

Adaptive finite element based tomography for fluorescence optical imaging in tissue

Amit Joshi,¹ Wolfgang Bangerth,² and Eva M. Sevick-Muraca¹

¹Photon Migration Laboratories, Texas A & M University
College Station, TX 77843-3255

²Center for Subsurface Modeling, Institute for Computational Engineering and Sciences,
and Institute for Geophysics, University of Texas at Austin, Austin, TX 78712

bangerth@ices.utexas.edu, eva-m-sevick@tamu.edu

Abstract: A three-dimensional fluorescence-enhanced optical tomography scheme based upon an adaptive finite element formulation is developed and employed to reconstruct fluorescent targets in turbid media from frequency-domain measurements made in reflectance geometry using area excitation illumination. The algorithm is derived within a Lagrangian framework by treating the photon diffusion model as a constraint to the optimization problem. Adaptively refined meshes are used to separately discretize maps of the forward/adjoint variables and the unknown parameter of fluorescent yield. A truncated Gauss-Newton method with simple bounds is used as the optimization method. Fluorescence yield reconstructions from simulated measurement data with added Gaussian noise are demonstrated for one and two fluorescent targets embedded within a 512ml cubical tissue phantom. We determine the achievable resolution for the area-illumination/area-detection reflectance measurement geometry by reconstructing two 0.4cm diameter spherical targets placed at a series of decreasing lateral spacings. The results show that adaptive techniques enable the computationally efficient and stable solution of the inverse imaging problem while providing the resolution necessary for imaging the signals from molecularly targeting agents.

© 2004 Optical Society of America

OCIS codes: (170.3010) Image Reconstruction Techniques; (170.5280) Photon Migration; (170.3880) Medical and Biological Imaging.

References and links

1. M. G. Pomper, "Molecular Imaging: An Overview," *Acad. Radiol.* **8**, 1141–1153 (2001).
2. M. A. O'Leary, D. A. Boas, B. Chance, and A. Yodh, "Reradiation and imaging of diffuse photon density waves using fluorescent inhomogeneities," *J. Luminescence* **60**, 281–286 (1994).
3. J. Wu, Y. Wang, L. Perleman, I. Itzkan, R. R. Desai, and M. S. Feld, "Time resolved multichannel imaging of fluorescent objects embedded in turbid media," *Opt. Lett.* **20**, 489–491 (1995).
4. E. L. Hull, M. G. Nichols, and T. H. Foster, "Localization of Luminescent Inhomogeneities in Turbid Media with Spatially Resolved Measurements of CW Diffuse Luminescence Emittance," *Appl. Opt.* **37**, 2755–2765 (1998).
5. J. C. Schotland, "Continuous wave diffusion imaging," *J. Opt. Soc. Am. A* **14**(275–279) (1997).
6. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.* **20**, 426–428 (1996).
7. E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution fluorescence molecular imaging system for small animal imaging," *Med. Phys.* **30**, 901–911 (2003).

8. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," *Opt. Lett.* **26**(12), 893–895 (2001).
9. V. Chernomordik, D. Hattery, I. Gannot, and A. H. Gandjbakhche, "Inverse method 3-D reconstruction of localized in vivo fluorescence-application to Sjögren syndrome," *IEEE J. Sel. Top. Quantum Electron.* **54**, 930–935 (1999).
10. H. Quan and Z. Guo, "Fast 3-D Optical Imaging With Transient Fluorescence Signals," *Opt. Express* **12**(3), 449–457 (2004). <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-12-3-449>.
11. D. Y. Paithankar, A. U. Chen, B. W. Pogue, M. S. Patterson, and E. M. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," *Appl. Opt.* **36**(10), 2260–2272 (1997).
12. R. Roy and E. M. Sevick-Muraca, "Truncated Newton's Optimization Schemes for Absorption and Fluorescence Optical Tomography: part(1) theory and formulation," *Opt. Express* **4**(10), 353–371 (1999). <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-4-10-353>.
13. H. Jiang, "Frequency-domain fluorescent diffusion tomography: a finite element based algorithm and simulations," *Appl. Opt.* **37**(22), 5337–5343 (1998).
14. M. J. Eppstein, D. E. Dougherty, T. L. Troy, and E. M. Sevick-Muraca, "Biomedical optical tomography using dynamic parametrization and Bayesian conditioning on photon migration measurements," *Appl. Opt.* **38**(10), 2138–2150 (1998).
15. M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. Sevick-Muraca, "Three dimensional near infrared fluorescence tomography with Bayesian methodologies for image reconstruction from sparse and noisy data sets," *Proc. Nat. Acad. Sci.* **99**, 9619–9624 (2002).
16. A. Godavarty, M. J. Eppstein, C. Zhang, S. Theru, A. B. Thompson, M. Gurfinkel, and E. M. Sevick-Muraca, "Fluorescence-enhanced optical imaging in large tissue volumes using a gain-modulated ICCD camera," *Phys. Med. Biol.* **48**, 1701–1720 (2003).
17. A. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. Boas, and R. P. Milane, "Fluorescence Optical Diffusion Tomography," *Appl. Opt.* **42**(16), 3061–3094 (2003).
18. J. P. Houston, S. Ke, W. Wang, C. Li, and E. M. Sevick-Muraca, "Optical and nuclear image quality analysis with in vivo NIR fluorescence and conventional gamma images acquired using a dual labeled tumor targeting probe," *J. Biomed. Opt.* (submitted) (2004).
19. W. Bangerth, "Adaptive Finite Element Methods for the Identification of Distributed Coefficients in Partial Differential Equations," Ph.D. thesis, University of Heidelberg (2002).
20. W. Bangerth, "A framework for the adaptive finite element solution of large inverse problems. I. Basic techniques," Tech. Rep. 04-39, ICES, University of Texas at Austin (2004).
21. E. M. Sevick-Muraca, E. Kuwana, A. Godavarty, J. P. Houston, A. B. Thompson, and R. Roy, *Near Infrared Fluorescence Imaging and Spectroscopy in Random Media and Tissues*, chap. 33, Biomedical Photonics Handbook (CRC Press, 2003).
22. A. Godavarty, D. J. Hawrysz, R. Roy, E. M. Sevick-Muraca, and M. J. Eppstein, "Influence of the refractive index-mismatch at the boundaries measured in fluorescence-enhanced frequency-domain photon migration imaging," *Opt. Express* **10**(15), 650–653 (2002). <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-10-15-653>.
23. A. B. Thompson and E. M. Sevick-Muraca, "NIR fluorescence contrast enhanced imaging with ICCD homodyne detection: measurement precision and accuracy," *J. Biomed. Opt.* **8**, 111–120 (2002).
24. A. Joshi, W. Bangerth, and E. Sevick-Muraca, "Adaptive finite element methods for fluorescence enhanced frequency domain optical tomography: Forward imaging problem," in *International Symposium on Biomedical Imaging*, pp. 1103–1106 (IEEE, 2004).
25. R. A. Adams, *Sobolev Spaces* (Academic Press, 1975).
26. A. N. Tikhonov and V. Y. Arsenin, eds., *Solution of Ill-Posed Problems* (Winston, Washington, DC, 1977).
27. S. R. Arridge and M. Schweiger, "A Gradient Based Optimization Scheme for Optical Tomography," *Opt. Express* **2**(6), 213–225 (1998). <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-2-6-213>.
28. L. Beilina, "Adaptive Hybrid FEM/FDM Methods for Inverse Scattering Problems," Ph.D. thesis, Chalmers University of Technology (2002).
29. R. Becker, H. Kapp, and R. Rannacher, "Adaptive Finite Element Methods for Optimal Control of Partial Differential Equations: Basic Concept," *SIAM J. Contr. Optim.* **39**, 113–132 (2000).
30. R. Luce and S. Perez, "Parameter identification for an elliptic partial differential equation with distributed noisy data," *Inverse Problems* **15**, 291–307 (1999).
31. D. G. Luenberger, *Optimization by Vector Space Methods* (John Wiley, 1969).
32. J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research (Springer, New York, 1999).
33. S. C. Brenner and R. L. Scott, *The Mathematical Theory of Finite Elements* (Springer, Berlin-Heidelberg-New York, 1994).
34. R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques* (Wiley/Teubner, New York, Stuttgart, 1996).
35. W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations* (Birkhäuser Verlag,

- 2003).
36. M. Molinari, S. J. Cox, B. H. Blott, and G. J. Daniell, "Adaptive Mesh Refinement techniques for Electrical Impedance Tomography," *Physiological Measurement* **22**, 91–96 (2001).
 37. M. Molinari, B. H. Blott, S. J. Cox, and G. J. Daniell, "Optimal Imaging with Adaptive Mesh Refinement in Electrical Impedance Tomography," *Physiological Measurement* **23**, 121–128 (2002).
 38. R. Becker and R. Rannacher, "An optimal control approach to error estimation and mesh adaptation in finite element methods," *Acta Numerica* **10**, 1–102 (2001).
 39. H. Ben Ameur, G. Chavent, and J. Jaffré, "Refinement and coarsening indicators for adaptive parametrization: application to the estimation of hydraulic transmissivities," *Inverse Problems* **18**, 775–794 (2002).
 40. A.-A. Grimstad, H. Krüger, T. Mannseth, G. Nævdal, and H. Urkedal, "Adaptive selection of parameterization for reservoir history matching," in *ECMOR VIII: 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany*, pp. E–46 (European Association of Geoscientists and Engineers (EAGE), 2002).
 41. R. Li, W. Liu, H. Ma, and T. Tang, "Adaptive finite element approximation for distributed elliptic optimal control problems," *SIAM J. Control Optim.* **41**, 1321–1349 (2002).
 42. D. W. Kelly, J. P. d. S. R. Gago, O. C. Zienkiewicz, and I. Babuška, "A posteriori error analysis and adaptive processes in the finite element method: Part I—Error Analysis," *Int. J. Num. Meth. Engrg.* **19**, 1593–1619 (1983).
 43. W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II Differential Equations Analysis Library, Technical Reference* (2004). <http://www.dealii.org/>.
 44. A. Thompson, "Development of a new optical imaging modality for detection of fluorescence enhanced disease," Ph.D. thesis, Texas A & M University (2003).
 45. M. Huang and Q. Zhu, "Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information," *Appl. Opt.* **43**(8), 1654–1662 (2004).
 46. X. Gu, Y. Xu, and H. Jiang, "Mesh-based enhancement schemes in diffuse optical tomography," *Med. Phys.* **30**(5), 861–869 (2003).

1. Introduction

Molecular Imaging is a rapidly advancing research area with the potential of providing early diagnosis and identification of the underlying biochemical causes of human diseases [1]. As near infrared (NIR) light can travel several centimeters in tissue, fluorescence enhanced NIR optical imaging promises to open new pathways for the characterization of biological processes in living animals at cellular and molecular levels.

In the past decade, several approaches have been proposed for fluorescence enhanced optical tomography involving the determination of the fluorophore yield and/or fluorescence lifetime distribution in the tissue from a finite number of boundary measurements. Due to the diffusive nature of photon propagation in tissue, fluorescence-enhanced tomography represents a non-trivial inverse problem. Initial efforts did not involve classical optimization but instead focused on approaches such as: perturbative localization [2, 3, 4], backprojection [5], Born Approximation [6, 7, 8], random walk theory [9], and more recently, fast fluorescence localization [10]. These approaches were successful in locating small fluorescent targets in otherwise homogeneous, small sample volumes. Other approaches to fluorescence tomography cast the image reconstruction problem as an optimization problem in which a least squares type minimization is performed in order to determine the fluorescence map which best predicts the measured boundary fluorescence distribution. Optimization approaches are more general in their scope and can handle heterogeneous backgrounds as well as large sample volumes albeit at increased computational cost. These approaches include algorithms based on Newton's or Newton-type optimization methods [11, 12, 13] and Bayesian nonlinear least squares approaches [14, 15, 16, 17].

The sensitivity of fluorescence enhanced imaging potentially rivals that of the conventional, but "gold-standard" molecular imaging using radiotracers [18]. Yet, the achievable resolution for fluorescence tomography is determined first by the signal to noise ratio, and secondly by the level of discretization. To date, the discretization level is selected *a priori* based on knowledge of the domain and/or computational constraints. Image quality can be improved by uniformly refining the level of discretization throughout the domain. However, this global refinement further increases the ill-posedness of the problem and results in insurmountable computational

requirements by increasing the number of unknowns. For example, to achieve a resolution of one millimeter in a volume of one liter would require the use of 10^9 mesh points, a number that is clearly not achievable with today's computational technologies. In contrast, adaptive mesh refinement provides fine mesh resolution around target locations with coarser resolution in other regions to improve image quality, while maintaining solution stability and computational economy.

The outline of this article is as follows: In Section 2 the fluorescence tomography algorithm is described in a continuous function space setting and its discrete implementation with an adaptive mesh strategy is presented. Sections 3 and 4 detail the computational experiments conducted to demonstrate the adaptive tomography algorithm for reconstructing single and resolving dual fluorescence targets embedded in turbid media with a target to background ratio of 100 : 1. Finally, the results are summarized in Section 5 and we show that Rayleigh resolution (i.e. the minimum distance between which two target can be resolved from one another) can approach the continuum limit using the adaptive strategy. To the best of the authors' knowledge, this contribution represents the first time that adaptivity has been used in optical tomography to address the issue of reconstructed image quality.

2. Methodology

In this section, the formulation for the nonlinear inverse problem of fluorescence tomography is developed in a continuous function space setting, allowing separate and independent discretization of the parameter map and the finite element mesh used to solve state/adjoint problems for the nonlinear update steps toward the optimal solution. The mesh refinement criteria and implementation described herein are based on the general framework developed in [19, 20].

2.1. Formulation

Under conditions of multiple scattering, the generation and propagation of diffuse fluorescence photon density waves from modulated, time-periodic sources can be described by the following coupled system of diffusion equations [21]:

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

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

where

$$D_{x,m} = \frac{1}{3(\mu_{ax,mi} + \mu_{ax,mf} + \mu'_{sx,m})}, \quad k_{x,m} = \frac{i\omega}{c} + \mu_{ax,mi}(\mathbf{r}) + \mu_{ax,mf}(\mathbf{r}), \quad \beta_{xm} = \frac{\phi\mu_{axf}}{1 - i\omega\tau(\mathbf{r})}.$$

Here, an index x denotes the excitation light field and m denotes the emission field; u, v are the complex-valued photon fluence fields at excitation and emission wavelengths, respectively; (Note that part of the literature uses the symbols Φ_x, Φ_m for these variables; u, v are used to avoid overly complicated expressions with many indices in the next sections.) $D_{x,m}$ are the photon diffusion coefficients; $\mu_{ax,mi}$ is the absorption coefficient due to endogenous chromophores; $\mu_{ax,mf}$ is the absorption coefficient due to exogenous fluorophores; $\mu'_{sx,m}$ is the reduced scattering coefficient; ω is the modulation frequency; ϕ is the quantum efficiency of the fluorophore; finally, τ is the fluorophore lifetime associated with first order fluorescence decay kinetics. The fluorescence generation mechanism is detailed in [21]. These equations are complemented by Robin-type boundary conditions on the boundary $\partial\Omega$ of the domain Ω :

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

where n denotes the outward normal to the surface and γ is a constant depending on the optical reflective index mismatch at the boundary [22]. $S(\mathbf{r})$ is the excitation boundary source. There is no source term for the emission boundary condition. Note that, here the NIR excitation source is modeled as a boundary condition. This is an approximation as the true isotropic source will be one scattering length below the illumination surface. This approximation is justified for a diffuse area illumination scheme [23] instead of the traditional point illumination by fiber optics and the finite element solutions of equations (1)–(3) match the experimentally observed boundary fluorescence [24]. The goal of fluorescence tomography is the reconstruction of a spatial map of coefficients $\mu_{axf}(\mathbf{r})$ and/or $\tau(\mathbf{r})$ from measurements of the fluences u, v on the boundary.

Instead of the (strong) formulation as a PDE above, equations (1)–(3) are posed in a weak variational form as is usual in finite element applications. The two equations are multiplied with arbitrary test functions $\zeta, \xi \in H^1$, integrated over Ω , and terms with second derivatives are integrated by parts. Here, H^1 is the Sobolev space of functions with integrable (weak) first derivatives (see Adams [25]). With this, the resulting variational equation reads:

$$A(q; [u, v])([\zeta, \xi]) = 0 \quad \forall \zeta, \xi \in H^1, \quad (4)$$

where the bilinear form A is nonlinear in its first set of arguments and linear in the test functions. q denotes the set of unknown parameters, i.e. μ_{axf} and/or τ . In this work, we only consider $q = \mu_{axf}$. With (\cdot, \cdot) denoting the L_2 inner product, the definition of A reads

$$\begin{aligned} A(q; [u, v])([\zeta, \xi]) &= (D_x \nabla u, \nabla \zeta)_\Omega + (k_x u, \zeta)_\Omega + \frac{\gamma}{2} (u, \zeta)_{\partial\Omega} + \frac{1}{2} (S, \zeta)_{\partial\Omega} \\ &\quad + (D_m \nabla v, \nabla \xi)_\Omega + (k_m v, \xi)_\Omega + \frac{\gamma}{2} (v, \xi)_{\partial\Omega} - (\beta_{xm} u, \xi)_\Omega. \end{aligned} \quad (5)$$

Note that $D_{x,m} = D_{x,m}(q)$, $k_{x,m} = k_{x,m}(q)$, $\beta_{xm} = \beta_{xm}(q)$.

The goal of the parameter identification problem is to find the set of parameters q for which the predicted boundary fluorescence measurements resulting from equation (4) best match actual measurements. The quantity to be minimized (i.e. the misfit between prediction and measurement) is then

$$J(q, v) = \frac{1}{2} \|v - z\|_\Sigma^2 + \beta r(q), \quad (6)$$

where the L_2 norm of the difference between the actual measurements, z , and the prediction of the emission fluence v on a part Σ of the boundary $\partial\Omega$ is minimized. In practice, z is interpolated between the pixels of an area detection system such as the gain modulated CCD camera. $r(q)$ is a Tikhonov regularization functional which penalizes certain undesirable aspects of solutions [26] and β is the regularization parameter. Information about measurement noise can be incorporated by using a different or weighted norm of the misfit [21].

In the past, optical tomography approaches have followed the output least squares formulation where the state variables u, v are taken to be dependent on the parameters q [27, 12]. In this work, however, we minimize the error functional J by treating $\{q, u, v\}$ as independent variables where their relationship is enforced by including the state equation as a constraint to the optimization problem [19, 20, 28, 29, 30]. We also deviate from the larger part of the literature in that we formulate the following steps in function spaces, rather than choosing the route of most optical tomography papers to first discretize state equation and objective function and then stating the optimization problem in the finite-dimensional space of matrices and vectors. While our approach leads to equations that at times look cumbersome, they are exactly equivalent to the matrix formulation if a fixed grid is used. However, they afford for the possibility of using

different grids for different nonlinear iterations; in other words, we do not yet want to settle on a fixed discretization and therefore cannot choose the more usual matrix formulation.

With these considerations, the minimization problem reads:

$$\min_{q,u,v} J(q,v) \quad \text{subject to} \quad A(q;[u,v])([\zeta,\xi]) = 0. \quad (7)$$

This constrained optimization problem can be considered in a Lagrangian framework where the Lagrangian is defined as:

$$L([u,v],[\lambda^{ex},\lambda^{em}],q) = J(q,v) + A(q;[u,v])([\lambda^{ex},\lambda^{em}]). \quad (8)$$

Here, $\lambda^{ex}, \lambda^{em}$ are the Lagrange multipliers corresponding to the excitation and emission diffusion equation constraints, respectively. For simplicity, the abbreviation of $x = \{u,v,\lambda^{ex},\lambda^{em},q\}$ is introduced so that the Lagrangian functional can be written as $L(x)$. The optimum is then characterized by a stationary point of the Lagrangian, i.e. the first order conditions:

$$L_x(x)(y) = 0 \quad \forall y = \{\varphi^{ex}, \varphi^{em}, \psi^{ex}, \psi^{em}, \chi\}, \quad (9)$$

where $L_x(x)(y)$ is the Fréchet differential [31] of $L(x)$ and y denotes possible test functions. Eq. (9) can be expanded for all components of x :

$$L_u(x)(\varphi^{ex}) = A_u(q;[u,v])(\varphi^{ex})([\lambda^{ex},\lambda^{em}]) = 0, \quad (10)$$

$$L_v(x)(\varphi^{em}) = J_v(q,v)(\varphi^{em}) + A_v(q;[u,v])(\varphi^{em})([\lambda^{ex},\lambda^{em}]) = 0, \quad (11)$$

$$L_{\lambda^{ex}}(x)(\psi^{ex}) = A(q;[u,v])([\psi^{ex},0]) = 0, \quad (12)$$

$$L_{\lambda^{em}}(x)(\psi^{em}) = A(q;[u,v])([0,\psi^{em}]) = 0, \quad (13)$$

$$L_q(x)(\chi) = J_q(q,v)(\chi) + A_q(q;[u,v])(\chi)([\lambda^{ex},\lambda^{em}]) = 0. \quad (14)$$

Subscripts q,u,v,λ^{ex} , and λ^{em} indicate first order partial Fréchet derivatives of J or A . Equations (12)–(13) are the state equations in variational form. Equation (12) can be solved to provide the excitation fluence u which is then used to solve equation (13) to obtain the emission fluence v . Equations (10)–(11) are the adjoint equations defining the Lagrange multipliers $[\lambda^{ex},\lambda^{em}]$. Finally, equation (14) is the control equation.

The above set of coupled nonlinear equations is solved by Newton's method. The update direction for the k th iteration, $\delta x_k = \{\delta u_k, \delta v_k, \delta \lambda_k^{ex}, \delta \lambda_k^{em}, \delta q_k\}$, is determined from

$$L_{xx}(x_k)(\delta x_k, y) = -L_x(x_k)(y) \quad \forall y, \quad (15)$$

where $L_{xx}(x_k)$ is the Hessian matrix of second derivatives of L at point x_k . These equations represent one condition for each variable in δx_k and when expanded read as follows [19, 31]:

$$\begin{aligned} A_u(q_k;[u_k,v_k])(\varphi^{ex})([\delta \lambda_k^{ex},0]) + A_u(q_k;[u_k,v_k])(\varphi^{ex})([0,\delta \lambda_k^{em}]) \\ + A_{uq}(q_k;[u_k,v_k])(\varphi^{ex},\delta q_k)([\lambda^{ex},\lambda^{em}]) = -L_u(x_k)(\varphi^{ex}), \end{aligned}$$

$$\begin{aligned} J_{vv}(q_k,v_k)(\delta v_k,\varphi^{em}) + A_v(q_k;[u_k,v_k])(\varphi^{em})([\delta \lambda_k^{ex},0]) + A_v(q_k;[u_k,v_k])(\varphi^{em})([0,\delta \lambda_k^{em}]) \\ + J_{vq}(q_k,v_k)(\delta q_k,\varphi^{em}) + A_{vq}(q_k;[u_k,v_k])(\varphi^{em},\delta q_k)([\lambda^{ex},\lambda^{em}]) = -L_v(x_k)(\varphi^{em}), \end{aligned}$$

$$A_u(q_k; [u_k, v_k])(\delta u_k)([\psi^{ex}, 0]) + A_v(q_k; [u_k, v_k])(\delta v_k)([\psi^{ex}, 0]) \\ + A_q(q_k; [u_k, v_k])(\delta q_k)([\psi^{ex}, 0]) = -L_{\lambda^{ex}}(x_k)(\psi^{ex}),$$

$$A_u(q_k; [u_k, v_k])(\delta u_k)([0, \psi^{em}]) + A_v(q_k; [u_k, v_k])(\delta v_k)([0, \psi^{em}]) \\ + A_q(q_k; [u_k, v_k])(\delta q_k)([0, \psi^{ex}]) = -L_{\lambda^{em}}(x_k)(\psi^{em}),$$

$$A_{qu}(q_k; [u_k, v_k])(\delta u_k, \chi)([\lambda^{ex}, \lambda^{em}]) + A_{qv}(q_k; [u_k, v_k])(\delta v_k, \chi)([\lambda^{ex}, \lambda^{em}]) \\ + J_{qv}(q_k, v_k)(\delta v_k, \chi) + J_{qq}(q_k, v_k)(\delta q_k, \chi) + A_q(q_k; [u_k, v_k])(\chi)([\delta \lambda^{ex}, 0]) \\ + A_q(q_k; [u_k, v_k])(\chi)([0, \delta \lambda^{em}]) = -L_q(x_k)(\chi).$$

From Eq.s (10)–(11) one can infer that the Lagrange multipliers $[\lambda^{ex}, \lambda^{em}]$ are proportional to $J_v(q; v)$. As a consequence, all terms involving $[\lambda^{ex}, \lambda^{em}]$ become negligible near the optimal solution under conditions of low noise and can be dropped from the Hessian of the Lagrange functional, resulting in the simplified Gauss-Newton method. The equations for the Gauss-Newton method are then the same as above with the exception that (i) the last terms on the left hand side of the first and second equations as well as (ii) the first and second term of the last equation are eliminated.

Once the search direction is computed from Eq. (15), the actual update is determined by calculating a safeguarded step length α_k :

$$x_{k+1} = x_k + \alpha_k \delta x_k. \quad (16)$$

The step-length α_k can be computed from one of several methods, such as the Goldstein-Armijo backtracking line search [12, 19, 20].

In many cases, bounds on the parameters q are available. For example, background or maximal uptake concentrations may be known. This information should be used in the inverse problem to stabilize its solution, as well as to enforce physically reasonable solutions. Thus, we incorporate bounds $q_0 \leq q(\mathbf{r}) \leq q_1$ using the scheme presented in [19, 20]. The method is a variation of the active set strategy, see [32]. At the beginning of each iteration, the set of parameters which lie at either of the bounds is identified. A first order approximation based on the gradient of the Lagrangian is used to determine whether the parameters are likely to move out of the feasible region. If so, then the update for these parameters is constrained to be zero. This method is computationally efficient since the determination of the likely direction uses only the information that is already available. Furthermore, enforcing a zero constraint on the update is equivalent to removing rows and columns from the Schur complement matrix described in the next section. In particular, no penalty terms need to be integrated into the objective function.

2.2. Discretization

In the previous section, we have formulated the Gauss-Newton method in function spaces. For carrying out actual computations, we discretize the Gauss-Newton equations with the finite element method and choose $\{\varphi_i\}$ as the basis functions for the state and adjoint variables u, v, λ^{ex} , and λ^{em} , and $\{\chi_i\}$ as the basis for the parameter q . Piecewise linear, continuous shape functions on hexahedral meshes for $\{\varphi_i\}$ and piecewise constant, discontinuous functions for $\{\chi_i\}$ are used. The discrete equations resulting from the Gauss-Newton modification of (15) are then

represented by the following block system:

$$\begin{bmatrix} M & 0 & P^T \\ 0 & R & C^T \\ P & C & 0 \end{bmatrix} \begin{bmatrix} \delta p_k \\ \delta q_k \\ \delta d_k \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}, \quad (17)$$

where the updates for the primal and dual variables are abbreviated as $\delta p_k = [\delta u_k, \delta v_k]^T$, $\delta d_k = [\delta \lambda_k^{ex}, \delta \lambda_k^{em}]^T$. The blocks of the matrix are defined as

$$M = \begin{bmatrix} 0 & 0 \\ 0 & (\varphi_i, \varphi_j)_\Sigma \end{bmatrix}_{ij}, \quad R = [\beta r''(q_k, \chi_i, \chi_j)]_{ij}, \quad P^T = \begin{bmatrix} A_{global}^{ex} & -B_{global}^{ex \rightarrow em} \\ 0 & A_{global}^{em} \end{bmatrix}. \quad (18)$$

Here, P is the representation of the discrete forward diffusion model; A_{global}^{ex} and A_{global}^{em} are the global stiffness matrices associated with the excitation and emission diffusion equations; and $B_{global}^{ex \rightarrow em}$ is the matrix which couples the excitation and emission linear systems. The subscripts i, j in the preceding equations iterate over all degrees of freedom. The matrix C is defined as $C^T = [C_1, C_2]$ with its components obtained by differentiating the semilinear form A in Eq. (5) with respect to the parameter q :

$$\begin{aligned} C_1 &= \left(\frac{\partial D_x(q_k)}{\partial q} \nabla u_k \cdot \nabla \psi_i, \chi_j \right)_{ij} + \left(\frac{\partial k_x(q_k)}{\partial q} u_k \psi_i, \chi_j \right)_{ij}, \\ C_2 &= \left(\frac{\partial D_m(q_k)}{\partial q} \nabla v_k \cdot \nabla \psi_i, \chi_j \right)_{ij} + \left(\frac{\partial k_m(q_k)}{\partial q} v_k \psi_i, \chi_j \right)_{ij} - \left(\frac{\partial \beta_{xm}(q_k)}{\partial q} u_k \psi_i, \chi_j \right)_{ij}. \end{aligned} \quad (19)$$

The right hand side terms in Eq. (17) are the discretized components of the negative gradient of the Lagrangian:

$$\begin{aligned} F_1 &= \begin{bmatrix} -(D_x(q_k) \nabla \lambda_k^{ex}, \nabla \varphi_i) - (k_x(q_k) \lambda_k^{ex}, \varphi_i) - \frac{\gamma}{2} (\lambda_k^{ex}, \varphi_i)_{\partial \Omega} + (\beta_{xm}(q_k) \lambda_k^{em}, \varphi_i) \\ -(v_k - z, \varphi_i)_\Sigma - (D_m(q_k) \nabla \lambda_k^{em}, \nabla \varphi_i) - (k_m(q_k) \lambda_k^{em}, \varphi_i) - \frac{\gamma}{2} (\lambda_k^{em}, \varphi_i)_{\partial \Omega} \end{bmatrix}_i, \\ F_2 &= \begin{bmatrix} -\beta r'(q_k, \chi_i) - \left(\frac{\partial D_x(q_k)}{\partial q} \nabla u_k \cdot \nabla \lambda_k^{ex}, \chi_i \right) - \left(\frac{\partial k_x(q_k)}{\partial q} u_k \lambda_k^{ex}, \chi_i \right) \\ + \\ - \left(\frac{\partial D_m(q_k)}{\partial q} \nabla v_k \cdot \nabla \lambda_k^{em}, \chi_i \right) - \left(\frac{\partial k_m(q_k)}{\partial q} v_k \lambda_k^{em}, \chi_i \right) + \left(\frac{\partial \beta_{xm}(q_k)}{\partial q} u_k \lambda_k^{em}, \chi_i \right) \end{bmatrix}_i, \\ F_3 &= \begin{bmatrix} -D_x(q_k) \nabla \psi_i, \nabla u_k) - (k_x(q_k) \psi_i, u_k) - \frac{1}{2} (S_x, \psi_i)_{\partial \Omega} - \frac{\gamma}{2} (\psi_i, u_k)_{\partial \Omega} \\ -D_m(q_k) \nabla \psi_i, \nabla v_k) - (k_m(q_k) \psi_i, v_k) - \frac{\gamma}{2} (\psi_i, v_k)_{\partial \Omega} + (\beta_{xm}(q_k) \psi_i, u_k) \end{bmatrix}_i. \end{aligned} \quad (20)$$

For large domains and fine discretization, the linear system Eq. (17) can be as large as several 100,000 unknowns. Furthermore, the matrix is generally indefinite, restricting the choice of available iterative linear solvers for determining the Gauss-Newton update directions. Thus, instead of solving Eq. (17) directly, we use an efficient solver based on the Schur complement developed in [19, 20]. By block elimination, Eq. (17) is reduced to the following sequence of three equations:

$$\{R + C^T P^{-T} M P^{-1} C\} \delta q_k = F_2 - C^T P^{-T} F_1 + C^T P^{-T} M P^{-1} F_3, \quad (21)$$

$$P \delta p_k = F_3 - C \delta q_k, \quad (22)$$

$$P^T \delta d_k = F_1 - M \delta p_k. \quad (23)$$

The matrix $R + C^T P^{-T} M P^{-1} C$ is the Schur complement matrix of the Gauss-Newton system. Under practical conditions it is symmetric positive definite [19] and can thus be inverted ef-

ficiently by iterative solvers such as the conjugate gradient method. Furthermore, the Schur complement matrix is comparatively small (at most a few 1,000 to 10,000) with its size equivalent to the number of discretized parameters, rather than the number of parameters, state, and adjoint variables combined. Since the Newton method updates only approximates the direction to the optimal solution, it is not necessary to solve Eq. (21) exactly for each Newton step. The conjugate gradient iteration is thus stopped once the l_2 residual falls below a certain tolerance, for example 10^{-3} times its initial value. Consequently, this method is a variant of truncated or inexact Gauss-Newton methods [32].

2.3. Adaptive Mesh Refinement

The accuracy of finite element solutions of partial differential equations depends on the mesh width. This is reflected in *a priori* error estimates comparing the exact solution u and the numerically computed finite element solution u_h . For the simple example of the Laplace equation, an *a priori* error estimate would be [33]:

$$\|u - u_h\| \leq C(u) h^2, \quad (24)$$

where h is the maximum mesh size and $C(u)$ is a constant that depends on the exact (and unknown) solution, the domain, and element shapes, but not upon the mesh width. Thus, the accuracy of the finite element solution can be increased by decreasing the maximum mesh size h . Similar estimates may be derived for the system in Eq.s (1)–(2). The drawback of *a priori* estimates is that the exact solution, and thus the numerical value of $C(u)$, is unknown, so that no quantitative bound on the error can be computed.

While Eq. (24) shows that it can be guaranteed that global mesh refinement reduces the error in the solution, this does not represent an efficient strategy in general. A fine mesh is only necessary where the solution varies greatly and where small cells are needed to accurately capture this variation. Since it is in general unknown in advance where these locations are, *a posteriori* error estimates have been derived in the mathematical community to provide criteria for local mesh refinement. These estimates use the computed solution u_h to not only provide a bound on the error $\|u - u_h\|$ without requiring knowledge of the exact solution u , but also to indicate on which cells the contribution to this error is largest. Thus, these estimates indicate the cells for which mesh refinement will be most beneficial, and conversely for which cells, mesh refinement will not yield a significant contribution to the reduction of the error. Using this process, the mesh on which u_h was computed can be appropriately refined, and the solution is computed again on the refined mesh; this process is iterated until either the error estimate indicates that the requested accuracy is obtained, or computational resources are exhausted. The method in which only selected cells are repeatedly refined based upon an error indicator is commonly referred to as *adaptive mesh refinement*. For nonlinear problems, an additional advantage is realized by performing the initial Gauss-Newton updates on coarse meshes: finer and thus computationally more expensive meshes are employed only near the solution which reduces both the ill-posedness of the problem and the computational work in the initial steps.

A posteriori error estimates and adaptive mesh refinement have been extensively studied in the last two decades. See [34, 35] and the references therein for an overview. However, most of the previous work is tailored to model equations rather than parameter estimation problems with the exception of Molinari *et al.* [36, 37] for adaptive mesh refinement in electrical impedance tomography (EIT); see also [19, 20, 29, 38, 28, 39, 40, 41] for approaches in other fields.

In this work, we use two separate meshes that are adaptively refined. The first one is used for the discretization of state and adjoint variables u, v and $\lambda^{ex}, \lambda^{em}$, while the second one is used for the discretization of the parameter field q . The state/adjoint mesh is finer than the parameter mesh in order to avoid stability problems with the saddle point problem Eq. (17). In addition,

and in accordance with the regularity requirements of the equation, we discretize state and adjoint variables using piecewise tri-linear finite elements, and the parameter with piecewise constant functions. Whenever Gauss-Newton iterations on these meshes have reduced the error function by a significant amount, both meshes are refined using *a posteriori* refinement criteria. In this work, the state and adjoint mesh is refined using a variation of the refinement criterion first derived by Kelly *et al.* [42]:

$$\eta_K^u = \frac{h}{24} \|[\partial_n u_h]\|_{\partial K}^2, \quad \eta_K^v = \frac{h}{24} \|[\partial_n v_h]\|_{\partial K}^2, \quad \eta_K = \alpha \eta_K^u + (1 - \alpha) \eta_K^v, \quad (25)$$

where $[\partial_n u_h]$ is the jump of the normal derivative of the finite element solution u_h across the element boundary ∂K , and α is chosen so that the errors in the two variables are roughly weighted equally. In our implementation, we refine those 35% of elements with highest error indicator η_K and coarsen those 5% of elements with lowest errors in each cycle. Some other elements are also refined to maintain numerical stability of the solution.

The mesh for q is refined by computing, for each cell, a discrete approximation to the gradient of q weighted by the local mesh width [19]. Note that the choice of two separate meshes means that the first mesh can be fine close to the source where the excitation fluence greatly varies, while the second mesh will only be fine close to the fluorescent target and coarse everywhere else. The detailed derivation of error estimates can be found elsewhere [19, 28, 35].

2.4. Software Implementation

A simplified block diagram of the tomography algorithm is shown in Fig. 1. At the beginning of the algorithm, all variables are initialized on coarse meshes (mesh level $m = 1$): state/adjoint variables are set to zero, while the parameter is set to the lower bound q_0 (the minimal background concentration). In each subsequent iteration, Gauss-Newton update directions are determined from Eq.s (21)–(23) and all variables are updated after computing the step length. The program is terminated if the number of Gauss-Newton iterations exceeds the maximum number of iterations k_{\max} or if the model misfit $\frac{1}{2} \|v - z\|_{\Sigma}^2$ is reduced below a prespecified threshold ε . For the results reported in this paper we used $k_{\max} = 40$ and $\varepsilon = 10^{-15}$.

Mesh refinement is triggered if either (i) the Gauss-Newton step length α_k is less than a prespecified minimum step length α_{\min} , or (ii) the nonlinear residual, $r_k = \|L_x(x_k)(\cdot)\|$, of the optimality condition (9) has been sufficiently reduced on the current mesh m , i.e. $r_k \leq \varepsilon_{\text{mesh}} r_{k_m^0}$, where k_m^0 denotes the first iteration on mesh level m . In this work $\alpha_{\min} = 0.15$ and the error reduction threshold is $\varepsilon_{\text{mesh}} = 10^{-4}$. Mesh refinement is followed by re-computation of synthetic measurements z on the new, finer mesh.

This scheme for fluorescence tomography with adaptive mesh refinement was implemented in C++ based on the deal.II finite element library [43]. Deal.II provides advanced object oriented design techniques and support for the complex data structures needed for adaptive finite element applications. While the current version of the software is used for an area excitation illumination geometry as previously published [23], the implementation is also suitable for point illumination geometries. Reconstructions were carried out on a dual processor (750MHz) Sun Sparc workstation with 3GB of RAM.

3. Computational Experiments

Two computational experiments were performed to test the efficacy of the proposed algorithm: image reconstruction of (i) a single fluorescent target, and (ii) of two closely spaced fluorescence targets of varying separation distance. The synthetic frequency-domain fluorescence data was generated on a simulated $8 \times 8 \times 8 \text{ cm}^3$ cube illuminated by a simulated expanded laser

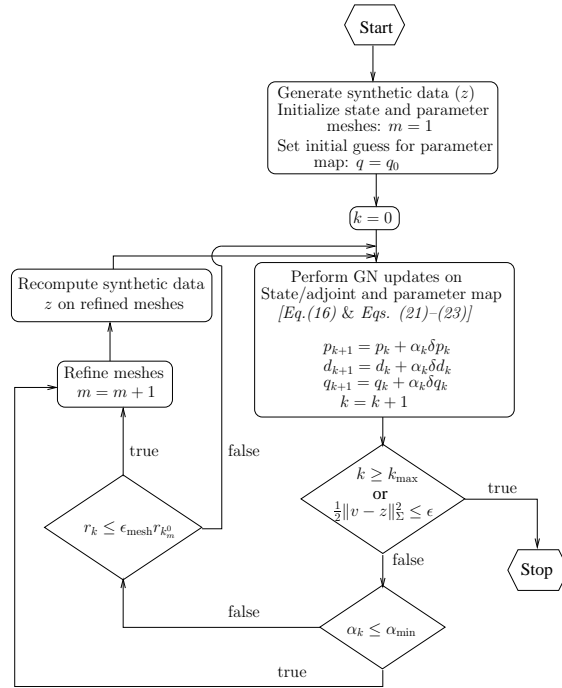


Fig. 1. Adaptive tomography algorithm. GN stands for Gauss-Newton; see Section 2.4 for a description of symbols.

beam with a Gaussian profile at the $x = 0$ plane, as shown in Fig. 2. This measurement geometry and phantom corresponds to experimental measurements previously reported [23]. Fluorescence phase and amplitude measurements were generated at the illumination surface in order to mimic the actual experimental data collection. Random Gaussian noise of zero mean and specified half width is then applied to these synthetic measurements.

3.1. Single Fluorescent Target

Data was generated for a 0.5cm diameter spherical target at a depth of 2.15cm from the illumination surface at an off-center position ($x = 2.15\text{cm}$, $y = 3.15\text{cm}$, $z = 3.15\text{cm}$). The absorption and isotropic scattering properties of 1% Liposyn solution were chosen to mimic the background of the phantom, corresponding to actual experimental measurements using a gain modulated ICCD camera system [23]. For the simulated background absorption coefficient, we chose $\mu_{axi} = 0.023\text{cm}^{-1}$ and $\mu_{ami} = 0.0289\text{cm}^{-1}$ [44]. The absorption coefficient due to fluorophore at the excitation wavelength was set to $\mu_{axf} = 0.5\text{cm}^{-1}$ in the target and 0.005cm^{-1} in the background. The emission wavelength absorption coefficient was $\mu_{amf} = 0.0506\text{cm}^{-1}$ in the target and 0.00506cm^{-1} in the background. The lifetime of the fluorophore was taken to be $\tau = 0.56\text{ns}$ and the quantum efficiency was $\phi = 0.016$ to match the corresponding properties of Indocyanine Green (ICG) dye used in experiments. The excitation wavelength for ICG

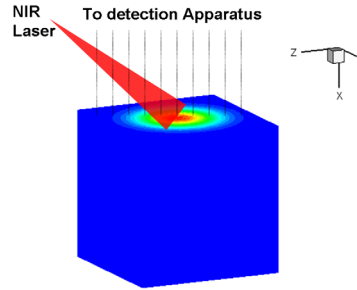


Fig. 2. Area illumination and area detection geometry employed by Thompson *et al.* [23]

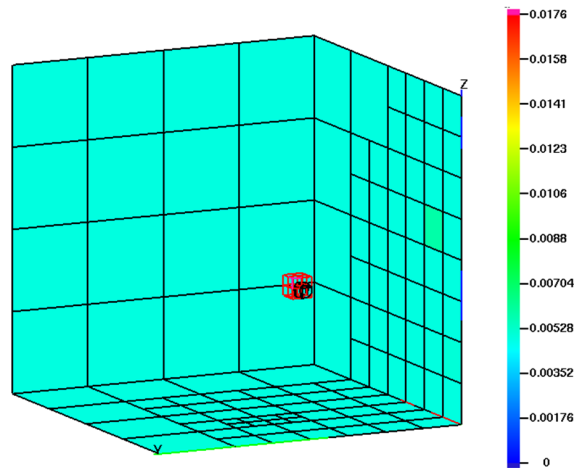


Fig. 3. Single target reconstruction: A black wire-frame depicts the actual target and colored blocks represent the reconstruction. Top 10% of the contour levels of μ_{axf} are shown.

is $785nm$ and the emission data is collected at $830nm$. The reduced scattering coefficient was taken as $\mu'_s = 9.84cm^{-1}$ in both the target and the background, and for this study was taken to be the same at excitation and emission wavelengths. Two percent random Gaussian noise was added to real and imaginary parts of the synthetic emission fluence solution at the measurement surface at $x = 0$.

3.2. Two Fluorescent Targets

Tomographic reconstructions were performed from synthetic data generated from the solution of Eq.s (1)–(3) for two closely spaced fluorescent targets to test the resolving ability of the adaptive inverse algorithm. Two spherical targets with diameter $0.4cm$ were placed at an edge to edge distance varying from $1cm$ to $0.1cm$, the latter of which is the continuum limit of the diffusion equation for the chosen optical properties. The depth of both targets was simulated to be $1.2cm$ from the illumination and measurement plane at $x = 0$. The optical properties of targets and background as well as the noise level remained the same as in the single target study.

4. Results

4.1. Single Fluorescent Target

Figure 3 illustrates the tomographic reconstruction for the single target case after 22 Gauss-Newton update steps utilizing approximately $2hrs$ of CPU time. The location of the target is

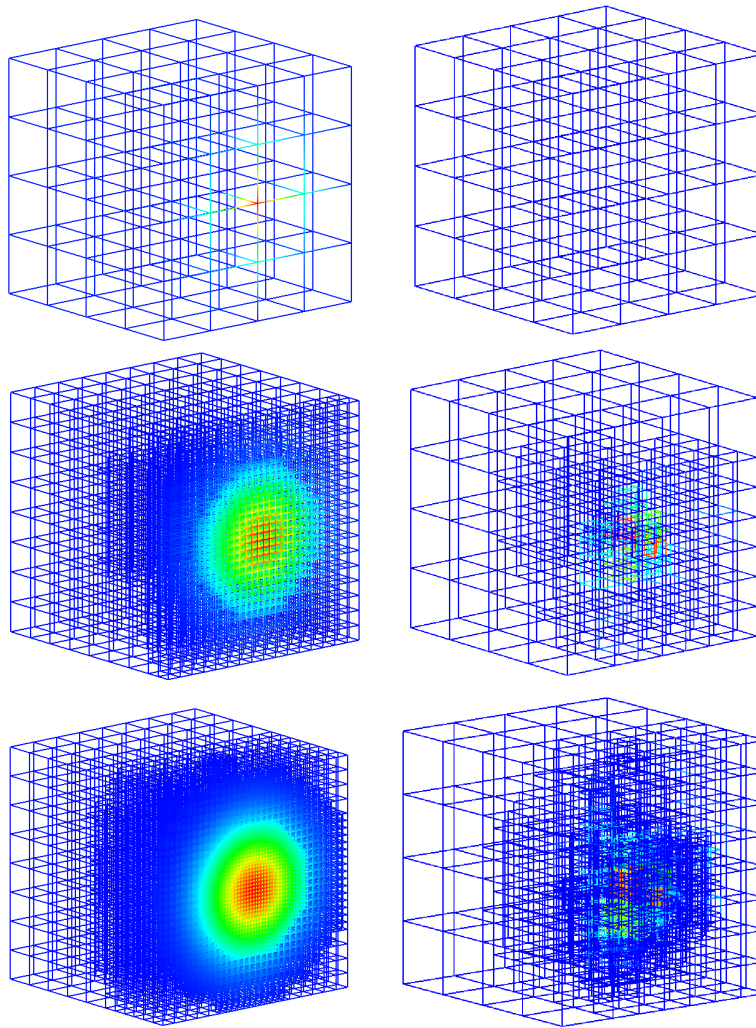


Fig. 4. Adaptive mesh evolution for state/adjoint (left) and parameter discretization (right). Meshes are shown at 1st, 11th and 22nd Gauss-Newton iterations.

reconstructed accurately. However, the magnitude 0.0176cm^{-1} of reconstructed μ_{axf} is lower than the actual value 0.5cm^{-1} , which is an artifact of the L_2 regularization used ($r(q) = \frac{1}{2}\|q\|^2$, with a fixed value $\beta = 10^{-3}$). Figure 4 shows the evolution of state/adjoint and parameter meshes during the reconstruction process. The algorithm started with coarse initial meshes with 64 hexahedral elements. Five automatic mesh refinements were carried out by the algorithm. The final state/adjoint mesh consisted of 116,936 nodes and was refined predominantly on the illumination plane to accurately resolve the Gaussian NIR excitation source. The final parameter mesh had 1016 elements, mostly located around the reconstructed fluorescent target. A uniformly refined mesh with the same parameter mesh resolution as the adaptive mesh surrounding the target would have $32^3 = 32,768$ unknowns. The advantages of adaptive refinement for the reduction of the total number of unknowns and the improvement of resolution are obvious.

d	Iter.	$\ q - q_{\text{true}}\ _2$	$\frac{1}{2} \ v - z\ _{\Sigma}^2$	$(x, y, z)_{\text{true}}$	$(x, y, z)_{\text{recovered}}$	N_q
1.0142	27	0.125	$2.58 \cdot 10^{-9}$	(1.2, 3.15, 3.15) (1.2, 4.15, 4.15)	(1.2, 3.13, 3.20) (1.2, 4.19, 4.22)	729
0.6607	18	0.127	$2.27 \cdot 10^{-9}$	(1.2, 3.625, 3.625) (1.2, 4.375, 4.375)	(1.4, 3.68, 3.625) (1.2, 4.30, 4.30)	1359
0.3071	27	0.119	$2.14 \cdot 10^{-9}$	(1.2, 4.23, 4.23) (1.2, 4.73, 4.73)	(1.1, 4.23, 4.27) (1.2, 4.73, 4.71)	1359
0.1657	28	0.118	$2.31 \cdot 10^{-9}$	(1.2, 3.8, 3.8) (1.2, 4.2, 4.2)	(1.3, 3.9, 3.8) (1.3, 4.1, 3.8)	638
0.1	25	0.121	$2.47 \cdot 10^{-9}$	(1.2, 3.82, 3.82) (1.2, 4.18, 4.18)	(1.2, 3.82, 3.82)	1247

Table 1. Summary of results for dual fluorescent target reconstructions. d is the edge to edge target separation in cm ; Iter. is the Gauss-Newton iteration for which the other results are reported; $\|q - q_{\text{true}}\|_2$ is the error in reconstructed parameter; $\frac{1}{2} \|v - z\|_{\Sigma}^2$ is the measurement error; N_q is the number of elements (unknowns) in the parameter mesh.

4.2. Two Fluorescent Targets

Figure 5 shows the tomographic reconstructions for targets separated by 1.0142, 0.6607, 0.3071, and 0.1657 cm . The reconstruction process was started with the same coarse state and parameter meshes as used in the single target case. A heuristic criterion was used to assess the reconstructions of the two targets: the top 10% contour levels of the μ_{axf} map were plotted and if two distinct maximum were visible, then the two targets were considered to be identified separately. The centroid of these two maxima was computed and treated as the reconstructed centroids. The centroids of the two targets was reconstructed accurately for target separations of 0.1657 cm and greater, when the reconstructed targets begin to appear as one.

At a target spacing of 0.1 cm , the reconstruction is biased towards one of the targets, and no maximum associated with a second target is resolved at this distance (see Fig. 6). This is consistent with the fact that this separation is approximately the mean isotropic scattering length, below which the continuum assumption necessary for the validity of the diffusion approximation is violated. No image resolution below this threshold can be expected.

Table 1 summarizes the computational results for the dual target reconstructions, demonstrating the efficacy of adaptive mesh refinement for determining the maximum achievable resolution for a given experimental configuration in the presence of measurement noise. The reconstructed targets appear slightly closer than the actual separation as the target spacing is reduced. This can be explained by the merging of the fluorescence emission fields with decreasing distance when using only the reflectance data for image reconstruction as we have done here.

5. Conclusions

In this work, we have demonstrated an efficient adaptive finite element algorithm for fluorescence optical tomography. The algorithm is formulated in function spaces independent of any *a priori* discretization in order to allow for meshes to change as nonlinear (Gauss-Newton) iterations progress. Independent meshes have been used for the state/adjoint variables and the parameter map, and physical information about upper and lower bounds on the unknown parameters have been incorporated to improve the stability of the algorithm. The choice of separately

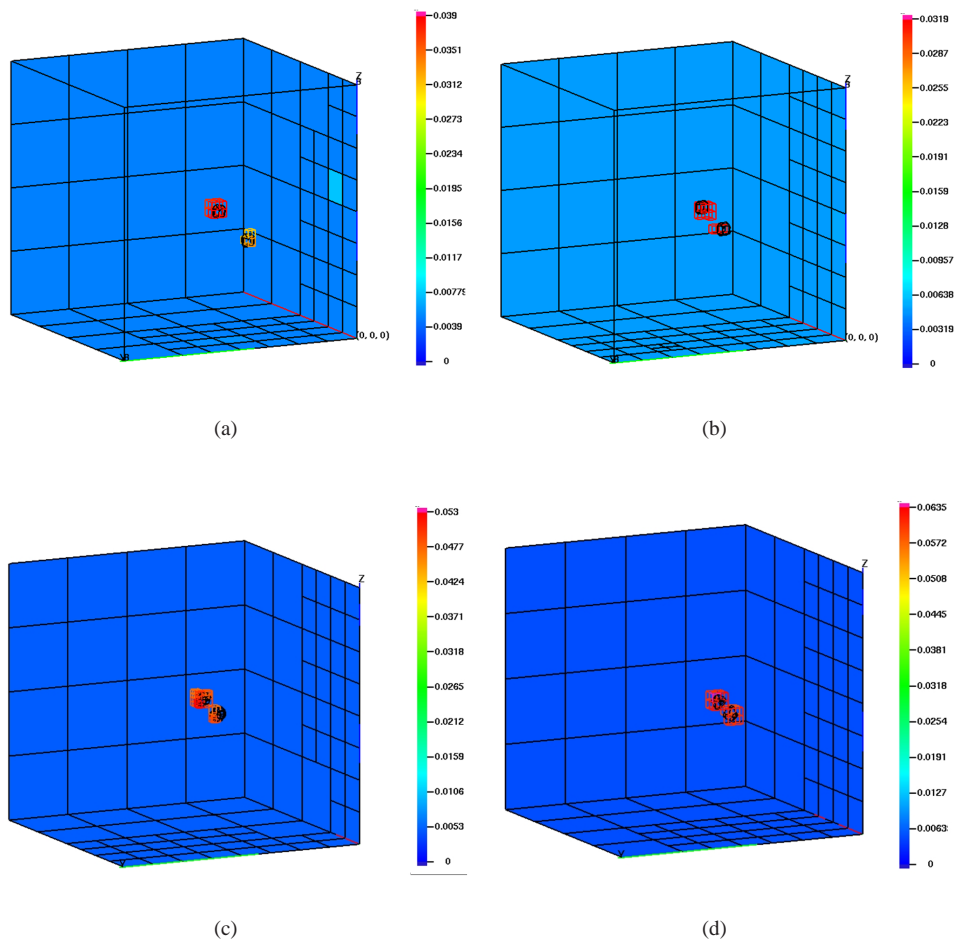


Fig. 5. Dual target reconstructions: A black wire-frame depicts the actual targets and colored blocks represent the reconstruction. Top 10% of the contour levels of μ_{axf} are shown. Edge to edge spacing: (a) 1.0142cm, (b) 0.6607cm, (c) 0.3071cm, and (d) 0.1657cm.

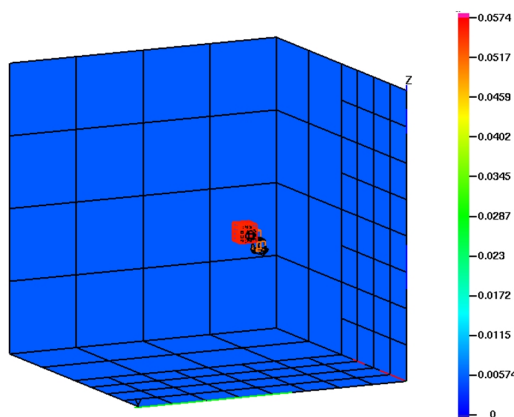


Fig. 6. Dual target reconstruction for 0.1cm target separation.

adapted meshes (with the parameter mesh being coarser), different shape functions for the state and parameter variables, as well as the inclusion of bounds on the parameter and a Tikhonov regularization term yield an algorithm that is able to cope with the well known ill-posedness of optical tomography for comparable high resolution imaging with acceptable numerical effort in the presence of noise. The discretization scheme can be contrasted with that recently proposed by Huang *et al.* [45], in which prior information about the subdivision (discretization) of the domain is obtained from ultrasound images. While their strategy requires that the target is structurally dissimilar to the surrounding tissue to obtain meaningful information via ultrasound or MRI, our algorithm is able to find high resolution discretizations automatically by using the solutions on coarse grids to generate finer grids where necessary. To reiterate, the final number of unknowns (between 729 and 1359) is 25 to 40 times less than that necessary with the use of conventional finite element based optical tomography schemes for large tissue volumes while still retaining the resolution in the regions of interest. In addition, the reduced number of unknowns in our adaptive finite element scheme greatly reduces the ill-posedness of the inverse problem and thus aids in the quality of the reconstruction. Jiang *et al.* [46] have also proposed a dual and adaptive meshing scheme for diffuse optical tomography, However their scheme doesn't implement automatic mesh adaptation with *a posteriori* error estimates.

We have demonstrated image reconstruction from frequency-domain fluorescence reflectance data arising from area-illumination. In comparison to circumferential point illumination measurements, reflectance measurements possess reduced information content for image reconstruction, but make the approach practical for patient imaging. This reflectance geometry and the associated 3-D adaptive finite element reconstruction algorithm could be pertinent for sentinel lymph node mapping as well as intraoperative localization of metastatic lesions (as used to track the progress of breast cancer). Our results provide the first report of a practical approach which may lead to high resolution molecular imaging in the clinic.

Acknowledgments

The authors acknowledge the support of the National Institutes of Health (NIH grant no. R01CA67176) and the Center for Subsurface Modeling at the Institute for Computational Engineering and Sciences, University of Texas at Austin.