Efficient, exact PDE solutions for MCMC

Colin Fox, Mathias Palm, Geoff Nicholls

Mathematics Department, The University of Auckland, Auckland, New Zealand

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. We develop a method for solving linear PDEs which is efficient and exact, trading off computation time against storage requirements. The method is based on using the present solution within the Woodbury formula for updating solutions away from changes in the trial image, or state. Hence the method merges well with typical Metropolis-Hastings algorithms 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.

Keywords: Markov chain Monte Carlo, Impedance tomography, Bayesian, Partial differential equation, Inverse problem

1. INTRODUCTION

1.1. Imaging from wave scattering

Non-invasive imaging as used in medical diagnostics and geophysics often uses scattering of waves to probe the object under investigation. These techniques are non-invasive because the wave scattering is measured remotely from the object, and often 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-absorbed 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.

1.2. Conductance Imaging

A canonical problem for imaging from strongly scattered waves is where the radiation has a sufficiently low frequency that the propagation is quasi-static. Then the measurables are the stationary current and the d.c. potential. In this long-skin-depth limit, the quantity imaged is the spatially varying conductance 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. We will focus on the problem of reconstructing a spatially varying electrical conductance for the remainder of this paper – though other imaging problems such as imaging from ultrasound backscatter, can be shown to be equivalent to a sequence of problems of this type.[1]

CF: e-mail: fox@math.auckland.ac.nz, MP: e-mail: palm@scitec.auckland.ac.nz, GKN: e-mail: nicholls@math.auckland.ac.nz
Further info at http://www.math.auckland.ac.nz/~nicholls/linkfiles/papers.html

1.2.1. Measurement procedure

Instrumentation for conductance 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.

Measurements are made by asserting vectors of currents at the electrodes $j = (j(x_1), j(x_2), \cdots, j(x_E))^T$ and measuring the voltages $v = (\phi(x_1), \phi(x_2), \cdots, \phi(x_E))^T$. If N current vectors are used, the set of measurements is the collection of current-voltage pairs $\{j^n, v^n\}_{n=1}^N$.

1.2.2. Forward map

It is convenient to describe the electrical field within the object using the scalar potential $\phi(x)$ from which the vector current is given by $-\sigma(x)\nabla\phi(x)$. The combined statement of Ohm's law and of conservation of current gives the differential equation $\nabla \cdot \sigma(x)\nabla\phi(x) = s(x)$ that relates the potential and conductivity being imaged. The function s gives the density of current *sinks* within the region. 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 surface of the object is usually insulated away from the electrodes, so that the current crossing the boundary $j(x)$ equals $j(x_i)$ at the i^{th} electrode and is otherwise zero.

Let Ω denoting the region occupied by the object. The relationship between the unknown image $\sigma(x)$ and the measurement procedure as the Neuman boundary value problem (BVP):

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

with the extra requirement that the potential be specified with respect to some value. Note that n denotes the unit outward normal vector on the boundary of the region $\partial\Omega$. The forward map, which goes from conductivity images to the measured potentials on the boundary, requires solution of the partial-differential equation with $s \equiv 0$ subject to the specified boundary conditions. Note that the boundary-value problem we solve is linear, but the relation between $\sigma$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.

1.2.3. Green's functions

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

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

which gives the potential due to a unit current source at position 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. The two boundary conditions are chosen to ensure that the boundary value problem is self-adjoint 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 our purposes it is important that the choice be made so that the Green's function is symmetric.

Solutions to 1 may then be written in terms of this Green's function as

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

with ϕ being measured with respect to its mean value on the boundary. Non-invasive measurements necessarily have $s \equiv 0$ and 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 linear. The inverse problem is to determine σ from measurements of the mapping Γ_σ measured on the boundary. The kernel of the integral transform depends implicitly on the unknown conductivity σ and, because that dependence in not linear, the inverse problem is non-linear.

1.2.4. 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 $\{j^n, v^n\}_{n=1}^N$. 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. In general we use upper-case to denote random variables and lower case to denote trial values or realizations.

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

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

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

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

The three field quantities can not be set independently since, given the conductivity and potential, the current density equals $R = -\Sigma\nabla\Phi$. Further, the internal potential is determined by the BVP given the boundary conditions which are the current density at 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 $\rho|_{\partial\Omega}$. Given these quantities, then $\phi = \Gamma_\sigma\left(\rho|_{\partial\Omega}\right)$ and $\rho = -\sigma\nabla\phi$. We will write this parametrization as Σ, R with the fact that we are actually only free to specify the current on the boundary being understood. The joint prior distribution for the field quantities is thus

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

Since we are only interested in reconstructions of the conductivity, we choose to stipulate the functional form of $\Pr\{\Sigma = \sigma\}$ only – usually a Markov random field – with the prior having no dependence on R^n 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

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

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

$$L\left(\sigma, \{\phi^n\}, \{\rho^n\}\right) = \Pi_{n=1}^N \Pr\left\{V^n = v^n | \Phi^n = \Gamma_\sigma\left(\rho^n|_{\partial\Omega}\right)\right\} \Pr\left\{J^n = j^n | R^n = \rho^n\right\}.$$

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

$$\Pr\left\{J = j|R = \rho\right\} \sim \mathrm{N}\left(\left(\rho\left(\mathrm{x}_1\right), \rho\left(\mathrm{x}_2\right), \cdots, \rho\left(\mathrm{x}_\mathrm{k}\right)\right)^\mathrm{T}, \mathrm{s}_\rho^2\right)$$

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

$$\Pr\left\{V = v|\Phi = \Gamma_\sigma\left(\rho|_{\partial\Omega}\right)\right\} \sim \mathrm{N}\left(\left(\Gamma_\sigma\left(\rho|_{\partial\Omega}\right)\left(\mathrm{x}_1\right), \Gamma_\sigma\left(\rho|_{\partial\Omega}\right)\left(\mathrm{x}_2\right), \cdots, \Gamma_\sigma\left(\rho|_{\partial\Omega}\right)\left(\mathrm{x}_\mathrm{k}\right)\right)^\mathrm{T}, \mathrm{s}_\phi^2\right)$$

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

1.2.5. 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 make good 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 means over the posterior using MCMC to draw samples from the posterior. We have previously demonstrated[2,3] calculation of mean images, along with sample variances, etc, from synthetic data for this inverse problem via a MCMC algorithm.

In MCMC, the forward problem is 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, $\Pr\left\{\Sigma = \sigma|\left\{J^n, V^n\right\} = \left\{j^n, v^n\right\}\right\}$, as t tends to infinity. In our implementations we use the Metropolis-Hastings method 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.[2,3] The moves are chosen so that the process is irreducible on allowable conductivity distributions 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 and so moves for the boundary current will not be considered here.

The MCMC algorithm is as follows: Given a state $(\Sigma, R)_t = (\sigma, \rho)$, and the likelihood of the state $L(\sigma, \{\rho^n\})$, 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 $(\sigma', \{\rho'^n\})$ according to the selected move.

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

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

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

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

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

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 generated.

1.3. Algorithms

The most frequently required calculation within the MCMC, that depends on the forward map, is the is 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. Since evaluating the likelihood requires a solution for each measurement vector, the step would require $O(N)\,O\left(M^2\right)$ 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 $O(E)$ 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, further $O\left(M^2\right)$ calculations are required to update the stored quantities. While this step is computationally expensive, it is still considerably cheaper than each Likelihood calculation in the direct implementation.

Since there are many possible samplers, each defined by the choice and probability of the moves, we seek 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 Neuman 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 – for finding the updated inverse of a matrix given 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 section along with the application of the Woodbury formula. However, 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 we used a *linearization* of the forward map $\sigma \to \{j^n, v^n\}$ 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 think that in fact 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) are more serious. That case tends to occur for moves such as pixel swapping that do not greatly change the likelihood. The linearization for such moves seems to have 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, it appears that the linear approximation has the effect of reducing mixing and hence efficiency.

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.

2. FINITE ELEMENT FORMULATION

In this section we give a discretization of the forward map based on a finite-element method 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 discretizations are based on the equivalence between the BVP 1 and the variational statement that

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

subject to the constraint

$$\int_{\partial\Omega} \phi(x)\,d\,l(x) = 0$$

corresponding to our using the mean over the boundary as the reference of potential. In the particular case of solving for a Green's function, since $j = 1/|\partial\Omega|$, we have 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.

2.1. Pixel FEM for EIT

We discretize Ω using a square mesh; The conductivity within each square, often referred to as a 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 over each pixel can be 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 σ. 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 of the nodal values

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

The energy integral over Ω_{M} is then

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

where

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

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. It is convenient to arrange the vector so that the values for nodes at electrodes appear as the first K components. 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 $\sigma_i K^l_{n_{ij} n_{ik}}$ for $j, k = 1, 2, 3, 4$ where i loops over the M elements. 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.

Using the global stiffness matrix, the variational form for the discretized Green's function for current source at node k is then

$$g^k = \arg\min g^{k\text{T}} K g^k + 2g^k_k$$

where g^k_k is the k^{th} component of g^k, i.e., the potential at the point of current injection and can be written as $\left(g^k\right)^{\text{T}} \delta^k$ where δ^k is the vector which is 1 at node k and is otherwise zero. The minimization is subject to the constraint that the mean potential on the boundary is zero, i.e.,

$$g^k \cdot \chi_{\text{b}} = 0$$

where χ_{b} is a vector that is 1 for nodes on the boundary and is otherwise zero.

2.1.1. Non-singular matrix form

Define the projection onto χ_{b} and the complementary projection

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

respectively. 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. Then g^k may be found by an unconstrained minimization of the modified quadratic form

$$Q = g^{k\mathrm{T}} \left(P_{\perp} K P_{\perp} + \gamma P_{\mathrm{b}}\right) g^k + 2 \left(g^k\right)^{\mathrm{T}} P_{\perp} \delta^k$$

where the term γP_{b} is a regularizing term to ensure that $P_{\mathrm{b}} g^k = 0$ and $\gamma > 0$ is otherwise arbitrary. The Hessian matrix, unlike K, is invertible and the normal equations then give

$$g^k = -A^{-1} \left(P_{\perp} \delta^k\right)$$

where $A = \left(P_{\perp} K P_{\perp} + \gamma P_{\mathrm{b}}\right)$ is the Hessian of the modified quadratic form. It is useful to see that g^k is the k^{th} column of $-A^{-1}P_{\perp}$, and since A is symmetric, g^k is also the k^{th} row. Note that the current vector $P_{\perp}\delta^k$ corresponds to unit current injected at node k and uniformly removed on the boundary.

2.1.2. Woodbury formula

On a change to the conductivity, the global stiffness matrix changes according to the assembly by elements outlined earlier. The Likelihood for a given set of measurement currents is determined by the linear map with kernels being the Green's functions with a pole at one of the electrodes. As we have seen, the Green's functions are the rows of 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 rows of 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 $M \times P$ matrices U and W, the Woodbury formula 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 $P \times P$ which is cheap when $P \ll M$, which is the case for localized changes to the conductivity.

The change to the global stiffness matrix K can always be written as an outer product of two matrices UW^{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 a pixel that does not have nodes on the boundary, then $P_{\perp}U = U$, and $P_{\perp}W = W$, and consequently, the change to the modified global stiffness matrix also equals UW^{T}.

2.1.3. Rejection step

Consider now the calculation required to evaluate the likelihood for a candidate state found by changing the conductivity at one pixel. The likelihood is easily calculated given the value at electrodes of the updated Green's functions for current sources at electrodes. That is, we require the components of the Green's functions g_l^k for $k, l = 1, 2, \cdots E$. But this is just the block of the inverse matrix A_{kl}^{-1} for $k, l = 1, 2, \cdots E$.

Changes to that block of the inverse require calculation of the term $\left[A^{-1}U \left(I + W^{\mathrm{T}} A^{-1} U\right)^{-1} W^{\mathrm{T}} A^{-1}\right]$ over the block. The kl element of this term is

$$\left[A^{-1}U \left(I + W^{\mathrm{T}} A^{-1} U\right)^{-1} W^{\mathrm{T}} A^{-1}\right]_{kl} = \left(g^k\right)^{\mathrm{T}} U \left(I + W^{\mathrm{T}} A^{-1} U\right)^{-1} W^{\mathrm{T}} g^l. \tag{5}$$

Hence we need to store the Green's functions for sources at electrodes, i.e., g^k for $k, = 1, 2, \cdots E$. Since the matrices U and W have non-zero entries only at rows corresponding to a node on the pixel being changed, this calculation

only requires the values of these Green's functions at the nodes on the pixel. The 4×4 matrix to be inverted, $\left(I + W^{\mathrm{T}} A^{-1} U\right)$, requires the components of A^{-1} for rows and columns corresponding to nodes on the pixel, since U and W are non-zero for those entries, only. Hence we also need to store the components of the Green's functions for source and field points at nodes on pixels, or pairs of pixels etc, that will be changed.

In summary, if we store the E Green's functions which have sources at electrodes, requiring $E \times M$ components stored, and components local to the source of all M Green's functions, requiring $9 \times M$ components for single pixel changes or $25 \times M$ for pixel swapping with a maximum gap of 2 pixels between those being swapped, then the likelihood corresponding to any single or double-pixel change may be computed exactly for the cost of a $P \times P$ matrix inversion plus two $P \times P$ matrix multiplications, where P nodes are on the pixels being changed. The storage and calculation requirements may be reduced somewhat by exploiting the symmetry of the Green's functions, but the saving is not substantial.

If the candidate state is rejected, 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.

2.1.4. 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 5 can be used to efficiently recalculate these values. Using the symmetry of the Green's functions we have $g_j^k = g_k^j$, and so changes to any component of A^{-1} may be computed given all values of the Green's functions for sources at nodes on the pixels that were changed. At present we can see no alternative to simply having to calculate these Green's functions following an acceptance. Since this requires P solutions of the BVP, $O(P)\, O\left(M^2\right)$ calculations are required to find these Green's functions. Note that it is not practical to store these Green's functions 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.

3. 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. This will enable improved Bayesian inference from current/voltage data. 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.

We can be fairly confident that experiments from synthetic data will be successful, since our earlier simulations achieved ergodic behaviour of the Markov chain in reasonable times, using a less efficient algorithm. Removal of the approximation inherent in using the linearized forward map is likely to allow 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 simulation for Bayesian inference from wave scattering data.

REFERENCES

1. C. Fox and G. K. Nicholls, "Physically-based likelihood for ultrasound imaging," in *Bayesian Inference for Inverse Problems*, A. Mohammad-Djafari, ed., vol. 3459 of *Proceedings of SPIE*, pp. 92–99, The International Society for Optical Engineering, July 1998.
2. C. Fox and G. K. Nicholls, "Sampling conductivity images via MCMC," in *The Art and Science of Bayesian Image Analysis*, K. Mardia, R. Ackroyd, and C. Gill, eds., Leeds Annual Statistics Research Workshop, pp. 91–100, University of Leeds, 1997.
3. G. K. Nicholls and C. Fox, "Prior modelling and posterior sampling in impedance imaging," in *Bayesian Inference for Inverse Problems*, A. Mohammad-Djafari, ed., vol. 3459 of *Proceedings of SPIE*, pp. 116–127, The International Society for Optical Engineering, July 1998.