INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **22** (2006) 1533–1559 doi:10.1088/0266-5611/22/5/002

Two-level domain decomposition methods for diffuse optical tomography

Kiwoon Kwon, Birsen Yazıcı and Murat Guven

Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, 110 8th St, Troy, NY 12180, USA

E-mail: yazici@ecse.rpi.edu

Received 26 March 2006
Published 24 July 2006
Online at stacks.iop.org/IP/22/1533

Abstract

Diffuse optical tomography (DOT) in the near infrared involves the reconstruction of spatially varying optical properties of a turbid medium from boundary measurements based on a forward model of photon propagation. Due to the nonlinear nature of DOT, high quality image reconstruction is a computationally demanding problem which requires repeated use of forward and inverse solvers. Therefore, it is desirable to develop methods and algorithms that are computationally efficient. In this paper, we develop two-level overlapping multiplicative Schwarz-type domain decomposition (DD) algorithms to address the computational complexity of the forward and inverse DOT problems. We use a frequency domain diffusion equation to model photon propagation and consider a nonlinear least-squares formulation with a general Tikhonov-type regularization for simultaneous reconstruction of absorption and scattering coefficients. In the forward solver, a two-grid method is used as a preconditioner to DD to enhance convergence. In the inverse solver, DD is initialized with a coarse grid solution to achieve local convergence. We show the strong local convexity of the nonlinear objective functional resulting from the inverse problem formulation and prove the local convergence of the DD algorithm for the inverse problem. We provide a computational cost analysis of the forward and inverse solvers and demonstrate their performance in numerical simulations.

1. Introduction

Diffuse optical image reconstruction based on the diffusion equation is a nonlinear ill-posed problem that calls for the use of nonlinear minimization methods with regularization to stabilize the solution [1].

Due to lack of analytical solutions for practical applications with arbitrary geometries, DOT image reconstruction is often posed as an optimization problem involving two coupled

steps, namely forward and inverse problems. Each step consists of an iterative solver whose solution is used as an input to the other solver. More precisely, the forward solver computes the photon density and its Jacobian with respect to the current optical coefficient estimates, and the inverse solver updates the optical coefficients based on the output of the forward step. The updated coefficients are then used in the forward solver to recompute the photon density and its Jacobian. As a result, the computational complexity of DOT image reconstruction quickly grows with the number of unknowns and dimension. Thus, real time computation of DOT requires numerical techniques to reduce the complexity of the problem.

In this paper, we develop two-level domain decomposition (DD) algorithms to address the computational complexity of the forward and inverse problems. More specifically, we develop an overlapping multiplicative Schwarz-type DD algorithm equipped with a two-grid preconditioner to solve the forward problem resulting from the finite element discretization of the frequency-domain diffusion equation. For simultaneous absorption and scattering coefficient reconstruction, we consider a nonlinear least-squares formulation with a general Tikhonov-type regularization. To solve the resulting optimization problem, we develop a two-level overlapping multiplicative Schwarz-type DD algorithm, where we use a trust region method for minimization. Finally, under the conditions that lead to the local strong convexity of the nonlinear objective functional considered for the inverse problem formulation, we prove the local convergence of the DD algorithm developed for the inverse problem.

DD methods originate from the Schwarz alternating procedure, which is known to be the first DD method applied for solving partial differential equations (PDEs) [32]. In the last two decades, motivated by the need for fast and efficient algorithms for solving large-scale, three-dimensional problems, DD methods have been extensively developed and applied in the area of numerical solution of PDEs [31, 33]. DD methods involve partitioning of bounded domains into two or more sub-domains, thereby dividing the original problem into a series of smaller sized sub-problems defined on the sub-domains. As a result, DD methods allow the parallel solution of the resulting sub-problems, making the computation extremely efficient.

Owing to the computational advantages they offer, DD methods have been applied to inverse problems as well. In [4–6, 36, 38], DD methods were applied to constrained convex minimization problems arising from variational inequalities, where the term 'space decomposition' was introduced for DD methods in the context of optimization problems. In [5, 36, 38], the convergence of DD methods has been shown for convex optimization problems. DD methods have also been applied to a number of application specific inverse problems, including Radon transform inversion in radar and x-ray tomography [8, 9], geophysics [14], parameter estimation problems [37], inverse heat conduction problem [29] and welding and metal cutting problems [23, 28]. In the area of DOT, a 'data driven zonation' method coupled with extended Kalman filtering was applied in conjunction with a DD method [15], where the convergence of the DD algorithm was shown empirically. In this work [15], instead of partitioning the main inverse problem formulation, independent local inverse problems were formulated on each non-overlapping sub-domain, using the measurement data due to the source–detector pairs that were physically present in the sub-domain. In the same work, a multi-grid algorithm was used to accelerate the finite difference solution of the forward problem; however, DD methods were not used. A similar multi-grid-based approach for the forward problem was also presented in [30]. Multi-grid algorithms in the context of optimization have been proposed for the solution of the inverse DOT problem as well [26, 40], however in both of these studies [26, 40], a multi-grid was used as an iterative solver and DD methods were not considered. In [16, 17], a fast adaptive composite-grid (FAC) algorithm was proposed for the linearized DOT inverse problem, which can be viewed as a fully overlapping Schwarz-type domain decomposition algorithm [13], to provide enhanced spatial resolution

in a region of interest. Most recently in [34], a multi-level overlapping Schwarz-type domain decomposition algorithm was proposed for the linearized inverse DOT problem, which was posed as a linear least-squares problem without regularization. In all these studies [16, 17, 34], DD methods were not pursued for the solution of the forward problem and convergence of the proposed algorithms was presented empirically.

In this work, we consider an overlapping partitioning of the optical domain to convert the finite element formulation of the frequency-domain diffusion equation into a number of smaller-sized sub-problems and use a multiplicative Schwarz-type DD algorithm. For the numerical solution of elliptic PDEs such as the frequency-domain diffusion equation, it has been shown that one-level domain decomposition methods are not efficient in conveying the information on one sub-domain to the others, which is attributed to the rapid decay of Green's function of these PDEs and the dependence of the solution on the boundary conditions [33]. An effective way to address this issue is to use multi-grid methods [33]. Therefore, we employ a two-grid preconditioner to enhance the global communication of the multiplicative Schwarz-type DD algorithm applied on the fine resolution level. We note that the multiplicative nature of the Schwarz algorithm provides a higher convergence rate as compared to that of the additive Schwarz algorithms (section 3.3). We refer to [10, 33] for a detailed discussion about the convergence analysis of DD methods with multi-grid preconditioners for the numerical solution of elliptic PDEs. For the inverse problem, we consider a nonlinear optimization problem resulting from the nonlinear least-squares formulation with a general Tikhonov-type regularization. Unlike the approach in [15], we formulate a single inverse problem on the whole domain using the boundary measurements due to all source–detector pairs. Then, we apply a two-level Schwarz-type multiplicative DD algorithm coupled with a trust region method to successively minimize the resulting objective functional on each overlapping sub-domain. In this case, the coarse level serves to provide a good initial guess for the fine level, where the inverse problem is formulated. Under some mild conditions, we show the local strong convexity of the nonlinear objective functional. Then, we prove the local convergence of the DD algorithm for the nonlinear inverse DOT problem by using the local strong convexity property of the nonlinear inverse problem formulation. In this context, the coarse-level solution used as the initial guess to the fine level is crucial in achieving the convergence. Note that we show the convergence properties of the DD algorithm for the inverse problem, independent of the optimization method. Therefore, the convergence properties do not change if the trust region method is replaced by another optimization method. We perform two sets of experiments to show the computational savings provided by the proposed DD algorithms as compared to non-DD algorithms.

The rest of the paper is organized as follows: in section 2, we define the forward problem and formulate the inverse problem, respectively. In section 3, we present the DD algorithms developed for the solution of the forward and inverse problems, and discuss the convergence properties and computational complexity of the proposed algorithms. Section 4 presents the numerical simulations and section 5 summarizes our results and conclusions. The paper includes an appendix for the proof of the local convergence of the DD algorithm developed for the inverse problem.

2. Diffuse optical tomography

2.1. Photon diffusion equation in the frequency domain

We model the propagation of light in biological tissue by the diffusion equation with Robin boundary conditions [1]. In the frequency domain, the photon diffusion equation is given as

follows:

$$\begin{aligned} -\nabla\cdot(\kappa\nabla\Phi)+\left(\mu_a+\frac{\mathrm{i}\omega}{c}\right)\Phi &= q \qquad \text{in} \quad \Omega \\ \Phi+2a\kappa\frac{\partial\Phi}{\partial\nu} &= 0 \qquad \text{on} \quad \partial\Omega \end{aligned} \tag{2.1}$$

where Ω is a Lipschitz domain in $\mathbb{R}^n$, $n = 2, 3$; $\partial\Omega$ is its boundary, Φ is the photon density, c is the speed of light, ω is the angular frequency of the source q, ν is the unit outward normal vector on the boundary, a is a parameter to account for the refractive index mismatch at the boundary, and μ_a, μ_s' and $\kappa = \frac{1}{3(\mu_a+\mu_s')}$ are the absorption, reduced scattering and isotropic diffusion coefficients, respectively. For the general anisotropic material, see [19].

The unique identification of the optical coefficients μ_a and κ in (2.1) when Dirichlet-to-Neumann map is given (or when infinite sources and infinite detectors are given) can be easily shown by using the uniqueness results for the isotropic case [35]. For the uniqueness of the optical coefficients when κ has an anisotropic anomalous region contained in a known background, see [18, 20, 22].

In [2, 3], numerical examples show that the Rytov approximation provides better DOT images as compared to the Born approximation. Therefore, we use the Rytov measurements [27] on the boundary:

$$\Gamma = \log\left(-\kappa\frac{\partial\Phi}{\partial\nu}\right) \tag{2.2a}$$

$$= \log\left(\frac{1}{2a}\Phi\right). \tag{2.2b}$$

Then, the Jacobian of Γ with respect to μ_a and μ_s' is given as follows:

$$\frac{\partial\Gamma}{\partial\mu_a}(r) = -\frac{1}{2a\Phi(r)}\int_\Omega [-3\kappa(r')^2\nabla G(r,r')\nabla\Phi(r')+G(r,r')\Phi(r')]\,\mathrm{d}r', \tag{2.3a}$$

$$\frac{\partial\Gamma}{\partial\mu_s'}(r) = -\frac{1}{2a\Phi(r)}\int_\Omega -3\kappa(r')^2\nabla G(r,r')\nabla\Phi(r')\,\mathrm{d}r' \tag{2.3b}$$

where $r\in\partial\Omega$, $\overline{\Phi}$ is the complex conjugate of Φ, and G is Green's function of (2.1).

2.2. Discretization and finite element method

Suppose that there are N_s sources located at r_j, $j = 1,\ldots,N_s$; and N_d detectors located on the boundary of Ω, at r_{i+N_s}, $i = 1,\ldots,N_d$. Let Φ_j be the solution of (2.1) for the point source $q^j(r) = \delta(r-r_j)$. Define

$$\Gamma_{i,j} = \log(\Phi_j(r_{i+N_s})), \tag{2.4}$$

for the ith detector and the jth source.

Consider the finite element space spanned by the bases $u_k, k = 1,\ldots,N_n$. Note that for a piecewise bilinear element, N_n is the same as the number of nodes. Let T_m denote the elements for $m = 1,\ldots,N_e$, where N_e is the number of the elements. Then, the finite element formulation for (2.1) in this finite element space for the point source q^j located at r_j, is given as follows:

$$\left[K+C+\frac{1}{2a}A\right]\Phi_j = f^j, \tag{2.5}$$

where K, C, A are $N_n \times N_n$ matrices and f^j is a $N_n \times 1$ column vector given by

$$K_{kl} = \int_\Omega \kappa \nabla u_k \nabla u_l$$
$$C_{kl} = \int_\Omega \left(\mu_a + \mathrm{i}\frac{\omega}{c}\right) u_k u_l$$
$$A_{kl} = \int_{\partial\Omega} u_k u_l$$
$$f_l^j = \begin{cases} 1 & \text{if} \quad l = j \\ 0 & \text{otherwise,} \end{cases}$$

for $k, l = 1, \ldots, N_n$.

Next, we define the following function spaces:

$$V^y = \{\eta \in L^2(\Omega) | L^y \leqslant |\eta(r)|_{r\in\Omega} \leqslant U^y\}, \tag{2.6a}$$

$$V_{N_e}^y = \{\eta \in V^y | \eta \text{ is constant at each } T_m, m = 1, \ldots, N_e\}, \tag{2.6b}$$

where y is either μ_a or μ_s', and L^y and U^y are positive constants. Note that we can choose different values for L^y and U^y depending on μ_a and μ_s'. Let $\eta \in V_{N_e}^y$ be $\eta = (\eta_1, \ldots, \eta_{N_e})$, then we define the norms in V^y and $V_{N_e}^y$ as follows:

$$\|\eta\|_{V^y} = \|\eta\|_{L^2(\Omega)} \text{ for } \eta \in V^y,$$

$$\|\eta\|_{V_{N_e}^y} = \sqrt{\sum_{m=1}^{N_e} \eta_m^2 |T_m|} \text{ for } \eta = (\eta_1, \ldots, \eta_{N_e}) \in V_{N_e}^y,$$

where $|T_m|$ is the area of T_m.

Assume that $\mu_a \in V_{Ne}^{\mu_a}$, $\mu_s' \in V_{Ne}^{\mu_s'}$. Let $\Phi_j(k)$ be the value of Φ_j at the kth node point, $G(k : j)$ be the value of Green's function at the kth node point due to the jth point source, and $\kappa(m)$ be the value of κ at the mth element T_m. By discretizing (2.3a), we obtain the value of the Jacobian of $\Gamma_{i,j}$ at the mth element as follows:

$$\frac{\partial \Gamma_{i,j}}{\partial \mu_a}(m) = -\frac{1}{2a\Phi_j(i)} \sum_{k,l=1}^{N_n} G(i : k)[-3\kappa(m)^2 E_m(k, l) + F_m(k, l)]\Phi_j(l), \tag{2.7a}$$

$$\frac{\partial \Gamma_{i,j}}{\partial \mu_s'}(m) = -\frac{1}{2a\Phi_j(i)} \sum_{k,l=1}^{N_n} G(i : k)[-3\kappa(m)^2 E_m(k, l)]\Phi_j(l), \tag{2.7b}$$

where

$$E_m(k, l) = \int_{T_m} \nabla u_k \nabla u_l, \tag{2.8a}$$

$$F_m(k, l) = \int_{T_m} u_k u_l. \tag{2.8b}$$

We note that in most practical applications, computing $G(i : k)$ for each point source located at the kth node and evaluating at the ith detector location is not feasible. Instead, we consider the adjoint problem associated with (2.1) [1], with the adjoint point source located at the ith detector position. Then, the solution $G^*(k : i)$ to the adjoint problem at the kth node for the adjoint source at the ith detector location satisfies $G^*(k : i) = G(i : k)$.

2.3. DOT as a nonlinear ill-posed optimization problem and the trust region method

Although the unique determination of μ_a and μ_s' from the measurement Γ at the boundary is possible for infinite sources and detectors when $\omega \neq 0$ and κ is known near the boundary [1], the unique solvability of inverse DOT problem for finite sources and detectors is not known. In addition, when the number of measurements is less than the total number of unknowns in the discretized inverse problem, the underdetermined problem challenges the determination of the optical coefficients. However, even in the overdetermined case, the inevitable ill posedness of the inverse problem may result in large perturbations in the reconstructed optical coefficients due to the small amount of noise in the measurements. In order to address the ill posedness, we use general Tikhonov-type regularization and consider the following nonlinear minimization problem:

$$(\mu_a, \mu_s') = \operatorname*{argmin}_{(\eta,\nu)\in V_{N_e}^{\mu_a}\times V_{N_e}^{\mu_s'}} F(\eta, \nu), \tag{2.9}$$

where

$$F(\eta, \nu) = \frac{1}{2}\sum_{j=1}^{N_s}\sum_{i=1}^{N_d}(\Gamma_{i,j}(\eta, \nu) - \mathcal{M}_{i,j})^2 + \alpha\Lambda(\eta, \nu). \tag{2.10}$$

In (2.10), $\mathcal{M}_{i,j}$ denotes the measurement at the ith detector due to the jth source, Λ is a non-negative operator from $V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$ into non-negative real numbers $\mathbb{R}^+ \cup \{0\}$, and $\alpha > 0$ is the regularization parameter. For example, Λ can be chosen as follows:

$$\Lambda(\eta, \nu) = \|L_1(\eta - \eta^*)\|^2_{V_{N_e}^{\mu_a}} + \|L_2(\nu - \nu^*)\|^2_{V_{N_e}^{\mu_s'}}, \tag{2.11}$$

where $(\eta^*, \nu^*) \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$ is a given prior for the optical coefficients, and L_1 and L_2 are $N_e \times N_e$ matrices. Note that depending on the choice of L_1 and L_2, (2.11) can be either zeroth- or first-order Tikhonov regularizer. For other regularization methods, see [19].

Note that (2.10) can be extended in a straightforward manner to include multi-frequency measurements. However, to simplify our notation, we will present the rest of our development for the single frequency measurements.

The optimization problem in (2.9) is composed of two steps; the step to determine the minimizing direction at the current coefficients (μ_a, μ_s'); and the step to perform a line search on those minimizing directions. In the Newtonian method, the minimizing direction is $-(F'')^{-1}F'$, where F' and F'' are the Jacobian and the Hessian of F with respect to (μ_a, μ_s'), respectively.

In this paper, we adopted the trust region method [12], one of the Newtonian approaches, which updates (μ_a, μ_s') in (2.9) iteratively as

$$(\mu_a, \mu_s') \leftarrow (\mu_a, \mu_s') + (\delta\mu_a, \delta\mu_s'),$$

by solving the following quadratic minimization problem formulated at the previous update of (μ_a, μ_s'):

$$(\delta\mu_a, \delta\mu_s') = \arg\min_{x\in K}\left\{\tfrac{1}{2}x^t F''(\mu_a, \mu_s')x + x^t F'(\mu_a, \mu_s')\right\}, \qquad \text{subject to} \quad \|Dx\| \leqslant T. \tag{2.12}$$

In (2.12), D is a scaling matrix, T is a trust region parameter and K is a subspace of $V^{\mu_a} \times V^{\mu_s'}$. The scaling matrix D is used to handle the constraints for the minimization. To stabilize the minimization, we control the trust region parameter T, similar to the way the parameter λ in the Levenberg–Marquardt method is controlled [7, 24, 25]. To avoid extensive computation,

the subspace K is chosen as the two-dimensional subspace composed of the gradient direction and the approximate Newton direction. This makes the trust region method suitable for large-scale constrained optimization problems like the one in (2.9).

Note that F' and F'' for a given (μ_a, μ_s') in (2.12) are given as follows:

$$F' = J^t b + \alpha \Lambda', \tag{2.13a}$$

$$F'' = J^t J + Hb + \alpha \Lambda'', \tag{2.13b}$$

where $b = \Gamma_{i,j}(\mu_a, \mu_s') - \mathcal{M}_{i,j}$, J and H are the Jacobian and the Hessian of the operator $\Gamma_{i,j}$, respectively. J is given by $J = J_{(i,j)}(m) = \left(\frac{\Gamma_{i,j}}{\partial \mu_a}(m), \frac{\Gamma_{i,j}}{\partial \mu_s'}(m)\right), i = 1, \ldots, N_d, j = 1, \ldots, N_s, m = 1, \ldots, N_e$ as in (2.7*a*). For the computation of H, see [21].

Taking the vanishing gradient point $x = (\delta\mu_a, \delta\mu_s') \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$ of the quadratic form $\frac{1}{2}x^t F''(\mu_a, \mu_s')x + x^t F'(\mu_a, \mu_s')$ in (2.12), we obtain

$$F''(\mu_a, \mu_s')(\delta\mu_a, \delta\mu_s') = -F'(\mu_a, \mu_s'). \tag{2.14}$$

Using (2.13*a*), we get

$$[J^t(\mu_a, \mu_s')J(\mu_a, \mu_s') + H(\mu_a, \mu_s')b + \alpha\Lambda''](\delta\mu_a, \delta\mu_s') = -[J(\mu_a, \mu_s')^t b + \alpha\Lambda']. \tag{2.15}$$

Thus, if $\delta\mu_a$ and $\delta\mu_s'$ are sufficiently small, the trust region method can be used to solve (2.12) at each iteration without considering the scaling matrix D.

We discuss in section 3.2 how we apply the two-level multiplicative DD method to solve the optimization problem in (2.9) and (2.10) by the trust region method.

3. Two-level domain decomposition methods for diffuse optical tomography

In this section, we describe the two-level domain decomposition methods considered in this paper, as applied to the forward and inverse problems. For simplicity and notational clarity, we will describe our notation and approach for the two-dimensional optical domain $\Omega = [a, b] \times [c, d] \subset \mathbb{R}^2$ and bilinear finite elements. Its extension to the three-dimensional domain is straightforward, by appropriate definition of operators and function spaces.

Let Ω^h denote the domain Ω that is uniformly divided by N_x times in the x-axis direction and N_y times in the y-axis direction. Thus, Ω^h has $N_n = (N_x + 1) \times (N_y + 1)$ nodes and $N_e = N_x \times N_y$ elements. We shall call Ω^H with $(N_x/2 + 1) \times (N_y/2 + 1)$ nodes and $N_x/2 \times N_y/2$ elements, the coarse level of Ω^h, assuming N_x and N_y are even.

Let Ω^h be decomposed into a disjoint union of d sub-domains $\Omega_l, l = 1, \ldots, d$ such that

$$\Omega^h = \bigcup_{l=1}^{d} \Omega_l. \tag{3.16}$$

Equation (3.16) describes the non-overlapping domain decomposition. For the overlapping domain decomposition, we define $\Omega_l^w, l = 1, \ldots, d$; an extension of Ω_l, recursively for all non-negative integers w as follows: $\Omega_l^0 = \Omega_l$, and Ω_l^w is the union of Ω_l^{w-1} and its adjacent elements in contact with the boundary of Ω_l^{w-1}, where w will be called the width of the overlapping region. Thus, the overlapping domain decomposition is given by

$$\Omega^h = \bigcup_{l=1}^{d} \Omega_l^w. \tag{3.17}$$

Figure 1 illustrates an overlapping and a non-overlapping domain decomposition for two sub-domains. Table 1 lists the acronyms and explanations for the algorithms developed in this paper.

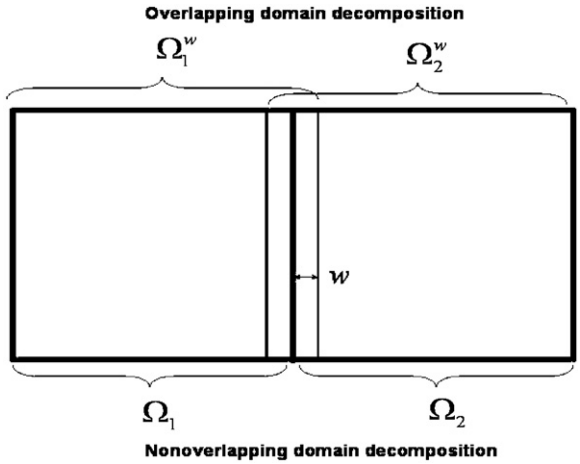

Figure 1. Overlapping domain decomposition and nonoverlapping domain decomposition.

Table 1. The acronyms and explanations of the algorithms.

Non-DDM	Non-domain decomposition method —trust region method for the inverse solver using (2.9) and (2.10). and finite element method for the forward solver without using domain decomposition and multigrid methods
MODDM	Multiplicative overlapping domain decomposition method
TMODDM	Two-level multiplicative overlapping domain decomposition method —the forward solver : multiplicative Schwarz method with two-grid preconditioner
MSDM	Multiplicative space decomposition method
TMSDM	Two-level multiplicative space decomposition method —the inverse solver : multiplicative Schwarz method with coarse-level initialization using (2.9) and (2.10).

3.1. Two-level multiplicative overlapping domain decomposition method for the forward problem

The forward problem is defined by the boundary value problem (2.1), which is approximated by the finite element formulation (2.5). In this work, we apply a two-level multiplicative overlapping domain decomposition method (TMODDM) to solve (2.5) in an attempt to reduce the computational complexity of the forward problem. In this context, the multi-level structure in TMODDM functions to speed up the convergence, while DD lets us formulate the forward problem as a sequence of smaller-sized problems, thereby reducing the computational requirements. Each sweep of TMODDM involves two steps: the *coarse-level correction* and the *sub-domain correction*. In the following discussion, we describe these steps in detail.

We list the notation and explanation for the variables used in the algorithm TMODDM in table 2.

3.1.1. The coarse-level correction: By formulating a smaller-sized problem on the coarse grid Ω^H, the coarse-level correction step provides an approximation to the error in the optical

Table 2. The notation used in the algorithm TMODDM.

Φ_j^0	Initial guess for the photon density for the jth source
$\Phi_j^{n-1+\frac{1}{d+1}}$	The update of Φ_j^{n-1} by coarse-level correction at the nth sweep of TMODDM
$\Phi_j^{n-1+\frac{p+1}{d+1}}$	The update of $\Phi_j^{(n-1)+\frac{1}{d+1}}$ after p successive sub-domain corrections on $\cup_{k=1}^{p}\Omega_k^w$ at the nth sweep of TMODDM

density estimate, which in turn is used to update the optical density on Ω^h. The coarse-level correction step can be explained in detail as follows:

Let Φ^n denote the current solution update for the optical density Φ on Ω^h after the nth sweep of TMODDM. The coarse-level correction in the $(n+1)$th sweep starts with computing the residual [11] on the fine grid Ω^h:

$$r_h \leftarrow \left(K + C + \frac{1}{2a}A\right)_h \Phi^n - f, \tag{3.18}$$

and restricting it onto the coarse grid Ω^H:

$$r_H \leftarrow R(r_h), \tag{3.19}$$

where $R : \mathbb{R}^{(N_x+1)\times(N_y+1)} \rightarrow \mathbb{R}^{(N_x/2+1)\times(N_y/2+1)}$ is the restriction operator. In this work, we use the full weighting restriction operator [39], which performs a nine-point weighted averaging. An approximation to the error in the solution update Φ^n can be obtained by solving the defect equation [39] formulated on Ω^H:

$$\left(K + C + \frac{1}{2a}A\right)_H e_H = r_H. \tag{3.20}$$

Then, the solution of (3.20) is used to update Φ^n, which completes the coarse-level correction:

$$\Phi^{n+\frac{1}{d+1}} \leftarrow \Phi^n + P(e_H), \tag{3.21}$$

where $P : \mathbb{R}^{(N_x/2+1)\times(N_y/2+1)} \rightarrow \mathbb{R}^{(N_x+1)\times(N_y+1)}$ is the prolongation operator. In this work, we use a bilinear prolongation operator [11].

3.1.2. Sub-domain correction. Following the coarse grid correction step, the sub-domain correction is performed on the fine grid Ω^h, by employing a multiplicative overlapping domain decomposition method (MODDM). In this step, the boundary value problem (2.1) is redefined on each of the overlapping sub-domains with appropriate boundary conditions. Then, the solution of the boundary value problem formulated on each sub-domain is used to update the optical density. In the following, we present the details of the sub-domain correction.

Let $\Phi^{n+\frac{p}{d+1}}$ be the current optical density estimate, obtained by updating $\Phi^{n+\frac{1}{d+1}}$ with the solutions of the boundary value problems formulated on the sub-domains $\Omega_1^w, \Omega_2^w, \ldots, \Omega_{p-1}^w$. Then, the boundary value problem on the sub-domain Ω_p^w is formulated as

$$-\nabla\cdot(\kappa\nabla v) + \left(\mu_a + \frac{\mathrm{i}\omega}{c}\right)v = q \qquad \text{in} \qquad \Omega_p^w \tag{3.22a}$$

$$v + 2a\kappa\frac{\partial v}{\partial v} = 0 \qquad \text{on} \qquad \partial\Omega_p^w \cap \partial\Omega, \tag{3.22b}$$

$$v = \Phi^{n+\frac{p}{d+1}} \qquad \text{on} \qquad \partial\Omega_p^w \backslash \partial\Omega. \tag{3.22c}$$

Then, the solution of (3.22*a*) is used to obtain the new optical density estimate on Ω as follows:

$$\Phi^{n+\frac{p+1}{d+1}} = \begin{cases} v & \text{in} \quad \Omega_p^w \\ \Phi^{n+\frac{p}{d+1}} & \text{in} \quad \Omega \backslash \overline{\Omega}_p^w. \end{cases} \tag{3.23}$$

Recursive application of this procedure for the rest of the sub-domains $\Omega_l^w : l = p+1, p+2, \ldots, d$ completes the sub-domain correction step and results in the optical density update Φ^{n+1} for the next TMODDM sweep. We refer to the pseudo-code in algorithm 1 for a summary of the coarse-level correction and sub-domain correction steps in TMODDM.

3.2. Two-level multiplicative overlapping space decomposition method for the inverse problem

In this work, we apply a two-level multiplicative overlapping space decomposition method (TMSDM) to solve the inverse problem in (2.9). We use the term 'space decomposition' to distinguish the DD method developed for the optimization problem (2.9) from TMODDM described in section 3.1. Although both methods involve domain decomposition; in the forward problem, the DD method is applied on a partial differential equation, whereas in the inverse problem, it is applied to an optimization problem.

In the forward problem, we obtain the value of the photon density at each node, whereas in the inverse problem, we are interested in the constant value of the optical coefficients on each element. Note that this is a direct consequence of the different discretization schemes followed in the forward and inverse problems. Even though TMSDM can be applied for any overlapping domain decomposition, in this work, we use the same domain decomposition in both forward and inverse problems to avoid introducing new notation.

Algorithm 1 TMODDM

N_s: The number of sources.
N_d: The number of detectors.
M_F: The maximum number of sub-domain correction sweeps.
for $j = 1, \ldots, N_s + N_d$ **do**
  Initialize Φ_j^0.
  for $n = 1, \ldots, M_F$ **do**
    Start coarse-level correction
    $r_h^j \leftarrow \left(K + C + \frac{1}{2a}A\right)_h \Phi_j^{n-1} - f^j$, {Compute the residual}
    $r_H^j \leftarrow R\left(r_h^j\right)$ Restrict the residual
    $e_H^j \leftarrow \left(K + C + \frac{1}{2a}A\right)_H^{-1} r_H^j$, {Compute the error on Ω^H}
    $\Phi_j^{n-1+\frac{1}{d+1}} \leftarrow \Phi_j^{n-1} + P\left(e_H^j\right)$ {Update the photon density}
    end coarse-level correction
    for $l = 1, \ldots, d$ **do**
      Update $\Phi_j^{n-1+\frac{l+1}{d+1}}$ at Ω_l^w by (3.22) and (3.23) {Sub-domain correction}
    end for
  end for
end for
Compute $\Gamma_{i,j}$ and $\frac{\partial \Gamma_{i,j}}{\partial \mu_a}(m)$, $\frac{\partial \Gamma_{i,j}}{\partial \mu_s'}(m)$, $m = 1, \ldots, N_e$ using (2.4) and (2.7).
{Post-processing}

Table 3. The notation used in the algorithm TMSDM.

μ^0	Initial guess for the optical coefficients $\mu = (\mu_a, \mu_s')$
$\mu^{n-1+\frac{1}{d+1}}$	The update of μ^{n-1} by the coarse-level inverse solver at the nth sweep of TMSDM
$\mu^{n-1+\frac{p+1}{d+1}}$	The update of $\mu^{(n-1)+\frac{1}{d+1}}$ after p successive sub-domain corrections on $\cup_{k=1}^{p}\Omega_k^w$ at the nth sweep of TMSDM

TMSDM consists of two steps: *coarse-level initiation* and *sub-domain correction*. The coarse-level initiation step provides an initial guess for the fine level, by solving (2.9) on the coarse level. The sub-domain correction involves the minimization of (2.9) on the fine level, iteratively on each sub-domain using a multiplicative space decomposition method (MSDM). In the following, we present the details of coarse-level initiation and sub-domain correction.

We list the notation and explanation for the variables used in the algorithm TMSDM in table 3.

3.2.1. Coarse-level initiation: By solving a smaller sized minimization problem formulated on the coarse grid Ω^H, the coarse-level initiation provides a computationally viable approximation for the optical coefficient estimates. This approximation is then used as an initial guess at the first sweep of the sub-domain correction step. In theorem 1, we show the local convergence of the sub-domain correction step under the condition that the initial guess is sufficiently close to the solution. Therefore, the coarse-level initiation is motivated to achieve the convergence of the sub-domain correction. The coarse-level initiation step can be described as follows:

Let $(\mu_a, \mu_s')^0$ denote the initial guess for the optical coefficients on Ω^h for the optical coefficients (μ_a, μ_s'). The coarse-level initiation starts with downsampling $(\mu_a, \mu_s')^0$ onto the coarse grid Ω^H:

$$(\mu_a, \mu_s')_H^0 \leftarrow D((\mu_a, \mu_s')^0), \tag{3.24}$$

where the downsampling operator $D : \mathbb{R}^{N_e} \to \mathbb{R}^{N_e/4}$ is defined by

$$\begin{aligned} D(h)(m_x, m_y) = \tfrac{1}{4}[&h(2m_x - 1, 2m_y - 1) + h(2m_x - 1, 2m_y) \\ &+ h(2m_x, 2m_y - 1) + h(2m_x, 2m_y)], \end{aligned} \tag{3.25}$$

for $1 \leqslant m_x \leqslant N_x/2$, $1 \leqslant m_y \leqslant N_y/2$, $h \in \mathbb{R}^{N_e}$, where $N_e/4 = N_x/2 \times N_y/2$ is the number of elements on the coarse level. Then, we formulate the inverse problem on the coarse grid with the initial guess $(\mu_a, \mu_s')_H^0$ and seek a solution to the following minimization problem:

$$(\mu_a, \mu_s')_H^{\frac{1}{d+1}} = \operatorname*{argmin}_{(\eta,\nu)\in V_{N_e/4}^{\mu_a}\times V_{N_e/4}^{\mu_s'}} F_H(\eta, \nu), \tag{3.26}$$

where $V_{N_e/4}^{y} = \{\eta \in V^y | \eta \text{ is constant at each } T_m, m = 1, \ldots, N_e/4\}$ for $y = \mu_a, \mu_s$ and $F_H : V_{N_e/4}^{\mu_a} \times V_{N_e/4}^{\mu_s'} \to \mathbb{R}$ is given by

$$F_H(\eta, \nu) = \frac{1}{2}\sum_{j=1}^{N_s}\sum_{i=1}^{N_d}(\Gamma_{i,j}(\eta, \nu) - \mathcal{M}_{i,j})^2 + \alpha\Lambda_H(\eta, \nu),$$

and $\Lambda_H : V_{N_e/4}^{\mu_a} \times V_{N_e/4}^{\mu_s'} \to \mathbb{R}^+ \bigcup\{0\}$. Then, the initial guess $(\mu_a, \mu_s')^{\frac{1}{d+1}}$ for the sub-domain correction step is obtained by upsampling $(\mu_a, \mu_s')_H^0$ to the fine grid Ω^h:

$$(\mu_a, \mu_s')^{\frac{1}{d+1}} \leftarrow U\big((\mu_a, \mu_s')_H^{\frac{1}{d+1}}\big), \tag{3.27}$$

where the upsampling operator $U : \mathbb{R}^{N_x/2 \times N_y/2} \to \mathbb{R}^{N_x \times N_y}$ in (3.27) is the bilinear prolongation operator.

3.2.2. Sub-domain correction: Following the coarse grid initiation step, the sub-domain correction is performed on the fine grid Ω^h. In this step, the inverse problem (2.9) is solved by successively minimizing the objective functional in (2.9) with respect to the unknown optical coefficients on each sub-domain. Thus, the optimization problem (2.9) is decomposed into a sequence of smaller minimization problems, using MSDM.

Before we give the details of the sub-domain correction step, we first introduce some notation. Recall that $(\eta, \nu) \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$. Then, for each $l = 1, \ldots, d$, we can decompose (η, ν) into orthogonal components as follows:

$$(\eta, \nu)_l = \begin{cases} (\eta, \nu) & \text{in} \quad \Omega_l^w \\ (0, 0) & \text{in} \quad \Omega \backslash \overline{\Omega}_l^w, \end{cases} \tag{3.28}$$

and

$$(\eta, \nu)_{l^\perp} = \begin{cases} (0, 0) & \text{in} \quad \Omega_l^w \\ (\eta, \nu) & \text{in} \quad \Omega \backslash \overline{\Omega}_l^w, \end{cases} \tag{3.29}$$

such that $(\eta, \nu) = (\eta, \nu)_l + (\eta, \nu)_{l^\perp}$.

The first sub-domain correction sweep is initiated by the update $(\mu_a, \mu_s')^{\frac{1}{d+1}}$ provided by the coarse-level initiation. Let $(\mu_a, \mu_s')^{n+\frac{p}{d+1}}$ be the estimate for the optical coefficients in the $(n+1)$th sub-domain correction sweep, obtained by updating $(\mu_a, \mu_s')^{n+\frac{1}{d+1}}$ with the solution estimates obtained on the sub-domains $\Omega_1^w, \Omega_2^w, \ldots, \Omega_{p-1}^w$. Then, the inverse problem on the sub-domain Ω_p^w reads

$$(\mu_a, \mu_s') = \underset{(\eta,\nu) \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}}{\operatorname{argmin}} F((\eta, \nu)_p + (\eta, \nu)_{p^\perp}), \tag{3.30}$$

subject to

$$(\eta, \nu)_{p^\perp} = (\mu_a, \mu_s')_{p^\perp}^{n+\frac{p}{d+1}},$$

where $F(\eta, \nu)$ is given by (2.10). Thus, the minimization is performed over only (η, ν) on Ω_p^w. Then, the solution of (3.30) is used to obtain the new optical coefficients update on Ω as follows:

$$(\mu_a, \mu_s')^{n+\frac{p+1}{d+1}} = \begin{cases} (\mu_a, \mu_s')_p & \text{in} \quad \Omega_p^w \\ (\mu_a, \mu_s')_{p^\perp}^{n+\frac{p}{d+1}} & \text{in} \quad \Omega \backslash \overline{\Omega}_p^w. \end{cases} \tag{3.31}$$

Successive application of this procedure for the rest of the sub-domains $\Omega_l^w : l = p+1, p+2, \ldots, d$ completes the $(n+1)$th sweep of the sub-domain correction step and results in the optical coefficient update $(\mu_a, \mu_s')^{n+1}$ for the next sub-domain correction sweep. We refer to the pseudo-code in algorithm 2 for a summary of the coarse-level initiation and sub-domain correction steps in TMSDM.

3.3. Convergence of TMODDM and TMSDM

Assume that the mesh size of the finite element formulation is $O(h)$ and that the sub-domains are of diameter $O(H)$ and the width of the overlapping region is $O(\delta H)$, where $0 \leqslant \delta < 1$. Then the following convergence behaviour is known for algorithm TMODDM (see section 2.5 in [33]).

(a) Convergence is poor if $\delta = 0$, but improves rapidly as δ increases.

(b) If δ is fixed, the number of sub-domain correction sweeps required for convergence is bounded independent of h, H and H/h.
(c) The number of sub-domain correction sweeps required for convergence is roughly half of that needed for the additive Schwarz method.

Algorithm 2 TMSDM

d: The number of sub-domains.
M_U: The maximum number of sub-domain correction sweeps.
Initialize $\mu^{\mathbf{0}}$ on Ω^h.
Start coarse-level initiation
$(\mu_a, \mu_s')_H^n \leftarrow D((\mu_a, \mu_s')^n)$
Solve $(\mu_a, \mu_s')_H^{n/(d+1)} = \arg\min_{(\eta,\nu)\in V_{N_e/4}^{\mu_a}\times V_{N_e/4}^{\mu_s'}} F_H(\eta, \nu)$
$(\mu_a, \mu_s')^{n+\frac{1}{d+1}} \leftarrow U\left((\mu_a, \mu_s')_H^{n+\frac{1}{d+1}}\right)$
end coarse-level initiation
for $n = 1, \ldots, M_U$ **do**
  for $l = 1, \ldots, d$ **do**
    $(\mu_a, \mu_s') = \arg\min_{(\eta,\nu)\in V_{N_e}^{\mu_a,l}\times V_{N_e}^{\mu_s',l}} F(\eta, \nu)$
    subject to $(\eta, \nu)_{l^\perp} = (\mu_a, \mu_s')_{l^\perp}^{n+\frac{l}{d+1}}$
    Update the optical coefficients by (3.31)
  end for
end for

Therefore, in terms of convergence, multiplicative Schwarz algorithms are more advantageous as compared to additive Schwarz algorithms. This follows from the fact that the multiplicative Schwarz algorithms take advantage of immediate use of the solution update in the successive neighbouring sub-problem. On the other hand, in additive Schwarz algorithms, one computes the solutions to the localized sub-problems independent of each other and uses the solution updates in the following sub-domain correction sweep. We refer to [31] and [33] for a detailed discussion of the Schwarz algorithms.

The local linear convergence of the algorithm MSDM is shown below, using the results in [5].

Theorem 1. *Let Λ be the regularization operator given in (2.11), and Λ' and Λ'' denote the first- and second-order Fréchet derivatives of Λ. Assume that there is a positive constant C_Λ such that $y^t\Lambda'' y \geqslant C_\Lambda y^t y$ for all $y \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$.*

Let $(\mu_a, \mu_s')^n$ be the nth step MSDM approximation of (μ_a, μ_s'). Assume that

$$\|(\mu_a, \mu_s') - (\mu_a, \mu_s')^q\|_{V_{N_e}^{\mu_a}\times V_{N_e}^{\mu_s'}} \leqslant \frac{\alpha C_\Lambda}{C_H C_J}, \qquad (q = 0, \ldots, n). \tag{3.32}$$

Then,

$$\|(\mu_a, \mu_s')^n - (\mu_a, \mu_s')\|_{V_{N_e}^{\mu_a}\times V_{N_e}^{\mu_s'}} \leqslant \left(\frac{8}{9}\right)^n \frac{\alpha^2 C_\Lambda}{C_H^2 C_J}[C_\Lambda C_J + \widetilde{C}_\Lambda C_H], \tag{3.33}$$

where $\widetilde{C}_\Lambda = \|\Lambda'\|_{V_{N_e}^{\mu_a}\times V_{N_e}^{\mu_s'}\to\mathbb{R}}$ and C_J, C_H are positive constants such that

$$\|J(\mu_a, \mu_s')\|_{V_{N_e}^{\mu_a}\times V_{N_e}^{\mu_s'}\to\mathbb{R}^{N_s\times N_d}} \leqslant C_J, \tag{3.34a}$$

$$\|H(\mu_a, \mu_s')\|_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'} \to \mathbb{R}^{N_s \times N_d}} \leqslant C_H. \tag{3.34b}$$

Proof. See the appendix for the proof of theorem 1.

Note that C_J and C_H depend on the frequency ω. For the multi-frequency measurement case with N_ω frequencies $\omega_1, \ldots, \omega_{N_\omega}$, let

$$C_J = \max_{i=1,\cdot,N_\omega} C_J^{\omega_i}, \qquad C_H = \max_{i=1,\cdot,N_\omega} C_H^{\omega_i}, \tag{3.35}$$

where $C_J^{\omega_i}$ and $C_H^{\omega_i}$ are the frequency-dependent norm bounds. Then, theorem 1 can be extended to the multi-frequency measurement case in a straightforward manner, by replacing the norm bounds in (3.34*a*) and (3.34*b*) by the bounds given in (3.35).

For the general Tikhonov regularizer, Λ is chosen as

$$\Lambda(\eta, \nu) = \|L_1(\eta - \eta^*)\|^2_{V_{N_e}^{\mu_a}} + \|L_2(\nu - \nu^*)\|^2_{V_{N_e}^{\mu_s'}}.$$

Therefore, if L_1 and L_2 are positive definite, the assumption of $y^t \Lambda'' y \geqslant C_\Lambda y^t y$ in theorem 1 holds since

$$(\eta, \nu)^t \Lambda''(\eta, \nu) = \eta^t L_1 \eta + \nu^t L_2 \nu.$$

Clearly, this assumption holds for the zeroth-order Tikhonov regularization. For the appropriate choice of the discrete approximation of the differential operator, the first-order Tikhonov regularizer also satisfies the assumption in theorem 1.

Note that in [21], we have shown that the first- and second-order Fréchet derivatives of the coefficient-to-measurement operator with respect to (μ_a, μ_s') exist and are bounded. Similarly, J and H, being finite approximations to the first- and second-order Fréchet derivatives, satisfy (3.34*a*) and (3.34*b*). □

3.4. Computational cost of the algorithms

In this subsection, the computational cost of the proposed algorithms is analysed. We will call the method solving (2.5) and (2.15) without any domain decomposition the non-DD method and compare it with the proposed methods.

The system matrices for equations (2.5) and (2.15) are $2N_n \times 2N_n$ and $2N_e \times 2N_e$, respectively. Assume that we require $O\left(N_n^q\right)$ and $O\left(N_e^q\right)$ $(1 \leqslant q \leqslant 3)$ floating point operations for each equation with the non-DD method. Note that for a full nonzero matrix, $q = 3$ and for a diagonal matrix, $q = 1$. Suppose that we use d sub-domains with equal nodes and elements for the proposed algorithms.

The coarse-level computation needs 4^{-q} times (in two dimensions) or 8^{-q} times (in three dimensions) the computations required on the fine level computation for the forward and inverse solvers. Thus, if we neglect the coarse-level computation, the computational cost of algorithm MODDM is $M_F d^{1-q}$ (for one computer) and $M_F d^{-q}$ (for d parallel computers) times that of the non-DD method, where M_F is the maximum number of sub-domain correction sweeps. Recall that M_F is the maximum sweep of all sub-domain corrections followed by the coarse grid correction. It is well known that for a given M_F, the convergence behaviour of TMODDM is independent of N_n, d and N_n/d (see (a)–(c) below subsection 3.3). Furthermore TMODDM has good convergence behaviour and it has better convergence behaviour for smaller M_F as compared to MODDM. In this work, we have chosen $M_F = 3$ when TMODDM is used in the inverse solver in section 4.2. This analysis shows that by using multiple sub-domains, algorithm TMODDM can achieve significant reduction in computational requirements.

Table 4. The comparison of the computational costs for the non-DD method, TMSDM on one, d, and d^2 computers. q is a constant between 1 and 3 depending on the sparsity of the system matrix. In the table, N_n is the number of nodes, N_e is the number of elements, d is the number of sub-domains, M_F is the number of sweeps for the forward solver and M_U is the number of sweeps for the inverse solver.

	Computational cost	
Method	Forward solver	Inverse solver
Non-DDM	$O(N_n^q)$	$O(N_e^q)$
TMSDM on 1 computer	$M_F d^{1-q} O(N_n^q)$	$M_U d^{1-q} O(N_e^q)$
TMSDM on d computers	$M_F d^{1-q} O(N_n^q)$	$M_U d^{-q} O(N_e^q)$
TMSDM on d^2 computers	$M_F d^{-q} O(N_n^q)$	$M_U d^{-q} O(N_e^q)$

The computational cost of algorithm TMSDM with algorithm TMODDM as the forward solver, and the computational cost of the non-DD method are tabulated in table 4, using 1, d and d^2 computers. When d parallel computers are used, the parallel computing is applied to the inverse problem only, whereas when d^2 computers are used, the parallel computing is applied to both the inverse and forward problems. Note that this comparison for the inverse solver is based on the assumption that the speed of data communication between parallel computers is sufficiently fast. Parallel computation is not treated in this paper, but the proposed algorithms combined with the parallel computing are expected to provide more efficient results.

4. Numerical simulations

In this section, we demonstrate the performance of the algorithms TMODDM and TMSDM using simulated optical data. In section 4.1, we conduct a series of experiments to test and compare the performance of TMODDM to that of the one-level DD method. In section 4.2, we test TMSDM in a number of experiments, where we consider μ_a-only, μ_s'-only and simultaneous μ_a and μ_s' reconstructions. For the summary of method descriptions, we refer to table 1.

In our experiments, we set the angular frequency $\omega = 2\pi \times 100$ MHz, and $a = 1$, except for the simultaneous absorption and reduced scattering imaging case, where we use two frequencies: $\omega_1 = 300$ MHz and $\omega_2 = 500$ MHz. Note that for the μ_a-only simulation, zero-mean Gaussian noise with standard deviation equal to 1% of the average of the measurements was added to the measurements. Similarly, for the μ_s-only case, we added zero-mean Gaussian noise with standard deviation equal to 0.5% of the average of the measurements. For the simultaneous μ_a and μ_s' imaging, we considered additive zero-mean Gaussian noise with standard deviation equal to 5% of the average of the measurements. For the inverse problem formulation, we chose the following zeroth-order Tikhonov regularizer with the regularization parameter $\alpha = 10^{-2}$:

$$\Lambda(\eta, \nu) = \|\eta\|^2_{V_{N_e}^{\mu_a}} + \|\nu\|^2_{V_{N_e}^{\mu_s'}}, \tag{4.36}$$

for $\eta \in V_{N_e}^{\mu_a}$ and $\nu \in V_{N_e}^{\mu_s'}$.

4.1. Algorithm TMODDM

In this experiment, we evaluate the performance of TMODDM, for an optical medium whose absorption and scattering coefficient distributions are shown in figures 2(a) and (b),

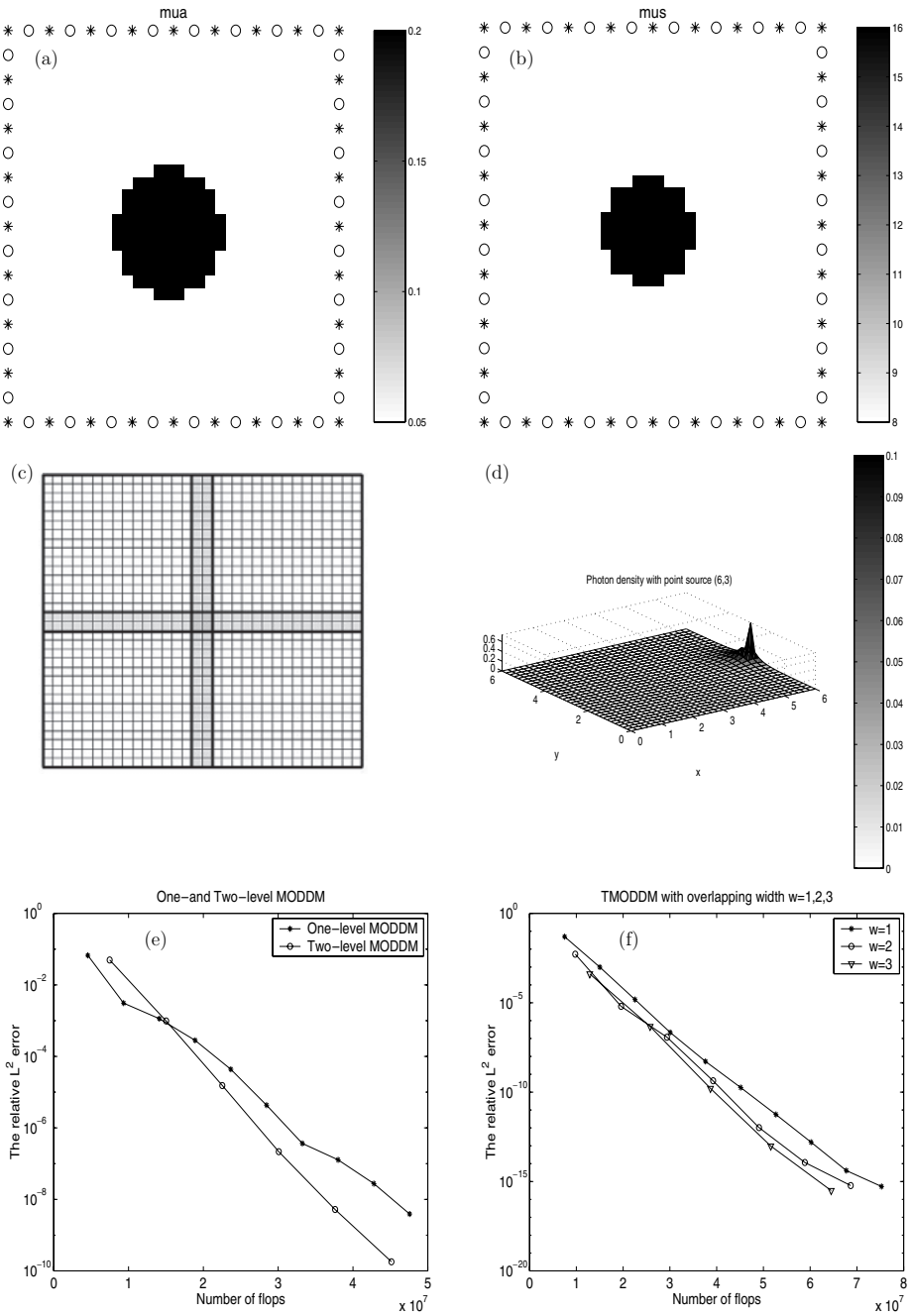

Figure 2. (a) The original μ_a image. The white region represents background tissue with μ_a value of 0.05 cm^{-1}, and the black region represents an anomaly with μ_a value of 0.2 cm^{-1}. (b) The original μ_s' image. The white region represents background tissue with μ_s' value of 8 cm^{-1}, and the black region represents an anomaly with μ_s' value of 16 cm^{-1}. (c) 2×2 domain decomposition of $\Omega = [0, 6] \times [0, 6]$. (d) The LINPACK solution of the frequency-domain diffusion equation for a point source located at $(6, 3)$. (e) Relative L_2 error versus number of floating point operations for one- and two-level MODDM. (f) Relative L_2 error versus number of floating point operations for TMODDM with domain decompositions with the overlap region width $w = 1, 2$ and 3 pixels.

respectively. The black region in figure 2(a) shows an anomaly with absorption coefficient $\mu_a = 0.2$ cm^{-1}, where the background has an absorption coefficient $\mu_a = 0.05$ cm^{-1}. Figure 2(b) shows the reduced scattering coefficient of the same medium, where the black region and the background have reduced scattering coefficients $\mu_s' = 16$ and $\mu_s' = 8$ cm^{-1}, respectively. We divide the square domain $\Omega = [0, 6] \times [0, 6]$ cm^2 into 32×32 uniform pixels. Next, we decompose the domain into four overlapping sub-domains with $2/3$ cm of overlap region, as shown in figure 2(c).

We consider the solution of the diffusion equation (2.1) for a point source located at (6, 3) and compare the performance of TMODDM and the one-level DD method (MODDM) using the following relative $L^2(\Omega)$ error:

$$\frac{\|\Phi^A - \Phi^L\|_{L^2(\Omega)}}{\|\Phi^L\|_{L^2(\Omega)}}, \tag{4.37}$$

where Φ^L is the LINPACK solution of (2.5), and Φ^A is either the TMODDM or MODDM solution. Figure 2(d) shows the LINPACK solution for the point source located at (6, 3). Figure 2(e) shows the relative $L^2(\Omega)$ error versus the number of floating point operations for TMODDM and MODDM. We observe that TMODDM converges faster than the one-level MODDM, with a smaller number of floating point operations.

Next, to show the effect of overlap width on the performance of TMODDM, we consider three cases where the overlap region consists of 1, 2 and 3 pixels, respectively. Figure 2(f) shows the relative $L^2(\Omega)$ error versus the number of floating point operations for each overlap width. We see that the convergence improves as the width gets larger. This verifies the convergence result stated in (a) below subsection 3.3.

4.2. Algorithm TMSDM

In this section, we evaluate the performance of TMSDM in two sets of experiments using simulated data. In the first experiment, we consider μ_a-only and μ_s'-only reconstructions and compare TMSDM to the non-DD method. In the second experiment, we consider simultaneous reconstruction of μ_a and μ_s' using TMSDM and compare the performance to that of the non-DD method. For the description of the non-DD method, we refer to table 1.

For the forward solver, we use TMODDM with the maximum number of sub-domain correction sweeps M_F set to 3. In order to evaluate the performance of the inverse solvers and compare TMSDM to the non-DD method, we define the following signal-to-noise-ratio (SNR):

$$\text{SNR} = 20 \log_{10} \frac{\|y\|_{L^2(\Omega)}}{\|y^A - y\|_{L^2(\Omega)}}, \tag{4.38}$$

where y^A is the actual optical coefficient (either μ_a or μ_s') and y denotes the reconstructed optical coefficients, obtained by using either TMSDM or the non-DD method.

4.2.1. Experiment 1. In this experiment, we consider two cases. In the first one, we assume that the reduced scattering coefficient of the medium is known and set it to 8 cm^{-1}, and we reconstruct the absorption coefficient of the medium. Figure 3(a) shows the circular heterogeneity with $\mu_a = 0.20$ cm^{-1} embedded in a background with $\mu_a = 0.05$ cm^{-1}.

In the second case, we assume that the absorption coefficient of the medium is known and is set to $\mu_a = 0.05$ cm^{-1} and we reconstruct the reduced scattering image. Figure 4(a) shows the circular heterogeneity with $\mu_s' = 16$ cm^{-1} embedded in a background with $\mu_s' = 8$ cm^{-1}.

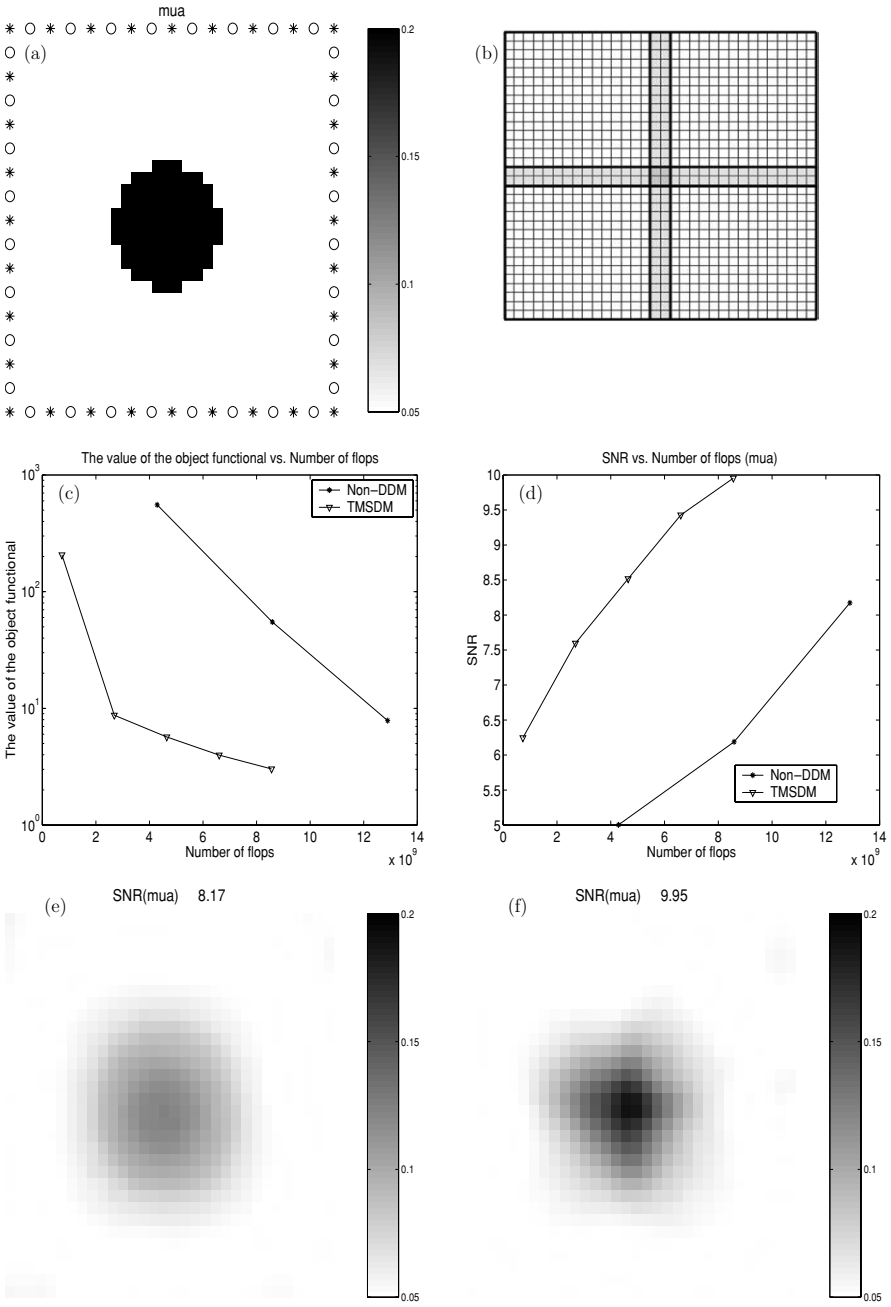

Figure 3. The set-up and the results regarding the μ_a-only reconstruction. (a) The original μ_a image and the source–detector configuration. The white region represents background tissue with μ_a value of 0.05 cm^{-1}, and the black region represents an anomaly with μ_a value of 0.2 cm^{-1}. (b) The 2×2 domain decomposition and uniform 32×32 discretization of Ω. (c) The value of the objective functional versus the number of floating point operations of the non-DD method and TMSDM. (d) The SNR versus the number of floating point operations of the non-DD method and TMSDM. (e) The absorption image reconstruction by the non-DD method after 12.893×10^9 floating point operations, with SNR = 8.174. (f) The reconstruction by TMSDM after 8.561×10^9 floating point operations with SNR = 9.955.

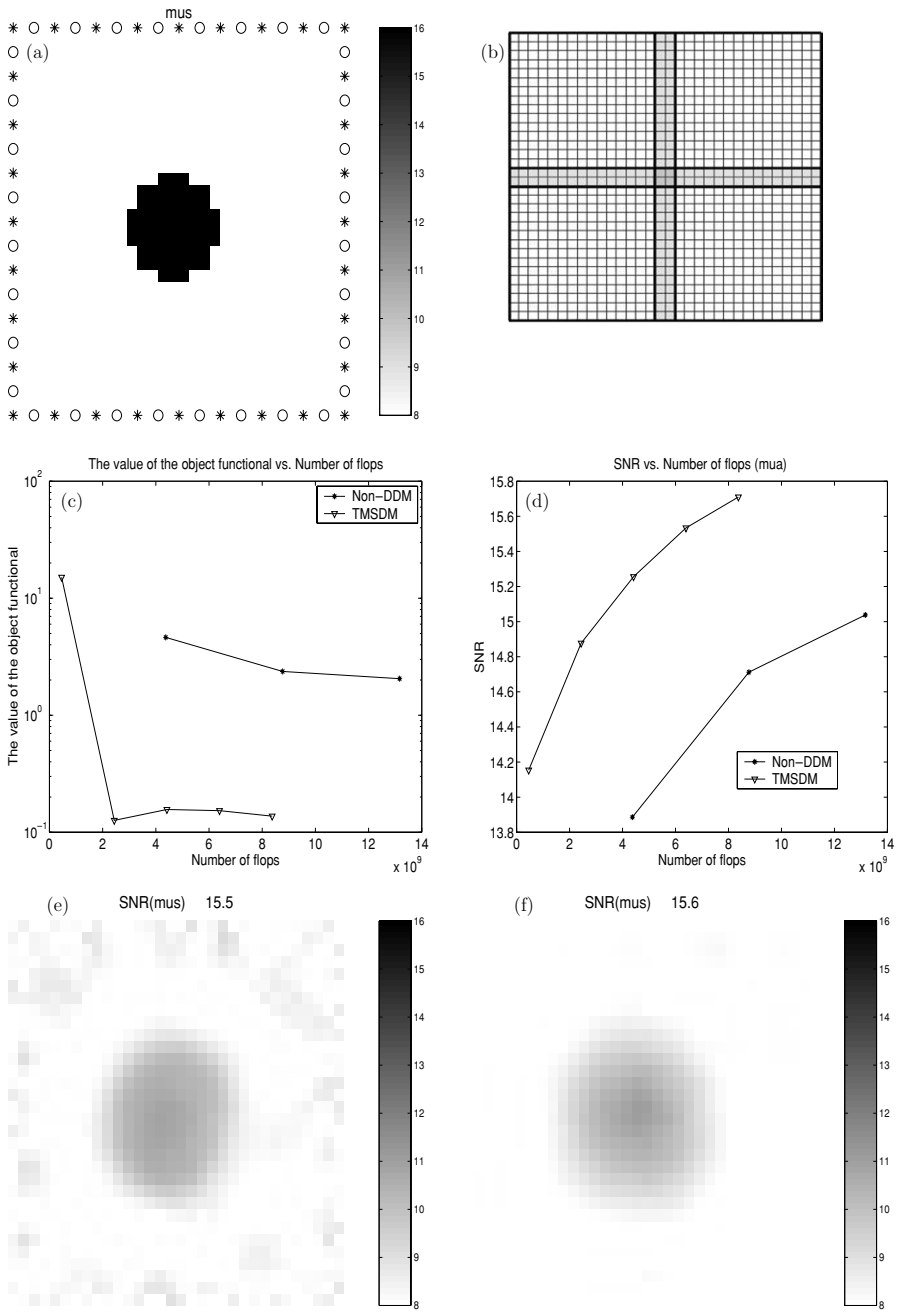

Figure 4. The set-up and the results regarding the μ_a-only reconstruction. (a) The original μ_s' image and the source–detector configuration. The white region represents background tissue with μ_s' value of 8 cm^{-1}, and the black region represents an anomaly with μ_s' value of 16 cm^{-1}. (b) The 2×2 domain decomposition and uniform 32×32 discretization of Ω. (c) The value of the objective functional versus the number of floating point operations of the non-DD method and TMSDM. (d) The SNR versus the number of floating point operations of the non-DD method and TMSDM. (e) The reduced scattering image reconstruction by the non-DD method after 13.165×10^9 floating point operations, with SNR $= 15.5$. (f) The reconstruction by TMSDM after 8.373×10^9 floating point operations with SNR $= 15.6$.

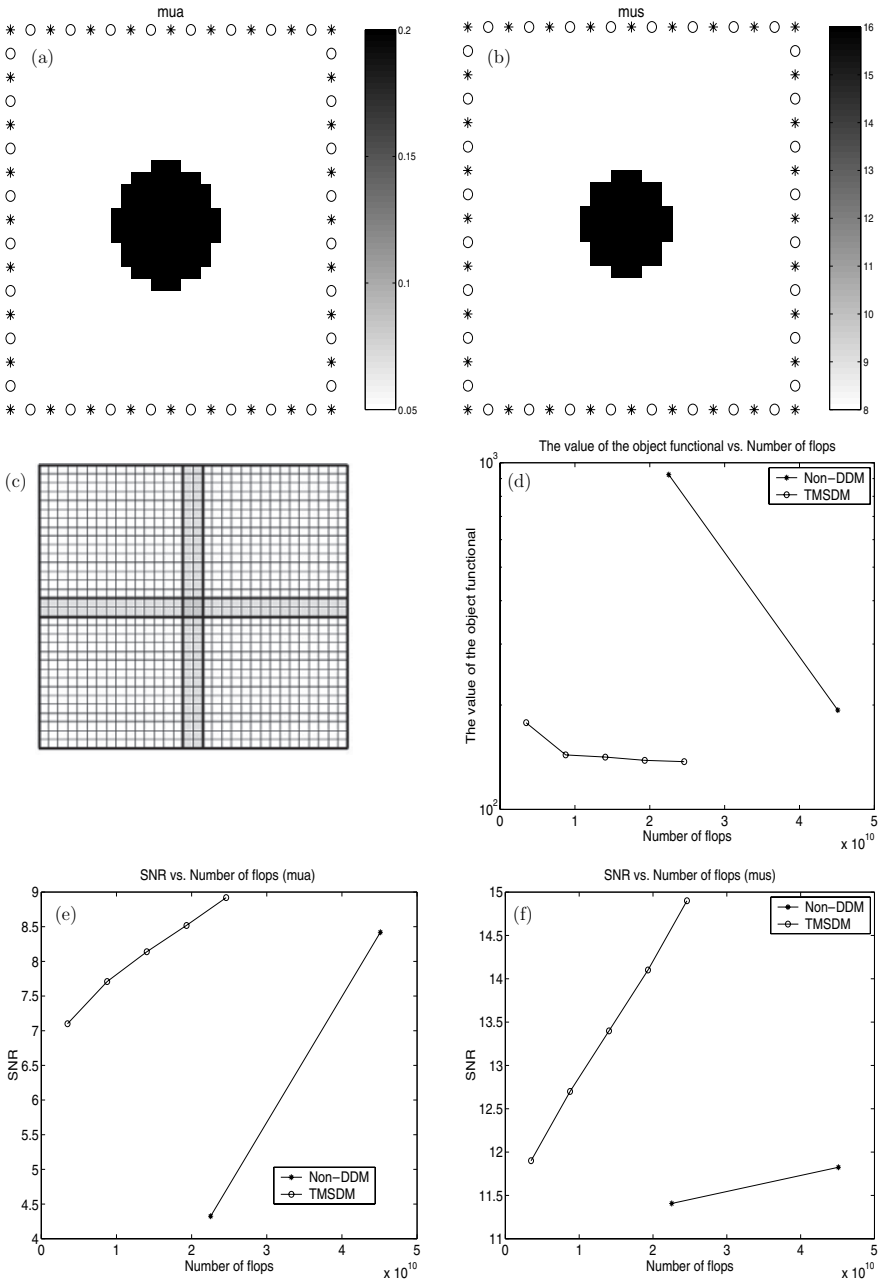

Figure 5. The set-up and the results regarding the simultaneous μ_a and μ_s' reconstruction. (a) The original μ_a' image and the source–detector configuration. The white region represents background tissue with μ_a value of 0.05 cm^{-1}, and the black region represents an anomaly with μ_a value of 0.2 cm^{-1}. (b) The original μ_s' image. The white region represents background tissue with μ_s' value of 8 cm^{-1}, and the black region represents an anomaly with μ_s' value of 16 cm^{-1}. (c) The 2×2 domain decomposition and uniform 32×32 discretization of Ω. (d) The value of the objective functional versus the number of floating point operations of non-DD method and TMSDM. (e) The SNR for the absorption coefficient reconstruction versus the number of floating point operations of non-DD method and TMSDM. (f) The SNR for the reduced scattering coefficient reconstruction versus the number of floating point operations of non-DD method and TMSDM.

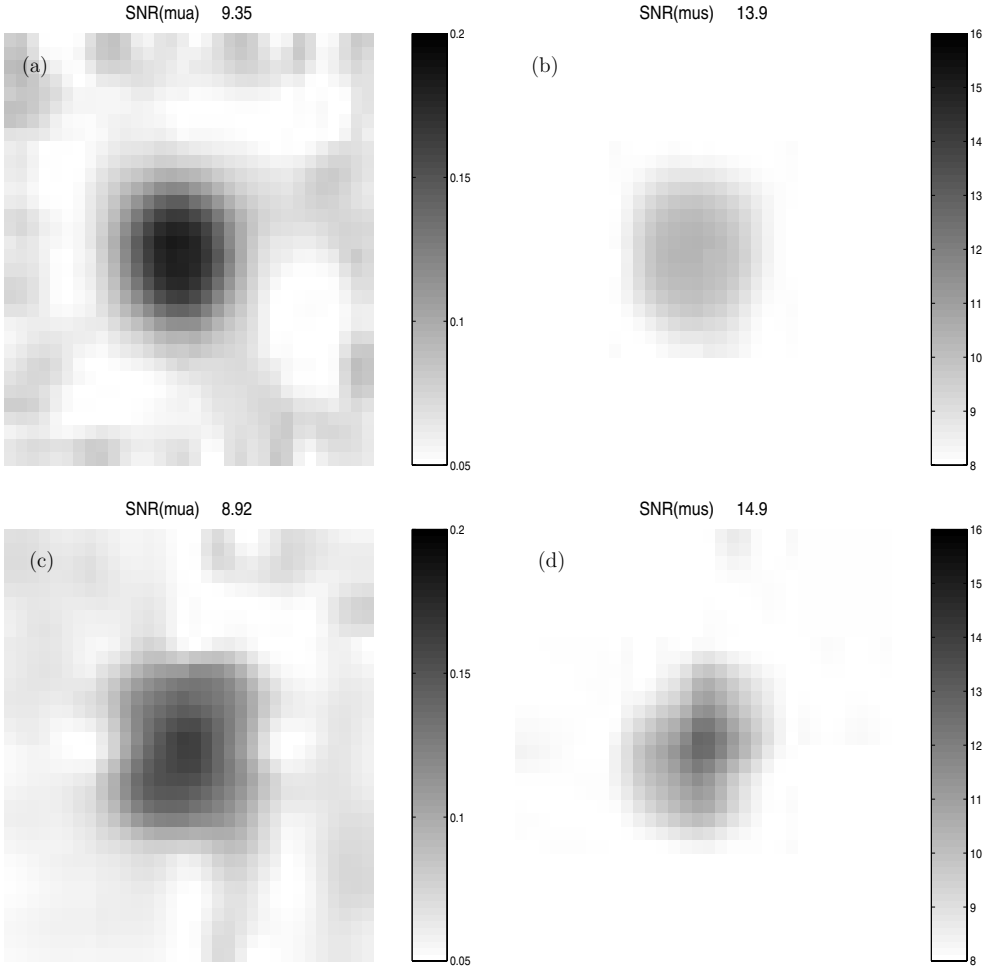

Figure 6. The simultaneous reconstruction results of the absorption and reduced scattering coefficients shown in figures 5(a) and (b). (a), (b) The simultaneous reconstruction of the absorption and reduced scattering coefficients by using the non-DD method after 6.780×10^{10} floating point operations with $\mathrm{SNR}(\mu_a) = 9.346$ and $\mathrm{SNR}(\mu_s') = 13.935$. (c), (d) The simultaneous reconstruction of the absorption and reduced scattering coefficients by using TMSDM after 2.459×10^{10} floating point operations, with $\mathrm{SNR}(\mu_a) = 8.920$ and $\mathrm{SNR}(\mu_s') = 14.899$.

Figures 3(c)–(d) and figures 4(c)–(d) show the value of the objective functional and the SNR versus the number of floating point operations. Figures 3(e), 4(e) and figures 3(f), 4(f) show the reconstructed images using non-DD and TMSDM methods, respectively.

Figures 3(c) and 4(c) show that at a given number of floating point operations, TMSDM achieves a lower objective functional value as compared to the non-DD method. Similarly, figures 3(d) and 4(d) show that at a given number of floating point operations, TMSDM achieves a higher SNR value.

For the μ_a'-only imaging, the reconstructed images shown in figures 3(e)–(f) indicate that TMSDM provides qualitatively and quantitatively better results as compared to the non-DD method. Note that the reconstructed μ_a image using TMSDM was obtained after 8.561×10^9

floating point operations, while the image using the non-DD method was obtained after 12.893×10^9 floating point operations. Similarly figures 4(e)–(f) show the reconstructed μ_s' images obtained after 13.165×10^9 and 8.373×10^9 floating point operations, by using the non-DD method and TMSDM, respectively. Similar to the μ_a-only imaging case, these images indicate that TMSDM outperforms the non-DD method.

4.2.2. Experiment 2. In this experiment, we test the performance of TMSDM in the simultaneous reconstruction of absorption and reduced scattering images and show the convergence of the algorithm experimentally.

Figure 5(a) shows an inclusion with $\mu_a = 0.20$ cm^{-1} in a background with $\mu_a = 0.05$ cm^{-1}. Figure 5(b) shows the reduced scattering coefficient of the same medium, where the circular inclusion corresponding to an object with $\mu_s' = 16$ cm^{-1} is embedded in a background with $\mu_s' = 8$ cm^{-1}. Figure 5(c) shows the domain decomposition with uniform discretization for the forward and inverse problems.

Figure 5(d) shows the value of the objective functional versus the number of floating point operations for both TMSDM and the non-DD method. We see that TMSDM achieves a lower objective functional value at a lower number of floating point operations.

Figures 5(e) and (f) show the SNR value achieved by TMSDM and the non-DD method versus the number of floating point operations, respectively for absorption and reduced scattering images. These images indicate that TMSDM outperforms the non-DD method, especially for the reduced scattering image. The reconstructed images shown in figure 6 are consistent with the performance numbers given in figures 5(d)–(f). We note that the images reconstructed by using TMSDM are obtained in 2.459×10^{10} floating point operations, while the images obtained by using the non-DD method require 6.780×10^{10} floating point operations.

5. Conclusion

In this work, we developed two-level overlapping domain decomposition algorithms to address the computational complexity of the forward and inverse problems associated with DOT imaging. We used the frequency-domain diffusion equation to model NIR light propagation. In order to address the ill-posed nature of the inverse problem, we used a nonlinear least-squares formulation with a general Tikhonov regularization term to recover both the absorption and scattering coefficients.

In the forward problem, we employed an overlapping domain decomposition algorithm with a two-grid preconditioner (TMODDM), and for the nonlinear inverse problem, we used an overlapping space decomposition algorithm with a coarse-level initiation (TMSDM). We proved the local convergence of the TMSDM method under the conditions that lead to the strict local convexity of the objective functional formulated for the inverse problem. For notational brevity, we described both DD algorithms for a 2D bounded optical domain and uniform discretization. Nevertheless, the extension of the algorithm for the 3D case and adaptive discretization is straightforward. This requires replacing the 2D finite elements with 3D finite elements and definition of 3D restriction (downsampling) and prolongation (upsampling) operators for the two-level algorithms.

Note that in our inverse problem solver, measurements from all source–detector pairs are used for each sub-domain. This accounts for the contribution of all sources to the boundary data. Therefore, this scheme does not impose any constraints on how the image domain is decomposed into sub-domains.

We analysed the computational complexity of both algorithms and demonstrated their performance in three different numerical simulations, where we considered simultaneous absorption and reduced scattering coefficient reconstruction as well as absorption only and reduced scattering only reconstructions. Our study shows that TMODDM provides lower relative error than the one-level MODDM for the same floating point operations and the relative error becomes much lower as the width of the overlapping region grows. Similarly, TMSDM provides lower objective functional values and higher SNR than the non-DD method for the same number of floating point operations, in all experiments including the simultaneous absorption and reduced scattering reconstruction. If parallel computers are used, the computational efficiency of TMSDM is expected to be further enhanced.

Finally, the local convergence properties of the algorithms do not change when they are implemented only for a region of interest (ROI). Therefore, if an ROI is identified either by *a priori* information provided by a secondary imaging modality such as magnetic resonance or x-ray; or by *a posteriori* information obtained from the coarse-level solution, the optical coefficients can be updated only in the sub-domains covering the ROI, providing further reduction in computational requirements.

Acknowledgments

This material is based upon work supported by NSF-BES-0353160, ONR-N00014-04-1-0694 and US Army Medical Research-W81XWH-04-1-0559. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Various portions of this research were supported by the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821), and Rensselaer Polytechnic Institute.

Appendix A. The proof of theorem 1

Lemma 1. *Let* $\mu = (\mu_a, \mu_s)$ *and* $\widetilde{\mu} = \mu + \delta\mu$, $\delta\mu = (\delta\mu_a, \delta\mu_s')$. *If* $||\delta\mu||_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \leqslant \frac{\epsilon + \alpha C_\Lambda}{C_H C_J}$ *for some* $\epsilon > 0$, *then* F'' *satisfies*

$$\epsilon \|\delta\mu\|_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \leqslant (\delta\mu)^t F''(\mu)(\delta\mu). \tag{A.1}$$

Proof. Let G and $\widetilde{G}$ be Green's function for the optical coefficients μ and $\widetilde{\mu}$, respectively. Then, we get

$$|b|_{l^2} = \left| \log\left(\frac{1}{2a}\widetilde{G}(i:j)\right) - \log\left(\frac{1}{2a}G(i:j)\right) \right|_{l^2} = \left| \log\left(\frac{|\widetilde{G}(i:j)|}{|G(i:j)|}\right) \right|_{l^2}$$
$$= \log\left(1 + \frac{|(\mathcal{R}\widetilde{G})(i:j)|}{|G(i:j)|}\right) \leqslant \frac{|(\mathcal{R}\widetilde{G})(i:j)|}{|G(i:j)|} \leqslant C_J \|\delta\mu\|_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}},$$

where $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$ and

$$\mathcal{R}_1\psi(r) = \int_\Omega G(r, r')\psi(r')\,\mathrm{d}r'.$$
$$\mathcal{R}_2\psi(r) = \int_\Omega \nabla G(r, r') \cdot \nabla\psi(r')\,\mathrm{d}r'.$$

For the norm boundedness of $\mathcal{R}$, see [21].

Therefore, using (2.13), (3.34), and (A.2), we get

$$\delta\mu^t F''\delta\mu = (J\delta\mu)^t(J\delta\mu) + \delta\mu^t Hb\delta\mu + \alpha\delta\mu^t\Lambda''\delta\mu \tag{A.2}$$

$$\geqslant (J\delta\mu)^t(J\delta\mu) + \alpha\delta\mu^t\Lambda''\delta\mu - C_H\delta\mu^t\delta\mu|b| \tag{A.3}$$

$$\geqslant (\alpha C_\Lambda - C_H C_J|\delta\mu|)\delta\mu^t\delta\mu. \tag{A.4}$$

Thus, we have proved the lemma. □

The inequality (A.1) is called the local strong convexity. With this property and the theorem in [5], we will prove theorem 1.

Proof of theorem 1. The proof of theorem 1 is based on the proof in theorem 3.1 [5]. Let $\mu = (\mu_a, \mu_s')$ and $\delta\mu = (\delta\mu_a, \delta\mu_s')$. Let the restrictions of $V_{N_e}^y$, $y = \mu_a, \mu_s'$ to the non-overlapping and overlapping sub-domain Ω_p be $W_{N_e}^{p,y}$ and $V_{N_e}^{p,y}$ for $p = 1, \ldots, d$, respectively, such that

$$W_{N_e}^{y,p} = \left\{x \in V_{N_e}^y | x = 0 \text{ on } \Omega\backslash\overline{\Omega_p}\right\}, \tag{A.5}$$

$$V_{N_e}^{y,p} = \left\{x \in V_{N_e}^y | x = 0 \text{ on } \Omega\backslash\overline{\Omega_p^w}\right\}, \tag{A.6}$$

where $y = \mu_a, \mu_s'$. Then, $W_{N_e}^{y,p}$, $p = 1, \ldots, d$ are mutually disjoint, $W_{N_e}^{y,p} \subset V_{N_e}^{y,p}$, and

$$V_{N_e}^y = V_{N_e}^{y,1} + \cdots + V_{N_e}^{y,d} = W_{N_e}^{y,1} + \cdots + W_{N_e}^{y,d}.$$

Since we did not consider the coarse-level correction step in this theorem, we will use p and d instead of $p+1$ and $d+1$ in (3.31). Let μ^n be the solution of MSDM after n sweeps of sub-domain correction and $\mu^{n+\frac{p}{d}}$ be the solution obtained by updating μ^n with the solution estimates obtained on the sub-domain $\Omega_1^w, \ldots, \Omega_p^w$ $(p \leqslant d)$. Define

$$z_p^n = \begin{cases} \mu - \mu^{n+\frac{p}{d}} & \text{in} \quad \Omega_p \\ (0,0) & \text{in} \quad \Omega\backslash\overline{\Omega_p} \end{cases}$$

and $e^{n+\frac{p}{d}} = \mu^{n+\frac{p}{d}} - \mu^{n+\frac{p-1}{d}} \in V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p}$. Then, $z_p^n \in W_{N_e}^{\mu_a,p} \times W_{N_e}^{\mu_s',p}$ and we obtain

$$\mu - \mu^n = \sum_{p=1}^d z_p^n, \qquad \|\mu - \mu^n\|_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} = \left(\sum_{p=1}^d \|z_p^n\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}}\right)^{1/2} \tag{A.7}$$

and

$$e^{n+\frac{p}{d}} = \underset{v_p \in V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p}}{\operatorname{argmin}} F\left(\mu^{n+\frac{p-1}{d}} + v_p\right), \tag{A.8}$$

Equation (A.8) implies that

$$\left\langle F'\left(\mu^{n+\frac{p-1}{d}} + e^{n+\frac{p}{d}}\right), v_p - e^{n+\frac{p}{d}}\right\rangle \geqslant 0 \qquad \text{for all } v_p \in V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p} \tag{A.9}$$

where $\langle\cdot,\cdot\rangle$ is the l^2 inner product in the space $V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p}$. Using (A.1) and the Taylor expansion for F, we get

$$F(w) - F(v) \geqslant \langle F'(v), w - v\rangle + \frac{\epsilon}{2}\|w - v\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}}, \tag{A.10}$$

for all $v, w \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}$ such that $\|v - w\|_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \leqslant \frac{\epsilon + \alpha C_\Lambda}{C_F^2}$. Inserting $w = \mu^{n+\frac{p-1}{d}}$ and $v = \mu^{n+\frac{p}{d}}$ in (A.10) and using (A.9) with $v_p = 0$, we get

$$F\left(\mu^{n+\frac{p-1}{d}}\right) - F\left(\mu^{n+\frac{p}{d}}\right) \geqslant \frac{\epsilon}{2}\left\|e^{n+\frac{p}{d}}\right\|^2_{V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p}} \geqslant 0$$

and

$$F(\mu^n) - F(\mu^{n+1}) \geqslant \sum_{p=1}^{d} \left(F\left(\mu^{n+\frac{p-1}{d}}\right) - F\left(\mu^{n+\frac{p}{d}}\right)\right)$$
$$\geqslant \frac{\epsilon}{2} \sum_{p=1}^{d} \left\| e^{n+\frac{p}{d}} \right\|^2_{V_{N_e}^{\mu_a,p} \times V_{N_e}^{\mu_s',p}} \geqslant 0. \tag{A.11}$$

Since $\mu = \operatorname{argmin}_{v \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} F(v)$, we get

$$\langle F'(\mu), v - \mu \rangle \geqslant 0 \text{ for all } v \in V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}, \tag{A.12}$$

as in (A.8). Inserting $w = \mu^n$ and $v = \mu$ into (A.10) and using (A.12), we get

$$F(\mu^n) - F(\mu) = \langle F'(\mu), \mu^n - \mu \rangle + \frac{\epsilon}{2} \|\mu^n - \mu\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}}$$
$$\geqslant \frac{\epsilon}{2} \|\mu^n - \mu\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \geqslant 0. \tag{A.13}$$

Finally, we obtain

$$0 \leqslant F(\mu^{n+1}) - F(\mu) \leqslant \left\langle F'(\mu^{n+1}), \mu^{n+1} - \mu \right\rangle \qquad \text{(using (A.10))}$$
$$= \left\langle F'(\mu^{n+1}), \sum_{p=1}^{d} e^{n+\frac{p}{d}} + \mu^n - \mu \right\rangle = \sum_{p=1}^{d} \left\langle F'\left(\mu^{n+1}, e^{n+\frac{p}{d}}\right) - z_p^n \right\rangle$$
$$\leqslant \sum_{p=1}^{d} \left\langle F'(\mu^{n+1}) - F'\left(\mu^{n+\frac{p}{d}}\right), e^{n+\frac{p}{d}} - z_p^n \right\rangle \qquad \text{(using (A.9))}$$
$$= \sum_{p=1}^{d} \sum_{l=p+1}^{d} \left\langle F'\left(\mu^{n+\frac{l}{d}}\right) - F'\left(\mu^{n+\frac{l-1}{d}}\right), e^{n+\frac{p}{d}} - z_p^n \right\rangle$$
$$= \sum_{p=1}^{d} \sum_{l=p+1}^{d} \left\langle F''(y_l^n) e^{n+\frac{l}{d}}, e^{n+\frac{p}{d}} - z_p^n \right\rangle$$
for some y_l^n which lies between $\mu^{n+\frac{l}{d}}$ and $\mu^{n+\frac{l-1}{d}}$
$$\leqslant \epsilon \left(\sum_{l=1}^{d} \left\| e^{n+\frac{l}{d}} \right\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \right)^{1/2} \left(\sum_{p=1}^{d} \left\| e^{n+\frac{p}{d}} - z_p^n \right\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \right)^{1/2} \qquad \text{(using (A.1))}$$
$$\leqslant \epsilon \left[\sum_{l=1}^{d} \left\| e^{n+\frac{l}{d}} \right\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} + \left(\sum_{p=1}^{d} \left\| e^{n+\frac{p}{d}} \right\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \right)^{1/2} \left(\sum_{p=1}^{d} \|z_p^n\|^2_{V_{N_e}^{\mu_a} \times V_{N_e}^{\mu_s'}} \right)^{1/2} \right]$$
$$\leqslant 2(F(\mu^n) - F(\mu^{n+1})) + 2\sqrt{F(\mu^n) - F(\mu^{n+1})}\sqrt{F(\mu^n) - F(\mu)}$$
(using (A.7), (A.11) and (A.13)).

Let $d_n = F(\mu^n) - F(\mu)$, then the above equation and the Cauchy–Schwarz inequality implies

$$d_{n+1} \leqslant 2(d_n - d_{n+1}) + 2\sqrt{d_n - d_{n+1}}\sqrt{d_n}$$
$$\leqslant \left(2 + \frac{1}{2\eta}\right)(d_n - d_{n+1}) + 2\eta d_n,$$

for all $\eta > 0$. Thus, we obtain

$$\frac{d_{n+1}}{d_n} \leqslant \frac{4\eta^2 + 4\eta + 1}{6\eta + 1}. \tag{A.14}$$

The left-hand side of (A.14) takes a minimum value $\frac{8}{9} < 1$ when $\eta = \frac{1}{6}$. Thus we have the following successive inequalities:

$$d_n \leqslant \frac{8}{9} d_{n-1} \leqslant \left(\frac{8}{9}\right)^2 d_{n-2} \leqslant \cdots \leqslant \left(\frac{8}{9}\right)^n d_0. \tag{A.15}$$

Using the mean-value theorem, (2.13*a*), (3.32), (3.34*a*) and (A.2), we can estimate d_0 as follows:

$$\begin{aligned} d_0 = F(\mu^0) - F(\mu) &\leqslant \sup_{0<\theta<1} |F'(\theta\mu^0 + (1-\theta)\mu)||\mu - \mu_0| \\ &\leqslant C_J^2 |\mu - \mu_0|^2 + \alpha \widetilde{C}_\Lambda |\mu - \mu_0| \leqslant \frac{\alpha^2 C_\Lambda^2}{C_H^2} + \frac{\alpha^2 C_\Lambda \widetilde{C}_\Lambda}{C_H C_J}. \end{aligned} \tag{A.16}$$

Using (A.15) and (A.16), we obtain (3.33). □

References

[1] Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93

[2] Arridge S R 2001 Methods for the inverse problem in optical tomography *Waves and Imaging Through Complex Media* ed P Sebbah and J-M Tualle (Dordrecht: Kluwer) pp 307–29

[3] Arridge S R, Schweiger M and Delpy D T 1992 Iterative reconstruction of near-infrared absorption images *Inverse Problems in Scattering and Imaging* (*Proc. SPIE* **1767**) ed M A Fiddy 372–83

[4] Badea L 1991 On the Schwarz alternation method with more than two subdomains for nonlinear monotone problems *SIAM J. Numer. Anal.* **28** 179–204

[5] Badea L, Tai X-C and Wang J 2003 Convergence rate analysis of a multiplicative Schwarz method for variational inequalities *SIAM J. Numer. Anal.* **41** 1052–73

[6] Badea L and Wang J 1999 An additive Schwarz method for variational inequalities *Math. Comput.* **69** 1341–54

[7] Bazaraa M S, Sherali H D and Shetty C M 1993 *Nonlinear Programming, Theory and Algorithms* 2nd edn (New York: Wiley)

[8] Boag A 2001 A fast multilevel domain decomposition algorithm for radar imaging *IEEE Trans. Antennas Propag.* **49** 666–71

[9] Boag A, Bresler Y and Michielssen E 2000 A multilevel domain decomposition algorithm for fast $O(n^2 \log n)$ reprojection of tomographic images *IEEE Trans. Image Process.* **9** 1573–82

[10] Bramble J, Pasciak J, Wang J and Xu J 1991 Convergence estimates for product iterative methods with applications to domain decomposition *Math. Comput.* **57** 1–21

[11] Briggs W L 1987 *A Multigrid Tutorial* (Philadelphia, PA: SIAM)

[12] Byrd R H, Schnabel R B and Shultz G A 1988 Approximate solution of the trust region problem by minimization over two-dimensional subspaces *Math. Program.* **40** 247–63

[13] McCormick S F 1989 *Multilevel Adaptive Methods for Partial Differential Equations* (Philadelphia, PA: SIAM)

[14] Eppstein M J and Dougherty D E 1998 Optimal 3-d travel time tomography *Geophys.* **63** 1053–61

[15] Eppstein M J, Dougherty D E, Hawrysz D J and Sevick-Muraca E M 2001 Three-dimensional Bayesian optical image reconstruction with domain decomposition *IEEE Trans. Med. Imaging* **20** 147–63

[16] Guven M, Yazici B, Intes X and Chance B 2003 An adaptive multigrid algorithm for region of interest diffuse optical tomography *Int. Conf. Image Processing* (*Proc. IEEE* **2**) pp 823–6

[17] Guven M, Yazici B, Intes X and Chance B 2003 An adaptive V-grid algorithm for diffuse optical tomography *Bioengineering Conf. (Proc. IEEE)* pp 95–96

[18] Isakov V 1990 On uniqueness in the inverse transmission scattering problem *Commun. Partial Diff. Eqns* **15** 1565–87

[19] Kaipio J and Somersalo E 2005 *Statistical and Computational Inverse Problems (Applied Mathematical Sciences* vol 160*)* (New York: Springer)

[20] Kwon K 2004 Identification of anisotropic anomalous region in inverse problems *Inverse Problems* **20** 1117–36

[21] Kwon K and Yazıcı B 2006 Born expansion and Fréchet derivatives in diffuse optical tomography *Inverse Problems* submitted

[22] Kwon K and Sheen D 2002 Anisotropic inverse conductivity and scattering problems *Inverse Problems* **18** 745–56
[23] Lai C-H, Ierotheou C S, Palansuriya C J and Pericleous K A 2001 Domain decomposition methods for welding problems *12th Int. Conf. Domain Decomposition Methods (Proc. IEEE* **43**) pp 411–20
[24] Levenberg K 1944 A method of the solution of certain problems in least squares *Q. Appl. Math.* **2** 164–68
[25] Marquardt D 1979 An algorithm for least-squares estimation of nonlinear parameters *CAS-26: IEEE Trans. Circuits and Systems (Sept., 1979)*
[26] Oh S, Milstein A B, Bouman C A and Webb K J 2005 A general framework for nonlinear multigrid inversion *IEEE Trans. Image Process.* **14** 125–40
[27] O'Leary M A, Boas D A, Chance B and Yodh A G 1995 Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography *Opt. Lett.* **20** 426–8
[28] Palansuriya C J 1999 Domain decomposition based algorithms for inverse problems *PhD Thesis* University of Greenwich
[29] Palansuriya C J, Lai C H and Ierotheou C H 1998 A domain decomposition based algorithm for nonlinear 2D inverse heat conduction problems *Contemp. Math.* **218** 515–22
[30] Pogue B W, Patterson M S, Jiang H and Paulsen K D 1995 Initial assessment of a simple system for frequency domain diffuse optical tomography *Phys. Med. Biol.* **40** 1709–29
[31] Quarteroni A and Valli A 1999 *Domain Decomposition Methods for Partial Differential Equations* (Oxford: Oxford Science Publications)
[32] Schwarz H A 1890 *Gesammelte Mathematische Abhandlungen* vol 2 (Berlin: Springer) pp 133–43
[33] Smith B, Bjørstad P and Gropp W 1996 *Domain Decomposition, Parallel Multilevel Methods for Elliptic Partial Differential Equations* (Cambridge: Cambridge University Press)
[34] Son I-Y, Guven M, Yazıcı B and Intes X 2003 A 2-level domain decomposition algorithm for inverse diffuse optical tomography *ICIP Proc. IEEE* **2** 823–6
[35] Sylvester J and Uhlmann G 1987 A global uniqueness theorem for an inverse boundary value problem *Ann. Math.* **125** 153–69
[36] Tai X-C 2003 Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities *Numer. Math.* **93** 755–86
[37] Tai X-C, Frøen J, Espedal M S and Chan T F 1998 Overlapping domain decomposition and multigrid methods for inverse problems *Contemp. Math.* **218** 523–29
[38] Tai X-C and Xu J 2001 Global and uniform convergence of subspace correction methods for some convex optimization problems *Math. Comput.* **7** 105–24
[39] Trottenberg U, Oosterlee C and Schuller A 2001 *Multigrid* (New York: Academic)
[40] Ye J C, Bouman C A, Millane R P and Webb K J 2001 Nonlinear multigrid algorithms for Bayesian optical diffusion tomography *IEEE Trans. Image Process.* **10** 909–22