirement that is specific to SEAGLE relies only on the storage of the NAGD iterates during the forward computation. Suppose that $K_{\text{NAGD}} \in \mathbb{N}$ iterations are necessary to compute the forward model with Eq. (9) and that the region Ω is sampled over $N \in \mathbb{N}$ pixels (voxels, in 3D). Since the total field $\mathbf{u}_p(\mathbf{f})$ computed by NAGD is complex-valued, each pixel is represented with 16 bytes (double precision for accurate computations). Hence, the difference of memory consumption between SEAGLE and our method is

$$\Delta_{\text{Mem}} = N \times K_{\text{NAGD}} \times 16\ [\text{bytes}], \tag{26}$$

which corresponds to the storage of the K_{NAGD} intermediate iterates of NAGD. Here, we assumed that $\boldsymbol{\nabla}\mathcal{D}$ was computed by sequentially adding the partial gradients $\boldsymbol{\nabla}\mathcal{D}_p$ associated to the P incident fields. Hence, once the partial gradient associated to one incident angle is computed by successively applying the forward model (NAGD) and the error-backpropagation procedure, the memory used to store the intermediate iterates can be recycled to compute the partial gradient associated to the next incident angle. However, when the parallelization strategy detailled in Section 2.2.3 is used, the memory requirement is mutiplied by the number $N_{\text{Threads}} \in \mathbb{N}$ of threads, so that

$$\Delta_{\text{Mem}} = N \times K_{\text{NAGD}} \times N_{\text{Threads}} \times 16\ [\text{bytes}]. \tag{27}$$

Indeed, since the threads of a single computer share memory, computing N_{Threads} partial gradients in parallel requires N_{Threads} times more memory.

For illustration, we give in Fig. 2 the evolution of Δ_{Mem} as a function of N for different values of K_{NAGD} and N_{Threads}. One can see with the vertical dashed lines that, for 3D volumes, the memory used by SEAGLE quickly reaches several tens of Megabytes, even for small volumes (*e.g.,* $128 \times 128 \times 128$), to hundreds of Gigabytes for the larger volumes that are typical of microscopy (*e.g.,* $512 \times 512 \times 256$). This shows the limitation of SEAGLE for 3D reconstruction in the presence of a shortage of memory resources and reinforces the interest of the proposed alternative.

Optics EXPRESS

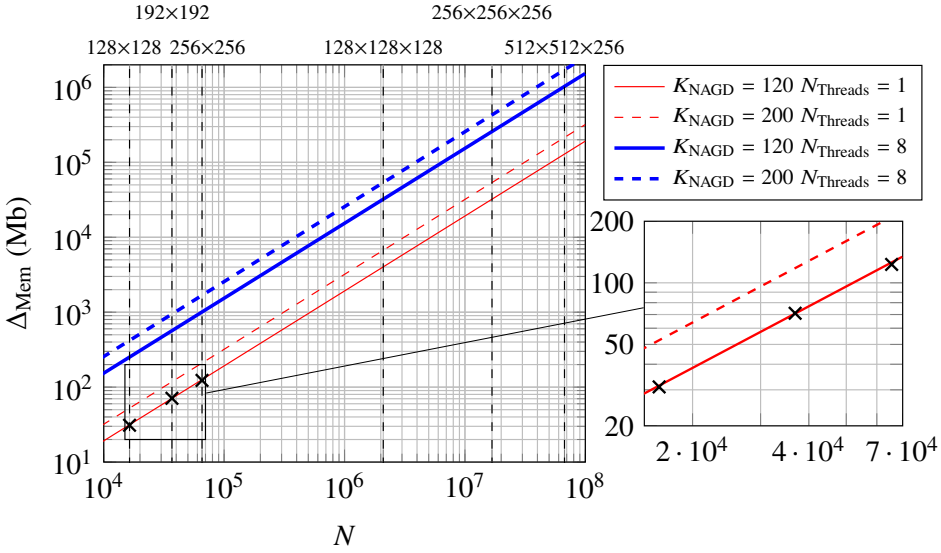

Fig. 2. Predicted evolution of Δ_{Mem} as function of the number N of points for two values of K_{NAGD} and N_{Threads}. The vertical dashed lines give examples of 2D and 3D volumes for a range of values of N. Finally, the three crosses correspond to values of Δ_{Mem} measured experimentally.

4.2. *Conjugate gradient vs. Nesterov accelerated gradient descent for Eq. (9)*

Due to Proposition 3.1, we can compute both Eq. (9) and $\mathbf{J}_{h_p}^H(\mathbf{f})$ using any state-of-the-art quadratic optimization algorithm. This contrasts with SEAGLE, where one must derive the error-backpropagation rule from the forward algorithm, which may limit its choice. We now provide numerical evidence that GC is more efficient than NAGD for solving Eq. (9). To this end, we consider a circular object (bead) of radius r_{bead} and refractive index n_{bead} immersed into water ($n_{\text{b}} = 1.333$), as presented in Fig. 3 (top-left). In such a situation, an analytic expression of the total field is provided by the Mie theory [31, 32]. Hence, at each iteration k, we compute the relative error ε_k of the current estimate $\mathbf{u}^k$ to the Mie solution $\mathbf{u}_{\text{Mie}}$ as

$$\varepsilon_k = \frac{\|\mathbf{u}^k - \mathbf{u}_{\text{Mie}}\|^2}{\|\mathbf{u}_{\text{Mie}}\|^2}. \tag{28}$$

In our experiment, the bead is impinged by a plane wave of wavelength $\lambda = 406$ nm. The region of interest is square with a side length of 16λ (see top-left panel of Fig. 3). It is sampled using 1,024 points along each side. We used a fine grid in order to limit the impact of numerical errors related to discretization. The wave source corresponds to the bottom border of this region. Then, as in [6, 7], we refer to the refractive index n_{bead} by its contrast with respect to the background medium, defined as $\max(|\mathbf{f}|)/(k_0^2 n_{\text{b}}^2)$. We show in Fig. 4 the evolution of k_{ε_0}, which is the number of iterations needed to let the relative error Eq. (28) fall below $\varepsilon_0 = 10^{-2}$. One can observe that CG is much more efficient than NAGD, in particular for large contrasts. This is not negligible since an evaluation of the forward model is required at each iteration of Algorithm 1 (line 4). Our comparison in terms of a number of iterations is fair because the computational cost of one iteration is the same for both algorithms. Note that the descent step of NAGD was adapted during the iterations following the same rule as in [6, 7].

Optics EXPRESS

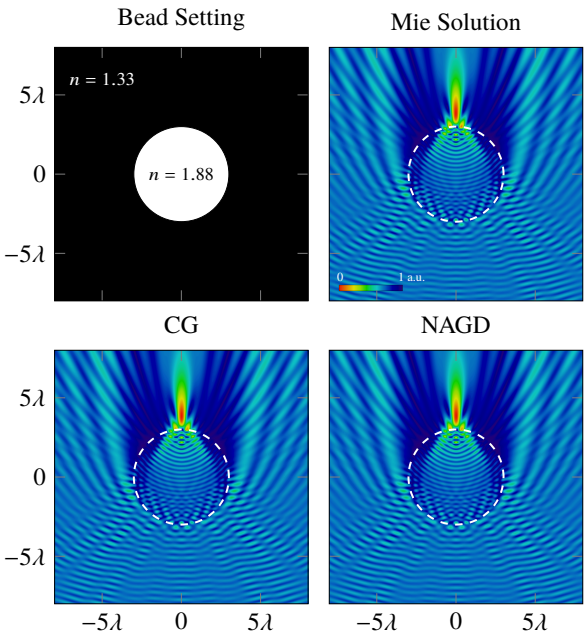

Fig. 3. Forward-model solution for a bead with radius 3λ and a contrast of 1 using CG (bottom-left) and NAGD (bottom-right), as well as the Mie solution (top-right). The setting used for this experiment is presented in the top-left panel. The colormap is the same for each figure.

Finally, the solution obtained with the two algorithms for $r_{\text{bead}} = 3\lambda$ and a contrast of 1 are shown in Fig. 3. The analytic Mie solution is also provided for comparison. From these figures, one can appreciate the high accuracy obtained by solving Eq. (9), as first demonstrated in [6, 7].

5. Numerical experiments

This section is devoted to numerical experiments that illustrate the two main advantages of the proposed method over SEAGLE, which consist in a reduced computational cost and a reduced memory consumption. The algorithms have been implemented using an inverse-problem library developed in our group [33] (*GlobalBioIm*: http://bigwww.epfl.ch/algorithms/globalbioim/). Hence, they share the implementation of the overall FISTA algorithm as well as inner procedures such as the computation of the proximity operator of $\mathcal{R}$ (see Appendix A). The only difference between the two methods resides in the computation of the forward model in Eq. (9) and $\mathbf{J}_{h_p}^H(\mathbf{f})$. For SEAGLE, this is performed using the NAGD algorithm and an error-backpropagation strategy. For our method, Eq. (9) and $\mathbf{J}_{h_p}^H(\mathbf{f})$ are computed using the CG algorithm, in accordance with Proposition 3.1. Note that no parallelization is used. Reconstructions are performed with MATLAB 9.1 (The MathWorks Inc., Natick, MA, 2000) on a PowerEdge T430 Dell computer (Intel Xeon E5-2620 v3).

5.1. Simulated data

5.1.1. Simulation settings

The Shepp-Logan phantom of Fig. 5 has the contrast $\max(|\mathbf{f}|)/(k_0^2 n_{\text{b}}^2) = 0.2$. It is immersed into water ($n_{\text{b}} = 1.333$). The wavelength of the incident plane waves is $\lambda = 406$ nm. We consider

Optics EXPRESS

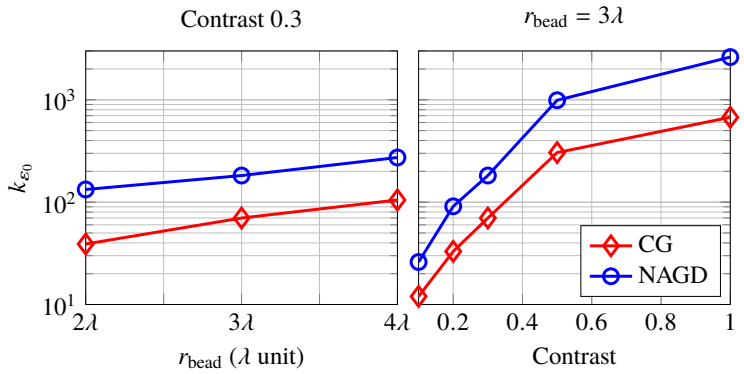

Fig. 4. Evolution of the number of iterations k_{ε_0} needed to let the relative error Eq. (28) fall below $\varepsilon_0 = 10^{-2}$ as function of bead radius (left) and bead contrast (right).

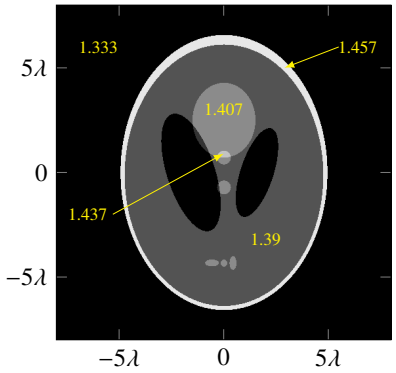

Fig. 5. Sheep-Logan phantom and refractive indices of the gray levels. The contrast is 20%.

thirty-one incident angles, from $-60°$ to $+60°$. The sources are placed at the bottom side of the sample, at a distance of 16.5λ from its center. Moreover, we consider two detectors placed on both top and bottom sides of the object, also at a distance of 16.5λ from its center. Hence, the overall region is a square of length 33λ per side. Data are simulated using a fine discretization of this region, with a (1024×1024) grid that leads to square pixels of surface $(3.223 \cdot 10^{-2}\lambda)^2$. We used a large number of CG iterations to get the accurate simulation mentioned in Section 4.2. Then, the measurements were extracted from the first and last rows of each total field associated to the incident fields. This lead to a total of $(31 \times 2 \times 1024)$ measurements. Finally, we defined three ODT problems by downsizing (using averaging) the $(31 \times 2 \times 1024)$ measurements to grids with size of $(31 \times 2 \times 512)$, $(31 \times 2 \times 384)$, and $(31 \times 2 \times 256)$.

This setting corresponds to an ill-posed and highly scattering situation. Moreover, the detector length is only two times larger than the object, which results in a loss of information for large incident angles. This makes the resulting inverse problem challenging.

5.1.2. Algorithm parameters

For each simulated OTD problem, we considered a square region of interest Ω with sides half the sources–detector distance. That corresponds to images of size (256×256) with pixels of area $(6.445 \cdot 10^{-2}\lambda)^2$, (192×192) with pixels of area $(8.839 \cdot 10^{-2}\lambda)^2$, and (128×128) with pixels of

Optics EXPRESS

Table 1. Proposed method *vs.* SEAGLE [6, 7] in terms of running time and memory consumption. The reconstructed refractive-index maps are presented in Fig. 6.

ROI Ω size	(128×128)		(192×192)		(256×256)	
Method	Ours	[7]	Ours	[7]	Ours	[7]
Time (min)	**9**	35	**12**	72	**19**	110
Memory (Mb)	**138**	169	**224**	295	**337**	460
SNR (dB)	43.96	43.76	45.44	45.48	46.96	46.99

area $(1.289 \cdot 10^{-1}\lambda)^2$. The support of the phantom is fully contained in Ω.

Then, to compute the gradient (stochastic-gradient strategy), we selected eight angles over the thirty-one that were available and changed this selection at each iteration (see Section 2.2.3).

The NAGD or CG forward algorithms are stopped either after hundred-twenty iterations or when the relative error between two iterates is below 10^{-4}. Finally, two-hundred iterations of FISTA are performed with a descent step fixed empirically to $\gamma = 5 \cdot 10^{-3}$. We used the regularization parameter $\mu = 3.3 \cdot 10^{-2}$.

5.1.3. Metrics

We compared the two methods in terms of running time and memory consumption, as measured by the peak memory (maximum allocated memory) reached by each algorithm during execution. The outcome is reported in Table 1. Once again, due to the use of our inverse-problem library [33], the comparison of the two methods is fair because their implementations differ only by the forward algorithm and by the computation of $\mathbf{J}_{h_p}^H(\mathbf{f})$. Moreover, CG and NAGD are implemented in the same fashion since they inherit the same optimization class of our inverse-problem library. Finally, we also provide the SNR of the reconstructed refractive index and observe that the computational gain comes at no cost in quality.

5.1.4. Discussion

Our proposed alternative to SEAGLE allows us to reduce both time and memory. Moreover, we have measured the peak memory difference Δ_{Mem} between the two methods and superimposed it on the predictions of Fig. 2 where the adequacy between the theoretical curves and the measured points is remarkable. Hence, although our experiments are restricted to 2D data, where the gap between the two algorithms is moderate, the evolution of Δ_{Mem} for 3D data can be extrapolated from Fig. 2. This shows the relevance of our method when size increases.

The SNR values given in Table 1 as well as the reconstructions presented in Fig. 6 suggest that the two methods perform similarly in terms of quality. This is not surprising since the overall algorithm is the same, the differences residing merely in the computation of the forward model in Eq. (9) and the Jacobian $\mathbf{J}_{h_p}(\mathbf{f})$. Moreover, one can observe that the quality of reconstruction decreases when the discretization grid becomes coarser. Indeed, the model is insufficiently accurate when the discretization is too poor. For instance, in the case of the (128×128) grid, one wavelength unit is discretized using eight pixels, which is clearly detrimental to the accuracy of the forward model.

Reconstructions for the (256×256) problem are presented in Fig. 6 for completeness. Besides the difficulty of the considered scenario, the two methods are able to retrieve most details of the object in comparison with the Rytov approximation. Artifacts are mainly due to the missing-cone problem and to the limited length of the detector. This corroborates the findings of [7].

Optics EXPRESS

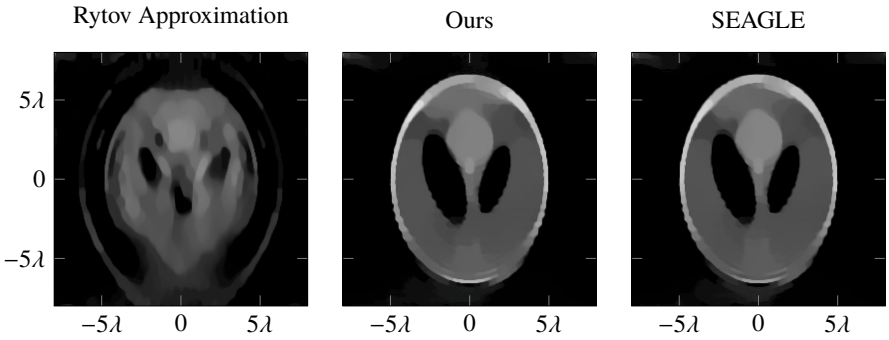

Fig. 6. Reconstructions obtained by the proposed method and by SEAGLE for the (256×256) ODT problem with $\mu = 3.3 \cdot 10^{-2}$. The colormap is identical to that of Fig. 5. For comparison, we provide the TV-regularized Rytov reconstruction with $\mu = 3 \cdot 10^{-3}$.

5.2. Real data

We evaluated our method using the *FoamDielExt* target (TM polarisation) of the Institut Fresnel's public database [34]. The data were collected for the two-dimensional inhomogeneous sample depicted in the left panel of Fig. 7. The permitivity of the ground truth was measured experimentally and is subject to uncertainties [34]. The object is fully contained in a square region of length 15 cm per side, which we discretize using a 256×256 grid. Sensors were distributed circularly around the object, at a distance of 1.67 m from its center, and with a step of 1°. Eight sources, uniformly distributed around the object, were sequentially activated. For each activated source, the sensors closer than 60° from the source were excluded. Thus, 241 detectors among the 360 available were used for each source. Frequencies from 2 to 10 GHz with a step of 1 GHz are available in the database but we used only the 3 GHz measurements (*i.e.,* $\lambda = 10$ cm).

The NAGD or CG forward algorithms are stopped either after two-hundred iterations or when the relative error between two iterates is below 10^{-6}. Hundred iterations of FISTA are performed with a descent step $\gamma = 5 \cdot 10^{-3}$. We used the regularization parameter $\mu = 1.6 \cdot 10^{-2}$.

In Fig. 7, we see that both methods provide good reconstructions that are essentially indistinguishable (see also SNR values provided in the caption of the figure). This corroborates the simulated numerical experiments of Section 5.1. The main point here is that, for this setting, the proposed method was 15 times faster than SEAGLE.

6. Conclusion

We have presented a refinement of the SEAGLE algorithm that was recently proposed in [6, 7] and that has shown unprecedented reconstructions for difficult configurations. However, the current limitation of SEAGLE is that its memory requirements increase excessively with the size of the problem, particularly in 3D. As an alternative, we have derived the explicit expression of the Jacobian matrix $\mathbf{J}_{h_p}(\mathbf{f})$ of the nonlinear Lippman-Schiwnger model and shown that it can be computed in a direct analogy with the computation of the forward model. This approach allows us to drastically reduce the memory consumption and opens the door to 3D reconstruction using desktop computers. Moreover, the proposed method is quite flexible in the sense that it can cope with any iterative algorithm employed to compute either the forward model or $\mathbf{J}_{h_p}^H(\mathbf{f})$. For instance, the conjugate-gradient algorithm proved its efficiency for this task. It allows a significant decrease of the computational time with respect to SEAGLE. Finally, these improvements in terms of speed and memory come at no loss in quality.

Optics EXPRESS

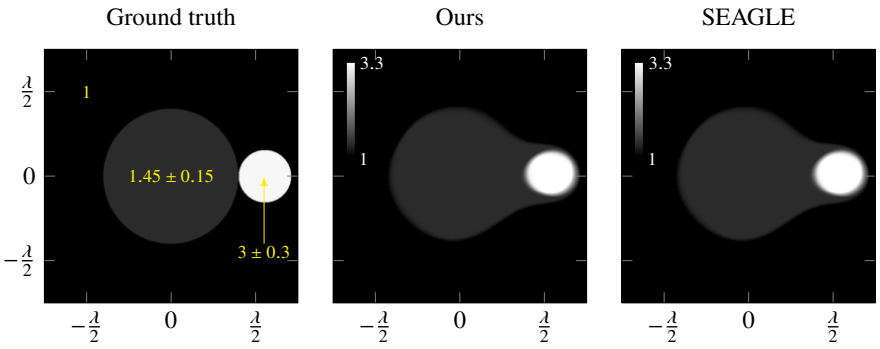

Fig. 7. Reconstructions (permittivity) obtained by the proposed method and by SEAGLE for the *FoamDielExt* target of the Institut Fresnel's database [34] with $\mu = 1.6 \cdot 10^{-2}$. The SNR values (computed from the experimentally measured permittivity of the ground truth) are 25.13 dB (Ours) and 25.15 dB (SEAGLE) while the computing times are respectively of 6 min and 93 min.

A. Proximity operator of $\mathcal{R}$

In this appendix, we describe how we compute the proximity operator of the regularization term $\mathcal{R}$ in Eq. (14) using the ADMM algorithm [28]. The proximity operator is defined [35] as the solution of the optimization problem

$$\text{prox}_{\mu\mathcal{R}}(\mathbf{v}) = \arg\min_{\mathbf{f}\in\mathbb{R}^N} \left(\frac{1}{2}\|\mathbf{f}-\mathbf{v}\|_2^2 + \mu\|\mathbf{f}\|_{\text{TV}} + i_{\geqslant 0}(\mathbf{f}) \right) \tag{29}$$

for $\mu > 0$. Let us start by reformulating Eq. (29) as

$$\begin{aligned} \text{prox}_{\mu\mathcal{R}}(\mathbf{v}) = \arg\min_{\mathbf{f}\in\mathbb{R}^N} & \left(\frac{1}{2}\|\mathbf{f}-\mathbf{v}\|_2^2 + \mu\|\mathbf{q}_1\|_{2,1} + i_{\geqslant 0}(\mathbf{q}_2) \right), \\ s.t. \quad & \mathbf{q}_1 = \boldsymbol{\nabla}\mathbf{f}, \\ & \mathbf{q}_2 = \mathbf{f}, \end{aligned} \tag{30}$$

which admits the augmented-Lagrangian form

$$\begin{aligned} \mathcal{L}(\mathbf{f},\mathbf{q}_1,\mathbf{q}_2,\mathbf{w}_1,\mathbf{w}_2) = \frac{1}{2}\|\mathbf{f}-\mathbf{v}\|_2^2 + \frac{\rho_1}{2}\left\|\boldsymbol{\nabla}\mathbf{f} - \mathbf{q}_1 + \frac{\mathbf{w}_1}{\rho_1}\right\|_2^2 \\ + \frac{\rho_2}{2}\left\|\mathbf{f} - \mathbf{q}_2 + \frac{\mathbf{w}_2}{\rho_2}\right\|_2^2 + \mu\|\mathbf{q}_1\|_{2,1} + i_{\geqslant 0}(\mathbf{q}_2), \end{aligned} \tag{31}$$

where ρ_1 and ρ_2 are positive scalars, and where $\mathbf{w}_1 \in \mathbb{R}^{N\times D}$ and $\mathbf{w}_2 \in \mathbb{R}^N$ are the Lagrangian multipliers. Then, one can minimize Eq. (31) using ADMM. The iterates are summarized in Algorithm 2.

For the sake of completeness, we provide in Eq. (32) and Eq. (33) the expressions

$$\forall \mathbf{q} \in \mathbb{R}^N, \left[\text{prox}_{i_{\geqslant 0}}(\mathbf{q})\right]_n = (\mathbf{q}_n)_+, \tag{32}$$

$$\forall \mathbf{q} \in \mathbb{R}^{N\times D}, \left[\text{prox}_{\gamma\|\cdot\|_{2,1}}(\mathbf{q})\right]_{n,d} = \mathbf{q}_{n,d}\left(1 - \frac{\gamma}{\|\mathbf{q}_{n,\cdot}\|_2}\right)_+, \tag{33}$$

of $\text{prox}_{i_{\geqslant 0}}$ and $\text{prox}_{\gamma\|\cdot\|_{2,1}}$ where

$$(x)_+ := \max(x, 0), x \in \mathbb{R}. \tag{34}$$

Optics EXPRESS

Algorithm 2 ADMM for solving Eq. (29).

Require: $\mathbf{f}^0 \in \mathbb{R}^N$, $\mu > 0$, $\rho_1 > 0$, $\rho_2 > 0$

$\mathbf{A} = \left((1+\rho_2)\mathbf{I} + \rho_1 \boldsymbol{\nabla}^T \boldsymbol{\nabla}\right)$

$\mathbf{q}_1^0 = \boldsymbol{\nabla}\mathbf{f}^0$, $\mathbf{q}_2^0 = \mathbf{f}^0$

$\mathbf{w}_1 = \mathbf{q}_1$, $\mathbf{w}_2 = \mathbf{q}_2$

$k = 1$

while (not converged) **do**

$\mathbf{q}_1^{k+1} = \text{prox}_{\frac{\mu}{\rho_1}\|\cdot\|_{2,1}}\left(\boldsymbol{\nabla}\mathbf{f}^k + \frac{\mathbf{w}_1^k}{\rho_1}\right)$

$\mathbf{q}_2^{k+1} = \text{prox}_{i_{\geqslant 0}}\left(\mathbf{f}^k + \frac{\mathbf{w}_2^k}{\rho_2}\right)$

$\mathbf{f}^{k+1} = \mathbf{A}^{-1}\left(\mathbf{v} + \rho_1 \boldsymbol{\nabla}^T\left(\mathbf{q}_1^{k+1} - \frac{\mathbf{w}_1^k}{\rho_1}\right) + \rho_2 \mathbf{q}_2^{k+1} - \mathbf{w}_2^k\right)$ {Fourier division}

$\mathbf{w}_1^{k+1} = \mathbf{w}_1^k + \rho_1(\boldsymbol{\nabla}\mathbf{f}^{k+1} - \mathbf{q}_1^{k+1})$

$\mathbf{w}_2^{k+1} = \mathbf{w}_2^k + \rho_2(\mathbf{f}^{k+1} - \mathbf{q}_2^{k+1})$

$k = k + 1$

end while

Funding

European Research Council (ERC) (692726)

Acknowledgment

This research was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, Grant Agreement no 692726 "GlobalBioIm: Global integrative framework for computational bio-imaging."

126.178.48.127 {ts '2017-09-05 11:28:25'}