

Research Article

Space-Dependent Sobolev Gradients as a Regularization for Inverse Radiative Transfer Problems

Y. Favennec,¹ F. Dubot,² D. Le Hardy,¹ B. Rousseau,¹ and D. R. Rousse²

¹Laboratoire de Thermocinétique de Nantes (LTN), UMR CNRS 6607, Université de Nantes, rue C. Pauc, BP 50609, 44306 Nantes Cedex 3, France

²Chaire de Recherche Industrielle en Technologies de l'énergie et en Efficacité Énergétique (T3E), Ecole de Technologie Supérieure, 1100 rue Notre-Dame Ouest, Montréal, Canada H3C 1K3

Correspondence should be addressed to Y. Favennec; yann.favennec@univ-nantes.fr

Received 12 January 2016; Accepted 7 March 2016

Academic Editor: Maria Gandarias

Copyright © 2016 Y. Favennec et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Diffuse optical tomography problems rely on the solution of an optimization problem for which the dimension of the parameter space is usually large. Thus, gradient-type optimizers are likely to be used, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, along with the adjoint-state method to compute the cost function gradient. Usually, the L_2 -inner product is chosen within the extraction procedure (i.e., in the definition of the relationship between the cost function gradient and the directional derivative of the cost function) while alternative inner products that act as regularization can be used. This paper presents some results based on space-dependent Sobolev inner products and shows that this method acts as an efficient low-pass filter on the cost function gradient. Numerical results indicate that the use of Sobolev gradients can be particularly attractive in the context of inverse problems, particularly because of the simplicity of this regularization, since a single additional diffusion equation is to be solved, and also because the quality of the solution is smoothly varying with respect to the regularization parameter.

1. Introduction

Diffuse optical tomography (DOT) consists in reconstructing the spatial radiative property distributions of an optically thick participating (absorbing and scattering) medium, \mathcal{D} , from optical measurements collected at the border of \mathcal{D} [1–3]. This inverse problem has been a surge of interest since the middle of the nineties, due to better theoretical understanding in radiative transfer; improvements in light detection and control of light sources; and the increase of the computing resource performances. This optical diagnostic technique is mainly driven by potential applications in biological tissues imaging although it can be used for nondestructive testing in industrial processes [4, 5] or the characterization of complex semitransparent media such as cellular foams with spatially varying porosities or fiber reinforced thermoplastic composites with inhomogeneity in the fibers distribution. In this domain, the characterization has often been limited to the reconstruction of constant absorption and scattering

coefficients and anisotropy factor of the phase function, considering the unidimensional radiative transfer equation (RTE) as forward model [6].

Biomedical applications of the DOT include imaging of finger joints [7], breast imaging [8, 9], neuroimaging of humans [10–12], or small-animal imaging [13] to name but a few. In this context, near-infrared light is used in order to take advantage of the so-called therapeutic window (600–900 nm) in which tissues have low absorption. According to the application, RTE has to be considered as forward model [14–19], diffuse approximation (DA) may be applied [8, 12, 20–23], or hybrid model is employed [24–28]. This paper deals with two very different forward models, namely, the bidimensional steady-state RTE and the three-dimensional frequency DA, in order to gauge the potential ability of the introduced regularization with the different models.

The evaluation of space-distributed physical properties from the knowledge of measurements is an inverse problem

as opposed to forward problems where the state is computed from the knowledge of properties, boundary conditions, and so forth. Such inverse problems are usually difficult to solve due to their ill-posed nature [29].

The solution of the inverse problem is obtained through the minimization of a cost function that measures the misfit between the predictions (solution of the forward problem) and some experimental measurements. Though zero-order optimization methods may be used for relatively similar problems [31, 32], gradient-type methods are dealt with in this paper because of the large dimension of the parameter space. Due to the ill-posed nature of the inverse problem, unstable solutions may be obtained if no regularization is performed.

In order to cope with such instabilities, the ordinary Tikhonov approach consists in assuming that the properties are subject to a minimal norm in a given vectorial space which leads to enhancing the *a priori* regularity of the properties [33, 34]. This approach relies on the penalization of the cost function adding prior information on the properties to be recovered. However, the weight associated with the penalization function must be chosen carefully. Indeed, a too small value for the penalization function leads to a small cost function but with high property variations, while too high values lead to small property variations about priors but with a high resulting cost function value at the optimum. The compromise minimizing both the cost function and the property variations may be performed according to the construction of the so-called L-curve as suggested by Hansen and O'Leary [35]. Even though, for appropriate penalization, the optimal Tikhonov parameter has been found to be quasi-independent of the parameter space dimension [36], the construction of such curve, searching its convex corner location, is likely to necessitate a large number of inverse problem solutions for plenty of penalization function weights. Furthermore, it is now well known that the Tikhonov regularization is to be used only together with matrix-inversion-based optimizers, but not with BFGS-like optimizers because these do not rely on matrix inversion [19, 23].

Appropriate parameterization for the space-dependent properties may also affect the solution regularity. As a matter of fact, the amount of measurements data is likely to be insufficient to retrieve the large amount of parameters when considering, at the end, the discrete optimization problem on a fine grid. Thus, regularization by reduction of the dimension of the parameter space could be employed, which consists in searching the physical properties projected on a coarse finite element grid, as suggested for instance in [37], and applied for optical tomography applications in [23].

This paper focuses on a different strategy to improve the reconstruction of space-distributed parameters when the data is corrupted by noise. The proposed strategy filters the effect of the noise, not on the data as it is done for instance when using the mollification method [29, 38], but on the cost function gradient itself. This means that the optimization is not modified at all since the original cost function remains the same. It also means that the data is

not modified: the whole information contained within the experimental data is integrated within the cost function to be minimized. As a matter of fact, several inner products are to be defined when considering the inverse optimization problem. According to Protas [39], each of them (cost function definition, adjoint identity, and gradient extraction after defining the space used in the definition relationship between the cost function derivative and its gradient) may somehow yield to a particular regularization opportunity. The noise present in the measurements data propagates through the adjoint problem and then to the cost function gradient from which the directions of descent are computed to supply the optimizer. The propagation of the noise induces high variations of the cost function gradient with respect to what the gradient would be without noise. The idea is thus to somehow apply a low-pass filter to the cost function gradient, exactly where the noise is the highest. This is done choosing the space-dependent weighted Sobolev inner product definition when extracting the cost function gradient rather than the ordinary Hilbert inner product. Note that space-independent Sobolev inner products have been used in other situations, in image treatment (e.g., [40]), in partial differential equations (PDEs), in energy minimization (e.g., [41]), and in boundary conditions reconstruction with accurate final time [42, 43].

The underlying application of this study is optical tomography, that is, the reconstruction of radiative properties from light intensity measurements on some parts of the boundary. The paper is therefore organized as follows. Section 2 gives the theoretical bases for extracting the cost function gradients when the cost function is expressed in terms of a state which is solution of some PDEs. This section is written in a concise way, following [44], such that the results can be easily transposed to other practical applications. Also, only the continuous equations are dealt with, for the states, the adjoints, and the gradients. This “differentiate-then-discretize” approach has the advantage of versatility when considering afterwards different discretization choices (one for the states, one for the physical properties, etc.). The cost function gradient is then extracted firstly in a general framework using the ordinary L_2 -inner product and then using space-independent and finally space-dependent Sobolev inner products. Section 3 presents application of space-dependent Sobolev gradients for the inverse problem of optical tomography in which the forward model is the RTE. The continuous cost function gradient is fully derived and numerical tests are performed for several levels of noise. As the Sobolev filter is chosen to follow an exponential function, only one parameter fully describes the regularization. Next, in Section 4, a very different model, namely, the frequency DA in the three-dimensional case, is considered as forward model of light propagation. For the two models, the quality of the reconstructions is measured quantitatively with the aid of both deviation and correlation factors. Without contest, reconstructions are greatly improved with the space-dependent Sobolev gradients. Finally, Section 5 naturally summarizes the main conclusions and perspectives of the current work.

2. Space-Dependent Sobolev Cost Function Gradient

Let us define the domain of interest $\mathcal{D} \subset \mathbb{R}^d$ and denote the cost function $\mathcal{F}(u)$ expressed in terms of the state $u : \mathcal{D} \rightarrow \mathbb{K}$. This cost function may be expressed as the half of the squared norm of a quadratic function v based on the discrepancy between some state predictions u and the related measurements u_d . The norm is expressed with a complex inner product $(\cdot, \cdot)_{\mathcal{X}}$ defined later on for specific applications. Cost function $\mathcal{F}(u)$ is actually to be minimized with respect to some space-distributed parameters $\psi \in \mathcal{P}$, with $\mathcal{P} \subseteq L_2(\mathcal{D})$, for example (depending on the forward model considered). The generic cost function is given by

$$j(\psi) := \mathcal{F}(u) = \frac{1}{2} \|f(u - u_d)\|_{\mathcal{X}}^2. \quad (1)$$

The state $u(\psi)$ is related to the parameters ψ through an operator (that is likely to be nonlinear) that combines the partial differential equations along with the boundary conditions, initial condition, and so forth. This operator is denoted as \mathcal{S} for the state problem. To be concise, one writes down

$$\mathcal{S}(u, \psi) = 0. \quad (2)$$

In order to use efficient optimization algorithms for large-scale inverse problems encountered in distributed parameter estimation, the cost function derivative or rather the cost function gradient has to be computed. To do so, the definition of the directional derivative of $j(\psi)$ towards the direction ϕ is used.

Definition 1 (directional derivative). Let a point ψ and a direction $\phi \in \mathcal{P}$. The directional derivative of j at point ψ in the direction ϕ is [45]

$$j'(\psi; \phi) := \lim_{t \rightarrow 0^+} \frac{j(\psi + t\phi) - j(\psi)}{t}. \quad (3)$$

The application of the directional derivative (3) on the cost function (1) leads to the equality

$$j'(\psi; \phi) = (f(u - u_d), f'(u) u'(\psi; \phi))_{\mathcal{X}} \quad \forall \phi \in \mathcal{P} \quad (4)$$

and, because of the linearity of $j'(\psi; \phi)$ with respect to ϕ ,

$$j'(\psi; \phi) = (\nabla j(\psi), \phi)_{\mathcal{X}} \quad \forall \phi \in \mathcal{P}. \quad (5)$$

The inner product $(\cdot, \cdot)_{\mathcal{X}}$ involved in the right-hand side of (5) will be also defined later on, when necessary. That is, in this inner product, regularization should be introduced.

The differentiation of the state problem in the direction ϕ yields the directional derivative $u'(\psi; \phi)$ that can further be injected into the inner product (4). The state equation being given by (2), the derivative of $S(u, \psi)$ towards the direction ϕ is written as

$$\mathcal{S}'_u(u, \psi) u'(\psi; \phi) + \mathcal{S}'_\psi(u, \psi) \phi = 0 \quad \forall \phi \in \mathcal{P}, \quad (6)$$

where \mathcal{S}'_u is the tangent operator, that is, the Jacobian of \mathcal{S} with respect to u , and \mathcal{S}'_ψ is the Jacobian of \mathcal{S} with respect to ψ . Using this basic approach, one needs $\dim \mathcal{P}$ integrations of (6) to access the full gradient $\nabla j(\psi) \in \mathbb{R}^{\dim \mathcal{P}}$ for a fully discretized problem. When $\dim \mathcal{P}$ is large, which is likely to be the case when recovering space-distributed parameters, this approach becomes inefficient even though the computation of the linear tangent operator $\mathcal{S}'_u(u, \psi)$ can be reused for all directions ϕ of the canonical basis of \mathcal{P} .

Following Lions [46] and C ea [47], the adjoint-state method computes the whole cost function gradient through the integration of a single additional adjoint-state problem. A new variable (the adjoint-state variable $u^*(\psi)$) is then introduced so that the cost function derivative also satisfies the relationship

$$j'(\psi; \phi) = (\mathcal{S}'_\psi(u, \psi) \phi, u^*)_{\mathcal{Y}} \quad \forall \phi \in \mathcal{P}, \quad (7)$$

where, again, the inner product $(\cdot, \cdot)_{\mathcal{Y}}$ should be defined later on when necessary. Combining (6) and (7), the cost derivative is rewritten as

$$j'(\psi; \phi) = -(\mathcal{S}'_u(u, \psi) u', u^*)_{\mathcal{Y}} \quad \forall \phi \in \mathcal{P}. \quad (8)$$

The adjoint-state problem is then identified through (4) and (8), such that it satisfies the equality

$$\begin{aligned} & -(\mathcal{S}'_u(u, \psi) u', u^*)_{\mathcal{Y}} \\ & = (f(u - u_d), f'(u) u'(\psi; \phi))_{\mathcal{X}} \quad \forall \phi \in \mathcal{P}. \end{aligned} \quad (9)$$

Next, if the adjoint problem (9) is satisfied (this means that the adjoint-state u^* is accessible), then the cost function gradient is given by the inner product (7). In practice, operators involved in (9) are transposed. The inner product property $(\mathcal{S}'_u(u, \psi) u', u^*)_{\mathcal{Y}} = (S^*(u, \psi) u^*, u')_{\mathcal{Y}}$ (with S^* being the transposed conjugate operator of \mathcal{S}'_u) gives the adjoint equation concisely written as

$$S^*(u, \psi) u^* + f'(u) f(u - u_d) = 0 \quad \forall \phi \in \mathcal{P}. \quad (10)$$

As reported by Protas [39], the gradient is identified in a given space for which the metric had been selected. To do so, the two following definitions are needed.

Definition 2 ($L_2(\mathcal{D})$ Hilbert space). Consider

$$L_2(\mathcal{D}) = \left\{ u : \mathcal{D} \rightarrow \mathbb{R} \mid \int_{\mathcal{D}} u(x) u(x) dx < \infty \right\}. \quad (11)$$

This functional space is associated with the inner product:

$$(u, w)_{L_2(\mathcal{D})} = \int_{\mathcal{D}} u(x) w(x) dx. \quad (12)$$

Definition 3 ($H^1(\mathcal{D})$ Sobolev space). Consider

$$\begin{aligned} & H^1(\mathcal{D}) \\ & := \{ \psi : \mathcal{D} \rightarrow \mathbb{R} \mid \psi \in L_2(\mathcal{D}); |\nabla \psi| \in L_2(\mathcal{D}) \}. \end{aligned} \quad (13)$$

This functional space is associated with the inner product:

$$(\psi, \phi)_{H^1(\mathcal{D})} := (\psi, \phi)_{L_2(\mathcal{D})} + (\nabla \psi, \nabla \phi)_{L_2(\mathcal{D})}. \quad (14)$$

Most often, the $L_2(\mathcal{D})$ real inner product for $(\cdot, \cdot)_{\mathcal{Z}}$ is used to extract the gradient. Combining (5) and (7) with the inner product of Definition 2 gives, after transposing within the right-hand side inner product, the following result:

$$(\nabla j(\psi), \phi)_{L_2(\mathcal{D})} = \left(\mathcal{S}'_{\psi}(u, \psi) u^*, \phi \right)_{\mathcal{Y}} \quad (15)$$

$$\forall \phi \in L_2(\mathcal{D}).$$

This way for writing down the cost function gradient is the ordinary one as first suggested by Lions [46] and C ea [47]. However, according to Protas [39], it is suggested to use other inner product definitions due to the poor scaling of the corresponding discrete optimization problem. As reported in [48], the incorporation of such derivatives into the inner product, instead of using the L_2 inner product, has the effect of scale-dependent filtering and allows one to extract smoother gradients, thereby preconditioning the optimization process. In some other studies, solving partial differential equations through optimization (see, e.g., the articles of Mahavier [49], Danaila and Kazemi [41], and Majid and Sial [50]), in image treatment (see, e.g., Renka [40]), in boundary conditions reconstruction (see, e.g., Nassiopoulou and Bourquin [42] and Bourquin and Nassiopoulou [43]), or in initial state reconstruction [39], the Sobolev space $H^1(\mathcal{D})$ is used for extracting the cost function gradient. This acts as a preconditioner modifying the direction of descent involved in the gradient-type optimization algorithm. It is shown, for instance, in the study of Danaila and Kazemi [41] that the cost function decrease when solving the PDEs problem can be much higher with the H^1 inner product than with the ordinary L_2 one.

In the present study, the choice of another inner product comes from the fact that we are faced with dealing with noisy measurements u_d . As a matter of fact, the measurements are noisy by nature. These measures are involved in the cost function definition (1), through a difference with the predictions. The noisy measurements are thus also involved in the forcing term $f'(u)f(u-u_d)$. The noise then propagates through the adjoint equation (10), and then to the cost function gradient (15).

In order to deal with the smoothing of the measurements noise that propagates in the adjoint system and then to the cost function gradient, the weighted Sobolev space is introduced.

Definition 4 ($H^{1(\ell)}(\mathcal{D})$ weighted Sobolev space inner product). Consider

$$(\psi, \phi)_{H^{1(\ell)}(\mathcal{D})} := (\psi, \phi)_{L_2(\mathcal{D})} + (\ell \nabla \psi, \ell \nabla \phi)_{L_2(\mathcal{D})}. \quad (16)$$

Remark 5. $\ell \in \mathbb{R}_+$ is the regularization parameter.

This strategy has been applied in bidimensional optical tomography applications based on the radiative transfer equation [19] where it was shown how the diffusion of the noise could attenuate fluctuations and thus give better reconstructions with a moderate parameter ℓ .

The strategy that is developed here goes much further: it consists in filtering the cost function gradient where, and only

where it is needed, the high frequency fluctuations exist, at the vicinity of the sensors.

In applications considered here, that is, in optical tomography, sensors are located on the boundaries, so that the effect of the noise appears on the boundaries before being diffused through the adjoint-state equation. Thus, the idea is to choose a filtering function whose value is high on the boundary and that continuously decreases within the medium. To do so, one uses the distance function definition.

Definition 6 (distance function). Let $\mathbf{x} \in \mathcal{D}$ and $\mathbf{y} \in \partial\mathcal{D}$; $d(\mathbf{x}, \partial\mathcal{D})$ is the Euclidean distance function to the set $\partial\mathcal{D}$ if

$$d(\mathbf{x}, \partial\mathcal{D}) = \min_{\mathbf{y} \in \partial\mathcal{D}} \|\mathbf{x} - \mathbf{y}\|. \quad (17)$$

Remark 7. As an example, the filtering function can be written as

$$\ell_v^2(\mathbf{x}) = \exp(-vd^2(\mathbf{x}, \partial\mathcal{D})). \quad (18)$$

This makes the construction of the space-dependent weighted Sobolev space possible.

Definition 8 ($H^{1(v)}(\mathcal{D})$ space-dependent weighted Sobolev space inner product). Consider

$$(\psi, \phi)_{H^{1(v)}(\mathcal{D})} := (\psi, \phi)_{L_2(\mathcal{D})} + (\ell_v(\mathbf{x}) \nabla \psi, \ell_v(\mathbf{x}) \nabla \phi)_{L_2(\mathcal{D})}. \quad (19)$$

Remark 9. With the choice of the filtering function (18), the regularization parameter becomes v .

Using the $H^{1(v)}(\mathcal{D})$ inner product for $(\cdot, \cdot)_{\mathcal{Z}}$ and if $\phi \in H^1(\mathcal{D})$, then there exists a unique $\nabla^{H^{1(v)}} j(\psi) \in H^{1(v)}(\mathcal{D})$ such that

$$j'(\psi; \phi) = \left(\nabla^{H^{1(v)}} j(\psi), \phi \right)_{H^{1(v)}(\mathcal{D})} \quad \forall \phi \in H^1(\mathcal{D}). \quad (20)$$

Taking into account that

$$\begin{aligned} & \left(\nabla \nabla^{H^{1(v)}} j(\psi), \nabla \phi \right)_{L_2(\mathcal{D})} \\ &= - \left(\Delta \nabla^{H^{1(v)}} j(\psi), \phi \right)_{L_2(\mathcal{D})} \\ &+ \left(\nabla \nabla^{H^{1(v)}} j(\psi) \cdot \mathbf{n}, \phi \right)_{L_2(\partial\mathcal{D})}, \end{aligned} \quad (21)$$

one obtains by identifying terms

$$\begin{aligned} \left((1 - \ell_v^2 \Delta) \nabla^{H^{1(v)}} j(\psi), \phi \right)_{L_2(\mathcal{D})} &= \left(\mathcal{S}'_{\psi}(u, \psi) u^*, \phi \right)_{\mathcal{Y}} \\ &\forall \phi \in H^1(\mathcal{D}) \end{aligned} \quad (22)$$

with a null flux boundary condition; that is, $\nabla \nabla^{H^{1(v)}} j(\psi) \cdot \mathbf{n} = 0$ on $\partial\mathcal{D}$. At this stage, the application being defined (measurements, state model, etc.), the \mathcal{Y} -norm is given accordingly, and the cost function gradient can be extracted.

Remark 10. Due to the inclusion $H^{1(v)}(\mathcal{D}) \subset L_2(\mathcal{D})$, the space-dependent weighted Sobolev gradient $\nabla^{H^{1(v)}} j(\psi)$ is indeed an acceptable direction of descent for the optimization problem that consists in minimizing (1).

3. Application on 2D RTE-Based Optical Tomography

3.1. Inverse Problem Statement. The inverse problem consists mathematically in minimizing a cost function which quantitatively measures the discrepancy between some optical measurements and related predictions over a set of sources, sensors, and angles of directional integration. In optical tomography, taking into account the reflection at the interface and assuming that the medium is convex, the measurable quantity is the exitance, or portion of it in a given solid angle $\omega(\mathbf{s}) \subseteq \{\mathbf{s} \cdot \mathbf{n} > 0\}$ centered on direction \mathbf{s} . The related prediction can be written as

$$E(\mathbf{x}, \mathbf{s}) = \int_{\omega(\mathbf{s})} (1 - R(\mathbf{x}, \mathbf{s})) I(\mathbf{x}, \mathbf{s}) \mathbf{s} \cdot \mathbf{n} d\omega(\mathbf{s}), \quad (23)$$

where I is the radiative intensity and the function $R(\mathbf{x}, \mathbf{s}) \in [0, 1]$ is the Fresnel reflection factor. In what follows in this section, the border of the medium is assumed to be transparent; that is, $R = 0$. Moreover, the solid angle $\omega(\mathbf{s})$ is assumed to be sufficiently small so that the approximation $E(\mathbf{x}, \mathbf{s}_r) \approx \omega_r I(\mathbf{x}, \mathbf{s}_r) \mathbf{s}_r \cdot \mathbf{n}$ holds. Therefore, the radiative intensity measurements $\check{I}(\mathbf{x}, \mathbf{s}_r)$ can be assessed and the cost function (24) mapping from $\mathcal{P} = L^\infty(\mathcal{D}) \times L^\infty(\mathcal{D})$ to \mathbb{R} uses, according to the schematic description of the problem (Figure 1), the specific Euclidean norm:

$$j(\kappa, \sigma) := \mathcal{F}(I) = \frac{1}{2} \left\| \frac{\mathbf{s} \cdot \mathbf{n}}{\check{I}} (I - \check{I}) \right\|_{\mathcal{X}}^2 \quad (24)$$

with $\|u\|_{\mathcal{X}} \triangleq \left(\sum_{k=1}^K \sum_{d=1}^D \sum_{r=1}^R u_k(\mathbf{x}_d, \mathbf{s}_r) \right)^{1/2}$,

where the index k refers to the source (test) number, the index d refers to the detection node, and the index r refers to the direction. In (24), $\mathcal{F}(I)$ is the cost function explicitly expressed in terms of the prediction I , while $j(\kappa, \sigma)$ is the cost function implicitly expressed in terms of the parameters. This latter function is often called the reduced cost function. According to (1), one has the straightforward correspondence $f(u - u_d) \leftrightarrow (\mathbf{s} \cdot \mathbf{n} / \check{I})(I - \check{I})$. Note that the scalar $\mathbf{s} \cdot \mathbf{n}$ in (24) is introduced in order to avoid infinite Dirichlet boundary condition in the adjoint-state problem. For a given source configuration, say the k th one, the prediction $I_k(\mathbf{x}_d, \mathbf{s}_r)$ at the

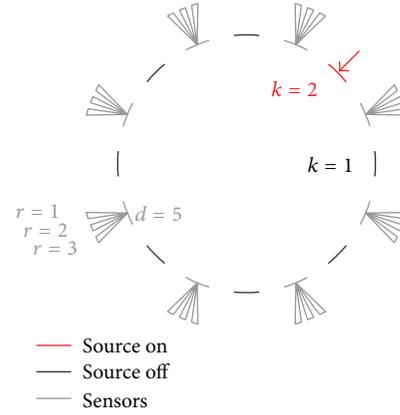


FIGURE 1: Schematic description of the experiment. Only one source is on at once, while others are off. For each source configuration number $k = 1, \dots, K$, the emerging intensity is measured on all sensors $d = 1, \dots, D$ in different directions $r = 1, \dots, R$. This is a theoretical extension of the experimental directional-directional spectral transmittances device using a Fourier transform infrared spectrometer described in [30].

sensor d and direction r is given through the solution of the forward steady-state radiative transfer problem formulated as

$$\begin{aligned} \text{Find } I_k(\mathbf{x}, \mathbf{s}) \quad \text{such that } \forall \mathbf{s} \in \mathcal{S}^{n-1}, \forall k = 1, \dots, K : \\ (\mathbf{s} \cdot \nabla + \kappa + \sigma) I_k(\mathbf{x}, \mathbf{s}) \\ = \sigma \oint_{\mathcal{S}^{n-1}} I_k(\mathbf{x}, \mathbf{s}') \Phi(\mathbf{s}, \mathbf{s}') d\omega(\mathbf{s}'), \quad \forall \mathbf{x} \in \mathcal{D}, \\ I_k(\mathbf{x}, \mathbf{s}) = I_k^0(\mathbf{x}) \times \mathbb{1}_{[\mathbf{x} \in \partial \mathcal{D}_k^0]} \times \mathbb{1}_{[\mathbf{s} = \mathbf{s}_k^0]}, \quad \forall \mathbf{x} \in \partial \mathcal{D}, \end{aligned} \quad (25)$$

where \mathbf{s} is the considered direction of propagation, $\Phi(\mathbf{s}, \mathbf{s}')$ is the phase function representing the probability that a photon arriving from the direction \mathbf{s}' is scattered to the direction \mathbf{s} , κ and σ are the absorption and scattering space-dependent functions, respectively, and $\mathbb{1}_{[\cdot]}$ denotes the indicator function.

3.2. Cost Function Gradient Derivation. The computation of the cost function gradient follows the generic methodology detailed in the previous section. The cost function derivative towards the direction ϕ is given differentiating (24):

$$\begin{aligned} j'(\kappa, \sigma; \phi) &= (\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} \\ &= \left(\left(\frac{\mathbf{s} \cdot \mathbf{n}}{\check{I}} \right)^2 (I - \check{I}), I' \right)_{\mathcal{X}} \quad \forall \phi \in \mathcal{P}. \end{aligned} \quad (26)$$

In order to derive the cost gradient for this specific application, \mathcal{Y} -based inner products are defined.

Definition 11. Consider

$$\begin{aligned} (u, w)_{\mathcal{Y}} &\triangleq \sum_k \oint_{\mathcal{S}^{n-1}} \int_{\mathcal{D}} uw \, dx \, d\omega(\mathbf{s}), \\ (u, w)_{\partial\mathcal{Y}} &\triangleq \sum_k \oint_{\mathcal{S}^{n-1}} \int_{\partial\mathcal{D}} uw \, dx \, d\omega(\mathbf{s}). \end{aligned} \quad (27)$$

The \mathcal{Y} -inner product satisfies Properties 1 and 2, both being useful for the cost function gradient derivation.

Property 1. Consider

$$(\mathbf{s} \cdot \nabla u, w)_{\mathcal{Y}} = -(\mathbf{s} \cdot \nabla w, u)_{\mathcal{Y}} + (\mathbf{s} \cdot \mathbf{n}w, u)_{\partial\mathcal{Y}}. \quad (28)$$

Property 2. Consider

$$\begin{aligned} & \left(\sigma \oint_{\mathcal{S}^{n-1}} u(\mathbf{x}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), w(\mathbf{x}, \mathbf{s}) \right)_{\mathcal{Y}} \\ &= \sum_k \oint_{\mathcal{S}^{n-1}} \int_{\mathcal{D}} \sigma \oint_{\mathcal{S}^{n-1}} u(\mathbf{x}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}') w(\mathbf{x}, \mathbf{s}) \, dx \, d\omega(\mathbf{s}) \\ &= \sum_k \int_{\mathcal{D}} \sigma \oint_{\mathcal{S}^{n-1}} \oint_{\mathcal{S}^{n-1}} w(\mathbf{x}, \mathbf{s}) u(\mathbf{x}, \mathbf{s}') \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}') \, d\omega(\mathbf{s}) \, dx \quad (29) \\ &= \sum_k \int_{\mathcal{D}} \oint_{\mathcal{S}^{n-1}} \sigma \oint_{\mathcal{S}^{n-1}} w(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}') u(\mathbf{x}, \mathbf{s}') \, dx \, d\omega(\mathbf{s}') \\ &= \left(\sigma \oint_{\mathcal{S}^{n-1}} w(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), u(\mathbf{x}, \mathbf{s}') \right)_{\mathcal{Y}}. \end{aligned}$$

The generic equation for the cost gradient (15) applied on both space-distributed parameters κ and σ gives, for direction $\phi = (\phi_\kappa; 0) \in \mathcal{P}$,

$$(\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} = (I(\mathbf{x}, \mathbf{s}) I^*(\mathbf{x}, \mathbf{s}), \phi)_{\mathcal{Y}} \quad (30)$$

and, for direction $\phi = (0; \phi_\sigma) \in \mathcal{P}$,

$$(\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} = \left(I(\mathbf{x}, \mathbf{s}) I^*(\mathbf{x}, \mathbf{s}) - \oint_{\mathcal{S}^{n-1}} I(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}') I^*(\mathbf{x}, \mathbf{s}), \phi \right)_{\mathcal{Y}}. \quad (31)$$

In these two equations, the direction $\phi \in \mathcal{P}$ is split into two terms ϕ_κ and ϕ_σ in order to be able to compute the cost function gradient relatively to each of the functions κ and σ separately.

On one hand, (30) is rewritten such that

$$(\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} = (\phi I(\mathbf{x}, \mathbf{s}), I^*(\mathbf{x}, \mathbf{s}))_{\mathcal{Y}} \quad (32)$$

and, on the other hand, Property 2 enables rewriting (31) such that

$$\begin{aligned} (\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} &= \left(\phi I(\mathbf{x}, \mathbf{s}) \right. \\ &\quad \left. - \phi \oint_{\mathcal{S}^{n-1}} I(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), I^*(\mathbf{x}, \mathbf{s}) \right)_{\mathcal{Y}}. \end{aligned} \quad (33)$$

Besides, the differentiation of the forward model enables rewriting previous equations too, for ϕ being either $(\phi_\kappa; 0)$ or $(0; \phi_\sigma)$:

$$\begin{aligned} (\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} &= - \left((\mathbf{s} \cdot \nabla + \kappa + \sigma) I'(\mathbf{x}, \mathbf{s}), I^* \right)_{\mathcal{Y}} \\ &\quad + \left(\sigma \oint_{\mathcal{S}^{n-1}} I'(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), I^* \right)_{\mathcal{Y}}. \end{aligned} \quad (34)$$

Properties 1 and 2 enable rewriting (34) such that

$$\begin{aligned} & (\nabla j(\kappa, \sigma), \phi)_{\mathcal{X}} \\ &= \left((\mathbf{s} \cdot \nabla - \kappa - \sigma) I^*(\mathbf{x}, \mathbf{s}), I' \right)_{\mathcal{Y}} - \left(\mathbf{s} \cdot \mathbf{n} I^*, I' \right)_{\partial\mathcal{Y}} \quad (35) \\ &\quad + \left(\sigma \oint_{\mathcal{S}^{n-1}} I^*(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), I' \right)_{\mathcal{Y}}. \end{aligned}$$

The adjoint problem is eventually obtained identifying (26) with (35):

Find $I_k(\mathbf{x}, \mathbf{s})$ such that $\forall \mathbf{s} \in \mathcal{S}^{n-1}, \forall k = 1, \dots, k :$

$$\begin{aligned} & (-\mathbf{s} \cdot \nabla + \kappa + \sigma) I_k^*(\mathbf{x}, \mathbf{s}) \\ &= \sigma \oint_{\mathcal{S}^{n-1}} I_k^*(\mathbf{x}, \mathbf{s}) \phi(\mathbf{s}, \mathbf{s}') \, d\omega(\mathbf{s}'), \quad \forall \mathbf{x} \in \mathcal{D}, \quad (36) \\ & -\mathbf{s} \cdot \mathbf{n} I_k^*(\mathbf{x}, \mathbf{s}) = \left(\frac{\mathbf{s} \cdot \mathbf{n}}{\check{I}_k} \right)^2 (I_k - \check{I}_k)(\mathbf{x}, \mathbf{s}) \\ &\quad \times \mathbb{1}_{[\mathbf{x} \in \partial\mathcal{D}_d]}(\mathbf{x}), \quad \forall \mathbf{x} \in \partial\mathcal{D}, \end{aligned}$$

where $\partial\mathcal{D}_d$ denotes the set of the pointwise sensors.

3.3. Numerical Results. The optimization relies on the BFGS algorithm described by Liu and Nocedal [51] coupled with a quadratic interpolation line search [52]. Generally speaking, this latter has proved to be particularly effective in these recent years in the area of inversion (see [53–55] to name but a few) due to its effectiveness in minimizing nonlinear problems [56].

The two-dimensional domain is a square of 2 cm length. Four collimated sources are located at the center of each side, while sensors are placed around sources towards twenty directions. The target properties to be reconstructed, κ_t and σ_t , are defined by the background for which $\kappa = 0.25 \text{ cm}^{-1}$ and $\sigma = 20 \text{ cm}^{-1}$ and two square inclusions, the former for which $\kappa = 0.15 \text{ cm}^{-1}$ and $\sigma = 10 \text{ cm}^{-1}$ and the latter for which $\kappa = 0.35 \text{ cm}^{-1}$ and $\sigma = 30 \text{ cm}^{-1}$. Both the forward and the adjoint models are solved using \mathcal{P}_1 discontinuous Galerkin finite elements for which the mesh convergence has been validated. Three distinct meshes have been created, one in order to build the synthetic data before adding random noise following Klose [57] (3721 nodes), one in order to project the forward and adjoint-state variables (1681 nodes), and one associated with the property functions with the aim of implicitly regularizing the inverse problem (441 nodes). Seeking to gauge the accuracy of the reconstructions, the deviation factor $e_1 \in [0, +\infty[$ and the correlation factor $e_2 \in [-1, +1]$ are used [58]. The former factor measures the deviation of the reconstructed image from the target image: a small value indicates a reconstructed image with high accuracy. The latter factor measures the linear correlation between the target and the reconstructed image: a large value shows a high detectability in the reconstructed image and indicates a reconstructed image with high accuracy.

These errors are defined as in [15], for the parameter ϑ being either κ or σ :

$$e_1^\vartheta = 100 \times \frac{\sqrt{\sum_{j=1}^N (\vartheta^j - \vartheta_t^j)^2}}{\sqrt{\sum_{j=1}^N (\vartheta_t^j)^2}}, \quad (37)$$

$$e_2^\vartheta = \frac{\sum_{j=1}^N (\vartheta^j - \bar{\vartheta})(\vartheta_t^j - \bar{\vartheta}_t)}{(N-1)\sigma_\vartheta\sigma_{\vartheta_t}},$$

where N is the dimension of the parameter space (the number of finite elements nodes), ϑ^j is the value of ϑ on the i th node, and $\bar{\vartheta}$ and $\bar{\vartheta}_t$ are the mean values of vectors ϑ and ϑ_t , and σ_ϑ , σ_{ϑ_t} are standard deviations of target and reconstructed images, respectively. Standard deviations are given by

$$\sigma_\vartheta = \sqrt{\frac{\sum_{j=1}^N (\vartheta^j - \bar{\vartheta})^2}{N-1}}, \quad (38)$$

$$\sigma_{\vartheta_t} = \sqrt{\frac{\sum_{j=1}^N (\vartheta_t^j - \bar{\vartheta}_t)^2}{N-1}}.$$

Numerical tests were performed for different levels of noise. First, a small noise of 20 dB is added to the synthetic data (this corresponds to 1% of noise). Then, a moderate noise of 15 dB is added to the data (approximately 3% of noise). Finally, the highest level of noise of 10 dB is considered (10% of noise).

Figure 2 presents the evolution of dimensionless cost function values along with iterations. Note that the main used stopping criterion has been based on the cost function stabilization. As expected, it is seen from this figure that the higher the level of noise, the higher the cost function value at stabilization. Moreover, the convergence to the solution is slightly faster with Sobolev gradients than with ordinary ones, even though, far at the end, the value of the cost function is the lowest for the ordinary gradient.

Note that the BFGS optimizer is not based on matrix inversion, as it is the case for Gauss-Newton method and its derivative. It has been shown in other works that such a choice implicitly regularizes the inverse problem. Another regularization tool comes from the parameter mesh that is chosen sufficiently coarse to avoid the presence of too many local minima [37]. These tools make the reconstructions possible without the need of adding a penalization term to the cost function.

Note also that the choice of the stopping criterion plays a crucial role in the effectiveness of the algorithm: a too weak stopping criterion may lead to useless solution, while a too severe stopping criterion may lead to a divergent solution. In fact, the stopping criterion based on cost function stabilization is another dummy regularization that is to be used in practical applications. The determination of an appropriate value for this criterion is a challenging task in theory because it highly depends on the application. For the needs of this paper, the value equal to 10^{-2} has been determined empirically.

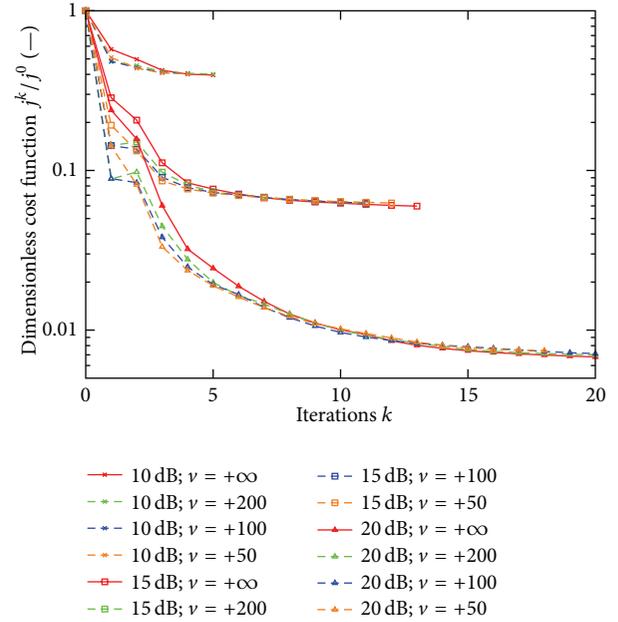


FIGURE 2: 2D RTE-based optical tomography. Evolution of the cost function with respect to iterations, for $\nu = +\infty$, 200, 100, and 50 and a 20 dB, 15 dB, and 10 dB noise. Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient. The algorithm stops at iteration k when $|J^k - J^{k-1}|/J^{k-1} < 10^{-2}$.

Figures 3, 4, and 5 present the evolutions, with respect to iterations, of deviation and correlation factors, respectively, to both parameters κ and σ , for the three levels of noise. These results show that when compared to ordinary gradients, the use of space-dependent Sobolev gradients decreases deviation factors and increases correlation factors. The optimum regularization parameter, ν , is about 100 in this particular case: when ν is too small, there is too much gradient diffusion, while when ν is too high, this yields the ordinary $L_2(\mathcal{D})$ gradient.

Figures 6 and 7 corroborate these conclusions. The Sobolev gradients filter the high frequency fluctuations and avoid the propagation of the noise to the cost function gradient and then to the reconstructions. It is observed that the regularization parameter $\nu = 100$ gives the best reconstructions as regards smoothness and precision of reconstructions.

4. Application on 3D DA-Based Optical Tomography

4.1. Inverse Problem Statement. The radiative transfer equation (RTE) provides an equation of light propagation valid in most of participating (absorbing and scattering) media as long as the independent scattering regime is fulfilled [59]. For specific applications, approximations of the RTE could be applied. Solutions of the RTE being computationally expensive, especially if the third dimension has to be considered, approximated models of the RTE should be employed when valid. Among the different models approaching the RTE,

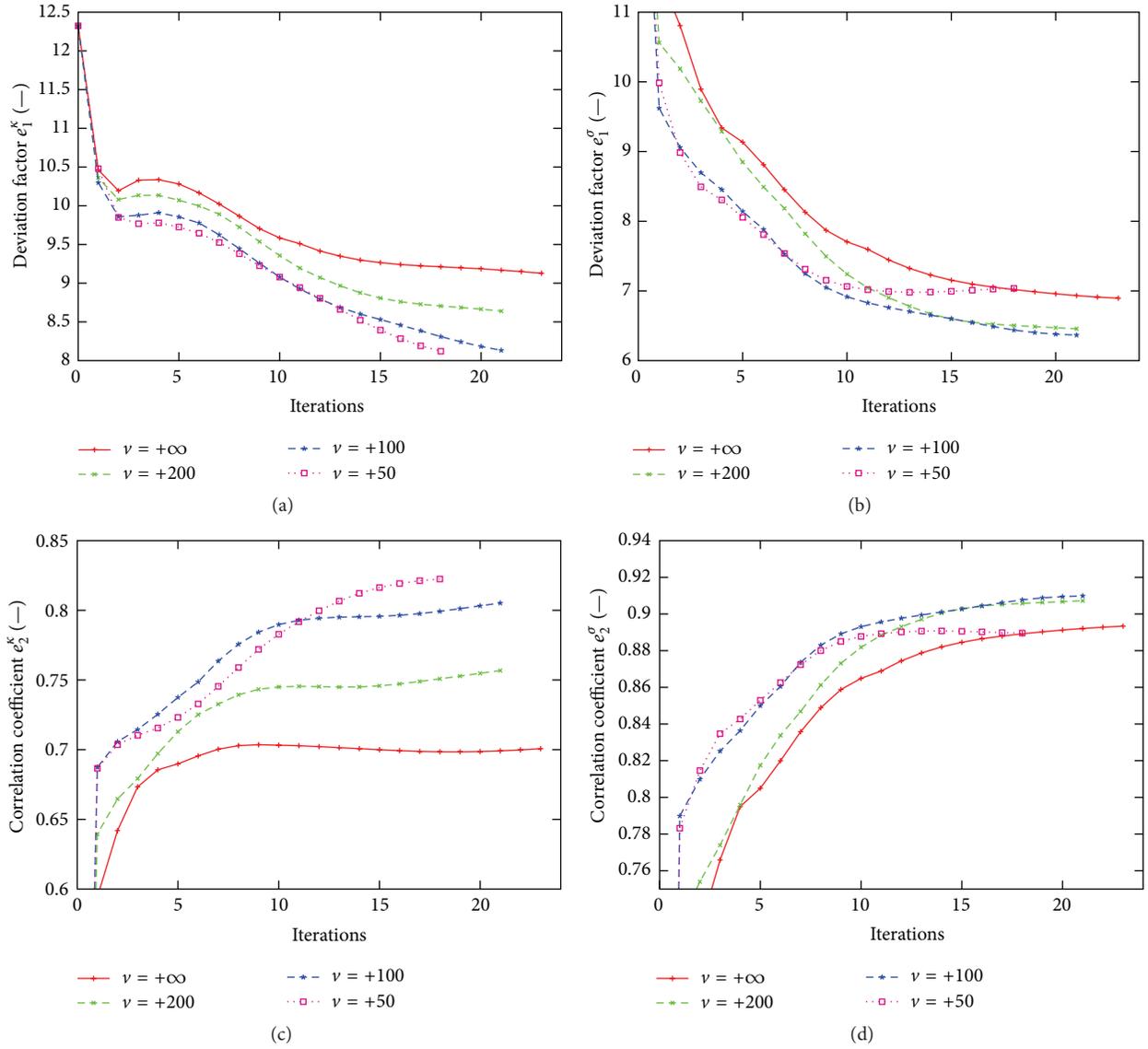


FIGURE 3: 2D RTE-based optical tomography. 20 dB noise. Evolution of errors e_1 ((a), (b)) and e_2 ((c), (d)) for κ ((a), (c)) and σ ((b), (d)), with respect to iterations, for $\nu = +\infty, 200, 100,$ and 50 . Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient. The algorithm stops at iteration k when $|j^k - j^{k-1}|/j^{k-1} < 10^{-2}$.

the diffuse approximation (DA) is the privileged one. The latter, which describes the photon density field $\varphi(\mathbf{x}) = \oint_{\mathcal{D}^{n-1}} I(\mathbf{x}, \mathbf{s}) d\omega(\mathbf{s})$ inside the medium, is given in the frequency domain by [1, 60]

$$\begin{aligned}
 & -\nabla \cdot [D(\mathbf{x}) \nabla \widehat{\varphi}_k(\mathbf{x}, f_0)] + \left[\kappa(\mathbf{x}) + \frac{2\pi i f_0}{c} \right] \widehat{\varphi}_k(\mathbf{x}, f_0) \\
 & = \widehat{X}_k(\mathbf{x}, f_0), \quad \forall \mathbf{x} \in \mathcal{D}, \\
 & 2\gamma \widehat{\varphi}_k(\mathbf{x}, f_0) + AD(\mathbf{x}) \nabla \widehat{\varphi}_k(\mathbf{x}, f_0) \cdot \mathbf{n} = 2\widehat{x}_k(\mathbf{x}, f_0), \\
 & \quad \quad \quad \forall \mathbf{x} \in \partial \mathcal{D}.
 \end{aligned} \tag{39}$$

Here, $\widehat{X}_k(\mathbf{x}, f_0)$ is the Fourier transform at the frequency f_0 of the internal source of radiation, X_k [W cm^{-3}]; $\widehat{x}_k(\mathbf{x}, f_0)$ is

the Fourier transform at the frequency f_0 of the temporal signal associated with the k th diffuse source, x_k [W cm^{-2}]; $\widehat{\varphi}_k : (\mathcal{D} \times \mathbb{C}) \mapsto \mathbb{R}$ is the Fourier transform of the k th photon density field, $\varphi_k : (\mathcal{D} \times \mathbb{R}^+) \mapsto \mathbb{R}$ [W cm^{-2}]; $D(\mathbf{x}) = (n_{\mathcal{D}}[\kappa(\mathbf{x}) + \sigma'(\mathbf{x})])^{-1}$ is the macroscopic scattering coefficient [cm]; $n_{\mathcal{D}}$ ($= 2, 3$) is the dimension of \mathcal{D} ; κ and $\sigma' = (1 - g)\sigma$ are the absorption and reduced scattering coefficients (σ is the scattering coefficient as previously defined) [cm^{-1}]; g is the asymmetry factor which equals the average cosine of the scattering angle; c is the speed of light in \mathcal{D} (constant) [cm s^{-1}]; \mathbf{n} is the unit normal vector to the boundary of \mathcal{D} ; γ is a constant parameter depending on $n_{\mathcal{D}}$ ($1/\pi$ for $n_{\mathcal{D}} = 2$, $1/4$ for $n_{\mathcal{D}} = 3$); and A is a parameter which characterizes the reflection on the border $\partial \mathcal{D}$ that can be derived from the Fresnel laws if specular reflection is considered [61, 62].

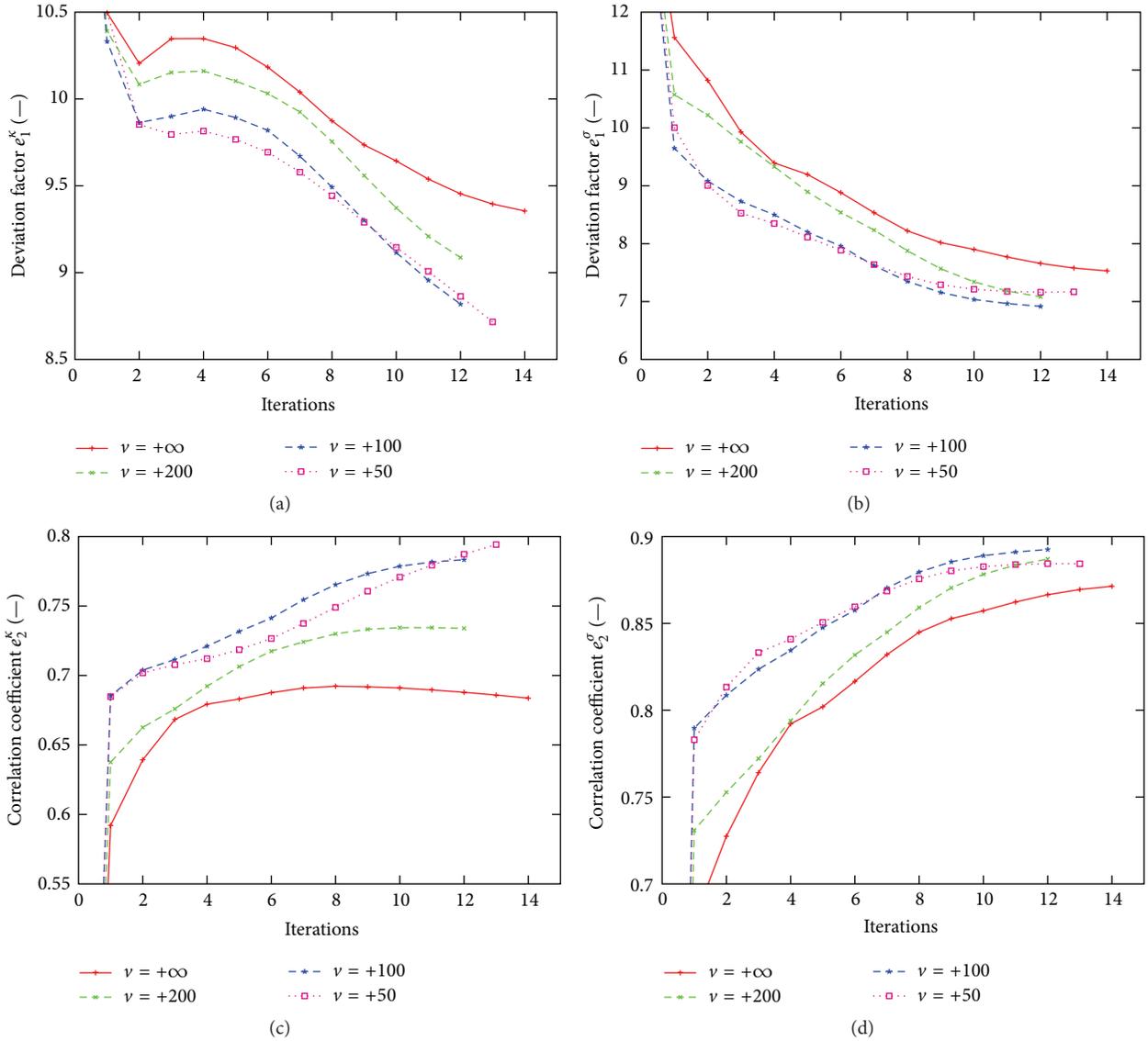


FIGURE 4: 2D RTE-based optical tomography. 15 dB noise. Evolution of errors e_1 ((a), (b)) and e_2 ((c), (d)) for κ ((a), (c)) and σ ((b), (d)), with respect to iterations, for $\nu = +\infty, 200, 100,$ and 50 . Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient. The algorithm stops at iteration k when $|j^k - j^{k-1}|/j^{k-1} < 10^{-2}$.

Similarly to Section 3.3, the inverse problem consists in solving the optimization problem:

$$\text{Find } (\kappa, \sigma') \in \mathcal{P} = L^\infty(\mathcal{D}) \times L^\infty(\mathcal{D})$$

$$\text{such that } (\kappa, \sigma') = \arg \min_{(\kappa, \sigma') \in \mathcal{P}} j(\kappa, \sigma'), \quad (40)$$

where

$$j(\kappa, \sigma') := \mathcal{F}(\hat{\varphi}) = \frac{1}{2} \left\| \frac{\hat{\varphi} - \hat{\varphi}}{\hat{\varphi}} \right\|_{\mathcal{X}}^2 \quad (41)$$

$$\text{with } \|u\|_{\mathcal{X}} \triangleq \left(\sum_{k=1}^K \sum_{d=1}^D |u_k(\mathbf{x}_d)|^2 \right)^{1/2}$$

and where the index k refers to the source number and the index d refers to the detection node. It should be noted that, under the assumption of the DA, the measurable quantity (23) is related to the photon density through the equality $E(\mathbf{x}) = E(\mathbf{x}, \{\mathbf{s} \cdot \mathbf{n} > 0\}) = 2\gamma\varphi(\mathbf{x})$ [60]. In the remainder of the section, the following values of the source terms are considered: $\hat{X}_k = 0 \text{ W cm}^{-2}$ and $\hat{x}_k = 0.1 \text{ W cm}^{-2}$ for all $k = 1, \dots, K$ (which corresponds to a temporal signal of the form $\cos(2\pi f_0 t)$).

4.2. Cost Function Gradients Derivation. The cost function gradients with respect to κ and σ' have been computed in [23] with the help of a continuous Lagrangian formulation.

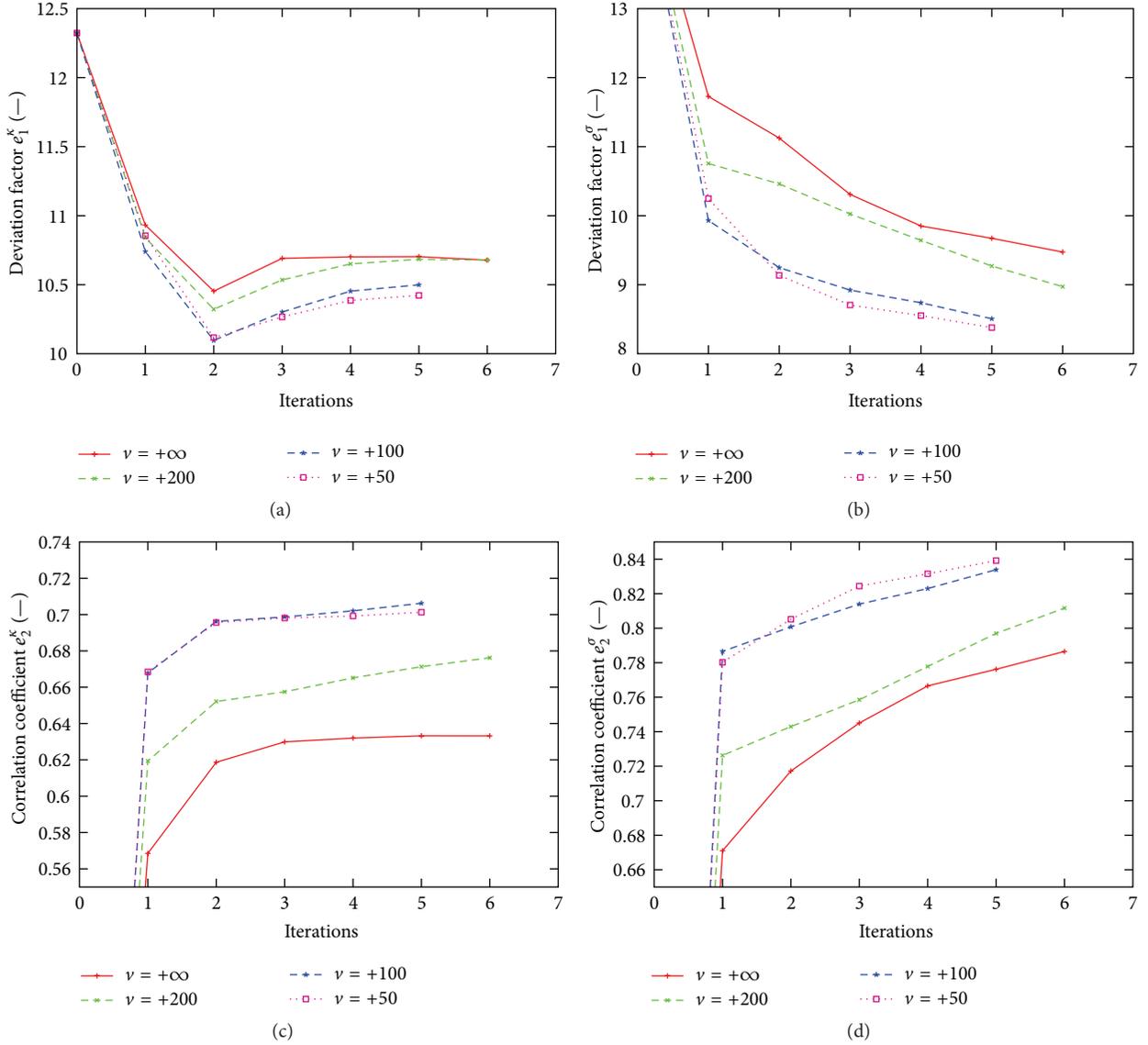


FIGURE 5: 2D RTE-based optical tomography. 10 dB noise. Evolution of errors e_1 ((a), (b)) and e_2 ((c), (d)) for κ ((a), (c)) and σ ((b), (d)), with respect to iterations, for $\nu = +\infty, 200, 100$, and 50 . Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient. The algorithm stops at iteration k when $|j^k - j^{k-1}|/j^{k-1} < 10^{-2}$.

In terms of (5), it can be shown that the directional derivatives of the cost function are equal to

$$\begin{aligned}
 & (\nabla j(\psi), \phi)_{\mathcal{F}} \\
 &= \sum_{k=1}^K \left(\Re \left(\widehat{\varphi}_k \overline{\widehat{\varphi}_k^*} - n_{\mathcal{D}} D^2 \nabla \widehat{\varphi}_k \cdot \nabla \overline{\widehat{\varphi}_k^*} \right), \phi \right)_{\mathcal{F}} \\
 & \quad \forall \phi = (\phi_\kappa; 0) \in \mathcal{P}, \quad (42)
 \end{aligned}$$

$$\begin{aligned}
 & (\nabla j(\psi), \phi)_{\mathcal{F}} = \sum_{k=1}^K \left(\Re \left(-n_{\mathcal{D}} D^2 \nabla \widehat{\varphi}_k \cdot \nabla \overline{\widehat{\varphi}_k^*} \right), \phi \right)_{\mathcal{F}} \\
 & \quad \forall \phi = (0; \phi_{\sigma'}) \in \mathcal{P},
 \end{aligned}$$

where \Re is the real part operator and the adjoint variable $\widehat{\varphi}_k^*$ involved in (15) is the solution to the adjoint problem:

$$-\nabla \cdot [D \nabla \widehat{\varphi}_k^*] + \left[\kappa - \frac{2\pi i f_0}{c} \right] \widehat{\varphi}_k^* = 0 \quad \forall \mathbf{x} \in \mathcal{D},$$

$$\frac{2\gamma}{A} \widehat{\varphi}_k^* + D \nabla \widehat{\varphi}_k^* \cdot \mathbf{n} = -\frac{1}{|\widehat{\varphi}_k|^2} (\widehat{\varphi}_k - \widetilde{\varphi}_k) \mathbb{1}_{[\mathbf{x} \in \partial \mathcal{D}_d]}(\mathbf{x}) \quad (43)$$

$$\forall \mathbf{x} \in \partial \mathcal{D}.$$

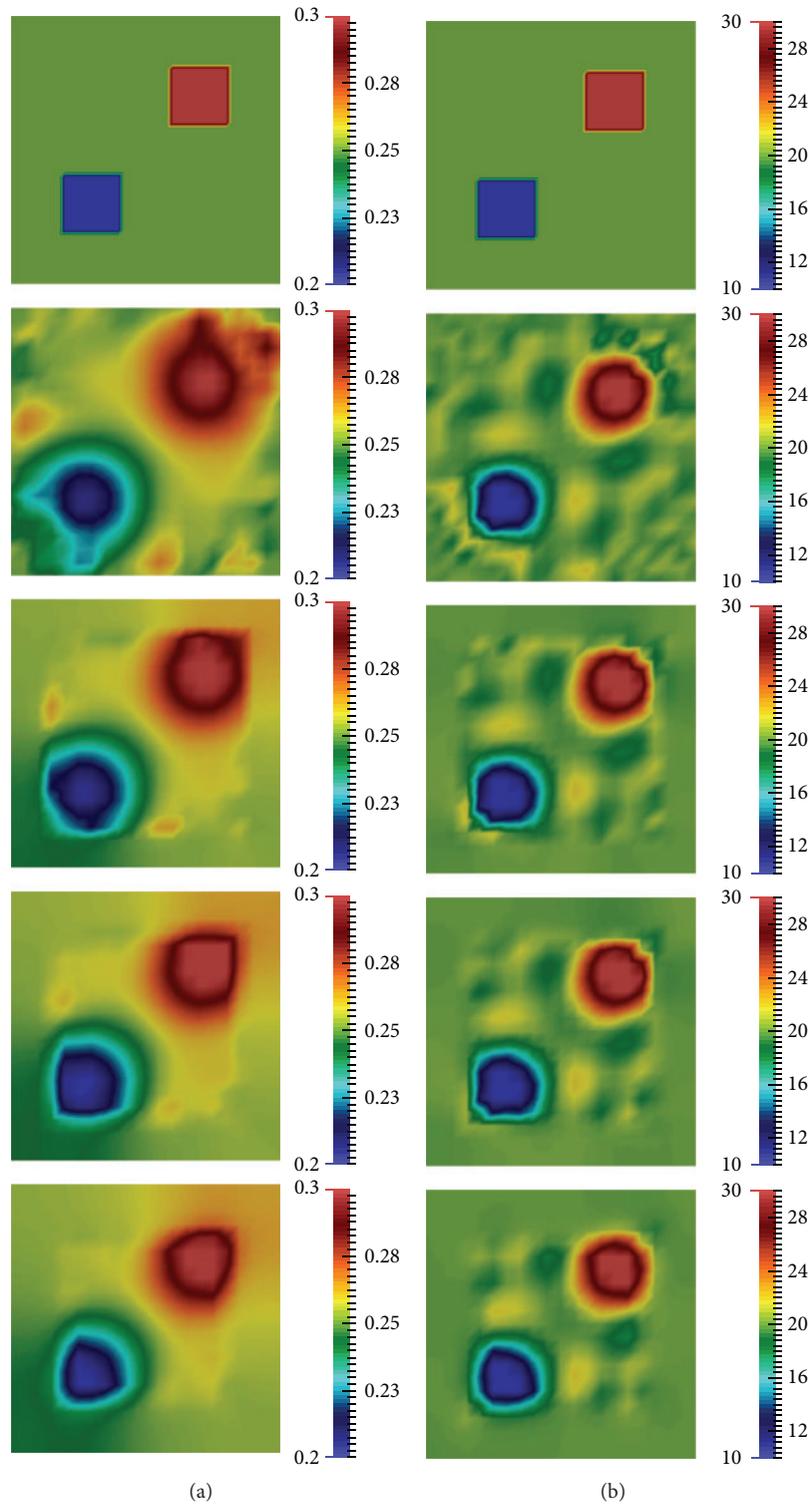


FIGURE 6: 2D RTE-based optical tomography. 20 dB noise. Reconstructions obtained for κ (a) and σ (b). 1st row: targets; then $\nu = +\infty, 200, 100$, and 50, respectively. Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient.

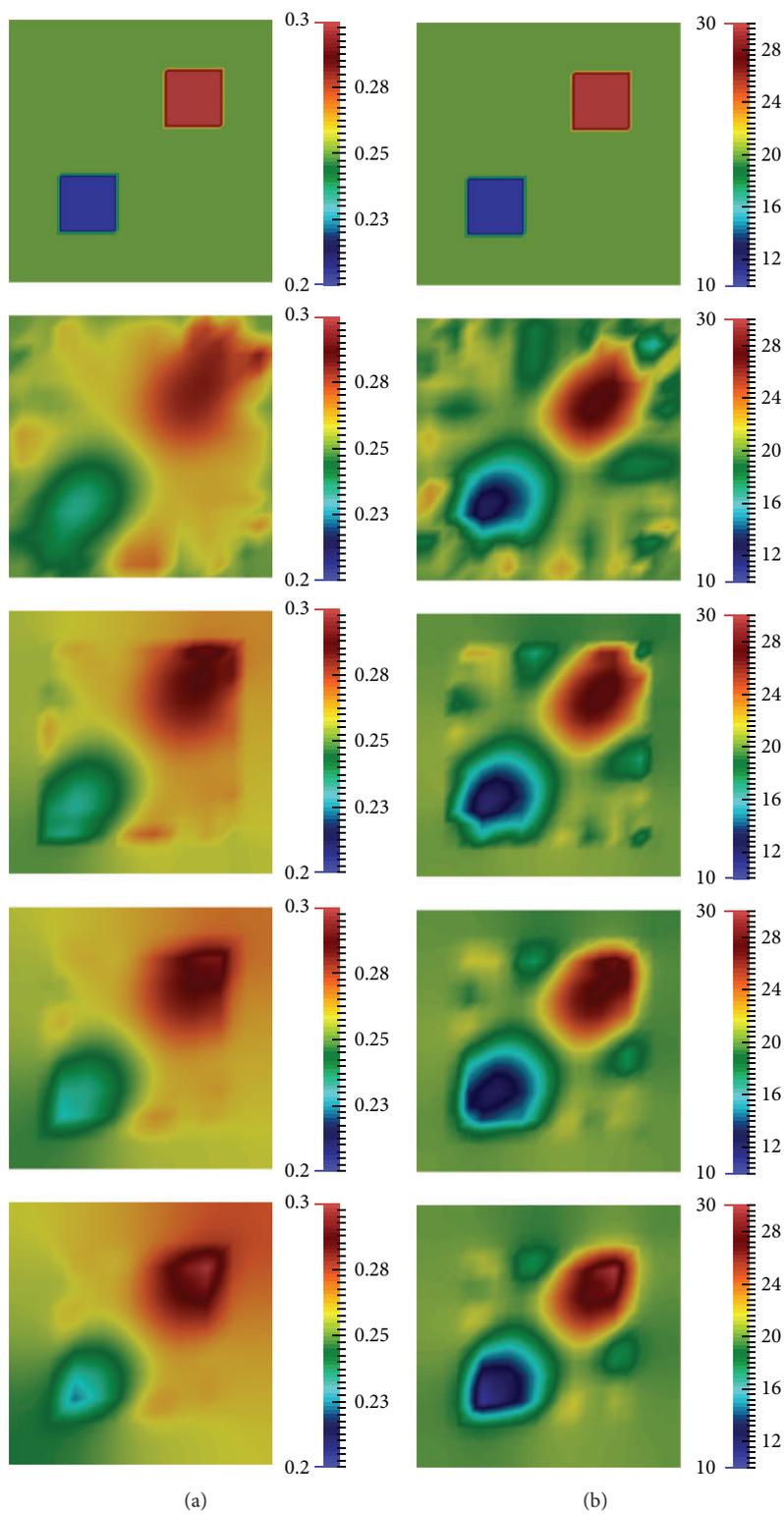


FIGURE 7: 2D RTE-based optical tomography. 10 dB noise. Reconstructions obtained for κ (a) and σ (b). 1st row: targets; then $\nu = +\infty, 200, 100, \text{ and } 50$, respectively. Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient.

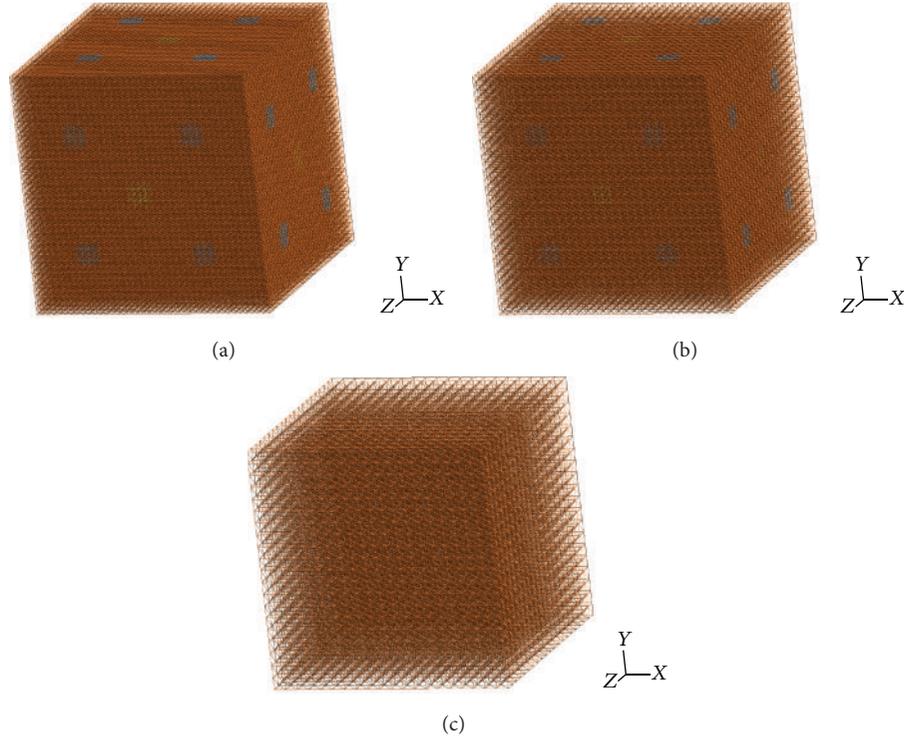


FIGURE 8: 3D DA-based optical tomography. Synthetic data mesh: 132 651 nodes (a); mesh of the state and adjoint variable: 68 921 nodes (b); mesh associated with radiative properties κ and σ' : 9 261 nodes (c).

Therefore, if the inner product $\mathcal{F} = L_2(\mathcal{D})$ is used to extract the cost function gradient in (5), the expressions for the continuous cost function gradients are such that

$$\begin{aligned} \nabla_{\kappa} j &= \sum_{k=1}^K \Re \left(\widehat{\varphi}_k \overline{\widehat{\varphi}_k^*} - n_{\mathcal{D}} D^2 \nabla \widehat{\varphi}_k \cdot \nabla \overline{\widehat{\varphi}_k^*} \right), \\ \nabla_{\sigma'} j &= \sum_{k=1}^K \Re \left(-n_{\mathcal{D}} D^2 \nabla \widehat{\varphi}_k \cdot \nabla \overline{\widehat{\varphi}_k^*} \right). \end{aligned} \quad (44)$$

Note that an adimensionalisation of the parameters by *a priori* radiative properties is used in the inversion process [23].

4.3. Numerical Results. The same algorithm as employed in Section 3 is used to reconstruct the target properties from initial estimates equal to values of those of the background. The domain considered, \mathcal{D} , is a square of 5 cm length ($\mathcal{D} = [0; 5]^2 \text{ cm}^2$). The target properties to be reconstructed, κ_t and σ'_t , are defined by

$$\begin{aligned} \kappa_t(\mathbf{x}) &= 0.1 + 0.05 \times \mathbb{1}_{[\mathbf{x} \in \mathcal{D}_1]}(\mathbf{x}) + 0.1 \\ &\quad \times \mathbb{1}_{[\mathbf{x} \in \mathcal{D}_2]}(\mathbf{x}) \text{ cm}^{-1}, \end{aligned} \quad (45)$$

$$\sigma'_t(\mathbf{x}) = 10 + 5 \times \mathbb{1}_{[\mathbf{x} \in \mathcal{D}_1]}(\mathbf{x}) + 10 \times \mathbb{1}_{[\mathbf{x} \in \mathcal{D}_2]}(\mathbf{x}) \text{ cm}^{-1},$$

where $\mathcal{D}_1 = [0.75; 1.75]^2 \text{ cm}^2$ and $\mathcal{D}_2 = [3.25; 4.25]^2 \text{ cm}^2$.

The stopping criteria are based on the cost function stabilization, with a critical value chosen equal to 10^{-3}

and a total number of iterations equal to 100. As in Section 3.3, these two criteria are used as dummy regularizations in order to get a low enough cost function value still without getting a divergent solution. Again, these parameters have been tuned empirically.

Physical properties involved in the forward model (39) are fixed to $f_0 = 100 \text{ MHz}$, $c = (3 \times 10^{10}/1.4) \text{ cm s}^{-1}$, and $\gamma = 0.25$. Synthetic data are considered for these numerical tests. These data, $\widehat{\varphi}$, which represent pseudoexperimental measurements, are built on a finer mesh (132 651 nodes) than the one that generates the predictions $\widehat{\varphi}$ (68 921 nodes) in order to avoid the inverse crime [60, 63]. Then, a 10 dB multiplicative white Gaussian noise is applied to $\widehat{\varphi}$ at the nodes of the sensors to simulate the noise inherent to experimental devices; that is, $\widehat{\varphi}_{\text{noisy}} = \widehat{\varphi}(1 + 0.1 \times m)$, where $m \sim \mathcal{N}(0, 1)$. Synthetic data, state, and parameter meshes can be seen in Figure 8. It can be seen from Figure 8 that 6 sources and 24 sensors, 0.5 cm^2 each (containing 16 discretization points), are used for this particular test. A reduction of the parameter space dimension is employed to improve the quality of the reconstructions obtained with the BFGS algorithm [23]: the mesh associated with the radiative properties κ and σ' contains only 9 261 nodes. Note that the total number of complex measurements is equal to 2 304.

Figure 9 presents the evolutions, with respect to iterations, of deviation and correlation factors, respectively, to both parameters. It is seen from this figure that when compared to ordinary L_2 gradients, the use of space-dependent Sobolev gradients decreases deviation factors and increases

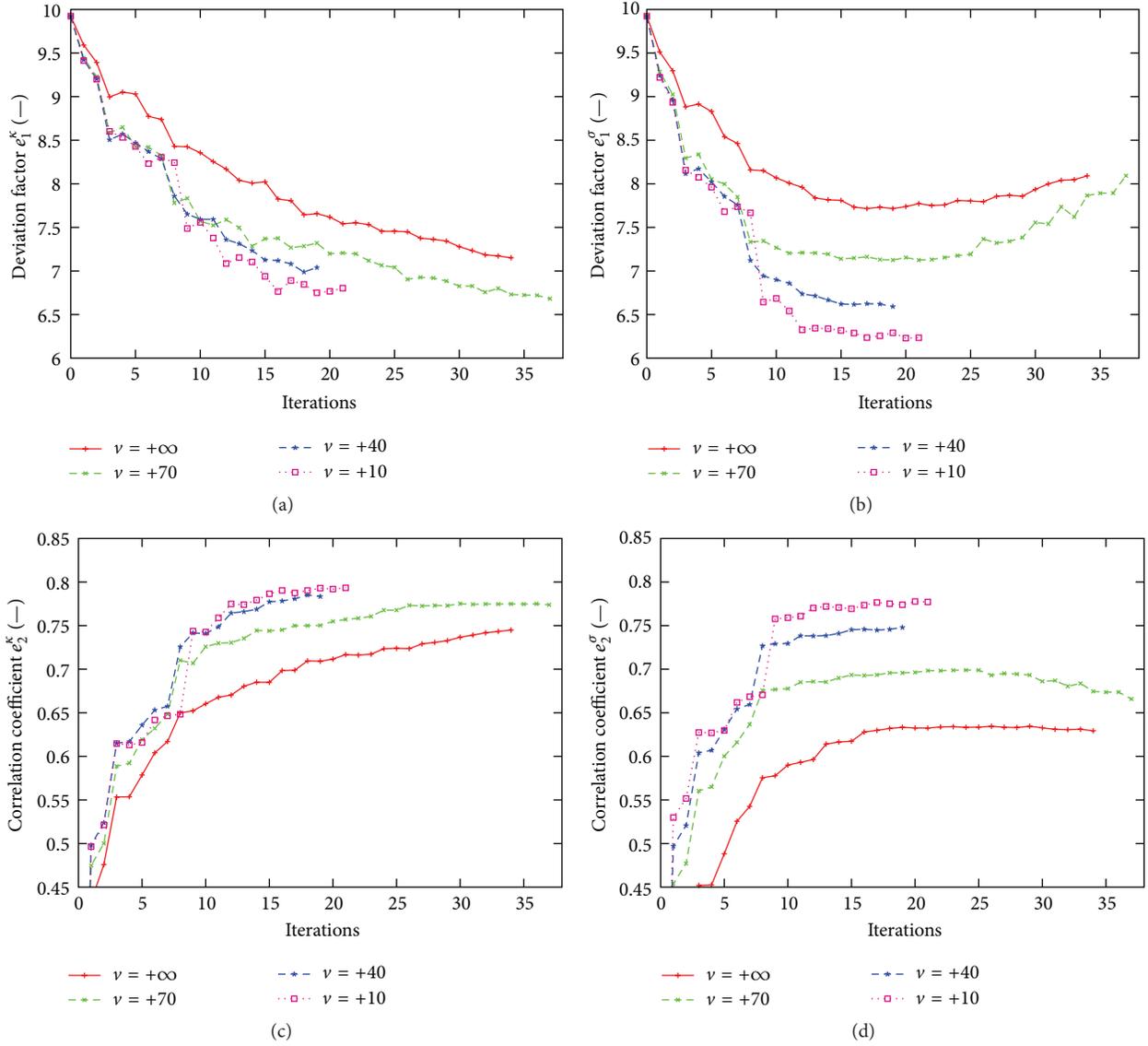


FIGURE 9: 3D DA-based optical tomography. 10 dB noise. Evolution of errors e_1 ((a), (b)) and e_2 ((c), (d)) for κ ((a), (c)) and σ' ((b), (d)), with respect to iterations, for $\nu = +\infty, 70, 40$, and 10. Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient. The algorithms stop at iteration k when $|j^k - j^{k-1}|/j^{k-1} < 10^{-3}$.

correlation factors. The optimum regularization parameter is about 10 in this particular case. Figure 10 corroborates these conclusions: the best reconstructions, in terms of smoothness and accuracy, are obtained for a value of the regularization parameter, ν , equal to 10. Finally, remark that Figure 9 allows observing the establishment of a cross talk effect [64, 65]: for $\nu = +\infty$ and $\nu = 70$, it is seen that the deviation factor e_1^κ decreases with the iterations, while the deviation factor $e_1^{\sigma'}$ increases. In these two cases, values of the cost function decrease with the iterations, while an improvement of the absorption coefficient reconstruction and a deterioration of the reduced scattering coefficient reconstruction are observed. This effect is prevented by the use of space-dependent Sobolev gradients with an appropriate regularization parameter weight.

5. Conclusion

In this paper, inverse models based on the BFGS algorithm have been developed to solve optical tomography problems based on two different forward models: the bidimensional steady-state radiative transfer equation and the three-dimensional frequency diffuse approximation. A Sobolev filter function was defined and space-dependent Sobolev inner products were used when extracting the cost function gradients instead of the L_2 -inner product. Two test cases consisting in reconstructing discontinuous radiative properties in a square and a cube have been considered. Numerical results obtained have shown that, for all levels of noise and all values of the regularization parameter tested, the use of Sobolev gradients drastically enhances the quality

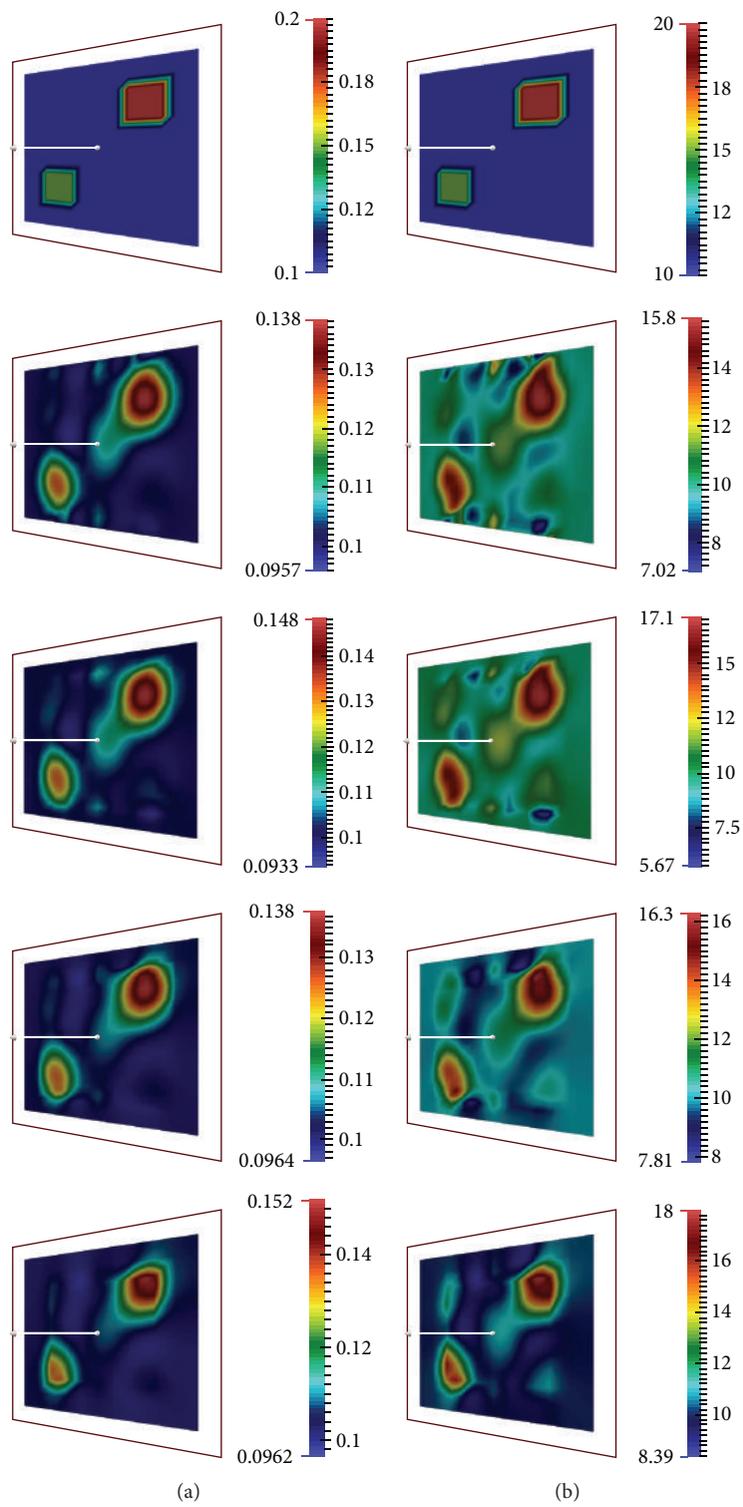


FIGURE 10: 3D DA-based optical tomography. 10 dB noise. Reconstructions obtained in the plane passing through the center of the cube of which the normal vector \mathbf{n} equals $(1, -1, 0)^T$, for κ (a) and σ' (b). 1st row: targets; then $\nu = +\infty, 70, 40$, and 10 , respectively. Note that the case $\nu = +\infty$ corresponds to the ordinary L_2 gradient.

of the reconstructions. Also, it has been observed that the space-dependent Sobolev gradients can allow avoiding the establishment of a crosstalk effect in the case of the frequency three-dimensional approximation diffuse based inversion.

Generally speaking, due to simplicity of implementation and effectiveness in filtering noise measurements locally that provides better reconstructions, space-dependent Sobolev gradients appear to be particularly attractive in the context of inverse problems. In the near future, the anisotropy factor will be integrated in the inverse problem based on the radiative transfer equation and the use of a logarithmic cost function will be considered as our more recent results suggest. Later, inversion numerical algorithms of the three-dimensional radiative transfer equation and diffuse approximation model will be integrated in a global numerical code. The latter will be based on the BFGS algorithm with the use of adimensionalisation of the radiative properties, reduction of the parameter space dimension, space-dependent Sobolev gradients, and a wavelet multiscale approach, as developed in another work.

Additional Points

Note that the proposed methods have been implemented within the FreeFem++ environment [66].

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper and regarding the funding that they have received.

Acknowledgments

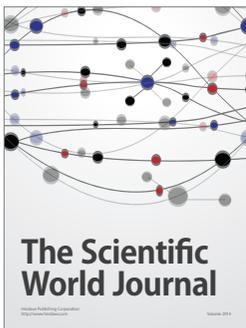
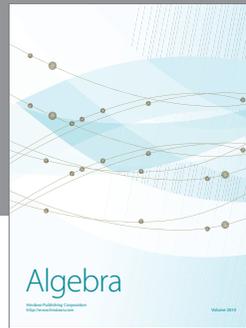
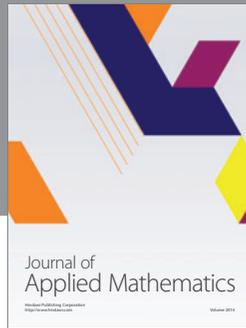
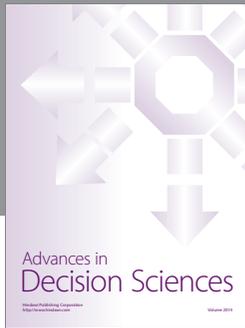
The authors thank the *Centre Régional de Calcul Intensif des Pays de la Loire* (CCIPL), financed by the French Research Ministry, the *Région Pays de la Loire* and the University of Nantes, for allowing computations on their supercomputers. The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research.

References

- [1] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems*, vol. 15, no. 2, pp. R41–R93, 1999.
- [2] A. H. Hielscher, A. Y. Bluestone, G. S. Abdoulaev et al., "Near-infrared diffuse optical tomography," *Disease Markers*, vol. 18, no. 5-6, pp. 313–337, 2002.
- [3] A. Charette, J. Boulanger, and H. K. Kim, "An overview on recent radiation transport algorithm development for optical tomography imaging," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 109, no. 17-18, pp. 2743–2766, 2008.
- [4] S. Z. M. Muji, R. A. Rahim, M. H. F. Rahiman et al., "Optical tomography: a review on sensor array, projection arrangement and image reconstruction algorithm," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 7, pp. 3839–3856, 2011.
- [5] M. T. M. Khairi, S. Ibrahim, M. A. M. Yunus, and M. Faramarzi, "A review on applications of optical tomography in industrial process," *International Journal on Smart Sensing and Intelligent Systems*, vol. 5, no. 4, pp. 767–798, 2012.
- [6] D. Baillis and J.-F. Sacadura, "Thermal radiation properties of dispersed media: theoretical prediction and experimental characterization," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 67, no. 5, pp. 327–363, 2000.
- [7] A. H. Hielscher, A. D. Klose, A. K. Scheel et al., "Sagittal laser optical tomography for imaging of rheumatoid finger joints," *Physics in Medicine and Biology*, vol. 49, no. 7, pp. 1147–1163, 2004.
- [8] B. W. Pogue, M. Testorf, T. McBride, U. Osterberg, and K. Paulsen, "Instrumentation and design of a frequency-domain diffuse optical tomography imager for breast cancer detection," *Optics Express*, vol. 1, no. 13, pp. 391–403, 1997.
- [9] D. R. Leff, O. J. Warren, L. C. Enfield et al., "Diffuse optical imaging of the healthy and diseased breast: a systematic review," *Breast Cancer Research and Treatment*, vol. 108, no. 1, pp. 9–22, 2008.
- [10] G. Strangman, D. A. Boas, and J. P. Sutton, "Non-invasive neuroimaging using near-infrared light," *Biological Psychiatry*, vol. 52, no. 7, pp. 679–693, 2002.
- [11] D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy," *NeuroImage*, vol. 23, supplement 1, pp. S275–S288, 2004.
- [12] H. Dehghani, B. R. White, B. W. Zeff, A. Tizzard, and J. P. Culver, "Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography," *Applied Optics*, vol. 48, no. 10, pp. D137–D143, 2009.
- [13] A. H. Hielscher, "Optical tomographic imaging of small animals," *Current Opinion in Biotechnology*, vol. 16, no. 1, pp. 79–88, 2005.
- [14] A. D. Klose, U. Netz, J. Beuthan, and A. H. Hielscher, "Optical tomography using the time-independent equation of radiative transfer. Part 1. Forward model," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 72, no. 5, pp. 691–713, 2002.
- [15] A. D. Klose and A. H. Hielscher, "Quasi-Newton methods in optical tomographic image reconstruction," *Inverse Problems*, vol. 19, no. 2, pp. 387–409, 2003.
- [16] K. Ren, G. Bal, and A. H. Hielscher, "Frequency domain optical tomography based on the equation of radiative transfer," *SIAM Journal on Scientific Computing*, vol. 28, no. 4, pp. 1463–1489, 2006.
- [17] T. Tarvainen, M. Vauhkonen, and S. R. Arridge, "Gauss-Newton reconstruction method for optical tomography using the finite element solution of the radiative transfer equation," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 109, no. 17-18, pp. 2767–2778, 2008.
- [18] O. Balima, J. Boulanger, A. Charette, and D. Marceau, "New developments in frequency domain optical tomography. Part II: application with a L-BFGS associated to an inexact line search," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 112, no. 7, pp. 1235–1240, 2011.
- [19] O. Balima, Y. Favennec, and D. Rousse, "Optical tomography reconstruction algorithm with the finite element method: an optimal approach with regularization tools," *Journal of Computational Physics*, vol. 251, pp. 461–479, 2013.
- [20] H. Jiang, *Diffuse Optical Tomography: Principles and Applications*, CRC Press, New York, NY, USA, 2010.

- [21] S. R. Arridge, M. Schweiger, and J. C. Schotland, "Inverse models of light transport," in *Handbook of Biomedical Optics*, chapter 17, pp. 319–336, CRC Press, 2011.
- [22] F. Dubot, Y. Favennec, B. Rousseau, and D. R. Rousse, "A wavelet multi-scale method for the inverse problem of diffuse optical tomography," *Journal of Computational and Applied Mathematics*, vol. 289, pp. 267–281, 2015.
- [23] F. Dubot, Y. Favennec, B. Rousseau, and D. R. Rousse, "Regularization opportunities for the diffuse optical tomography problem," *International Journal of Thermal Sciences*, vol. 98, pp. 1–23, 2015.
- [24] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, and J. P. Kaipio, "Hybrid radiative-transfer-diffusion model for optical tomography," *Applied Optics*, vol. 44, no. 6, pp. 876–886, 2005.
- [25] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, S. R. Arridge, and J. P. Kaipio, "Coupled radiative transfer equation and diffusion approximation model for photon migration in turbid medium with low-scattering and non-scattering regions," *Physics in Medicine and Biology*, vol. 50, no. 20, pp. 4913–4930, 2005.
- [26] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, J. P. Kaipio, and S. R. Arridge, "Utilizing the radiative transfer equation in optical tomography," *Piers Online*, vol. 4, no. 6, pp. 655–660, 2008.
- [27] H. K. Kim and A. H. Hielscher, "A diffusion-transport hybrid method for accelerating optical tomography," *Journal of Innovative Optical Health Sciences*, vol. 3, no. 4, pp. 293–305, 2010.
- [28] T. Tarvainen, V. Kolehmainen, S. R. Arridge, and J. P. Kaipio, "Image reconstruction in diffuse optical tomography using the coupled radiative transport-diffusion model," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 112, no. 16, pp. 2600–2608, 2011.
- [29] K. Woodbury, *Inverse Engineering Handbook*, CRC Press, Boca Raton, Fla, USA, 2003.
- [30] D. Baillis, M. Arduini-Schuster, and J. F. Sacadura, "Identification of spectral radiative properties of polyurethane foam from hemispherical and bi-directional transmittance and reflectance measurements," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 73, no. 2–5, pp. 297–306, 2002.
- [31] K. W. Kim, S. W. Baek, M. Y. Kim, and H. S. Ryou, "Estimation of emissivities in a two-dimensional irregular geometry by inverse radiation analysis using hybrid genetic algorithm," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 87, no. 1, pp. 1–14, 2004.
- [32] A. K. Parwani, P. Talukdar, and P. M. V. Subbarao, "Performance evaluation of hybrid differential evolution approach for estimation of the strength of a heat source in a radiatively participating medium," *International Journal of Heat and Mass Transfer*, vol. 56, no. 1–2, pp. 552–560, 2013.
- [33] A. Tikhonov and V. Arsenin, *Solution of Ill-Posed Problems*, John Wiley & Sons, New York, NY, USA, 1977.
- [34] V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer, Berlin, Germany, 1984.
- [35] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.
- [36] F. Dubot, Y. Favennec, B. Rousseau, Y. Jarny, and D. R. Rousse, "Quasi-optimal Tikhonov penalization and parameterization coarseness in space-dependent function estimation," *Inverse Problems in Science and Engineering*, vol. 24, no. 3, pp. 465–481, 2016.
- [37] G. Chavent, *Nonlinear Least Squares for Inverse Problems*, Springer, Berlin, Germany, 2009.
- [38] D. A. Murio, *The Mollification Method and the Numerical Solution of Ill-Posed Problems*, John Wiley & Sons, 1993.
- [39] B. Protas, "Adjoint-based optimization of PDE systems with alternative gradients," *Journal of Computational Physics*, vol. 227, no. 13, pp. 6490–6510, 2008.
- [40] R. J. Renka, "Image segmentation with a Sobolev gradient method," *Nonlinear Analysis. Theory, Methods & Applications. An International Multidisciplinary Journal. Series A: Theory and Methods*, vol. 71, no. 12, pp. e774–e780, 2009.
- [41] I. Danaila and P. Kazemi, "A new Sobolev gradient method for direct minimization of the gross-pitaevskii energy with rotation," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2447–2467, 2010.
- [42] A. Nassiopoulou and F. Bourquin, "Fast three-dimensional temperature reconstruction," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 49–52, pp. 3169–3178, 2010.
- [43] F. Bourquin and A. Nassiopoulou, "Inverse reconstruction of initial and boundary conditions of a heat transfer problem with accurate final state," *International Journal of Heat and Mass Transfer*, vol. 54, no. 15–16, pp. 3749–3760, 2011.
- [44] D. Maillet, Y. Favennec, and P. Le Masson, "Eurotherm spring school METTI 2011: thermal measurements and inverse techniques," Tech. Rep., Société Française de Thermique, Roscoff, France, 2011, <http://www.sft.asso.fr>.
- [45] G. Allaire, *Numerical Analysis and Optimization*, Oxford Science Publications, 2007.
- [46] J.-L. Lions, *Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles*, Dunod, Paris, France, 1968.
- [47] J. Céa, *Optimisation, Théorie et Algorithmes*, Dunod, 1971.
- [48] B. Protas, T. R. Bewley, and G. Hagen, "A computational framework for the regularization of adjoint analysis in multiscale PDE systems," *Journal of Computational Physics*, vol. 195, no. 1, pp. 49–89, 2004.
- [49] W. T. Mahavier, "A numerical method utilizing weighted Sobolev descent to solve singular differential equations," *Nonlinear World*, vol. 4, no. 4, pp. 435–455, 1997.
- [50] A. Majid and S. Sial, "Application of Sobolev gradient method to Poisson-Boltzmann system," *Journal of Computational Physics*, vol. 229, no. 16, pp. 5742–5754, 2010.
- [51] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 3, pp. 503–528, 1989.
- [52] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer, New York, NY, USA, 2007.
- [53] Z. Wang, Y. Liu, G. Wang, and L. Sun, "Elastography method for reconstruction of nonlinear breast tissue properties," *Journal of Biomedical Imaging*, vol. 2009, Article ID 406854, 9 pages, 2009.
- [54] F. Mohebbi and M. Sellier, "Aerodynamic optimal shape design based on body-fitted grid generation," *Mathematical Problems in Engineering*, vol. 2014, Article ID 505372, 22 pages, 2014.
- [55] G. Jin, Q. Liu, and X. Lv, "Inversion study of vertical eddy viscosity coefficient based on an internal tidal model with the adjoint method," *Mathematical Problems in Engineering*, vol. 2015, Article ID 915793, 14 pages, 2015.
- [56] C.-S. Liu, "Optimal algorithms and the BFGS updating techniques for solving unconstrained nonlinear minimization problems," *Journal of Applied Mathematics*, vol. 2014, Article ID 324181, 14 pages, 2014.

- [57] A. Klose, *Optical tomography based on the equation of radiative transfer [Ph.D. thesis]*, Department of Physics, Freie Universität Berlin, Berlin, Germany, October 2001.
- [58] B. P. Flannery, W. H. Press, S. A. Teukolsky, and W. Vetterling, *Numerical Recipes in C*, vol. 24, Press Syndicate of the University of Cambridge, New York, NY, USA, 1992.
- [59] J.-F. Sacadura, "Thermal radiative properties of complex media: theoretical prediction versus experimental identification," *Heat Transfer Engineering*, vol. 32, no. 9, pp. 754–770, 2011.
- [60] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160, Springer, New York, NY, USA, 2005.
- [61] M. Keijzer, W. M. Star, and P. R. Storchi, "Optical diffusion in layered media," *Applied Optics*, vol. 27, no. 9, pp. 1820–1824, 1988.
- [62] R. C. Haskell, L. O. Svaasand, T.-T. Tsay, T.-C. Feng, M. S. McAdams, and B. J. Tromberg, "Boundary conditions for the diffusion equation in radiative transfer," *Journal of the Optical Society of America A*, vol. 11, no. 10, pp. 2727–2741, 1994.
- [63] D. L. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, vol. 93 of *Applied Mathematical Sciences*, Springer, 1992.
- [64] S. R. Arridge and W. R. B. Lionheart, "Nonuniqueness in diffusion-based optical tomography," *Optics Letters*, vol. 23, no. 11, pp. 882–884, 1998.
- [65] T. McBride, B. W. Pogue, U. L. Osterberg, and K. D. Paulsen, "Separation of absorption and scattering heterogeneities in nir tomographic imaging of tissue," in *Biomedical Optical Spectroscopy and Diagnostics*, Paper TuC2, Optical Society of America, 2000.
- [66] F. Hecht, "New development in freefem++," *Journal of Numerical Mathematics*, vol. 20, no. 3-4, pp. 251–266, 2012.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

