Taking Advantage of Multiple Scattering for Optical Reflection Tomography

Thomas Wasik, Victor Barolle, Alexandre Aubry, and Josselin Garnier

***Abstract*—Optical Diffraction Tomography (ODT) is a powerful non-invasive imaging technique widely used in biological and medical applications. While significant progress has been made in transmission configuration, reflection ODT remains challenging due to the ill-posed nature of the inverse problem. We present a novel optimization algorithm for 3D refractive index (RI) reconstruction in reflection-mode microscopy. Our method takes advantage of the multiply-scattered waves that are reflected by uncontrolled background structures and that illuminate the foreground RI from behind. It tackles the ill-posed nature of the problem using weighted time loss, positivity constraints and Total Variation regularization. We have validated our method with data generated by detailed 2D and 3D simulations, demonstrating its performance under weak multiple scattering conditions and with simplified forward models used in the optimization routine for computational efficiency. In addition, we highlight the need for multi-wavelength analysis and the use of regularization to ensure the reconstruction of the low spatial frequencies of the foreground RI.**

***Index Terms*—Reflection optical phase tomography, ill-posed inverse problem, multiple scattering, total variation regularization, temporal loss function, stochastic proximal-gradient.**

I. Introduction

OPTICAL diffraction tomography (ODT) is a non-invasive quantitative imaging technique [1], [2], [3] that has shown promising applications in biology and medical imaging [4], [5]. This technique aims to reconstruct the 3D refractive index (RI) map of a sample by illuminating it from various angles, recovering the complex scattered field using holography techniques [6], and numerically solving the inverse scattering problem of light. This inverse problem is known as an ill-posed problem in transmission and in reflection mode configuration.

ODT was first applied to weakly scattering systems, which can be described using the first Born and Rytov approximations [1], [7]. Under first Born hypothesis, the measurements present a filtering effect [1]. In a transmission configuration, the optical system acts as a low-pass filter of the spatial frequencies of the object along the optical axis, corresponding to the missing-cone problem, while it acts as a high-pass filter in a reflection configuration [8]. In recent years, ODT techniques have primarily been developed for transmission configurations, with significant efforts focused on simulating stronger scattering regimes and mitigating the missing-cone problem by developing new forward models and iterative techniques for solving the inverse problem [9], [10], [11], [12], [13].

The application of ODT methods to reflection microscopy would represent a significant breakthrough in the field of non-invasive imaging. Reflection microscopy, in particular, allows for the imaging of thicker samples and enables in vivo imaging [14]. ODT in reflection presents substantial challenges due to the ill-posed nature of the associated inverse problem. In the single scattering regime, only the high spatial frequencies of the object along the optical axis can be reconstructed [8]. On the other hand, in the strongly scattering regime, the forward model becomes non-linear with respect to the RI map [15] and the associated optimization problem is non-convex, providing no guarantee of convergence when employing gradient descent methods, which are commonly used in the literature [9], [10], [11], [12], [13]. Several techniques have been developed for reflection ODT; however, they are limited by their conditions of application, either requiring a 1D reconstruction [16] or relying on specific setups that do not allow for in vivo imaging [17].

Here, we propose a novel reconstruction algorithm to solve the 3D inverse problem, applicable to interferometric imaging in a reflection configuration with multi-wavelength incident wave-fields [18], [19], [20]. The originality of our method lies in the necessity to account for multiple scattering phenomena in order to accurately reconstruct the RI map. It is intuitively based on the idea that, in the multiple scattering regime, the scattering medium itself becomes a source, making it possible to reconstruct a foreground object using the light reflected from the background. The solution to the inverse problem relies on an iterative optimization method that focuses on the low-frequency information of the foreground object. To guide the optimization, special time gating and regularization [21] on the loss will be

Received 26 May 2025; revised 31 October 2025; accepted 8 December 2025. Date of publication 15 December 2025; date of current version 27 February 2026. The work of Alexandre Aubry was supported by European Research Council (ERC) through the European Union's Horizon 2020 Research and Innovation Program under Grant 819261. The work of Josselin Garnier was supported by Agence de l'Innovation de Défense (AID) through Centre Interdisciplinaire d'Études pour la Défense et la Sécurité (CIEDS) project PRODIPO. The associate editor coordinating the review of this article and approving it for publication was Prof. Andreas Hauptmann. *(Corresponding author: Victor Barolle.)*

Thomas Wasik is with the Institut Langevin at ESPCI Paris, CNRS, PSL University, 75005 Paris, France, and also with the CMAP at Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France (e-mail: thomas.wasik@polytechnique.edu).

Victor Barolle is with OWLO SAS, 75014 Paris, France (e-mail: victor@owlo.io).

Alexandre Aubry is with the Institut Langevin at ESPCI Paris, CNRS, PSL University, 75005 Paris, France (e-mail: alexandre.aubry@espci.fr).

Josselin Garnier is with the CMAP at Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France (e-mail: josselin.garnier@polytechnique.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TCI.2025.3644225, provided by the authors.

Digital Object Identifier 10.1109/TCI.2025.3644225

applied. Our method shares similarities with techniques already explored in microscopy [17], [22], ultrasound [23], [24] and seismology [25], [26], where multiply-scattered waves enable the illumination of structures from behind and can give access to their low-spatial frequency RI distribution in a reflection configuration. In contrast with the existing reflection-ODT method, our approach reconstructs the medium without relying on a known reflector, making it a promising candidate for in vivo imaging. It differs from other reflection-based approaches, which reconstruct layers progressively, imaging deeper structures step by step [27].

II. Multiple-Scattering and Filtering Effect

Reconstructing RI maps is an ill-posed problem: for similar measured fields, many different RI distributions can be valid solutions. In transmission, this leads to the missing cone problem, which elongates the reconstructed structures. In reflection, the measured signal corresponds mainly to the high spatial frequencies of the sample: the average RI may no longer be detectable.

In this section, we show that, while this limitation holds under the first-order Born approximation, multiple scattering mitigates this effect (even at second order) and constrains the set of possible solutions of our inverse problem. In particular, we demonstrate that reflections from background structures provide additional information about foreground objects, enabling them to be reconstructed more accurately.

A. Measurement as a Filtering Process

In this paper, we will consider the following experimental configuration [18], [19], [20]: (*i*) a biological medium is illuminated through a microscope objective (MO) by a set of incident wavefields at different wavelengths; (*ii*) for each illumination, the backscattered wave field is collected through the same MO and interferes with a reference beam on a camera; (*iii*) the amplitude and phase of each scattered wavefield is then extracted by an interferometric technique (phase shifting interferometry, on-axis/off-xis holography *etc.*). Our goal is to retrieve from the set of recorded wave-fields the RI distribution inside the biological medium.

A biological medium can be described by an inhomogeneous scattering potential $V(x,y,z) = k_0^2(n^2(x,y,z) - n_\mathrm{b}^2)$, where k_0 is the wavenumber in vacuum, n_b is the background RI and $n(x,y,z)$ is the RI map that we want to recover. The total field E resulting from interaction with the RI map $n(x,y,z)$ satisfies Helmholtz equation:

$$(\nabla^2 + k_0^2 n^2(x,y,z))E(x,y,z) = 0 \tag{1}$$

Let us express the complex wavefield as $E = E_\mathrm{in} + E_\mathrm{s}$ [1], a sum of incident wave solution of the homogeneous Helmholtz equation E_in and a scattered field E_s solving:

$$(\nabla^2 + k_0^2 n_\mathrm{b}^2)E_\mathrm{s}(x,y,z) = U(x,y,z) \tag{2}$$

with Sommerfeld radiation condition. The equation (2) is the homogeneous Helmholtz equation with an internal source term $U(x,y,z) = V(x,y,z)E(x,y,z)$ [8] and it can be solved with the Green's theorem which links the scattered field with the internal source thanks to the homogeneous three-dimensional Green's function $G(x,y,z) = -e^{jn_\mathrm{b}k_0\sqrt{x^2+y^2+z^2}}/(4\pi\sqrt{x^2+y^2+z^2})$:

$$E_\mathrm{s}(x,y,z) = \iiint_{\mathbb{R}^3} U(x',y',z') \times G(x-x', y-y', z-z')\,\mathrm{d}x'\mathrm{d}y'\mathrm{d}z' \tag{3}$$

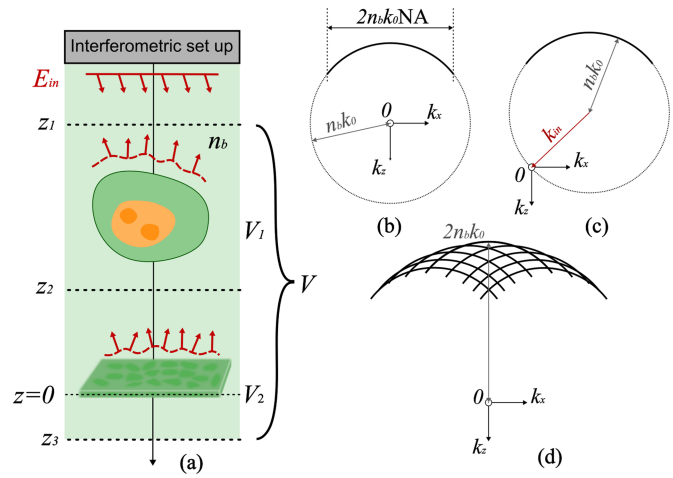


Fig. 1. Schematic representation of the medium to be imaged and illustration of the filtering effects. (a) Illustration of the reflection configuration: the measurement system is positioned at $z < 0$, while the incident waves propagate along the positive z-direction. The medium is separated into 2 parts: an object to be imaged with potential V_1 and a reflector with potential V_2. (b) Description of the filtering operation in the (spatial) Fourier domain: only frequencies of $U = VE$ within the semicircle contribute to the measurement. The limited numerical aperture (NA) of the system restricts the acquisition to a portion of the Ewald half-sphere. (c) Filtering effect on V in reflection under the first order Born approximation: the filter is located at high z-frequencies domain in a reflection configuration. (d) Effect of multiple illuminations on V reconstruction under first order Born hypothesis: by multiplying the number of incident waves, filter functions are superposed, providing more information on the object. However, low spatial frequencies remain inaccessible.

In order to express the filtering process operated by the optical system, it is common to apply the first order Born approximation directly to U ($U = VE_\mathrm{in}$) in order to obtain a linear relation between the scattered wavefield E_s and the scattering potential V. The filtering effect of the numerical aperture (NA) then directly applies to V. However, in the multiple scattering regime, we emphasize that this filtering operation shall be applied to U and not directly to V.

As a convention and as illustrated in Fig. 1(a), the optical axis is here the z-axis and the measurement of the scattered wavefield is carried out at negative z. If evanescent waves are neglected, the expression of the scattered field (3) can be simplified (Appendix A):

$$\mathcal{F}_{2D}[E_\mathrm{s}(\cdot,z)](k_x,k_y) \propto \tilde{U}\left(k_x, k_y, \mathrm{sgn}(z)\sqrt{n_\mathrm{b}^2k_0^2 - k_x^2 - k_y^2}\right) \tag{4}$$

where $\mathcal{F}_{2D}$ corresponds to the 2D Fourier transform and the symbol $\tilde{\cdot}$ stands for the 3D Fourier transform. Equation (4) means that it is only possible to recover information in the Ewald sphere of the source U in Fourier space (Fig. 1(b)). In our case, the

measurement system is placed in negative z: only the lower half sphere can be retrieved.

Experimentally, it is not possible to measure the entire spatial frequency spectrum of $E_s(\cdot, z)$, meaning its high-frequency components remain inaccessible in practice. This results in a filtering effect on $E_s(., z)$ dictated by the numerical aperture (NA) of the measurement system. In the following, we will consider V_{filt}, the reconstructible (i.e., filtered) potential, defined as follows:

$$\begin{aligned}\tilde{V}_{\text{filt}}(k_x, k_y, k_z) &= W_{\text{NA}}(k_x, k_y, k_z)\tilde{U}(k_x, k_y, k_z) \\ &= W_{\text{NA}}(k_x, k_y, k_z)\mathcal{F}_{3D}(VE)(k_x, k_y, k_z)\end{aligned} \tag{5}$$

$$\begin{aligned}W_{\text{NA}}(k_x, k_y, k_z) &= H_{\text{step}}(-k_z - n_b k_0 \text{NA}) \\ &\quad \times \delta(\sqrt{k_x^2 + k_y^2 + k_z^2} - n_b k_0)\end{aligned} \tag{6}$$

where H_{step} is the Heaviside step function. W_{NA} is the transfer function of the imaging system. Its support is limited by the NA. It corresponds to the part of the Ewald sphere that can be grasped by the measurement system. Note that we have here neglected the aberrations induced by the optical system.

B. Information Recovered Under Born Series Approximation

In presence of multiple scattering, the source term $U = EV$ in (2) is too complex to be expressed directly. In order to demonstrate how a multiple scattering system can broaden the spatial filtering effect in comparison with the first-order Born hypothesis, we will make the second-order Born hypothesis. For this proof-of-concept, we will consider a specific type of medium containing a foreground object V_1 that we will intend to image and a background object V_2 that we will exploit to reach this goal. We thus assume a medium of scattering potential $V = V_1 + V_2$ as described in Fig. 1(a).

1) First-Order Born Transfer Function: Under Born's first-order assumption, there is no cross-interactions between the two objects due to the linearity of the model:

$$\begin{aligned}U(x, y, z) &= (V_1(x, y, z) + V_2(x, y, z))E_{\text{in}}(x, y, z) \\ &= V_1(x, y, z)E_{\text{in}}(x, y, z) + V_2(x, y, z)E_{\text{in}}(x, y, z) \\ &= U_1(x, y, z) + U_2(x, y, z)\end{aligned} \tag{7}$$

Let us consider incident plane waves [18], [19], [20]: $E_{\text{in}}(x, y, z) = e^{j(k_{\text{in},x}x + k_{\text{in},y}y + k_{\text{in},z}z)}$. We can then apply the convolution theorem to get [1]:

$$\tilde{U}_1(k_x, k_y, k_z) = \tilde{V}_1(k_x - k_{\text{in},x}, k_y - k_{\text{in},y}, k_z - k_{\text{in},z}) \tag{8}$$

Because of the reflection configuration $k_{\text{in},z} > 0$ and only a part of the high z-frequencies of V can be recovered by the measurements as illustrated in Fig. 1(c). This filtering phenomenon can nevertheless be mitigated by illuminating the sample with other plane waves at different angles. However, as illustrated in Fig. 1(d), this multi-illumination scheme still does not allow our system to recover the low spatial frequencies of the potential.

Under the first Born approximation, only the potential's high $z-$frequencies can be recovered by the optical system. The problem is ill-posed in such a way that it is impossible to retrieve the average value of V_1, for instance. It therefore becomes mandatory to consider a multiple scattering model in order to recover the missing information.

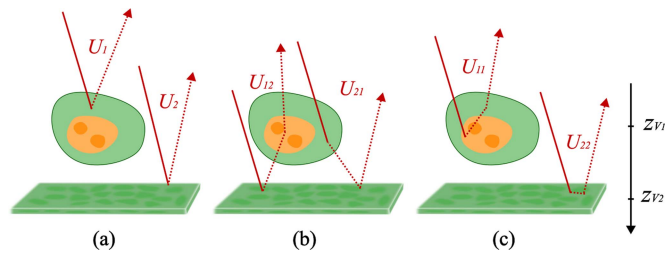


Fig. 2. Decomposition of U according to the different possible light scattering paths under second Born approximation. (a) Representation of single scattering contributions. (b) Crossed terms U_{12} and U_{21} corresponding to successive interactions with the reflector and the target object. (c) Uncrossed double scattering contributions.

2) Second-Order Born Transfer Function: In order to exploit multiply-scattered waves and analytically evaluate their impact on the transfer function, we assume in this subsection that $V_2(x, y, z) = \delta(z)v_2(x, y)$ and that the spatial support of V_1 is localized within a layer $[z_1, z_2]$, with $z_1 < z_2 < 0$. We choose V_2 as planar because we will consider it to be a reflector in our optimization problem, although in practice it never really is and it does not need to be a planar reflector as explained in detail in Section III-C4. We only need to be able to represent its reflectivity as the one of a plane reflector since this assumption allows us to express the second-order Born term analytically.

We model wave propagation by the second-order Born approximation which includes the contribution of double scattering events occurring within the medium:

$$\begin{aligned}U(x, y, z) &= U_1(x, y, z) + U_2(x, y, z) + U_{11}(x, y, z) \\ &\quad + U_{22}(x, y, z) + U_{12}(x, y, z) + U_{21}(x, y, z)\end{aligned} \tag{9}$$

where U_{ij} is defined as:

$$\begin{aligned}U_{ij}(x, y, z) &= V_i(x, y, z) \iiint_{\mathbb{R}^3} G(x - x', y - y', z - z') \\ &\quad \times V_j(x', y', z')E_{\text{in}}(x', y', z')\, dx'dy'dz'.\end{aligned} \tag{10}$$

We note that light scattering can here be expressed as the sum of three types of contributions: (a) single-scattering (or first-order Born) terms (U_1 and U_2), which encode the high spatial frequencies of the objects, (c) double-scattering terms (U_{11} and U_{22}), corresponding to multiple scattering events occurring within each object and (b) cross-interaction terms (U_{12} and U_{21}), which describe scattering paths involving a scattering event with respect to V_1 and a scattering event with respect to V_2. All these contributions are illustrated in Fig. 2.

In the following, we will demonstrate that cross-interaction terms encode the low z-frequencies of V_1. As the V_2-reflector is in the background relative to the V_1-medium, the V_2-reflector will act as a secondary source of waves propagating along the negative z-direction, which are "transmitted" waves from the observer's point of view. In the following, we will refer to this

as indirect transmission, and our goal is to exploit this effect to reconstruct V_1.

The calculation of these cross terms is detailed in Appendix B. Assuming that the supports of V_1 and V_2 are far enough from each other to neglect evanescent waves, we obtain:

$$\tilde{U}_{12}(k_x,k_y,k_z) = \iint_{\mathbb{R}^2} \frac{j\tilde{v}_2(k'_x - k_{\text{in},x}, k'_y - k_{\text{in},y})}{2\sqrt{n_\text{b}^2 k_0^2 - {k'_x}^2 - {k'_y}^2}} \times \tilde{V}_1(k_x - k'_x, k_y - k'_y, k_z + \sqrt{n_\text{b}^2 k_0^2 - {k'_x}^2 - {k'_y}^2})\, dk'_x dk'_y \tag{11}$$

$$\tilde{U}_{21}(k_x,k_y,k_z) = \iint_{\mathbb{R}^2} \frac{j\tilde{v}_2(k_x - k'_x, k_y - k'_y)}{2\sqrt{n_\text{b}^2 k_0^2 - {k'_x}^2 - {k'_y}^2}} \times \tilde{V}_1(k'_x - k_{\text{in},x}, k'_y - k_{\text{in},y}, \sqrt{n_\text{b}^2 k_0^2 - k_x^2 - k_y^2} - k_{\text{in},z})\, dk'_x dk'_y \tag{12}$$

where $\tilde{v}_2$ is the 2D Fourier transform of v_2. The two terms $\tilde{U}_{12}$ and $\tilde{U}_{21}$ are weighted convolutions between the reflector $\tilde{v}_2$ and the scattering potential $\tilde{V}_1$. These convolution products describe the fact that the presence of v_2 drastically enhances the spectral diversity of the field illuminating V_1. The amount of information about V_1 encoded in the components U_{21} and U_{12} is therefore increased compared to the first-order Born term U_1 (8). In particular, the spectral diversity provided by v_2 allows the emergence of the low $z-$frequency components of V_1 in the measured wavefield (5). Thus, the information encoded in cross-interaction terms demonstrate that the measurement is sensitive to the low spatial frequencies of the object. A further analysis of the double-scattering contributions (U_{11} and U_{22} and higher other terms) is therefore unnecessary here, as our goal was only to show this dependence.

To make this dependence on low spatial frequencies more explicit, we consider specific reflector configurations that allow closed-form expressions for (11) and (12). For instance, if the reflector is uniform (*i.e.* $v_2(x,y) = v_{2o}$), we obtain:

$$\tilde{U}_{12}(k_x,k_y,k_z) = \frac{jv_{2o}}{2\sqrt{n_\text{b}^2 k_0^2 - k_{\text{in},x}^2 - k_{\text{in},y}^2}} \times \tilde{V}_1(k_x - k_{\text{in},x}, k_y - k_{\text{in},y}, k_z + k_{\text{in},z}) \tag{13}$$

$$\tilde{U}_{21}(k_x,k_y,k_z) = \frac{jv_{2o}}{2\sqrt{n_\text{b}^2 k_0^2 - k_x^2 - k_y^2}} \times \tilde{V}_1(k_x - k_{\text{in},x}, k_y - k_{\text{in},y}, -\sqrt{n_\text{b}^2 k_0^2 - k_x^2 - k_y^2} + k_{\text{in},z}) \tag{14}$$

This equation can also be found directly using the method of images [28]. Applying (6), it can be noted that encoded information U_{21} is symmetrical with respect to the $k_z = 0$ axis of the encoded information U_{12}, the filter of this configuration is presented in Fig. 3(a).

We can also consider the case of a sinusoidal reflector: $v_2(x,y) = v_{2o}\cos(\gamma x)$ with Fourier transform $\tilde{v}_2(k_x,k_y) = \frac{v_{2o}}{2}(\delta(k_x - \gamma) + \delta(k_x + \gamma))$ with a normal incident wave

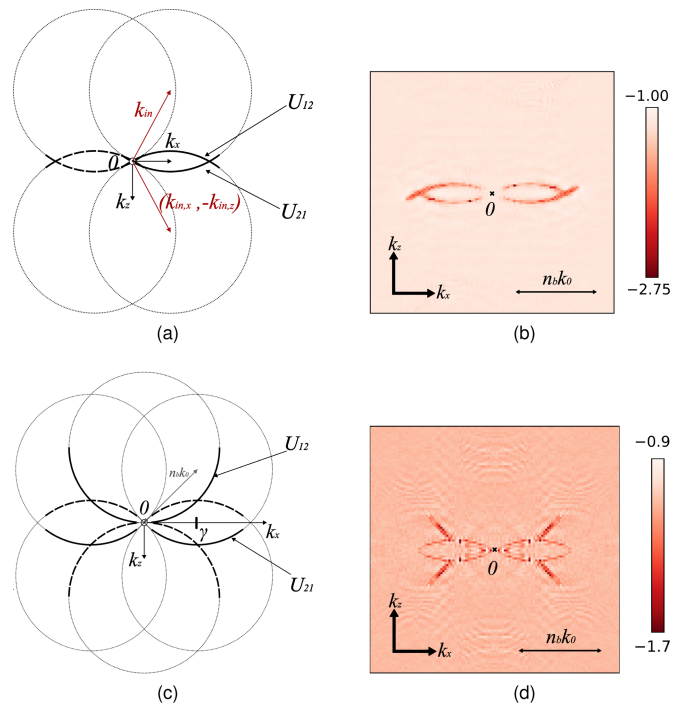


Fig. 3. Support of the spatial frequency spectrum of the object V_1 extracted thanks to the background reflector V_2. (a) Sketch of these spectral components for a uniform reflector (13) and an incident plane wave of wave vector $\mathbf{k}_\text{in} = (k_{\text{in},x}, k_{\text{in},z})$. (b) Reconstruction error in the spatial Fourier domain between the ground truth V_1 and its estimator $\hat{V}_1$ deduced from the optimization process in 2D (17) (with red indicating better results). (c) Sketch of the accessible spectral components of V_1 for a cosine-like reflector of frequency γ (15) and a normal incident plane wave ($k_{\text{in},x} = 0$). (d) Same as in panel (b) but for the cosine-like reflector considered in panel (c). In panels (a) and (c), the dashed lines are deduced from the real-value character of V_1 which implies a symmetric spectrum in the Fourier domain. Panels (b) and (d) are numerically obtained in 2D by considering a Dirac potential V_1 whose Fourier spectrum is uniform.

$E_\text{in}(k_x,k_y,k_z) = e^{jn_\text{b}k_0 z}$. For simplicity, the following expression is evaluated for $\tilde{v}_2(k_x,k_y) = v_{2o}\delta(k_x - \gamma)$ (Fig. 3(c)):

$$\tilde{U}_{12}(k_x,k_y,k_z) = \frac{jv_{2o}\tilde{V}_1(k_x - \gamma, k_y, -k_z + \sqrt{n_\text{b}^2 k_0^2 - \gamma^2})}{2\sqrt{n_\text{b}^2 k_0^2 - \gamma^2}} \tag{15}$$

$$\tilde{U}_{21}(k_x,k_y,k_z) = \frac{jv_{2o}}{2\sqrt{n_\text{b}^2 k_0^2 - (k_x - \gamma)^2 - k_y^2}} \times \tilde{V}_1(k_x - \gamma, k_y, \sqrt{n_\text{b}^2 k_0^2 - (k_x - \gamma)^2 - k_y^2} - n_\text{b} k_0) \tag{16}$$

In the following section, we use these two peculiar configurations to illustrate how the background reflector V_2 can be leveraged to retrieve the low $z-$frequencies of the foreground structure V_1.

C. Filtering Effect for Particular Configurations

To that aim, we describe the inverse problem of light scattering as the following optimization problem:

$$\hat{\mathbf{V}} = \underset{\mathbf{V}\in\mathbb{R}^N}{\arg\min}\, \mathcal{L}(\mathbf{V}) \tag{17}$$

$$\mathcal{L}(\mathbf{V}) = \|\mathbf{y}_{\text{model}}(\mathbf{V}, \mathbf{E}_{\text{in}}) - \mathbf{E}_{\text{m}}\|^2 \tag{18}$$

Here, $\mathbf{V}$ and $\mathbf{E}_{\text{in}}$ represent discrete approximations of the potential V and the incident field E_{in} described respectively as vectors in $\mathbb{R}^N$ and $\mathbb{C}^N$ (the square domain is discretized as a uniform grid of size $N = N_x \times N_y$). The objective is to minimize a loss function over one illumination to obtain an estimate of the potential V consistent with (13) and (15). Equation (17) represents the fidelity term for a single measurement, quantifying the similarity between the field predicted by the forward model, $\mathbf{y}_{\text{model}}$, and the measured scattered field $\mathbf{E}_{\text{m}}$.

As a forward model, we will use the second order Born in order to take into account the predominant multiple scattering paths. The computed total field becomes:

$$\mathbf{E}_{\text{Born-2}}(\mathbf{V}, \mathbf{E}_{\text{in}}) = \mathbf{E}_{\text{in}} + \mathbf{G}\text{diag}(\mathbf{V})\mathbf{E}_{\text{in}} + \mathbf{G}\text{diag}(\mathbf{V})\mathbf{G}\text{diag}(\mathbf{V})\mathbf{E}_{\text{in}} \tag{19}$$

$$\mathbf{y}_{\text{Born-2}}(\mathbf{V}, \mathbf{E}_{\text{in}}) = \mathbf{H}\text{diag}(\mathbf{V})\mathbf{E}_{\text{Born-2}}(\mathbf{V}, \mathbf{E}_{\text{in}}) \tag{20}$$

where $\mathbf{H} \in \mathbb{C}^{N_c \times N}$ is the transfer function describing wave propagation from the sample to the N_c pixels of the measurement system and $\text{diag}(\mathbf{V}) \in \mathbb{C}^{N \times N}$ the square diagonal matrix whose diagonal elements are given by the components of $\mathbf{V}$. In real space, this model involves convolution (or matrix) products, which are computationally expensive. However, this can be bypassed by moving to Fourier space and using Fourier transforms of Green's functions: $\mathbf{G} = \mathbf{F}^\dagger \text{diag}(\tilde{\mathbf{G}})\mathbf{F} \in \mathbb{C}^{N \times N}$. Here, $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the Fourier transform operator, which is efficiently computed in practice using the Fast Fourier Transform (FFT) algorithm. The computational cost of the forward model is then reduced as it only involves a sequence of term-by-term products and FFT operations. Because of the singularity of the Green's function in Fourier space, we use modified versions $\tilde{\mathbf{G}}^L \in \mathbb{C}^N$, as defined in Refs. [12], [29], [30] to ensure computation stability. This modified Green's function correspond to cropped version in real space, and a twofold zero-padding is required to ensure accurate computation.

In the following, we will attempt to solve (17) using the gradient descent algorithm. However, the forward model (19) is a second-order polynomial function of V, and gradient descent does not have theoretical guarantees to converge to a global solution of (17). To circumvent this problem and to illustrate the information that can be recovered from the system without dealing with convergence issues (which will be addressed in the next sections), we will fix the background reflector in the simulation as the ground truth.

The model used to calculate the ground truth field is the Lippmann-Schwinger forward model described in Ref. [29]. For sake of illustration of the filtering effect operated by our imaging method, the scattering potential V_1 that we consider here is a Dirac point, as it gives rise to a uniform spectrum in Fourier space. Fig. 3(b) and (d) show the reconstruction error between the estimated potential $\hat{V}_1$ and its ground truth V_1 in Fourier space for the uniform and sinusoidal reflectors V_2. A quantitative agreement is found between $\hat{V}_1$ and V_1 over the spatial frequencies captured by the imaging system. The recovered spatial frequency spectra of V nicely match with our analytical predictions (13) and (15) sketched in Fig. 3(a) and (c), respectively. In particular, it illustrates how the presence of reflector V_2 enables multiply scattered trajectories that can be leveraged for retrieving the low $z-$frequency components of the foreground object V_1.

III. Reconstruction Algorithm

As discussed in the previous section, an internal background reflector helps to retrieve the low-frequency information from the object in the foreground. However, this information can only be obtained by analyzing the multiply scattered components of the reflected wavefield because the single scattering contribution does not contain the low spatial frequencies of the foreground object. As a result, for reflection-mode ODT, the forward model cannot be linear with respect to the scattering potential, making the optimization problem non-convex. Gradient descent algorithms, such as stochastic gradient descent (SGD) or even FISTA [31], therefore lack theoretical guarantees for convergence.

In this section, we propose a reconstruction algorithm that enables good reconstruction in reflection configurations despite the non-convexity of the problem. Our approach is based on (i) the introduction of a new loss function that focuses on the low-frequency components of the foreground object to be reconstructed, (ii) the use of Total Variation (TV) regularization (a commonly applied method [21], [32] that constrains the reconstruction to block-wise behavior) in the foreground region only (i.e. between z_1 and z_2 in Fig. 1), and (iii) a proximal algorithm based on the Adam optimizer [33] ensuring a good convergence despite the non-convexity of the problem.

A. A Temporal Loss Function to Focus on Low-Frequency Information

As illustrated in (9), the total information recovered is a sum of terms which encode different pieces of information: U_1 encodes high $z-$frequencies of the object V_1 (8) while U_{12} (11) and U_{21} (12) encode its low $z-$frequencies. In the ideal case, if V_2 is known, then we can reconstruct both low and high frequencies of the object. However, in practice, the orders of magnitudes of these two terms can be very different, leading to difficulties for the reconstruction of the low $z-$frequencies of the foreground object if V_2 is too small. We could assume that V_2 is on average much larger than V_1, but this is not necessarily the case in reality. Ideally, the terms related to direct reflection with the object should be separated from those containing cross interactions between the reflector and the object to be imaged. This can be done by considering the reflected wavefield in the time domain, as it can be done in ultrasound with broadband signals [23].

In optical microscopy, the time-dependence of the reflected wavefield can be retrieved by low coherence interferometry [34]. An alternative strategy is to perform a polychromatic measurement of the wavefield using a spectrometer [35], [36] or a swept source laser [37], [38] as in spectral domain OCT. In this way, $\mathbf{E}_{\text{m}}$ can be measured at different temporal frequencies ω. In post-processing, an inverse Fourier transform with respect to ω can then provide the time-dependent wavefield $\mathbf{E}_{\text{m}}(t)$ that would

be obtained if the sample was illuminated by a coherent incident wave-packet.

With V_1 in the foreground and V_2 in the background, the ballistic term U_1 and crossed terms U_{12} and U_{21} can be associated with different central times-of-flight $t_1 \sim 2n_\mathrm{b} z_{V_1}/c$ and $t_2 \sim 2n_\mathrm{b} z_{V_2}/c$, with z_{V_1} and z_{V_2}, the central depths of the foreground object V_1 and the background reflector V_2 (Fig. 2). This assumption is valid in a weakly scattering regime in which the echo time t is directly proportional to the depth z of the associated reflector. Otherwise, high-order multiple scattering would lead to a large temporal dispersion of the incident wave packet [39], [40], [41], making their discrimination impossible on a time-of-flight basis.

By considering the ith Born order forward model, we have:

$$\mathbf{E}_\mathrm{m}(t) = \mathcal{F}_\omega^{-1}[\mathbf{E}_\mathrm{m}(\omega)](t) \tag{21}$$

$$\mathbf{y}_{\text{Born-}i}(\mathbf{V}, \mathbf{E}_\mathrm{in}, t) = \mathcal{F}_\omega^{-1}[\mathbf{y}_{\text{Born-}i}(\mathbf{V}, \mathbf{E}_\mathrm{in}, \omega)](t) \tag{22}$$

We then define a new temporal loss function over the N_in incident illuminations [20]:

$$\mathcal{L}(\mathbf{V}, t) = \sum_{j=1}^{N_\mathrm{in}} \mathcal{L}_j(\mathbf{V}, t) \tag{23}$$

$$\mathcal{L}_j(\mathbf{V}, t) = \|\mathbf{y}_{\text{Born-}i}(\mathbf{V}, \mathbf{E}_{\mathrm{in},j}, t) - \mathbf{E}_{\mathrm{m},j}(t)\|^2, \tag{24}$$

and the fidelity term of the loss function becomes

$$\mathcal{L}_\mathrm{fid}(\mathbf{V}, \gamma) = \mathcal{L}(\mathbf{V}, t_2) + \gamma \mathcal{L}(\mathbf{V}, t_1) \tag{25}$$

The purpose of this loss function is to treat the information from direct reflection of the foreground object (associated with time t_1) and the reflection from the background reflector and the cross interaction terms (associated with time t_2) with different weights. If the background reflector has a low reflectivity and $\gamma = 1$, then the minimization algorithm of the loss would prioritize the reconstruction of the high spatial frequencies of the foreground object, which may cause the whole optimization process to become inefficient. When γ is small, the impact of the direct reflection from the object is reduced. As a result, the information about the object to be imaged becomes a nonlinear function of its spatial low frequencies, effectively "forcing" the model to focus on a solution depending on the object's low-frequency components.

Thus, by considering the problem from a temporal perspective, we can weight the two effects. It becomes even possible, by adjusting γ, to recover both high and low spatial frequencies of the foreground object, improving the reconstruction quality of the sample. We will explain in Appendix C how γ is selected. Moreover, t_1 and t_2 can be considered as specific moments or as time intervals. In practice, the use of time intervals leads to better reconstruction results.

It is important to note that minimizing (25) remains a non-convex problem, which can be challenging even for classical gradient descent algorithms. We address this issue in the next section.

B. Regularization and Optimization Algorithm

The non-convexity of the loss function can lead to multiple local minima or saddle points in the loss landscape in (25). Using a regularization term introduces a priori information that helps reduce the number of possible solutions and improves the convergence of gradient-based methods. The objective function to minimize in our problem becomes:

$$\mathcal{L}_\mathrm{rODT}(\mathbf{V}, \gamma, \tau) = \mathcal{L}(\mathbf{V}, t_2) + \gamma \mathcal{L}(\mathbf{V}, t_1) + \tau \mathcal{R}(\mathbf{V}) \tag{26}$$

in which $\mathcal{R}$ denotes the regularization term and τ is the regularization parameter.

A commonly used regularization term is the Total Variation loss (TV) [32], which is today commonly used for ODT in transmission configuration [10], [21]. This regularization encourages the reconstruction of piecewise constant potentials and avoid noisy reconstructions. Noise can arise, for example, if the reflector is not smooth, making the reconstruction task more difficult. A widely used optimization approach for ODT in transmission configuration is proximal algorithms [42], which are well suited for TV regularization by allowing the iterative optimization of loss functions composed of differentiable and non-differentiable terms. Let $\mathbf{V}^{(k)}$ be the estimator of $\mathbf{V}$ at iteration k. The idea is to split each optimization step into two parts:

i) A standard gradient descent step is first performed:

$$\mathbf{V}^{(k)} \leftarrow \mathbf{V}^{(k)} - \alpha \nabla \mathcal{L}_\mathrm{fid}(\mathbf{V}^{(k)}, \gamma).$$

ii) The proximal operator is then applied:

$$\mathbf{V}^{(k+1)} \leftarrow \mathrm{prox}_{\alpha\tau\mathcal{R}}(\mathbf{V}^{(k)}),$$

which corresponds to the minimization of:

$$\mathrm{prox}_{\alpha\tau\mathcal{R}}(\mathbf{V}^{(k)}) = \underset{\mathbf{V}\in\mathbb{R}^N}{\arg\min} \left\{ \frac{1}{2}\|\mathbf{V}^{(k)} - \mathbf{V}\|^2 + \alpha\tau\mathcal{R}(\mathbf{V}) \right\}$$

The proximal operator for TV regularization has the advantage of being well-documented [21], [43].

FISTA and variants are gradient descent or stochastic gradient descent based algorithms. This class of algorithms are designed for convergence of convex loss functions. Reflection ODT being a non convex problem, it was natural to search for a proximal algorithm based on an optimizer efficient for non-convex optimization. In that respect, an adequate tool is the Adam optimizer [44] which is an adaptive learning rate method. This optimizer is suitable for stochastic gradient descent which allows a reduction of the memory requirement for the optimization process. Moreover, Adam showed better efficiency in non-convex optimization compared to Nestervov stochastic gradient descent algorithm [45], [46].

The challenge with adaptive algorithms lies in the fact that the learning rate depends on the parameter being optimized, thereby requiring a proximal operator that accounts for this dependency. Ref. [33] introduces a method called *ProxGen* that extends proximal gradient algorithms to adaptive optimization techniques such as Adam. Based on this approach, we will use the proximal algorithm described in Algorithm 1.

Algorithm 1: Reflection ODT Algorithm: *ProxGen*

input: $\mathbf{V}^{(0)}$, α, ϵ, β_1, β_2, γ and τ
Initialize $\mathbf{m}_0 = 0$ and $\mathbf{v}_0 = 0$
for $k = 1, 2, \ldots, N_{\text{epoch}}$ **do:**
  Choose randomly an index $i_k \in \{0, \ldots, N_{\text{in}}\}$
  $\mathbf{g}_k \leftarrow \nabla \mathcal{L}_{\text{fid},i_k}(\mathbf{V}^{(k-1)}, \gamma)$
  $\mathbf{m}_k \leftarrow \beta_1 \mathbf{m}_{k-1} + (1-\beta_1)\mathbf{g}_k$
  $\mathbf{v}_k \leftarrow \beta_2 \mathbf{v}_{k-1} + (1-\beta_2)(\mathbf{g}_k \circ \mathbf{g}_k)$
  $\mathbf{V}^{(k)} \leftarrow \arg\min_{\mathbf{V}\in\mathbb{R}^N} \{\langle \mathbf{V}, \mathbf{m}_k\rangle + \tau\mathcal{R}(\mathbf{V}) + \frac{\|\mathbf{V}-\mathbf{V}^{(k-1)}\|^2_{\sqrt{\mathbf{v}_k}+\epsilon}}{2\alpha}\}$
return The prediction $\mathbf{V}^{(N)}$

Algorithm 2: Proximal Calculus

input: $\mathbf{V}^{(k-1)}$, α, ϵ, $\mathbf{m}_k$, $\mathbf{v}_k$ and τ
$\mathbf{V}^{(k,0)}_{\text{prox}} = \mathbf{V}^{(k-1)}$
$\mathbf{z}_k = \mathbf{V}^{(k-1)} - \frac{\alpha}{\sqrt{\mathbf{v}_k}+\epsilon}\mathbf{m}_k$
for $t = 1, 2, \ldots, T$ **do:**
  $\mathbf{g}_t \leftarrow (\sqrt{\mathbf{v}_k}+\epsilon)(\mathbf{V}^{(k,t-1)}_{\text{prox}} - \mathbf{z}_k)$
  $\mathbf{V}^{(k,t)}_{\text{prox}} \leftarrow \text{prox}_{\tau\alpha/\max(\sqrt{\mathbf{v}_k}+\epsilon)\mathcal{R}}(\mathbf{V}^{(k,t-1)}_{\text{prox}} - \frac{\mathbf{g}_t}{\max(\sqrt{\mathbf{v}_k}+\epsilon)})$
return $\mathbf{V}^{(k)} = \mathbf{V}^{(k,T)}_{\text{prox}}$

In Algorithm 1 we have introduced a new norm as follows: $\|\mathbf{x}\|^2_{\mathbf{a}} = \langle \mathbf{x}, \mathbf{a}\circ\mathbf{x}\rangle$ where $\circ$ denotes the element-wise (Hadamard) product. The hyperparameters of the algorithm are: the learning rate α, the momentum parameters β_1 and β_2, the weight factor γ, and the regularization parameter τ. The last line of Algorithm 1 corresponds to a minimization problem, which can be reformulated using a new proximal operator:

$$\mathbf{V}^{(k)} \in \arg\min_{\mathbf{V}\in\mathbb{R}^N} \left\{ \langle \mathbf{V}, \mathbf{m}_k\rangle + \tau\mathcal{R}(\mathbf{V}) + \frac{\|\mathbf{V}-\mathbf{V}^{(k-1)}\|^2_{\sqrt{\mathbf{v}_k}+\epsilon}}{2\alpha} \right\}$$
$$= \text{prox}^{\sqrt{\mathbf{v}_k}+\epsilon}_{\alpha\tau\mathcal{R}}\left(\mathbf{V}^{(k-1)} - \frac{\alpha}{\sqrt{\mathbf{v}_k}+\epsilon}\mathbf{m}_k\right) \quad (27)$$

with $\text{prox}^{\mathbf{a}}_h(\mathbf{z}) = \arg\min_{\mathbf{x}}\{h(\mathbf{x}) + \frac{1}{2}\|\mathbf{x}-\mathbf{z}\|^2_{\mathbf{a}}\}$ and the division is performed element-wise. This minimization problem is convex and can be solved by the proximal algorithm [47]. The gradient of the objective function is $(\sqrt{\mathbf{v}_k}+\epsilon)(\mathbf{x}-\mathbf{z})$ and the Lipschitz constant is $L_k = \max(\sqrt{\mathbf{v}_k}+\epsilon)$.

An efficient method to compute $\mathbf{V}^{(k)}$is then described in Algorithm 2. The parameter T defines the number of iterations for Algorithm 2, thereby controlling its convergence behavior. In practice, we set T to be large and apply a stopping criterion which corresponds to the relative change from one iteration to the next:

$$\frac{\|\mathbf{V}^{(k,t)}_{\text{prox}} - \mathbf{V}^{(k,t-1)}_{\text{prox}}\|}{\|\mathbf{V}^{(k,t)}_{\text{prox}}\|} \leq \delta \quad (28)$$

with $\delta = 10^{-6}$ in all following simulations.

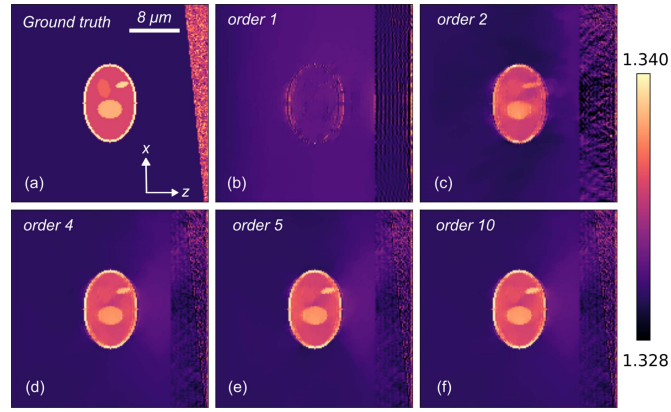


Fig. 4. Reconstruction of a sample composed of a cell and a reflector for different forward models used in the *ProxGen* algorithm with the weighted temporal loss function ($\gamma = 0.1$). A minimum regularizer is used to improve the quality of reconstruction. The reconstruction is impossible by using first order Born approximation.

C. Simulation Results

To validate our reconstruction algorithm, we applied it to 2D simulations for computational efficiency. 3D-simulated medium results are presented at the end of the section. For both 2D and 3D cases, to ensure consistency with the theoretical framework—which assumes the convergence of the Born series—the RI deviations are set to $\Delta n = 0.01$. This RI contrast, although not representative of all biological media, is comparable to the RI contrasts of biological samples considered in recent work [48], [49], [50]. The surrounding medium has a RI $n_{\text{b}} = 1.33$. We consider incident plane waves with angles confined within a numerical aperture of NA $= 0.8$, and wavelengths ranging from 800 to 875nm.

For both 2D and 3D, the ground truth is, as above, generated using a Lippmann-Schwinger forward model with a modified Green's function $\tilde{\mathbf{G}}^L$ [12], [30]. This model accurately accounts for all multiple-scattering effects and is a well-established method for simulating light propagation in complex media. The use of these modified Green's functions requires adding a padding region with a RI n_{b}, whose length matches that of the sample. The gradient for our reconstruction is calculated by using the autograd functionality of torch library [51]. For each reconstruction, the reflector is not known in advance and the starting RI map is always a uniform medium with a RI equal to n_{b}.

1) 2D Simulation: To evaluate the effectiveness of our approach, we tested our reconstruction algorithm on 2D simulated samples with a resolution of $dx = dz = 200$ nm in a square of size $N_x = N_z = 161$. Samples under study mimic the RI distribution of a cell. We have taken $N_{\text{in}} = 80$ incident waves taken uniformly and sampled over 11 wavelengths to guarantee a sufficient temporal resolution and discriminate the direct object echo and the background reflector components. The reflector placed in our simulation behind the sample corresponds to an interface with average RI $n = 1.335$, to which we add Gaussian noise with standard deviation $\delta n = 0.002$. For optimal convergence, all 2D hyperparameters have been carefully tuned for

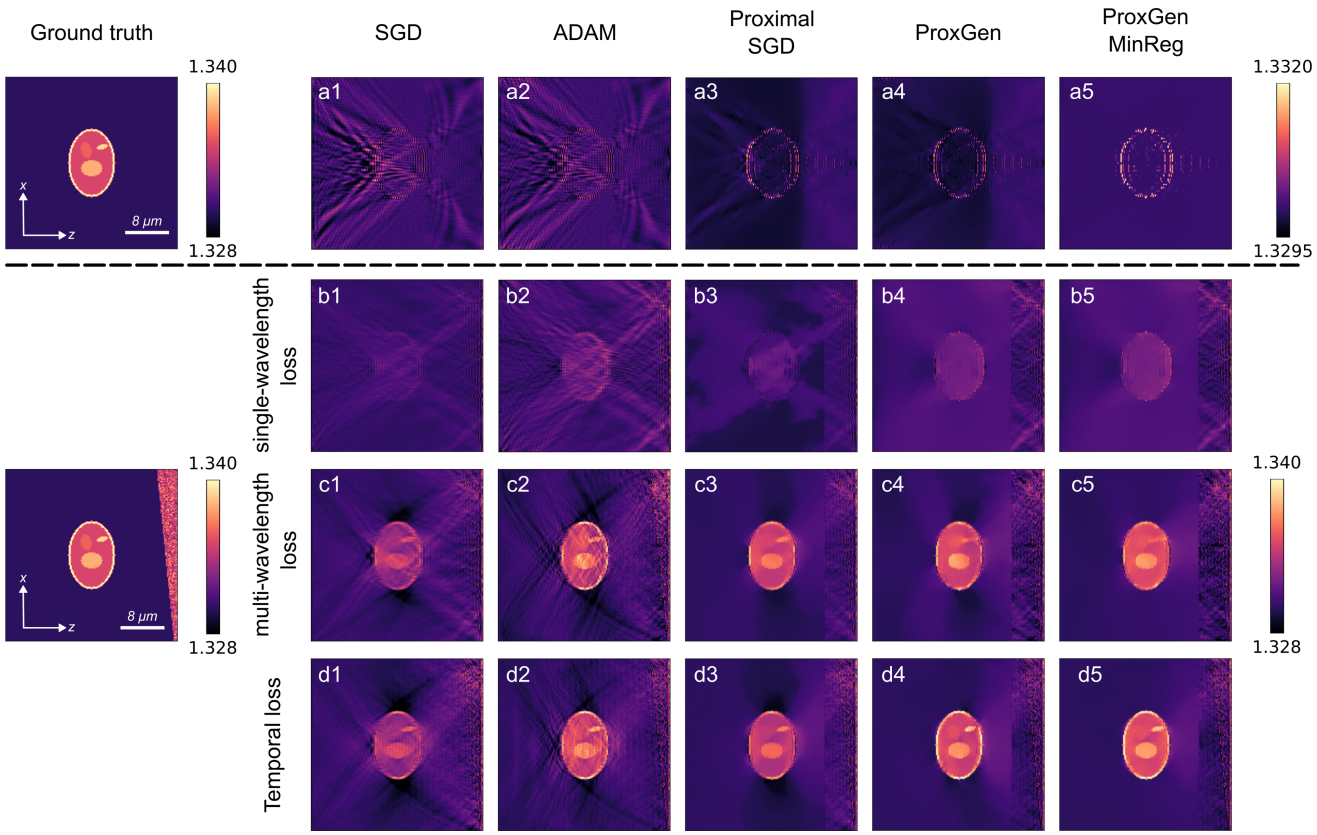


Fig. 5. 2D reconstruction results for different configurations with order 5 Born forward model. Each column corresponds to an optimization algorithm. Row (a) shows the results for the single-wavelength loss without a reflector. Since the results for different loss functions without reflector are similar, they are not displayed here. The last three rows present the reconstructions using different loss functions with ground truth including a reflector. The unweighted temporal loss is not shown, as it produces results similar to the multi-wavelength loss (row (c)): the Fourier transform does not significantly impact the results. The weighted temporal loss shown here corresponds to (23) with a coefficient $\gamma = 0.1$.

Figs. 4–7. We also validated our approach using simulated noisy observations. Further details can be found in Appendix C.

2) Choice of the Forward Model: Choosing an appropriate forward model for the optimization algorithm is essential for improving the quality of the RI reconstruction. Indeed, this model should closely account for the multiple-scattering contribution in order to accurately reconstruct the sample. Fig. 4 illustrates this fact by displaying the result of our optimization algorithm for different forward models. Each model involves a different order of the Born approximation. Not surprisingly, the first-order Born model (Fig. 4(b)) only grasps the high spatial frequencies of the sample and only provides an image of the external surface of the cell. Remarkably, the second-order Born model reveals the intra-cellular components (Fig. 4(c)). Nevertheless, the reconstruction remains imperfect due to discrepancies between the real light propagation and the forward model used in the optimization algorithm. Those imperfections can be mitigated by increasing the degree of complexity of the forward model (Fig. 4(d)–(f)). The 5th-order Born series forward model (Fig. 4(e)) offers a good trade-off between the accuracy of the RI map reconstruction and the computational complexity. This forward model includes more scattering events than in our theoretical analysis. One might think that the reconstruction quality also depends on other multiple-scattering terms, not just the cross-interactions U_{12} and U_{21}. Yet this effect remains minor compared to the contribution of the cross-interaction terms U_{12} and U_{21}, as detailed in Appendix D.

3) Regularization Strategies: In the following, total variation (TV) regularization is applied only in the region of the object V_1 that we aim to reconstruct. Since the reflector is difficult to reconstruct due to the ill-posed nature of the problem, we choose not to apply any regularization or constraint to it. This allows the reconstruction to produce a RI that, while not physically accurate, effectively mimics the reflector's reflectivity. Applying regularization only in foreground also improves the reconstruction, as shown in Appendix C.

Additionally, a differentiable regularization term, referred to as *Minimum regularization*, is introduced. This term prevents predictions below the RI n_{b}, acting as a Ridge regression for points with a RI lower than n_{b}. Since this regularization is differentiable, it can be easily integrated into the temporal loss function. This additional regularization corresponds to a positivity constraint. Choi et al. [2] actually demonstrate that it reduces the underestimation of the RI in a transmission configuration. For the same reason as the TV regularization, minimum regularization is then only applied to the V_1-object region (contained in the layer $[z_1, z_2]$ in Fig. 1(a)) but not to the V_2-reflector (layer $[z_2, z_3]$ in Fig. 1(a)).

Fig. 5 illustrates the improvement in the object reconstruction achieved by applying the aforementioned strategies. The first

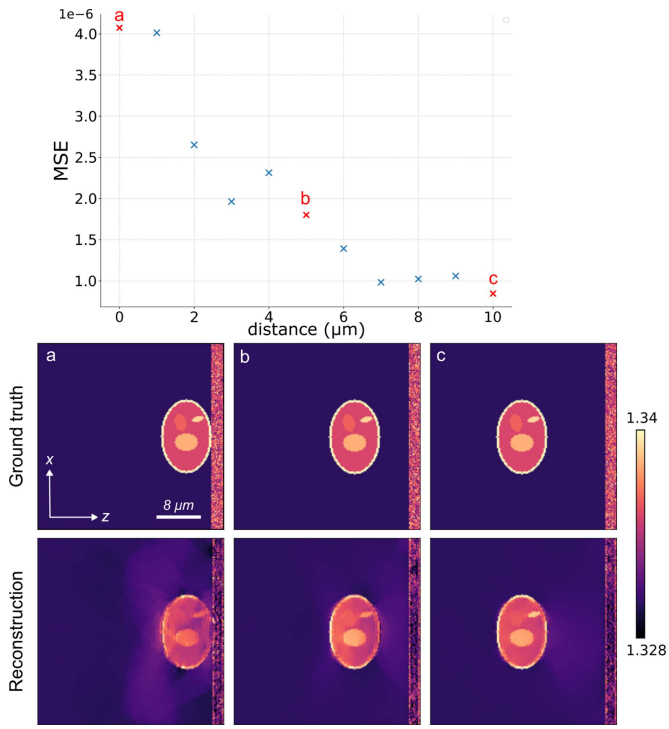


Fig. 6. Effect of the distance between reflector and cell on reconstruction quality. The mean squared error (MSE), calculated only inside the cell, shows how the reconstruction differs from the ground-truth cell as a function of distance. Reconstructions at different distances are shown to illustrate how reconstruction quality decreases as the cell and reflector get closer.

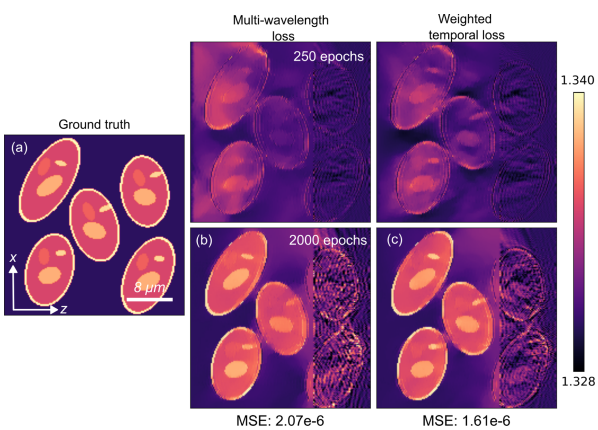


Fig. 7. Reconstruction of complex samples thanks to the *ProxGen algorithm* by using different loss functions. Here the reflector corresponds to the background cells at the right. The regularization (TV by proximal and minimum regularization) are applied at the left of the image and not for the reflector. As in previous cells the initialization corresponds to uniform medium. The reconstruction is shown both at an early stage of training (250 epochs) and at the final stage (2000 epochs), highlighting that the temporal loss improves reconstruction stability and consequently reduces the dependence on regularization. The MSE represents the mean squared error computed inside the cells at the end of training.

row (Fig. 5(a)) highlights the challenge of reconstructing the sample without a reflector, even when using various iterative optimization techniques. In this case, the scattering potential remains trapped in a local minimum or flat region, preventing accurate reconstruction. As previously highlighted, the low-spatial frequencies of the object can be hardly retrieved in such a configuration. However, the second row (Fig. 5(b)) shows that simply adding a reflector is not sufficient if the temporal aspect is not considered. In a single-wavelength experimental setup, the singly-scattered term U_1 and U_2, as well as the second-order crossed terms U_{12} and U_{21}, cannot be disentangled. This mixing prevents the convergence of the optimization algorithms. The last two rows of Fig. 5 demonstrate that temporal resolution is essential for accurate cell reconstruction. The multi-wavelength loss (Fig. 5(c)) corresponds to a broadband loss function that incorporates multiple wavelengths:

$$\mathcal{L}_B(\mathbf{V}) = \sum_{\omega} \|\mathbf{y}_{\text{Born-}i}(\mathbf{V}, \mathbf{E}_{\text{in}}, \omega) - \mathbf{E}_{\text{m}}(\omega)\|^2. \tag{29}$$

Due to Parseval's theorem, $\mathcal{L}_B(\mathbf{V})$ is actually equivalent to a temporal loss integrated over all times of flight:

$$\mathcal{L}_B(\mathbf{V}) = \sum_{t} \|\mathbf{y}_{\text{Born-}i}(\mathbf{V}, \mathbf{E}_{\text{in}}, t) - \mathbf{E}_{\text{m}}(t)\|^2 \tag{30}$$

All time points are here weighted equally. While this approach enables good cell reconstruction, further improvement can be achieved using a weighted temporal loss, as described in (25) and illustrated by Fig. 5(d). Thus, employing a multi-wavelength configuration and measuring the amplitude and phase of the reflected wavefield (and not just the intensity) are critical to reducing the ill-posedness of the inverse problem. Indeed, it gives access to the time-dependence of each reflected wave-field which greatly facilitate the reconstruction.

The TV regularization is here essential to reconstruct correctly the sample as the reconstruction is noisy without it. While SGD and Adam-based gradient descent methods yield similar results for the multi-wavelength loss ((Fig. 5(c3) and (c4)), Adam-based methods perform better when using the weighted temporal loss ((Fig. 5(d3) and (d4)). The results improve further when applying minimum regularization, which prevents the RI from falling below n_{b}.

In our reconstruction framework, the foreground object and the background reflector are spatially separated for two main reasons. First, our theoretical model assumes that evanescent waves from the reflector are negligible. Second, the temporal separation helps the optimization algorithm converge more reliably as mentioned above. However, the object can still be reconstructed (even if less accurately) when the spatial separation is limited, as described in Fig. 6. As expected, the reconstruction quality decreases as the reflector gets closer, dropping below 1μm. Since the measured signal is included in $[800nm, 875nm]$, the temporal resolution of the reconstructed wave packet is $d_{z,\text{temp}} \approx 3.5\mu\text{m}$. At distances around $1\mu m$, the cross-interaction and direct reflection terms from (9) begin to overlap, causing mixing that limits reconstruction performance.

4) Importance of a Weighted Temporal Loss: Unlike alternative methods that reconstruct the RI of an object using a deterministic reflector [17], our approach does not require any prior knowledge on it. Indeed, this reflector can belong to the inspected medium. The RI distribution can be reconstructed, even when the background reflector is highly irregular and initially unknown.

Additionally, this algorithm can be applied to more complex samples. For instance, we can consider a 2D sample made

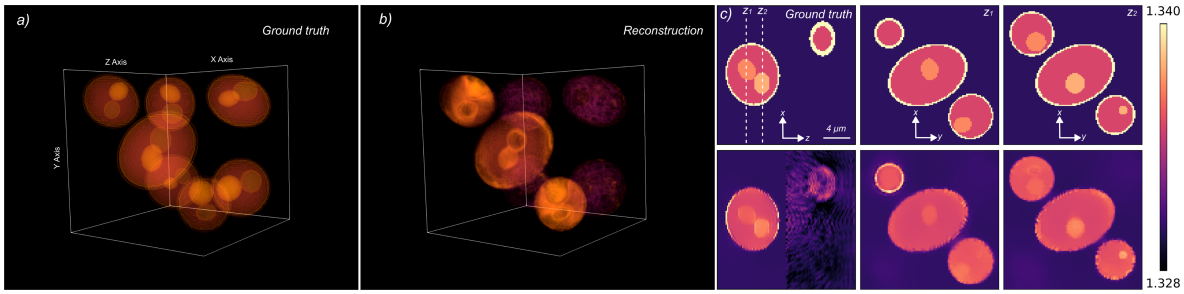


Fig. 8. 3D Reconstruction with the proposed method of $20 \times 20 \times 20\,\mu m$ sample made of several cells. Figure (a) shows the ground truth 3D rendering of the RI map, figure (b) corresponds to the reconstruction and figure (c) shows a series of ground truth and prediction 2D cuts. The inhomogeneity of the reconstructed cells in figure (b) is due to the fact that the reconstructed background cells correspond to an equivalent reflectivity of lower optical index than the ground truth. An animation is available in the supplementary material.

up of multiple cells, and try to reconstruct the upstream cells by using the background cells as reflectors. Examples of such reconstructions are shown in Fig. 7, where we can observe that the weighted temporal loss (Fig. 7(c)) leads to more satisfactory results and shows noticeably better reconstruction even at early stages of convergence compared to the multi-wavelength loss (Fig. 7(b)) [see the comparison with ground truth displayed in Fig. 7(a)]. For both losses (multi-illumination and weighted temporal), the low and high spatial frequency components are effectively separated in time, thanks to Parseval's theorem. However, the multi-illumination loss treats all times equally, because the energy received at both echo times is similar, the model struggles to reconstruct the low-frequency components of the object, which are mainly carried by the weaker U_{12} and U_{21} terms rather than by the dominant U_1 term. On the contrary, the weighted temporal loss enhances the strength of backscattered echoes and leads to a better image the sample.

Thus, despite being far from the ideal case of a planar reflector, our method enables high-quality reconstruction of RI maps. Moreover, the use of a weighted loss allows for the reconstruction of both low and high spatial frequencies, effectively mitigating the missing cone problem commonly encountered in transmission. However, as shown in Fig. 5, achieving a good reconstruction requires the presence of a reflector (and therefore a complex medium), as well as a temporal loss (hence a multi-illumination device), unlike in a transmission configuration.

5) 3D Reconstruction: To demonstrate the scalability of our model, we also simulate 3D reconstructions. The sample under study is displayed in Fig. 8(a). The simulation domain is a cube of size $N_x = N_y = N_z = 101$ of resolution $dx = dy = dz = 200$ nm. The simulated illumination sequence involves 80 illuminations spirally distributed within a pupil of numerical aperture $\mathrm{NA} = 0.8$. Each illumination is repeated for 11 wavelengths uniformly distributed between 800 and 875nm. The forward model used in our optimization process is still a 5th order Born model since we remain in a weakly multiple scattering regime. The optimization is performed over 1000 epochs to ensure convergence, although a good reconstruction can already be achieved after 400 epochs. The reconstruction was efficiently carried out on an RTX 6000 Ada GPU with 48GB of VRAM. Each epoch took 30 seconds, which is longer than for an equivalent transmission reconstruction, as wave propagation needs to be computed for multiple wavelengths.

Due to the long training time, we fine-tuned our optimization algorithm on a smaller cube and applied the same hyperparameters to the larger one. We used a learning rate of $\alpha = 2 \times 10^{-5}$, momentum parameters $(\beta_1, \beta_2) = (0.9, 0.999)$, and a weight factor $\gamma = 0.1$ to generate the figure. The result of the optimization process is displayed in Fig. 8. It appears that in 3D, it is still possible to obtain good reconstructions of cells in the foreground, by using cells in the background. The average values of the various substructures remain consistent with the ground truth. However, the reconstruction quality is not as good as in 2D, due to the larger number of parameters to optimize and the increased difficulty in tuning the hyperparameters.

IV. Conclusion

We have introduced a novel optimization algorithm for 3D refractive index reconstruction in reflection-mode microscopy. Our method is based on the exploitation of multiple light scattering. More precisely, our approach consists in leveraging the light backscattered by deep reflectors to map the refractive index distribution in the foreground. The background reflector is a priori unknown, and its reflectivity is simultaneously reconstructed as part of the training process. This approach is therefore a promising candidate to reach quantitative in vivo optical imaging.

The ill-posed character of the problem is circumvented by a weighted temporal loss in order to focus on the low frequency part of the object and by using positivity constraint and sparsity regularization to converge to a good minimum. The resulting reconstructions closely match the ground truth across both low and high spatial frequencies, highlighting the effectiveness of the weighted temporal loss.

We implemented a 2D and 3D version to demonstrate its effectiveness in simulations. To remain consistent with the theory, we restricted our study to weakly scattering media and used relatively simple forward models compared to transmission ODT [52]. However, in practice, this method could involve more advanced forward models for larger refractive index contrast as Multi-layer Born [10], modified Born series [53] or Lippmann-Schwinger Model [12]. Whether it can be adapted to a more strongly scattering regime still needs to be demonstrated, either in simulation or experimentally.

We have also demonstrated the necessity of multi-wavelength analysis in reflection to achieve the required temporal approach. As a result, ODT in reflection will be more time-expensive than transmission approaches but will be rewarding since we will have access to both the low and high spatial frequencies of the sample. This also helps overcome the missing cone issue inherent to transmission configurations.

Beyond optical microscopy, the proposed approach can be leveraged to map the local wave speed in any field of wave physics for which a frequency- or time-resolved measurement of the reflection matrix is possible [54], [55]. Indeed, not only is the local wave speed a quantitative marker for ultrasound diagnosis [56] but it is also, for instance, an important monitoring parameter of tectonic motion [57] and volcanic eruption [58] in geophysics. In each of this field, multiple scattering is far from being negligible [59], [60] and its inclusion in the forward model would be extremely rewarding for quantitative imaging. This is the principle of full wave-form inversion [61] but we here show how such brute force strategies would benefit from physically-grounded loss functions and simplifications of the forward model in order to guarantee the convergence of the inversion process.

Appendix A
Filtering Operation on Source Term

This section is a proof of the filtering effect of measurement in transmission and reflection configuration. The proof is the same as in *Emil Wolf* [1] work, but without the first order Born hypothesis.

We express the scattered field as a function of the source term, as shown in (3).

Since the object to be imaged is of finite size, we assume that V is supported in $[z_1, z_3]$ and that we measure the scattered field for $z \in (-\infty, z_1) \cup (z_3, +\infty)$.

Since we are measuring E_s at a fixed z, it is natural to look for calculations based on the 2D Fourier transform of the 3D Green's function. Thus we have [62]:

$$\iint_{\mathbb{R}^2} G(x,y,z) e^{-j(k_x x + k_y y)} dx dy = \frac{j}{2\kappa(k_x,k_y)} e^{j\kappa(k_x,k_y)|z|}$$

$$G(x,y,z) = \iint_{\mathbb{R}^2} \frac{j}{2\kappa(k_x,k_y)} e^{j(k_x x + k_y y + \kappa(k_x,k_y)|z|)} dk_x dk_y \tag{31}$$

with κ defined as:

$$\kappa(k_x,k_y) = \begin{cases} \sqrt{n_\mathrm{b}^2 k_0^2 - k_x^2 - k_y^2} & \text{if } n_\mathrm{b}^2 k_0^2 \geq k_x^2 + k_y^2, \\ j\sqrt{k_x^2 + k_y^2 - n_\mathrm{b}^2 k_0^2} & \text{if } n_\mathrm{b}^2 k_0^2 < k_x^2 + k_y^2. \end{cases}$$

κ is the z-wavelength of the Green function with two modes: a propagative mode and an evanescent one.

In the following, we rewrite (31) taking into account the fact that $z \in (-\infty, z_1) \cup (z_3, +\infty)$ and $z' \in \mathrm{supp}(V)$

$$G(x-x', y-y', z-z') = \iint_{\mathbb{R}^2} \frac{j}{2\kappa(k_x,k_y)} \times e^{j(k_x(x-x') + k_y(y-y') + \mathrm{sgn}(z)\kappa(k_x,k_y)(z-z'))} dk_x dk_y \tag{32}$$

By injecting (32) into (3) we obtain:

$$E_\mathrm{s}(x,y,z) = \iint_{\mathbb{R}^2} A(k_x,k_y,z) \times e^{j(k_x x + k_y y + \mathrm{sgn}(z)\kappa(k_x,k_y)z)} dk_x dk_y \tag{33}$$

$$A(k_x,k_y,z) = \frac{j}{2\kappa(k_x,k_y)} \iiint_{\mathbb{R}^3} U(x',y',z') \times e^{-j(k_x x' + k_y y' + \mathrm{sgn}(z)\kappa(k_x,k_y)z')} dx' dy' dz' \tag{34}$$

By developing (33), we can express A by the 2D Fourier transform of $E_\mathrm{s}(.,z)$

$$\mathcal{F}_{2D}(E_\mathrm{s}(\cdot,z))(k_x,k_y) e^{-\mathrm{sgn}(z) j\kappa(k_x,k_y)z} = A(k_x,k_y,z) \tag{35}$$

The equation is difficult to express in simple terms, because κ is a complex number. However, we assume that we are measuring with z far enough away from the support of V that the evanescent waves are negligible (i.e. $A(k_x,k_y,k_z) \approx 0$ if $n_\mathrm{b}^2 k_0^2 < k_x^2 + k_y^2$). Consequently, κ is real, and we can express A directly as the Fourier transform of U:

$$A(k_x,k_y,z) = \frac{j\tilde{U}(k_x,k_y,\mathrm{sgn}(z)\sqrt{n_\mathrm{b}^2 k_0^2 - k_x^2 - k_y^2})}{2\sqrt{n_\mathrm{b}^2 k_0^2 - k_x^2 - k_y^2}} \tag{36}$$

And by combining (35) and (36) we obtain (4).

We finally found that we can only measure the Ewald sphere of the source U as shown in [8] under the Fraunhofer approximation.

Appendix B
Filtering Operation Under Second Born Approximation

Let us begin by examining the U_{12} term, which represents an interaction with the reflector V_2 followed by an interaction with the potential V_1. We can express U_{12} as:

$$\begin{aligned} U_{12}(x,y,z) &= V_1(x,y,z) H_2(x,y,z) \\ H_2(x,y,z) &= \iiint_{\mathbb{R}^3} G(x-x', y-y', z-z') \\ &\quad \times V_2(x',y',z') E_\mathrm{in}(x',y',z')\, dx' dy' dz' \\ &= \iint_{\mathbb{R}^2} G(x-x', y-y', z) \\ &\quad \times v_2(x',y') E_\mathrm{in}(x',y',0)\, dx' dy' \end{aligned} \tag{37}$$

H_2 corresponds to a new incident wave from the reflector V_2.

By injecting the 2D Fourier transform of the Green function [62] (31), assuming that z is small enough to neglect evanescent waves (since $\mathrm{supp}(V_1) = [z_1, z_2]$ we assume here that z_2 is far enough from 0) and by using convolution theorem, we can express H_2 in a 2D integral form:

$$\mathcal{F}_{2D}(H_2(.,z)) = \frac{j e^{j\sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2}|z|}}{2\sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2}} \tilde{v_2}(k_x - k_{\text{in},x}, k_y - k_{\text{in},y})$$

$$H_2(x,y,z) = \iint_{\mathbb{R}^2} \frac{j\,\tilde{v_2}(k_x - k_{\text{in},x}, k_y - k_{\text{in},y})}{2\sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2}} e^{j(k_x x + k_y y)} \times e^{j\sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2}|z|}\, dk_x\, dk_y \quad (38)$$

By applying a Fourier transform to (37), substituting (38), and simplifying the absolute value due to the support of $V_1(.,z')$ being restricted to $(-\infty, 0)$, we obtain (11).

The second term U_{21} corresponds to the interaction with V_1 followed by the interaction with the reflector V_2.

$$U_{21}(x,y,z) = V_2(x,y,z) H_1(x,y,z)$$

$$H_1(x,y,z) = \iiint_{\mathbb{R}^3} G(x-x', y-y', z-z') \times V_1(x',y',z') E_{\text{in}}(x',y',z')\, dx'dy'dz' \quad (39)$$

By applying a Fourier transform to U_{21} we have:

$$\begin{aligned}\tilde{U}_{21}(k_x,k_y,k_z) &= \iiint_{\mathbb{R}^3} V_2(x,y,z) H_1(x,y,z) \times e^{j(k_x x + k_y y + k_z z)} dxdydz \\ &= \iint_{\mathbb{R}^2} v_2(x,y) H_1(x,y,0) e^{j(k_x x + k_y y)} dxdy \\ &= \iint_{\mathbb{R}^2} \tilde{v_2}(k_x - k_x', k_y - k_y') \times \mathcal{F}_{2D}(H_1(.,z=0))(k_x', k_y') dk_x' dk_y'\end{aligned}$$

In order to get an analytical expression for the Fourier transform U_{12}, we need to express the 2D Fourier transform of H_1:

$$\begin{aligned}&\mathcal{F}_{2D}(H_1(.,z=0))(k_x,k_y) \\ &= \iint_{\mathbb{R}^2} H_1(x,y,0) e^{-j(k_x x + k_y y)} dxdy \\ &= \iiint_{\mathbb{R}^3} \iint_{\mathbb{R}^2} G(x-x', y-y', -z') V_1(x',y',z') \times e^{j(k_{\text{in},x} x' + k_{\text{in},y} y' + k_{\text{in},z} z') - j(k_x x + k_y y)} dxdydx'dy'dz' \\ &= \iiint_{\mathbb{R}^3} \iint_{\mathbb{R}^2} G(x'', y'', -z') e^{-j(k_x x'' + k_y y'')} dx''dy'' \times V_1(x',y',z') e^{j[(k_{\text{in},x} - k_x)x' + (k_{\text{in},y} - k_y)y' + k_{\text{in},z} z')} dx'dy'dz' \\ &= \int_{\mathbb{R}} dz' \left[\iint_{\mathbb{R}^2} G(x'', y'', -z') e^{-j(k_x x'' + k_y y'')} dx''dy''\right] \times \left[\iint_{\mathbb{R}^2} V_1(x',y',z') e^{j[(k_{\text{in},x} - k_x)x' + (k_{\text{in},y} - k_y)y' + k_{\text{in},z} z']} dx'dy'\right]\end{aligned}$$

The first term between square brackets is the 2D Fourier transform of the Green function and defined in (31). The second term between square brackets is:

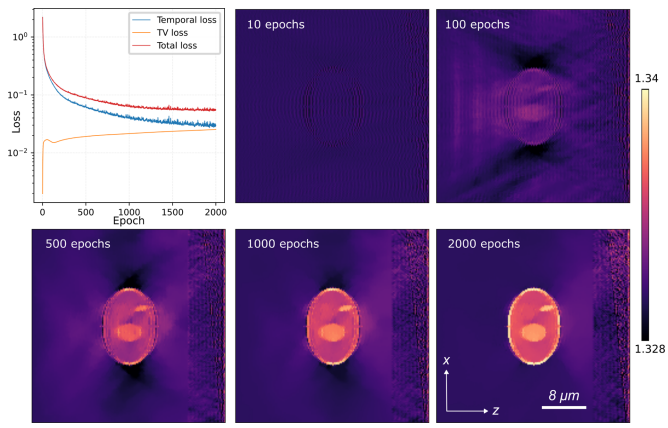


Fig. 9. Analysis of the convergence process. During the first epochs, only the high-frequency components are reconstructed, while the low-frequency information of the object gradually appears around the 500th epoch. The TV regularization parameter is small to allow the recovery of as much information as possible before the regularization term becomes dominant, leading to a better reconstruction.

$$\iint_{\mathbb{R}^2} V_1(x',y',z') e^{j[(k_{\text{in},x} - k_x)x' + (k_{\text{in},y} - k_y)y' + k_{\text{in},z} z']} dx'dy' = \mathcal{F}_{2D}(V_1(.,z'))(k_x - k_{\text{in},x}, k_y - k_{\text{in},y}) e^{j k_{\text{in},z} z'}$$

Since the support of $V_1(.,z')$ is in $(-\infty, 0)$, the integral in z' is concentrated on $(-\infty, 0)$ and:

$$\mathcal{F}_{2D}(H_1(.,z=0))(k_x,k_y) = \int_{\mathbb{R}} \frac{j\mathcal{F}_{2D}(V_1(.,z'))(k_x - k_{\text{in},x}, k_y - k_{\text{in},y})}{2\kappa(k_x,k_y)} e^{j(k_{\text{in},z} - \kappa(k_x,k_y))z'} dz'$$

which gives (by ignoring evanescent waves):

$$\mathcal{F}_{2D}(H_1(.,z=0))(k_x,k_y) = \frac{j}{2\sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2}} \times \tilde{V_1}(k_x - k_{\text{in},x}, k_y - k_{\text{in},y}, \sqrt{n_b^2 k_0^2 - k_x^2 - k_y^2} - k_{\text{in},z}) \quad (40)$$

Finally, we obtain (12) for U_{21}.

Appendix C
Noise Analysis, Selection of Hyperparameters and Regularization

Since we had to solve an ill-posed inverse problem, longer training times were necessary. To ensure better convergence for method comparison, we fine-tuned the learning rate and trained until 2000 epochs for Figs. 4–7. However, the reconstruction quality becomes acceptable around 500 epochs, even though the loss continues to decrease. An illustration of the convergence is presented Fig. 9.

The optimal learning rate was generally found to be around $\alpha = 10^{-5}$. For adaptive methods, the coefficients were set as recommended in Kingma's et al. work [44] $(\beta_1, \beta_2) = (0.9, 0.999)$. The weight factor was set to $\gamma = 0.1$ for all simulations and kept unchanged across different cases, as this ensured consistently good performance, but could be fine tuned depending on the medium to be reconstructed. The principle is that one tries

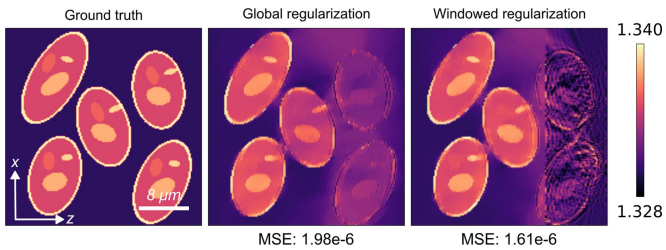


Fig. 10. Reconstruction of a multi-cell medium under different regularization strategies. The comparison is made between a reconstruction with regularization applied to the entire medium and applied only to the foreground region. The Mean Squared Error (MSE) of the reconstruction inside the foreground cells is displayed.

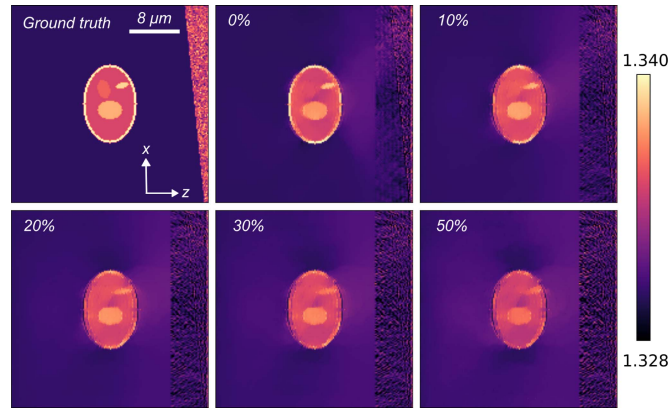


Fig. 11. Evolution of reconstruction as a function of the level of Gaussian noise added to the simulated measurements. Despite the loss of precision, the 2D reconstruction of the foreground cell remains consistent with the ground truth. Reconstructions are performed with weighted temporal loss ($\gamma = 0.1$) with *ProxGen* algorithm and minimum regularization.

to balance the contributions of the direct reflection from the foreground object (that contains the high spatial frequencies on the foreground object) and the cross-interaction terms (that contains the low spatial frequencies on the foreground object). Decreasing γ reduces the weight of the direct-reflection component, which corresponds to information that is reconstructed rapidly during optimization (see Fig. 9), thereby helping the optimization focus on the more challenging cross-interaction terms. Keeping $\gamma > 0$ is generally preferable, since the high-frequency information helps mitigate the missing-cone problem by adding complementary data. TV regularization and minimum regularization are not applied on the reflector because it allows better reconstruction of foreground objects as shown in Fig. 10.

The reconstruction quality is slightly better when no regularization is applied to the reflector, as this avoids constraining on it. In this case, the reconstructed RI of the reflector corresponds to an effective value that produces better reflectivity, rather than to its true physical index. Because the reflector's low-frequency components are difficult to retrieve, this effective value mainly acts as a local minimum that mimics the reflective behavior. Adding regularization, in this case, limits the reflector without bringing useful physical information about it.

Despite the use of the weighted temporal loss to focus on the low-frequency signal, the optimization problem is not straightforward and requires more training time than in transmission. Thus our optimization algorithm could be sensitive to noisy measurement. To verify this, we conducted 2D tests by adding noise to the ground truth measurements. As shown in Fig. 11, our model remains robust to additive Gaussian noise up to 50%. The noise level is here defined relatively to the mean of the reflected absolute amplitude.

Appendix D
Spectral Study of the Reconstruction

In order to validate the importance of cross-terms U_{12} and U_{12} described in (12) and (11), we display in Fig. 12 the spectral representation of the reconstructed potential V.

As expected, using the 5th-order Born model allows us to recover both the high and low-frequency components of the object in the foreground (Fig. 12(a), (b)), which is consistent with the combination of multi-illumination transmission and reflection terms described in Coupland et al.'s work [8].

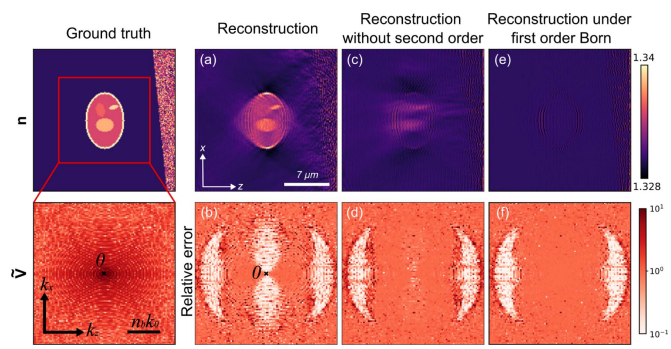


Fig. 12. RI reconstruction and relative error of the reconstruction $\tilde{V}$ in Fourier space. The reconstruction are performed without regularization, with Adam optimizer and the temporal loss. (a,b) The ground truth is generated using the Lippmann–Schwinger model and the forward model used for reconstruction is the 5th-order Born model. (c,d) The second-order (double scattering) term, including the cross terms U_{12} and U_{21}, is removed from both the ground truth and the forward model. (e,f) Both the ground truth and the forward model are computed in the single-scattering regime using the first-order Born approximation.

When the second-order term is removed (Fig. 12(c), (d)), the reconstruction quality decreases since the low-frequency information encoded in the cross terms disappears. Some of these components can still be partially recovered through higher-order terms, although their contribution remains limited, justifying the main contribution of second-order term in reconstruction. Finally, removing all multiple-scattering terms (Fig. 12(e), (f)) makes it impossible to retrieve the low-frequency information, even with a reflector in the background.

Moreover, Fig. 12 provides an estimate of the maximum resolution, as the reconstruction can only recover information within a finite region of the object's spatial Fourier space. Following Coupland et al. [8], the lateral frequency range is bounded by $dk_x = |4n_b k_0 \text{NA}|$, while the axial range (corresponding to direct reflection information) extends over $dk_z = |4n_b k_0|$. For our microscopy configuration with incident wavelengths within $[800\,\text{nm}, 875\,\text{nm}]$, the maximum spatial resolution is about $150\,\text{nm}$. The figure shown here corresponds to a $140\,\text{nm}$

resolution in order to clearly highlight the information which can be recovered in Fourier space.

REFERENCES

[1] E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Opt. Commun.*, vol. 1, no. 4, pp. 153–156, 1969.

[2] W. Choi et al., "Tomographic phase microscopy," *Nature Methods*, vol. 4, no. 9, pp. 717–719, 2007.

[3] D. Jin, R. Zhou, Z. Yaqoob, and P. T. So, "Tomographic phase microscopy: Principles and applications in bioimaging," *JOSA B*, vol. 34, no. 5, pp. B64–B77, 2017.

[4] P. Y. Liu et al., "Cell refractive index for cell biology and disease diagnosis: Past, present and future," *Lab Chip*, vol. 16, no. 4, pp. 634–644, 2016.

[5] Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nature Photon.*, vol. 12, no. 10, pp. 578–589, 2018.

[6] M. K. Kim, "Principles and techniques of digital holographic microscopy," *SPIE Rev.*, vol. 1, no. 1, 2010, Art. no. 018005.

[7] A. Devaney, "Inverse-scattering theory within the Rytov approximation," *Opt. Lett.*, vol. 6, no. 8, pp. 374–376, 1981.

[8] J. M. Coupland and J. Lobera, "Holography, tomography and 3D microscopy as linear filtering operations," *Meas. Sci. Technol.*, vol. 19, no. 7, 2008, Art. no. 074012.

[9] X. Ma, W. Xiao, and F. Pan, "Optical tomographic reconstruction based on multi-slice wave propagation method," *Opt. Exp.*, vol. 25, no. 19, pp. 22595–22607, 2017.

[10] M. Chen, D. Ren, H.-Y. Liu, S. Chowdhury, and L. Waller, "Multi-layer born multiple-scattering model for 3D phase microscopy," *Optica*, vol. 7, no. 5, pp. 394–403, 2020.

[11] M. Lee, H. Hugonnet, and Y. Park, "Inverse problem solver for multiple light scattering using modified born series," *Optica*, vol. 9, no. 2, pp. 177–182, 2022.

[12] T.-a. Pham, E. Soubies, A. Ayoub, J. Lim, D. Psaltis, and M. Unser, "Three-dimensional optical diffraction tomography with Lippmann-Schwinger model," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 727–738, 2020.

[13] U. S. Kamilov, H. Mansour, and B. Wohlberg, "A plug-and-play priors approach for solving nonlinear imaging inverse problems," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1872–1876, Dec. 2017.

[14] R. Weissleder, "A clearer vision for in vivo imaging," *Nature Biotechnol.*, vol. 19, no. 4, pp. 316–317, 2001.

[15] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[16] S. Kang, R. Zhou, M. Brelen, H. K. Mak, P. T. So, and Z. Yaqoob, "Reflection-mode optical diffraction tomography for label-free imaging of thick biological specimens," 2022, *arXiv:2202.13668*.

[17] T. Li, J. Zhu, Y. Shen, and L. Tian, "Reflection-mode diffraction tomography of multiple-scattering samples on a reflective substrate from intensity images," *Optica*, vol. 12, no. 3, pp. 406–417, Mar. 2025. [Online]. Available: https://opg.optica.org/optica/abstract.cfm?URI=optica-12-3-406

[18] Y.-R. Lee, D.-Y. Kim, Y. Jo, M. Kim, and W. Choi, "Exploiting volumetric wave correlation for enhanced depth imaging in scattering medium," *Nature Commun.*, vol. 14, no. 1, Apr. 2023, Art. no. 1878. [Online]. Available: https://www.nature.com/articles/s41467-023-37467-z

[19] Y. Zhang, D. Minh, Z. Wang, T. Zhang, T. Chen, and C. W. Hsu, "Deep imaging inside scattering media through virtual spatiotemporal wavefrontshaping," 2023, *arXiv:2306.08793*.

[20] P. Balondrade et al., "Multi-spectral reflection matrix for ultrafast 3D label-free microscopy," *Nature Photon.*, vol. 18, no. 10, pp. 1097–1104, 2024.

[21] U. S. Kamilov et al., "Optical tomographic image reconstruction based on beam propagation and sparse regularization," *IEEE Trans. Comput. Imag.*, vol. 2, no. 1, pp. 59–70, Mar. 2016.

[22] S. Kang et al., "Tracing multiple scattering trajectories for deep optical imaging in scattering media," *Nature Commun.*, vol. 14, no. 1, 2023, Art. no. 6871.

[23] T. D. Weber, N. Khetan, R. Yang, and J. Mertz, "Ultrasound differential phase contrast using backscattering and the memory effect," *Appl. Phys. Lett.*, vol. 118, no. 12, Mar. 2021, Art. no. 124103.

[24] P. Stähli, M. Frenz, and M. Jaeger, "Bayesian approach for a robust speed-of-sound reconstruction using pulse-echo ultrasound," *IEEE Trans. Med. Imag.*, vol. 40, no. 2, pp. 457–467, Feb. 2021.

[25] P. Mora, "Inversion=Migration+tomography," in *Proc. Shell Conf., Parallel Comput.*, 1989, pp. 78–101.

[26] A. E. Malcolm, B. Ursin, and M. V. De Hoop, "Seismic imaging and illumination with internal multiples," *Geophysical J. Int.*, vol. 176, no. 3, pp. 847–864, 2009.

[27] Y. Wang and Y. Rao, "Reflection seismic waveform tomography," *J. Geophysical Res., Solid Earth*, vol. 114, no. B3, 2009, Art. no. B03304, doi: 10.1029/2008JB005916.

[28] G. Taraldsen, "The complex image method," *Wave Motion*, vol. 43, no. 1, pp. 91–97, 2005.

[29] E. Soubies, T.-A. Pham, and M. Unser, "Efficient inversion of multiple-scattering model for optical diffraction tomography," *Opt. Exp.*, vol. 25, no. 18, pp. 21786–21800, 2017.

[30] F. Vico, L. Greengard, and M. Ferrando, "Fast convolution with free-space green's functions," *J. Comput. Phys.*, vol. 323, pp. 191–203, 2016.

[31] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009, doi: 10.1137/080716542.

[32] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D, Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016727899290242F

[33] J. Yun, A. C. Lozano, and E. Yang, "Adaptive proximal gradient methods for structured neural networksx," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 24365–24378. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2021/file/cc3f5463bc4d26bc38eadc8bcffbc654-Paper.pdf

[34] A. Badon, G. Lerosey, A. C. Boccara, M. Fink, and A. Aubry, "Retrieving time-dependent green's functions in optics with low-coherence interferometry," *Phys. Rev. Lett.*, vol. 114, no. 2, Jan. 2015, Art. no. 023901.

[35] M. Choma, M. Sarunic, C. Yang, and J. Izatt, "Sensitivity advantage of swept source and fourier domain optical coherence tomography," *Opt. Exp.*, vol. 11, no. 18, pp. 2183–2189, Sep. 2003.

[36] S. H. Yun, G. J. Tearney, B. E. Bouma, B. H. Park, and J. F. de Boer, "High-speed spectral-domain optical coherence tomography at 1.3 μm wavelength," *Opt. Exp.*, vol. 11, no. 26, pp. 3598–3604, Dec. 2003. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-11-26-3598

[37] D. Hillmann, G. Franke, C. Lührs, P. Koch, and G. Hüttmann, "Efficient holoscopy image reconstruction," *Opt. Exp.*, vol. 20, no. 19, pp. 21247–21263, Sep. 2012. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-20-19-21247

[38] B. Považay, A. Unterhuber, B. Hermann, H. Sattmann, H. Arthaber, and W. Drexler, "Full-field time-encoded frequency-domain optical coherence tomography," *Opt. Exp.*, vol. 14, no. 17, pp. 7661–7669, Aug. 2006. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-14-17-7661

[39] A. Derode, P. Roux, and M. Fink, "Robust acoustic time reversal with high-order multiple scattering," *Phys. Rev. Lett.*, vol. 75, no. 23, pp. 4206–4209, Dec. 1995.

[40] A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," *Nature Photon.*, vol. 6, no. 5, pp. 283–292, May 2012.

[41] R. K. Wang, "Signal degradation by multiple scattering in optical coherence tomography of dense tissue: A Monte Carlo study towards optical clearing of biotissues," *Phys. Med. Biol.*, vol. 47, no. 13, pp. 2281–2299, Jun. 2002, doi: 10.1088/0031-9155/47/13/307.

[42] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, 2014, doi: 10.1561/2400000003.

[43] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, 2015. [Online]. Available: https://arxiv.org/abs/1412.6980

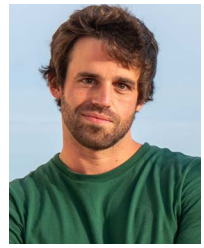
[45] S. De, A. Mukherjee, and E. Ullah, "Convergence guarantees for RMSProp and adam in non-convex optimization and an empirical comparison to nesterov acceleration," 2018, *arXiv:1807.06766*.

[46] M. Zaheer, S. J. Reddi, D. Sachan, S. Kale, and S. Kumar, "Adaptive methods for nonconvex optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 9815–9825. [Online]. Available: https://papers.nips.cc/paper/2018/hash/8186-adaptive-methods-for-nonconvex-optimization-8186.pdf

[47] P. Melchior, R. Joseph, and F. Moolekamp, "Proximal adam: Robust adaptive update scheme for constrained optimization," 2020, *arXiv:1910.10094*.

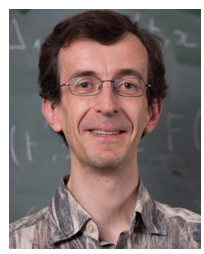
[48] S. Chowdhury et al., "High-resolution 3D refractive index microscopy of multiple-scattering samples from intensity images," *Optica*, vol. 6, no. 9, pp. 1211–1219, Sep. 2019. [Online]. Available: https://opg.optica.org/optica/abstract.cfm?URI=optica-6-9-1211
[49] A. B. Ayoub, A. Roy, and D. Psaltis, "Optical diffraction tomography using nearly in-line holography with a broadband led source," *Appl. Sci.*, vol. 12, no. 3, 2022, Art. no. 951. [Online]. Available: https://www.mdpi.com/2076-3417/12/3/951
[50] J. M. Soto, J. A. Rodrigo, and T. Alieva, "Partially coherent optical diffraction tomography toward practical cell study," *Front. Phys.*, vol 9, 2021, Art. no. 666256. [Online]. Available: https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2021.666256
[51] A. Paszke et al., "Automatic differentiation in PyTorch," in *Proc. NIPS Workshop Automat. Differentiation*, 2017. [Online]. Available: https://openreview.net/pdf?id=BJJsrmfCZ
[52] U. S. Kamilov, D. Liu, H. Mansour, and P. T. Boufounos, "A recursive born approach to nonlinear inverse scattering," *IEEE Signal Process. Lett.*, vol. 23, no. 8, pp. 1052–1056, Aug. 2016.
[53] G. Osnabrugge, S. Leedumrongwatthanakun, and I. M. Vellekoop, "A convergent born series for solving the inhomogeneous helmholtz equation in arbitrarily large media," *J. Comput. Phys.*, vol. 322, pp. 113–124, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999116302595
[54] F. Bureau et al., "Three-dimensional ultrasound matrix imaging," *Nature Commun.*, vol. 14, 2024, Art. no. 6793.
[55] E. Giraudat, A. Burtin, A. L. Ber, M. Fink, J.-C. Komorowski, and A. Aubry, "Matrix imaging as a tool for high-resolution monitoring of deep volcanic plumbing systems with seismic noise," *Commun. Earth Environ.*, vol. 5, 2024, Art. no. 509.
[56] P. Stähli et al., "First-in-human diagnostic study of hepatic steatosis with computed ultrasound tomography in echo mode," *Commun. Med.*, vol. 3, no. 1, Dec. 2023, Art. no. 176.
[57] S. Solarino et al., "Seismic tomography reveals contrasting styles of subduction-channel and mantle-wedge exhumation controlled by upper plate divergent motion," *Gondwana Res.*, vol. 136, pp. 169–182, Dec. 2024.
[58] L. D'Auria et al., "Rapid magma ascent beneath la palma revealed by seismic tomography," *Sci. Rep.*, vol. 12, no. 1, Oct. 2022, Art. no. 17654.
[59] J. Chaput et al., "Multiple scattering from icequakes at erebus volcano, antarctica: Implications for imaging at glaciated volcanoes," *J. Geophysical Res., Solid Earth*, vol. 120, no. 2, pp. 1129–1141, Feb. 2015.
[60] A. Goïcoechea et al., "Reflection measurement of the scattering mean free path at the onset of multiple scattering," *Phys. Rev. Lett.*, vol. 133, no. 17, Oct. 2024, Art. no. 176301.
[61] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *Geophysics*, vol. 74, no. 6, pp. WCC1–WCC26, Nov. 2009.
[62] H. Weyl, "Ausbreitung elektromagnetischer wellen über einem ebenen leiter," *Annalen der Physik*, vol. 365, no. 21, pp. 481–500, 1919.

Thomas Wasik received the master's degree in applied mathematics from École Polytechnique, Palaiseau, France, in 2022, and the second master's degree in machine learning from Institut Polytechnique de Paris, Palaiseau, in 2023. He has been working toward the Ph.D. degree with the Department of Applied Mathematics, École Polytechnique, in collaboration with the CNRS/Langevin Institute in Paris, since 2024. His research interests include computational imaging, inverse problems, machine learning, deep learning, and optimization.



Victor Barolle studied at the École Polytechnique, Palaiseau, France. He received the Ph.D. degree in physics from Institut Langevin, Paris, France, in 2019. Under the supervision of Alexandre Aubry, he laid the foundation of a matrix approach for optical coherence tomography. He then joined Greenerwave, a spin-off of Institut Langevin, where he developed innovative wave control solutions for satellite communications. He is currently the Chief Science Officer and Co-Founder of OWLO, a company developing a real-time, non-invasive 3D microscope with applications ranging from embryology to ophthalmology and cell culture imaging.

Alexandre Aubry studied at the Ecole Supérieure de Physique et Chimie Industrielles in Paris. He received the Ph.D. degree from Université Pierre et Marie Curie, Paris, France, in 2008. He spent a postdoctoral period with Imperial College London, before joining the Langevin Institute in Paris as a CNRS Research Fellow in 2012. Since 2012, he has been developing a matrix approach to wave control through complex media. In 2019, he was awarded an ERC Consolidator grant, which enabled him to develop a universal approach to imaging complex media, with applications ranging from ultrasound to seismology and optical microscopy. He has authored more than sixty-five papers in these areas and twenty-five patents. In 2024, he co-founded the OWLO Company dedicated to the applications of matrix imaging to optical microscopy. In 2025, he was the recipient of the Jean Jerphagnon Prize from the Société Française d'Optique.



Josselin Garnier he has been a Professor of applied mathematics, since 2001, first in Toulouse, then in Paris with Université Paris Diderot and Ecole Normale Supérieure, and with Ecole Polytechnique, since 2016. He has been the Head of the Department of Applied Mathematics, Ecole Polytechnique, since 2023. He has authored 300 papers in his research interests which include the interface of probability theory and analysis, concerns the modeling and study of random phenomena, in particular wave propagation and imaging in complex media, the quantification of uncertainties in numerical simulation, and stochastic algorithms. areas and has written five research books and three textbooks. He was the recipient of the Blaise Pascal Prize in 2007 and the Felix Klein Prize in 2008. He was an Invited Speaker at the International Congress of Mathematicians in 2018. In 2023, he was elected a member of the French Academy of Sciences.