Diffuse optical tomographic imaging of biological media by time-dependent parabolic SP_N equations: a two-dimensional study

Jorge Bouza Domínguez
Yves Bérubé-Lauzière

Diffuse optical tomographic imaging of biological media by time-dependent parabolic SP$_N$ equations: a two-dimensional study

Jorge Bouza Domínguez[a] and Yves Bérubé-Lauzière[a,b]
[a]Université de Sherbrooke, Laboratoire TomOptUS, Département de Génie Electrique et de Génie Informatique, Faculté de génie, 2500 Boulevard Université, Sherbrooke, Québec, J1K 2R1, Canada
[b]Centre d'Imagerie Moléculaire de Sherbrooke (CIMS)—Centre de Recherche Clinique Étienne-Le Bel du CHUS, 3001, 12e Avenue Nord, Sherbrooke, Québec, J1H 5N4, Canada

Abstract. We investigate the problem of retrieving the optical properties (absorption and scattering) of biological tissue from a set of optical measurements. A diffuse optical tomography (DOT) algorithm that incorporates constrained optimization methods is implemented. To improve image quality, the DOT algorithm exploits full time-domain data. The time-dependent parabolic simplified spherical harmonics equations (TD-pSP$_N$) are used as the forward model. Time-dependent adjoint variables are resorted to in the calculation of the gradient of the objective function. Several numerical experiments for small geometric media with embedded inclusions that mimic small animal imaging are performed. In the experiments, optical coefficient values are varied in the range of realistic values for the near-infrared spectrum, including high absorption values. Single and multiparameter reconstructions are performed with the diffusion equation and higher orders of the TD-pSP$_N$ equations. The results suggest the DOT algorithm based on the TD-pSP$_N$ model outperforms the DE, and accurately reconstructs optical parameter distributions of biological media both spatially and quantitatively. © *2012 Society of Photo-Optical Instrumentation Engineers (SPIE).* [DOI: 10.1117/1.JBO.17.8.086012]

Keywords: diffuse optical tomography; near-infrared; time-domain; simplified spherical harmonics equations; partial differential equation-constrained optimization; adjoint variables.

Paper 11764 received Dec. 16, 2011; revised manuscript received Jul. 2, 2012; accepted for publication Jul. 12, 2012; published online Sep. 4, 2012.

1 Introduction

Optical radiation in the near-infrared (NIR) spectrum is used to probe biological tissues with quite significant success in structural and functional imaging.[1–12] Usually, light from a distribution of sources which has propagated through a tissue is measured at the tissue's boundary. These measurements are then used to reconstruct a map of the internal optical properties. This medical imaging modality, known as diffuse optical tomography (DOT), has undergone fast progress pertaining to models of light propagation in biological media and algorithms for solving the inverse problem.[1,2] DOT experimental systems have been simultaneously developed with measurements being made in either the continuous wave (CW),or the time-resolved (TR) regimes, with the latter comprising frequency-domain (FD) or time-domain (TD) measurements as subsets. These different measurement schemes have collected numerous data types.[13–15] FD and TD data are frequently used when both scattering and absorption properties of tissues are needed. On the other hand, the uniqueness of CW-based solutions to the DOT and fluorescence DOT (FDOT) inverse problems is a topic of current discussion.[16–18] It is generally believed that TD measurements can provide the richest information content about the local distribution of optical properties in tissue.[19,20]

Image reconstructions based on the distinct features of TD data are able to separate absorption from scatter maps, while reducing the crosstalk effect.[21–24] Moreover, DOT and FDOT reconstruction algorithms that use full TD data show promising results in terms of improving spatial resolution and their ability to quantitatively determine optical coefficients maps.[25–27] Despite the former results, DOT reconstruction algorithms are limited by the accuracy of the forward model. In the presence of small geometries and high absorption values, the diffusion equation (DE) is no longer valid as a forward model.[28] To overcome limitations of the DE, and computational expenses of other approaches resorting to the radiative transfer equation (RTE), or the P_N approximation thereof, low-order transport models incorporating the simplified spherical harmonics approximation (SP$_N$) were derived for biomedical optics applications.[29–32] Analytical solutions to SP$_N$-based models have appeared in the literature for simple geometries with homogeneous optical properties.[33,34] Numerical solutions for complex media were also found for simulating light propagation from external sources and fluorescent probes.[29–32,35–37] These results led to the development of SP$_N$ model-based image reconstruction algorithms for solving inverse problems in intrinsic, bioluminescence, and fluorescence DOT imaging.[38–41] However, as far as we know, no image reconstruction method or algorithms based on a time-dependent SP$_N$ model was ever attempted. Such an approach could benefit from both the accuracy of the SP$_N$ models and the richness of TD data. The subject of this

Address all correspondence to: Yves Bérubé-Lauzière, Université de Sherbrooke, Département de génie électrique et de génie informatique, Laboratoire TomOptUS, Faculté de génie, 2500 Boulevard Université, Sherbrooke, Québec, J1K 2R1, Canada. Tel: +1 819-821-8000; ext: 63753; Fax: +1 819-821-7937; E-mail: yves .berube-lauziere@usherbrooke.ca

paper is to provide such an approach, and assess the accuracy of the TD SP_N equations as a forward model (see below) for recovering the internal optical properties of biological media.

With that goal, an image reconstruction algorithm for solving the inverse DOT imaging problem is implemented. As the forward model, which is reviewed in Sec. 2, a discretized formulation of the time-dependent parabolic SP_N equations (the TD-pSP_N model) is used, which was derived and studied in depth in a previous work.[32] The spatial discretization of the model is performed through the finite element method (FEM), and the temporal discretization is carried out with an implicit finite difference scheme. In this way, it is possible to deal with complex geometries and heterogeneous distributions of the optical coefficients. Following the latest trends in model-based image reconstruction algorithms, which resort to iterative optimization techniques,[42–46] the inverse problem is cast here as a partial differential equation (PDE)-constrained optimization scheme. This will be described in Sec. 3. To aid in reducing the ill-posedness and improve the stability of the algorithm, the inverse problem solution is further constrained by imposing bounds on the optical coefficient values. The resulting optimization problem possesses a high number of unknown variables, and has to deal with a large amount of data owing to the use of TD data. To reduce complexity, a nested analysis and design (NAND) method is employed.[47] Further, the problem is framed in a sequential quadratic programming (SQP) algorithm.[48] By the introduction of adjoint variables, an iterative numerical scheme for calculating the gradient values of the objective function (OF) at each time step is provided, with the OF measuring the discrepancy between the predicted and actual measurements. The former steps (i.e., the use of NAND, SQP, and adjoint variables) taken in solving the inverse problem reduce the computation time considerably.

To examine the DOT algorithm's performances and the accuracy of the TD-pSP_N model in the reconstructions, a series of numerical experiments resembling small animal intrinsic imaging are carried out, as described in Sec. 4. Single parameter reconstructions are performed for the case of a homogeneous medium with an embedded absorptive inclusion, for which increasing values of the absorption coefficient are employed. High absorption values which simulate practical situations encountered in DOT and FDOT are considered. The image reconstructions for SP_N orders $N > 1$ yield better quantitative results than the DE-based reconstructions. Particularly, the reconstructed images using the TD-pSP_3 approximation provide the best results. Next, a scattering inclusion is added in the previous situations, and the simultaneous reconstruction of the absorption and the diffusion coefficient maps is performed. As before, image reconstructions for $N > 1$ better estimate the optical properties of the inclusions than the DE (TD-pSP_1).

As will be seen, the results obtained support the TD-pSP_N-based DOT algorithm presented herein as an accurate approach for obtaining spatial distributions of internal optical properties of biological media.

2 Forward Problem

In Ref. 32, a new time-dependent model based on the SP_N approximation for describing light propagation in scattering and absorbing media was described and validated. The model is represented by a set of $(N + 1)/2$-coupled parabolic partial differential equations (PDEs) referred to as the time-domain parabolic SP_N (TD$-pSP_N$) equations. The TD-pSP_N equations

can be written in matrix notation in terms of the vector of composite moment functions $\mathbf{\Phi}$ as[32]

$$\left[\mathbf{C} + \frac{\eta}{c}\frac{\partial}{\partial t}\mathbf{T} \right] \Phi(\mathbf{r}, t) + D_r \Phi(\mathbf{r}, t) = \mathbf{Q}(\mathbf{r}, t), \qquad (1)$$

where η is the refractive index of the medium, and c is the speed of light in a vacuum. The explicit expression for the matrices $\mathbf{C}$, $\mathbf{T}$, the matrix operator D_r, and the source vector $\mathbf{Q}$ are reproduced in Appendix A. FEM spatial discretization and an implicit finite differencing method (FDM) for the time variable lead to the following FEM-FDM numerical representation of the TD-pSP_N equations[32]

$$\tilde{\mathbf{W}}\tilde{\mathbf{\Phi}}^{(m)} = \left(\frac{\eta}{c} \right) \tilde{\mathbf{T}}\tilde{\mathbf{\Phi}}^{(m-1)} + \tilde{\Upsilon}^{(m)}, \qquad (2)$$

where

$$\tilde{\mathbf{W}} = \Delta t(\tilde{\mathbf{K}} + \tilde{\mathbf{M}} + \tilde{\mathbf{\Pi}}) + \left(\frac{\eta}{c} \right)\tilde{\mathbf{T}}, \qquad (3)$$

and

$$\tilde{\Upsilon}^{(m)} = \Delta t[\tilde{\mathbf{F}}^{(m)} + \tilde{\Gamma}^{(m)}]. \qquad (4)$$

In Eqs. (2)–(4), $\tilde{\mathbf{\Phi}}^{(m)}$ represents the vector of composite moments of the radiance at time $t_m = m\Delta t$, where Δt is the time-step employed in the finite difference scheme. The remaining terms appearing in Eqs. (3) and (4) are also described in Appendix A, along with their dependencies on the absorption μ_a, scattering μ_s, and anisotropy g coefficients. As the terms given in Eqs. (3) and (4) are critical in describing the solution to the inverse problem, Appendix A and Ref. 32 should be consulted as necessary. Next, the notation associated with Eq. (2) is discussed.

2.1 Notation

The SP_N equations come naturally with odd orders N.[29,32] To simplify the notation in the sequel, the constant $K = (N + 1)/2$ is defined. The following notation will also be used. First, the quantity $\Phi(\mathbf{r}, t)$ considered in the TD-pSP_N model is given by a vector of space- and time-dependent functions as follows

$$\Phi(\mathbf{r}, t) = [\, \varphi_1(\mathbf{r}, t) \quad \varphi_2(\mathbf{r}, t) \quad \cdots \quad \varphi_K(\mathbf{r}, t)\,]^T, \quad (5)$$

where $[\,]^T$ means the transpose; the functions $\varphi_k(\mathbf{r}, t)$ $k = 1, \ldots, K$ are called the composite moments, and they constitute what are referred to as the state variables when image reconstruction is considered later. When the time variable is discretized at timesteps $t_m = m\Delta t$, $m = 1, \ldots, M - 1$, the following quantities are defined

$$\Phi^{(m)}(\mathbf{r}) \equiv \Phi(\mathbf{r}, t_m). \qquad (6)$$

Using the FEM, the solutions $\Phi(\mathbf{r}, t)$ to the TD-pSP_N equations are approximated over a mesh of elements (e.g., triangles) by a piecewise polynomial and continuous function $\Phi^h(\mathbf{r}, t)$. Let L be the number of nodes in the mesh (e.g., the different vertices of the mesh triangles). Then, the solution may be written as

$$\Phi^h(\mathbf{r}, t) = \sum_{i=1}^{L} \Phi_i(t) u_i(\mathbf{r}), \qquad u_i(\mathbf{r}) \in \Omega^h, \qquad (7)$$

where Ω^h is a finite-dimensional subspace spanned by the polynomial basis functions $u_i(\mathbf{r})$, $i = 1, \ldots, L$, and where h denotes the characteristic size of the elements. Hence, the problem is reduced to finding the nodal values $\tilde{\Phi}(t) = \{\Phi_i(t)\}_{i=1,\ldots,L}$ from which the solution can be obtained everywhere through the interpolation rule given in Eq. (7). The simplest choice, and the one made here, is a piecewise linear basis over a mesh of triangles in which the nodal values of the basis functions are defined by $u_i(\mathbf{r}_j) = \delta_{i,j}$, $i, j = 1, \ldots, L$. Let $i = 1, \ldots, L$ be an index over the set of all nodes in the FEM mesh of triangles that will be denoted by $\mathcal{D} = \{\mathbf{r}_i\}_{i=1,\ldots L}$, with $\mathbf{r}$ denoting the position of a point in space. The subset of interior nodes in $\mathcal{D}$ will be denoted by $\mathcal{D}_I$, and L_I will denote the number of interior nodes. The remaining nodes form the set of boundary nodes denoted by $\mathcal{D}_B$ containing a total of L_B nodes, with, of course, $L_I + L_B = L$.

The discretized version of $\Phi(\mathbf{r}, t)$ is

$$\tilde{\Phi}^{(m)} = [\tilde{\varphi}_1(\mathbf{r}_1, t_m) \tilde{\varphi}_1(\mathbf{r}_2, t_m) \cdots \tilde{\varphi}_1(\mathbf{r}_L, t_m) \tilde{\varphi}_2(\mathbf{r}_1, t_m) \cdots$$
$$\tilde{\varphi}_2(\mathbf{r}_L, t_m) \cdots \tilde{\varphi}_K(\mathbf{r}_1, t_m) \cdots \tilde{\varphi}_K(\mathbf{r}_L, t_m)]^T,$$

or in more abbreviated form

$$\tilde{\Phi}^{(m)} = [\; \tilde{\varphi}_1^{(m)} \quad \tilde{\varphi}_2^{(m)} \quad \cdots \quad \tilde{\varphi}_K^{(m)} \;]^T. \qquad (8)$$

Light source configurations are indexed with the subscript s. A source configuration is used for illuminating the medium and making a set of measurements. Such a set of measurements for a given source configuration is what will be called a "tomographic projection". A source configuration may, for instance, consist of a set of different sources simultaneously illuminating the object. For a given source configuration s, the associated solution of the discretized TD-pSP$_N$ equations at a given time m will be written as $\tilde{\Phi}_s^{(m)}$. Hereon, $\{\tilde{\Phi}_s^{(m)}\}$ will denote the set containing the solution at all time steps for all source configurations.

Note that the discretized TD-pSP$_N$ equations are solved for each $s = 1, \ldots, S$ (i.e., for each source configuration illuminating the object) in order to predict the measurements made for that configuration. This means that the forward problem needs to be solved as many times as there are source configurations.

In the following, $\boldsymbol{\mu}$ will denote the vector containing the set of optical coefficients evaluated at the nodes of the FEM mesh and to be retrieved through the solution of the inverse problem. Each optical coefficient is expanded along the same type of basis functions of the FEM expansion, see Eq. (7). In accordance with the previous notation, if for example the optical coefficients μ_a and D are considered, as will be done later, where D is the diffusion coefficient given by $D = 1/[3(\mu_a + (1 - g)\mu_s)]$, then

$$\boldsymbol{\mu} = [\mu_a(\mathbf{r}_1)\mu_a(\mathbf{r}_2) \cdots \mu_a(\mathbf{r}_L)D(\mathbf{r}_1)D(\mathbf{r}_2) \cdots D(\mathbf{r}_L)]^T \quad (9)$$

represents the nodal values of the optical coefficients.

2.2 Measurements Modeling

The type of measurements depends on the experimental conditions, and can take diverse forms.[22] Here, measurements are taken as the full time-dependence of the outgoing normal component of the radiant current density vector $J_{n,s,d}^{(m)}$ (or excitance) at a finite collection of detector positions $\mathbf{r}_d$, $d = 1, \ldots, D$ for each source configuration s at m different times. The values of $J_{n,s,d}^{(m)}$ can be calculated using $\tilde{\Phi}_s^{(m)}$ and the optical coefficients evaluated at the detector positions via

$$J_{n,s,d}^{(m)} = [\mathbf{j}_1 - \mathbf{j}_2(\mathbf{B})^{-1}\mathbf{A}]\tilde{\Phi}_s^{(m)}(\mathbf{r}_d, t) = \mathbf{V}_d \tilde{\Phi}_s^{(m)}(\mathbf{r}_d, t),$$
$$s = 1, \ldots, S, \quad d = 1, \ldots, D, \quad m = 1, \ldots, M, \qquad (10)$$

where the expressions for vectors $\mathbf{j}_1$ and $\mathbf{j}_2$ are given in Appendix B, and $\mathbf{A}$ and $\mathbf{B}$ are matrices appearing in the boundary conditions,[29,32] see Appendix A. The elements of both matrices depend on the values of the optical coefficients of the medium at the boundary nodes. The vector $\mathbf{V}_d$ is a compact representation of the "measurement operator" $\mathbf{j}_1 - \mathbf{j}_2(\mathbf{B})^{-1}\mathbf{A}$. Note that if the outward flux is collected in a detection area, and that if geometrical factors and the numerical aperture of a real detector are taken into account (see Sec. 7-6 in Ref. 49), then other types of measurements can be expressed as a weighted sum of the values of $J_{n,s,d}^{(m)}$.

3 Inverse Problem

DOT imaging is an inverse problem that can be considered as a parameter-estimation task. It involves finding parameter values, which are optimal in that they minimize an objective—or functional—that measures the misfit between a given set of measurements and predictions thereof. A judicious choice of functional improves the quality of the final image and accelerates the convergence of the reconstruction algorithm, even in the presence of measurement noise. In addition, a suitable functional reduces the effect of cross variations on the optical parameter subsets, an effect known as crosstalk.[22] Several functionals and selection of measurement types are proposed in the literature to eliminate functional valley structure, and achieve less eccentricity in the functional characteristic surface.[22,50] Even so, simulated and experimental reconstructions still exhibit poor spatial resolution and quantitativeness.[19,21,22] Some authors recommend exploiting the full profiles of TD measurements in nonlinear reconstruction schemes.[25,26] The main idea behind the mentioned approaches is that the amplitude of the TD profile is influenced by both absorption and scattering coefficients (the influence of the absorption coefficient is stronger in these respects), while the shape of the TD profile is more dependent on the scattering coefficient values.[19] Still, exploiting TD data in image reconstruction algorithms and optimal functional selection is a subject of active research.[50–52]

3.1 Optimization Problem

Following the previous discussion, a least-squares estimation is cast to the inverse problem using an objective function resembling that proposed in Ref. 25. Then, the minimum of the difference between the TD measurement data $M_{s,d}^{(m)}$, $s = 1, \ldots, S$, $d = 1, \ldots, D$, $m = 1, \ldots, M$ and the predictions $J_{n,s,d}^{(m)}$ obtained from the TD-pSP$_N$ forward model is sought. Thus, the following objective function (OF) is considered

$$f = \frac{1}{2} \sum_{s=1}^{S} \sum_{d=1}^{D} \sum_{m=1}^{M} \left[\frac{M_{s,d}^{(m)} - J_{n,s,d}^{(m)}}{\sigma_{s,d}^{(m)}} \right]^2, \qquad (11)$$

where the magnitudes $\sigma_{s,d}^{(m)}$ quantify the uncertainty in measurements, which are mainly influenced by photon noise. Because of the Poisson distribution nature of the photon noise present in TD data (shot noise is considered as the main contribution here); the uncertainty is assumed to be proportional to the square root of the measured signal,[51] i.e., $\sigma_{s,d}^{(m)} \sim \sqrt{M_{s,d}^{(m)}}$. To gain clarity for the sequel, the OF is re-written directly in terms of $\bar{\Phi}_s^{(m)}$

$$f = \frac{1}{2} \sum_{s=1}^{S} \sum_{d=1}^{D} \sum_{m=1}^{M} \frac{\left[M_{s,d}^{(m)} - \sum_{k=1}^{K} V_{k,d} \mathbf{P}_{k,d} \bar{\Phi}_s^{(m)} \right]^2}{M_{s,d}^{(m)}}, \quad (12)$$

where the measurement operator $\mathbf{V}_d$ appearing in Eq. (10) was decomposed into K components ($k = 1, \ldots K$). The projection operator (or vector) $\mathbf{P}_{k,i}$ is represented by a row vector of length K times L which has zero components except for a value equal to 1 at position $(k-1)L + i$. Thus, for any m, [refer back to Eq. (8)]

$$\mathbf{P}_{k,d} \tilde{\Phi}_s^{(m)} = \tilde{\varphi}_{k,s,d}^{(m)}. \quad (13)$$

Now, the inverse problem can be posed as finding

$$\boldsymbol{\mu} = \arg \min_{\{\mu^h\}} \frac{1}{2} \sum_{s=1}^{S} \sum_{d=1}^{D} \sum_{m=1}^{M} \frac{\left(M_{s,d}^{(m)} - \sum_{k=1}^{K} V_{k,d} \mathbf{P}_{k,d} \tilde{\Phi}_s^{(m)} \right)^2}{M_{s,d}^{(m)}},$$

$$\text{subject to} \begin{cases} \tilde{\mathbf{W}} \tilde{\Phi}_s^{(m)} = \left(\frac{\eta}{c} \right) \tilde{\mathbf{T}} \tilde{\Phi}_s^{(m-1)} + \tilde{\Upsilon}_s^{(m)}, \\ \boldsymbol{\mu}_{\text{low}} \leq \boldsymbol{\mu} \leq \boldsymbol{\mu}_{\text{up}}, \end{cases}$$

$$m = 1, \ldots, M, s = 1, \ldots, S,$$

$$(14)$$

where arg min stands for argument of the minimum. The vectors $\boldsymbol{\mu}_{\text{low}}$ and $\boldsymbol{\mu}_{\text{up}}$ are lower and upper bounds over the set of optical coefficients $\boldsymbol{\mu}$.

The inverse problem posed in Eq. (14) is a constrained optimization problem, specifically a PDE-constrained optimization problem. In general, PDE-constrained optimization refers to the optimization of systems whose states are governed by PDEs. The forward problem consists in solving the PDEs for the state variables (e.g., temperature, electric field, or in the present case, the optical field $\{\tilde{\Phi}_s^{(m)}\}$), given appropriate data (e.g., geometry, boundary conditions, initial conditions, source functions, and in DOT, the set of optical coefficients $\boldsymbol{\mu}$). The optimization problem seeks to determine some of these data, called the decision or design variables, given performance goals in the form of an OF [first line of Eq. (14)], and possibly inequality or equality constraints on the behaviour and parameters of the system. Since the system's behavior is modeled by PDEs, the latter appear as equality constraints in the optimization problem. The PDE constraints are usually referred as the state equations. There is an extra feature in the present problem that increases its complexity. If the state equations are evolutionary in nature (that is, dependent on time), the optimality conditions (necessary conditions for the existence of a minimum) for Eq. (14) constitutes a boundary-value problem in the space–time domain. For this reason, the optimization problem considered here is significantly more difficult to solve than steady-state problems. Posing the inverse problem as a PDE-constrained optimization problem

is a relatively new approach in biomedical optics[42–46] that benefits from considerable experience gained in the area of large-scale optimization.[47,48,53,54]

3.2 Nested Analysis and Design

The total number of variables (state and design variables) in the present case dealing with TD data considerably surpasses those in CW and FD inverse problems. The optimization problem becomes more complex because of the time-dependence of the forward model. The reconstruction algorithm has to deal with a large amount of variables and involves sparse matrices whose size and manipulation are memory demanding, especially for a fine mesh and small time-steps. Typically, a TD measurement can contain up to a few thousands of points. For the present purposes, on the order of a hundred points on each TD curve are used, which makes the number of state variables around $100 \times K$ times that of design variables. Motivated by these considerations, it was decided to employ a "nested analysis and design" (NAND) method to solve the inverse problem. The NAND method, also called the "black-box" approach, reduces the number of unknown parameters in a constrained optimization problem[47] by considering the state variables (in the present case, $\{\tilde{\Phi}_s^{(m)}\}$) as implicit functions of the design variables $\boldsymbol{\mu}$. Thus, the states $\{\tilde{\Phi}_s^{(m)}\}$ are obtained by explicitly solving the PDE [given by Eq. (2)], and substituted in the OF, leading to a reduction of the dimensionality of the problem. In mathematical terms

$$\tilde{\Phi}_s^{(m)} = \tilde{\Phi}_s^{(m)}(\boldsymbol{\mu}) \Rightarrow f(\tilde{\Phi}_s^{(m)}, \boldsymbol{\mu}) = f(\tilde{\Phi}_s^{(m)}(\boldsymbol{\mu}), \boldsymbol{\mu}) = f(\boldsymbol{\mu}). \quad (15)$$

Consequently, the problem is reduced to an optimization problem which is unconstrained with respect to the PDE, but constrained with respect to the optical properties to be recovered because of the lower and upper bounds imposed thereon. These upper and lower bound constraints are treated in the algorithm as linear constraints with an associated l_1 merit function; see below. Descriptions of the algorithm and implementation details are provided next.

3.3 Sequential Quadratic Programming

To find the solution, the optimization problem posed in Eq. (14) is cast into a sequential quadratic programming (SQP) algorithm. SQP basically consists in locally approximating the objective function at each iteration step by a quadratic function. This significantly simplifies the optimization problem, because a simpler quadratic function needs to be considered, rather than a nonlinear one.[48] SQP is widely used in large-scale optimization for solving nonlinear constrained optimization problems because of the reduction in the computational cost and its fast convergence. Details about the implementation of this kind of schemes can be found elsewhere. For example, see Refs. 48, 53, and 54 for extensive documentation about the topic, and Refs. 44–46 and 54 for applications of inverse problems in electromagnetism, DOT, and fluorescence DOT. A description of the SQP method used in the context of the present DOT algorithm will now be made.

For the line-search at every iterate $\boldsymbol{\mu}_k$ of the optical properties at step k during the optimization process, the search direction $\mathbf{d}_k$ in the SQP scheme is obtained by solving the quadratic subproblem[48]

$$\min_{\mathbf{d} \in \mathbf{R}^n} \frac{1}{2} \mathbf{d}_k^T \mathbf{H}_k \mathbf{d}_k + \mathbf{g}_k \mathbf{d}_k, \quad \text{subject to } \boldsymbol{\mu}_{\text{low}} \leq \boldsymbol{\mu} \leq \boldsymbol{\mu}_{\text{up}}, \quad (16)$$

where $\mathbf{g}_k$ denotes the gradient of the OF at step k, which is obtained through the forward model (see next subsection), and $\mathbf{H}_k$ denotes the Hessian (with respect to $\boldsymbol{\mu}_k$) of the Lagrangian function $L(\boldsymbol{\mu}, \lambda) = f(\boldsymbol{\mu}) - \boldsymbol{\beta}^T \mathbf{c}(\boldsymbol{\mu})$ [$\boldsymbol{\beta}$ is the vector of Lagrange multipliers determined by the Karush-Khun-Tucker (KKT) conditions]. Here, the bound constraints $\boldsymbol{\mu}_{\text{low}} \leq \boldsymbol{\mu} \leq \boldsymbol{\mu}_{\text{up}}$ are rewritten as linear constraints represented by $\mathbf{c}(\boldsymbol{\mu}) \geq 0$. The updated vector is given by

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + \alpha_k \mathbf{d}_k, \quad (17)$$

where α_k is a step-length parameter that provides sufficient reduction in the l_1 merit function

$$\phi_\rho = f(\boldsymbol{\mu}) + \rho \|v(\boldsymbol{\mu})\|_1. \quad (18)$$

Here, ρ is the penalty parameter defined to enforce a descent of φ_ρ, and $v(\boldsymbol{\mu})$ measures the violation of the constraints; the function v is defined as

$$\|\nu(\boldsymbol{\mu})\|_1 = \sum_{n=1}^{N_v} \{| \max(0, \mu_n - \mu_{n,\text{up}})| + | \max(0, \mu_{n,\text{low}} - \mu_n)|\}, \quad (19)$$

where N_v is the number of elements in the vector $\boldsymbol{\mu}$. Usually in such SQP schemes, the Hessian is updated by the damped Broyden–Fletcher–Goldfarb–Shanno (BFGS) method to avoid computing second derivatives,[48] and this method was used here. For a representation of the general algorithm and the SQP scheme, see the flowchart appearing in Fig. 1.

In the following, some focus is put on the necessary calculation of the gradient of the OF. The technique based on adjoint variables avoids using finite difference schemes of the OF and associated inherent errors and extremely high computational cost.
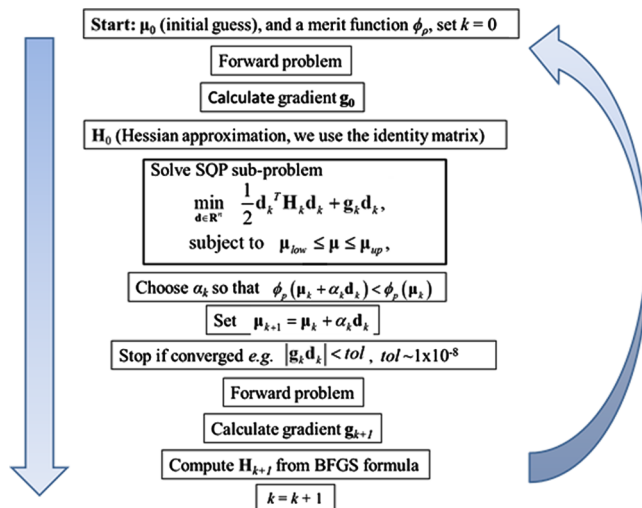
Fig. 1 Flowchart of the general algorithm and sequential quadratic programming (SQP) scheme.

3.4 Gradient of the Objective Function

Calculating the gradient of the OF is undoubtedly the most difficult part of an inverse problem. This will now be tackled for the specific problem considered herein. Here, the gradient of the OF $\mathbf{g}$ can be expanded as a sum of independent terms for each source configuration: $\mathbf{g} = \sum_{s=1}^{S} \mathbf{g}_s$. Considering the implicit dependence of $\tilde{\Phi}_s^{(m)}$ on $\boldsymbol{\mu}$, the j'th component of the gradient $g_{s,j}$ can be written as

$$g_{s,j} = \sum_{m=1}^{M} \frac{df}{d\tilde{\Phi}_s^{(m)}} \frac{\partial \tilde{\Phi}_s^{(m)}}{\partial \mu_j}, \quad (20)$$

where $df/d\tilde{\Phi}_s^{(m)}$ denotes the derivative of the OF with respect to $\tilde{\Phi}_s^{(m)}$. The explicit variation of $\tilde{\Phi}_s^{(m)}$ with $\boldsymbol{\mu}$ can be obtained from Eq. (2) as

$$\frac{\partial \tilde{\Phi}_s^{(m)}}{\partial \mu_j} = -\tilde{\mathbf{W}}^{-1} \left(\frac{\partial \tilde{\mathbf{W}}}{\partial \mu_j} \right) \tilde{\Phi}_s^{(m)}. \quad (21)$$

Note that in obtaining Eq. (21), $\tilde{\Phi}_s^{(m-1)}$, which appears in Eq. (2), was not differentiated with respect to μ_j, because, as shown below, backwards differentiation will be resorted to, and in this framework, $\tilde{\Phi}_s^{(m-1)}$ is considered dependent on $\tilde{\Phi}_s^{(m)}$. Next, an expression for $df/d\tilde{\Phi}_s^{(m)}$ is derived. Note that because of Eq. (2), a simple explicit derivative of the OF with respect to $\tilde{\Phi}_s^{(m)}$ does not consider the dependence of $\tilde{\Phi}_s^{(m)}$ on previous time values of the field which comes from the FDM scheme. To resolve that issue, the core idea of adjoint differentiation (or reverse differentiation) expounded in Ref. 55 for a time-dependent DE-based inverse problem is followed. The idea is to apply the chain rule of differentiation backwards from time $m + 1$ to m to the OF to yield

$$\frac{df}{d\tilde{\Phi}_s^{(m)}} = \frac{df}{d\tilde{\Phi}_s^{(m+1)}} \frac{d\tilde{\Phi}_s^{(m+1)}}{d\tilde{\Phi}_s^{(m)}} + \frac{\partial f}{\partial \tilde{\Phi}_s^{(m)}}, \quad (22)$$

with

$$\frac{df}{d\tilde{\Phi}_s^{(M)}} = \frac{\partial f}{\partial \tilde{\Phi}_s^{(M)}}. \quad (23)$$

Equations (22) and (23) now take into account the time-dependence of $\tilde{\Phi}_s^{(m)}$. From Eqs. (2) and, the following terms can be calculated

$$\frac{d\tilde{\Phi}_s^{(m+1)}}{d\tilde{\Phi}_s^{(m)}} = \tilde{\mathbf{W}}^{-1} \left(\frac{\eta}{c} \right) \tilde{\mathbf{T}}, \quad (24)$$

$$\frac{\partial f}{\partial \tilde{\Phi}_s^{(m)}} = \sum_{d=1}^{D} \frac{\left(\sum_{k=1}^{K} V_{k,d} \mathbf{P}_{k,d} \tilde{\Phi}_s^{(m)} - M_{s,d}^{(m)} \right)}{M_{s,d}^{(m)}} \sum_{k=1}^{K} V_{k,d} \mathbf{P}_{k,d}. \quad (25)$$

Substituting Eq. (21) into Eq. (20), and Eq. (24) into Eq. (22), leads to

$$g_{s,j} = -\sum_{m=1}^{M} \frac{df}{d\tilde{\Phi}_s^{(m)}} \tilde{\mathbf{W}}^{-1} \left(\frac{\partial \tilde{\mathbf{W}}}{\partial \mu_j}\right) \tilde{\Phi}_s^{(m)},$$

$$\frac{df}{d\tilde{\Phi}_s^{(m)}} = \frac{df}{d\tilde{\Phi}_s^{(m+1)}} \tilde{\mathbf{W}}^{-1} \left(\frac{\eta}{c}\right) \tilde{\mathbf{T}} + \frac{\partial f}{\partial \tilde{\Phi}_s^{(m)}}. \quad (26)$$

Equation (26) suggests the introduction of the adjoint variables $\lambda_s^{(m)}$ as auxiliary variables

$$\tilde{\mathbf{W}}^T \lambda_s^{(m)} = -\left(\frac{df}{d\tilde{\Phi}_s^{(m)}}\right)^T, \quad (27)$$

which allows rewriting Eq. (26) as

$$g_{s,j} = \sum_{m=1}^{M} (\lambda_s^{(m)})^T \left(\frac{\partial \tilde{\mathbf{W}}}{\partial \mu_j}\right) \tilde{\Phi}_s^{(m)}, \quad (28)$$

$$\frac{df}{d\tilde{\Phi}_s^{(m)}} = -(\lambda_s^{(m+1)})^T \left(\frac{\eta}{c}\right) \tilde{\mathbf{T}} + \frac{\partial f}{\partial \tilde{\Phi}_s^{(m)}}. \quad (29)$$

Introducing the adjoint variables $\lambda_s^{(m)}$ reduces the number of computations to be carried out because it avoids the need to compute the partial derivatives of the field $\tilde{\Phi}_s^{(m)}$ with respect to $\boldsymbol{\mu}$. As seen from Eq. (28), the adjoint variables $\lambda_s^{(m)}$ need to be computed in order to obtain the gradient. The adjoint variables are obtained from Eqs. (27) and (29) by starting with $m = M$. In this case,

$$\tilde{\mathbf{W}}^T \lambda_s^{(M)} = -\left(\frac{\partial f}{\partial \tilde{\Phi}_s^{(M)}}\right)^T. \quad (30)$$

Then, for $m < M$ the derivative $df/d\tilde{\Phi}_s^{(m)}$ is obtained from Eq. (29) and used in Eq. (27) to get $\lambda_s^{(m)}$. From Eq. (27) it is seen that the matrix $\tilde{\mathbf{W}}^T$ needs to be repeatedly inverted in this process. The computation load of these inversions, which are always the same for a given algorithm iteration, can be reduced by first storing information on the matrix (e.g., some form of factorization). It should be pointed out that in solving the forward model, the inversion of $\tilde{\mathbf{W}}$ is needed as well. At that point, the information alluded to above regarding $\tilde{\mathbf{W}}$ can be stored.

As a final note, further applications of Eq. (29) are: 1. the construction of sensitivity maps (for each order N) to changes in the medium optical properties, and 2. the generation of single-perturbation error maps for custom data types and functional optimal selection.[22]

4 Results of Numerical Experiments

4.1 Reconstruction of Absorption Coefficient Maps

The numerical experiments in this subsection were performed for a circular geometry with a radius of 1.5 cm with homogeneous optical properties. This medium mimics a two dimensional (2D) slice of a small volume of biological tissue, see Fig. 2.

The values of the optical coefficients of the homogeneous medium are $\mu_a = 0.01$ cm^{-1}, $\mu_s = 100$ cm^{-1}, and $g = 0.9$.
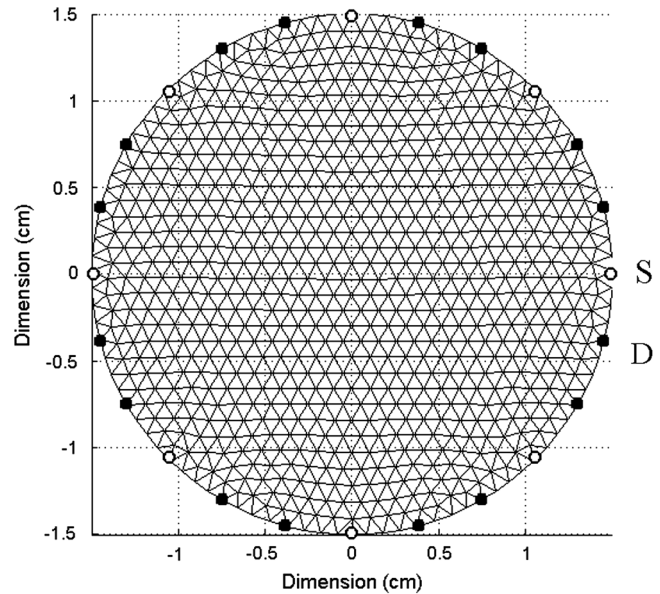
Fig. 2 Mesh of the circular geometry. Sources and detector (optodes) positions are represented by white spots (S) and dark squares (D), respectively.

The refractive index of the medium is 1.4, resembling fat tissue, and the medium is considered as surrounded by air. These values are representative of small animal tissue samples in the NIR spectrum, e.g., abdomen, brain, and blood vessels in rats.[29,56] Next, Dirac delta functions in time collimated laser sources are placed at eight positions located at the boundary of the circular geometry, see Fig. 2. The sources are modeled as exponentially damped line sources distributed inside the medium and are fired sequentially. To use the FEM, the medium was divided into triangular elements (816 nodes and 1534 elements) using a coarse regular Delaunay triangulation as shown in Fig. 2. As a second step, an irregular-shaped inclusion is positioned inside the medium. It has the same properties as the background medium, except for the absorption coefficient μ_a, for which the following increasing values were examined: 0.05, 0.1, and 1 cm^{-1} (Fig. 3). These μ_a values cover possible situations appearing in DOT and FDOT, including regions with high absorption, such as vascularized tissues and/or fluorescent inclusions.[56,57]

Using the FEM-FDM representation of the TD-pSP$_N$ equations, the time-dependent distribution of fluence in the circular geometry is calculated for the three aforementioned μ_a values and for the orders $N = 1, 3, 5$, and 7 of the SP$_N$. A partially delimited darker region with depressed values of the fluence was noted at the location of the inclusion in all the cases. The darkening increased with the increase of the inclusion's μ_a. In addition, synthetic boundary data was generated by collecting the time-dependent values of the excitance. Measurements are gathered at eight different detector positions located at the boundary of the circular geometry (Fig. 2). Image reconstruction is then carried out with the DOT algorithm using the generated TD data and the same mesh used for the forward problem. At this point, it should be stressed that the inverse crime is strictly committed in the calculations (forward and inverse calculations employed the same mesh). This is done because the aim is to study the accuracy of the model (at each order N) in determining the corresponding set of optical parameters, without considering other sources of error, such as those
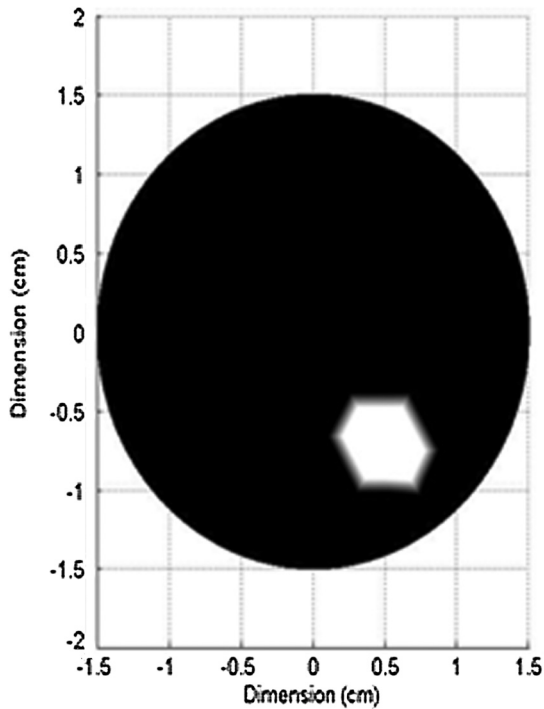
Fig. 3 Circular geometry with an embedded absorptive inclusion (in white).

introduced by using different mesh resolutions and/or noise in the measurements. To evaluate the image reconstructions, we introduce the relative error ε_r (in percent) in the maximum of the reconstructed values at the region of interest (inclusion or background medium) μ_{max}^{re}, with respect to the original value μ_{max}^{o} as $\varepsilon_r = 100 \cdot |\mu_{max}^{re} - \mu_{max}^{o}|/\mu_{max}^{o}$. The relative error ε_r is also used in the multi-parameter reconstructions (next section) with the same purpose.

Figure 4 consecutively displays the solution to the inverse problem posed in Eq. (14) for the medium shown in Fig. 3. In all calculations, the absorption coefficient lower and upper bounds [see Eq. (14)] are taken to be 0.01 and 5 cm^{-1}, respectively. For the three values of the inclusion's μ_a, 0.05, 0.1, and 1 cm^{-1} (left to right), the inverse problem solution is plotted for the orders $N = 1$, 3, and 5 (upper, middle, and lower rows, respectively). Order $N = 7$ was omitted from the exposed results because it did not introduce significant changes compared to order $N = 5$. In all cases, the algorithm used the optical properties of the homogeneous medium as the initial guess; no *a priori* information was employed.

Figure 4(a)–4(c) shows the reconstructed images using the SP$_1$ approximation can provide an accurate localization of the inclusion for the whole range of absorption coefficient values, and the background optical properties are accurately reconstructed; however, the absorption coefficient values of the heterogeneity are overestimated in all cases. The relative error ε_r values range from 10% to 27% (left to right cases). This result is in accordance with the known fact that DE-based algorithms start failing in regions with high absorption values because the diffusion approximation on which the DE is based does not hold in such regions.[32]

Figure 4(d)–4(f) shows the reconstructed images using the SP$_3$ approximation also accurately locate the position of the inclusion. The background optical properties are accurately recovered as well. The absorption coefficient values of the heterogeneity are underestimated in cases (d) and (f) and overestimated in case (e). The relative error ε_r values vary from 5% (middle figure) to 9%, and 13% (left and right figures).
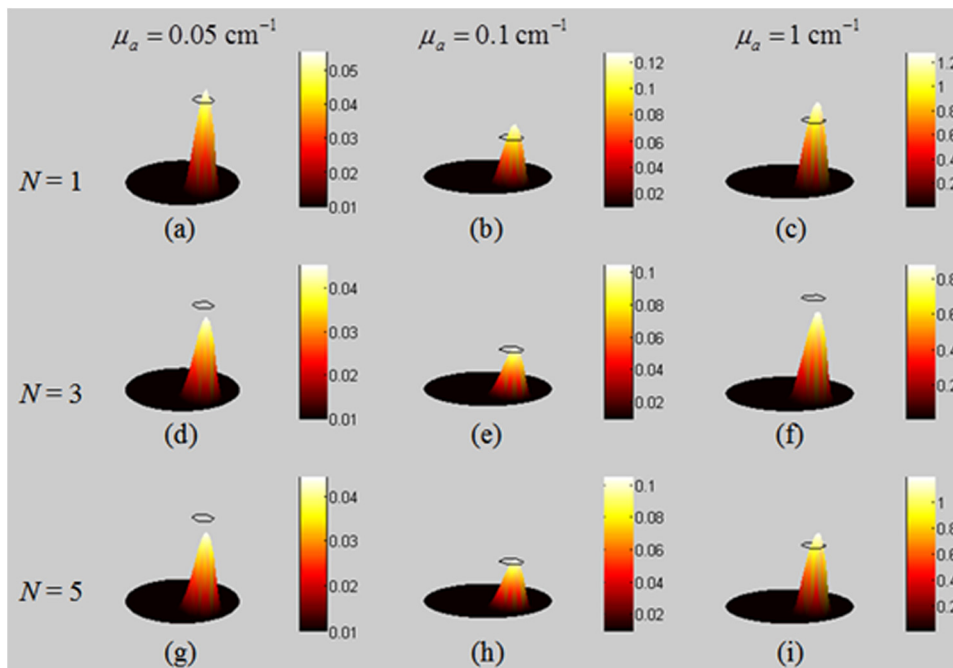

Fig. 4 Solution of the inverse problem for a homogeneous medium with an absorptive inclusion. The true values of the absorption coefficient for the inclusion are: 0.05 (1st column), 0.1 (2nd column), and 1 cm^{-1} (3rd column). Absorption of the background is 0.01 cm^{-1}. Solutions are plotted for the orders $N = 1$, 3, and 5 (upper, middle, and lower rows, respectively). For comparison, the black ring located at the top of the figures represents the inclusion's shape at a height corresponding to its absorption coefficient value. See text for precise numbers on the differences between reconstructed and true values.

Finally, Fig. 4(g)–4(i) shows the reconstructed images using the SP$_5$ approximation accurately locate the position of the inclusion and reconstruct the background optical properties. The reconstructed values of the inclusion's absorption coefficient are overestimated in cases (h) and (i) and underestimated in case (g). The relative errors are 10%, 5%, and 19% (from left to right).

From the numerical experiments, it can be deduced that SP_N orders higher than $N = 1$ (which corresponds to the DE) provide accurate localization of the absorptive heterogeneity and retrieval of the background optical properties. Moreover, they are able to better estimate high values of the absorption coefficient of the heterogeneity than the DE. Particularly, the reconstructed images using the SP$_3$ approximation provide the best estimates, with a reduction in the relative errors compared to the DE.

4.2 Simultaneous Reconstruction of Absorption and Diffusion Maps

The simultaneous reconstruction of absorption and diffusion coefficient maps will now be discussed. The numerical experiments are performed with the same homogeneous circular geometry as previously (Fig. 2), and the optodes (sources and detectors) are in the same configuration as before. The optical coefficient values of the homogeneous medium are $\mu_a = 0.01$ cm^{-1}, $\mu_s = 80$ cm^{-1}, and $g = 0.9$. These values correspond to a diffusion coefficient D of 0.0416 cm. As before, the refractive index of the medium is taken as 1.4, and the medium is considered as surrounded by air. Besides pursuing the same goal as before, the inverse crime is again committed.

Next, two irregular inhomogeneities are embedded in the medium, see Fig. 5(a) and 5(c) (the inhomogeneities are shown in two different maps but they are really in the same medium; this is because the scales of absorption and scattering
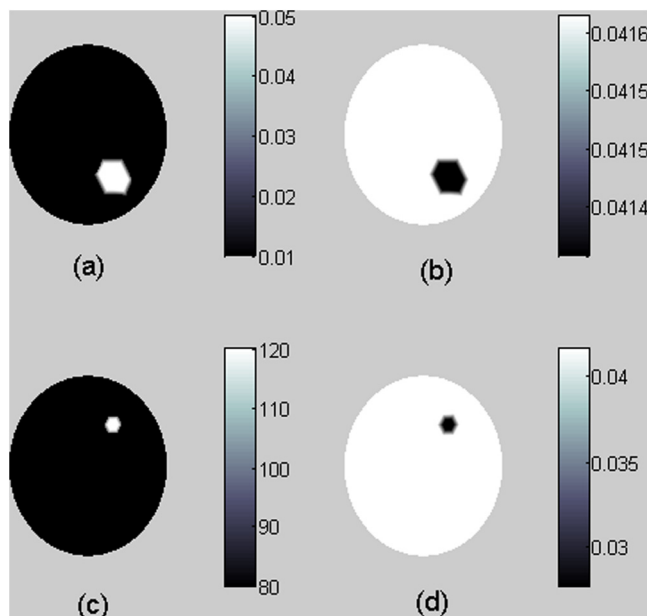
Fig. 5 (a) Absorption map depicting the absorptive inclusion ($\mu_a = 0.05$ cm^{-1}). (b) Diffusion map (D-map) for the absorptive inclusion displayed in (a). (c) Scattering coefficient distribution maps for the scattering inclusion ($\mu_s = 120$ cm^{-1}). (d) D-map for the scattering inclusion displayed in (c).

coefficients widely differ). The absorption coefficient of the first inclusion takes the increasing values of 0.05, 0.1, and 1 cm^{-1} [Fig. 5(a) shows the case for $\mu_a = 0.05$ cm^{-1}] with remaining optical properties identical to those of the homogeneous background medium. These values correspond to diffusion coefficient values of 0.0414, 0.0412, and 0.037 cm, respectively.

The second inclusion [Fig. 5(c)] has the same properties as the background medium, except for the scattering coefficient μ_s, which is assigned a value of 120 cm^{-1}. This value corresponds to $D = 0.0277$ cm. Note that the differences in the diffusion coefficient maps between the background and values at the inclusions are much higher for the scattering inclusion than the absorptive inclusion. As a consequence, for small absorption coefficient values, diffusion coefficient maps (D-maps) have to be carefully analyzed to appreciate the effect of absorptive inclusions. All optical coefficient values reported in this subsection are typical in small animal imaging (e.g., of commonly used in experimental sets).[58]

As in the previous subsection, the time-dependent fluence distribution is calculated for all the physical situations and the orders $N = 1, 3, 5$, and 7. Sources are modelled as previously. Differences in the fluence profiles partially delineate the zones where the inclusions are located. Excitance values are collected at eight detector positions located at the boundary of the circular geometry (Fig. 2). At this point, the absorption and the diffusion coefficient distributions are simultaneously reconstructed with the generated TD data. In the calculations, the absorption coefficient values [see Eq. (14)] are bound between 0.01 and 5 cm^{-1}, which is the range of typical values encountered in practice. In a similar way, the diffusion coefficient is bound to lie between 0.0001 and 0.0416 cm (same value as the background medium), respectively. Reported error values for the optical parameters in this section are calculated as in the absorption-only reconstructions.

Figure 6 shows the reconstructed absorption coefficient maps (μ_a-maps). For the three values of μ_a, 0.05, 0.1, and 1 cm^{-1} (left to right columns of images), the inverse problem solution is plotted for the orders $N = 1, 3, 5$, and 7 (first, second, third and fourth rows, respectively). In all cases, the algorithm uses the optical properties of the homogeneous medium as the initial guess; no *a priori* information is employed. In general, it can be seen that all orders ($N = 1, 3, 5$, and 7) provide an accurate localization of the inclusion for the whole range of μ_a values presented. In addition, background optical properties are acceptably recovered, with an almost indistinguishable (less than 1%) relative error at few nodes (the reference used to calculate this error is the background absorption coefficient value).

For the absorptive inclusion with $\mu_a = 0.05$ cm^{-1} (Fig. 6, first column of images), it is observed that the SP$_1$ and SP$_7$ approximations underestimate the value of μ_a at the inclusion with errors of 19% and 12%, respectively. On the other hand, the SP$_3$ and SP$_5$ approximations overestimate the value of μ_a at the inclusion with errors of 0.1% and 1%, respectively.

For $\mu_a = 0.1$ cm^{-1} (Fig. 6, second column of images), the SP$_1$ approximation underestimates the value of μ_a at the inclusion with a 16% error. In this physical situation, the orders $N = 3, 5$, and 7 overestimate the value of μ_a at the inclusion with errors of 8%, 12%, and 13%, respectively.

For $\mu_a = 1$ cm^{-1} (Fig. 6, last column of images), only the SP$_1$ approximation overestimates the absorption coefficient value at the inclusion with an error of 8%. The SP$_3$, SP$_5$,
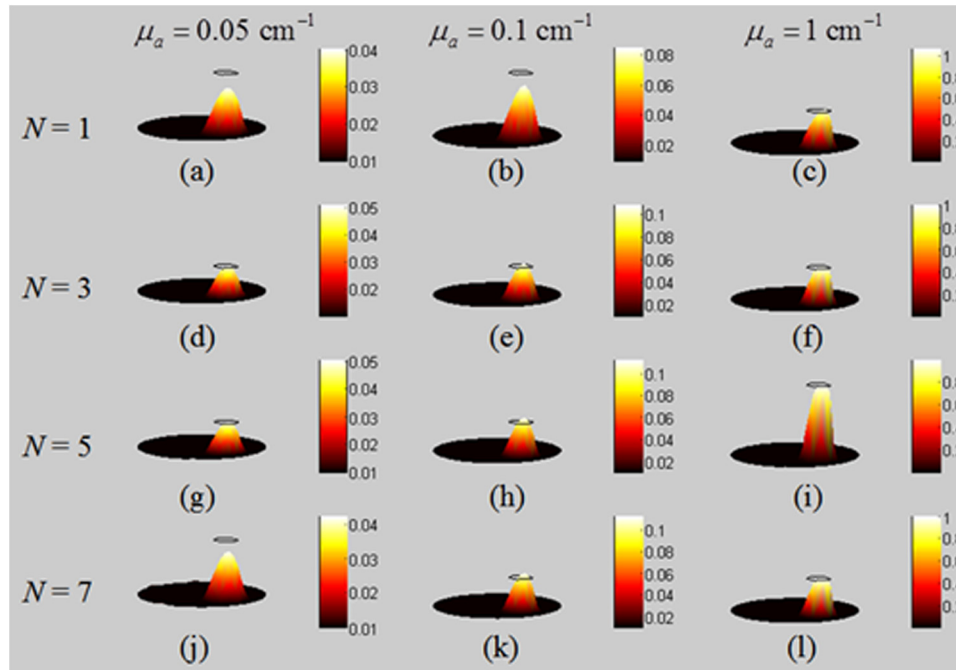
Fig. 6 Solution of the inverse problem(absorption coefficient) for the multiparametric case. Values of the absorption coefficient are: 0.05, 0.1, and 1 cm^{-1} (left to right columns of images). Images are plotted for the orders $N = 1, 3, 5$, and 7 (first, second, third, and fourth rows, respectively). For comparison, the black ring located at the top of the figures represents the inclusion's shape at a height corresponding to its absorption coefficient value. See text for precise numbers on the differences between reconstructed and true values.

and SP$_7$ approximations underestimate the absorption coefficient value at the inclusion, but with errors below 1%.

From the results, it may be inferred that in the case of multiparameter image reconstructions, orders higher than $N = 1$ provide an accurate localization of the absorptive heterogeneity and an acceptable quantitative retrieval of the background optical properties. In addition, they better quantify the inclusion's μ_a values than the DE. Reconstructed images using the SP$_3$ approximation give the best estimates, closely followed by the higher orders.

Figure 7 shows the reconstructed D-maps (refer back to Fig. 6(b) and 6(d) for the true diffusion coefficient distribution). For the absorptive heterogeneity values of: 0.05, 0.1, and 1 cm^1 (left to right), the inverse problem solution is plotted for the orders $N = 1, 3, 5$, and 7 (first, second, third, and fourth rows, respectively). The diffusion coefficient reconstructions will now be described for each value of the absorption coefficient of the absorptive inclusion.

For $\mu_a = 0.05$ cm^{-1} (Fig. 7, first column of images), all SP$_N$ orders image reconstructions are able to accurately localize the scattering inclusion with a 5% to 6% error. The SP$_3$ gives better results, but only by a 1% better error than the others. The presence of the absorptive inclusion appears in the reconstructed image with a moderate intensity (in light gray in the image; only the peak of the distribution is clearly seen). Errors at this location are 9%, 7%, 9%, and 10% for the orders $N = 1, 3, 5$, and 7, respectively. The background value of D (0.0416 cm) is acceptably recovered by all the SP$_N$ orders. Some artifacts can be seen near the image boundary after an inspection of the image. Artifacts are more pronounced in the case of the SP$_1$ reconstructed image and in the high absorption case (1 cm^{-1}).

For $\mu_a = 0.1$ cm^{-1} (Fig. 7, second column of images), all SP$_N$ image reconstructions are able to accurately localize the

scattering inclusion. Errors for SP$_1$, SP$_3$, and SP$_5$ are the same: 6%. SP$_7$ shows a 20% error but with a better spatially delimited inclusion. As before, the presence of the absorptive inclusion appears in the reconstructed image with moderate intensity, although this time the visible spot is bigger due to a higher value of the absorptive inclusion. These results show the crosstalk effect between the absorption and diffusion coefficients and how it varies with the SP$_N$ order (the effect is strongest for $N = 7$). Errors at this location are 16%, 14%, 15%, and 3% for $N = 1, 3, 5$, and 7, respectively. Small artifacts can be observed near the boundary, which again are more pronounced in the SP$_1$ reconstructed image.

For $\mu_a = 1$ cm^{-1} (Fig. 7, last column of images), all SP$_N$-based image reconstructions are able to accurately localize the scattering inclusion. However, the reconstructed images using orders $N = 3, 5$, and 7 qualitatively provide a better shaped, recovered scattering inclusion in comparison with the original distribution. The errors in the inclusion's scattering coefficient values are 6% for SP$_3$ and SP$_5$ and 15% for SP$_7$. In the case of SP$_1$ reconstruction, the error is greater than 40%, showing its failure to accurately quantify the effect of the scattering inclusion in the presence of high absorption values. An examination of the presence of the absorptive inclusion in the D-maps shows that the visible area has increased in size because of the large μ_a value of the absorptive inclusion (cross-talk). The area is better delimited by the SP$_3$ and SP$_7$ reconstructions, followed by SP$_5$ reconstruction. Finally, SP$_1$ shows a highly irregular distribution at the absorptive inclusion. Errors at this location are difficult to quantify because of the irregular distribution of recovered optical coefficient values. The following estimates are obtained: 18%, 27%, and 8% for SP$_3$, SP$_5$, and SP$_7$ reconstructions, respectively. In the case of SP$_1$ reconstruction, the error is greater than 57%. As in the previous results, all images present
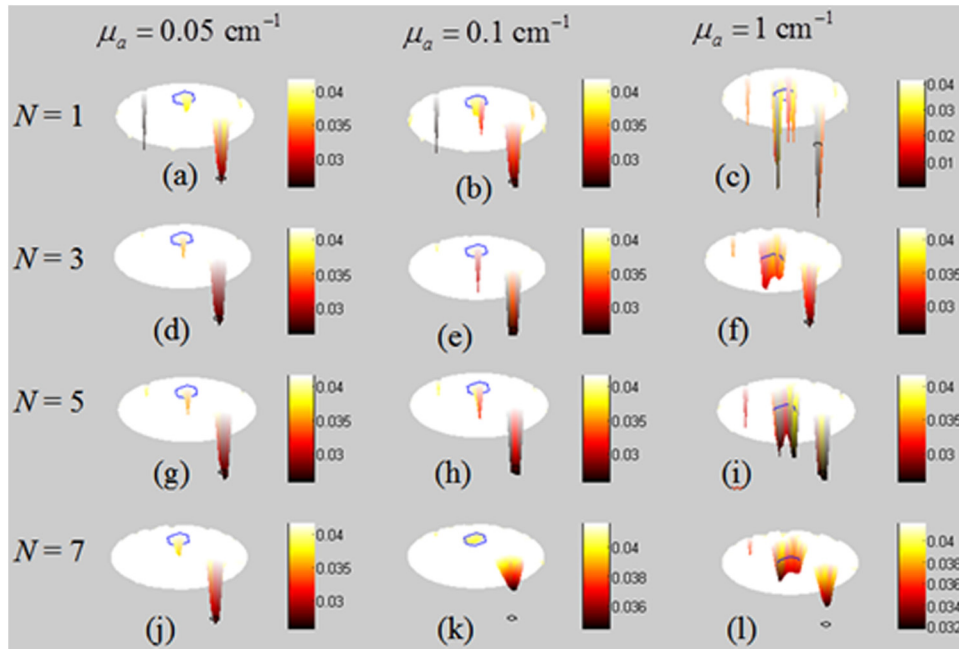
Fig. 7 Solution of the inverse problem(diffusion coefficient) for the multiparametric case. Images are plotted for the orders $N = 1$, 3, 5, and 7 (first, second, third, and fourth rows, respectively) where the absorptive heterogeneity takes the following values: 0.05, 0.1, and 1 cm^{-1} (left to right). These values are in turn associated with the following values of D for the absorptive inclusion: 0.0414, 0.0412, and 0.0370 cm, respectively. As regards the scattering inclusion, the value of D that should be recovered is 0.0277 cm. For comparison, a blue and a black ring are located in the figure to represent the absorption and the scattering inclusions' shape at a height corresponding to its diffusion coefficient values. See text for precise numbers on the differences between reconstructed and true values.

small artifacts near the boundary, with more pronounced artifacts in the SP$_1$ reconstructed image.

From the previous results, it may be inferred that in the case of multiparameter image reconstructions, orders higher than $N = 1$ provide an accurate localization of the scattering heterogeneity and acceptable retrieval of the background optical properties. The SP$_N$ diffusion coefficient reconstructions provide images with better shaped inclusions for $N > 1$ when compared with the original distributions. Also, SP$_N$ ($N > 1$) image reconstructions better quantify the diffusion coefficient values at the scattering inclusion than the DE reconstruction. In particular, the reconstructed images using the SP$_3$ approximation provide the best estimates for the diffusion coefficient values at the scattering inclusion. Furthermore, SP$_N$ ($N > 1$) image reconstructions better quantify the absorptive heterogeneity contribution in the image than the DE. Particularly, SP$_7$-based reconstructions give the best estimates for the diffusion coefficient values where the absorptive inclusion is located, but the size of the recovered inclusion is generally overestimated.

5 Comparison with Other Approaches

The results for multiparametric reconstructions obtained in the previous section will now be shortly compared with previously published works, Refs. 39, 42, 44, and 59. This is by no means an attempt to be exhaustive, but it will nevertheless give an idea about how our approach compares to others.

In Ref. 39, a DOT algorithm based on the FD SP$_N$ equations is presented. The forward model is discretized using the FEM (this builds on the SP$_N$-based forward model presented in Ref. 31). The DOT algorithm uses an L_2-norm Tikhonov minimization approach and a Levenberg-Marquardt procedure for updating the vector of optical properties. Multiparametric reconstructions (absorption and scattering coefficients) are performed with

synthetic (noiseless) data. The inverse crime is strictly and purposely committed in the numerical experiments. The multi-parametric reconstructions appearing in this work can be appreciated in details in the presented figures. Comparing the results presented in Sec. 3.4 of Ref. 39 with those presented herein, it is found that inclusion optical properties are better estimated by the algorithm presented here. In addition, the algorithm fails to recover the scattering heterogeneity position when the order $N = 7$ is employed. The presence of cross-talk effects and image artifacts at the boundary is also mentioned by the authors. For $N = 7$, image artifacts are located thorough the domain. Finally, high absorption regimes (~ 1 cm^{-1}) are not studied in Ref. 39.

In Ref. 59, the time-dependent radiative transfer equation (RTE) is used in a model-based DOT algorithm. The RTE is spatially discretized on a structured grid and the angular dependence is approximated by a discrete ordinates procedure called the piecewise parabolic method (PPM), which ensures numerical stability.[60] Adjoint methods are employed in the calculation of the gradient of an objective square error function, similar to f used here. The Polak–Ribière version of the nonlinear conjugate gradient method is used in the reconstructions, see details in the reference section of Ref. 59. 3D multiparameter reconstructions (absorption and scattering coefficients) using noiseless data are conducted. Although in two-dimensional (2D), the present algorithm provides better estimates of the inclusions optical properties for similar experimental cases. Hence, it may be expected that the present DOT algorithm can compete with that proposed in Ref. 59 in 3D cases. However, more conclusive experiments (same experimental conditions) would be needed, which is beyond the scope of the present work.

An analysis of the results presented in Refs. 42 and 44, where the RTE for FD problems is employed as the forward model and the inverse problem is posed as a PDE-constrained optimization,

support the competitiveness of our DOT algorithm. In these works, the solution of the inverse problem is sought through an augmented Lagrangian[42] and the reduced SQP[44] method. In Ref. 42, image artifacts and cross-talk effects which appear in the image reconstructions are difficult to evaluate for noise-free cases, mainly due to mesh effects (a different mesh is used for generating synthetic data and in the inverse problem solution). In Ref. 44, in the case of an absorptive inclusion and noise-free data, the background optical properties are not uniformly reconstructed, with large zones where the errors are around 20%. Finally, a considerable amount of image artifacts and cross-talk effects appears in the multiparametric reconstructions when 20 dB noise-added synthetic data is used. Hence, future works should be conducted to evaluate noise effects in our DOT algorithm and compared with these results in Ref. 44.

6 Conclusions

In this manuscript, a novel approach for solving inverse problems in time-domain diffuse optical tomography is presented. A finite element-finite difference discretized formulation of the time-dependent parabolic SP_N equations is used as the forward model. In this way, complex geometries and heterogeneous distributions of the optical coefficients can be imaged. To achieve image reconstruction, an optimization method which includes bounds on the optical coefficient values as additional linear constraints is implemented. The algorithm incorporates some necessary features for dealing with large-scale TD data. First, the implementation employs the nested analysis and design method to reduce the dimensionality of the problem. Therefore, even when TD data are dealt with, only the set of optical coefficients evaluated at the mesh nodes has to be recovered. Second, adjoint variables are introduced in the calculations. These reduce the computation time and provide an exact formula for the calculation of the gradient of the OF.

Several numerical experiments were performed for a small geometric medium with homogeneous background optical properties that mimic small animal imaging conditions. In a first round of experiments, an absorptive inclusion with increasing absorption coefficient values of 0.05, 0.1, and 1 cm^{-1} was positioned into an otherwise homogeneous medium. These values were chosen to simulate practical situations encountered in DOT and FDOT small animal imaging. Synthetic TD data were generated in the simulations and used to solve the inverse problem for all the SP_N orders studied in the literature. The inverse crime was committed in the numerical experiments, as the goal here was to evaluate the accuracy of each SP_N order in determining the optical coefficient distributions. Noise-free TD data are employed in the calculations for the same reason. An analysis of the results showed that, in all the physical situations considered, the image reconstructions based on the SP_N ($N > 1$) provide an accurate localization of the absorptive heterogeneity and retrieval of the background optical properties. In addition, the SP_N ($N > 1$) better estimated the absorption coefficient values of the heterogeneity than the DE. In the simulations, the reconstructed images using the SP_3 approximation provide the best estimates of absorption coefficient distributions. The DE consistently provided the worst estimates with an error of near 20% for high absorption values at the inclusion.

In a second round of experiments, an absorptive inclusion with increasing absorption coefficient values of 0.05, 0.1, and 1 cm^{-1} was placed in the homogeneous medium along with a second inclusion about half the size of the first, with optical properties identical to those of the homogeneous medium, except for the scattering coefficient to which a value of 120 cm^{-1} was assigned. Using synthetic data collected at the boundary of the medium, a simultaneous reconstruction of the absorption and diffusion coefficient distributions was performed. The inverse crime was committed again in the numerical experiments. The absorption coefficient reconstructions showed that orders higher than $N = 1$ provide accurate localizations of the absorptive heterogeneity and an acceptable retrieval of the background optical properties. Moreover, they better quantify the absorption coefficient values of the inclusion than the DE. The SP_3 approximation stood outas the model providing the best results, followed by $N = 5$ and 7. The cross-talk effect appeared in the reconstructed images, reinforcing the absorption coefficient values of the inclusion. The cross-talk effect increases with increasing order N. Similar results were gathered for the diffusion coefficient reconstructions. The image reconstructions based on the SP_N ($N > 1$) approximations provide accurate localizations of the scattering heterogeneity and acceptable retrieval of the background optical properties. In addition, they better quantify the values of the diffusion coefficient at the second inclusion (except for the SP_7), and the effect of the absorptive heterogeneity in the diffusion coefficient maps. Particularly, SP_3 approximation provides the best estimates for diffusion coefficient values at the scattering inclusion. The SP_7 approximation gives the best estimates for diffusion coefficient values at the absorptive inclusion when $\mu_a = 0.1$ and 1 cm^{-1}. The estimates of the optical properties of the absorptive inclusion by the DE reconstructions resulted in the worst values of all the orders, especially in the case of high absorption. In that case, the errors hamper any kind of estimation of the optical properties of scattering or absorptive inclusions. Cross-talk effects appeared in the reconstructed images in the form of small artifacts at the boundary, with these artifacts being more or less pronounced depending on the order. In the DE reconstructions, artifacts have a more significant contribution to the image.

The results obtained in this work indicate that our DOT algorithm can accurately substitute DE algorithms, especially in physical situations where the DE fails. The methods developed here can be used in further studies to 1. analyze the effect of other optical parameters in DOT reconstructions, such asthe anisotropy parameter, and perhaps eventually the refractive index, and 2. as a starting point for more elaborate optimization algorithms with other types offunctionals and additional constraints, and 3. it could eventually be reused with more accurate forward models.

Acknowledgments

Appendix A

In this Appendix, the structure of the matrices $\mathbf{C}$, $\mathbf{T}$, the matrix operator D_r and the source vector $\mathbf{Q}$ are provided up to the order $N = 7$, which is the highest SP_N order studied in the literature. In addition, the FEM matrices and vectors appearing in Eqs. (3) and (4) are given.

Matrix **C** has the form (Matlab notation is used for specifying the columns)

$$\mathbf{C}(:,1) = \begin{bmatrix} \mu_0(\mathbf{r}) \\ -\frac{2}{3}\mu_0(\mathbf{r}) \\ \frac{8}{15}\mu_0(\mathbf{r}) \\ -\frac{16}{35}\mu_0(\mathbf{r}) \end{bmatrix}, \quad \mathbf{C}(:,2) = \begin{bmatrix} -\frac{2}{3}\mu_0(\mathbf{r}) \\ \frac{4}{9}\mu_0(\mathbf{r}) + \frac{5}{9}\mu_2(\mathbf{r}) \\ -\frac{16}{45}\mu_0(\mathbf{r}) - \frac{4}{9}\mu_2(\mathbf{r}) \\ \frac{32}{105}\mu_0(\mathbf{r}) + \frac{8}{21}\mu_2(\mathbf{r}) \end{bmatrix},$$

$$\mathbf{C}(:,3) = \begin{bmatrix} \frac{8}{15}\mu_0(\mathbf{r}) \\ -\frac{16}{45}\mu_0(\mathbf{r}) - \frac{4}{9}\mu_2(\mathbf{r}) \\ \frac{64}{225}\mu_0(\mathbf{r}) + \frac{16}{45}\mu_2(\mathbf{r}) + \frac{9}{25}\mu_4(\mathbf{r}) \\ -\frac{128}{525}\mu_0(\mathbf{r}) - \frac{32}{105}\mu_2(\mathbf{r}) - \frac{54}{175}\mu_4(\mathbf{r}) \end{bmatrix},$$

$$\mathbf{C}(:,4) = \begin{bmatrix} -\frac{16}{35}\mu_0(\mathbf{r}) \\ \frac{32}{105}\mu_0(\mathbf{r}) + \frac{8}{21}\mu_2(\mathbf{r}) \\ -\frac{128}{525}\mu_0(\mathbf{r}) - \frac{32}{105}\mu_2(\mathbf{r}) - \frac{54}{175}\mu_4(\mathbf{r}) \\ \frac{256}{1225}\mu_0(\mathbf{r}) + \frac{64}{245}\mu_2(\mathbf{r}) + \frac{324}{1225}\mu_4(\mathbf{r}) + \frac{13}{49}\mu_6(\mathbf{r}) \end{bmatrix}.$$

(A1)

The matrix **T**, and consequently $\mathbf{T}^{-1}$, do not depend on the medium optical coefficients and have the following structure

$$\mathbf{T} = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} & -\frac{16}{35} \\ 0 & \frac{1}{3} & -\frac{4}{15} & \frac{8}{35} \\ 0 & 0 & \frac{1}{5} & -\frac{6}{35} \\ 0 & 0 & 0 & \frac{1}{7} \end{bmatrix}, \mathbf{T}^{-1} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 7 \end{bmatrix}. \quad (A2)$$

The matrix operator D_r is a diagonal 4×4 matrix operator whose diagonal elements are given by

$$\text{diag}(D_r)$$

$$= -\left[\nabla\left(\frac{1}{3\mu_1}\nabla\right), \quad \nabla\left(\frac{1}{7\mu_3}\nabla\right), \quad \nabla\left(\frac{1}{11\mu_5}\nabla\right), \quad \nabla\left(\frac{1}{15\mu_7}\nabla\right) \right],$$

(A3)

where the expression diag() is used to list the diagonal elements. Finally, the source vector $\mathbf{Q}(\mathbf{r}, t)$ contains the information about the source distribution $Q(\mathbf{r}, t)$

$$\mathbf{Q}(\mathbf{r}, t) = Q(\mathbf{r}, t) \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} & -\frac{16}{35} \end{bmatrix}^T. \quad (A4)$$

In the FEM, the domain of interest V is partitioned into l nonoverlapping elements τ_j, $j = 1, \ldots l$ joined at d vertex nodes. The boundary of V, denoted as ∂V, will be assumed to have d_b nodes. In the following, the index $l_N = (N+1)/2$ where N is the order of the SP_N approximation is used for convenience.

The square sparse matrix $\tilde{\mathbf{K}}$ is a block diagonal matrix composed of elemental "stiffness" matrices $\{\tilde{\mathbf{K}}_k\}_{k=1,\ldots,l_N}$ whose entries are given by

$$\tilde{K}_k(i, j) = \int_V \frac{1}{(4k-1)\mu_{2k-1}} \nabla u_i(\mathbf{r}) \cdot \nabla u_j(\mathbf{r}) dV,$$

$$k = 1, \ldots, l_N, i, j = 1, \ldots, d,$$

(A5)

where $\mu_k = \mu_a + \mu_s(1 - g^k)$ are the transport coefficients, see Ref. 32. The square sparse matrices $\tilde{\mathbf{M}}$, $\tilde{\Pi}$, and $\tilde{\mathbf{T}}$ are composed of $l_N \times l_N$ block "mass" matrices $\{\tilde{\mathbf{M}}_{k_1,k_2}\}_{k_1,k_2=1,\ldots,l_N}$, $\{\tilde{\Pi}_{k_1,k_2}\}_{k_1,k_2=1,\ldots,l_N}$, $\{\tilde{\mathbf{T}}_{k_1,k_2}\}_{k_1,k_2=1,\ldots,l_N}$ whose entries are, respectively given by

$$\tilde{M}_{k_1,k_2}(i, j) = \int_V C(k_1, k_2) u_i(\mathbf{r}) u_j(\mathbf{r}) dV,$$

$$k_1, k_2 = 1, \ldots, l_N, \quad i, j = 1, \ldots, d,$$

(A6)

$$\tilde{\Pi}_{k_1,k_2}(i, j) = \int_{\partial V} \frac{\Theta(k_1, k_2)}{(4k_1 - 1)\mu_{2k_1-1}} u_i(\mathbf{r}) u_j(\mathbf{r}) d\sigma,$$

$$k_1, k_2 = 1, \ldots, l_N, \quad i, j = 1, \ldots, d,$$

(A7)

$$\tilde{T}_{k_1,k_2}(i, j) = \int_V T(k_1, k_2) u_i(\mathbf{r}) u_j(\mathbf{r}) dV,$$

$$k_1, k_2 = 1, \ldots, l_N, \quad i, j = 1, \ldots, d.$$

(A8)

In Eq. (A6) and (A8), $C(k_1, k_2)$ and $T(k_1, k_2)$ are the elements of the matrices **C** and **T** given in Eqs. (A1) and (A2), respectively. In Eq. (A7), the term $\Theta = (k_1, k_2)$ represents the elements of the matrix $\Theta = (\mathbf{B})^{-1}\mathbf{A}$ which originates from the SP_N boundary conditions, see Ref. 32. The matrices **A** and **B** have the following form

$$\mathbf{A} = \begin{bmatrix} 1/2 + A_1 & -1/8 - C_1 & 1/16 - E_1 & -5/128 - G_1 \\ -1/8 - C_2 & 7/24 + A_2 & -41/384 - E_2 & 1/16 - G_2 \\ 1/16 - C_3 & -41/384 - E_3 & 407/1920 + A_3 & -233/2560 - G_3 \\ -5/128 - C_4 & 1/16 - E_4 & -233/2560 - G_4 & 3023/17920 + A_4 \end{bmatrix}, \quad (A9)$$

$$\mathbf{B} = \begin{bmatrix} (1 + B_1)/3\mu_1 & -D_1/\mu_3 & -F_1/\mu_5 & -H_1/\mu_7 \\ -D_2/3\mu_1 & (1 + B_2)/7\mu_3 & -F_2/\mu_5 & -H_2/\mu_7 \\ -D_3/3\mu_1 & -F_3/\mu_3 & (1 + B_3)/11\mu_5 & -H_3/\mu_7 \\ -D_4/3\mu_1 & -F_4/\mu_3 & -H_4/\mu_5 & (1 + B_4)/15\mu_7 \end{bmatrix}. \quad (A10)$$

The explicit expressions for the coefficients $(A_1, \ldots, H_1, A_4, \ldots, H_4)$ can be found in Appendix A of Ref. 29. The vectors $\tilde{\mathbf{F}}^{(m)}$ and $\tilde{\mathbf{\Gamma}}^{(m)}$ are composed of $\{\tilde{\mathbf{F}}_k^{(m)}\}_{k=1,\ldots,l_N}$ and $\{\tilde{\mathbf{\Gamma}}_k^{(m)}\}_{k=1,\ldots,l_N}$ "elemental load vectors" evaluated at the time step m. Their entries are

$$\tilde{\mathbf{F}}_k^{(m)}(i) = \int_V Q^{(m)}(k)u_i(\mathbf{r})\mathrm{d}V, \quad k = 1, \ldots, l_N, \quad i = 1, \ldots, d, \tag{A11}$$

$$\tilde{\boldsymbol{\Gamma}}_k^{(m)}(i) = \int_{\partial V} \frac{G^{(m)}(k)}{(4k-1)\mu_{2k-1}} u_i(\mathbf{r})\mathrm{d}\sigma, \quad k = 1, \ldots, l_N, \quad i = 1, \ldots, d. \tag{A12}$$

In Eq. (A11), $Q^{(m)}(k)$ are the components of the source vector Eq. (A4) evaluated at time step m. In Eq. (A12), $G^{(m)}(k)$ are the elements of the vector $\mathbf{G} = \mathbf{B}^{-1}\mathbf{S}$ where the column vector $\mathbf{S}$ is related to the transmitted contribution of the exterior source (the exterior source power multiplied by the transmission coefficient) $B_T(\mathbf{r}, \hat{\mathbf{s}}, t)$ as

$$\mathbf{S} = \begin{bmatrix} \int_{\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}>0} B_T(\mathbf{r}, \hat{\mathbf{s}}, t)2|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|\mathrm{d}\Omega \\ \int_{\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}>0} B_T(\mathbf{r}, \hat{\mathbf{s}}, t)\left[5|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^3 - 2|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|\right]\mathrm{d}\Omega \\ \int_{\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}>0} B_T(\mathbf{r}, \hat{\mathbf{s}}, t)\left[\frac{63}{4}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^5 - \frac{35}{2}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^3 + \frac{15}{4}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|\right]\mathrm{d}\Omega \\ \int_{\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}>0} B_T(\mathbf{r}, \hat{\mathbf{s}}, t)\left[\frac{429}{8}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^7 - \frac{693}{8}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^5 + \frac{315}{8}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|^3 - \frac{35}{8}|\hat{\mathbf{s}}\cdot\hat{\mathbf{n}}|\right]\mathrm{d}\Omega \end{bmatrix}, \tag{A13}$$

where $\hat{\mathbf{s}}$ and $\hat{\mathbf{n}}$ are unit vectors denoting direction and the normal to the boundary ∂V, respectively.

Appendix B

The vectors $\mathbf{j}_1$ and $\mathbf{j}_2$ have the following expressions

$$\begin{aligned}\mathbf{j}_1 = [&1/4 + J_0, (1/4 + J_0)(-2/3) \\ &+ (5/16 + J_2)(1/3), (1/4 + J_0)(8/15) \\ &+ (5/16 + J_2)(-4/15) + (-3/32 + J_4)(1/5), \ldots \\ &(1/4 + J_0)(-16/35) + (5/16 + J_2)(8/35) \\ &+ (-3/32 + J_4)(-6/35) + (13/256 + J_6)(1/7)], \end{aligned} \tag{B1}$$

$$\mathbf{j}_2 = \left[-\left(\frac{0.5 + J_1}{3\mu_1}\right), \left(-\frac{J_3}{7\mu_3}\right), \left(-\frac{J_5}{11\mu_5}\right), \left(-\frac{J_7}{15\mu_7}\right)\right], \tag{B2}$$

where the coefficients $\{J_i\}_{i=1,\ldots,4}$ linearly depend on the angular moments of the angle-dependent Fresnel reflection coefficient $R_F(\theta)$.[61] The explicit expressions for the coefficients $J_0, J_{1,\ldots},J_7$ can be found in Appendix A of Ref. 29.

References

1. H. Dehghani et al., "Numerical modelling and image reconstruction in diffuse optical tomography," *Phil. Trans. Royal Soc. A.* **367**(1900), 3073–3093 (2009).
2. S. R. Arridge and J. C. Schotland, "Optical tomography: forward and inverse problems," *Inverse Probl.* **25**(12), 123010 (2009).
3. F. S. Azar and X. Intes, Eds., *Translational Multimodality Optical Imaging*, Artech House Inc, Boston, MA (2008).
4. D. R. Leff et al., "Diffuse optical imaging of the healthy and diseased breast: a systematic review," *Breast Cancer Res. Treat.* **108**(1), 9–22 (2008).
5. H. Zhao et al., "Time-resolved diffuse optical tomography and its application to in vitro and in vivo imaging," *J. Biomed. Opt.* **12**(6), 062107 (2007).
6. J. C. Hebden, "Advances in optical imaging of the newborn infant brain," *Psychophysiology* **40**(4), 501–510 (2003).
7. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Phys. Med. Biol.* **50**(4), R1–R43 (2005).
8. R. Choe, A. Corlu, and K. Lee, "Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: a case study with comparison to MRI," *Med. Phys.* **32**(4), 1128–1139 (2005).
9. Q. Zhang and H. Jiang, "Three-dimensional diffuse optical tomography of simulated hand joints with a 64×64-channel photodiodes-based optical system," *J. Opt. A: Pure Appl. Opt.* **7**(5), 224–231 (2005).
10. D. A. Boas et al., "Imaging the body with diffuse optical tomography," *IEEE Sign. Process. Mag.* **18**(6), 57–75 (2001).
11. A. Bluestone et al., "Three-dimensional optical tomography of hemodynamics in the human head," *Opt. Express* **9**(6), 272–286 (2001).
12. M. Maris, E. Gratton, and J. Maier, "Functional near-infrared imaging of deoxygenated hemoglobin during exercise of the finger extensor muscles using the frequency-domain technique," *Bioimaging* **2**(4), 174–83 (1994).
13. D. K. Joseph et al., "Diffuse optical tomography system to image brain activation with improved spatial resolution and validation with functional magnetic resonance imaging," *Appl. Opt.* **45**(31), 8142–8151 (2006).
14. J. Wang et al., "Spectral tomography with diffuse near-infrared light: inclusion of broadband frequency domain spectral data," *J. Biomed. Opt.* **13**(4), 041305 (2008).
15. H. Zhao et al., "Time-resolved diffuse optical tomographic imaging for the provision of both anatomical and functional information about biological tissue," *Appl. Opt.* **44**(10), 1905–1916 (2005).
16. S. R. Arridge and W. R. B. Lionheart, "Nonuniqueness indiffusion-based optical tomography," *Opt. Lett.* **23**(11), 882–884 (1998).
17. B. Harrach, "On uniqueness in diffuse optical tomography," *Inverse Probl.* **25**(5), 055010 (2009).
18. L. Hervé, A. Koenig, and J.-M. Dinten, "Non-uniqueness in fluorescence-enhanced continuous wave diffuse optical tomography," *J. Opt.* **13**(1), 015702 (2011).
19. M. Schweiger et al., "Application of the finite element method for the forwardmodel in infrared absorption imaging," *Proc. SPIE* **1768**, 97–108 (1992).
20. L. Wang and H. Wu, *Biomedical Optics: Principles and Imaging*, 1st ed., Wiley-Interscience, Hoboken, NJ (2007).
21. F. Gao et al., "The forward and inverse models in time-resolved optical tomography imaging and their finite-element method solutions," *Image Vision Comput.* **16**(9–10), 703–712 (1998).
22. M. Schweiger and S. R. Arridge, "Application of temporal filters to time-resolved data in optical tomography," *Phys. Med. Biol.* **44**(7), 1699–1717 (1999).
23. F. E. W. Schmidt et al., "A 32-channel time-resolved instrument for medical optical tomography," *Rev. Sci. Instrum.* **71**(1), 256–265 (2000).

24. D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed., Springer Verlag, New York, p. 133 (1998).

25. F. Gao, H. Zhao, and Y. Yamada, "Improvement of quality in diffuse optical tomography by use of full time-resolved data," *Appl. Opt.* **41**(4), 778–791 (2002).

26. H. Zhao et al., "Time-resolved diffuse optical tomography and its application to in vitro and in vivo imaging," *J. Biomed. Opt.* **12**(6), 062107 (2007).

27. F. Gao et al., "Simultaneous fluorescence yield and lifetime tomography from time-resolved transmittances of small-animal-sized phantom," *Appl. Opt.* **49**(16), 3163–3172 (2010).

28. A. H. Hielscher, R. E. Alcouffe, and R. L. Barbour, "Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues," *Phys. Med. Biol.* **43**(5), 1285–1302 (1998).

29. A. Klose and E. Larsen, "Light transport in biological tissue based on the simplified spherical harmonics equations," *J. Comput. Phys.* **220**(1), 441–470 (2006).

30. Z. Yuan, H. Xin-Hua, and H. Jiang, "A higher order diffusion model for three-dimensional photon migration and image reconstruction in optical tomography," *Phys. Med. Biol.* **54**(1), 65–88 (2009).

31. M. K. Chu et al., "Light transport in biological tissue using three-dimensional frequency-domain simplified spherical harmonics equations," *Phys. Med. Biol.* **54**(8), 2493–2509 (2009).

32. J. Bouza Domínguez and Y. Bérubé-Lauzière, "Diffuse light propagation in biological media by a time-domain parabolic simplified spherical harmonics approximation with ray-divergence effects," *Appl. Opt.* **49**(8), 1414–1429 (2010).

33. A. Liemert and A. Kienle, "Analytical solutions of the simplified spherical harmonics equations," *Opt. Lett.* **35**(20), 3507–3509 (2010).

34. A. Liemert and A. Kienle, "Comparison between radiative transfer theory and the simplified spherical harmonics theory for the semi-infinite geometry," *Opt. Lett.* **36**(20), 4041–4043 (2011).

35. Y. Lu et al., "A parallel adaptive finite element simplified spherical harmonics approximation solver for frequency domain fluorescence molecular imaging," *Phys. Med. Biol.* **55**(16), 4625–4645 (2010).

36. J. Bouza Domínguez and Y. Bérubé-Lauzière, "Light propagation from fluorescent probes in biological tissues by coupled time-dependent parabolic simplified spherical harmonics equations," *Biomed. Opt. Express* **2**(4), 817–837 (2011).

37. L. D. Montejo, H. K. Kim, and A. H. Hielscher, "A finite volume algorithm for modeling light transport with the time-independent simplified spherical harmonics approximation to the equation of radiative transfer," *Proc. SPIE* **7896**, 78960J (2011).

38. Y. Lu et al., "Spectrally resolved bioluminescence tomography with the third-order simplified spherical harmonics approximation," *Phys. Med. Biol.* **54**(21), 6477–6493 (2009).

39. M. Chu and H. Dehghani, "Image reconstruction in diffuse optical tomography based on simplified spherical harmonics approximation," *Opt. Express* **16**(12), 17780–17791 (2009).

40. J. Tian et al., "Evaluation of the simplified spherical harmonics approximation in bioluminescence tomography through heterogeneous mouse models," *Opt. Express* **18**(20), 20988–21002 (2010).

41. D. Han et al., "Sparsity-promoting tomographic fluorescence imaging with simplified spherical harmonics approximation," *IEEE Tran. Biomed. Eng.* **57**(10), 2564–2567 (2010).

42. G. Abdoulaev, K. Ren, and A. H. Hielscher, "Optical tomography as a PDE-constrained optimization problem," *Inverse Probl.* **21**(5), 1507–153 (2005).

43. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Non-contact fluorescence optical tomography with scanning patterned illumination," *Opt. Express* **14**(14), 6516–6534 (2006).

44. H. K. Kim and A. H. Hielscher, "A PDE-constrained SQP algorithm for optical tomography based on the frequency-domain equation of radiative transfer," *Inverse Probl.* **25**(1), 015010 (2009).

45. H. K. Kim et al., "PDE-constrained multispectral imaging of tissue chromophores with the equation of radiative transfer," *Biomed. Opt. Express* **1**(3), 812–824 (2010).

46. H. K. Kim, J. H. Lee, and A. H. Hielscher, "PDE-constrained fluorescence tomography with the frequency-domain equation of radiative transfer," *IEEE J. Sel. Top. Quant.* **16**(4), 793–803 (2010).

47. S. B. Hazra, "PDE-constrained optimization methods," Chapter 3 in *Large-Scale PDE-Constrained Optimization in Applications, Lecture Notes in Applied and Computational Mechanics*, Vol. **49**, pp. 19–25, Springer-Verlag, Berlin, Germany (2010).

48. J. Nocedal and S. Wright, "Sequential quadratic programming," Chapter 18, in *Numerical Optimization*, pp. 529–561, Springer, New York, NY (2006).

49. A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, New York, NY (1978).

50. A. V. Patil, S. Mukherji, and U. B. Desai, "Optimal objective functional selection for image reconstruction indiffuse optical tomography," in *International Conference on Computing: Theory and Applications, 2007. ICCTA '07*, pp. 698–704 (2007).

51. N. Ducros, "Tomographie optique de fluorescence dans les milieux diffusants: apport de l'information temporelle," PhD Thesis, Université Claude Bernard (2009).

52. F. Nouizi et al., "Improvement of absorption and scattering discrimination by selection of sensitive points on temporal profile in diffuse optical tomography," *Opt. Express* **19**(13), 12843–12854 (2011).

53. L. T. Biegler et al., "Numerical experience with a reduced Hessian method for large scale constrained optimization," *Comput. Optim. Appl.* **15**(1), 45–67 (2000).

54. J. L. Hu et al.,"Sequential quadratic programming method for solution of electromagnetic inverse problems," *IEEE Trans. Antennas Propag.* **53**(8), 2680–2687 (2005).

55. A. H. Hielscher, A. D. Klose, and K. M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography," *IEEE Trans. Med. Imaging* **18**(3), 262–271 (1999).

56. D. C. Comsa, T. J. Farrell, and M. S. Patterson, "Quantitative fluorescence imaging of point-like sources in small animals," *Phys. Med. Biol.* **53**(20), 5797–5814 (2008).

57. J. Mobley and T. Vo-Dinh, "Optical properties of tissue," in *Biomedical Photonics Handbook*, T. Vo-Dinh, Ed., pp. 20–95, CRC Press, Boca Raton, FL (2003).

58. G. Alexandrakis, F. R. Rannou, and A. F. Chatziioannou, "Tomographic bioluminescence imaging by use of a combined optical-PET (OPET) system: a computer simulation feasibility study," *Phys. Med. Biol.* **50**(17), 4225–4241 (2005).

59. J. Boulanger and A. Charette, "Numerical developments for short-pulsednear-infrared spectroscopy. Part II: inverse treatment," *JQSRT* **91**(3), 297–318 (2005).

60. J. Boulanger and A. Charette, "Numerical developments for short-pulsed near-infrared spectroscopy. Part I: direct treatment," *JQSRT* **91**(2), 189–209 (2005).

61. M. Born and E. Wolf, *Principles of Optics*, 7th ed., Cambridge University Press, Cambridge, UK (2003).