Nonlinear inversion for optical tomography

Wolfgang Bangerth* and Amit Joshi†

*Department of Mathematics, Texas A&M University, College Station, TX 77845, USA
†Department of Radiology, Baylor College of Medicine, Houston, TX 77030, USA

Abstract. Optical tomography is a recent addition to the available set of computerized imaging methods available to characterize live or inanimate matter. It uses light in the optical and near-infrared ranges and is thus particularly suited to study samples that exhibit significant variation in their optical properties in this wavelength range. This includes, for example, biological tissues as well as some inhomogenous fluids of industrial interest. We will focus here on the use of the method for biomedical imaging.

One of the advantages of optical tomography over many of the established methods is that it is relatively fast, inexpensive, and does not use ionizing radiation. The latter is particularly important when studying live tissue as the near infrared radiation does not induce ionization. Another advantage is that the variant of optical tomography described here, namely fluorescence enhanced optical tomography, can be made molecularly targeted, i.e. it can be used to determine the three-dimensional distribution of biochemical events of interest, as long as they can be targeted by antibodies or other targeting moeties.

In contrast to most traditional imaging methods such as X-ray tomography, optical tomography does not use a linear mapping from the desired description of an object to the detected signal. Consequently, we can not hope for explicit inversion formulas such as the inverse Radon transform, and have to resort to numerical procedures for imaging. We will explain and demonstrate such a procedure in this contribution.

Keywords: Inverse problems, partial differential equations, numerical inversion, finite element methods, optical tomography.
PACS: 42.30.Wb,81.70.Tx

INTRODUCTION

Fluorescence enhanced optical tomography is a recent addition to the set of biomedical imaging methods that is currently intensely researched. It attempts to reconstruct interior tissue optical properties using light in the red and infrared range. Its main current use is for biological tissues since these are highly scattering but not strongly absorbing in the NIR (700-900nm) window.

For tissues, optical tomography is developed as a tool for imaging up to depths of several centimeters (some 20–40 scattering length scales), which includes in particular important applications to breast and cervix cancer detection and staging, lymph node imaging, as well as imaging of the brains in neonates.

Optical tomography improves on a number of shortcomings of established cancer imaging techniques. In particular, most currently available techniques only image tissue structure variation created by tumors such as calcification of blood vessels (X-rays), density and stiffness differences (ultrasound) or water content (MRI) of tissues. In biomedical imaging, such effects are often associated with tumors. However, since they are not *specific* to the presence of actual tumor cells, imaging these secondary effects

frequently lead to both false positive and false negative assessments. In addition, X-ray imaging uses ionizing radiation and is therefore harmful and potentially cancer-inducing. In contrast, optical tomography is a functional imaging method that (i) does not use harmful radiation, and (ii) can be made specific to the presence of certain cell types on the molecular level, distinguishing proteins and other molecules that are only expressed in certain tissues we are interested in (for example tumor cells, or lymph nodes if the goal is to track the spread of a tumor). Similar specificity may be possible in applications to inanimate matter as well. Since the main focus of the authors' work is on biomedical imaging, we shall in the following only address this application; however, translation to other kinds of objects under investigation should be readily possible with the mathematical framework outlined below.

Originally, optical tomography [1] was envisaged as illuminating the tissue surface with a known laser source. The light will diffuse and be absorbed in the tissue. By observing the light flux exiting the tissue surface, one hopes to recover the spatial structure of absorption and scattering coefficients inside the sample, which in turn is assumed to coincide with anatomical and pathological structures. This rests on the observation that hemoglobin concentration, blood oxygenation levels, and water content affect optical tissue properties, all of which are correlated with the presence of tumors. As a result diffuse optical tomography (DOT) can image the *in vivo* metabolic environment [2, 3, 4, 5].

However, during the 1990s, it was recognized that DOT is a method that is hard to implement because it does not produce a very large signal to noise ratio. This follows from the fact that a relatively small tumor, or one that does not have a particularly high absorption contrast, does not produce much dimming of the light intensity on the surface, in particular in reflection geometry, i.e. where illumination and measurement surfaces coincide. DOT is also not specific: it detects areas of high light absorption, but does not distinguish the reasons for absorption; for example, it can not distinguish between naturally dark tissues as compared to invading dark diseased tissues.

Since then, a number of approaches have been developed that attempt to avoid these drawbacks. One is fluorescence-enhanced optical tomography, in which a fluorescent dye (or "fluorophore") is injected that specifically targets certain tissue types, for example tumor or lymph node cells. The premise is that in living tissues, the dye is naturally washed out from the rest of the tissue and is only left at sites where it specifically attaches. If the dye is excited using light of one wave length, then we will get a fluorescent signal at a different wavelength, typically in the infrared, from areas in which dye is present (i.e. where the tumor or a lymph node is). In other words, if we illuminate the skin with a red laser, light will travel into the tissue, be absorbed by the dye, and will be re-emitted at a different wave length (in the infrared range). This secondary light is then detected again at the skin: here, a bright infrared spot on the surface indicates the presence of a high dye concentration underneath, which is then indicative of the presence of the tissue kind the dye is specific to.

As the signal is now the presence of a different kind of light, not a faint dimming of the incident light intensity, the signal to noise ratio of fluorescence optical tomography is much better than in DOT. It is also much better than, for example, in positron emission tomography (PET) because dyes can be excited millions or billions of times per second, each time emitting an infrared photon. In addition, the specificity of dyes can be used for molecularly targeted imaging, i.e. we can really image the presence

of diseased cells, not only secondary effects. This has been used in recent years for small animal imaging applications [6] as well as for the clinical imaging of large tissue volumes [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20].

The challenge in optical tomography, as compared to more traditional imaging methods such as X-ray tomography or MRI is that the imaging process, i.e. the mapping from the three-dimensional dye concentration to the two-dimensional images in the infrared range that we can record at the tissue surface is not linear. This means that we can not hope for the existence of an inversion formula, such as the inverse Radon transform, that would recover the dye concentrations from surface images explicitly or through solving a single linear integral equation. Rather, we will have to develop numerical iterative methods that can solve this problem.

Herein, we will therefore review the numerical techniques necessary to solve nonlinear inverse problems on adaptively refined meshes, using a realistic optical tomography testcase. The general approach to solving the problem is similar as used in work by other researchers [21, 22, 23]. However, we will present adaptivity as a central component to our strategy. Our exposition will follow the framework laid out in [24, 25] and applied in [26, 27]. The rest of the paper is structured as follows: we will first present the mathematical formulation of the optical tomography problem, then discuss the numerical algorithm used to solve it, and then show an example of numerical reconstructions using synthetic data before presenting conclusions and an outlook.

OPTICAL TOMOGRAPHY PROBLEM FORMULATION

Nonlinear inverse problems such as fluorescence optical tomography are typically performed in a model-based framework, wherein a partial differential equation (here the photon transport model) is used to predict measurements for a given set of absorption or scattering parameters in the tissue interior. This distributed map of the absorption owing to fluorophore is then iteratively updated until the predicted measurements match the experimentally observed ones. This iterative process will be driven by a Newton-type method that attempts to minimize the difference between prediction and actual measurements.

Let us start with the model, a set of partial differential equations: For time-periodic sources modulated at a frequency ω, the following set of coupled diffusion equations accurately describes the complex-valued photon fluences $u = u(\mathbf{r})$ at the excitation wavelength and $v = v(\mathbf{r})$ at the fluorescent wavelength (u, v describe amplitude and phase-shift relative to the source of the photon waves at all points $\mathbf{r}$):

$$-\nabla \cdot [D_x(\mathbf{r})\nabla u(\mathbf{r})] + k_x u(\mathbf{r}) = 0, \tag{1}$$

$$-\nabla \cdot [D_m(\mathbf{r})\nabla v(\mathbf{r})] + k_m v(\mathbf{r}) = \beta_{xm} u(\mathbf{r}), \tag{2}$$

where

$$D_{x,m} = \frac{1}{3(\mu_{ax,mi} + \mu_{ax,mf} + \mu'_{sx,m})},$$

$$k_{x,m} = \frac{i\omega}{c} + \mu_{ax,mi}(\mathbf{r}) + \mu_{ax,mf}(\mathbf{r}), \qquad \beta_{xm} = \frac{\phi\,\mu_{axf}}{1 - i\omega\tau(\mathbf{r})}.$$

Subscripts x and m denote material properties at excitation and emission wavelengths, respectively: $D_{x,m}$ are the photon diffusion coefficients; $\mu_{ax,mi}$ the absorption coefficients due to endogenous chromophores; $\mu_{ax,mf}$ the absorption coefficients due to exogenous fluorophore; $\mu'_{sx,m}$ the reduced scattering coefficients; ϕ the quantum efficiency of the fluorophore; and finally, τ is the fluorophore lifetime associated with first order fluorescence decay kinetics. These equations are solved with Robin-type boundary conditions on the boundary $\partial\Omega$ of the domain Ω:

$$2D_x\frac{\partial u}{\partial n}+\gamma u+S(\mathbf{r})=0, \qquad 2D_m\frac{\partial v}{\partial n}+\gamma v=0, \tag{3}$$

where n denotes the outward normal to the surface and γ is a constant depending on the optical reflective index mismatch at the boundary. The complex-valued function $S(\mathbf{r})$ is the excitation boundary source.

The goal of fluorescence tomography is to reconstruct the spatially variable coefficients $\mu_{axf}(\mathbf{r})$ and/or $\tau(\mathbf{r})$ from measurements of the emission fluence v on the boundary. Here, we will focus on the recovery of only $\mu_{axf}(\mathbf{r})$. For notational brevity, we set $\mu_{axf}=q$ in the following paragraphs. The remaining optical properties were assumed known, with values corresponding to a 2% Liposyn solution [28].

The model laid out above is able to predict measurements $v(\mathbf{r})$ at the measurements part Γ of the boundary whenever the source term $S(\mathbf{r})$ is specified. In our experiments, we use a shaped laser beam that scans across M positions on the tissue surface, representing sources $S^i(\mathbf{r}), i=1,2,\ldots,M$. We can then predict $v^i(\mathbf{r})$ given these sources. In the experiment, we take fluorescence measurements z^i at the measurement surface Γ for each source position i.

A mathematical description of the imaging problem is then: find that coefficient $q(\mathbf{r})$ for which the predicted values $v^i|_\Gamma$ are closest in some sense to $z^i|_\Gamma$. This problem can be posed as a constrained optimization problem wherein an L_2 norm-based error functional of the distance between boundary fluorescence measurements $\mathbf{z}=\{z^i, i=1,2,\ldots,M\}$ and predictions $\mathbf{v}=\{v^i, i=1,2,\ldots,M\}$ is minimized by variation of the parameter q. The diffusion model above connecting q and v^i is used as an explicit constraint. In a function space setting this minimization problem reads as:

$$\begin{aligned} &\min_{q,\mathbf{u},\mathbf{v}} && J(q,\mathbf{v}) \\ &\text{subject to} && A^i(q;[u^i,v^i])([\zeta^i,\xi^i])=0,\ i=1,\ldots,M. \end{aligned} \tag{4}$$

Here, the error functional $J(q,\mathbf{v})$ incorporates a least square error term over Γ and a Tikhonov regularization term:

$$J(q,\mathbf{v})=\sum_{i=1}^{M}\frac{1}{2}\left\|v^i-\sigma z^i\right\|_\Gamma^2+\beta r(q). \tag{5}$$

σ is a factor that describes the relationship between actual fluorescent fluence at the tissue surface and the CCD camera signal. The constraint $A^i(q;[u^i,v^i])([\zeta^i,\xi^i])=0$ is the weak or variational form of the coupled photon diffusion equations in frequency

domain with partial current boundary conditions for the i^{th} excitation source, and with test functions $[\zeta, \xi] \in H^1(\Omega)$:

$$A^i(q;[u^i,v^i])([\zeta^i,\xi^i]) = (D_x\nabla u^i, \nabla\zeta^i)_\Omega + (k_x u^i, \zeta^i)_\Omega + \frac{\gamma}{2}(u^i,\zeta^i)_{\partial\Omega} + \frac{1}{2}(S^i,\zeta^i)_{\partial\Omega} + (D_m\nabla v^i, \nabla\xi^i)_\Omega + (k_m v^i, \xi^i)_\Omega + \frac{\gamma}{2}(v^i,\xi^i)_{\partial\Omega} - (\beta_{xm}u^i,\xi^i)_\Omega.$$

To solve this problem, we use that the solution of the constrained minimization problem (4) is a stationary point of the Lagrangian [29]

$$L(x) = J(q,\mathbf{v}) + \sum_{i=1}^{M} A^i(q;[u^i,v^i])([\lambda_i^{ex},\lambda_i^{em}]). \tag{6}$$

Here, $\lambda_i^{ex}, \lambda_i^{em}$ are the Lagrange multipliers corresponding to the excitation and emission diffusion equation constraints for the i^{th} source, respectively, and we use the abbreviation $x = \{\mathbf{u}, \mathbf{v}, \lambda^{ex}, \lambda^{em}, q\}$ for brevity. The stationary point is found using the Gauss-Newton method which computes an update direction $\delta x_k = \{\delta\mathbf{u}_k, \delta\mathbf{v}_k, \delta\lambda_k^{ex}, \delta\lambda_k^{em}, \delta q_k\}$ by solving the linear system

$$L_{xx}(x_k)(\delta x_k, y) = -L_x(x_k)(y) \qquad \forall y, \tag{7}$$

where $L_{xx}(x_k)$ is the Gauss-Newton approximation to the Hessian matrix of second derivatives of L at point x_k, and y are test functions.

NUMERICAL ALGORITHM

Equations (7) are partial differential equations of their own, defining the update $\delta x_k(\mathbf{r})$ whenever the previous iterates $x_k(\mathbf{r})$ are given. While linear, these are a complicated set of coupled equations with non-constant coefficients (through the dependence on x_k) for which we can not expect to find analytical solutions.

To make a practical scheme out of this, we will therefore have to discretize the equations. In our work, we use the finite element method. State and adjoint variables u^i, v^i, λ_i^{ex}, and λ_i^{em} for each excitation source are discretized and solved for on individual meshes with continuous finite elements, while the unknown parameter map q is discretized on a separate mesh with discontinuous finite elements. Hence for M sources, $M+1$ finite element meshes are employed. After this discretization step, we obtain a large but sparse linear system of equations of the following structure:

$$\begin{pmatrix} \mathbf{M} & \mathbf{A}^T & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{C} \\ \mathbf{0} & \mathbf{C}^T & \beta\mathbf{R} \end{pmatrix} \begin{pmatrix} \delta u_{k,h} \\ \delta\lambda_{k,h} \\ \delta q_{k,h} \end{pmatrix} = \begin{pmatrix} F_u \\ F_\lambda \\ F_q \end{pmatrix}, \tag{8}$$

where $\delta u_{k,h}, \delta\lambda_{k,h}, \delta q_{k,h}$ are the vectors of nodal values of updates for $[u,v], [\lambda^{ex},\lambda^{em}], q$, respectively. The size of this linear system equals the number of all M solution vectors

and Lagrange multipliers plus the number of unknowns in the parameter. It is therefore very large already on modestly refined three-dimensional meshes and can be on the order of several ten millions. To make things worse, the linear system is indefinite and usually very badly conditions, with condition numbers often exceeding 10^{12}. The solution of this linear system is therefore not possible with direct solvers, and not straightforward with the usual set of iterative linear solvers such as Conjugate Gradients or GMRES [30].

In order to compute the Newton updates defined by (8) we therefore re-formulate this set of linear equations using block elimination. We then arrive at the following sequence of equations that need to be solved and whose solution is equivalent to the one above:

$$\mathbf{S}\ \ \delta q_{k,h} = \mathbf{F}_q - \sum_{i=1}^{N} \mathbf{C}^{i^T} \mathbf{A}^{i^{-T}} (\mathbf{F}_u^i - \mathbf{M}^i \mathbf{A}^{i^{-1}} \mathbf{F}_\lambda^i), \tag{9}$$

$$\mathbf{A}^i\ \delta u_{k,h}^i = \mathbf{F}_\lambda^i - \mathbf{C}^i \delta q_{k,h}, \tag{10}$$

$$\mathbf{A}^{i^T}\ \delta \lambda_{k,h}^i = \mathbf{F}_u^i - \mathbf{M}^i \delta q_{k,h}, \tag{11}$$

where $\mathbf{S}$ denotes the Schur complement

$$\mathbf{S} = \beta \mathbf{R} + \mathbf{C}^T \mathbf{A}^{-T} \mathbf{M} \mathbf{A}^{-1} \mathbf{C}. \tag{12}$$

This sequence of linear systems is much simpler to solve (i) since the first equation that defines the update $\delta q_{k,h}$ requires only the inversion of a symmetric and positive matrix $\mathbf{S}$ that is furthermore relatively small, involving only the number of unknowns in the parameter $q_{k,h}$; (ii) because the rest of the matrices to be inverted are either $\mathbf{A}^i$ or $(\mathbf{A}^i)^T$, i.e. matrices that stem from the discretization of the ith forward model that computes the fluxes u^i, v^i from a source term, or from the adjoint operator. While these matrices can still be large (with sizes in the 100,000s to a few millions), devising forward and adjoint solvers for partial differential equations is a well-understood process for which good solvers and preconditioners are readily available. The overall solution process has therefore been reduced to applying a CG method to $\mathbf{S}$ and to applying traditional solvers for the forward and adjoint operators. While the resulting process is not trivial, it is also not too complicated to implement on a parallel machine and a Newton iteration can be performed in a few minutes even on relatively fine meshes with many unknowns.

After solving a few Newton iterations as outlined above, our algorithm examines the solution and refines the $M+1$ meshes we use to discretize the M forward and adjoint variables and the one parameter. We then adaptively refine each of the meshes to obtain new grids that are better adjusted to where the solution is not smooth. The pictures shown in the next section will demonstrate this.

AN IMAGE RECONSTRUCTION EXAMPLE

In the following, let us illustrate our imaging scheme using a synthetic example. Examples using experimental data can be found in [31]. The computations shown here were implemented using the Open Source finite element library deal.II [32]. More details on the numerical methods can be found in References [25, 24].

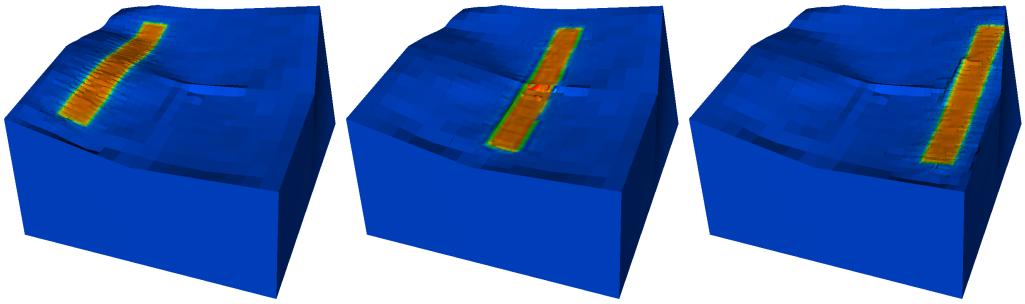

FIGURE 1. *Solutions for measurements $i = 1,4,7$ for a laser line scanning over the top of the tissue sample.*

The example we show uses a realistic geometry obtained stereographically using a pair of cameras trained to the groin region of a pig (the experiment was performed to image the lymph nodes in this area). Fig. 1 shows a simulated laser source at three positions while scanning over the surface of this region (a total of $M = 8$ source positions is used in this example). At each of these positions, we take measurements z^i of the fluorescent light intensity and relative phase emitted by the dye that has been injected at that is predominantly attached to the tissue kind under consideration here, i.e. lymphatic cells. For the purpose of illustration, measurements z^i are computed numerically (rather than obtained experimentally) using a separate computer program that employs a different numerical method to avoid an inverse crime.

Using these measurements, we then employ the inversion algorithm outlined in the previous sections. As mentioned, we refine each of the involved meshes every few Newton iterations. For the first of the source positions shown in the Fig. 1, we show a sequence of meshes used after zero, two, and four adaptive refinements. It is easy to see that the mesh becomes gradually finer around those locations where a high mesh resolution is necessary to resolve features of the solution, whereas it remains coarse where resolution is not needed. Using this adaptive scheme allows us to focus on those areas where it is necessary while keeping the overall number of degrees of freedom, and consequently the numerical effort involved in all steps, as small as possible. In cases like the ones shown here, the use of adaptivity reduces the size of the involved problems by factors of 10–100, and is consequently indispensible to make the solution of such inverse problems feasible.

Similarly, the unknown parameter $q(\mathbf{r})$ is discretized on a sequence of meshes that are also adapted successively. Fig. 3 shows these meshes at the same iterations as Fig. 2 for the state and adjoint variables. As can be seen, the mesh is refined towards an object at the center of the domain to provide high resolution there. Fig. 4 illustrates that the reason for this refinement pattern is that the reconstructed parameter $q(\mathbf{r})$ has a high dye concentration at the center of the domain, indicative of a lymph node or tumor. The refinement pattern is clearly appropriate for this purpose.

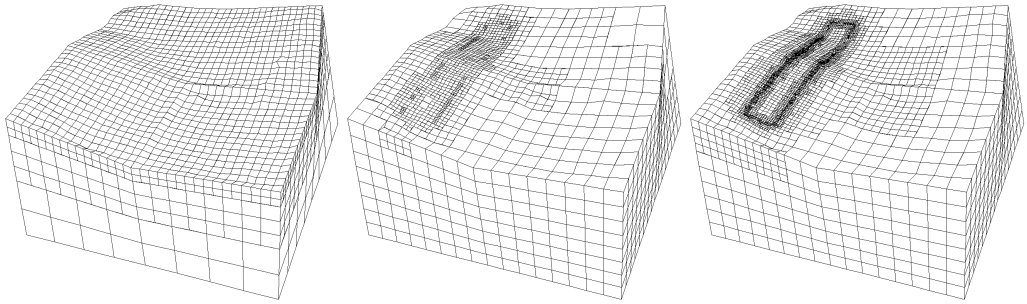

FIGURE 2. *Meshes for experiment $i = 1$ after zero, two, and four refinement cycles. Note that the mesh density is localized around where it is necessary to resolve the structure of the solution.*

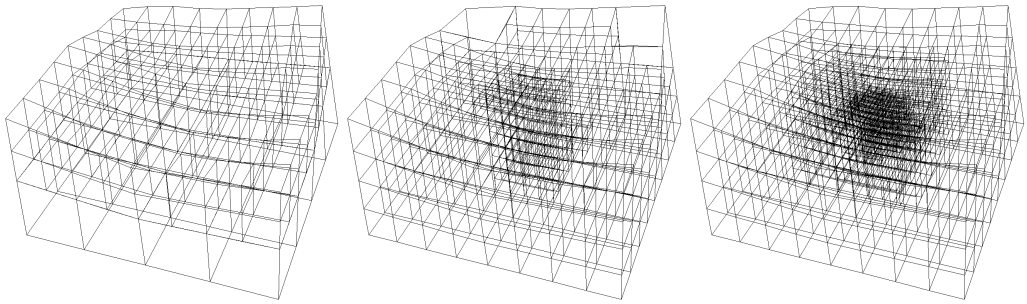

FIGURE 3. *Meshes on which the parameter q is discretized after zero, two, and four refinement cycles.*

CONCLUSIONS AND OUTLOOK

In this paper, we have given a brief overview of the various techniques necessary and available for the solution of nonlinear inverse problems, illustrated using a recent biomedical imaging technique: fluorescence enhanced optical tomography. For this and similar cases, uniformly refined meshes can not deliver the necessary resolution within compute times that are clinically acceptable because they lead to nonlinear optimization problems that are orders of magnitude too large for today's hardware. Our approach to this problem is to introduce adaptively refined meshes for solving the forward/adjoint problems and the unknown parameter updates. They are able to focus numerical effort to regions in the domain where high resolution is actually necessary. Other advantages of such schemes are that they also regularize the inverse problem and in particular make the initial Gauss-Newton iterations extremely cheap since we can compute on coarse meshes while we are still far away from the solution.

Using multiple excitation source patterns is essential for acquiring high information content measurements for nonlinear inverse problems (as it is necessary for high resolution in linear inverse problems such as X-ray tomography). We have integrated the treatment of multiple experiments into our approach by using a mathematical formulation that considers all measurements at once. Our algorithmic implementation exploits the availability of multiprocessor computers and Beowulf clusters to rapidly solve the resulting problem. Finally, we have demonstrated the application of the developed algorithms and methods to a problem of clinical importance, namely locating lymph nodes

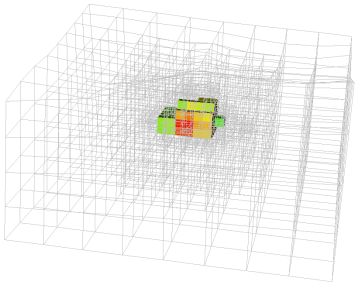

FIGURE 4. *Identified parameter $q(\mathbf{r})$ after 25 Newton iterations. The cells shown are those where the reconstructed dye concentration is more than 50% of the maximum identified value.*

marked with fluorescent contrast agents.

Although the results shown here and elsewhere [25, 24, 31] demonstrate that we are able to efficiently solve inverse fluorescence tomography problems with practically sufficient resolution, further progress is necessary in several areas to improve the numerical performance. This includes improving linear and nonlinear solvers, regularization, and stabilization by imposing additional constraints on the solution. For practical applicability, numerical methods also have to work in the presence of significant background heterogeneity, unknown or large noise levels, systematic measurement bias, and other practical constraints. Systematic testing of reconstructions for statistically sampled scenarios with Objective Assessment of Image Quality (OAIQ) methods is therefore necessary to achieve clinical recognition for fluorescence optical tomography.

Looking beyond the application use in this paper, adaptive finite elements and many of the other techniques demonstrated here can also be used for a wide variety of other nonlinear inverse problems. This includes, among many others, electrical impedance tomography (EIT), eddy current imaging, diffuse optical tomography, and magnetoresistivity, all of which are used in areas outside biomedical imaging for the characterization of materials, nondestructive testing, or in geophysical applications. In each of these cases, the model is a partial differential equation in which the predicted oberservable quantity depends nonlinearly on the parameter that we would like to identify. In all these cases, an iterative nonlinear algorithm and a discretized version of the partial differential equation is necessary, and the methods shown herein are immediately applicable to achieve this goal.

ACKNOWLEDGMENTS

Part of this work was supported by NSF grant DMS-0604778 and NIH grant R01 CA 67176.

REFERENCES

1. S. R. Arridge, *Inverse Problems* **15**, R41–R93 (1999).
2. B. Chance, J. S. Leigh, H. Miyake, D. S. Smith, S. Nioka, R. Greenfeld, M. Finander, K. Kauffmann, W. Levy, M. Young, P. Cohen, H. Yoshioka, and R. Borestsky, *Proc. Nat. Acad. Sciences* **85**, 4791–4975 (1988).
3. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, *Proc. Nat. Acad. Sciences* **100**, 12349–12354 (2003).
4. J. C. Hebden, A. Gibson, R. M. Yusof, N. Everdell, E. M. Hillman, D. T. Delpy, S. R. Arridge, T. Austin, J. H. Meek, and J. S. Wyatt, *Phys. Med. Biol.* **47**, 4155–4166 (2002).
5. J. C. Hebden, A. Gibson, T. Austin, R. M. Yusof, N. Everdell, D. T. Delpy, S. R. Arridge, J. H. Meek, and J. S. Wyatt, *Phys. Med. Biol.* **49**, 1117–1130 (2004).
6. E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, *Med. Phys.* **30**, 901–911 (2003).
7. E. M. Sevick-Muraca, and C. L. Burch, *Optics Letters* **19**, 1928–1930 (1994).
8. E. M. Sevick-Muraca, G. Lopez, T. L. Troy, J. S. Reynolds, and C. L. Hutchinson, *Photochemistry and Photobiology* **66**, 55–64 (1997).
9. X. D. Li, M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, *Applied Optics* **35**, 3746–3758 (1996).
10. X. D. Li, B. Chance, and A. G. Yodh, *Applied Optics* **37**, 6833–6844 (1998).
11. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, *J. Luminescence* **60**, 281–286 (1994).
12. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, *Opt. Lett.* **20**, 426–428 (1996).
13. J. C. Schotland, *J. Opt. Soc. Am. A* **14**, 275–279 (1997).
14. V. Chernomordik, D. Hattery, I. Gannot, and A. H. Gandjbakhche, *IEEE J. Sel. Top. Quantum Electron.* **54**, 930–935 (1999).
15. J. Wu, Y. Wang, L. Perleman, I. Itzkan, R. R. Desai, and M. S. Feld, *Opt. Lett.* **20**, 489–491 (1995).
16. E. L. Hull, M. G. Nichols, and T. H. Foster, *Appl. Opt.* **37**, 2755–2765 (1998).
17. A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. Boas, and R. P. Milane, *Appl. Opt.* **42**, 3061–3094 (2003).
18. M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. Sevick-Muraca, *Proc. Nat. Acad. Sci.* **99**, 9619–9624 (2002).
19. R. Roy, A. B. Thompson, A. Godavarty, and E. M. Sevick-Muraca, *IEEE Transactions on Medical Imaging* **24**, 137–154 (2005).
20. A. Godavarty, M. J. Eppstein, C. Zhang, S. Theru, A. B. Thompson, M. Gurfinkel, and E. M. Sevick-Muraca, *Phys. Med. Biol.* **48**, 1701–1720 (2003).
21. G. Biros, and O. Ghattas, *SIAM J. Sci. Comput.* **27**, 687–713 (2005).
22. G. Biros, and O. Ghattas, *SIAM J. Sci. Comput.* **27**, 714–739 (2005).
23. G. S. Abdoulaev, K. Ren, and A. H. Hielscher, *Inverse Problems* **21**, 1507–1530 (2005).
24. W. Bangerth, and A. Joshi, *Inverse Problems* **24**, 034011/1–22 (2008).
25. W. Bangerth, *SIAM J. Sc. Comput.* (2008, accepted).
26. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, *Optics Express* **14**, 6516–6534 (2006).
27. A. Joshi, W. Bangerth, K. Hwang, J. C. Rasmussen, and E. M. Sevick-Muraca, *Med. Phys.* **33**, 1299–1310 (2006).
28. A. Joshi, W. Bangerth, K. Hwang, J. Rasmussen, and E. M. Sevick-Muraca, *Med. Phys.* **33**, 1299–1310 (2006).
29. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, *Optics Express* **12**, 5402–5417 (2004).
30. R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, E. V, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, 1993.
31. A. Joshi, W. Bangerth, and E. Sevick, "Non-Contact Fluorescence Optical Tomography with Adaptive Finite Element Methods," in *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, edited by Y. Censor, M. Jiang, and A. K. Louis, Birkhäuser, 2008.
32. W. Bangerth, R. Hartmann, and G. Kanschat, *ACM Trans. Math. Softw.* **33**, 24/1–24/27 (2007).