A wavelet multi-scale method for the inverse problem of diffuse optical tomography

Fabien Dubot[a,b,*], Yann Favennec[b], Benoit Rousseau[b], Daniel R. Rousse[a]

[a] *Chaire de recherche industrielle en technologies de l'énergie et en efficacité énergétique (t3e), École de technologie supérieure, 1100, rue Notre-Dame Ouest, Montréal, Canada H3C 1K3*

[b] *Laboratoire de Thermocinétique de Nantes (LTN), UMR CNRS 6607, Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France*

Abstract

This paper deals with the estimation of optical property distributions of participating media from a set of light sources and sensors located on the boundaries of the medium. This is the so-called diffuse optical tomography problem. Such a non-linear ill-posed inverse problem is solved through the minimization of a cost function which depends on the discrepancy, in a least-square sense, between some measurements and associated predictions. In the present case, predictions are based on the diffuse approximation model in the frequency domain while the optimization problem is solved by the L-BFGS algorithm. To cope with the local convergence property of the optimizer and the presence of numerous local minima in the cost function, a wavelet multi-scale method associated with the L-BFGS method is developed, implemented, and validated. This method relies on a reformulation of the original inverse problem into a sequence of sub-inverse problems of different scales using wavelet transform, from the largest scale to the smallest one. Numerical results show that the proposed method brings more stability with respect to the ordinary L-BFGS method and enhances the reconstructed images for most of initial guesses of optical properties.

Keywords: Optical tomography, Wavelet multi-scale method, Inversion, L-BFGS algorithm, Optical properties
2000 MSC: 65M32, 65T60, 35Q93

*Corresponding author
Email address: `Fabien.Dubot@univ-nantes.fr` (Fabien Dubot)

1. Introduction

The diffuse optical tomography (DOT) problem is an inverse problem in which the spatial distributions of absorption and scattering coefficients of a participating medium are sought from a set of light sources and sensors located on the frontier of the probed medium [1]. The main application of the DOT is imaging biological tissues based on the fact that knowledge of the optical properties provides information on the physiopathological condition of these tissues. A thorough review of work on applications of the DOT in the biomedical domain can be found in [2].

Such a non-linear ill-posed inverse problem is solved through the minimization of a cost function which depends on the discrepancy, in a least-square sense, between some measurements and associated predictions. The methods employed to solve this problem include the non-linear conjugate-gradient method [3], Gauss-Newton based methods [4, 5, 6, 7], the L-BFGS method [8, 9, 10, 11, 12], shape-based reconstruction method [13, 14] or, in a Bayesian framework, the approximation error method [15, 16]. Regarding the first three listed methods, some problems remain to be overcome such as the stability with respect to the initial guesses and the blurring effect of the reconstructed images due to the need for relatively strong regularization tools [17, 14].

A relatively new method has emerged in the field of inversion, namely the wavelet multi-scale method [18, 19, 20, 21, 22, 23, 24, 25]. This method relies on a reformulation of the original inverse problem into a sequence of sub-inverse problems of different scales using wavelet transform, from the largest scale to the smallest one. Successful applications of this method include the inversion of the Maxwell equations [19], the identification of space-dependent porosity in fluid-saturated porous media [20, 21], the inversion of the two-dimensional acoustic wave equation [22], the reconstruction of permittivity distribution by the inversion of the electrical capacitance tomography model [23], the parameter estimation of elliptical partial differential equations [18, 24] and the identification of space-dependent permeability in a nonlinear diffusion equation [25]. It is shown in these papers that the wavelet multi-scale method can enhance stability of inversion, accelerate convergence and cope with the presence of local minima to reach the global minimum. The rationale behind this success is that the cost function shows stronger

convexity and has less minima at larger scale such that the global minimum can be achieved. Best results can thus be obtained, using this solution to the initialization of the optimization problem, at the lower scale until finding the minimum at the original scale.

This paper is dedicated to the application of a wavelet multi-scale method associated with the L-BFGS method to the specific inverse problem of diffuse optical tomography (DOT) and aims to improve the efficiency of the ordinary L-BFGS algorithm. To the best knowledge of the authors, this is the first time that the L-BFGS algorithm is coupled with the multi-scale method in solving the DOT inverse problem.

The paper is organized as follows: section 2 shortly summarizes the models used in the area of optical tomography, namely the radiative transfer equation (RTE) and the related diffuse approximation (DA) model; section 3 introduces the cost function, provides the expression of the continuous cost function gradient and defines the optimization problem; section 4 describes the L-BFGS algorithm; section 5 presents an introduction to the wavelet theory and the proposed wavelet multi-scale method; and section 6 provides selected numerical results to assess the proposed method.

But before presenting the work carried-out in the paper, some clarification is mandatory to fully understand the originality of the work. Indeed, although some papers found in the literature deal with the DOT problem using the wavelet tool such as [26, 27, 28, 29], these works show significant differences with the current research:

- In [26], the wavelet transform is used to denoise and compress noisy experimental data before performing the reconstruction. In this paper, the use of the wavelet transform is not limited to the data filtering, the wavelet transform is used in all stages of the inverse problem.

- In [27, 28, 29], the DOT problem is solved through the solution of a linear perturbation equation in which the main ingredient is a matrix of weights that are essentially the derivatives of the detector readings with respect to the optical coefficients in the reference medium. Several iterative algorithms have been used to solve this matrix system, including the conjugate gradient descent [27] and the total least squares [28, 29]. In all these algorithms, the multi-scale approach consists in multiplying both terms of the matrix system by wavelet transform matrices at each scale to obtain solutions at the different scales. In the present paper, the methodology is totally different: the DOT problem

is solved through the minimization of a least squares cost function by the L-BFGS algorithm for which the main ingredient is the cost function gradient. Thus, the multi-scale approach is mainly based on the representation of this gradient at the different scales.

- In [27, 28, 29], the steady-state diffuse approximation (DA) was considered while the DA in the frequency domain is used in the present case. As a result, only the absorption coefficient was reconstructed in [27, 28, 29] assuming the reduced scattering coefficient known within the media, while the simultaneous reconstruction of both the absorption and reduced scattering coefficients is performed in this paper, which represents an additional difficulty.

As a final comment about the current paper, it is worth mentioning that a wavelet multi-scale method associated with the L-BFGS algorithm is compared to the ordinary L-BFGS method. Comparaisons with other methods found in the literature such as the Gauss–Newton method is out of the scope of the paper. However, let us point-out that the L-BFGS method has proven its worth in the field of the optical tomography as shown by the following papers [8, 9, 10, 11].

2. The diffuse optical tomography problem

In the framework of radiative heat transfer in participating medium, the quantity of interest is the radiative intensity I. The equation describing the spectral radiative intensity distribution in space and time at the frequency ν, I_ν, in the medium is the RTE [30, 31]:

$$\left(\frac{n}{c_0}\frac{\partial}{\partial t} + \mathbf{s}\cdot\nabla\right) I_\nu(\mathbf{r},\mathbf{s},t) = -(\kappa(\mathbf{r}) + \sigma(\mathbf{r}))I_\nu(\mathbf{r},\mathbf{s},t) + \sigma(\mathbf{r})\int_{\mathcal{S}^2} I_\nu(\mathbf{r},\mathbf{s}',t)\Psi(\mathbf{s},\mathbf{s}')\,d\mathbf{s}' + S_\nu(\mathbf{r},\mathbf{s},t) \quad \forall \mathbf{r}\in\Omega \quad (1)$$

where $I_\nu : (\Omega\times\mathcal{S}^2\times\mathbb{R}) \mapsto \mathbb{R}$, t is the time variable, $\mathbf{r} = (x,y,z)$ is the space variable, $\mathbf{s}$ is the direction, n is the refractive index (constant), c_0 is the speed of light in vacuum, κ, $\sigma : \Omega \mapsto \mathbb{R}^+$ are the absorption and scattering coefficients, respectively, Ψ is the scattering phase function, which describes the probability that a ray from the direction $\mathbf{s}'$ will be scattered

into the direction $\mathbf{s}$, and S_ν represents volumetric source terms such as the spontaneous emission.

The RTE is an integro-differential equation that requires heavy computation to get accurate solutions. Thus, an estimation of optical property maps based on this forward model leads to a very cumbersome inverse problem to solve. The DOT problem consists in using the DA of the RTE as the forward model of the inverse problem [1, 32]:

$$\begin{aligned} &\frac{n}{c_0}\frac{\partial\varphi}{\partial t}(\mathbf{r},t) - \nabla\cdot[D(\mathbf{r})\nabla\varphi(\mathbf{r},t)] + \kappa(\mathbf{r})\varphi(\mathbf{r},t) = 0 \quad \forall\mathbf{r}\in\Omega \\ &\varphi(\zeta,t) + \frac{A}{2\gamma}D(\zeta)\nabla\varphi(\zeta,t)\cdot\mathbf{n} = \frac{f(\zeta,t)}{\gamma} \qquad\qquad \forall\zeta\in\partial\Omega \end{aligned} \tag{2}$$

where $\varphi(\mathbf{r},t) = \int_{\mathcal{S}^2} I_\nu(\mathbf{r},\mathbf{s},t)\,d\mathbf{s}$ is the photon density, $\varphi : (\Omega\times\mathbb{R}) \mapsto \mathbb{R}$, $D(\mathbf{r}) = (n_\Omega(\kappa+\sigma'))^{-1}$ is the macroscopic scattering coefficient, n_Ω is the dimension of Ω, $\sigma' = (1-g)\sigma$ is the reduced scattering coefficient, g is the asymmetry factor of scattering, which equals the average cosine of the scattering angle, $\mathbf{n}$ is the unit normal vector to the surface $\partial\Omega$, γ is a fixed parameter dependent on n_Ω, A is a parameter which characterizes the reflection at the frontier of Ω and f is a diffuse source.

When derivating the DA from the RTE, it is shown that the DA model is a good approximation of the RTE as soon as the medium is highly scattering and satisfies $0 \ll \kappa$, $0 \ll \sigma$ and $\kappa \ll \sigma'$ [1, 32, 33].

Next, using amplitude modulated diffuse source at the frequency ω, f_ω, the DA model in the frequency domain writes:

$$\begin{aligned} &-\nabla\cdot[D(\mathbf{r})\nabla\varphi(\mathbf{r})] + \left[\kappa(\mathbf{r}) + \frac{\mathrm{i}\omega n}{c_0}\right]\varphi(\mathbf{r}) = 0 \quad \forall\mathbf{r}\in\Omega \\ &\varphi(\zeta) + \frac{A}{2\gamma}D(\zeta)\nabla\varphi(\zeta)\cdot\mathbf{n} = \frac{f_\omega(\zeta)}{\gamma} \qquad\qquad \forall\zeta\in\partial\Omega \end{aligned} \tag{3}$$

where $\varphi : \Omega \mapsto \mathbb{C}$.

Equation (3) leads to compute the state variable φ that depends on space-dependent optical properties κ, σ', and which values at the nodes of the sensors will be called predictions. Such modeling is the forward model used in the remainder of this paper which is known to be well-posed mathematically, for example using the complex version of the Lax-Milgram theorem [34].

3. Cost function and its gradients

Let $\partial\Omega_s = \partial\Omega_s^1 \cup \ldots \partial\Omega_s^L$ be a set of L surfacic sensors on $\partial\Omega$ and let us consider K diffuse sources at different locations along $\partial\Omega$. For each source, which constitutes a test, let $\breve{\varphi}_i : \Omega \mapsto \mathbb{C}$ be equal to measurements of the amplitude and phase shift of the transmitted photon density. Predictions φ_i from the forward model (3) and measurements $\breve{\varphi}_i$ are integrated to the cost function to be minimized:

$$j(\kappa, \sigma') = \frac{1}{2} \sum_{i=1}^{K} \int_{\partial\Omega_s} \left| \frac{\varphi_i - \breve{\varphi}_i}{\breve{\varphi}_i} \right|^2 d\zeta \tag{4}$$

The inverse problem can then be formulated as a constrained optimization problem of the form:

$$\text{Find } (\bar{\kappa}, \bar{\sigma'}) = \arg\min_{\kappa,\sigma'} j(\kappa, \sigma') \text{ subject to } \varphi_i \text{ solution of (3)}, \forall i \tag{5}$$

As a gradient-based method will be used to solve the optimization problem (5), gradients of the cost function relative to κ and σ' have to be computed. For such a large scale optimization problem, the gradient has to be computed, for CPU-time considerations, by the very fast adjoint-state method. It can be shown that [1]:

$$\begin{aligned} \nabla_\kappa j &= \sum_{i=1}^{K} \mathrm{Re}\left(\varphi_i \bar{\varphi}_i^* - n_\Omega D^2 \nabla\varphi_i \cdot \nabla\bar{\varphi}_i^*\right) \\ \nabla_{\sigma'} j &= \sum_{i=1}^{K} \mathrm{Re}\left(-n_\Omega D^2 \nabla\varphi_i \cdot \nabla\bar{\varphi}_i^*\right) \end{aligned} \tag{6}$$

where $\mathrm{Re}(u)$ and $\bar{u}$ stand for the real part and complex conjugate of the expression u, respectively, and φ_i^* is the adjoint variable of the state variable φ_i, φ_i^* being solution of the adjoint model:

$$\begin{aligned} &-\nabla \cdot (D(\mathbf{r})\nabla\varphi_i^*(\mathbf{r})) + (\kappa(\mathbf{r}) - \frac{i\omega n}{c_0})\varphi_i^*(\mathbf{r}) = 0 && \forall \mathbf{r} \in \Omega \\ &\frac{2\gamma}{A}\varphi_i^*(\zeta) + D(\zeta)\nabla\varphi_i^*(\zeta) \cdot \mathbf{n} = -\frac{1}{|\breve{\varphi}_i(\zeta)|^2}\left(\varphi_i(\zeta) - \breve{\varphi}_i(\zeta)\right) \mathbb{1}_{[\zeta \in \partial\Omega_s]} && \forall \zeta \in \partial\Omega \end{aligned} \tag{7}$$

where $\mathbb{1}_{[\cdot]}$ is the indicator function. In this paper, the forward and adjoint models (3)-(7) will be solved by the finite element method using the P1-finite

element discretization. The reader is invited to consult [35, 1] for details concerning variational formulations and associated matrix systems to solve (3)-(7) with the finite element method.

The two optical properties to be retrieved from data are different in nature, and their order of magnitude also differ. As a consequence, the cost function gradient parts associated with both optical properties also differ by roughly the same order of magnitude, which is very bad for the convergence. In order to speed-up the iterative convergence to the local minimum, the proposed strategy consists in performing a scaling on the parameters at the beginning of the optimization problem. Choosing an a priori function for both optical properties, say κ_{ap} and σ'_{ap}, one searches parameters that fluctuate about unity. This adimensionalization leads to recover both $\varkappa(\mathbf{x}) = \kappa(\mathbf{x})/\kappa_{ap}(\mathbf{x})$ and $\varsigma(\mathbf{x}) = \sigma'(\mathbf{x})/\sigma'_{ap}(\mathbf{x})$ for which magnitude is of order one approximately for both coefficients. It results that the considered cost function becomes $j(\varkappa, \varsigma)$ and the gradients become:

$$\nabla_{\varkappa} j = \kappa_{ap} \nabla_{\kappa} j \quad ; \quad \nabla_{\varsigma} j = \sigma'_{ap} \nabla_{\sigma'} j \tag{8}$$

4. Optimization

Although lots of algorithms solving DOT problems in the literature are based on the Gauss-Newton method [4, 5, 6, 7], the L-BFGS algorithm is used in this paper. The rationale behind this choice was shown in a previous work by the authors [12]: compared to the Gauss-Newton algorithm, the L-BFGS yields much better reconstructions at lower computational price and is much less sensitive to the ill-posed character of the DOT problem. Indeed, in our applications, regularization may be not compulsory when using the L-BFGS algorithm, although its use may however enhance regularity of the solutions. However, penalization type regularization is absolutely compulsory when considering optimization algorithms that rely on matrix inversion, such as with the Gauss-Newton method.

Generally speaking, two ingredients are important in gradient methods: the descent direction and the step-size. At the iteration m, the descent direction d_m of the L-BFGS method is given by:

$$d_m = -H_m^{-1} g_m \tag{9}$$

where g_m is the gradient of the cost function and H_m^{-1} is computed by the

update formula:

$$H_{m+1}^{-1} = H_m^{-1} + \frac{s_m s_m^\top}{y_m^\top s_m}\left[\frac{y_m^\top H_m^{-1} y_m}{y_m^\top s_m} + I_d\right] - \frac{1}{y_m^\top s_m}\left[s_m y_m^\top H_m^{-1} + H_m^{-1} y_m s_m^\top\right] \tag{10}$$

where $s_m = \left(\kappa_{m+1} - \kappa_m, \sigma'_{m+1} - \sigma'_m\right)$, $y_m = g_{m+1} - g_m$ and H_0 is a positive-definite matrix (identity matrix in this study). According to [36], the matrix H_m^{-1} shall not explicitly be computed in the algorithm for the sake of efficiency. Vectors s_m and y_m are stored during the iterations and allow the computation of the direction update (9).

Being given a descent direction d, the step-size is computed with the help of a line-search method which consists in solving the one-dimensional optimization problem:

$$\alpha^* \approx \arg\min_{\alpha>0} j((\kappa, \sigma') + \alpha d) \tag{11}$$

A comparative study for the choice of the line-search with the L-BFGS method has been realized and presented in [37]. The golden-section search (algorithm 4.2 in [38]), the quadratic interpolation search (algorithm 4.3 in [38]) and an inexact line-search due to Fletcher [39] have been considered. The conclusion is that the quadratic interpolation method is the most efficient, in the proposed DOT applications, among the three tested methods. The implementation of the L-BFGS method used in this paper is presented in Algorithm 1. A maximum iteration number N_M is given in order to ensure that the algorithm stops after a finite number of iterations. It has been observed that a maximum iteration number of a few hundred is a good choice when the algorithm does not end by other stopping criteria ($N_M = 300$ in this paper). In particular, if an higher maximum iteration number is given, the cost function could continue to decrease but the quality of reconstructions will deteriorate. With gradient-based methods such as the L-BFGS, the iteration number plays the role of regularization in the context of Alifanov's iterative regularization [40]. Concerning initial guesses of the optical properties $\gamma^0 = (\varkappa^0, \varsigma^0)$, they are generally determined by a fitting algorithm applied to the data averaged over all sources making the assumption of a homogeneous medium [41]. In this paper, several choices of guessed optical properties will be systematically considered in the numerical results section 6.

Algorithm 1: The L-BFGS algorithm.

Input: $k \leftarrow 1$, $\gamma^0 = (\varkappa^0, \varsigma^0)$, N_M, ϵ_1, ϵ_2;

Compute $j(\gamma^0)$, $\nabla j(\gamma^0)$;
$d^0 \leftarrow -\nabla j(\gamma^0)$;
$\alpha \leftarrow$ Line-Search$(\gamma^0, d^0, j(\gamma^0))$;
$\gamma^1 \leftarrow \gamma^0 + \alpha d^0$;
Compute $j(\gamma^1)$, $\nabla j(\gamma^1)$;
$s(0) \leftarrow \gamma^1 - \gamma^0$; $y(0) \leftarrow \nabla j(\gamma^1) - \nabla j(\gamma^0)$;
while $k \leq N_M$ and $\frac{\|\nabla j(\gamma^k)\|}{\|\nabla j(\gamma^0)\|} > \epsilon_1$ and $\frac{|j(\gamma^k) - j(\gamma^{k-1})|}{j(\gamma^k)} > \epsilon_2$ **do**
  $d^k \leftarrow$ Direction-Update$(s, y, \nabla j(\gamma^k), k)$ following [36];
  $\alpha \leftarrow$ Line-Search$(\gamma^k, d^k, j(\gamma^k))$;
  $\gamma^{k+1} \leftarrow \gamma^k + \alpha d^k$;
  Compute $j(\gamma^{k+1})$, $\nabla j(\gamma^{k+1})$;
  $s(k) \leftarrow \gamma^{k+1} - \gamma^k$; $y(k) \leftarrow \nabla j(\gamma^{k+1}) - \nabla j(\gamma^k)$;
  $k \leftarrow k + 1$;
return γ^k;

5. Wavelet multi-scale method

In the first part of this section, an introduction to the wavelet theory is provided from [42]. Then, in the second part, the strategy used for the development of a wavelet multi-scale method in the context of the proposed inverse problem is presented.

5.1. Orthogonal wavelet bases

Definition 1. *The sequence of closed subspaces of $L^2(\mathbb{R})$, $\{V_j\}_{j\in\mathbb{Z}}$, is a multiresolution approximation if the following 6 properties are satisfied [42]:*

1. $\forall (j,k) \in \mathbb{Z}^2,\ f(x) \in V_j \Longleftrightarrow f(x - 2^j k) \in V_j$
2. $\forall j \in \mathbb{Z},\ V_{j+1} \subset V_j$
3. $\forall j \in \mathbb{Z},\ f(x) \in V_j \Longleftrightarrow f(x/2) \in V_{j+1}$
4. $\lim_{j\to+\infty} V_j = \cap_{j=-\infty}^{+\infty} V_j = \{0\}$
5. $\lim_{j\to-\infty} V_j = \overline{\cup_{j=-\infty}^{+\infty} V_j} = L^2(\mathbb{R})$
6. *There exists θ such that $\{\theta(x-n)\}_{n\in\mathbb{Z}}$ is a Riesz basis of V_0.*

The approximation of f at the scale 2^j is defined as the orthogonal projection $P_{V_j} f$ on V_j. To compute this projection, one has to find an orthonormal basis of V_j for all j. The following theorem constructs this basis from the Riesz basis $\{\theta(x-n)\}_{n\in\mathbb{Z}}$ [42]:

Theorem 1. *Let $\{V_j\}_{j\in\mathbb{Z}}$ be a multiresolution approximation and Φ be the scaling function whose Fourier transform is*

$$\hat{\Phi}(w) = \frac{\hat{\theta}(w)}{\left(\sum_{k=-\infty}^{+\infty} \left|\hat{\theta}(w+2k\pi)\right|^2\right)^{1/2}} \tag{12}$$

Then, the family

$$\left\{\Phi_{j,n}(x) = 2^{-j/2}\Phi(2^{-j}(x-n))\right\}_{n\in\mathbb{Z}} \tag{13}$$

is an orthonormal basis of V_j for all $j \in \mathbb{Z}$.

The conjugate mirror filter h, which plays an important role in the fast wavelet transform algorithm, is now defined, with the inner product on $L^2(\mathbb{R})$ noted $\langle\cdot,\cdot\rangle$:

Definition 2. *Let $\Phi \in L^2(\mathbb{R})$ be an integrable scaling function. The conjugate mirror filter h is defined as*

$$h[n] = \langle 2^{-1/2}\Phi(x/2), \Phi(x-n)\rangle \tag{14}$$

The sequence $h[n]$ may be interpreted as a lowpass discrete filter which smooths the data.

The approximations of f at scales 2^j and 2^{j-1} are equal to their orthogonal projections on V_j and V_{j-1}, respectively. Let us note $P_{V_j} f$ and $P_{V_{j-1}} f$ these projections. As $V_j \subset V_{j-1}$, the orthogonal complement W_j of V_j in V_{j-1} can be considered. Then we have:

$$V_{j-1} = V_j \bigoplus W_j \tag{15}$$

$$\Rightarrow \forall f \in L^2(\mathbb{R}), \quad P_{V_{j-1}} f = P_{V_j} f + P_{W_j} f \tag{16}$$

$$\sum_{n=-\infty}^{+\infty} \langle f, \Phi_{j-1,n}\rangle \Phi_{j-1,n} = \sum_{n=-\infty}^{+\infty} \langle f, \Phi_{j,n}\rangle \Phi_{j,n} + \sum_{n=-\infty}^{+\infty} \langle f, \Psi_{j,n}\rangle \Psi_{j,n} \tag{17}$$

$$\sum_{n=-\infty}^{+\infty} a_{j-1}[n]\Phi_{j-1,n} = \sum_{n=-\infty}^{+\infty} a_j[n]\Phi_{j,n} + \sum_{n=-\infty}^{+\infty} d_j[n]\Psi_{j,n} \tag{18}$$

where $\{\Psi_{j,n}(x)\}_{n\in\mathbb{Z}}$ has to be an orthonormal basis of W_j for all $j \in \mathbb{Z}$. As written in [42]: "the complement $P_{W_j}f$ provides the "details" of f that appear at the scale 2^{j-1} but which disappear at the coarser scale 2^j". The following theorem due to Mallat and Meyer proves that one can construct an orthonormal basis of W_j by scaling and translating a wavelet Ψ [42].

Theorem 2. *Let Φ be a scaling function and h the corresponding conjugate mirror filter. Let Ψ be the function whose Fourier transform is*

$$\hat{\Psi}(w) = 2^{-1/2}\hat{g}(w/2)\hat{\Phi}(w/2) \tag{19}$$

with

$$\hat{g}(w) = e^{-iw}\overline{\hat{h}}(w+\pi) \tag{20}$$

Let us denote

$$\Psi_{j,n}(x) = 2^{-j/2}\Psi(2^{-j}(x-2^j n)) \tag{21}$$

For any scale 2^j, $\{\Psi_{j,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of W_j. For all scales, $\{\Psi_{j,n}\}_{(j,n)\in\mathbb{Z}^2}$ is an orthonormal basis of $L^2(\mathbb{R})$.

It can be shown that the inverse Fourier transform of $\hat{g}$ (see (20)) is equal to:

$$\begin{aligned} g[n] &= \langle 2^{-1/2}\Psi(t/2), \Phi(x-n)\rangle && (22)\\ &= (-1)^{1-n}h[1-n] && (23) \end{aligned}$$

The sequence g is the conjugate mirror filter associated to the wavelet Ψ and may physically be interpreted as a highpass discrete filter which picks out the details of the signal.

Finally, the Mallat's fast orthogonal wavelet transform is given, which allows the fast computation of scaling and wavelet coefficient vectors a_j and d_j (see (18)), respectively, from a_{j-1} [42].

Theorem 3. *At the decomposition, coefficients are computed such that:*

$$a_{j+1}[p] = \sum_{n\in\mathbb{Z}} h[n-2p]a_j[n] \tag{24}$$

$$d_{j+1}[p] = \sum_{n\in\mathbb{Z}} g[n-2p]a_j[n] \tag{25}$$

At the reconstruction, coefficients are computed such that:

$$a_j[p] = \sum_{n\in\mathbb{Z}} h[p-2n]a_{j+1}[n] + \sum_{n\in\mathbb{Z}} g[p-2n]d_{j+1}[n] \tag{26}$$

Although the fast bi-dimensional orthogonal wavelet transform will be used in the multi-scale reconstruction method, the reader is invited to consult [42] to obtain explanations about wavelets in higher dimensions. Generally speaking, multiresolutions in higher dimensions are direct extensions of the one-dimensional case using tensor product spaces.

5.2. Multi-scale reconstruction method

For the sake of simplicity, the wavelet multi-scale method is presented: (i) in the bidimensional case, (ii) with the domain Ω discretized on a grid of $2^{-N} \times 2^{-N}$ points, $N \leq 0$, (iii) with each source and sensor constituted of 2^{-M} pointwise measurements on the grid points, $0 \geq M \geq N$, and (iv) with a P1-finite element discretization for all functions involved in the inverse method in order to have an obvious correspondence between function values on grid points and their corresponding scale coefficients. Indeed, because a subsampling of a factor 2 is required when applying equations (24) and (25), the Mallat's algorithm, Theorem 3, is designed for signal which length is an integer power of two. When it is not the case, algorithms exist to apply the fast wavelet transform, e.g. by adding zeros to the signal so that its length is increased to the next integer power of two [43].

At the beginning of the algorithm, the sources and experimental measurements (or synthetic data) are decomposed on the smallest scale 2^M (i.e. belong to the V_M space). This can be achieved by making a correspondence between localizations of grid points in the mesh and scale coefficients in scale coefficient vectors, and setting the scale coefficients to the values of the functions multiplied by $2^{M/2}$ [42]. Similarly, the initial optical properties are decomposed on the smallest scale 2^N by setting the scale coefficients to the values of the functions multiplied by 2^N (due to the dimension two). Inverse multiplications, respectively by $2^{-M/2}$ and 2^{-N}, are performed to pass from the scale coefficients to the function values at the grid points. The same procedure is considered when the wavelet transform is applied to the forward and adjoint fields.

The wavelet multi-scale method consists in applying the L-BFGS method until convergence, from the largest scale (2^{N-M}) to the smallest one (2^N), with, at a given scale, initializing the optical properties to the solution of the optimizer at the upper scale. As the two main ingredients of the L-BFGS algorithm are the computation of the cost function and its gradient, the way these functions are computed for a given scale is detailed hereafter in Algorithms 2 and 3.

Three wavelets are considered in this paper. The first is the Haar wavelet ($h = [h_0, h_1] = [2^{-1/2}, 2^{-1/2}]$), the second is the Daubechies D4 wavelet ($h = [h_i]_{i=0}^{3}$) and the third is the Daubechies D6 wavelet ($h = [h_i]_{i=0}^{5}$). The coefficients of the conjugate mirror filters h for the D4 and D6 wavelets are not given here but can be found in [42] for instance. It should be noted that when the Haar wavelet is used, no particular attention has to be given when applying the fast wavelet transform. However, an edge problem appears when considering equations (24), (25) and (26) with a finite length sequence a_j and a wavelet for which the conjugate mirror filter length is greater than two. Among the known and widely used approaches to deal with this problem, let us mention the extension by zeros, the periodic extension, and the symmetric extension [42]. Let us note that all of these methods can occur undesirable edge effects in the transformed signal. These three approaches have been implemented and tested applying the wavelet multi-scale method with the Daubechies D4 and D6 wavelets. It has been observed that the extension by zeros and the periodic extension provide very poor results when compared to the symmetric extension. The latter approach was therefore chosen.

Algorithm 2: Computation of the cost function at scale 2^{N-l}, $M \leq l < 0$.

Input: Experimental measurements at scale 2^{M-l};

for $i \leftarrow 1$ **to** K **do**

- Compute the state variable φ_i at the smallest scale 2^N;
- Extract the complex values of the photon density φ_i at the nodes of the sensors, for each sensor $\partial\Omega_s^1, \ldots, \partial\Omega_s^L$, to obtain L vectors of length 2^{-M};
- Perform the fast wavelet transform following equation (24) from scale 2^M to scale 2^{M-l}, for each one of the L vectors previously built;
- Add the contribution of the test i to the cost function at scale 2^{N-l}, computing equation (4) with experimental measurements and values of the state variable at the nodes of the sensors at the scale 2^{M-l};

return *The value of the cost function at scale* 2^{N-l};

Algorithm 3: Computation of the gradient of the cost function at scale 2^{N-l}, $M \leq l < 0$.

Input: Experimental measurements at scale 2^{M-l};

for $i \leftarrow 1$ **to** K **do**

Compute the state variable φ_i at the smallest scale 2^N;
Compute the source term of the adjoint model (right term of equation (7)) at scale 2^{M-l}, following an analog procedure of steps 2 and 3 of the cost function computation algorithm;
Compute the adjoint state variable φ_i^* at the smallest scale 2^N;
Perform the fast wavelet transform of φ_i, φ_i^*, $\partial_x(\varphi_i)$, $\partial_y(\varphi_i)$, $\partial_x(\varphi_i^*)$ and $\partial_y(\varphi_i^*)$ from scale 2^N to scale 2^{N-l};
Add the contribution of the test i to the gradient of the cost function at scale 2^{N-l}, computing equation (6) with the help of φ_i, φ_i^*, $\partial_x(\varphi_i)$, $\partial_y(\varphi_i)$, $\partial_x(\varphi_i^*)$ and $\partial_y(\varphi_i^*)$ previously computed at the scale 2^{N-l};

return *The gradient of the cost function at scale* 2^{N-l};

6. Numerical results

6.1. Tests presentation

The proposed method has been implemented in the freefem++ environment [44] and some results are presented in this section. The wavelet multi-scale method is applied and compared with the ordinary L-BFGS method to the reconstruction of two bidimensional optical property maps, $\kappa(\mathbf{r})$ and $\sigma'(\mathbf{r})$, where the domain considered, Ω, is a square of 4 cm in length ($\Omega = [-2, 2]^2$). The target properties to be reconstructed, κ_t and σ'_t, are defined by

$$\begin{array}{llll} \kappa(\boldsymbol{r}) = 0.1\,\mathrm{cm}^{-1}, & \sigma'(\boldsymbol{r}) = 30\,\mathrm{cm}^{-1}, & \boldsymbol{r} \in \Omega_1, & \Omega_1 = [-1.5, -0.5]^2 \\ \kappa(\boldsymbol{r}) = 0.06\,\mathrm{cm}^{-1}, & \sigma'(\boldsymbol{r}) = 10\,\mathrm{cm}^{-1}, & \boldsymbol{r} \in \Omega_2, & \Omega_2 = [0.5, 1.5]^2 \\ \kappa(\boldsymbol{r}) = 0.08\,\mathrm{cm}^{-1}, & \sigma'(\boldsymbol{r}) = 20\,\mathrm{cm}^{-1}, & \boldsymbol{r} \in \Omega \setminus (\Omega_1 \cup \Omega_2) & \end{array} \tag{27}$$

for the first test, and by

$$\begin{array}{lll} \kappa(\boldsymbol{r}) = 0.08 + 0.06\cos(\sqrt{\frac{x^2+y^2}{2}})\mathrm{cm}^{-1}, & \sigma'(\boldsymbol{r}) = 20 + 15\cos(\sqrt{\frac{x^2+y^2}{2}})\mathrm{cm}^{-1}, & \boldsymbol{r} \in \Omega_3 \\ \kappa(\boldsymbol{r}) = 0.08\,\mathrm{cm}^{-1}, & \sigma'(\boldsymbol{r}) = 20\,\mathrm{cm}^{-1}, & \boldsymbol{r} \in \Omega \setminus \Omega_3 \end{array} \tag{28}$$

where $\Omega_3 = \left\{(x,y) \in \mathbb{R}^2 \text{ such that } x^2+y^2 \leq \left(\frac{\pi}{2}\right)^2\right\}$, for the second test. The regular spatial discretization associated to Ω is composed of 128×128 points ($N = -7$). On each side of the square, one source and two sensors are located. Let us note $\varepsilon = 1/127$. Then, with respect to ε, the source is located on the interval $[-2+56\times\varepsilon, -2+71\times\varepsilon]$ while the two sensors are respectively located on intervals $[-2+21\times\varepsilon, -2+36\times\varepsilon]$ and $[-2+91\times\varepsilon, -2+106\times\varepsilon]$ ($M = -4$).

The physical parameters involved in (3) are fixed to: $n = 1.4$, $\omega = 100 \times 10^6$ s^{-1}, $c_0 = 3 \times 10^8$ m.s^{-1}, $f_\omega = 1$ W.m^{-2}, $\gamma = \pi^{-1}$ and A is derived from Fresnel's law (see [6]). Synthetic data are considered for these numerical tests. These data, $\breve{\varphi}$, representing the pseudo-experimental measurements, are built on a finer mesh (255×255 points) than that of generating the predictions φ in order to avoid the inverse crime [45, 16]. Then, a multiplicative white gaussian noise is applied to $\breve{\varphi}$ at the nodes of the sensors to simulate the noise inherent to experimental devices, i.e $\breve{\varphi}_{\text{noisy}} = \breve{\varphi}\,(1 + 0.01 \times m)$ where $m \sim \mathcal{N}(0,1)$.

The fast wavelet transform is applied with the Haar, D4 or D6 wavelets to decompose the sources and synthetic data at scales $-3, \ldots, 0$ and thus, to obtain a series of inverse problems for scales $-7, \ldots, -3$ following Algorithms 2 and 3. Concerning the comparison of the wavelet multi-scale method and the ordinary L-BFGS method, the choice was made to allow the same total number of iterations for the two methods. Moreover, larger is the scale, faster is the convergence of the solution toward the local minimum, and thus the allowed maximal number of iterations will be reduced. Therefore, the inputs of the ordinary L-BFGS method, Algorithm 1, are the following: $N_M = 300$, $\epsilon_1 = 10^{-4}$ and $\epsilon_2 = 10^{-8}$. The same inputs are considered for the wavelet multi-scale method, except the maximum number of iterations which is fixed to $N_M = 20$, 40, 60, 80 and 100 for scales -3, -4, -5, -6 and 7, respectively.

6.2. Definition of reconstruction errors

In order to gauge accuracy of the reconstructions, two errors are introduced. Let ϑ^j and ϑ_t^j be either the reconstructed and target absorption or

reduced scattering coefficients at the node j, respectively, and $N_s = 128 \times 128$ be the dimension of the state and control spaces. The first error is the deviation factor, e_1^{ϑ}, which measures the deviation of the reconstructed image from the target image and is defined as:

$$e_1^{\vartheta} = 100 \times \frac{\sqrt{\sum_{j=1}^{N_s} \left(\vartheta^j - \vartheta_t^j\right)^2}}{\sqrt{\sum_{j=1}^{N_s} \left(\vartheta_t^j\right)^2}} \tag{29}$$

A small value of e_1^{ϑ} indicates a reconstructed image with high accuracy. The second error is the correlation coefficient [46, 8], e_2^{ϑ}, which measures the linear correlation between the target and the reconstructed image and is defined as:

$$e_2^{\vartheta} = \frac{\sum_{j=1}^{N_s} \left(\vartheta^j - \bar{\vartheta}\right)\left(\vartheta_t^j - \bar{\vartheta}_t\right)}{(N_s - 1)\sigma_{\vartheta}\sigma_{\vartheta_t}} \tag{30}$$

where $\bar{\vartheta}$ 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, given by

$$\sigma_{\vartheta} = \sqrt{\frac{\sum_{j=1}^{N_s} \left(\vartheta^j - \bar{\vartheta}\right)^2}{N_s - 1}} \tag{31}$$

and

$$\sigma_{\vartheta_t} = \sqrt{\frac{\sum_{j=1}^{N_s} \left(\vartheta_t^j - \bar{\vartheta}_t\right)^2}{N_s - 1}} \tag{32}$$

A large value of e_2^{ϑ} shows a high detectability in the reconstructed image and indicates a reconstructed image with high accuracy.

6.3. Test 1: reconstruction of discontinuous target properties

For this first test, several reconstructions of target properties, eq. (27), are investigated depending on initial guesses for the optical properties. The control variables $\varkappa(\mathbf{r})$ and $\varsigma(\mathbf{r})$ are initialized to unity while values taken by a priori functions κ_{ap} and σ'_{ap} are fixed to those of the background and in the range of more of less ten percent of those of the background. Tables 1, 2, 3 and 4 give errors in matrix form $\begin{pmatrix} e_1^{\kappa} & e_1^{\sigma'} \\ e_2^{\kappa} & e_2^{\sigma'} \end{pmatrix}$, for all considered initial guesses, obtained at the end of the inversion with the ordinary L-BFGS method or the wavelet multi-scale method using the Haar, D4 or D6 wavelets. On each

$\kappa_{ap} \setminus \sigma'_{ap}$	18	20	22
0.072	$\begin{pmatrix} 7.37 & 12.92 \\ 0.73 & 0.71 \end{pmatrix}$	$\begin{pmatrix} 7.38 & 11.65 \\ 0.68 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.48 & 11.31 \\ 0.63 & 0.77 \end{pmatrix}$
0.08	$\begin{pmatrix} 5.27 & 12.12 \\ 0.80 & 0.74 \end{pmatrix}$	$\begin{pmatrix} 5.23 & 10.80 \\ 0.81 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.46 & 12.07 \\ 0.80 & 0.74 \end{pmatrix}$
0.088	$\begin{pmatrix} 7.70 & 12.16 \\ 0.74 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.64 & 11.70 \\ 0.74 & 0.75 \end{pmatrix}$	$\begin{pmatrix} 7.26 & 12.77 \\ 0.73 & 0.70 \end{pmatrix}$

Table 1: Errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix}$ for the test 1 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the ordinary L-BFGS method.

table the values of κ_{ap} are in the first column (0.072-0.088) and those for σ'_{ap} are along the first row (18-22). The other cells contain errors. Results show that the wavelet multi-scale method using the Haar wavelet outperforms all of the other methods, errors e_1^κ and $e_1^{\sigma'}$ being smaller and errors e_2^κ and $e_2^{\sigma'}$ being larger for all the investigated cases. Overall, it is observed that the wavelet multi-scale method using the D6 wavelet outperforms the ordinary L-BFGS method while the wavelet multi-scale method using the D4 wavelet outperforms most of the time the wavelet multi-scale method using the D6 wavelet. Figure 1 presents the target properties and the obtained reconstructions with the four proposed inverse methods for initial guesses $\kappa_{ap} = 0.08$ and $\sigma'_{ap} = 20$. It is seen that the wavelet multi-scale method using the Haar wavelet allows to access better reconstructions compared to the three other methods. Specifically, two major improvements should be noted. Firstly, it is observed that the wavelet multi-scale method using the Haar wavelet reduces significatively the presence of inclusions at the edge of the domain Ω that have no reason to exist, unlike what is found for the ordinary L-BFGS method (Figure 1-(d)). Secondly, one notes that the location and shape of inclusions are more accurately determined, unlike images obtained with the ordinary L-BFGS method, which, as in classical approaches, are blurred [17, 14]. These improvements are also observed to a lesser degree from reconstructions based on the wavelet multi-scale method using the D4 and D6 wavelets when compared to those of the ordinary L-BFGS method.

6.4. Test 2: reconstruction of continuous target properties

In this section, in order to more deeply compare the ordinary L-BFGS and wavelet multi-scale methods, reconstructions of continuously varying prop-

$\kappa_{ap} \setminus \sigma'_{ap}$	18	20	22
0.072	$\begin{pmatrix} 6.02 & 10.64 \\ 0.78 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 6.63 & 9.77 \\ 0.75 & 0.83 \end{pmatrix}$	$\begin{pmatrix} 7.17 & 9.97 \\ 0.71 & 0.82 \end{pmatrix}$
0.08	$\begin{pmatrix} 4.76 & 10.13 \\ 0.84 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 4.77 & 9.61 \\ 0.84 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 4.89 & 9.76 \\ 0.84 & 0.83 \end{pmatrix}$
0.088	$\begin{pmatrix} 6.73 & 9.91 \\ 0.77 & 0.83 \end{pmatrix}$	$\begin{pmatrix} 6.16 & 9.41 \\ 0.78 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 5.92 & 10.13 \\ 0.78 & 0.81 \end{pmatrix}$

Table 2: Errors $\begin{pmatrix} e_1^{\kappa} & e_1^{\sigma'} \\ e_2^{\kappa} & e_2^{\sigma'} \end{pmatrix}$ for the test 1 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the Haar wavelet.

$\kappa_{ap} \setminus \sigma'_{ap}$	18	20	22
0.072	$\begin{pmatrix} 6.47 & 12.24 \\ 0.75 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 7.20 & 10.70 \\ 0.75 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.53 & 10.66 \\ 0.71 & 0.80 \end{pmatrix}$
0.08	$\begin{pmatrix} 5.33 & 11.29 \\ 0.81 & 0.78 \end{pmatrix}$	$\begin{pmatrix} 5.02 & 10.06 \\ 0.82 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 5.20 & 10.88 \\ 0.81 & 0.79 \end{pmatrix}$
0.088	$\begin{pmatrix} 7.19 & 10.56 \\ 0.75 & 0.81 \end{pmatrix}$	$\begin{pmatrix} 6.54 & 10.01 \\ 0.74 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 6.21 & 10.80 \\ 0.76 & 0.79 \end{pmatrix}$

Table 3: Errors $\begin{pmatrix} e_1^{\kappa} & e_1^{\sigma'} \\ e_2^{\kappa} & e_2^{\sigma'} \end{pmatrix}$ for the test 1 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the D4 wavelet.

$\kappa_{ap} \setminus \sigma'_{ap}$	18	20	22
0.072	$\begin{pmatrix} 6.47 & 12.18 \\ 0.73 & 0.73 \end{pmatrix}$	$\begin{pmatrix} 6.67 & 10.74 \\ 0.73 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.43 & 12.43 \\ 0.71 & 0.74 \end{pmatrix}$
0.08	$\begin{pmatrix} 5.39 & 10.67 \\ 0.80 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.25 & 10.75 \\ 0.81 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 5.40 & 11.54 \\ 0.80 & 0.75 \end{pmatrix}$
0.088	$\begin{pmatrix} 7.88 & 10.82 \\ 0.71 & 0.79 \end{pmatrix}$	$\begin{pmatrix} 7.55 & 11.48 \\ 0.73 & 0.76 \end{pmatrix}$	$\begin{pmatrix} 7.34 & 13.88 \\ 0.66 & 0.62 \end{pmatrix}$

Table 4: Errors $\begin{pmatrix} e_1^{\kappa} & e_1^{\sigma'} \\ e_2^{\kappa} & e_2^{\sigma'} \end{pmatrix}$ for the test 1 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the D6 wavelet.

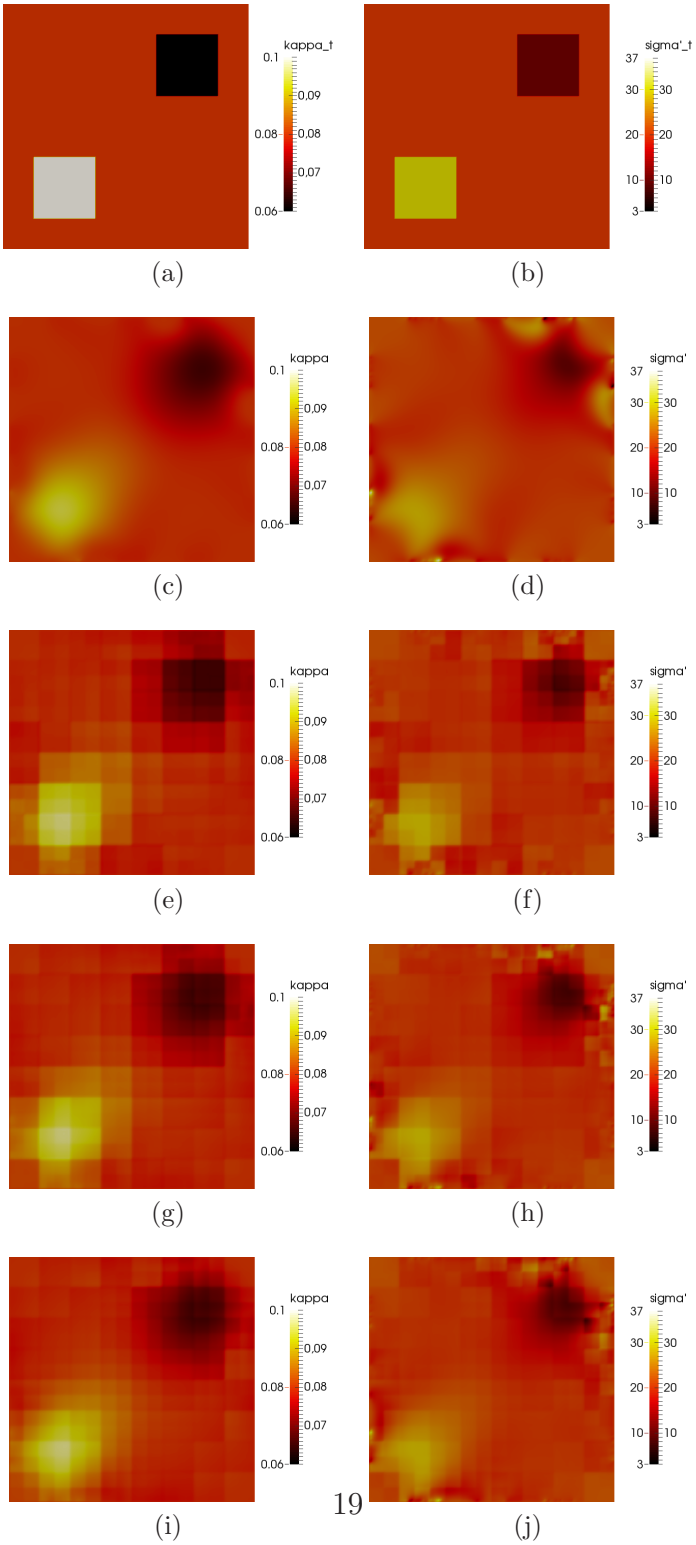

Figure 1: Test 1. Reconstruction of optical properties with $\kappa_{ap} = 0.08$ and $\sigma'_{ap} = 20$. Left-hand side, results for κ; right-hand side, results for σ'. (a-b) target properties; (c-d) ordinary L-BFGS method; (e-f) Haar wavelet; (g-h) D4 wavelet; (i-j) D6 wavelet.

erties (28) which occupy a circle-shaped subdomain in Ω are investigated. Indeed, wavelet multi-scale methods gave better results for the test 1 partly because the original targets contained rectangular inclusions. But the bidimensional wavelet transform has the property to extract vertical and horizontal edges very well. Especially, because the bidimensional Haar scaling function and the inclusions have similar shapes. For this second test, a priori functions κ_{ap} and σ'_{ap} are fixed to those of the background ($\kappa_{ap} = 0.08$, $\sigma'_{ap} = 20$), to approximately the mean values of target properties ($\kappa_{ap} = 0.1$, $\sigma'_{ap} = 25$) or between these two values ($\kappa_{ap} = 0.09$, $\sigma'_{ap} = 22.5$). Tables 5, 6, 7 and 8 give errors in matrix form, for all considered initial guesses, obtained at the end of the inversion with the ordinary L-BFGS method and the wavelet multi-scale method using the Haar, D4 or D6 wavelets. First, it is observed that all the tests converged for the wavelet multi-scale method using the Haar wavelet while one and five tests diverged for the wavelet multi-scale method using the D4 and D6 wavelets, respectively, and two tests diverged for the ordinary L-BFGS method. Second, it is observed that reconstructions of the reduced scattering coefficient with the wavelet multi-scale method using the Haar wavelet are much more accurate than reconstructions based on the ordinary L-BFGS method. It is also observed that reconstructions of the absorption coefficient with the wavelet multi-scale method using the Haar wavelet are much more accurate than reconstructions based on the ordinary L-BFGS method in term of the error e_1^κ (except the test $\kappa_{ap} = 0.09$, $\sigma'_{ap} = 22.5$ which gives a close result) but slightly less accurate in term of the error e_2^κ. It is found that the wavelet multi-scale method using the Haar wavelet clearly outperforms the wavelet multi-scale method using the D4 wavelet and that, overall, the latter gives better results than the wavelet multi-scale method using the D6 wavelet. The Figure 2 presents the target properties and the obtained reconstructions with the four proposed inverse methods for initial guesses $\kappa_{ap} = 0.1$ and $\sigma'_{ap} = 22.5$. Although this is one of the worst results for reconstructions based on the Haar wavelet, it is observed that the latter gives few false fluctuations at the edge of the domain when compared to the other methods, specially for the reduced scattering coefficient. The Figure 3 presents the target properties and reconstructions obtained with the wavelet multi-scale method using the Haar wavelet for the two divergent cases of the ordinary L-BFGS method. It is seen that accurate reconstructions are obtained. Finally, in order to verify that the ordinary L-BFGS method would not give better results if other stopping criteria were chosen, the reconstruction process was stopped just before the divergence of the al-

$\kappa_{ap} \setminus \sigma'_{ap}$	20	22.5	25
0.08	$\times$	$\times$	$\begin{pmatrix} 12.58 & 15.10 \\ 0.97 & 0.85 \end{pmatrix}$
0.09	$\begin{pmatrix} 10.59 & 14.13 \\ 0.97 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 7.79 & 12.50 \\ 0.97 & 0.85 \end{pmatrix}$	$\begin{pmatrix} 10.66 & 12.41 \\ 0.97 & 0.86 \end{pmatrix}$
0.1	$\begin{pmatrix} 15.49 & 14.18 \\ 0.97 & 0.85 \end{pmatrix}$	$\begin{pmatrix} 10.03 & 8.80 \\ 0.94 & 0.92 \end{pmatrix}$	$\begin{pmatrix} 11.35 & 10.73 \\ 0.95 & 0.88 \end{pmatrix}$

Table 5: Errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix}$ for the test 2 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the ordinary L-BFGS method. A cross indicates the divergence of the algorithm.

$\kappa_{ap} \setminus \sigma'_{ap}$	20	22.5	25
0.08	$\begin{pmatrix} 6.81 & 7.25 \\ 0.96 & 0.95 \end{pmatrix}$	$\begin{pmatrix} 6.89 & 8.49 \\ 0.96 & 0.94 \end{pmatrix}$	$\begin{pmatrix} 7.14 & 10.98 \\ 0.96 & 0.91 \end{pmatrix}$
0.09	$\begin{pmatrix} 8.17 & 6.84 \\ 0.95 & 0.96 \end{pmatrix}$	$\begin{pmatrix} 7.94 & 7.39 \\ 0.96 & 0.95 \end{pmatrix}$	$\begin{pmatrix} 7.98 & 10.25 \\ 0.96 & 0.91 \end{pmatrix}$
0.1	$\begin{pmatrix} 9.89 & 6.56 \\ 0.92 & 0.96 \end{pmatrix}$	$\begin{pmatrix} 9.88 & 7.60 \\ 0.93 & 0.94 \end{pmatrix}$	$\begin{pmatrix} 9.99 & 9.93 \\ 0.94 & 0.91 \end{pmatrix}$

Table 6: Errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix}$ for the test 2 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the Haar wavelet.

gorithm for these two cases. For instance, errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix} = \begin{pmatrix} 15.32 & 19.00 \\ 0.92 & 0.53 \end{pmatrix}$ were obtained for the case $\kappa_{ap} = 0.08$ and $\sigma'_{ap} = 20$, which indicates much less accurate reconstructions than the wavelet multi-scale method using the Haar wavelet.

6.5. Discussion

In conclusion, one can state from these two numerical studies that: (i) the wavelet multi-scale method using the Haar wavelet brings more stability with respect to the ordinary L-BFGS method, (ii) the wavelet multi-scale method using the Haar wavelet provides the best reconstructions in most of the tested cases, (iii) the wavelet multi-scale method using the D4 wavelet brings a little more stability with respect to the ordinary L-BFGS method

$\kappa_{ap} \setminus \sigma'_{ap}$	20	22.5	25
0.08	$\times$	$\begin{pmatrix} 8.66 & 14.56 \\ 0.95 & 0.86 \end{pmatrix}$	$\begin{pmatrix} 9.62 & 15.79 \\ 0.95 & 0.86 \end{pmatrix}$
0.09	$\begin{pmatrix} 8.70 & 10.28 \\ 0.95 & 0.90 \end{pmatrix}$	$\begin{pmatrix} 8.52 & 11.35 \\ 0.95 & 0.89 \end{pmatrix}$	$\begin{pmatrix} 8.75 & 12.55 \\ 0.95 & 0.87 \end{pmatrix}$
0.1	$\begin{pmatrix} 11.01 & 9.56 \\ 0.92 & 0.91 \end{pmatrix}$	$\begin{pmatrix} 12.11 & 10.58 \\ 0.94 & 0.89 \end{pmatrix}$	$\begin{pmatrix} 10.94 & 11.38 \\ 0.93 & 0.88 \end{pmatrix}$

Table 7: Errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix}$ for the test 2 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the D4 wavelet. A cross indicates the divergence of the algorithm.

$\kappa_{ap} \setminus \sigma'_{ap}$	20	22.5	25
0.08	$\times$	$\times$	$\times$
0.09	$\begin{pmatrix} 12.55 & 16.03 \\ 0.94 & 0.72 \end{pmatrix}$	$\begin{pmatrix} 11.47 & 14.31 \\ 0.95 & 0.77 \end{pmatrix}$	$\times$
0.1	$\times$	$\begin{pmatrix} 11.51 & 12.36 \\ 0.95 & 0.84 \end{pmatrix}$	$\begin{pmatrix} 12.83 & 13.04 \\ 0.96 & 0.82 \end{pmatrix}$

Table 8: Errors $\begin{pmatrix} e_1^\kappa & e_1^{\sigma'} \\ e_2^\kappa & e_2^{\sigma'} \end{pmatrix}$ for the test 2 and different values of the a priori functions κ_{ap} and σ'_{ap} (cm^{-1}) with the wavelet multi-scale method using the D6 wavelet. A cross indicates the divergence of the algorithm.

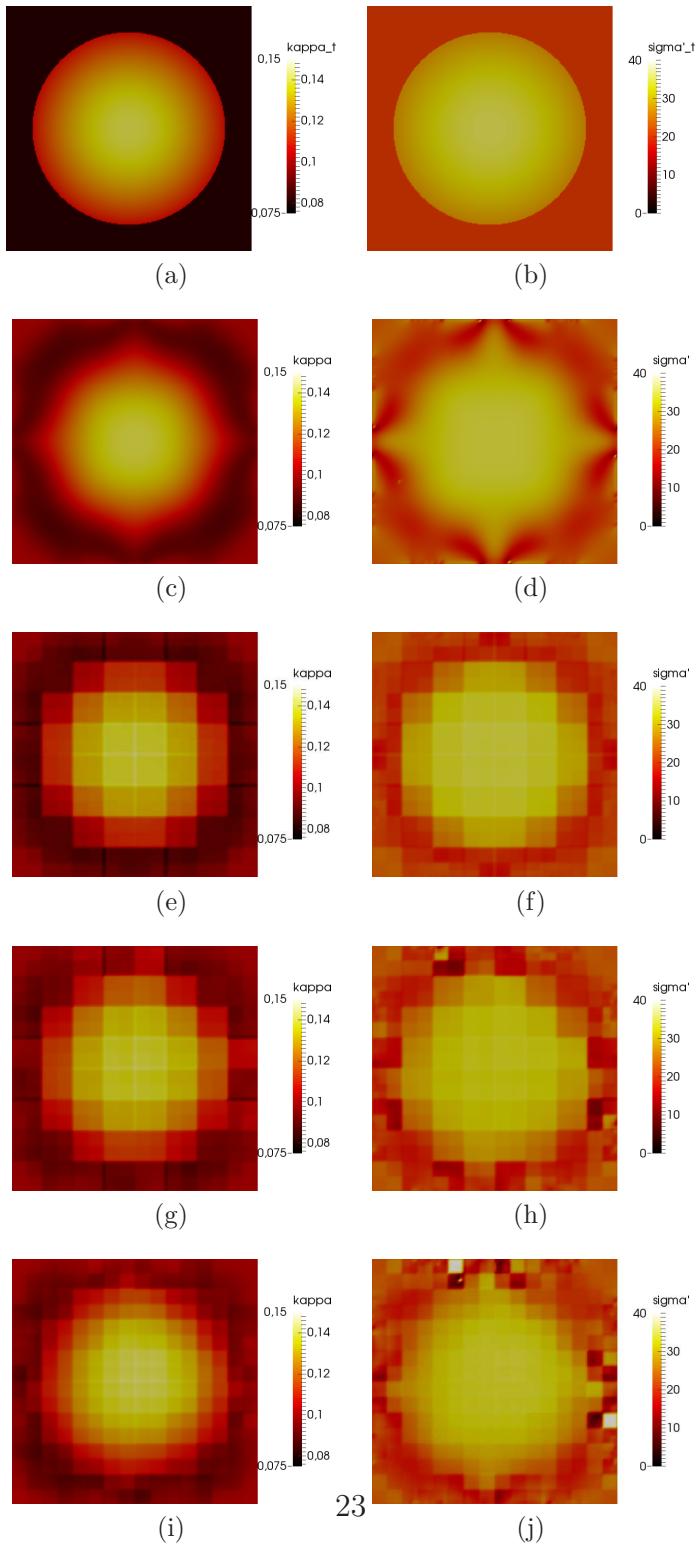

Figure 2: Test 2. Reconstruction of optical properties with $\kappa_{ap} = 0.1$ and $\sigma'_{ap} = 22.5$. Left-hand side, results for κ; right-hand side, results for σ'. (a-b) target properties; (c-d) ordinary L-BFGS method; (e-f) Haar wavelet; (g-h) D4 wavelet; (i-j) D6 wavelet.

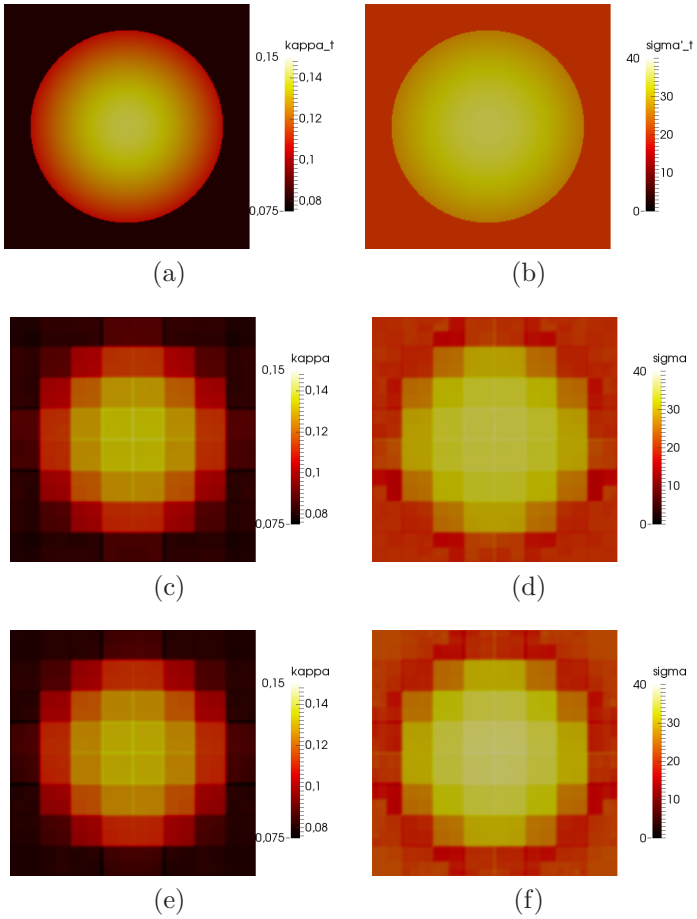

Figure 3: Test 2. Reconstruction of optical properties with the Haar wavelet. Left-hand side, results for κ; right-hand side, results for σ'. (a-b) target properties; (c-d) $\kappa_{ap} = 0.08$ and $\sigma'_{ap} = 20$; (e-f) $\kappa_{ap} = 0.08$ and $\sigma'_{ap} = 22.5$.

while the wavelet multi-scale method using the D6 wavelet is the least stable among the four proposed inverse methods, and (iv) the wavelet multi-scale method using the D4 and D6 wavelets does not really enhance the quality of the reconstructions when compared to the ordinary L-BFGS method. Finally, let us note that the reason why the wavelet multi-scale method using the D4 and D6 wavelets fail to give better results should come from undesirable edge effects in the transformed sequences and images. Indeed, it is of crucial importance to not exhibit edge effects in the one-dimensional transformed experimental measurements (or synthetic data) and adjoint source terms as well as in the bidimensional transformed forward and adjoint fields, the latter being maximal at the edges of the domain. Thus, it might be interesting to consider a discrete wavelet transform without edge effects using wavelet extrapolation as described in [47].

7. Conclusion

A wavelet multi-scale method based on the L-BFGS algorithm using the Haar, Daubechies D4 and D6 wavelets has been designed, implemented and validated for the two-dimensional inverse problem of diffuse optical tomography in the frequency domain. The numerical results indicate the effectiveness of the method using the Haar wavelet with respect to the ordinary L-BFGS method. In particular, a reduction of the blurring effect has been observed in a numerical test containing rectangular inclusions and a strong stability of the latter method has been showed on a numerical test containing continuously varying target properties which occupy a circle-shaped subdomain. Upcoming work includes an extension of the method to the three-dimensional case and/or in the time domain, together with the use of a discrete wavelet transform without edge effects using wavelet extrapolation in order to improve obtained results with the Daubechies D4 and D6 wavelets. The application of the proposed method to the inversion of the full radiative transfer equation is also considered.

Acknowledgements

The authors would like to thank the t3e chair industrial partners for their financial support to the industrial research chair t3e. The University of Nantes is also thanked for funding in part the stay of the corresponding author at the conference ACOMEN. The authors also 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, allowing computations be performed on their supercomputers. Finally, the authors would like to sincerely thank the reviewers for their constructive and relevant comments formulated on the first version that significantly enhanced the quality of this paper.

[1] S. R. Arridge, Optical tomography in medical imaging, Inverse problems 15 (2) (1999) R41.

[2] A. Gibson, J. Hebden, S. R. Arridge, Recent advances in diffuse optical imaging, Physics in medicine and biology 50 (4) (2005) R1.

[3] S. Arridge, M. Schweiger, A gradient-based optimisation scheme for optical tomography, Optics Express 2 (6) (1998) 213–226.

[4] M. Schweiger, S. R. Arridge, I. Nissilä, Gauss–Newton method for image reconstruction in diffuse optical tomography, Physics in medicine and biology 50 (10) (2005) 2365.

[5] H. Niu, P. Guo, L. Ji, Q. Zhao, T. Jiang, Improving image quality of diffuse optical tomography with a projection-error-based adaptive regularization method, Optics express 16 (17) (2008) 12423–12434.

[6] H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, K. D. Paulsen, Near infrared optical tomography using NIRFAST: Algorithm for numerical model and image reconstruction, Communications in numerical methods in engineering 25 (6) (2009) 711–732.

[7] T. Tarvainen, V. Kolehmainen, S. Arridge, J. Kaipio, Image reconstruction in diffuse optical tomography using the coupled radiative transport-diffusion model, J. Quant. Spectrosc. Radiat. Transfer 112 (2011) 2600–2608.

[8] A. D. Klose, A. H. Hielscher, Quasi-newton methods in optical tomographic image reconstruction, Inverse Problems 19 (2) (2003) 387.

[9] K. Ren, G. Bal, A. H. Hielscher, Frequency domain optical tomography based on the equation of radiative transfer, SIAM Journal on Scientific Computing 28 (4) (2006) 1463–1489.

[10] O. Balima, J. Boulanger, A. Charette, 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 112 (7) (2011) 1235–1240.

[11] O. Balima, Y. Favennec, D. Rousse, Optical tomography reconstruction algorithm with the finite element method: An optimal approach with regularization tools, Journal of Computational Physics 251 (2013) 461–479.

[12] Y. Favennec, F. Dubot, B. Rousseau, D. R. Rousse, Mixing regularization tools for enhancing regularity in optical tomography applications, in: O. Fudym, J. L. Battaglia, G. S. Dulikravich (Eds.), IPDO 2013 : 4th Inverse problems, design and optimization symposium, Albi, 2013.

[13] A. D. Zacharopoulos, S. R. Arridge, O. Dorn, V. Kolehmainen, J. Sikora, Three-dimensional reconstruction of shape and piecewise constant region values for optical tomography using spherical harmonic parametrization and a boundary element method, Inverse Problems 22 (5) (2006) 1509.

[14] S. R. Arridge, O. Dorn, V. Kolehmainen, M. Schweiger, A. Zacharopoulos, Parameter and structure reconstruction in optical tomography, in: Journal of Physics: Conference Series, Vol. 135, IOP Publishing, 2008, p. 012001.

[15] S. Arridge, J. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, M. Vauhkonen, Approximation errors and model reduction with an application in optical diffusion tomography, Inverse Problems 22 (1) (2006) 175.

[16] J. Kaipio, E. Somersalo, Statistical and computational inverse problems, Vol. 160, Springer, 2005.

[17] N. Cao, A. Nehorai, M. Jacobs, Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm, Optics express 15 (21) (2007) 13695–13708.

[18] J. Liu, A multiresolution method for distributed parameter estimation, SIAM Journal on Scientific Computing 14 (2) (1993) 389–405.

[19] L. Ding, B. Han, J.-q. Liu, A wavelet multiscale method for inversion of maxwell equations, Applied Mathematics and Mechanics 30 (2009) 1035–1044.

[20] X. Zhang, K. Liu, J. Liu, The wavelet multiscale method for inversion of porosity in the fluid-saturated porous media, Applied mathematics and computation 180 (2) (2006) 419–427.

[21] Y. He, B. Han, A wavelet adaptive-homotopy method for inverse problem in the fluid-saturated porous media, Applied Mathematics and Computation 208 (1) (2009) 189–196.

[22] H. Fu, B. Han, G. Gai, A wavelet multiscale–homotopy method for the inverse problem of two-dimensional acoustic wave equation, Applied mathematics and computation 190 (1) (2007) 576–582.

[23] J. Lei, S. Liu, Z. Li, M. Sun, X. Wang, A multi-scale image reconstruction algorithm for electrical capacitance tomography, Applied Mathematical Modelling 35 (6) (2011) 2585–2606.

[24] H. Fu, B. Han, H. Liu, A wavelet multiscale iterative regularization method for the parameter estimation problems of partial differential equations, Neurocomputing 104 (0) (2013) 138–145.

[25] J. Zhao, T. Liu, S. Liu, Identification of space-dependent permeability in nonlinear diffusion equation from interior measurements using wavelet multiscale method, Inverse Problems in Science and Engineering 22 (4) (2014) 507–529.

[26] S. K. Nath, R. Vasu, M. Pandit, Wavelet based compression and denoising of optical tomography data, Optics communications 167 (1) (1999) 37–46.

[27] W. Zhu, Y. Wang, Y. Deng, Y. Yao, R. L. Barbour, A wavelet-based multiresolution regularized least squares reconstruction approach for optical tomography, Medical Imaging, IEEE Transactions on 16 (2) (1997) 210–217.

[28] W. Zhu, Y. Wang, Y. Yao, J. Chang, H. L. Graber, R. L. Barbour, Iterative total least-squares image reconstruction algorithm for optical

tomography by the conjugate gradient method, JOSA A 14 (4) (1997) 799–807.

[29] W. Zhu, Y. Wang, J. Zhang, Total least-squares reconstruction with wavelets for optical tomography, JOSA A 15 (10) (1998) 2639–2650.

[30] M. Modest, Radiative Heat Transfer, McGraw-Hill, Inc., 2003.

[31] J. Howell, R. Siegel, M. Mengüç, Thermal Radiation Heat Transfer, 5th Edition, CRC Press, 2011.

[32] T. Tarvainen, Computational methods for light transport in optical tomography, Ph.D. thesis, University of Kuopio (2006).

[33] A. Klose, U. Netz, J. Beuthan, A. Hielscher, Optical tomography using the time-independent equation of radiative transfer - part 1 : forward model, J. Quant. Spectrosc. Radiat. Transfer 72 (2002) 691–713.

[34] V. Brattka, A. Yoshikawa, Towards computability of elliptic boundary value problems in variational formulation, Journal of Complexity 22 (6) (2006) 858–880.

[35] S. R. Arridge, M. Schweiger, Photon-measurement density functions. part 2: Finite-element-method calculations, Applied Optics 34 (34) (1995) 8026–8037.

[36] J. Nocedal, Updating quasi-Newton matrices with limited storage, Mathematics of computation 35 (151) (1980) 773–782.

[37] F. Dubot, Y. Favennec, B. Rousseau, D. R. Rousse, Paramétrisation des variables de contrôle et méthodes de recherche linéaire dans un code d'inversion de l'approximation de diffusion basé sur le L-BFGS, in: Comptes-Rendus Congrès de la Société Française de Thermique, Lyon, France, 2014.

[38] A. Antoniou, W.-S. Lu, Practical optimization: algorithms and engineering applications, Springer, 2007.

[39] R. Fletcher, Practical methods of optimization, vol. 1, unconstrained optimization, British Library Cataloguing in Publication Data (1980) 126.

[40] O. Alifanov, Solution of an inverse problem of heat conduction by iteration methods, Journal of Engineering Physics and Thermophysics 26 (4) (1974) 471–476.

[41] T. O. McBride, B. W. Pogue, S. Jiang, U. L. Oesterberg, K. D. Paulsen, A parallel-detection frequency-domain near-infrared tomography system for hemoglobin imaging of the breast in vivo, Review of Scientific Instruments 72 (3) (2001) 1817–1824.

[42] S. Mallat, A wavelet tour of signal processing, Academic press, 1999.

[43] P. S. Addison, The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance, CRC Press, 2010.

[44] F. Hecht, New development in freefem++, Journal of Numerical Mathematics 20 (3-4) (2012) 251–266.

[45] D. Colton, R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, Springer, 1998.

[46] B. P. Flannery, W. H. Press, S. A. Teukolsky, W. Vetterling, Numerical recipes in c, Press Syndicate of the University of Cambridge, New York.

[47] J. R. Williams, K. Amaratunga, A discrete wavelet transform without edge effects using wavelet extrapolation, Journal of Fourier Analysis and Applications 3 (4) (1997) 435–449.