INVERSE BIOMEDICAL IMAGING USING SEPARATELY ADAPTED MESHES FOR PARAMETERS AND FORWARD MODEL VARIABLES[1]

Wolfgang Bangerth,[†] *Amit Joshi,*[*] *and Eva Sevick*[*]

[†]Department of Mathematics, Texas A&M University, College Station, TX 77840
[*]Department of Radiology, Baylor College of Medicine, Houston, TX 77030

ABSTRACT

Many important existing and upcoming biomedical imaging modalities lead to nonlinear relationships between state variables from which measurements result and the tissue properties one would like to reconstruct, and typically involve partial differential equations. For such cases, exact reconstruction formulas are rarely known and their solution requires numerical techniques such as the finite element method that approximates all involved variables on a mesh.

Traditionally, the same mesh is used for all involved variables. In this contribution, we argue that this is inappropriate and meshes should be chosen independently of each other for the various variables involved. We support this claim through a numerical experiment.

1. INTRODUCTION

In a number of current and upcoming biomedical imaging methods, the relationship between the quantities that can be measured and the tissue properties one would like to reconstruct are nonlinear, and typically involve partial differential equations that relate the two. Examples for this are electrical impedance tomography (EIT, see [1]) as well as several recently developed optical tomography (OT) schemes (see, for example, [2, 3]).

In these inverse problems, exact reconstruction formulas are not usually known due to the nonlinearity of the problem. Their solution therefore requires the use of numerical techniques that convert the underlying partial differential equation into a finite-dimensional system of equations. The most widely used approach for this is the finite element method that replaces all distributed functions by piecewise polynomial approximations with finitely many degrees of freedom, and thereby replaces the nonlinear inverse problem by a nonlinear system of (possibly very many) algebraic equations.

Traditionally, the same mesh is taken for all variables involved in this. These are the state variable(s) that describe how the system reacts to external forcing (in EIT the electric potential, in OT the light field amplitude), the parameter we want to identify (in EIT the electric conductivity, in OT the light absorption and scattering coefficients or the concentration of fluorescent dye molecules), and possibly Lagrange multiplier(s) coupling the various equations. The number of variables increases if we make multiple experiments; in OT, for example, we may illuminate a body with differently patterned light sources and get a number of different light field amplitudes inside the body from measurements of all of which we hope to reconstruct our parameter fields [4]. The traditional approach is to also use the same mesh for all of these fields, even though each of these light fields may be localized in a different part of the domain.

In this contribution, we argue that this approach is inappropriate. In particular, efficient discretizations will lead to meshes for different variables that are chosen independently of each other, in order to capture the characteristics and localization of the variable discretized on it. In the above example, the mesh used for the light field in each of the experiments should be fine where the patterned source illuminates the body, whereas it may be coarse elsewhere (possibly including locations where we illuminate the body in a different experiment). In addition, the mesh for the parameter field that we want to recover should be independent of the meshes used for the state variables and should only reflect the properties of the recovered parameter.

Such an approach in which all meshes are chosen independently has two main advantages:

- Each mesh is adapted to the variables it is used for and can deliver high accuracy at an optimal cost.
- The mesh for the parameter we want to recover can be made as coarse as possible in regions where the parameter is smooth or even constant (e.g. in the background); this coarseness reduces the size of the problem, but also acts as an additional regularization that both suppresses unwanted oscillations as well as make the problem simpler to solve.

In the following, we will demonstrate our approach using a simple model problem akin to electrical impedance tomography. More complex applications of this technique to optical tomography can be found in [4, **?**].

[1]Part of this work has been funded through NIH grant no. R01 CA112679.

2. METHODS

Let us consider a simple model similar to electrical impedance tomography. In this model, electric currents are induced in a body by applying certain voltage patterns $g^i(\mathbf{r}), i = 1, \ldots, M$ to the surface. For each source pattern i, the electric potential $u^i(\mathbf{r})$ in the body then satisfies the equations

$$-\nabla \cdot q\nabla u^i = 0 \quad \text{in } \Omega, \qquad u^i = g^i \quad \text{on } \partial\Omega, \tag{1}$$

where $q = q(\mathbf{r})$ is the spatially varying electric conductivity we want to recover from measurements of the electric potential u or of currents $\mathbf{j^i} = q\nabla u^i$. Let us further consider the model situation where we are able to measure the electric potentials u^i everywhere; this corresponds to certain newer variants of electric impedance tomography such as acousto-electric tomography. Similarly, newer variants of optical tomography, such as ultrasound modulated optical tomography are able to probe the light intensity not only at the surface, but everywhere in the body.[1]

Under this assumption, our aim is to find a parameter $q(\mathbf{r})$ such that the predicted potentials u^i for all source patterns g^i match the actual measurements. Let us denote measurements for the i-th experiment by z^i. Then our goal is to find a parameter q that minimizes the misfit functional

$$J(\mathbf{u}, q) = \frac{1}{2}\sum_{i=1}^{M} \|u^i - z^i\|^2 + \beta\|q\|^2, \tag{2}$$

where the last term acts as a regularization term, $\|v\|^2 = \int_\Omega v^2 \; dx$, and the u^i are related to q and g^i through equation (1). If we consider (1) as a constraint to the minimization task (2), then the problem can be posed in the framework of constrained minimization [5] by introducing Lagrange multipliers λ^i for the state equations (1) and forming the Lagrangian

$$L(\mathbf{u}, \lambda, q) = J(\mathbf{u}, q) + \sum_{i=1}^{M} (q\nabla u^i, \nabla\lambda^i), \tag{3}$$

in addition to the constraint $u^i = g^i$ on $\partial\Omega$ which we enforce separately. Here, $(v, w) = \int_\Omega v\, w \; dx$. Note that we have incorporated the Laplace equation in its weak form, i.e. by multiplying the equation $-\nabla \cdot q\nabla u^i = 0$ by λ^i, integrating over Ω, and integrating by parts once.

The desired solution is then a stationary point of this Lagrangian, i.e. a point $(\mathbf{u}, \lambda, q)$ where all (variational) derivatives of the Lagrangian L vanish. The equations that define this point form a system of nonlinear coupled partial differential equations that can be solved using a Newton-type method safeguarded by a linear search.

[1] In contrast, in traditional EIT only measurements of $\mathbf{j}$ on the boundary $\partial\Omega$ are available, as is the light intensity u in traditional optical tomography. This lack of data leads to mathematically harder, more ill-posed problems. We consider the present, simpler situation in order to separate the effects of meshing and ill-posedness, the former being the focus of this paper.

Instead of going into the details of such an algorithm (for this, we refer to [6, 7]), let us here comment on the question of how to discretize the various variables involved. In particular, we note that the variables that need to be discretized are: (a) the M different state variables $u^i(\mathbf{r})$, (b) the M different Lagrange multipliers $\lambda^i(\mathbf{r})$, and finally (c) the parameter $q(\mathbf{r})$.

It turns out to be advantageous to choose $M + 1$ different meshes for the M state and adjoint variables on the one hand, and the parameter on the other hand. The reason for this choice is that these $M + 1$ sets of variables may have entirely different and independent properties. For example, if boundary sources g^i, g^j describe independent experiments where we have changed the location of electrodes through which we inject current, then it is clear that the current patterns for these two experiments need not have much in common. In fact, currents may flow through entirely different parts of the body, and a mesh that accurately resolves these current patterns for measurement i need not be good for measurement j. In order to focus numerical work to where it is necessary, we should therefore choose meshes independently adapted to each of these experiments. Because the Lagrange multiplier λ^i tracks how well u^i satisfies its state equation, it is reasonable to also discretize it on the same mesh.

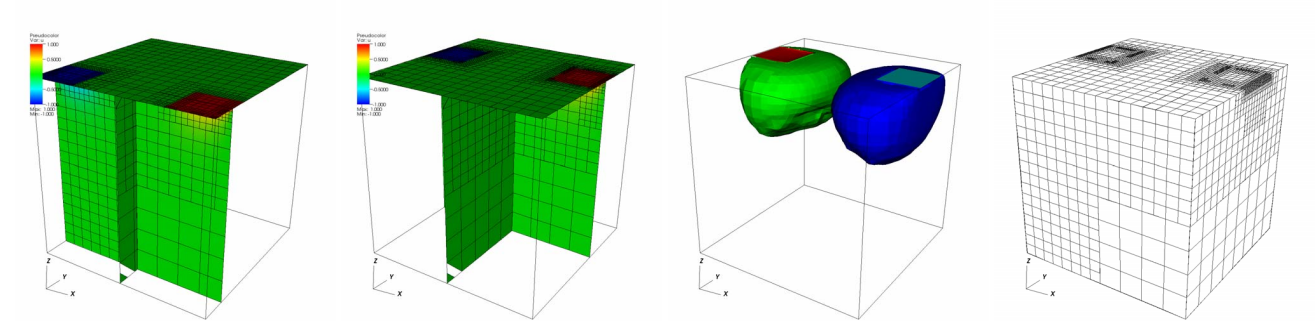
Likewise, the mesh used to resolve the parameter q should be fine where q jumps or has otherwise large variation. However, these locations do not necessarily have to coincide with where any of the state variables varies strongly. In particular, the latter often have their strongest variation close to the surface, due to the exponential decay with depth inherent in solutions of the Laplace equation, whereas, for example, if we are interested in tumor detection, we want to identify parameters that are discontinuous deep inside the body.

Our approach therefore is to start with a relatively coarse, uniform mesh for all these variables and perform a number of Newton steps towards the solution of the nonlinear inverse problem. We then evaluate the different functions using a smoothness indicator [7, 4], and refine each mesh independently and with the goal of adequately capturing the characteristics of the functions discretized on it. We then repeat this sequence of performing a few iterations followed by mesh refinement until we are satisfied with the results.

3. RESULTS

We present results for a test case where four different voltage patterns are applied to the top surface of the three-dimensional box $[-1, 1]^3$. The patterns are meant to imitate pairs of electrodes at positive and negative voltage attached to opposite sides of the top surface, and are moved over the surface for successive measurements. The top surfaces of Fig. 1a/b show the boundary conditions for the first and third measurement, respectively, i.e. g^1, g^3, along with the respective solutions u^1, u^3 on three planes through the domain.

Using these boundary conditions, we acquire "synthetic"



(a) (b) (c) (d)

Fig. 1. *Numerical solutions: state variables. (a,b) Boundary conditions on the top surface and cuts through the solution for experiments $i = 1$ and $i = 3$. (c) Isocontour surfaces for the electric potential for the third pair of electrodes. (d) The mesh used in the discretization of the electric potential for the third pair of electrodes.*

measurements $z^i, i = 1, \ldots, M$ by numerically solving

$$-\nabla \cdot q^* \nabla z^i = 0 \quad \text{in } \Omega, \qquad z^i = g^i \quad \text{on } \partial\Omega, \quad (4)$$

for the electric conductivity field $q^*(\mathbf{r})$ with value 8 inside a sphere of radius $\frac{1}{2}$ around the origin, and value 1 in the rest of the domain. Using these synthetic measurements, we then invoke the procedure described above to recover the parameter $q(\mathbf{r})$. Ideally, this reconstruction would be close to or equal to the "exact" value q^*, although we can not usually expect this for solutions of such ill-posed problems, in particular in areas of the domain far away from the sources. The reconstruction procedure is of course unaware of the exact q^*.

Fig. 1c shows iso-surfaces of the solution u^3. The banana-shaped region through which most of the current flows from one electrode to the other is easily imagined by recalling that current flows perpendicular to these iso-surfaces. The effect of the high-conductivity region at the center of the domain is clearly seen at the bottom of these surfaces.

Fig. 1d depicts the mesh used to discretize u^3. It was obtained after 5 refinement steps. It clearly shows the areas where the boundary condition g^3 has a jump and where the solution is correspondingly rough. This mesh has approximately 7,000 cells, which is also roughly the number of cells we obtain for the meshes used in the discretization of the other three experiments $i = 1, 2, 4$. We note that a uniformly refined mesh after 5 refinement would have 32,768 cells.

By comparing the numerical solutions u^i with the previously obtained synthetic measurements z^i, and varying the parameter field q until the mismatch between the two becomes minimal, we obtain a numerical approximation of the electric conductivity q. Fig. 2 shows the results of this reconstruction process and the mesh used for it.[2] The reconstruction clearly recovers the high-conductivity zone at the center and approximates its shape and size well. Its elongation is due to the fact that all electrode pairs are arranged left-right; a completely symmetric reconstruction can be obtained by adding front-back electrode measurements as well.

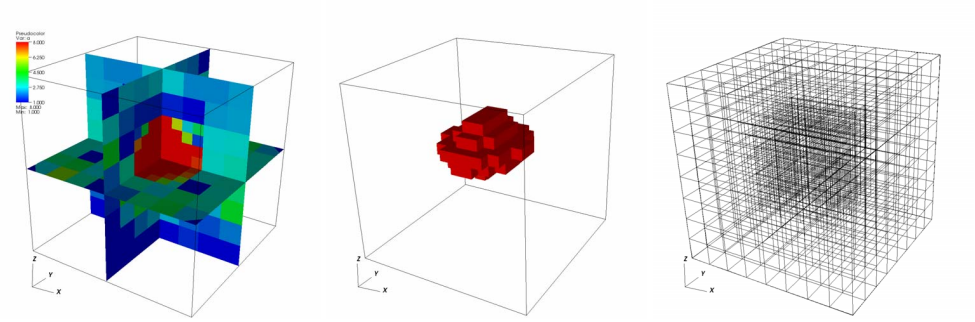
Since the smoothness properties of the reconstructed parameter field q (i.e. the discontinuities of the parameter deep in the domain) have little to do with the smoothness of each of the solutions u^i (which are rough primarily in the vicinity of the electrodes), it is clear that none of the meshes used for the u^i is relevant for the discretization of q. As discussed before, we therefore choose a completely independent mesh for it, shown in Fig. 2c. This mesh is fine only in the vicinity of the jump in the values of q, and coarse everywhere else, including in the regions where the state variables u^i are rough. It has 904 only cells, again much less than the 32,768 cells a uniformly refined mesh would have that was everywhere at least as fine as our adaptive mesh.

4. DISCUSSION

Comparing the number of cells in each of the meshes discussed above with those that would be necessary if one were to use a uniformly refined mesh makes it clear that our approach leads to large savings both in terms of memory consumption and run time of our reconstruction algorithm. Even though we have to deal with $M + 1$ different meshes that do not share a common memory representation or matrix structure, the savings are clearly large enough to still be beneficial.

Of particular importance is the vastly smaller number of mesh cells necessary to discretize the parameter $q(\mathbf{r})$, compared to a uniformly refined mesh or a hypothetical mesh that could be constructed by using the finest cells of all meshes involved. The importance of this lies in the fact that even with the best linear solvers, the numerical work to solve the inverse problem will grow at least linearly with the number of parameter unknowns, but also with the condition number of the problem. As mentioned in the previous section, the num-

[2]These reconstructions were obtained with a program based on the deal.II finite element library [8]. Run times on a laptop equipped with an Intel Core processor and 1GB of memory were roughly 5min for reconstructions with $M = 4$ independent measurements.



(a) (b) (c)

Fig. 2. *Numerical reconstructions of the electric conductivity. (a) Cross sections through the recovered parameter. (b) Volume at which the parameter is above 90% of the maximum value. (c) The mesh for the parameter after two refinement steps.*

ber of cells is already smaller by a factor of 35. The change in condition number is less tractable, but it is well known that for inverse problems the use of overly fine meshes for the parameter leads to badly conditioned problems and results in spurious oscillations. The relevance of adaptive meshes in this context is that the question whether a mesh is "overly fine" rests on the amount of information available at a location; in particular, the amount of information decreases exponentially with depth and with the distance from the sources.

It is therefore important to choose a coarse mesh deep in the body and far away from the source patterns. The resulting decreased condition number of the linear subproblems solved in each Newton step consequently leads not only to faster iterative solvers, but also avoids additional spurious oscillations of the reconstructed parameter and thereby reduces the need for artificial and non-physical regularization strategies.

5. CONCLUSIONS

Through numerical examples, we have shown that our approach of using separate meshes for the solutions u^i of the state equations as well as for the parameter q that we want to reconstruct, leads to significant savings in the number of unknowns. This is important for both reducing the computational requirements for numerically solving the inverse problem, as well as for the accuracy of solutions and for the suppression of spurious oscillations.

Although the numerical results presented here are for a problem akin to electrical impedance tomography, the techniques are applicable to similar inverse problems as well. In particular, applications to fluorescence-enhanced optical tomography can be found in [7, 4], with savings on a similar order of magnitude.

6. REFERENCES

[1] L. Borcea, "Electrical impedance tomography," *Inverse Problems*, vol. 18, pp. R99–R136, 2002.

[2] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems*, vol. 15, pp. R41–R93, 1999.

[3] 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*, chapter 33, Biomedical Photonics Handbook. CRC Press, 2003.

[4] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Non-contact fluorescence optical tomography with scanning patterned illumination," *Optics Express*, vol. 14, no. 14, pp. 6516–6534, 2006.

[5] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research. Springer, New York, 1999.

[6] W. Bangerth, "A framework for the adaptive finite element solution of large inverse problems. I. Basic techniques," Tech. Rep. 04-39, Institute for Computational Engineering and Sciences (ICES), University of Texas at Austin, 2004.

[7] W. Bangerth, A. Joshi, and E. M. Sevick-Muraca, "Adaptive finite element methods for increased resolution in fluorescence optical tomography," *Progr. Biomed. Optics Imag.*, vol. 6, pp. 318–329, 2005.

[8] W. Bangerth, R. Hartmann, and G. Kanschat, `deal.II` *Differential Equations Analysis Library, Technical Reference*, 2006, `http://www.dealii.org/`.