Swap moves for MCMC applied to electrical impedance tomography

Erfang Ma[1] and Colin Fox[2]

[1] Xi'an Jiaotong-Liverpool University, Suzhou, China
[2] Electronics Group, Department of Physics, University of Otago, PO Box 56, Dunedin, New Zealand

Email: affinema@yahoo.com.cn & fox@physics.otago.ac.nz

Abstract: Electrical impedance tomography (EIT) is a non-invasive means of imaging the conductivity (or permittivity) distribution within an object. In the Bayesian framework, computation of 'solutions' is performed using Markov chain Monte Carlo (MCMC). Despite 20 years of development, the simplest single-site random walk update still is viewed as most efficient in some present MCMC implementations for EIT. We report on experiments utilizing swap moves in MCMC that significantly improve on the single-site update.

Keywords: Electrical impedance tomography, inverse problem, computational inference, Markov chain Monte Carlo, single-site update, swap moves

1. INTRODUCTION

Electrical impedance tomography (EIT) is a canonical inverse problem in which the electrical property within an object is inferred from electrical measurements on the surface of the object. This leads to many potential applications, when structures of interest have a contrast in electrical properties, such as scanning for breast cancers and measuring flow in milk-drying plants.

We denote the unknown electrical conductivity by x and the (noisy) surface measurements of potential at electrodes by $u^{\rm n}$. EIT is known to be a severely ill-posed inverse problem [1]; that is, direct inversion to give x in terms of $u^{\rm n}$ gives a solution that is highly sensitive to errors on measured data, and to model errors. This can be understood in terms of the physics of electrical currents, since each measurement of surface potential depends on all of the unknown conductivity; this follows from Ohm's law that states that current follows *all* paths, in proportion to conductance of each path[1]. Hence, measurements are most sensitive to bulk properties of the conductivity, such as the mean conductivity, while localized or 'high spatial frequency' variations in conductivity have little influence on measurements.

This sensitivity to bulk processes actually makes EIT a good imaging mode for those situations where we wish to non-invasively measure bulk properties. One such application is the measurement of *total* flow of the various constituents in milk powder flowing in a pipe, which is an application we are developing at Otago. In a sense, we can think of EIT as a set of scales – that measures the total mass of an object on the scales – though with some low resolution spatial discrimination.

Here we consider a simple observation process, in which noise-free measurements at electrodes u are a known function of the conductivity x and measurement noise e is additive and independent of x, that is

$$u^{\rm n} = u + e. \tag{1}$$

The functional relationship $x \mapsto u$ is called the *forward map* for EIT. In this paper we address the problem of computing statistical answers to questions about x given measurements $u^{\rm n}$, a known forward map, and knowledge of the (statistical) distribution over noise e.

2. A TEST PROBLEM

Figure 1 shows a phantom conductivity in a square region that we will attempt to recover, from simulated noisy measurements. The conductance is constant within each square pixel in a 24×24 pixel grid. Measurements consist of injecting (unit) current into each of 16 point electrodes in turn, with current removed uniformly on all electrodes, and measuring the potential ate each electrode with respect to the average potential on electrodes. This stylized problem has been studied previously in [2, 3].

[1] The common saying that that electricity travels "by the path of least resistance" is incorrect.

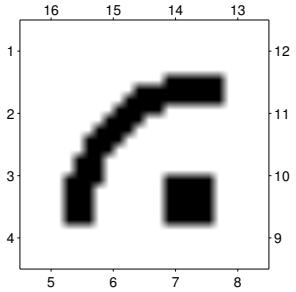

Figure 1: Phantom conductivity distribution, having conductance 3 (white) and 4 (black) in arbitrary units. Sixteen point electrodes on the boundary are numbered 1,... 16.

We use the same numerical computation of the forward map for simulating measurements and performing the inverse problem; hence we are certainly committing every possible 'inverse crime' and so recovered conductivity distributions need to be viewed as overly optimistic. However, the speed of performing the inference is our primary concern here, and that is not affected by the inverse crime.

3. COMPUTING THE FORWARD MAP

Simulation of forward map $x \mapsto u$ involves solving the boundary value problem (BVP)

$$-\nabla \cdot (x(s)\nabla u(s)) = 0 \qquad s \in \Omega, \tag{2a}$$

$$-x(s)\frac{\partial u}{\partial n}(s) = j(s) \qquad s \in \partial\Omega. \tag{2b}$$

We solve for $u(s)$ that is the potential at location s, while $x(s)$ is the that is a given conductivity distribution. We have not yet stated a potential reference, though that is necessary to ensure the solution of this BVP is unique. The current crossing the boundary $\partial\Omega$ is denoted $j(s)$, which must satisfy $\int_{\partial\Omega} j(s)\mathrm{d}s = 0$ since current is conserved.

We solve the BVP (2) using a standard finite element method (FEM) in which the pixels are taken as elements, and potential is interpolated using bilinear interpolation (see [2, 3] for details). FEM discretization results in a symmetric sparse system of linear equations

$$Ku = f \tag{3}$$

where the global stiffness matrix K is assembled (by elements) as

$$K(x) = \sum_{i=1}^{n} x_i K^i \tag{4}$$

where x_i is the conductivity in pixel i and K^i is the local stiffness matrix for pixel i. We solve the sparse system (3) using a sparse Cholesky factorization of the stiffness matrix K followed by solution of the resulting triangular system.

4. PRIOR MODELING AND POSTERIOR DISTRIBUTION

We follow [3] and model the conductivity as continuous valued in each pixel, restricted to the interval $[2.5, 4.5]$. The 'blocky' nature of the image is modeled by the prior distribution

$$\pi_p(x) \propto \exp\left\{\beta \sum_{i\sim j} v(x_i - x_j)\right\} \mathbb{I}_{[2.5,4.5]^n}(x) \tag{5}$$

where

$$v(s) = \begin{cases} \frac{1}{w}\left(1 - |s/w|^3\right)^3 & \text{if } |s| < w \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

and $\mathbb{I}_{[2.5,4.5]^n}(x)$ is the indicator function for each component of x being in the interval $[2.5, 4.5]$. The sum is over all nearest neighbours on the pixel lattice, denoted by $i \sim j$. This prior encourages neighbouring x_i and x_j to have similar values, but once x_i and x_j are more than w apart, the penalty does not grow. This allows occasional large shifts between neighboring pixels. Thus, the effect is to prefer images with connected regions with constant value, separated by distinct jumps in conductivity. The lumping parameter β is set to 0.5 and the separation value w is set to 0.3.

The likelihood function is

$$l(u^{\mathrm{n}}|x) \propto \exp\left\{-\frac{1}{2\sigma^2}\|u^{\mathrm{n}} - u\|^2\right\} \tag{7}$$

when the noise e in Eqn 1 is iid (independent identically distributed) zero-mean Gaussian with variance σ^2. As in [3], we set σ so that the resulting SNR (signal to noise ratio) in the measurement process is close to 1000:1 to mimick practical EIT measurements.

By Bayes' rule, the posterior distribution over conductivity images is

$$\pi(x) \propto l(u^{\mathrm{n}}|x)\pi_p(x). \tag{8}$$

It is not possible to summarize the posterior distribution by analytic means, so we employ MCMC (Markov chain Monte Carlo) to sample from this distribution and evaluate summary statistics.

Figure 2 shows five samples drawn from the posterior distribution. These samples are close to each other and all resemble the original conductivity distribution, indicating that the posterior distribution is supported around the true conductivity distribution. Figure 3 shows the posterior

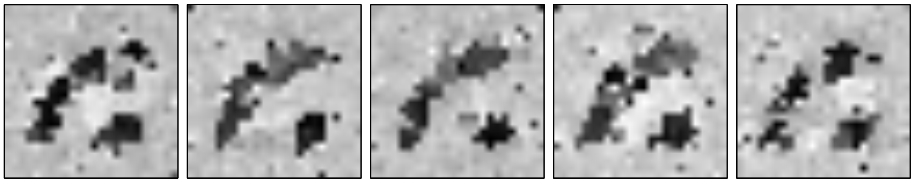

Figure 2: Five samples of the conductivity x drawn from the posterior distribution.

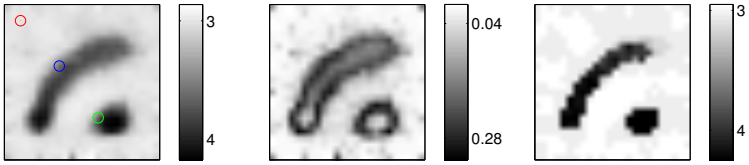

Figure 3: Posterior mean (left), posterior variance (middle), and marginal posterior mode (right).

mean, posterior variance, and the marginal posterior mode (MPM) (see [4] for a discussion on the MPM). As can be seen, the marginal posterior mode gives a useful estimate of the true conductivity distribution. The variance image shows that the primary uncertainty is in the location of boundaries of the regions of constant conductivity, while the conductivity values are well determined.

The results presented in this section were evaluated using samples drawn by the standard single-site Metropolis MCMC as in [3] that had run for $1.2 \times 10^5 \times 576$ iterations. In the following section we consider improved proposals.

5. SWAP MOVES

As the name implies, the swap moves simply swap the conductivity value in a pair of nearby pixels. We consider three specific swap moves as shown in Figure (4). They are (b) 1×2 swap between the two-nearest neighbours, (c) swap between the diagonal pixels within a 2×2 block, and (d) swap between two components within a 3×3 block as shown. Swap move (b) was used previously in [5, 2], while moves (c) and (d) are special cases of moves used in [5, 6], though in the context of a discrete representation. Also shown is (a) the single site update move.

Each of these swap moves provide a localized update that leaves approximately invariant the posterior distribution in Eqn 8. When these moves are chosen with probabilities p_1, p_2, p_3, p_4, and the Metropolis-Hastings (MH) accept/reject rule applied, the resulting transition kernel is

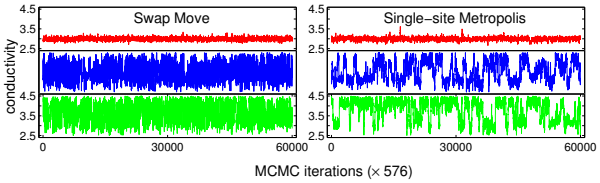

Figure 5: Traces of conductivity at three representative pixels for the swap move (with probabilities $(0.1, 0.1, 0.4, 0.4)$) and single-site Metropolis.

$$A(x, y) = \sum_{i=1}^{4} p_i A_i(x, y) \tag{9}$$

where $A_i(x, y)$ is the transition kernel resulting from a MH algorithm using just the i-th proposal. Since kernel $A_1(x, y)$ associated with the single-site move is irreducible and we set $p_1 > 0$, the resulting chain is also irreducible.

We tested eight sets of move probabilities listed in Table 1 and ran the resulting algorithms for $1.2 \times 10^5 \times 576$ iterations. During each run, the state-variable was recorded every 10×576 updates, while log-likelihood and log-prior were recorded every 576 updates. As in [3], three representative pixels were selected (marked by red, blue and green circles in Figure 3) and their trace plots were used for qualitative assessment of how well the Markov chain mixed. From the traces shown in Figure 5 it is clear that using the swap moves significantly improves performance of the MCMC. As in [3], we also evaluated the integrated autocorrelation time (IACT) [7] at the 'blue' pixel as a quantitative measure of the performance of the

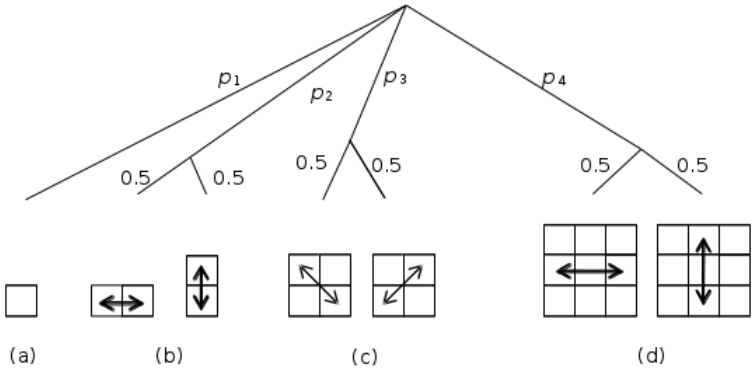

Figure 4: Probability tree for moves: (a) random scan single-site update, (b) 1×2 swap, (c) 2×2 swap and (d) 3×3 swap, that are selected with probabilities p_1, p_2, p_3, p_4, respectively.

Table 1: Estimated IACT at 'blue' pixel τ_{blue} with various swap move probabilities.

(p_1, p_2, p_3, p_4)	$\tau_{\text{blue}}(\times 10 \times m)$
(0.1,0.1,0.7,0.1)	9.5 ± 2.1
(0.1,0.1,0.1,0.7)	8.3 ± 1.8
(0.1,0.7,0.1,0.1)	7.5 ± 1.5
(0.1,0.1,0.4,0.4)	5.6 ± 1.0
(0.1,0.4,0.4,0.1)	6.5 ± 1.2
(0.1,0.4,0.1,0.4)	11 ± 2.6
(0.1,0.3,0.3,0.3)	7.3 ± 1.5
(0.1,0.9,0.0,0.0)	20 ± 6

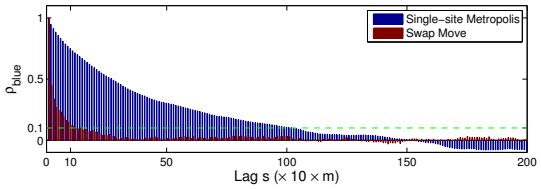

Figure 6: Estimated normalized ACF at x_{blue} using swap move and single-site Metropolis.

MCMC sampler, also listed in Table 1.

We were not able to accurately estimate IACT at the 'blue' pixel using the Markov chain with just the single-site update, even after $1.6 \times 10^5 \times 576$ iterations. Instead we show in Figure 6 the normalized ACF at at the 'blue' pixel using singl-site update and the swap move (as in Figure 5). As can be seen, the normalized ACF decays to zero about 10 time faster for the swap move algorithm compared to the single-site algorithm. This confirms the qualitative assessment, and shows that the swap move is about ten times more efficient than the single-site Metropolis.

Acknowledgements

EM thanks the NZIMA for its generous support of his PhD study.

6. REFERENCES

[1] C. Fox, *Conductance Imaging*. PhD thesis, University of Cambridge, 1989.

[2] J. D. Moulton, C. Fox, and D. Svyatskiy, "Multilevel approximations in sample-based inversion from the Dirichlet-to-Neumann map," *Journal of Physics: Conference Series*, vol. 124, no. 1, 2008.

[3] D. Higdon, C. S. Reese, J. D. Moulton, J. A. Vrugt, and C. Fox, "Posterior exploration for computationally intensive forward models," in *Handbook of Markov Chain Monte Carlo*, Chapman & Hall/CRC, 2011.

[4] C. Fox and G. K. Nicholls, "Exact map states and expectations from perfect sampling: Greig, Porteous and Seheult revisited," in *Bayeisan Inference and Maximum Entropy Methods in Science and Engineering, 20th International Workshop*, pp. 252–263, 2001.

[5] G. K. Nicholls and C. Fox, "Prior modeling and posterior sampling in impedance imaging," in *Bayesian Inference for Inverse Problems*, 1998.

[6] J. A. Christen and C. Fox, "MCMC using an approximation," *Journal of Computational and Graphical Statistics*, vol. 14, no. 4, pp. 795–810, 2005.

[7] U. Wolff, "Monte Carlo errors with less errors," *Computer Physics Communications*, vol. 156, pp. 143–153, 2004.