September 1999

Efficient Solution of Boundary-Value Problems for Image Reconstruction via Sampling

Colin Fox, Geoff Nicholls, Mathias Palm
Mathematics Department, The University of Auckland

Abstract

Non-invasive imaging based on wave scattering remains a difficult problem in those cases where the forward map can only be adequately simulated by solving the appropriate partial-differential equation (PDE) subject to boundary conditions. We develop a method for solving these linear boundary-value problems (BVP) which is efficient and exact, trading off storage requirements against computation time. The method is based on using the present solution within the Woodbury formula for updating solutions given changes in the trial image, or state. Hence the method merges well with the Metropolis-Hastings algorithm using localized updates. The scaling of the method as a function of image size and measurement set size is given. We conclude that this method is considerably more efficient than earlier algorithms that we have used to demonstrate sampling for inverse problems in this class. We give examples of sampling for imaging electrical conductivity from a simple synthetic data set. Full Bayesian inference is demonstrated with expectations calculated over the posterior for Potts type prior distributions.

I. Introduction

A. Imaging from Wave Scattering

Non-invasive imaging for medical diagnostics and geophysics often uses scattering of waves to probe the object under investigation. These techniques are non-invasive when the wave scattering is measured remotely from the object, and the object is irradiated with waves using a source remote to the object. A common example is simple x-ray imaging where a person is irradiated from one side using a point x-ray source and the partially-transmitted waves are measured using a photographic plate on the opposite side to the source. In this paper we are particularly interested in cases where the scattering within the object is strong, i.e., where the path that the waves take is significantly affected by propagation through the object. The example of x-ray imaging is not such a case since the rays propagate through tissue in essentially straight-line paths, largely independent of the particular tissue being imaged. This feature of x-ray imaging is related to the high energy, ionizing, nature of x-rays which as a side-effect causes damage to the tissue through which it propagates – or more correctly by which it is absorbed. Lower-energy, non-ionizing, radiation is therefore a potentially safer source of waves for the purpose of imaging for medical diagnostics. However, the scattering process for such waves is strong and the path that each ray takes is implicitly determined by the propagation medium.

This makes the imaging problem far more difficult as the path of propagation as well as the medium must be reconstructed in the imaging step. The implicit dependence of the wave path makes the imaging problem non-linear in contrast to imaging from weak scattering where the forward map is linear.[1]

Three examples of imaging from wave scattering, corresponding to three different types of energy being propagated, are given in the following table.

quantity being imaged	governing PDE	PDE classification
electrical conductivity	$\boldsymbol{\nabla}\cdot(\sigma\nabla\phi) = s$	elliptic
acoustic impedance	$\boldsymbol{\nabla}\cdot(\sigma\nabla p) = \frac{\sigma}{c^2}\ddot{p}$	hyperbolic
thermal conductivity	$\boldsymbol{\nabla}\cdot(\sigma\nabla u) = \dot{u}$	parabolic

In each case the quantity being imaged is denoted by σ and appears as the spatially-varying coefficient in the PDE governing the propagation of energy, or 'waves'. The measurements are made of the boundary values of electrical potential ϕ,[2,3] ultra-sound pressure p,[4] and temperature u, respectively in the three cases. In each case the measurement process is simulated by solving the PDE subject to boundary conditions that correspond to the wave irradiation. The simplest case is imaging electrical conductivity, though it shares with the others the primary computational difficulty that the space part of the PDE is the elliptic, Laplacian-like, operator which has no characteristics and hence must be solved using an 'implicit' method. For the remainder of this paper we concentrate on imaging electrical conductivity and show how that implicit calculation

may be solved efficiently within an inferential computation using Markov chain Monte Carlo. Formally, the other imaging problems, such as imaging from ultrasound backscatter, can be shown to be equivalent to a sequence of problems of this type.[5]

B. Conductivity Imaging

We now focus on the canonical problem for imaging from scattered electromagnetic waves where the radiation has a sufficiently low frequency that the propagation is quasi-static. Then the measurables are the surface values of the stationary current and the constant potential. In this long-skin-depth limit, the quantity imaged is the spatially varying conductivity of the object, which we denote $\sigma(x)$. The waves now have (infinitely) long wavelength and the ray picture of propagation is not valid. Indeed, the current does not follow any single path but rather a distribution of paths.

1. Idealized Measurement Procedure

Instrumentation for conductivity imaging typically uses a number of electrodes placed on the surface of the object with the ability to assert currents into the electrodes and then measure the resulting voltages at the electrodes. We will denote the electrode positions by $x_1, x_2, \cdots, x_E$, there being E electrodes. For our purposes we consider point electrodes and assume that electrodes for current and voltage are the same. In practical instrumentation, electrodes are often finite sized and differ for current and voltages. These complications may be taken into account and do not fundamentally affect our results here.

It is usual to describe the electrical field within the object using the scalar potential $\phi(x)$ from which the vector current is given by $\rho(x) = -\sigma(x)\nabla\phi(x)$. The vector of currents $j = \left(\int_{x_1} \rho \cdot n, \int_{x_2} \rho \cdot n, \cdots, \int_{x_E} \rho \cdot n\right)^T$ is asserted at the electrodes and the voltages $v = (\phi(x_1), \phi(x_2), \cdots, \phi(x_E))^T$ are measured. Here $\int_{x_i} \rho \cdot n$ denotes the integral of the component of the current density in the direction of the unit normal to the boundary, n, over the i^{th} electrode, and hence is the total current passing through the i^{th} electrode. If L different current vectors are used, the total set of measurements is the collection of current-voltage pairs $\left\{j^l, v^l\right\}_{l=1}^{L}$ in which j^l is the l^{th} asserted vector of currents and v^l is the corresponding vector of measured voltages. For brevity, we use the shorthand $\{j, v\}$ for $\left\{j^l, v^l\right\}_{l=1}^{L}$.

2. Forward Map

Let Ω denoting the region occupied by the object. The relationship between the unknown image $\sigma(x)$ and the measurement procedure is given by the

Neumann boundary value problem (BVP):

$$\begin{aligned} \boldsymbol{\nabla} \cdot \sigma(x) \nabla \phi(x) &= s(x) && x \in \Omega \\ \sigma(x) \frac{\partial \phi(x)}{\partial n(x)} &= j(x) && x \in \partial\Omega \end{aligned} \tag{1}$$

along with the definition of the potential reference. Here n denotes the unit outward normal vector on the boundary of the region $\partial\Omega$. The function s gives the density of current *sinks* within the region as is zero for non-invasive measurements. In typical applications σ can be bounded above and away from zero, i.e., $\exists \sigma_{\min}, \sigma_{\max}$ such that $0 < \sigma_{\min} \leq \sigma \leq \sigma_{\max} < \infty$, though in all cases $\sigma > 0$.

The forward map, which goes from conductivity images to the measured potentials on the boundary, requires solution of this BVP with $s \equiv 0$. Note that the BVP is linear given σ, but the relation between σ and the measurements is not linear. Later we will see that if σ is compiled into a certain matrix, then the measurement procedure corresponds to measuring certain elements of the inverse matrix.

3. Green's Function

Solutions of the boundary-value problem 1 can be written in terms of the Neumann Green's function $g(x|\xi)$ which, for each $\xi \in \Omega$, is the solution of the auxiliary problem

$$\begin{aligned} \boldsymbol{\nabla} \cdot \sigma(x) \nabla g(x|\xi) &= \delta(x - \xi) && \forall x \in \Omega \\ \sigma(x) \frac{\partial g(x|\xi)}{\partial n(x)} &= \frac{1}{|\partial\Omega|} && \forall x \in \partial\Omega \ . \\ \int_{\partial\Omega} g(x|\xi) \, dl(x) &= 0 \end{aligned} \tag{2}$$

Hence, $g(x|\xi)$ gives the potential at x due to a unit current sink at the point ξ. Here $|\partial\Omega|$ denotes the length of the boundary and dl is the length element on the boundary. The first boundary condition is the statement that the current injected at ξ is removed uniformly around the boundary, while the second boundary condition effectively states that potentials are measured with respect to the mean potential on the boundary. The two boundary conditions are chosen to ensure that the boundary value problem is self-adjoint[6] and hence the Green's function is symmetric, i.e.,

$$g(x|\xi) = g(\xi|x) \, .$$

Pairs of boundary conditions can be found that correspond to any other reference for the potential. For computational efficiency it is important that the choice be made so that the Green's function is symmetric.

4. Inverse Problem

For non-invasive measurements $s(x) \equiv 0$, and solutions to 1 may then be written in terms of the Neumann Green's function as

$$\phi(x) = \int_{\partial\Omega} g(x|\xi)\, j(\xi)\; dl(\xi) \tag{3}$$

with ϕ being measured with respect to its mean value on the boundary. We denote the resulting map from boundary current to potential by $\Gamma_\sigma : j \to \phi$.

It is clear from 3 that Γ_σ is a linear functional defined by its kernel $g(x|\xi)$. Since we are free to choose the currents $j(\xi)$, complete measurements therefore correspond to measuring $g(x|\xi)$ for $x, \xi \in \partial\Omega$. Practical measurements made at electrodes are complete when the values $g(x|\xi)$ for $x, \xi \in \{x_1, x_2, \cdots, x_E\}$ are effectively measured. These values of the Green's function define the mapping between boundary currents and boundary voltages and, hence, also define the noise-free data. The imaging, or inverse, problem is to determine σ from these measurements of the Green's function. Since the Green's function depends implicitly on the unknown conductivity σ and, because that dependence is not linear, the inverse problem is non-linear.

5. Bayesian Formulation

The physical quantities in the model for the forward map are the conductivity σ, the potential ϕ, and the current density ρ – all being functions throughout the region. The measured data consists of the current-voltage vector pairs $\left\{j^l, v^l\right\}_{l=1}^{L}$. Because the measured voltages and currents are subject to measurement errors, we distinguish between the measured boundary values and the field quantities involved in the forward map. We use upper-case to denote random variables and lower case to denote trial values or realizations.

	random variable	realization
current in Ω	R	ρ
potential in Ω	Φ	ϕ
voltage at electrodes	V	v
current at electrodes	J	j
conductivity	Σ	σ

We add a superscript l to denote values corresponding to the l^{th} current vector. The set of realizations of potentials for all L measurement pairs is denoted $\{\phi\}$ as shorthand for $\left\{\phi^l\right\}_{l=1}^{L}$, i.e., that ϕ ranges over the vectors with index $l = 1, 2, \cdots L$ is taken implicitly. We use analogous shorthand for all other field quantities except the conductivity σ which does not vary with l.

Even though all quantities take continuum values, we use notation for discrete variables to avoid technical obfuscation. Then the joint posterior distribution for the field quantities given measurements may be expressed using Bayes'

equality as

$$\begin{aligned}
&\Pr\left\{\Sigma=\sigma,\{\Phi\}=\{\phi\},\{R\}=\{\rho\}\,|\,\{J,V\}=\{j,v\}\right\}\\
&\quad=\quad \Pr\left\{\,\{J,V\}=\{j,v\}\,|\Sigma=\sigma,\{\Phi\}=\{\phi\},\{R\}=\{\rho\}\right\}\\
&\qquad\times\Pr\left\{\Sigma=\sigma,\{\Phi\}=\{\phi\},\{R\}=\{\rho\}\right\}.
\end{aligned}$$

The three field quantities parametrizing the Likelihood and prior can not be set independently since, given the conductivity and potential, the current density is determined by Ohm's law that $\rho=-\sigma\nabla\phi$. Further, the internal potential is determined by the Neumann BVP 1 given the boundary conditions which are the current density crossing the boundary and one value of the potential. We have chosen the latter by setting the mean potential on the boundary to be zero always. Hence, the field parameters that we are free to set are the conductivity σ and the current crossing the boundary $(\rho\cdot n)\,|_{\partial\Omega}$. Given these quantities, then $\phi=\Gamma_\sigma\left((\rho\cdot n)\,|_{\partial\Omega}\right)$ and $\rho=-\sigma\nabla\phi$. We will write this parametrization as σ,ρ, understanding that we are actually only free to specify the current crossing the boundary and not all of ρ. The joint prior distribution for the field quantities is thus

$$\Pr\left\{\Sigma=\sigma,\{\Phi\}=\{\phi\},\{R\}=\{\rho\}\right\}=\Pr\left\{\Sigma=\sigma,\{R\}=\{\rho\}\right\}.$$

Since we are only interested in reconstructions of the conductivity, we choose to stipulate the functional form of the prior on the conductivity only, i.e., specify $\Pr\{\Sigma=\sigma\}$ only, – usually a Markov random field – with the prior having no dependence on R^l since we do not seek to make images of the currents.

The Likelihood function for the field quantities, once the measurements have been made, can be simplified using conditional independence between current and voltage measurements to give

$$\begin{aligned}
\mathcal{L}\left(\sigma,\{\phi\},\{\rho\}\right) &= \Pr\left\{\,\{J,V\}=\{j,v\}\,|\Sigma=\sigma,\{\Phi\}=\{\phi\},\{R\}=\{\rho\}\right\}\\
&= \Pr\left\{\,\{V\}=\{v\}\,|\,\{\Phi\}=\{\Gamma_\sigma\left(\rho\right)\}\right\}\Pr\left\{\,\{J\}=\{j\}\,|\,\{R\}=\{\rho\}\right\}
\end{aligned}$$

since j^l is the measured values of $\rho^l\cdot n$ at the electrodes and v^l is ϕ^l measured at the electrodes. When the errors between measurement sets are independent, we may further simplify the Likelihood as

$$\mathcal{L}\left(\sigma,\{\phi\},\{\rho\}\right)=\Pi_{l=1}^{L}\Pr\left\{V^l=v^l|\Phi^l=\Gamma_\sigma\left(\rho^l\right)\right\}\Pr\left\{J^l=j^l|R^l=\rho^l\right\}.$$

The two densities on the right-hand side depend on the distribution of errors in the instrumentation. For example, if the current measurement has normally-distributed additive errors then

$$\Pr\left\{J=j|R=\rho\right\}\sim\mathrm{N}\left(\left(\int_{x_1}\rho\cdot n,\int_{x_2}\rho\cdot n,\cdots,\int_{x_E}\rho\cdot n\right)^T,s_\rho^2\right)$$

where N denotes a multi-variate normal distribution with the mean being the vector of internal currents that flows into each electrode and with some variance (or covariance matrix) s_ρ^2. If the errors in measuring potential were also normally distributed then

$$\Pr\left\{V = v|\Phi = \Gamma_\sigma(\rho)\right\} \sim \mathrm{N}\left(\left(\Gamma_\sigma(\rho)(x_1), \Gamma_\sigma(\rho)(x_2), \cdots, \Gamma_\sigma(\rho)(x_k)\right)^T, s_\phi^2\right)$$

where now the mean is the vector of potentials at electrodes given by solving the BVP with boundary current $\rho \cdot n$, and the covariance matrix is s_ϕ^2. In both these distributions we have dropped the l superscript, assuming that the functional form does not depend on l.

The majority of the computational work required for calculating the Likelihood occurs in calculating the terms $\left\{\Gamma_\sigma\left(\rho^l\right)\right\}_{l=1}^L$ since this requires L solutions of the BVP 1.

6. Markov Chain Monte Carlo

Because the forward map depends non-linearly on σ, the posterior distribution is skewed and modes of the posterior do not necessarily make optimal reconstructions. For these high-dimensional problems there is no reason that the mode be representative of the bulk of probability which is dominated by metric factors rather than the value of the posterior. Instead, we seek to calculate expectations over the posterior using Markov chain Monte Carlo (MCMC) to draw samples from the posterior. We have previously demonstrated[7,8] calculation of mean images, along with sample variances, etc., from synthetic data for this inverse problem via a MCMC algorithm with an approximate method for solving the BVP.

In MCMC, the forward problem is repeatedly simulated and used to generate a Markov chain $\{(\Sigma, \{R\})_t\}_{t=0}^\infty$ of random variables taking values from allowable conductivity distributions and boundary currents, with a t-step distribution $\Pr\left(\Sigma_t = \sigma|\Sigma_0 = \sigma^{(0)}\right)$ which tends to the marginal posterior distribution, $\Pr\{\Sigma = \sigma|\{J, V\} = \{j, v\}\}$, as t tends to infinity.

In our implementations we use Metropolis-Hastings dynamics defined by a separate reversible transition probability for each of several *move* types. We will only give a very brief outline of the method here; interested readers can find details in our earlier papers.[7,8] The moves are chosen so that the process is irreducible on allowable states and behaves ergodically within useful time scales, but the moves are otherwise arbitrary. Evolution of the boundary current requires straightforward sampling from a normal distribution. Efficient evolution of the conductivity distribution is achieved by moves which pick a pixel at random or which concentrate on pixels near an 'update edge' which is an internal boundary between different conductivity values. Briefly the moves are: flip the conductivity at a randomly chosen pixel, flip the conductivity at a pixel near an update edge, swap conductivities at a pair of nearby pixels across an update edge, flip conductivities at a pair of nearby pixels across an update edge. The

first move guarantees irreducibility of the chain, and is dominant during burn-in, while the other moves are necessary for efficient mixing in equilibrium. A key feature of all these moves is that changes to the conductivity are localized, occurring at one or two nearby pixels.

The MCMC algorithm proceeds as follows: Given a state $(\Sigma, \{R\})_t = (\sigma, \{\rho\})$, and the likelihood of the state $L(\sigma, \{\rho\})$, the next state $(\Sigma, \{R\})_{t+1}$ is generated by:

Step I Select a move p with probability ζ_p. Generate a candidate new state $(\sigma', \{\rho'\})$ according to the selected move.

Step II Calculate the likelihood of the candidate state $L((\sigma', \{\rho'\}))$.

Step III Compute the acceptance probability for the transition, $(\sigma, \{\rho\}) \rightarrow (\sigma', \{\rho'\})$.

Step IV Accept or reject the candidate state $(\sigma', \{\rho'\})$:

IV.i If $(\sigma', \{\rho'\})$ is rejected, set $(\Sigma, \{R\})_{t+1} = (\sigma, \{\rho\})$.

IV.ii If $(\sigma', \{\rho'\})$ is accepted, set $(\Sigma, \{R\})_{t+1} = (\sigma', \{\rho'\})$.

Steps I to IV are repeated to generate a chain of any desired length. Expectations over the posterior are then calculated using Monte Carlo integration, the integral being estimated by a normalized sum of the desired quantity over the Markov chain.

C. Algorithms

The most frequently required calculation within the MCMC, that depends on the forward map, is the evaluation of the likelihood of the candidate state in step II . We seek an efficient algorithm by ensuring that this calculation may be performed as cheaply as possible. In a direct implementation of the forward map, calculation of the likelihood requires solution of the BVP, once for each measurement current vector. For an image with M pixels, each solution requires $O\left(M^2\right)$ calculations. Evaluating the likelihood requires a solution for each measurement vector, making $O\left(LM^2\right)$ calculations in all. Since typically many thousands, or millions, of states are required for accurate sample expectations, this direct implementation is prohibitively expensive. In contrast, the algorithm that we propose uses stored information about the present state in a way that allows calculation of the Likelihood in $O\left(LE^2\right)$ operations for E electrodes. This is a huge improvement over direct implementation. When the candidate state is rejected, no further calculations are required for this time step. However, if the candidate state is accepted, we presently require further $O\left(M^2\right)$ calculations to update the stored quantities. While this step is computationally expensive, it is still cheaper than each Likelihood calculation in the direct implementation.

There are many possible samplers, each defined by the choice and probability of the moves. We have found a combination of sampler and method for solving

the BVP that gives an efficient sampling algorithm. Our present scheme is based on restricting the moves to local changes in the conductivity state, either changing the conductivity at a single pixel or at pairs of nearby pixels across an update edge. By writing solutions of the BVP as an integral transform with kernel being the Neumann Green's functions, we can identify a minimal set of quantities that need to be stored so that the potential at the electrodes, and hence the likelihood, may be calculated with relatively few operations. This calculation corresponds to application of the Woodbury formula, which gives the updated inverse of a matrix for a low-rank change, in a discrete formulation of the forward problem. One suitable discrete formulation is a finite-element discretization of the problem, which we outline in the next sections along with the application of the Woodbury formula. The same calculation may be posed in terms of the differential operators directly, though that route is technically more demanding.

In our previous sampling algorithms for conductance imaging[7,8] we used a *linearization* of the forward map $\sigma \to \{j, v\}$ to allow a cheap *approximation* to the likelihood for a candidate state. That approximation had the consequence that, for our particular simulations, about 14% of acceptances were falsely made and hence detailed balance could not have been achieved. However, this did not seem to affect the reconstructions significantly. We now understand that actually the false rejections (candidate states that were rejected on the basis of the approximate likelihood but would have been accepted on the basis of the exact likelihood) have a greater effect on efficiency of the sampler. That case occurs for moves such as pixel swapping that do not greatly change the likelihood. The linearization for such moves has a good absolute error but poor relative error; the latter being important when Likelihood changes are small. And since moves that keep the Likelihood nearly unchanged are important for efficient mixing in equilibrium, the linear approximation has the effect of reducing mixing and hence efficiency. For moves involving changes in mean conductivity, i.e. flipping single or pairs of pixels, the linearization remains an efficient and accurate method for 'solving' the BVP, though we no longer use it.

Our main result here is to show that the linear approximation can be replaced with an exact calculation requiring the same order of operations. The exact calculation requires extra storage, with the consequence that more updating of stored values is required on an acceptance. However, the same scheme that allows us to perform cheap calculations of the likelihood also allows us to update the stored values more cheaply than in previous algorithms.

II. Finite Element Formulation

A. Variational Form

In this section we give a discretization of the forward map based on a finite-element method (FEM) in which pixels are taken as elements. The method for efficient calculation of the likelihood does not depend on this particular

discretization and we could equally have used a traditional triangle-based FEM or finite-difference discretization. The same pixel discretization is used for both the conductivity distribution as well as the field quantities. This feature is not necessary, nor desirable when higher-level prior models are used, but it allows us to display the algorithm simply.

Finite element methods utilize the equivalence between the BVP 1 and the variational statement that

$$\phi(x) = \arg\min \int_{\Omega} \left(\sigma(x) |\nabla \phi(x)|^2 + 2s(x)\phi(x) \right) dx - \int_{\partial\Omega} j(x)\phi(x)\ dl(x) \tag{4}$$

subject to the constraint

$$\int_{\partial\Omega} \phi(x)\, dl(x) = 0.$$

The Neumann boundary condition in the BVP 1, of given boundary current, is included here as the 'natural boundary condition' for this variational form. However the second boundary condition, corresponding to our using the mean over the boundary as the reference of potential, necessarily appears as a constraint.

The particular case of solving for a Green's function, in BVP 2, gives the reduced form

$$g(x|\xi) = \arg\min \int_{\Omega} \sigma(x) |\nabla g(x|\xi)|^2 dx + 2g(\xi|\xi)$$

subject to

$$\int_{\partial\Omega} g(x|\xi)\ dl(x) = 0.$$

Note that all derivatives and integrals are with respect to the first, x, variable.

B. Pixel FEM

We discretize Ω using a square mesh. The conductivity within each square, or 'pixel', is taken to be constant and the nodal values of the potential, at the vertices of each square, are used to define the potential. Denote the region occupied by the i^{th} pixel by Ω_i. Then $\Omega = \cup_{i=1}^{M} \Omega_i$ when there are M pixels in all. The integrals required in equation 4 can then be written as a sum of integrals over each pixel. The integral over each pixel is conventionally performed by transforming coordinates onto the master square $\Omega_{\text{m}} = [0,1] \times [0,1]$ in the (x, y) plane that has constant conductivity $\sigma = 1$. Given the potential $\phi(x, y)$ at the four corners $\phi(0,0) = \phi_{00}$, $\phi(1,0) = \phi_{10}$, $\phi(0,1) = \phi_{01}$, $\phi(1,1) = \phi_{11}$, we take the potential within the pixel to be the bilinear interpolation[9] of the 'local' nodal values

$$\phi(x,y) = (1-y)\left((1-x)\phi_{00} + x\phi_{01}\right) + y\left((1-x)\phi_{10} + x\phi_{11}\right).$$

The energy integral over Ω_{m} is then

$$\int_0^1 \int_0^1 \left(\sigma \left|\nabla\phi\right|^2\right) dx\, dy = \left(\phi_{00}, \phi_{01}, \phi_{10}, \phi_{11}\right)^{\mathrm{T}} K_{\mathrm{l}} \left(\phi_{00}, \phi_{01}, \phi_{10}, \phi_{11}\right)$$

where

$$K_{\mathrm{l}} = \frac{1}{6} \begin{pmatrix} 4 & -1 & -1 & -2 \\ -1 & 4 & -2 & -1 \\ -1 & -2 & 4 & -1 \\ -2 & -1 & -1 & 4 \end{pmatrix}$$

is the 'local stiffness matrix'.

The nodal values of potential can be arranged as a vector which we denote by ϕ, i.e., using the same symbol as the continuous counterpart. The position of a node within this vector is called the global nodal number. For each element Ω_i we need to know the mapping between global nodal numbers and the corresponding local nodal number, which in this case take the values $\{1, 2, 3, 4\}$. Let n_{ij} be the global node number of the node that is mapped to local node j when element i is transformed to the master element. Then, the $M \times M$ global stiffness matrix K can be assembled (by elements) by adding to the initially zero matrix the components $\sigma_i \left(K_{\mathrm{l}}\right)_{n_{ij} n_{ik}}$ for $j, k = 1, 2, 3, 4$ where i loops over the M elements. Note that we have not needed to include the Jacobian of the transformation[9] to the master square since this is a constant for all pixels. Thus, each pixel contributes 16 components at the intersections of the rows and columns with global nodal numbers on the element, the components are the local stiffness matrix multiplied by the conductance. The global stiffness matrix K is therefore very sparse with at most 9 non-zero elements per row. Note that the constant vector is in the null space of K_{l}, and so the global stiffness matrix has a zero eigenvector in the direction of the constant vector. Taking $s = 0$, the discrete form of the variational statement is

$$\phi = \arg\min \phi^{\mathrm{T}} K \phi - j^{\mathrm{T}} \phi \qquad \text{subject to} \qquad \phi \cdot \chi_{\mathrm{b}} = 0 \tag{5}$$

where χ_{b} is a vector that is 1 for nodes on the boundary and is otherwise zero. The variational form for the discretized Green's function for current source at node k is then

$$g^k = \arg\min \left(g^k\right)^{\mathrm{T}} K g^k + 2 \left(g^k\right)^{\mathrm{T}} \delta^k \qquad \text{subject to} \qquad g^k \cdot \chi_{\mathrm{b}} = 0 \tag{6}$$

where δ^k is the vector which is 1 at node k and is otherwise zero.

III. Efficient BVP solution

The normal equations for the discrete quadratic forms 5 and 6 are difficult to manipulate since they require the inverse of the global stiffness matrix which is singular. However the minimum is well defined in the subspace defined by the constraint. To work around the difficulty we define a modified quadratic

form by projecting the potential onto the subspace defined by the constraint and adding a non-zero curvature orthogonal to the subspace. The potential is then given by an unconstrained optimization and, hence, the modified normal equations are non-singular. The non-singular form is then suitable for direct application of the Woodbury formula, which is the essence of the efficient BVP solution method.

A. Non-singular matrix form

Define the projection onto χ_{b} and the complementary projection[10]

$$P_{\mathrm{b}} = \frac{\chi_{\mathrm{b}}\chi_{\mathrm{b}}^{\mathrm{T}}}{\|\chi_{\mathrm{b}}\|^2}, \quad P_{\perp} = I - P_{\mathrm{b}},$$

respectively, where I is the identity matrix. Note that $P_{\perp}$ has the action of subtracting the mean over the boundary from nodes on the boundary, and is the identity for interior nodes. Let $A = (P_{\perp}KP_{\perp} + \gamma P_{\mathrm{b}})$. The term γP_{b} is a regularizing term to ensure that $P_{\mathrm{b}}g^k = 0$ with $\gamma > 0$ being otherwise arbitrary. Then g^k may be found by an unconstrained minimization of the modified quadratic form

$$g^k = \arg\min \left(g^k\right)^{\mathrm{T}} Ag^k + 2\left(g^k\right)^{\mathrm{T}} P_{\perp}\delta^k.$$

The Hessian matrix A, unlike K, is invertible and the normal equations then give

$$g^k = -A^{-1}P_{\perp}\delta^k. \tag{7}$$

Since $(P_{\perp})^2 = P_{\perp}$, $(P_{\mathrm{b}})^2 = P_{\mathrm{b}}$ and $P_{\perp}P_{\mathrm{b}} = 0$, the following results are straightforward to establish

$$P_{\mathrm{b}}g^k = 0, \qquad P_{\perp}g^k = g^k, \qquad g^k = \left(-P_{\perp}A^{-1}P_{\perp}\right)\delta^k.$$

From equation 7, it follows that g^k is the k^{th} column of $-A^{-1}P_{\perp}$ and, since $\left(-P_{\perp}A^{-1}P_{\perp}\right)$ is a symmetric matrix, g^k is also the k^{th} row. Hence we can write the kl component of $-A^{-1}P_{\perp}$ as

$$\left(-A^{-1}P_{\perp}\right)_{kl} = g_l^k = g_k^l.$$

As observed in section 4, the Green's function with source and field-point at electrodes defines the forward map. These are the components g_l^k for k and l being nodes corresponding to electrodes.

B. Application of the Woodbury Formula

Consider now a change to the conductivity, $\Delta\sigma$, occurring at a few pixels. Denote by $k_1, k_2, \cdots, k_R$ the global nodal numbers for the nodes on the conductivity change and let $C = \{k_1, k_2, \cdots, k_R\}$ be the set of all these indices. For

a single pixel change to the conductivity $R = 4$ indices, while for a two pixel change in conductivity $R \leq 8$. The corresponding change in the global stiffness matrix is the matrix ΔK which is just the global stiffness matrix for $\Delta\sigma$. That is, $\sigma + \Delta\sigma$ has the global stiffness matrix $K + \Delta K$, where K and ΔK are the global stiffness matrices for σ and $\Delta\sigma$, respectively. Since ΔK is non-zero only at the intersection of rows and columns in C, a single pixel change to the conductivity changes 16 components of K, while a two pixel change in conductivity changes at most 64 components of K.

The forward map, and consequently the Likelihood, for a given set of measurement currents is determined by the Green's function components g_l^k for k and l being nodes corresponding to electrodes. Since the Green's functions are the columns of $-A^{-1}P_\perp$ requiring the inverse of the modified global stiffness matrix, consequently the Likelihood can be calculated cheaply if the appropriate rows of the inverse can be easily calculated. The Woodbury formula gives an efficient way of updating the inverse matrix requiring a matrix inversion of the size of the number of nodes at which a change is made.

For an arbitrary $M \times M$ matrix A, and arbitrary $M \times R$ matrices U and W, the Woodbury formula[11] is

$$\left(A + UW^{\mathrm{T}}\right)^{-1} = A^{-1} - \left[A^{-1}U\left(I + W^{\mathrm{T}}A^{-1}U\right)^{-1}W^{\mathrm{T}}A^{-1}\right].$$

The matrix inversion required is of size $R \times R$ which is cheap when $R \ll M$.

The change to the global stiffness matrix can always be written as an outer product of two matrices $\Delta K = UW^{\mathrm{T}}$. For example, when the conductivity at the single pixel i is changed by $\Delta\sigma_i$, let U be the 4 non-zero columns of the $M \times M$ matrix that is the identity over the nodes on pixel i, and zero elsewhere, while W is the 4 non-zero columns of the global stiffness matrix for a conductivity $\Delta\sigma_i$ at pixel i and is zero elsewhere. Many choices of appropriate matrices can be made; one possibility that exploits the defective rank of the local stiffness matrix is to use the eigenvectors of that matrix to find $M \times 3$ matrices.

If the conductivity change occurs at pixels that do not have nodes on the boundary, then both U and W are zero for nodes on the boundary and it follows that

$$P_\perp U = U \qquad \text{and} \qquad P_\perp W = W. \tag{8}$$

The change to the modified Hessian matrix A then also equals UW^{T}. We restrict our attention to this case, where conductivity changes occur away from the boundary, since typically the conductivity near the boundary is well determined by measurements and evolution of the conductivity solely consists of changes away from the boundary. Using the relationships 8 to substitute for U and W in the Woodbury formula, and that $\left(P_\perp\right)^{\mathrm{T}} = P_\perp$, the change to the kl element of $-A^{-1}P_\perp$ is

$$\begin{aligned}
\Delta g_l^k &= -\left[A^{-1}P_\perp U\left(I + W^{\mathrm{T}}P_\perp A^{-1}P_\perp U\right)^{-1}W^{\mathrm{T}}P_\perp A^{-1}P_\perp\right]_{kl} \\
&= -\left(g^k\right)^{\mathrm{T}}U\left(I + W^{\mathrm{T}}P_\perp A^{-1}P_\perp U\right)^{-1}W^{\mathrm{T}}g^l.
\end{aligned}$$

Since U and W are non-zero only for the R rows in C, we can further reduce the number of calculations required to evaluate this expression. We use the (Matlab-like) notation $W_{C,:}$ to denote the rows of W with indices in C, giving the $R \times R$ matrix of non-zero entries of W. Similarly by $U_{C,:}$ denote the $R \times R$ matrix of non-zero elements of U. Also, let g_C^C denote the $R \times R$ matrix of components of $-A^{-1}P_\perp$ at the intersection of rows and columns in C, and g_C^k the R-vector of rows of g^k with indices in C. Then

$$\Delta g_l^k = -\left(g_C^k\right)^{\mathrm{T}} U_{C,:} \left(I - \left(W_{C,:}\right)^{\mathrm{T}} g_C^C U_{C,:}\right)^{-1} \left(W_{C,:}\right)^{\mathrm{T}} g_C^l. \tag{9}$$

Hence calculation of the Likelihood requires 4 $R \times R$ matrix multiplications, a single $R \times R$ inversion, and $2E^2$ dot products of length R to find the kernel of the transform in 3. Further calculations of order E^2 may be required, depending on the current vectors used. The cost may be reduced further by using the symmetry of all the matrices, though the saving is not substantial.

The calculation in 9 requires storage of various components of the Green's function. The terms $\left(g^k\right)^{\mathrm{T}}$ and g^l require the E Green's functions which have sources at electrodes, requiring $E \times M$ components stored. The term g_C^C requires the Green's function at each node in C when the source point is also in C. Given that the change may occur anywhere in the image, this term requires $16 \times M$ values stored for single pixel changes or $64 \times M$ for two pixel moves with a maximum gap of 2 pixels between pixels being changed. Again, symmetry may be used to approximately halve these storage requirements, though then indexing in the implementation becomes complicated.

1. Rejection step

When a candidate state is proposed in step I of the MCMC algorithm, the Likelihood of the candidate required in step II may be calculated using the scheme described in the previous section. If the candidate state is rejected in step IV, no further calculations are required. In particular, no adjustment to the stored Green's functions is required as the subsequent state is unchanged from the present state. Thus, rejection steps are very cheap computationally.

2. Acceptance step

When a candidate conductivity state is accepted the global stiffness matrix must be updated and the consequent changes in the stored Green's functions must be computed. Again the Woodbury formula in 9 can be used to efficiently recalculate these values. Using the symmetry of the Green's function $g_j^k = g_k^j$, changes to any component of A^{-1} may be computed given all values of the Green's function for sources at nodes on the pixels that were changed. That is, all values g_l^k $\forall l \in C$ and $\forall k = 1, 2, \cdots, M$. It is not practical to store all these components of the Green's function as then we would need to store all of the matrix A^{-1} which exceeds the typical storage capacities for all but

the smallest images. At present we can see no alternative to simply having to calculate these Green's functions following an acceptance. These R solutions of the BVP require $O\left(RM^2\right)$ calculations in all, and make an acceptance step computationally expensive.

IV. Numerical Experiments

We now give two examples of sampling conductivity images from the posterior distribution for simple synthetic data sets. The images are square with 16 point electrodes placed evenly around the boundary, i.e., with 4 electrodes per side. With this number of electrodes there are 120 independent measurements available[7] which exceeds the roughly 103 independent measurables available via the continuous forward map[12] for the noise levels we will use. Thus the measurement set is effectively a complete measurement of the mapping in 3.

The two examples have discrete-valued conductivity distributions: the first being a 24×24 pixel image of a two-valued conductivity, the second is a 25×25 pixel image of a three-valued conductivity. The same discretization is used for generating the synthetic data as the reconstruction. In both cases the number of pixels exceeds the number of measurables and so prior information is required. We use Potts prior distributions in both cases with the lumping constant set to 1.

These examples are not intended to be definitive uses of prior information or of reconstruction. Rather they are given to demonstrate that sampling is tractable for forward problems requiring solution of a BVP and with prior information that gives a discrete, discontinuous, image space.

Figure 1 shows results for the first example. In subfigures a, b, and c, black and white correspond to a conductivities of 3 and 4, respectively. Grey levels in subfigure c are proportional to the conductivity between these two extreme values. The synthetic data set has zero-mean Gaussian noise added with standard deviation of 0.003 times the r.m.s. measurements. This signal to noise ratio of 50dB is typical for good biological measurements. Subfigure a shows the true image. Subfigure b is a sample from the chain in the long term, with the chain having been started from a random state. The similarity of the sample and the true image indicates that the chain has reached the equilibrium, posterior, distribution. Subfigure c is the mean calculated over the entire chain, including burn-in, and it is clear that the mean provides a reasonable reconstruction of the true image. Subfigure d is the sample variance of the mean with black indicating large variance while white indicates zero variance. It is interesting to note that there is large variance around the edges of the two conductivity inclusions while the background conductivity and the pixels internal to the inclusions are well determined. The low level non-zero variance between the two inclusions is a remnant of the samples occurring during burn-in; this feature disappears in the limit of infinitely many samples since, as can be seen in subfigure b, samples in equilibrium have the correct background conductivity in this area.

Figure 2 shows results for a three-level example with conductivity levels of 1 (black) 2 (half grey) and 3 (white). Again subfigures a, b, and c, give the true image, an equilibrium sample, and the mean image, respectively. For these images the synthetic data has additive Gaussian noise with standard deviation of 0.01 times the r.m.s. measurement level – or a signal to noise ratio of 40dB. As in the previous example the sampler was started from a random image, and the similarity of the state in subfigure c and the true image shows that the chain has reached equilibrium. The variance of the mean is given in subfigure d with black indicating large variance and white indicating zero variance. Figure 3 shows a trace of the log-likelihood, up to an additive constant, and the log-prior. It is interesting to see that the likelihood (lower trace) reaches its equilibrium value very quickly, while the prior reaches equilibrium more slowly. This indicates that there are many states, particularly during burn-in, which have high likelihood but low prior values and hence do not make good reconstructions. This is a consequence of there being many more pixels in the image than are effectively constrained by the measurements at the noise level used. Hence we see that the prior is critical in forming a clean image, with few regions of constant conductivity.

V. Conclusions

We have laid out a formalism for the efficient simulation of the posterior distribution of the conductance imaging problem. In particular we have outlined an algorithm that ensures that the steps requiring solution of the BVP can be performed rapidly. We expect that the same computational structure will also lead to efficient algorithms for other inverse problems where the forward map requires solution of a PDE.

Experiments from synthetic data achieved ergodic behavior of the Markov chain in reasonable times, demonstrating that sampling from this class of problems is tractable and hence it is feasible to calculate expectations over the posterior distribution. Removal of the approximation inherent in using the linearized forward map allowed us to investigate wider ranges of measurement noise, and achieve better mixing in equilibrium, compared to our previous algorithms.

Whilst this is cause for satisfaction, it remains unclear that the algorithm we have given will be effective for real data. However, given the optimal way in which the Woodbury formula evaluates the likelihood, we expect that the algorithm we have given will play some part in a full resolution of the problem of image reconstruction from wave scattering data via Bayesian inference.

Acknowledgement 1 *The authors gratefully acknowledge financial support for this work from the Auckland University Research Committee and the Marsden Fund.*

References

[1]D. F. Jackson and K. Kouris, "The Role and Potential of New Imaging Methods,", chapter in *Imaging with Non-Ionizing Radiations*, Surrey University Press, pp 1–49 (1983).

[2]D. C. Barber, "A review of image reconstruction techniques for electrical impedance tomography,", *Med. Phys.* 16, 162–169 (1989).

[3]M. Cheney, D. Isaacson, and J. C. Newell, "Electrical impedance tomography", *SIAM Review* 41, 85–101 (1999).

[4]J. F. Greenleaf, "Computerized tomography with ultrasound", *Proceedings of the IEEE* 71, 330–337 (1983).

[5]C. Fox and G. K. Nicholls, "Physically-based likelihood for ultrasound imaging," in *Bayesian Inference for Inverse Problems*, ed. A. Mohammad-Djafari, *Proc. SPIE* 3459, 92–99 (1998).

[6]I. Stakgold, *Green's Functions and Boundary Value Problems*, John Wiley & Sons, Inc., second edition (1998).

[7]C. Fox and G. K. Nicholls, "Sampling conductivity images via MCMC," in *The Art and Science of Bayesian Image Analysis*, ed. K. Mardia, R. Ackroyd, and C. Gill, *Leeds Annual Statistics Research Workshop*, 91–100, University of Leeds (1997).

[8]G. K. Nicholls and C. Fox, "Prior modelling and posterior sampling in impedance imaging," in *Bayesian Inference for Inverse Problems*, ed. A. Mohammad-Djafari, *Proc. SPIE* 3459, 116–127 (1998).

[9]O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems*, Academic Press, Inc., Orlando, Florida, (1984).

[10]G. Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Wellesley, MA (1993).

[11]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, second edition (1992).

[12]C. Fox, *Conductance Imaging*, PhD thesis, University of Cambridge (1989).

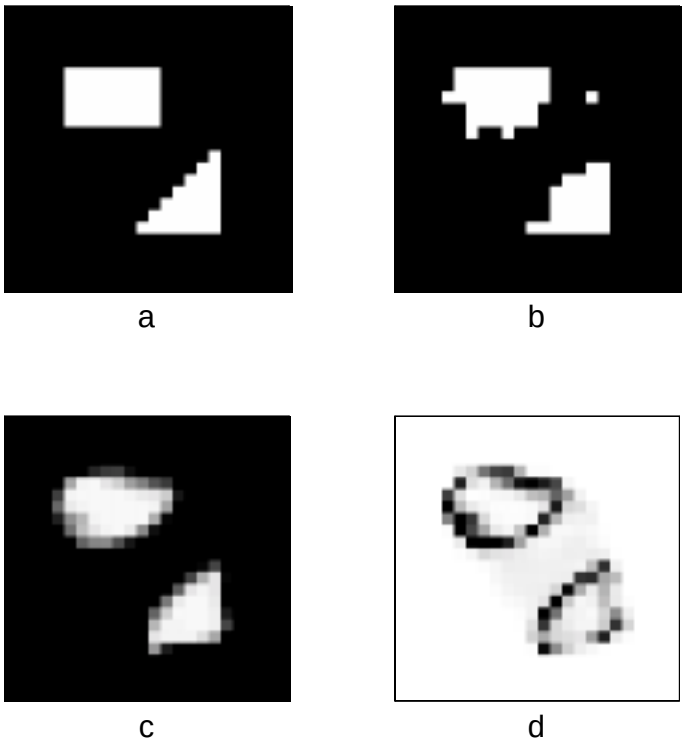

Figure 1: Numerical results for reconstruction of a two-level conductivity distribution with conductivities 3 (background) and 4 (inclusions). Subfigure a is the true image, subfigure b is an example of a sample from the posterior in equilibrium, subfigure c is the mean image over the entire chain, subfigure d is the sample variance of the mean.

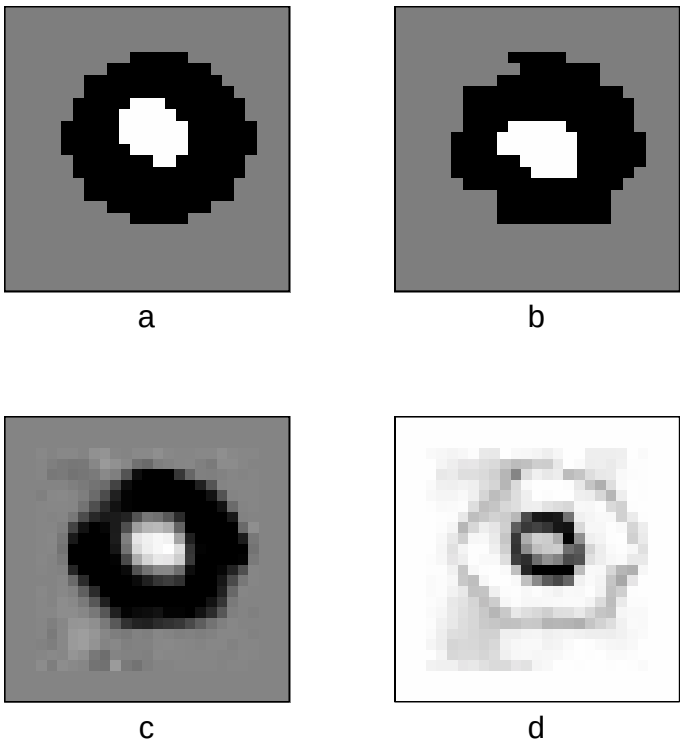

Figure 2: Numerical results for reconstruction of a three-level conductivity distribution with conductivities 1 (black), 2 (grey), and 3 (white). Subfigure a is the true image, subfigure b is a sample from the posterior in equilibrium, figure c is the mean image over the entire chain, figure d is the sample variance of the mean.

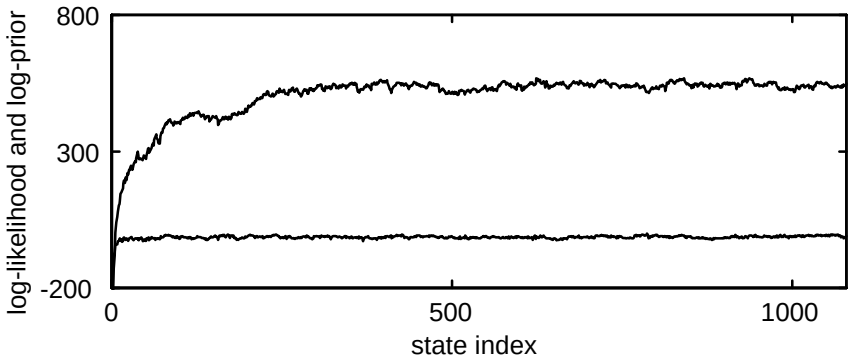

Figure 3: Trace of log-likelihood (lower trace) and log-prior (upper trace) for three-level conductivity example. Vertical scales are the same, though log-likelihood is shifted by an additive constant.