A Bayesian Approach to Crack Detection in Electrically Conducting Media

Kim E. Andersen†*, Stephen P. Brooks‡ and Martin B. Hansen†

† Aalborg University, Denmark

‡ University of Cambridge, U.K.

Abstract. In this paper, we review powerful new computational techniques which facilitate the Bayesian approach to statistical inference and discuss how they may be used to solve general inverse problems. Their power and flexibility is illustrated by the problem of detecting a finite set of linear non-intersecting perfectly insulating cracks in a homogeneously electrically conducting media. In this case, efficient algorithms only exist if the number of cracks is known a priori. However, in this paper we demonstrate how uncertainty about the number of cracks can be incorporated into the modelling process and assessed together with crack locations.

AMS classification scheme numbers: 60J25, 65N21

1. Introduction

It is often desirable to determine the interior structure of a given body in a non-destructive manner. Common non-destructive evaluation systems are based upon tomographic principles where the reconstruction is an inverse problem. Primary interest is in some object Ψ but data are only accessible about some transform $F\Psi$. The data are observed at m discrete points t_i and are corrupted by noise, so that they take the form

$$\Phi(t) = F\Psi(t) + \epsilon(t), \quad t = t_1, \ldots, t_m, \tag{1}$$

where $\epsilon(t)$ denotes a random noise process, typically assumed to be Gaussian.

The most obvious approach for reconstructing Ψ is to estimate Φ by $\widehat{\Phi}$ using standard statistical methods (Efromovich, 1999) and then use $\widehat{\Psi} = F^{-1}\widehat{\Phi}$ as an estimate of Ψ. However, such problems are generally ill-posed (Hadamard, 1923) which means that either the inverse operator F^{-1} does not exist, or it is an unbounded operator. In the latter case, even small perturbations of the data, and thereby $\widehat{\Phi}$, may result in unacceptably large distortions of $\widehat{\Psi}$. In order to avoid this instability Tikhonov regularisation (Tikhonov, 1963*a*,*b*) is often used. This is a method of balancing

* Address for correspondence: Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg Ø, Denmark, e-mail: emil@math.auc.dk

the degree of misfit of the solution Ψ and a functional $J(\Psi)$ which measures the 'appropriateness' of the function Ψ. The specification of J is crucial and is usually chosen on the basis of prior information about the solution space. The Bayesian approach has therefore enjoyed some popularity as a tool for formalising the incorporation of prior information and for properly accounting for uncertainty (Tarantola, 1987). To add to this popularity, powerful new simulation techniques are now available which allow even more complicated and realistic models, see Mosegaard and Tarantola (1995), Mosegaard (1998), Dahl-Jensen et al. (1998), Nicholls and Fox (1998) and Mosegaard and Rygaard-Hjalsted (1999). In this paper we review some of the recent progress in Bayesian computation, focussing upon Markov chain Monte Carlo methods, and indicate how they may be used to solve general inverse problems. For illustration, we will consider the problem of detecting a finite set of linear non-intersecting perfectly insulating cracks in homogeneously conducting media.

A variety of different imaging algorithms exist for non-destructive evaluation on the basis of electrical impedance tomography data. Many are general-purpose algorithms which seek to reconstruct the entire conductivity profile within the object (see e.g. Barber and Brown, 1986; Santosa and Vogelius, 1990; Cheney et al., 1998). Such algorithms typically base all computations for the underlying conductance problem upon a 2–dimensional finite element formulation. However, as illustrated by Bryan and Vogelius (1994) the efficiency and versatility of these algorithms can be improved dramatically when structural prior information about the profile is available.

Bryan and Vogelius (1994) develop a crack detection algorithm based upon a variation of Newton's method in which an initial crack configuration Ψ_0 is iterated until a solution is met. However, this method has the drawback that the solution obtained is critically dependent upon the initial crack configuration, as the initial number of cracks remains fixed throughout the iterative estimation procedure. Since the number of cracks is rarely known a priori, *ad hoc* techniques are generally used to overcome this problem.

In this paper, we show how problems associated with the a priori determination of the number of cracks may be overcome by explicitly incorporating this source of uncertainty within the modelling process. In this way we determine not only the number of cracks but also the crack locations in the media under study.

In section 2 we summarise the mathematical model for the voltage potential in the presence of cracks in an electrically conducting media. In section 3 we introduce the Bayesian approach to statistical inference and the computational techniques required for its implementation. Specific implementational issues within the reconstruction algorithm are addressed in section 4. In section 5 we discuss the behaviour and performance of our approach and finally in section 6 we discuss in the light of the achieved results further potential developments.

2. Modelling boundary measurements in the presence of cracks

Suppose that any crack lies within a 2–dimensional simply connected conductor Ω and is modelled as a perfectly insulating line segment. In this paper we assume that a given crack ψ is parameterised by

$$\boldsymbol{\psi}(t) = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a}), \qquad 0 \leq t \leq 1,$$

where $\boldsymbol{a} = (x_1, y_1)$ and $\boldsymbol{b} = (x_2, y_2)$ denote ψ's endpoints in Ω. We let $\Psi = \cup_{k=1}^{n} \psi_k$ denote a finite collection of n non-intersecting cracks in Ω (see figure 1). At low frequencies and with a given background conductivity $\gamma(\boldsymbol{x})$ satisfying $0 < \gamma_0 \leq \gamma(\boldsymbol{x}) \leq \gamma_1$ for $\boldsymbol{x} \in \Omega \setminus \Psi$, Maxwell's steady-state conductance equations for the voltage potential u read

$$\begin{aligned} &\nabla \cdot (\gamma^{-1} \nabla u) = 0 \quad \text{in} \;\; \Omega \setminus \Psi \\ &u = c_k \qquad\qquad\quad \text{on} \;\; \psi_k, \;\; k = 1, \ldots, n, \end{aligned}$$

for a particular set of constants $c_k, k = 1, \ldots, n$. If we let $\boldsymbol{\nu}$ denote the normal field to $\partial\Omega$, then the Neumann boundary condition takes the form

$$\gamma^{-1} \frac{\partial u}{\partial \boldsymbol{\nu}} = \chi \qquad \text{on } \partial\Omega,$$

which is used to model the applied current. As current flows into the object at electrode $\boldsymbol{x}_{\text{in}}$ and exits at electrode $\boldsymbol{x}_{\text{out}}$ (see figure 1) and the electrodes are modelled as point sources with diameter h, then the function χ is defined in terms of characteristic functions, i.e.

$$\chi(\boldsymbol{x}) = \delta(\boldsymbol{x}, \boldsymbol{x}_{\text{in}}) - \delta(\boldsymbol{x}, \boldsymbol{x}_{\text{out}}), \qquad \boldsymbol{x} \in \partial\Omega,$$

where

$$\delta(\boldsymbol{x}, \boldsymbol{x}_{\text{in}}) = \begin{cases} 1/h & \text{if } |\boldsymbol{x} - \boldsymbol{x}_{\text{in}}| < h/2, \\ 0 & \text{else.} \end{cases}$$

The function $\delta(\boldsymbol{x}, \boldsymbol{x}_{\text{out}})$ is defined similarly.

By application of standard potential theory arguments, Bryan and Vogelius (1994) develop an integral equation formulation of this boundary value problem based upon the assumption that the object Ω has constant background conductivity $\gamma \equiv 1$. As this formulation depends upon singular boundary data u, a more useful formulation is obtained by constructing a perturbation $v = u - u_0$ which is smooth up to $\partial\Omega$. Here u_0 satisfies the conductance equations for $\boldsymbol{x} \in \partial\Omega$ when no cracks are present. As v is smooth up to $\partial\Omega$, any problems associated with the lack of regularity of u are avoided.

For the purposes of this paper, we let $\boldsymbol{\Psi}$ denote the collection of allowable crack configurations, i.e., $\boldsymbol{\Psi} = \cup_{n=0}^{\infty} \boldsymbol{\Psi}_n$, where $\boldsymbol{\Psi}_n$ is the set of n non-intersecting cracks fully within Ω. Furthermore, we let $S : \boldsymbol{\Psi} \to H^1(\Omega)$ denote the solution operator of the derived integral equation, where $H^1(\Omega)$ is a Sobolev space of order 1, for more details consult Bryan and Vogelius (1994).

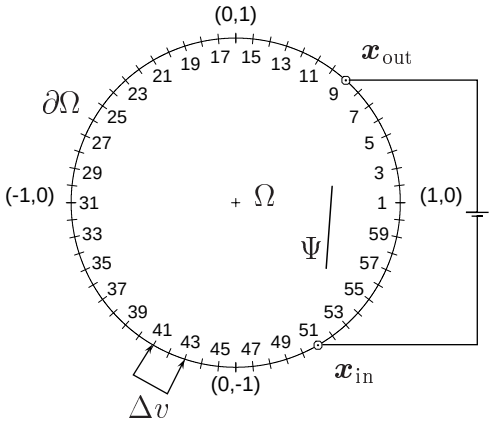

Figure 1. Schematic representation of the experimental setup. The object Ω has an interior crack Ψ which is to be determined. Current is flowing in at electrode $\boldsymbol{x}_{\text{in}} = 51$ and out at electrode $\boldsymbol{x}_{\text{out}} = 9$. Corresponding potential voltage differences are recorded.

To discretise the integral formulation we let $\boldsymbol{z}(t)$ parameterise $\partial\Omega$ for t-values $0 \leq t < 1$ and parameterise ψ_k for values $k \leq t < k+1$ for any $k = 1, \ldots, n$. Furthermore we let

$$K(t_i, t_j) = \frac{\partial}{\partial \boldsymbol{\nu}_y} \Gamma[\boldsymbol{z}(t_i), \boldsymbol{z}(t_j)] |\boldsymbol{z}'(t_j)|,$$
$$G(t_i, t_j) = \Gamma[\boldsymbol{z}(t_i), \boldsymbol{z}(t_j)] |\boldsymbol{z}'(t_j)|,$$

where $\boldsymbol{\nu}_y$ on $\partial\Omega$ denotes the normal outward derivative with respect to the y variable and Γ is the Green's function

$$\Gamma(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2\pi} \log |\boldsymbol{x} - \boldsymbol{y}|, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2.$$

Finally, we let

$$\phi_{kj} = \begin{cases} \left. \dfrac{\partial u(\boldsymbol{x}(t))}{\partial \boldsymbol{\nu}} \right|_{t=t_j} & \text{on } \psi_k \ (k = 1, \ldots, n), \\ v[\boldsymbol{z}(t_j)] & \text{on } \partial\Omega \ (k = 0), \end{cases}$$

where $\partial/\partial\boldsymbol{\nu}$ on ψ_k denotes the normal flux across the crack.

Now, by application of Nyström's method (see e.g. Atkinson, 1976), the integrals in the integral formulation of Bryan and Vogelius (1994) are discretised by quadrature rules to obtain a set of linear equations. Hence, if we let m denote the number of nodes to be used in the quadrature rule, we obtain the following linear system

$$-\frac{1}{2}\phi_{0i} + \sum_{j=1}^{m} K(t_i, t_j)\omega_j \phi_{0j} - \sum_{k=1}^{n}\sum_{j=1}^{m} G(t_i, k + t_j)\omega_j \phi_{kj} = 0,$$
$$\sum_{j=1}^{m} K(l + t_i, t_j)\omega_j \phi_{0j} - \sum_{k=1}^{n}\sum_{j=1}^{m} G(l + t_i, k + t_j)\omega_j \phi_{kj} - c_l = -u_0[\boldsymbol{z}(l + t_i)],$$

$$\sum_{j=1}^{m} \phi_{lj}\omega_j |\boldsymbol{z}'(l+t_j)| = 0 \quad \text{and}$$
$$\sum_{j=1}^{m} \phi_{0j}\omega_j |\boldsymbol{z}'(t_j)| = 0,$$

for $i = 1, \ldots, m$ and $l = 1, \ldots, n$. Due to linear dependencies in the first m equations, any one of them may be ignored, leaving a quadratic linear system of $mn + m + n$ equations.

The remaining linear system is well-conditioned due to the presence of the logarithmic singularity along the first part of the diagonal $s = t$. The singularity can be dealt with by a simple form of product integration (Atkinson, 1976) based upon the applied quadrature rule. Of course, the quadrature rule used on the boundary need not be the same as the one used on the cracks. As the solution is smooth on $\partial\Omega$, a simple trapezoidal rule can be used with weights $w_i = 1/m$ for $i = 1, \ldots, m$. However, on the cracks care must be taken due to singularities at the endpoints. Bryan and Vogelius (1994) propose using a quadrature rule that places many points close to, but not at, the endpoints. The corresponding weights are chosen to accommodate Nyström's method corresponding to a midpoint rule with variably spaced nodes.

In practice, the crack detection algorithm of Bryan and Vogelius (1994) begins with an initial guess Ψ_0 for which the corresponding boundary voltage data can be predicted by solving the linear system described above. This crack configuration is then iteratively improved with respect to the residual sum of squares using a Newtonian updating scheme. We refer the reader to Bryan and Vogelius (1994) for further details. Their algorithm was tested on a large number of simulated data sets with linear cracks and it appeared to work well when the number of cracks in the initial configuration was the same as the number in the true configuration. Bryan and Vogelius (1994) also consider the problem associated with estimating the number of cracks. Using simulated data with e.g. three 'true' cracks, they proceed as illustrated in figure 2 and described as follows. The reconstruction procedure is begun by attempting to fit the simulated data with just one crack and with endpoints situated at $(-0.2, 0.2)$ and $(-0.2, 0.6)$. A root for the linear system is found, but at this point it is impossible to say whether or not the simulated data come from just a single crack or several cracks. The middle 1/10th of the estimated crack is cut out and two new cracks are formed and used as the initial guess for a new run of the algorithm. This process is repeated so that at each stage the largest crack of those in the current configuration is split and this new configuration is used as an initial value for repeating the process with an additional crack. The estimated cracks are shown in figure 2(b) – (d).

Bryan and Vogelius (1994) point out that three things are likely to occur when splitting a crack to obtain a new crack configuration: (1) The cracks find different locations, (2) the cracks remain essentially where they are, or (3) one of the split cracks shrinks and the other cracks remain where they were previously estimated to be. The behaviour described in (2) or (3) usually indicates that the previous number of cracks

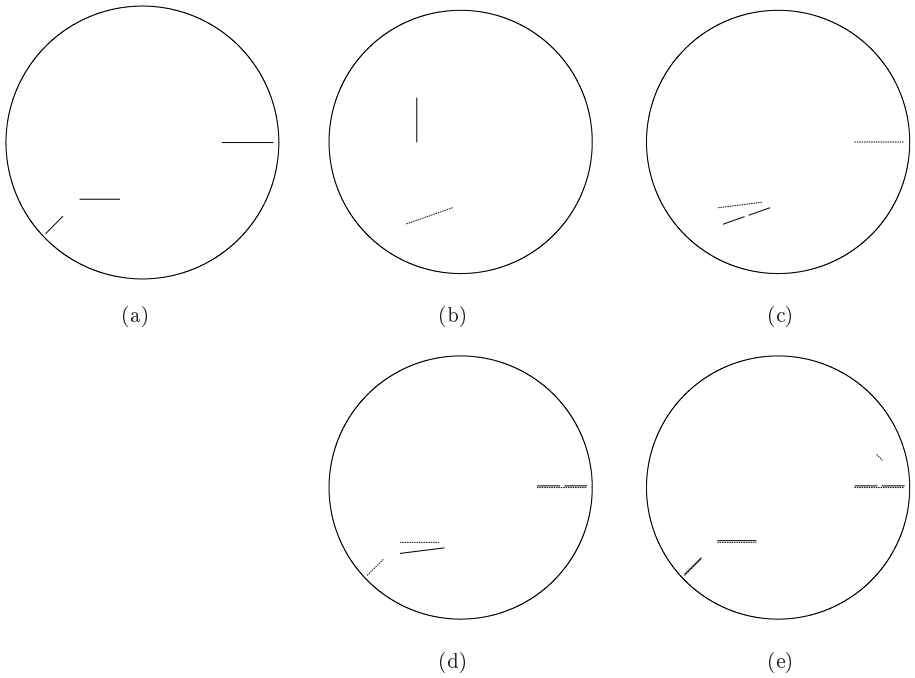

Figure 2. The deterministic reconstruction algorithm applied to unperturbed simulated crack data: (a) The experimental setup; and (b) – (e): crack estimates for one, two, three and four cracks fitted to data (initial guess:——, estimate:· · · · · ·).

used was the right number to fit to the data. When fitting four cracks to the data behavior (3) is observed. From this, Bryan and Vogelius (1994) conclude that three cracks is the right number of cracks to fit to the data.

In this paper, we provide a more rigorous framework for determining the number and position of the cracks. We begin by obtaining the likelihood function, L, derived from (1) under the assumption that the errors $\epsilon(t)$ are independently Gaussian distributed with mean zero and variance σ^2, i.e.

$$L(\Psi, \sigma \,|\, \Phi) \propto \frac{1}{\sigma^m} \exp\left\{-V(\Psi, \sigma)\right\}, \tag{2}$$

where

$$V(\Psi, \sigma) = \sum_{i=1}^{m} [\Phi(t_i) - F\Psi(t_i)]^2 / 2\sigma^2,$$

$F\Psi(t_i) = S\Psi(t_i) - S\Psi(t_{i-1})$ is the recorded voltage differences and S is the solution operator introduced in section 2. The classical statistical approach to models described in such a manner is to seek the configuration Ψ and parameter σ^2 that maximises this function. However, greater flexibility is obtained by adopting a Bayesian approach, as described in the following section.

3. The Bayesian approach to statistical inverse problems

If we assume that the observed data are described by the model given in (1), then the statistical problem is to estimate Ψ and σ^2 given Φ. For this we need to determine the *posterior distribution* of $\Theta = (\Psi, \sigma^2)$, denoted by $\pi(\Theta \,|\, \Phi)$, which describes our beliefs about the possible crack configurations and variance parameter after having observed the data Φ. Let us assume also that our beliefs about the possible configurations Θ before having observed any data can be described by a *prior distribution* denoted by $p(\Theta)$. Bayes' formula may then be used to derive the posterior distribution as a function of the prior distribution and the data represented through the likelihood function. We thus have that

$$\pi(\Theta \,|\, \Phi) \propto p(\Theta) L(\Phi \,|\, \Theta).$$

In practice, statistical inference is obtained through the calculation of posterior moments, for example the posterior mean of Θ given by

$$\mathbb{E}_\pi(\Theta) = \int \Theta \, \pi(\Theta \,|\, \Phi) \, \mathrm{d}\Theta. \tag{3}$$

The problem therefore reduces to the computational task of performing this integration, typically over a large parameter space. Often, explicit evaluation is impossible and traditionally one would use numerical integration or analytic approximation methods. However, Markov chain Monte Carlo (MCMC) methods provide an alternative integration technique whereby posterior means, for example, are estimated by using the sample mean from a series of random draws from the posterior distribution. These random draws are obtained by constructing a Markov chain $\{\Theta_1, \Theta_2, \ldots\}$ with state space $\boldsymbol{\Theta}$ and with stationary distribution $\pi(\Theta \,|\, \Phi)$. MCMC sampling was first introduced by Metropolis et al. (1953) and was subsequently adapted by Hastings (1970). Over the past ten years such methods have enjoyed widespread popularity within the statistical literature and there exist various standard techniques for constructing the necessary chains. See Brooks (1998) and Robert and Casella (1999), for example.

For the purposes of this paper, we shall distinguish between three separate simulation algorithms that will be combined to provide a technique suitable for the crack detection problem. We begin by describing the Metropolis–Hastings updating scheme for updating a crack configuration without altering the number of cracks. We then describe the reversible jump MCMC update which allows us to update the number of cracks, and finally we introduce the simulated tempering method as a means of increasing the computational efficiency of the overall simulation process.

3.1. Metropolis–Hastings updates

Metropolis–Hastings updates are used to move around the parameter space by proposing moves which are subsequently either accepted or rejected. Suppose that we are currently

in configuration Θ, then we draw a new configuration Θ' from some proposal density $q(\Theta, \Theta')$. This proposal is then accepted with probability

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{\pi(\Theta' \mid \Phi) q(\Theta', \Theta)}{\pi(\Theta \mid \Phi) q(\Theta, \Theta')}\right\}. \tag{4}$$

However, if the proposal is rejected, the chain remains in the current state. The choice of proposal distribution q is essentially arbitrary, though several forms possess useful analytic properties. For example, when the proposal distribution q is symmetric, i.e. $q(\Theta, \Theta') = q(\Theta', \Theta)$, the acceptance function in (4) reduces to

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{\pi(\Theta' \mid \Phi)}{\pi(\Theta \mid \Phi)}\right\},$$

which is essentially the original Metropolis update proposed by Metropolis et al. (1953). Another commonly used proposal distribution arises when $q(\Theta, \Theta')$ is a function only of $|\Theta - \Theta'|$, in which case the chain reduces to a random walk.

The Metropolis–Hastings updating scheme takes the simple algorithmic form

Initialise Θ_0; set $t = 0$
for $l = 1$ **to** L **do**
  Sample Θ' from q given Θ_{l-1}
  Compute the acceptance probability
    $\alpha_l = \alpha(\Theta_{l-1}, \Theta')$
  Sample uniform random variable u_l on $(0, 1)$
  if $u_l < \alpha_l$ **then** $\Theta_l = \Theta'$
    otherwise $\Theta_l = \Theta_{l-1}$
end

In practice the Metropolis–Hastings updating scheme can be used either to update the entire state vector or individual elements. Since 'large' jumps tend to have correspondingly small acceptance probabilities, typical MCMC algorithms consist of a sequence of updates focussing upon each element of the state vector in turn. We shall discuss this approach, known as single component Metropolis–Hastings, in more detail in section 4.

3.2. Reversible jump MCMC

The Metropolis–Hastings updates are used to update the state vector essentially moving the positions of the current cracks. In order to move between configurations with different numbers of cracks, we require what are known as reversible jump MCMC (RJMCMC) updates, since such updates involve moving between states of different dimensions. For example, the introduction of a new crack to the current crack configuration will increase the dimensionality of the state vector, since additional parameters will be needed to describe the position of the new crack. For these moves RJMCMC updates (Green, 1995) provide a general framework, extending the basic Metropolis–Hastings algorithm to general state spaces, so that π becomes a general

measure, rather than a density. Also, the proposal density $q(\Theta, \Theta')$ is replaced by the proposal distribution $q(\Theta, \mathrm{d}\Theta')$.

Suppose a dimension-changing move m is proposed and the proposal Θ' is generated by a deterministic invertible function $f(\Theta, u)$, where u is a continuous random variable. Then Green (1995) shows, that if $r_m(\Theta)$ denotes the probability of choosing move type m when in state Θ, $q(u)$ denotes the density function of u and $\pi(\Theta \mid \Phi)$ denotes the posterior density of Θ, then the corresponding acceptance probability becomes

$$\alpha(\Theta, \Theta') = \min\left\{1, \frac{\pi(\Theta' \mid \Phi) r_{m'}(\Theta')}{\pi(\Theta \mid \Phi) r_m(\Theta) q(u)} \left|\frac{\partial f(\Theta, u)}{\partial(\Theta, u)}\right|\right\}, \tag{5}$$

where m' denotes the reverse move to m. Note, that the final term in the above ratio is the Jacobian arising from the change of variables associated with moving from one space to the other. Algorithmically, the reversible jump updating procedure proceeds identically to that for Metropolis–Hastings updates.

The Metropolis–Hastings and reversible jump updates will produce a Markov chain with the required stationary distribution. However, the resulting chain may be slow to move around the state space so that large run lengths are required in order to obtain reliable inference. To improve the speed with which the state space is traversed (often termed the mixing rate), Marini and Parisi (1992) and Geyer and Thompson (1995) suggest the use of simulated tempering.

3.3. Simulated tempering

Simulated tempering is based upon using a series of stationary distributions $\pi_1, \ldots, \pi_L$ and augmenting the state vector to include an indicator variable signalling which distribution is being used at any time. If we let $\pi_1 = \pi$ and choose $\pi_l, l = 2, \ldots, L$ so that movement around the state space becomes easier as l increases, then we can run a chain on the augmented state space and base inference only upon observations in the chain attributed to distribution π_1. Since movement is easier for larger l, we obtain a more rapidly mixing chain and movement within the target distribution π_1 is facilitated by brief tours in other temperatures.

The chain swaps between temperatures i and j at time t and in state Θ_t with probability

$$\min\left\{1, \frac{\pi_j(\Theta_t \mid \Phi) c_j p_{j\to i}}{\pi_i(\Theta_t \mid \Phi) c_i p_{i\to j}}\right\},$$

where $p_{i\to j}$ denotes the probability of proposing the chain to move from sampler i to sampler j. The c_i's are approximate normalising constants that ensure that the chain divides its time roughly equally among the L different samplers (Geyer and Thompson, 1995).

4. Implementation

In this section, we provide a detailed description on the parameterisation of the models and on the updating mechanisms required within the MCMC simulation for the crack detection problem. For simplicity we also assume that Ω is the unit circle.

In order to ensure that the posterior is dominated by the likelihood, we adobt a vague prior distribution on the set of allowable crack configurations $\mathbf{\Psi}$ so that the number of cracks are Poisson distributed with mean λ. Then for any number of cracks n, we place a prior distribution on the possible crack configurations, $\mathbf{\Psi}_n$, by assuming that any configuration in $\mathbf{\Psi}_n$ is equally likely. These considerations can be formalised by assuming that the prior distribution p on $\mathbf{\Psi}$ has a density (with respect to the unit rate Poisson process on the product space $\Omega \times \Omega$) which satisfies

$$p(\Psi) \propto \lambda^{n(\Psi)} \mathbb{1}(\Psi \in \mathbf{\Psi}),$$

where $n(\Psi)$ is the number of cracks in configuration Ψ and $\mathbb{1}$ denotes the indicator function. The interested reader is referred to Daley and Vere-Jones (1988) for further details. For our purposes it is sufficient to note that the density is well-defined.

In order to move freely around the state space, we propose using two within-model moves, that allow for the movement of cracks within Ω, and one between-model move that adds and deletes cracks to/from the current configuration. The first move which we refer to as move-type A involves taking the coordinates of the endpoints of any particular crack and sampling a new value for each within a neighbourhood of the current value. This simultaneously translates and scales the chosen crack. Move type B involves rotating a chosen crack about its centre and, finally, move type C enables us to add a crack by proposing a new crack randomly placed within Ω. The reverse move in which a crack is deleted will be referred to as move type C'. These moves are illustrated in figure 3.

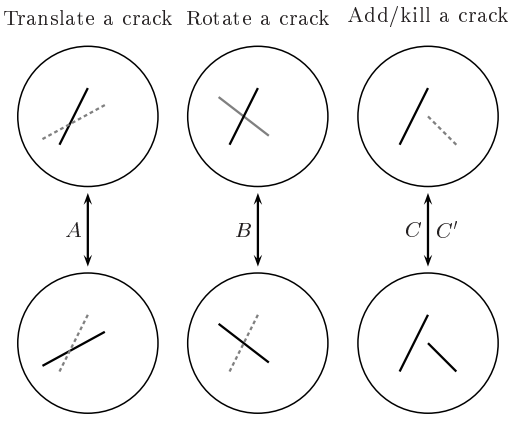

Figure 3. The proposed updates we use in the MCMC algorithm: From left to right we illustrate the different possible transitions for translating (A), rotating (B), and adding (C) and deleting (C') cracks.

Within each iteration, we begin by deciding whether we will perform a between-

model or within-model move. With probability ρ_{bd} we will perform a between-model move which will be either a death (i.e., we delete a crack – this is proposed with probability ρ_{d}) or a birth (i.e., we add a new crack – this is proposed with probability $1-\rho_{\mathrm{d}}$). If we choose to perform a within-model move, then with probability ρ_{r} we will choose to rotate all cracks, else we translate all cracks. These updates involve taking each crack in turn and performing the selected move type. The updating scheme is graphically illustrated in figure 4. We now consider each move type in turn and explain how each is performed.

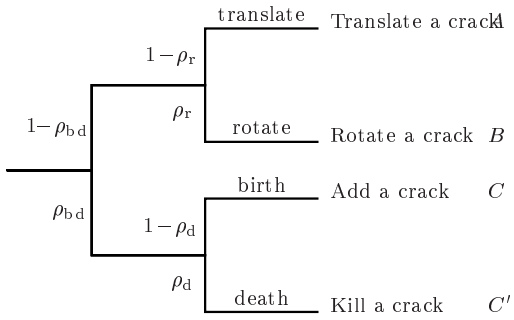

Figure 4. Schematic representation of the tree that the transition kernel ascends to pick a transition type.

4.1. Translating a crack

Suppose that we have decided to attempt a move of type A. We do this by considering in turn the four parameters (x_1, y_1, x_2 and y_2) of each crack and updating each parameter by selecting a new value within some small region of the current value. More specifically, to update any one of the four parameters, we sample a value uniformly in a symmetric interval around the current position with width $2\Delta l$. In proposing a change to any one of these parameters, we propose a move from Ψ to Ψ', with $q_A(\Psi, \Psi') = 1/2\Delta l$. Since this proposal is symmetric, the acceptance probability in (4) reduces to

$$\alpha_A(\Psi, \Psi') = \min\left\{1, \mathbb{1}(\Psi' \in \mathbf{\Psi}) \exp[V(\Psi, \sigma) - V(\Psi', \sigma)]\right\}.$$

Once the first of the four parameters has had a new value proposed and subsequently either accepted or rejected, we move on to the second. This continues until all four parameters have undergone this procedure. When all four parameters have been updated, we move on to the next crack until all cracks have been updated in turn which completes the MCMC iteration.

4.2. Rotating a crack

Again, let ψ denote an arbitrarily chosen crack in Ψ and assume that it is rotated counter-clockwise about its centre according to a rotation parameter θ'. We parameterise ψ by (x_1, y_1, x_2, y_2) with its centre given by $(c_1, c_2) = (x_1 + x_2, y_1 + y_2)/2$. Let θ denote

the angular direction from (c_1, c_2) to (x_1, y_1) with respect to the horizontal axis, then the crack is rotated counter-clockwise according to a rotation parameter θ' uniformly sampled in $[-\Delta\theta, \Delta\theta]$. Hence, the new crack ψ' is given by its new endpoints

$$\begin{aligned}(a_1', a_2') &= (c_1, c_2) + \lambda[\cos(\theta + \theta'), \sin(\theta + \theta')]/2,\\ (b_1', b_2') &= (c_1, c_2) - \lambda[\cos(\theta + \theta'), \sin(\theta + \theta')]/2,\end{aligned}$$

where $\lambda = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$ denotes the length of ψ. Thus, $q_B(\Psi, \Psi') = 1/2\Delta\theta$ and α_B is identical to α_A above.

4.3. Adding/deleting a crack

Assume that configuration Ψ consists of n cracks then, in the birth move, a new crack ψ is proposed by sampling two points $\boldsymbol{a} = (x_1, y_1)$ and $\boldsymbol{b} = (x_2, y_2)$ uniformly in Ω. By $\boldsymbol{u}$ we denote the vector of random variables (x_1, x_2, y_1, y_2) for the new crack position so that $q_C(\boldsymbol{u}) = 1/\pi^2$. Here, the probability of proposing move type C is given by $r_C = \rho_{\mathrm{bd}}(1 - \rho_{\mathrm{d}})$ and the probability of proposing the reverse is given by $r_{C'} = \rho_{\mathrm{bd}}\rho_{\mathrm{d}}/(n + 1)$. The latter being derived from the product of the probabilities of picking a between-model move, then a death and finally picking crack ψ from $n + 1$ cracks in Ψ'.

The Jacobian term in the acceptance probability given by $|\partial f(\Psi, \boldsymbol{u})/\partial(\Psi, \boldsymbol{u})|$ is simply one, since $f(\Psi, \boldsymbol{u}) = (\Psi, \boldsymbol{u})$. Hence,

$$\alpha_C(\Psi, \Psi') = \min\left\{1, \mathbb{1}(\Psi' \in \boldsymbol{\Psi})\frac{\pi^2\rho_{\mathrm{d}}}{(1 - \rho_{\mathrm{d}})(n + 1)}\exp[V(\Psi, \sigma) - V(\Psi', \sigma)]\lambda\right\}.$$

The death of a crack is performed similarly, with one of the n cracks being proposed for deletion, a move which is subsequently accepted with probability

$$\alpha_{C'}(\Psi, \Psi') = \min\left\{1, \frac{(1 - \rho_{\mathrm{d}})n}{\pi^2\rho_{\mathrm{d}}}\exp[V(\Psi, \sigma) - V(\Psi', \sigma)]/\lambda\right\}.$$

4.4. Updating the variance

The moves that we have so far described deal with movements in $\boldsymbol{\Psi}$. However, the error variance σ^2 is also a parameter that requires updating. We assume a priori that the precision $\tau = 1/\sigma^2$ is Gamma distributed with parameters κ_1 and κ_2, in which case the posterior conditional density for τ is given by

$$\pi(\tau \mid \Psi, \Phi) = \tau^{\kappa_1 + n - 1}\exp\left[-\tau\left(\kappa_2 + \sum_{i=1}^{m}[\Phi(t_i) - F\Psi(t_i)]^2/2\right)\right]$$

i.e., the conditional distribution of τ is Gamma distributed with parameters $\kappa_1 + n$ and $\kappa_2 + \sum_{i=1}^{m}[\Phi(t_i) - F\Psi(t_i)]^2/2$. It is easy to show that if we choose this distribution as our proposal, then the corresponding acceptance probability is identically equal to one. This is in fact a special case of the Metropolis–Hastings algorithm, known as a Gibbs sampling update, see Geman and Geman (1984) and Gelfand and Smith (1990). We update the variance in this manner at each iteration.

4.5. Simulated tempering updates

Pilot runs of the simulation appeared to exhibit rather slow mixing behaviour, primarily due to the existence of local maxima in the likelihood so that new cracks can be slow to move to more likely positions. To overcome this problem, the state space of the chain can be augmented to introduce simulated tempering updates which greatly improves sampling performance. These updates were attempted after every 10 iterations so that in temperature τ_i we propose a move to $\tau_{i\pm1}$ with probabilities $p_{i\to i+1} = p_{i\to i-1} = \frac{1}{2}$ for $i = 2, \ldots, L-1$, and $p_{1\to 2} = p_{L\to L-1} = 1$ (see also table 1). This change is accepted with probability

$$\min\left[1, \frac{c_j p_{j\to i}}{c_i p_{i\to j}} \left\{\sigma^{-m} \exp[V(\Psi, \sigma)]\right\}^{1/\tau_j - 1/\tau_i}\right],$$

where τ_j denotes the proposed new temperature ($j = i \pm 1$).

The Markov chain we have proposed here for crack detection was implemented and tested (experimental code-checking simulation can be done by letting $V \equiv 0$, so that the chain is entirely prior driven). Several simulations were performed varying only the ρ-parameters and, since estimates from these different runs are similar, we assume that all are sampling from the same (and correct) stationary distribution. Note that additional diagnostics may be employed, see for example Brooks and Roberts (1998).

5. Results

In this section, we consider the performance of our approach on simulated data similar to that used by Bryan and Vogelius (1994) and discussed in section 2. The simulated data comes from three cracks of which two are only 0.05 units away from the boundary to which 60 electrodes have been attached. The data is perturbed by Gaussian noise with zero mean and variance $\sigma^2 = 0.25$ approximately corresponding to a 10% noise-to-signal ratio. For the simulations, we chose $\rho_{\mathrm{bd}} = 1/4, \rho_{\mathrm{r}} = 2/15$ and $\rho_{\mathrm{d}} = 1/2$, so that the chain spends more time translating and rotating cracks than inserting and deleting them. A tempering scheme based on the temperatures $1, 1.5, 2, 4, 7.5$ and ∞ are used and a temperature change is proposed after every ten iterations. For the prior we choose $\lambda = 0.37$. To efficiently explore $\boldsymbol{\Theta}$ we suggest proposing 'bigger' moves in higher temperatures and different values of Δl and $\Delta\theta$ are therefore used. Table 1 provides these values together with the prior temperature probabilities which are based upon a series of pilot runs where the probabilities are adjusted from run to run to obtain values that make the chain visit the different temperatures roughly equally often, see Geyer and Thompson (1995). The temperature scheme used here and described in table 1 has been tailored (via pilot tuning) to this particular problem and may not be suitable for other data sets. In general some degree of pilot tuning is necessary in order for the simulations to work well, see Geyer and Thompson (1995) who describe a variety of more sophisticated methods for choosing both the prior temperature probabilities and the spacing between successive temperatures. These methods may be of most use when handling problems with densely packed non-intersecting cracks. The MCMC algorithm

was started with $\sigma^2 = 1$ in the same initial crack configuration used by Bryan and Vogelius (1994) and ran for 50,000 iterations.

Table 1. Prior temperatures and corresponding proposal steps and prior probabilities used in the crack detecting MCMC algorithm. Shown are also an example of observed frequencies after a run of the chain for 50,000 iterations.

Temperature	Δl	$\Delta\theta$	Prior probability	Observed frequencies
1	0.01	$\pi/48$	0.3627	0.162
1.5	0.02	$\pi/24$	0.1874	0.161
2	0.05	$\pi/12$	0.1471	0.188
4	0.10	$\pi/6$	0.1192	0.177
7.5	0.15	$\pi/4$	0.1009	0.150
∞	0.25	$\pi/2$	0.0827	0.162

In the following interpretation of the simulation results we need a partial ordering of the cracks in order to uniquely identify the different cracks in any given configuration. Assume that Ψ is a finite collection of n linear non-intersecting cracks where crack ψ_i is uniquely parameterised by its endpoints (x_1^i, y_1^i) and (x_2^i, y_2^i). A partial ordering of the cracks is then obtained by ordering each crack individually, so that $x_2^i \geq x_1^i$ for $i = 1, \ldots, n$. Moreover, the cracks are ordered globally so that $x_1^1 \geq x_1^2 \geq \cdots \geq x_1^n$.

The output from the chain whilst in the cold distribution is of particular interest, since it is only these that are drawn from $\pi(\Theta \,|\, \Phi)$. Figure 5(a) shows the sampled crack configurations from π, whereas figure 5(b) shows the corresponding mean crack configuration with 95% credible intervals. As the cracks are clearly split into three seperate clusters, mean and credible intervals are constructed within each cluster. This method is sufficient for the problem at hand, but for even higher noise levels one may encounter problems with 'overlapping' clusters of cracks. In such cases an intensity plot of the posterior probabilities for the cracks passing any point in Ω may be more appropriate. Each figure is based upon 8100 samples from π. The corresponding posterior means are given in table 2 together with 95% credible intervals.

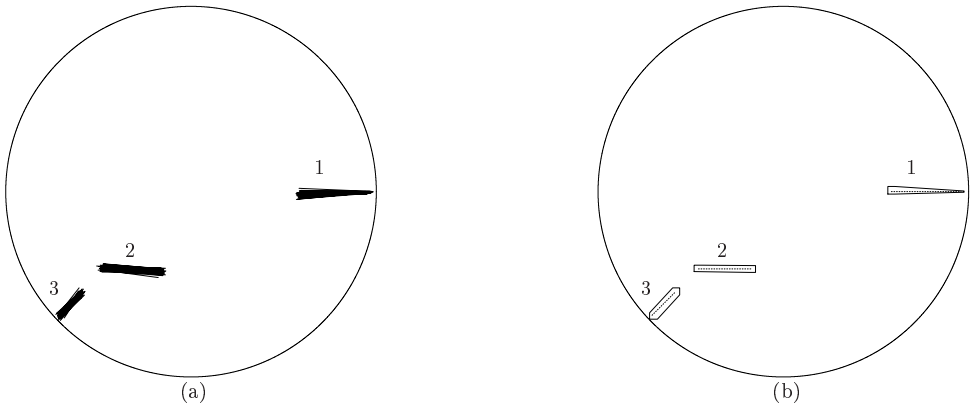

Figure 5. (a) Samples from $\pi(\Psi \,|\, \Phi)$; and (b) estimated cracks with 95% credible intervals (credible intervals:——, true cracks:· · · · · ·).

Table 2. The resulting posterior means, standard deviances and credible intervals for the locations of the three cracks, together with the error variance.

			95% C.I.		
Parameter	Estimate	St.d.	Lower	Upper	Truth
x_1^1	0.5827	0.0103	0.5685	0.6089	0.5833
y_1^1	0.0076	0.0108	−0.0383	0.0039	0.0000
x_2^1	0.9475	0.0093	0.9456	0.9820	0.9500
y_2^1	0.0002	0.0032	−0.0087	0.0037	0.0000
x_1^2	−0.4608	0.0103	−0.4957	−0.4555	−0.4583
y_1^2	−0.4140	0.0104	−0.4339	−0.3932	−0.4167
x_2^2	−0.1678	0.0099	−0.1790	−0.1401	−0.1667
y_2^2	−0.4174	0.0101	−0.4502	−0.4106	−0.4167
x_1^3	−0.7009	0.0097	−0.7244	−0.6863	−0.7083
y_1^3	−0.6715	0.0098	−0.6866	−0.6482	−0.6667
x_2^3	−0.5783	0.0093	−0.6088	−0.5725	−0.5833
y_2^3	−0.5381	0.0098	−0.5634	−0.5250	−0.5417
σ^2	0.2785	0.0363	0.2073	0.3496	0.2500

In figure 6 we give the trace plots for the temperatures, the number of cracks, the first coordinate of the first crack (x_1^1) and for the variance σ^2. From figure 6(a) it is apparent that the chain divides its time roughly equally between the 6 different samplers. A trace plot for the number of cracks is shown in figure 6(b). We can see that the chain appears to move fairly freely between having zero and three cracks. Actually, the posterior distribution is concentrated on configurations with three cracks. Figure 6(c) shows the trace plot for the first coordinate of the first crack. Finally, figure 6(d) shows the trace plots for the variance which exhibits excellent mixing properties.

6. Discussion

In this paper we have developed a pratical Bayesian approach for the reconstruction of an unknown set of linear cracks based upon electrostatic boundary measurements. The reconstruction technique is an extension of a very efficient algorithm developed by Bryan and Vogelius (1994). The advantage of the Bayesian approach is that uncertainty about the number of cracks is a part of the modelling process and thereby the use of ad hoc techniques is avoided. Moreover, the method allows for statistical inference, in the sense that uncertainty about the estimates can be evaluated. The Bayesian reconstruction technique has been tested on a substantial number of different crack configurations data and relatively precise reconstructions were provided, even under high noise levels.

Consequently the Bayesian reconstruction technique seems rather promising and very robust and it would therefore be interesting to test it on real crack data, as well as apply it to other problems. However, we note that the method is computationally intensive. For comparison it took 40 minutes to run $50,000$ simulations on a shared 400 MHz RISC processor with 1024 MB memory. Nevertheless, by constructing a proposal

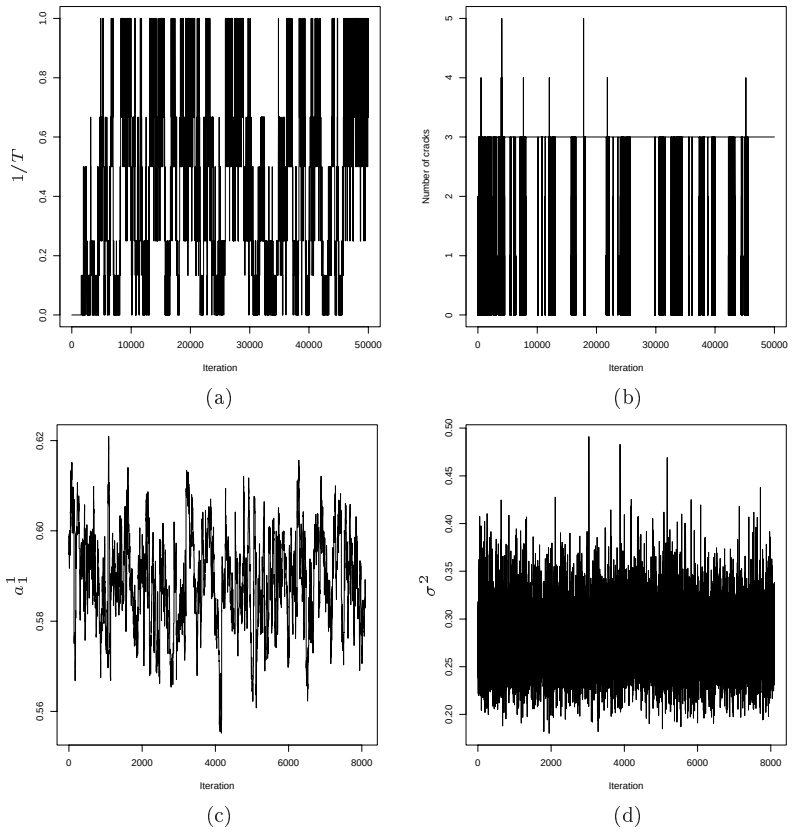

Figure 6. Trace plots from the Markov chain: (a) The visited temperatures; (b) number of cracks, (c) the trace of the first coordinate in the first crack's endpoint in the cold distribution; and (d) the error variance.

distribution based on the gradient of the forward map much more efficient Monte Carlo methods may be derived. Since the gradient already has been provided by Bryan and Vogelius (1994), this modified algorithm would be fairly easy to implement.

Acknowledgement

The work of MBH was (partially) supported by MaPhySto — Centre for Mathematical Physics and Stochastics, funded by a grant from the Danish National Research Foundation.

References

Atkinson, K.E. (1976), *A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind*, SIAM, Philadelphia, PA.

Barber, D. and Brown, B. (1986), Recent developments in applied potential tomography – APT, *in* S. Bacharach, ed., 'Information Processing in Medical Imaging', (Nijhoff, Amsterdam), pp. 106 – 121.

Brooks, S.P. (1998), 'Markov chain Monte Carlo method and its application', *The Statistician* **47**, 69 – 100.

Brooks, S.P. and Roberts, G.O. (1998), 'Assessing convergence of Markov chain Monte Carlo algorithms', *Statistics and Computing* **8**, 319 – 335.

Bryan, K. and Vogelius, M. (1994), 'A computational algorithm to determine crack locations from electrostatic boundary measurements. The case of multiple cracks', *Int. J. Engng Sci.* **32**, 579 – 603.

Cheney, M., Isaacson, D. and Newell, J.C. (1998), 'Electrical impedance tomography', *SIAM Review* **41**(1), 85 – 101.

Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Jensen, A. W. and Balling, N. (1998), 'Past temperatures directly from the Greenland ice sheet', *Science* **282**, 268–271.

Daley, D. J. and Vere-Jones, D. (1988), *An Introduction to the Theory of Point Processes*, Springer Verlag, New York.

Efromovich, S. (1999), *Nonparametric Curve Estimation*, Springer Verlag, New York.

Gelfand, A.E. and Smith, A.F.M. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**, 398 – 409.

Geman, S. and Geman, D. (1984), 'Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721 – 741.

Geyer, C. J. and Thompson, E. A. (1995), 'Annealing Markov chain Monte Carlo with applications to ancestral inference', *J. Am. Statist. Ass.* **90**, 909 – 920.

Green, P. (1995), 'Reversible jump Markov chain Monte Carlo computation and Bayesian model determination', *Biometrika* **82**, 711 – 32.

Hadamard, J. (1923), 'Lectures on the Cauchy Problem in Linear Partial Differential Equations', Yale University Press, New Haven.

Hastings, W. K. (1970), 'Monte Carlo sampling methods using Markov chains and their applications', *Biometrika* **57**, 97 – 109.

Marini, E. and Parisi, G. (1992), 'Simulated tempering: a new Monte Carlo scheme', *Europhys. Lett.* **19**, 451 – 458.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953), 'Equations of state calculations by fast computing machines', *J. of Chem. Phys.* **21**, 1087 – 1092.

Mosegaard, K. (1998), 'Resolution analysis of general inverse problems through inverse monte carlo sampling', *Inverse Problems* **14**, 405–426.

Mosegaard, K. and Rygaard-Hjalsted, C. (1999), 'Probalistic analysis of implicit inverse problems', *Inverse Problems* **15**, 573–583.

Mosegaard, K. and Tarantola, A. (1995), 'Monte carlo sampling of solutions to inverse problems', *Journal of Geophysical Research* **B 100**, 12431–12447.

Nicholls, G. K. and Fox, C. (1998), Prior modelling and posterior sampling in impedance imaging, *in* A. Mohammad-Djafari, ed., 'Proceeding of SPIE', Vol. 3459, pp. 116–127.

Robert, C. P. and Casella, G. (1999), *Monte Carlo Statistical Methods*, Springer-Verlag, New York.

Santosa, F. and Vogelius, M. (1990), 'A backprojection algorithm for electrical impedance imaging', *SIAM J. Appl. Math* **50**, 216 – 243.

Tarantola, A. (1987), *Inverse Problem Theory*, Elsevier, Amsterdam.

Tikhonov, A. N. (1963*a*), 'Regularization of incorrectly posed problems', *Sov. Doklady* **4**, 1624 – 1627.

Tikhonov, A. N. (1963*b*), 'Solution of incorrectly formulated problems and the regularization method', *Sov. Doklady* **4**, 1035 – 1038.