



UNIVERSITY OF LEEDS

This is a repository copy of *A fully Bayesian approach to shape estimation of objects from tomography data using MFS forward solutions*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/79318/>

Version: Accepted Version

Article:

Aykroyd, RG, Lesnic, DL and Karageorghis, A (2015) A fully Bayesian approach to shape estimation of objects from tomography data using MFS forward solutions. *International Journal of Tomography and Simulation*, 28 (1). 1 - 21. ISSN 0972-9976

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A fully Bayesian approach to shape estimation of objects from tomography data using MFS forward solutions

Robert G. Aykroyd¹, Daniel Lesnic² and Andreas Karageorghis³

¹Department of Statistics
University of Leeds
Leeds, LS2 9JT, UK
r.g.aykroyd@leeds.ac.uk

²Department of Applied Mathematics
University of Leeds
Leeds, LS2 9JT, UK
D.Lesnic@leeds.ac.uk

³Department of Mathematics and Statistics
University of Cyprus
Nicosia, Cyprus
andreak@ucy.ac.cy

ABSTRACT

It is possible to characterize the aim of many practical inverse geometric problems as one of identifying the shape of an object within some domain of interest using non-intrusive measurements collected on the boundary of the domain. In the problem considered here the object is a rigid inclusion within a homogeneous background medium of constant conductivity, and the data are potential and current flux measurements made on the boundary of the region. The rigid inclusion is described using a geometric parametrization in terms of a star-shaped object. A Bayesian modelling approach is used to combine data likelihood and prior information, and posterior estimation is based on a Markov chain Monte Carlo algorithm which provides measures of uncertainty, as well as point estimates. This means that the inverse problem is never solved directly, but the cost is that instead the forward solution must be found many thousands of times. The forward problem is solved using the method of fundamental solutions (MFS) which is an efficient meshless alternative to the more common finite element or boundary element methods. This paper is the first to apply Bayesian modelling to a problem using the MFS, with numerical results demonstrating that for appropriate choices of prior distributions accurate results are possible. Further, it demonstrates that a fully Bayesian approach is possible where all prior smoothing parameters are estimated. It is important to note that the geometric modelling and statistical estimation approach are not limited to this example and hence the general technique can be easily applied to other inverse problems. A great benefit of the approach is that it allows an intuitive model description and directly interpretable output. The methods are illustrated using numerical simulations.

1 INTRODUCTION

Inverse problems occur in a wide range of practical applications in geophysics, industry and medicine – see Stuart (2010) for a Bayesian perspective of inverse problems. For example in electrical tomography, voltages are recorded between multiple electrode-pairs attached to the boundary and the aim is to reconstruct the interior conductivity distribution – a review of statistical modelling for such examples can be found in Watzenig and Fox (2009). The standard method of analysis involves domain discretization and the use of the finite element method. This, however, inevitably leads to an ill-posed inverse problem demanding regularization. For examples of this approach to electrical impedance tomography (EIT), see West et al. (2004; 2005) and references therein. In the following sections an alternative approach is proposed. A parametric model of the inclusion will be defined and brief details of the method of fundamental solutions (MFS) will be given. Then, Bayesian statistical modelling will be discussed with specific examples given and an outline of the Markov chain Monte Carlo (MCMC) method presented – for a detailed theoretical discussion of the MCMC method see, for example, Geyer (2011) and Brooks et al. (2011). To demonstrate the proposed approach a series of numerical simulations are described which highlight the flexibility of the modelling and estimation procedures.

2 MATHEMATICAL MODELLING

Suppose that there is an unknown two-dimensional star-shaped object, D , represented by radii, $\mathbf{r} = (r_i)_{i=1, \overline{M}}$, at fixed angles, $\boldsymbol{\theta} = (\theta_i)_{i=1, \overline{M}}$, which is compactly contained in a given body Ω such that $\Omega \setminus D$ is connected – see Figure 1. This model has been used for similar problems in Aykroyd and Cattle (2006; 2007). The data, looking ahead to our example, consist of potential, $\mathbf{y} = (y_j)_{j=1, \overline{N}}$, and current flux measurements, $\mathbf{z} = (z_j)_{j=1, \overline{N}}$, recorded on the outer fixed boundary, $\partial\Omega$, at locations $\mathbf{x} = (x_j)_{j=1, \overline{N}}$. The aim of the method is to use the measured data, (\mathbf{y}, \mathbf{z}) , to estimate the unknown radii, \mathbf{r} .

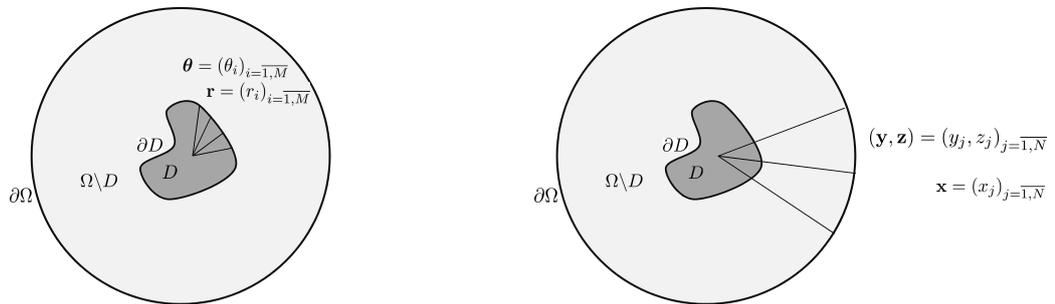


Figure 1: Diagram of star-shaped object model (left) and data measurements (right).

The data model defines the measurements on $\partial\Omega$ in terms of exact values of the potential, u , and the current flux, $\partial u / \partial \mathbf{n}$, combined with additive Gaussian noise, that is,

$$y_j = u(x_j) + \epsilon_j, \quad z_j = \frac{\partial u}{\partial \mathbf{n}}(x_j) + \zeta_j, \quad j = \overline{1, \overline{N}}, \quad (2.1)$$

where \mathbf{n} is the outer unit normal to the boundary $\partial\Omega$, and the noise $(\epsilon_j)_{j=\overline{1,N}}$ and $(\zeta_j)_{j=\overline{1,N}}$ follow independent normal distributions, with zero means and variances σ_y^2 and σ_z^2 , respectively, and u satisfies the Laplace equation in $\Omega \setminus D$. Further, if D is a rigid inclusion then $u = 0$ on ∂D , otherwise if D is a cavity then $\partial u / \partial \mathbf{n} = 0$ on ∂D . We can also have that D is an inclusion with a different conductivity than that of the background $\Omega \setminus D$ in which case transmission conditions are applied at the interface ∂D .

The values of the potential and current flux on $\partial\Omega$ are calculated using the MFS, see Borman et al. (2009) and Karageorghis et al. (2011; 2013), as a linear combination of fundamental solutions of the governing Laplace equation

$$u(\mathbf{c}, \boldsymbol{\xi}, x_j) = \sum_{k=1}^{2M} c_k G(\xi_k, x_j), \quad \frac{\partial u}{\partial \mathbf{n}}(\mathbf{c}, \boldsymbol{\xi}, x_j) = \sum_{k=1}^{2M} c_k \frac{\partial G}{\partial \mathbf{n}}(\xi_k, x_j), \quad j = \overline{1, N}, \quad (2.2)$$

where $G(\xi, x) = -\frac{1}{2\pi} \log |\xi - x|$ is the fundamental solution in two-dimensions of the governing Laplace equation and $\boldsymbol{\xi} = (\xi_k)_{k=\overline{1, 2M}}$ are sources which are located on pseudo-boundaries inside the rigid object D and outside the outer fixed boundary $\partial\Omega$. We also need to impose that D is a rigid inclusion, that is $u = 0$ on ∂D , which can be rewritten as

$$\sum_{k=1}^{2M} c_k G(\xi_k, (r_i \cos(\theta_i), r_i \sin(\theta_i))) = 0, \quad i = \overline{1, M}. \quad (2.3)$$

Notice that the MFS introduces an additional $2M$ unknown coefficients, $\mathbf{c} = (c_k)_{k=\overline{1, 2M}}$, which must be estimated in addition to the M radii, $\mathbf{r} = (r_i)_{i=\overline{1, M}}$ from the system given by equations (2.3) and those obtained by fitting (2.2) to match the Cauchy data measurements (2.1), that is,

$$\sum_{k=1}^{2M} c_k G(\xi_k, x_j) = y_j, \quad j = \overline{1, N}, \quad (2.4)$$

and

$$\sum_{k=1}^{2M} c_k \frac{\partial G}{\partial \mathbf{n}}(\xi_k, x_j) = z_j, \quad j = \overline{1, N}. \quad (2.5)$$

A geometric nonlinear constraint that D is compactly contained in Ω can also be imposed. Altogether, equations (2.3)–(2.5) form a system of $(2N + M)$ equations with $3M$ unknowns. Out of these equations, (2.4) and (2.5) are linear in \mathbf{c} , whilst equation (2.3) represents nonlinear equations. The tomographic inverse rigid inclusion problem is nonlinear and ill-posed, but provided $u|_{\partial\Omega} \neq 0$ the solution is unique (Haddar and Kress, 2005). The solution may not exist, but even if the solution exists it is not stable with respect to the noise in the Cauchy data measurements defined in equation (2.1).

3 STATISTICAL MODELLING

In this section models for the noise process and for prior knowledge will be proposed. These will define a likelihood and a prior distribution, which are combined using Bayes theorem to produce a posterior distribution which is the basis for estimation. For background to Bayesian

modelling, see Gelman et al. (2003), and for applications of Bayesian modelling in electrical tomography problems, see West et al. (2004; 2005) and Aykroyd and Cattle (2006; 2007).

With data (\mathbf{y}, \mathbf{z}) , and assuming (conditional) independence of \mathbf{y} and \mathbf{z} given \mathbf{r} and \mathbf{c} , then the appropriate form of the likelihood is:

$$l(\mathbf{y}, \mathbf{z} | \mathbf{r}, \mathbf{c}) = l(\mathbf{y} | \mathbf{r}, \mathbf{c}) \times l(\mathbf{z} | \mathbf{r}, \mathbf{c}). \quad (3.1)$$

The likelihood quantifies both the inaccuracies in the measuring equipment and other uncontrolled influences. From (3.1), the likelihood of \mathbf{y} given \mathbf{r} and \mathbf{c} is

$$l(\mathbf{y} | \mathbf{r}, \mathbf{c}) = (2\pi\sigma_y^2)^{-N/2} \exp \left\{ -\frac{1}{2\sigma_y^2} \|\mathbf{y} - \hat{\mathbf{y}}(\mathbf{r}, \mathbf{c})\|^2 \right\}, \quad (3.2)$$

where $\hat{\mathbf{y}}(\mathbf{r}, \mathbf{c}) = (\hat{y}_j(\mathbf{r}, \mathbf{c}))_{j=1, \dots, N}$ are fitted values assuming inclusion radii \mathbf{r} and MFS coefficients \mathbf{c} . The structure of the likelihood of \mathbf{z} given \mathbf{r} and \mathbf{c} is identical to (3.2), except that \mathbf{z} replaces \mathbf{y} , $\hat{\mathbf{z}}$ replaces $\hat{\mathbf{y}}$ and σ_z^2 replaces σ_y^2 .

Estimating from the likelihood alone may not be possible due to the non-linear relationship between the radii, \mathbf{r} , and the data, and the ill-posed nature of the problem in terms of the MFS coefficients, \mathbf{c} . In a standard approach, progress can be made by imposing regularization. This leads to a numerical approach which will produce point estimates, but there will be no information about confidence, that is, about the precision of the point estimates. Here an alternative approach is adopted based on the widely used Bayesian modelling framework. The key addition to the modelling is to consider *prior distributions* for the model parameters which quantify specific expert opinion or more vague knowledge of the relative ranking of the various alternatives.

It is assumed that there is some knowledge of the values, or relationship between, the model parameters \mathbf{r} and \mathbf{c} . In the examples considered here we expect the boundary to vary gently around the object, which suggests smoothing, leading to a prior distribution such as

$$\pi(\mathbf{r} | \beta_r) = \frac{1}{(2\pi\beta_r^2)^{M/2}} \exp \left\{ -\frac{1}{2\beta_r^2} \|\nabla \mathbf{r}\|_2^2 \right\}, \quad (3.3)$$

which uses a 2-norm, and hence corresponds to a Gaussian distribution, or

$$\pi(\mathbf{r} | \beta_r) = \frac{1}{(2\beta_r)^M} \exp \left\{ -\frac{1}{\beta_r} \|\nabla \mathbf{r}\|_1 \right\}, \quad (3.4)$$

which uses the 1-norm, and hence gives a Laplace distribution. In each case, $\|\nabla \mathbf{r}\|_q^q = \sum |r_i - r_{i-1}|^q$, with periodic boundary, and β_r defines the amount of variability between adjacent radii.

The whole prior modelling can be repeated for the MFS coefficients producing prior distribution, $\pi(\mathbf{c} | \beta_c)$, for \mathbf{c} . Here, the same first-order smoothing prior distribution will be used, but the range of alternatives is still available and there is no requirement for this to be of the same type as for the radii.

Bringing the likelihood functions and prior distributions together gives the corresponding *posterior distribution* as the product of likelihood and prior distribution

$$\pi(\mathbf{r}, \mathbf{c} | \mathbf{y}, \mathbf{z}) \propto l(\mathbf{y} | \mathbf{r}, \mathbf{c}) l(\mathbf{z} | \mathbf{r}, \mathbf{c}) \times \pi(\mathbf{r} | \beta_r) \pi(\mathbf{c} | \beta_c). \quad (3.5)$$

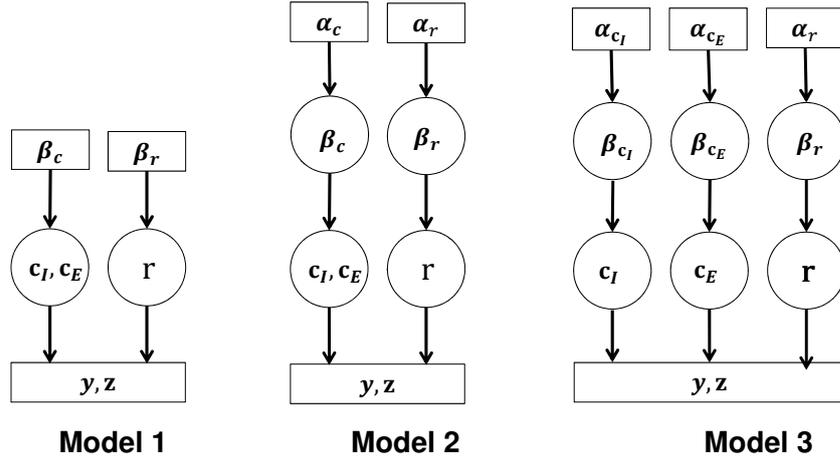


Figure 2: Hierarchical relationship between data, model parameters and hyper-parameters.

The hierarchical structure of this model is represented in the directed graph in Figure 2 (left). The boxed variables are fixed data and prior parameters whereas the circled variables are to be estimated. The arrows indicate causal relationships.

Now, of course, the prior parameters, β_r and β_c , are also unknown and hence should be included in the modelling process. Here the *hyper-prior distribution* for β_r is taken as

$$\pi(\beta_r) = \alpha_r^2 \exp \{ -\alpha_r^2 / \beta_r^2 \}. \quad (3.6)$$

This is an example of the, widely used, inverse-gamma prior for a variance parameter (Gelman, 2006). The value of the *hyper-parameter*, α_r , can be fixed at a reasonable value chosen during initial trials. In addition, there will be a similar prior, $\pi(\mathbf{c}|\beta_c)$, for \mathbf{c} and hyper-prior distribution for β_c with a hyper-parameter α_c . This leads to the full posterior distribution as the product of likelihood, prior and hyper-prior distributions

$$\pi(\mathbf{r}, \mathbf{c}, \beta_r, \beta_c | \mathbf{y}, \mathbf{z}) \propto l(\mathbf{y}|\mathbf{r}, \mathbf{c})l(\mathbf{z}|\mathbf{r}, \mathbf{c}) \times \pi(\mathbf{r}|\beta_r)\pi(\beta_r) \times \pi(\mathbf{c}|\beta_c)\pi(\beta_c). \quad (3.7)$$

Figure 2 (centre) illustrates the hierarchical relationship between the model variables.

Taking the modelling one final step further, it is entirely reasonable to allow separate prior distributions, $\pi(\mathbf{c}_I|\beta_{c_I})$ and $\pi(\mathbf{c}_E|\beta_{c_E})$, for the two sets of MFS coefficients in (2.2), that is, $\mathbf{c}_I = (c_k)_{k=\overline{1, M}}$, those associated with the interface ∂D , and $\mathbf{c}_E = (c_k)_{k=\overline{M+1, 2M}}$, those associated with the outer boundary $\partial\Omega$. This then also suggests corresponding separate hyper-prior distributions $\pi(\mathbf{c}_I)$ and $\pi(\mathbf{c}_E)$, with separate hyper-prior parameters, α_{c_I} and α_{c_E} . Again these hyper-prior parameters will be fixed at reasonable values chosen during initial trials. The resulting posterior distribution is again the product of likelihood, prior and hyper-prior distributions

$$\pi(\mathbf{r}, \mathbf{c}, \beta_r, \beta_{c_I}, \beta_{c_E} | \mathbf{y}, \mathbf{z}) \propto l(\mathbf{y}|\mathbf{r}, \mathbf{c})l(\mathbf{z}|\mathbf{r}, \mathbf{c}) \times \pi(\mathbf{r}|\beta_r)\pi(\beta_r) \times \pi(\mathbf{c}_I|\beta_{c_I})\pi(\beta_{c_I}) \times \pi(\mathbf{c}_E|\beta_{c_E})\pi(\beta_{c_E}). \quad (3.8)$$

The hierarchical structure of this model is illustrated in Figure 2.

4 MARKOV CHAIN MONTE CARLO ESTIMATION

The Markov chain Monte Carlo (MCMC) approach is now widely used for many Bayesian statistical estimation problems in situations where model complexity and parameter dimensionality make other procedures infeasible – see, for example, Gamerman and Lopes (2006) and Liu (2008). The procedure has come to mean much more than an alternative numerical method. In particular, the approach allows a deeper exploration of the posterior distribution than permitted by other approaches.

The MCMC approach gives a framework which can be used to design tailor-made iterative algorithms for many estimation problems. In particular, a resulting algorithm is used to produce a correlated sample from some target statistical distribution – usually the posterior distribution in a Bayesian analysis. Specifically, the transitions in the Markov chain are designed so that an equilibrium distribution exists and is equal to the target distribution. If the transitions are designed well, then after an initial transient period, referred to as burn-in, the remaining sample will have the same statistical properties as a sample obtained directly from the posterior distribution. The only exception is that, by the very nature of a Markov chain, there will be correlation within the sample which must be taken into account when the algorithm output is summarised. If transitions are designed badly however, then the initial transient period could be long and the within sample correlation could be high. This means that the algorithm is inefficient and would require larger samples to achieve acceptable accuracy and precision.

Our particular implementation is now described. Suppose, that all the model parameters are stored in a single vector, $\Theta = (\Theta_i)_{i=\overline{1,p}}$. Examples of this are, $\Theta = (\mathbf{r}, \mathbf{c})$, $\Theta = (\mathbf{r}, \mathbf{c}, \beta_r, \beta_c)$ and $\Theta = (\mathbf{r}, \mathbf{c}, \beta_r, \beta_{c_I}, \beta_{c_E})$ – these are the three cases illustrated in Figure 2. Starting from an arbitrary value, Θ^0 , K random walk transition steps are performed based on Gaussian perturbations. At each step, $k = \overline{1, K}$, the proposed value is accepted with a probability which depends on a posterior ratio. The algorithm is summarised in Figure 3. The statement and implementation of the algorithm are straightforward and a sensible choice for the variance, τ^2 , in the proposal distributions can be made from initial experimentation.

```

Set an initial value for  $\Theta = (\Theta)_{i=\overline{1,p}}$ , call this  $\Theta^0$ 
Repeat the following steps for  $k = \overline{1, K}$ 
  Repeat the following steps for  $i = \overline{1, p}$ 
    Propose new value  $\Theta_i^k = \Theta_i^{k-1} + \mathcal{N}(0, \tau^2)$ 
    Evaluate  $\alpha = \min \{1, \pi(\Theta^k | \mathbf{y}, \mathbf{z}) / \pi(\Theta^{k-1} | \mathbf{y}, \mathbf{z})\}$ 
    Generate  $u$  from a uniform distribution,  $U(0, 1)$ 
    If  $\alpha > u$  then accept the proposal, otherwise reject and set  $\Theta^k = \Theta^{k-1}$ 
  End repeat
End repeat
Discard initial values and use remainder to make inference.

```

Figure 3: Random-walk Metropolis-Hastings algorithm.

As this is a very simple estimation problem, an equally simple random walk proposal is very likely to work well. When considering more complex estimation problems, particularly with many parameters, more careful consideration may be needed. The efficiency of the algorithm, however, is heavily dependent on the choice of the proposal scheme.

When choosing a value for τ^2 , it is important to realise that both low and high values lead to long transient periods and highly correlated samples and hence unreliable estimation. A reasonable proposal variance can be chosen adaptively during the early burn-in period, and it has been proven theoretically that for a wide variety of high-dimensional problems an acceptance rate of 23.4% (Roberts et al., 1997) is optimal. Further, if different types of parameter are being estimated, then it may be appropriate to have a separate proposal variance for each type. Further, it is wise to also check Markov chain paths and to calculate sample autocorrelation functions. For good estimation the paths should look “random” and the autocorrelation functions be close to zero for all except small lags. For suggestions on judging the appropriate size of MCMC samples, and other convergence issues, see Raftery and Lewis (1995), Cowles and Carlin (1996) and Geyer (2011).

Once the sample has been generated from the posterior distribution, a number of possible estimators are available. One choice is the posterior mean, which can be estimated by the mean of the sample collected after a suitable burn-in period to allow the chain to reach equilibrium. The whole MCMC sampling ethos encourages the investigation of a variety of summary measures, and not only mean and variance. Instead the sample can be used to calculate interval estimates using sample percentiles, or in fact the whole of the posterior distribution can be examined. Also, it is usual not to assume normality of the sampling distributions of the various quantities being estimated, but instead the sample histogram is used to estimate the unknown distribution. In the following numerical results section a variety of output will be shown, but as a minimum it is usual to examine the histogram of the sampling distributions and to form credible intervals using the percentage points of the corresponding sampling distribution. For applications of MCMC methods to electrical tomography, see West et al. (2004; 2005) and Aykroyd and Cattle (2006; 2007).

5 NUMERICAL RESULTS

5.1 Preliminary

In this section part of a series of numerical experiments based on simulated data will be reported. Three true object geometries for D will be considered, namely: (i) a circle of radius 0.5 centred at the origin given by the radial parameterization

$$r(\theta) = 0.5, \quad \theta \in [0, 2\pi); \quad (5.1)$$

(ii) a bean-shaped obstacle given by the radial parameterization (Ivanyshyn and Kress, 2006),

$$r(\theta) = \frac{0.5 + 0.4 \cos(\theta) + 0.1 \sin(2\theta)}{1 + 0.7 \cos(\theta)}, \quad \theta \in [0, 2\pi); \quad (5.2)$$

and (iii) a round-cornered rectangle given by the radial parameterization (Ivanyshyn, 2007)

$$r(\theta) = \frac{2}{3} \left[\sin^{10}(\theta) + \left(\frac{2}{3} \cos(\theta) \right)^{10} \right]^{-0.1}, \quad \theta \in [0, 2\pi); \quad (5.3)$$

each of these being contained in the unit disc Ω .

First we determine the current flux data, $\partial u / \partial n$, on $\partial\Omega$ by solving the direct Dirichlet problem

$$\nabla u = 0 \quad \text{in } \Omega \setminus D, \quad (5.4a)$$

$$u = 0 \quad \text{on } \partial D \quad (5.4b)$$

$$u(1, \theta) = \exp(-\cos^2(\theta)) \quad \text{on } \partial\Omega = \{(1, \theta) \mid \theta \in [0, 2\pi)\}, \quad (5.4c)$$

using the MFS with $M = 500$ degrees of freedom. The boundary potential and current flux measurements were then selected at $N = 30$ equally-spaced points on the outer fixed boundary $\partial\Omega$. Data, as defined in equations (2.1), was then produced by addition of Gaussian noise with $\sigma_y = \sigma_z = 0.01$ (corresponding to a signal-to-noise ratio of 1%). We also take $M = 50$ such that the discretised problem defined in equations (2.3)–(2.5) is underdetermined as it contains $M + 2N = 110$ equations with $3M = 150$ unknowns. Of course, by increasing n to 50 or beyond we obtain the determined and the overdetermined situations. For more details of applying the MFS to these three scenarios, see Smyrlis and Karageorghis (2009). The contraction and dilation parameters, η and χ , entering the boundaries $\{\chi r(\theta) \mid \theta \in [0, 2\pi)\}$ and $\eta \partial B(0, 1)$, on which the sources $(\xi_k)_{k=1, 2M}$ are positioned, are taken to be $\eta = 1.8$ and $\chi = 0.9$.

The first section below reports a pilot study to understand the effects of smoothing on the estimation of the MFS coefficients, as well as on the rigid inclusion shape. Then the second section considers full model estimation using Gaussian prior distributions and the final section shows results of full estimation using Laplace prior distributions.

5.2 Understanding the influence of the prior distribution

In the first set of examples the true object, D , is taken as the disk of radius 0.5 centred at the origin as parameterised by equation (5.1). The simplest possible model includes a single unknown radius, r , along with unknown MFS coefficients, c . Figure 4 shows the object reconstructed (left), without prior information, using the radius estimated as the mean of the posterior sample. Also shown are the estimated MFS coefficients (centre and right) surrounded by 95% credible intervals. In all the relevant figures these coefficients are plotted as functions of θ .

The posterior estimate of the radius is 0.4997, compared to the true value of 0.5, with an estimated standard deviation of 0.00107. The estimated MFS coefficients follow the true values, which were obtained from the MFS direct problem solution and are shown in all relevant figures as a continuous dark line, but clearly those associated with the interface source points (centre) show substantially more variability between values and greater uncertainty in the estimates than those associated with the outer boundary points (right).

Now consider the case where prior information on the smoothness of the MFS coefficients is included in the estimation. Figure 5 shows the object reconstructed using fixed prior param-

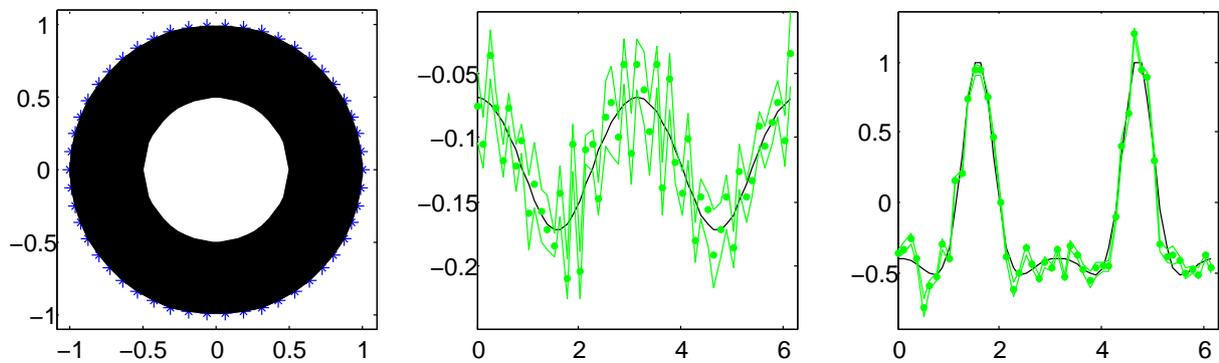


Figure 4: Circular inclusion and circle model fitted with no prior information: fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

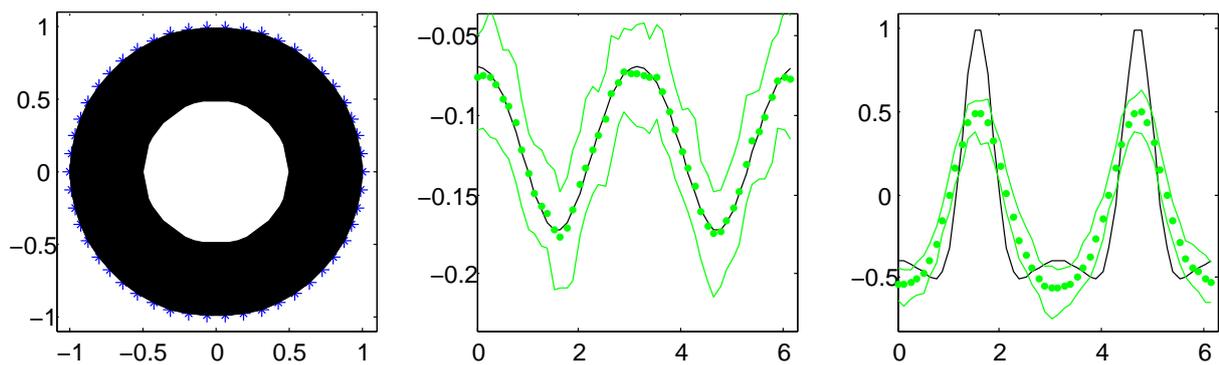


Figure 5: Circular inclusion and circle model with strong prior information ($\beta_{c_I} = \beta_{c_E} = 0.01$): fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

eters $\beta_{c_I} = \beta_{c_E} = 0.01$. The reconstructed object (left) is indistinguishable from the previous reconstruction, but the estimated MFS coefficients are very different to those without prior smoothing. The coefficients for the interface (centre) very closely follow the true values and the credible intervals are reasonably constant in width. Notice, however, that overall the width of the credible intervals has not changed dramatically. For the coefficients associated with the outer boundary (right) the variability between estimates has reduced (for example, focus on the region between 2 and 4), but there is a dramatic bias in the estimated values. In particular, the width and height of the peaks is lost. In summary, the coefficients for the interface are well-estimated with this choice of smoothing parameters, but those for the outer boundary are over-smoothed.

The obvious suggestion is to reduce the amount of smoothing by reducing the value of the smoothing parameters. In another experiment, not shown here, the values $\beta_{c_I} = \beta_{c_E} = 0.1$ were used. The estimates of the coefficients for the interface, however, resemble those without smoothing even though those coefficients for the outer boundary are well estimated. The conclusion from these two experiments is that the coefficients for the interface benefit from more smoothing than those for the outer boundary.

Figure 6 shows the object reconstructed using fixed prior parameters $\beta_{c_I} = 0.01$ and $\beta_{c_E} = 0.1$. Clearly, for both sets of coefficients the estimates closely follow the true values and have narrow credible intervals. As well as producing good object reconstruction the process has also produced accurate coefficient estimates which could be easily described. Hence, we conclude that smoothing of the MFS coefficients is worthwhile, but that it is not appropriate to use the same degree of smoothing for the interface and outer boundary coefficients.

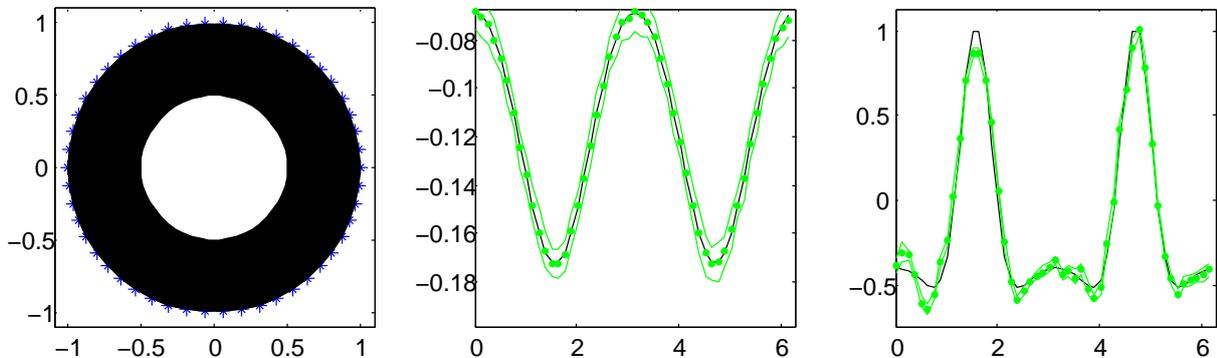


Figure 6: Circular inclusion and circle model with separate prior information ($\beta_{c_I} = 0.01$, $\beta_{c_E} = 0.1$): fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

5.3 Full estimation using Gaussian prior distributions

Consider now the full estimation incorporating the hyper-prior distributions and hence including estimation of the prior parameters β_{c_I} and β_{c_E} . For this, we must specify values for the hyper-prior parameters α_{c_I} and α_{c_E} . Here the values of the fixed smoothing parameter values from the previous experiments have been used, and so $\alpha_{c_I} = 0.01$ and $\alpha_{c_E} = 0.1$. Figures 7 and 8

show summaries from the MCMC estimation.

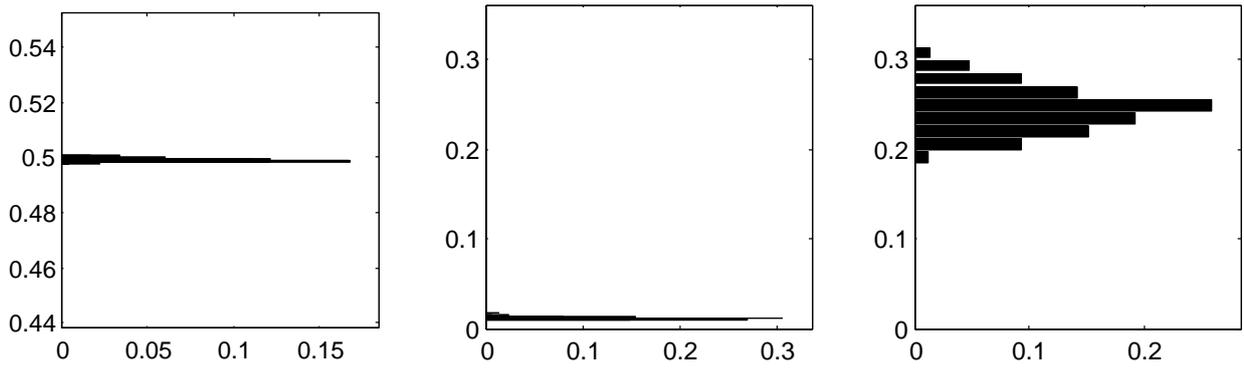


Figure 7: Circular inclusion and circle model with full posterior distribution and separate prior information ($\alpha_{c_I} = 0.01$ and $\alpha_{c_E} = 0.1$): histograms, showing the posterior relative frequency, for radius (left) and MFS interface (centre) and outer boundary coefficients (right).

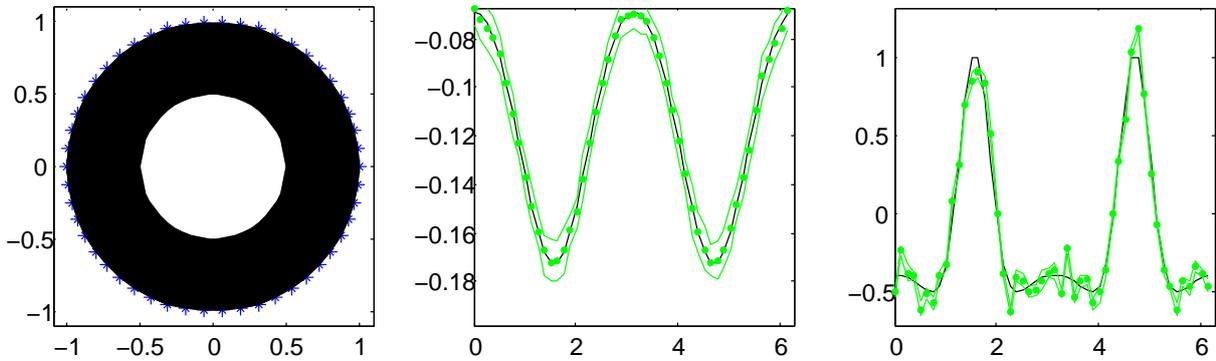


Figure 8: Circular inclusion and circle model with full posterior distribution and separate prior information ($\alpha_{c_I} = 0.01$ and $\alpha_{c_E} = 0.1$): fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

Figure 7 shows posterior histograms for the object radius, r , and for the prior parameters β_{c_I} and β_{c_E} . As summaries of this information, posterior estimates (with standard deviations) are $\hat{r} = 0.4992$ (0.0024), $\hat{\beta}_{c_I} = 0.0116$ (0.0045), and $\hat{\beta}_{c_E} = 0.2457$ (0.0891). Clearly, the variation in the radius is very small indicating that it can be well estimated. Similarly, the smoothing parameter, β_{c_I} , for the interface coefficients has low variability. In contrast, the smoothing parameter for the outer boundary, β_{c_E} , has higher variability and slight positive skew. It is worth noting that the posterior estimate of the smoothing parameter for the interface, β_{c_I} , is close to the prior mean and hence the likelihood has little effect. In contrast, the posterior estimate of the smoothing parameter for the outer boundary, β_{c_E} , is not sensitive to the prior mean value, α_{c_E} .

Figure 8 shows the reconstructed object and estimates of the MFS coefficients with credible intervals. The inclusion remains well estimated and, although the MFS coefficient smoothing parameters are being estimated, the estimates of MFS coefficients have not significantly changed. This demonstrates that it is possible to successfully estimate the prior smoothing parameters and the inclusion shape together without loss of accuracy.

In the next experiment the star-shaped model is used which contains $M = 50$ radii at equally-

spaced angles. The prior parameter for the radii smoothing, β_r was fixed at 1.0, and the prior parameters for the MFS coefficient smoothing were fixed at the posterior estimates from the previous example, that is $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$.

Figure 9 shows the reconstructed inclusion and estimated MFS coefficients. The dataset used is based on a circular true object and so the reconstruction is very accurate. The mean of the posterior radii (with standard deviation) is 0.5012 (0.0012). Similarly, the MFS coefficients are well estimated. It is worth noting that the object reconstruction is not sensitive to the value of the prior parameter, β_r , but the reconstruction is significantly worse if this smoothing is removed completely from the modelling. Accuracy and variability in the object reconstruction are shown in Figure 10. The estimation errors (left), defined as the difference between the estimated and true radii, are indicated by the very thin region around the inner circle. This shows that the estimation errors are very small and are reasonably evenly distributed around the circle. A circular histogram (centre) and circular credible interval (right) aim to describe estimation variability. In the histogram the darker regions indicate the higher frequencies and in the credible interval the thickness of the region indicates the amount of variability. From this, it is clear that the circular histogram tends to exaggerate the slightly non-circular shape of the reconstructed object and hence perhaps the credible interval gives a more reliable representation. These results show that fitting of the star-shaped model to data from a circle truth has been successful.

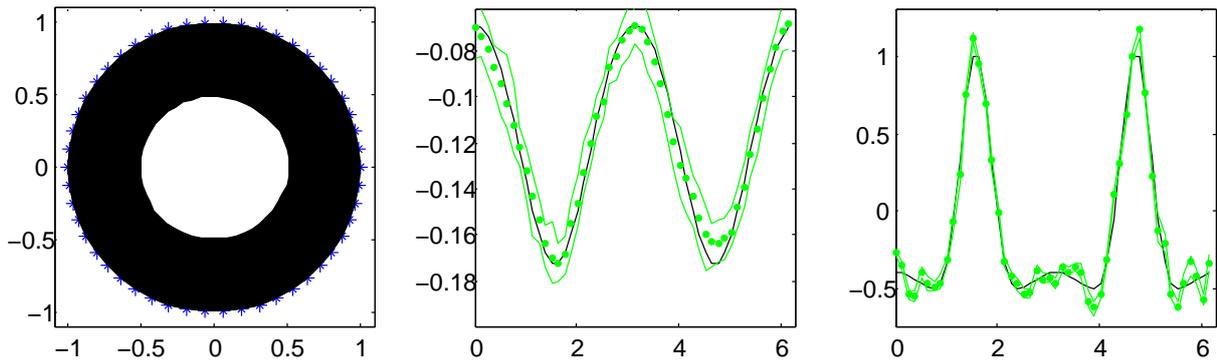


Figure 9: Circular inclusion and star-shaped model with separate prior information ($\beta_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

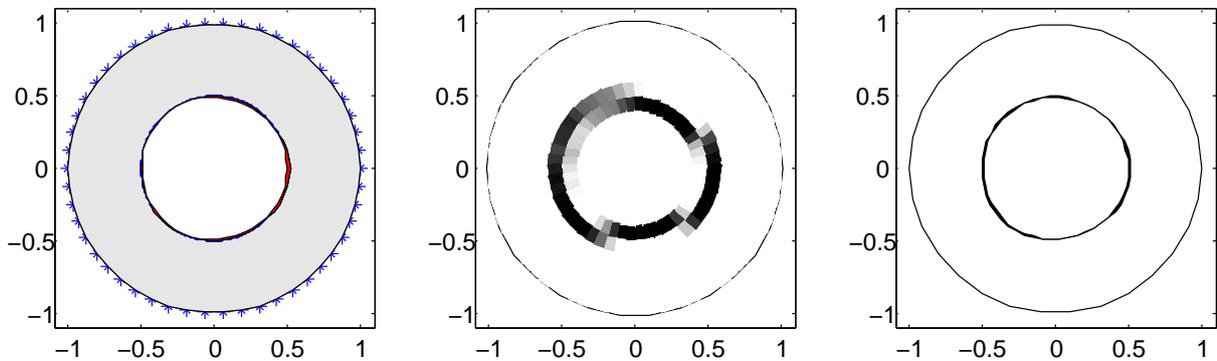


Figure 10: Circular inclusion and star-shaped model with separate prior information ($\beta_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

In the next experiment the star-shaped model is applied for recovering the bean-shaped truth, as defined in equation (5.2), with the prior parameters kept fixed as before. Figure 11 shows the reconstructed object and estimated MFS coefficients for the interface and outer boundary. The rigid bean-shaped inclusion is clearly recovered and MFS coefficients are well estimated without the need for any adjustments.

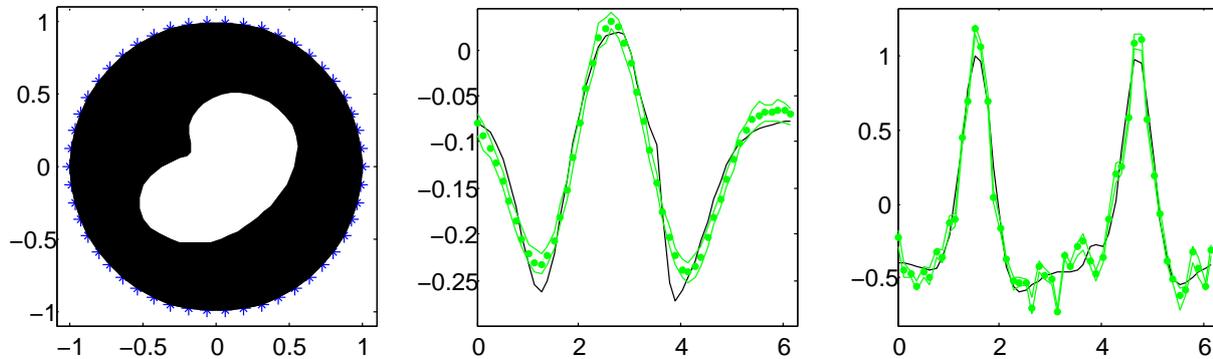


Figure 11: Bean-shaped inclusion and star-shaped model with separate prior information ($\alpha_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

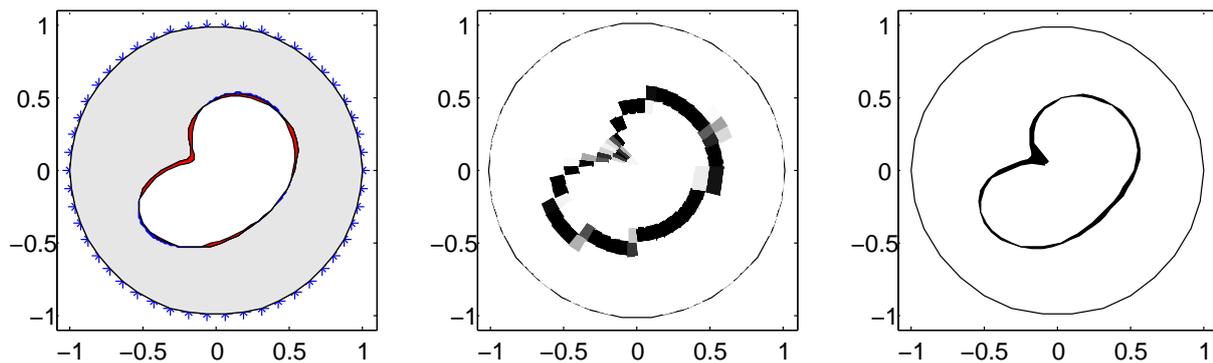


Figure 12: Bean-shaped inclusion and star-shaped model with separate prior information ($\alpha_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

Accuracy and variability in the object reconstruction are shown in Figure 12. The estimation errors (left) are small and are reasonable evenly spread around the boundary. The circular histogram (centre) and circular credible interval (right) indicate that there is greater variability at the “cusp” than elsewhere. This is, however, a very difficult feature to reconstruct accurately and so this estimate can be considered more than acceptable. Overall, the estimation of the star-shaped model to data from the bean-shaped truth has also been very successful.

Now consider the full estimation problem, that is, including estimation of the smoothing parameters of the MFS coefficients. This requires choice of the hyper-prior parameters α_r , α_{c_I} and α_{c_E} and then the estimation of β_r , β_{c_I} and β_{c_E} in addition to the radii and MFS coefficients. In pilot runs, not reported here, it was found that if these hyper-parameters are chosen small enough then good estimation is possible – for example using $\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$.

Figure 13 shows the posterior histograms of the prior parameters, with posterior means (and standard deviations): $\hat{\beta}_r = 0.0383$ (0.0070), $\hat{\beta}_{c_I} = 0.0754$ (0.0322) and $\hat{\beta}_{c_E} = 0.2384$ (0.0651). Figures 14 and 15 show estimated of the inclusion shape and MFS coefficients which clearly indicate less accuracy and greater variability Hence, in this case allowing estimation of the prior parameters has produced a less accurate reconstruction of the shape of the object.

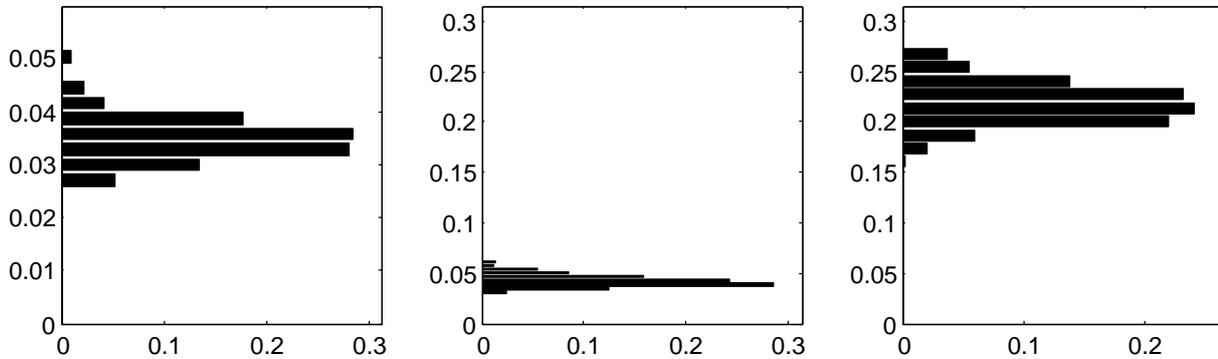


Figure 13: Bean-shaped inclusion and star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$): histograms, showing the posterior relative frequency, for radius (left) and MFS interface (centre) and outer boundary coefficients (right).

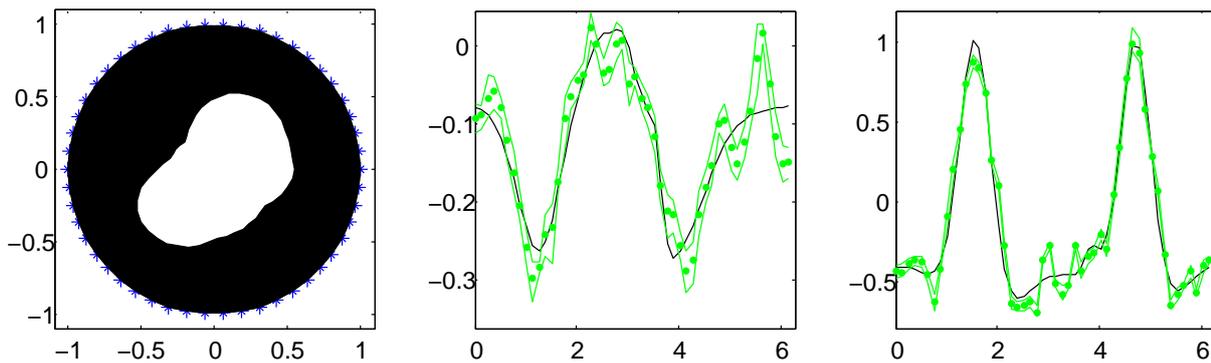


Figure 14: Bean-shaped inclusion and star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$): fitted shape (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

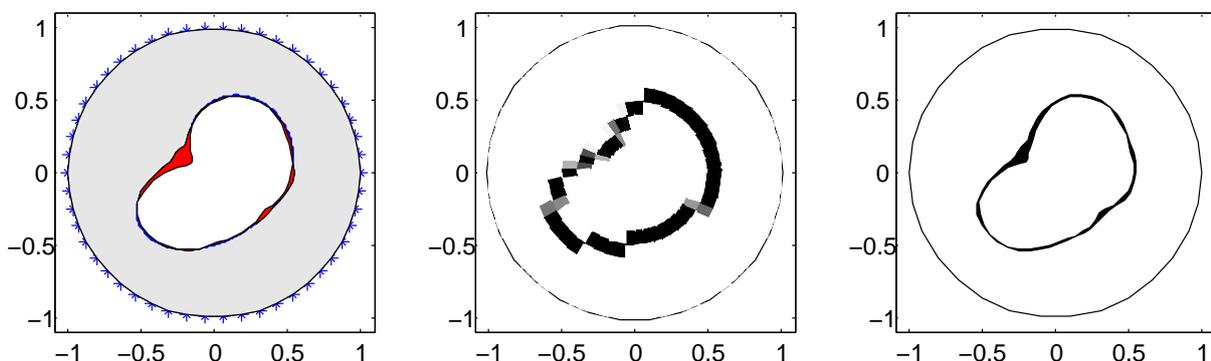


Figure 15: Bean-shaped inclusion and star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$): estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

In the final experiment the star-shaped model is applied for recovering the round-cornered rect-

angle defined in (5.3)—most numerical methods will find this problem challenging because of the relatively sharp corners to the shape. Figure 16 shows the reconstructed object and estimated MFS coefficients for the interface and outer boundary with prior parameters kept fixed as before. The most dramatic change is in the pattern of true MFS coefficient for the interface which is caused by the rounded corners of the rectangle. In spite of this, the reconstruction resembles the truth and the MFS coefficients for the outer boundary are well estimated. The MFS coefficient estimates for the interface however are not good, though they do following the general pattern of the interface coefficients. The posterior mean (and standard deviations) of the radii smoothing parameter is $\hat{\beta}_r = 0.028886$ (0.02519). Accuracy and variability in the object reconstruction are shown in Figure 17. The estimation errors (left) clearly show over-rounding at the corners and bulging in between – the reconstruction is too circular. The circular histogram (centre) and circular credible interval (right) indicate that variability in the posterior distribution is small.

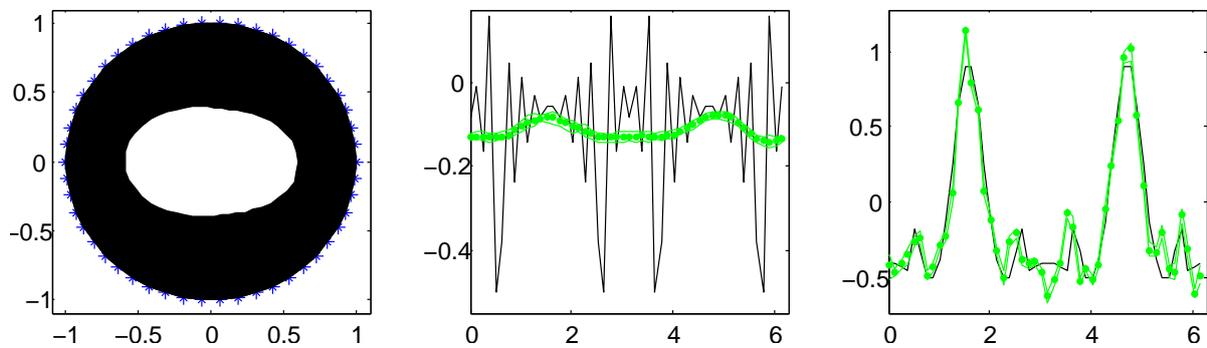


Figure 16: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): fitted shape (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

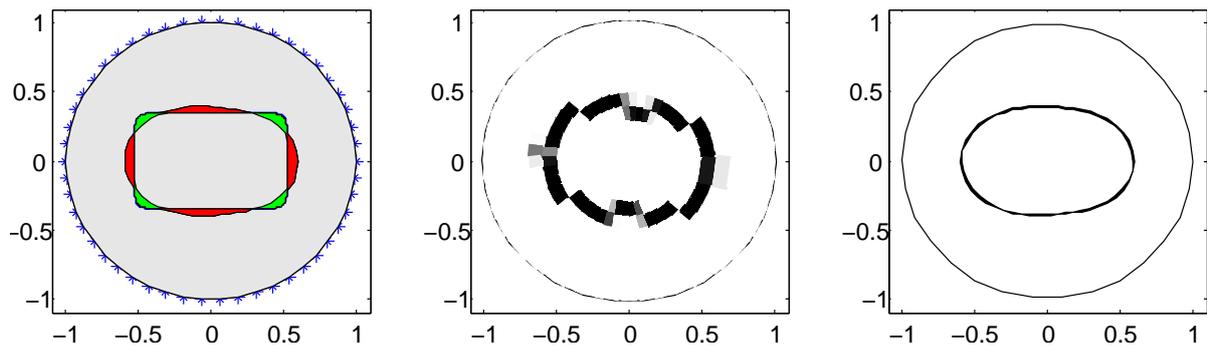


Figure 17: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 1.0$, $\beta_{c_I} = 0.0116$ and $\beta_{c_E} = 0.2457$): estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

Finally, consider the full estimation, including the smoothing parameters of the MFS coefficients, that is, fixing values for hyper-prior parameters α_r , α_{c_I} and α_{c_E} and allowing the estimation of β_r , β_{c_I} and β_{c_E} in addition to the radii and MFS coefficients. Figure 18 shows the reconstructed object and estimated MFS coefficients for the interface and outer boundary. The

rectangular object reconstruction is slightly better, with a useful improvement in the estimation of the interface MFS coefficients, and the coefficients on the outer boundary remain well estimated. The posterior mean estimates of the prior parameters (and standard deviations) are: $\hat{\beta}_r = 0.0307$ (0.0426), $\hat{\beta}_{c_I} = 0.0322$ (0.0301) and $\hat{\beta}_{c_E} = 0.2331$ (0.0941). Accuracy and variability in the object reconstruction are shown in Figure 19. The estimation errors (left) show a slight improvement but still the reconstruction is too circular. The circular histogram (centre) and circular credible interval (right) indicate that the posterior distribution is concentrated, hence this time allowing estimation of the prior parameters has produced a slightly more accurate reconstruction of the shape of the object. As with the bean-shape, this is also a very difficult feature to reconstruct accurately and so this estimate can be considered more than acceptable. Overall, the estimation of the round-cornered rectangular truth has been very successful.

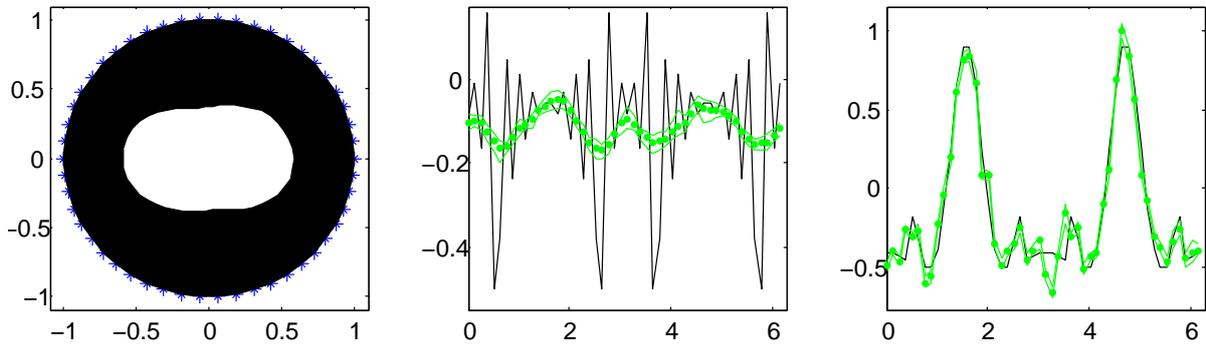


Figure 18: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$): fitted shape (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

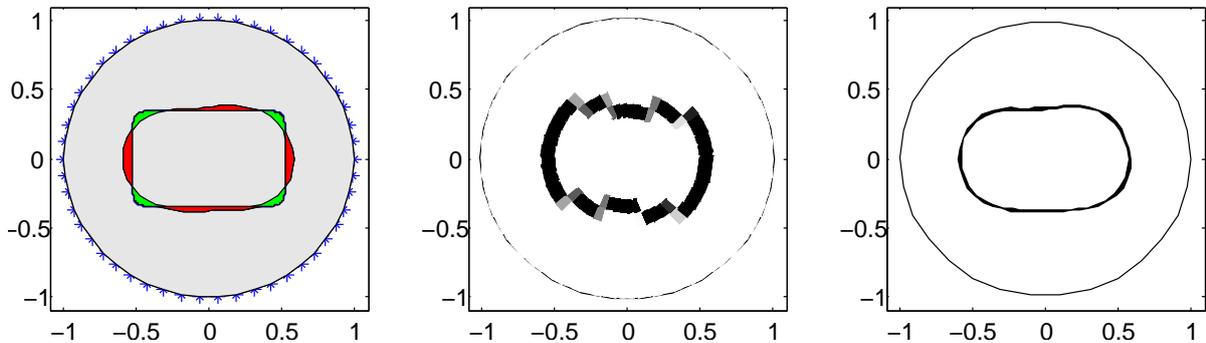


Figure 19: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$): estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

5.4 Full estimation using Laplace prior distributions

In this section, consider the full estimation using the Laplace distribution, for all prior distributions, in place of the Gaussian distribution. The hyper-prior parameter values are fixed at the values $\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$ and allowing estimation of the smoothing

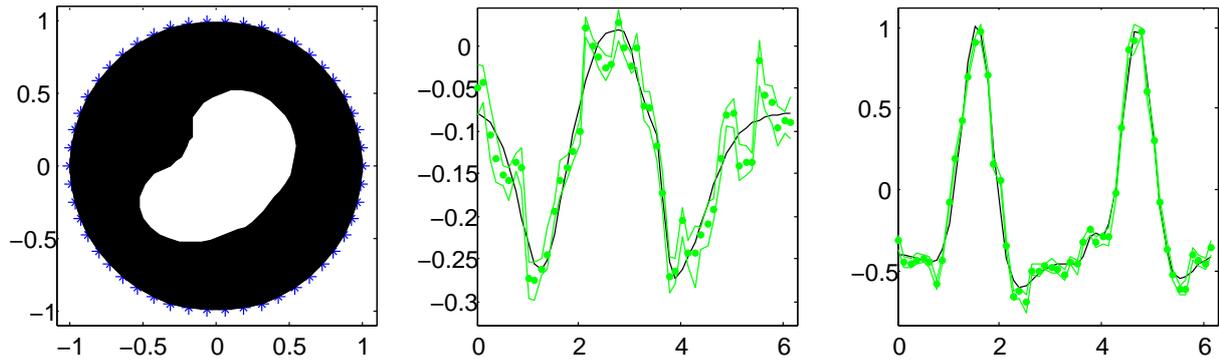


Figure 20: Bean-shaped inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$) and Laplace prior distributions: fitted shape (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

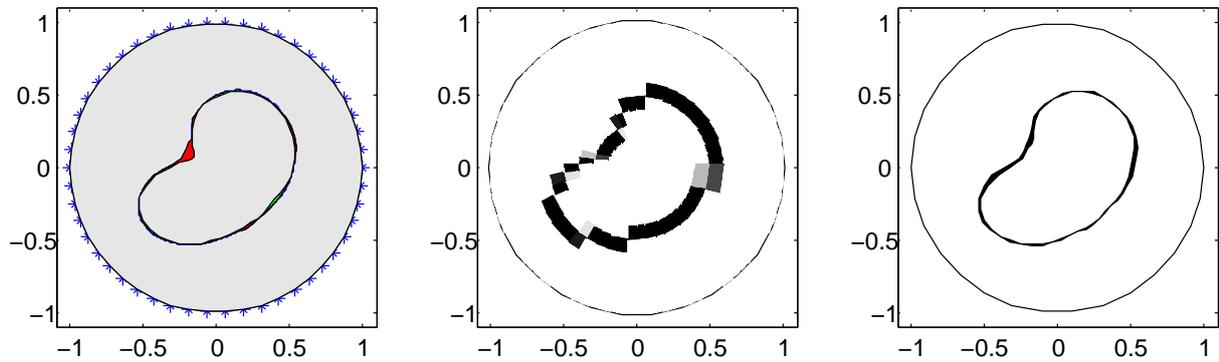


Figure 21: Bean-shaped inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$) and Laplace prior distributions: estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

parameters, β_r , β_{c_I} and β_{c_E} , as well as the MFS coefficients and the radii.

For the bean-shaped object, Figure 20 shows the reconstructed object and estimated MFS coefficients and Figure 21 shows the variability summaries. There has been little change, compared to the corresponding results using the Gaussian prior distribution (see Figures 14 and 15). However, there is a slight improvement in the accuracy of the “cusp” which reflects the Laplace distributions ability to better model abrupt changes.

Finally, consider the estimation for the round-cornered rectangular object. Figures 22 and 23 show the estimates and variability summaries. From Figure 22 it can be seen that the object outline is very well recovered with slight over rounding at the corners and irregular sides, but these are minor compared to those when the Gaussian prior distribution is used (see Figures 18 and 19). The MFS coefficients are also well estimated with substantial improvement in those associated with the interface compared to the earlier cases. These results demonstrate that the star-shaped model with a Laplace prior distribution has been very successful.

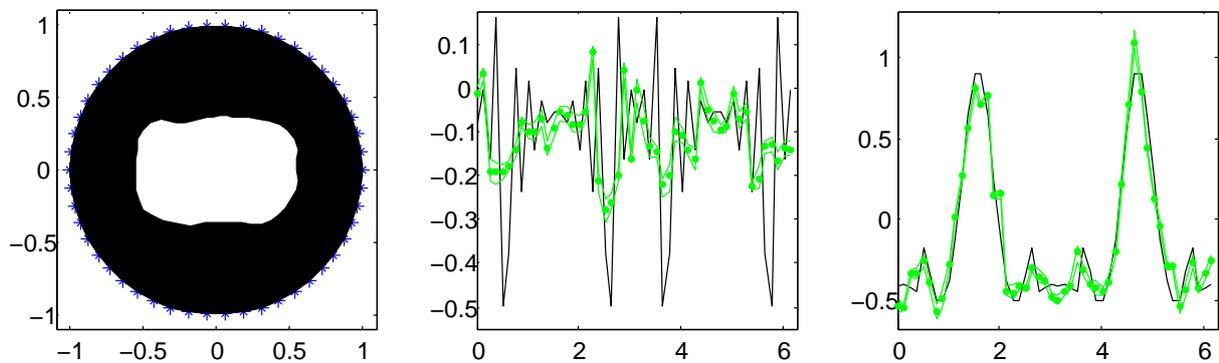


Figure 22: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$) and Laplace prior distributions: fitted circle (left) and MFS coefficients (with credible intervals) associated with the interface (centre) and outer boundary (right).

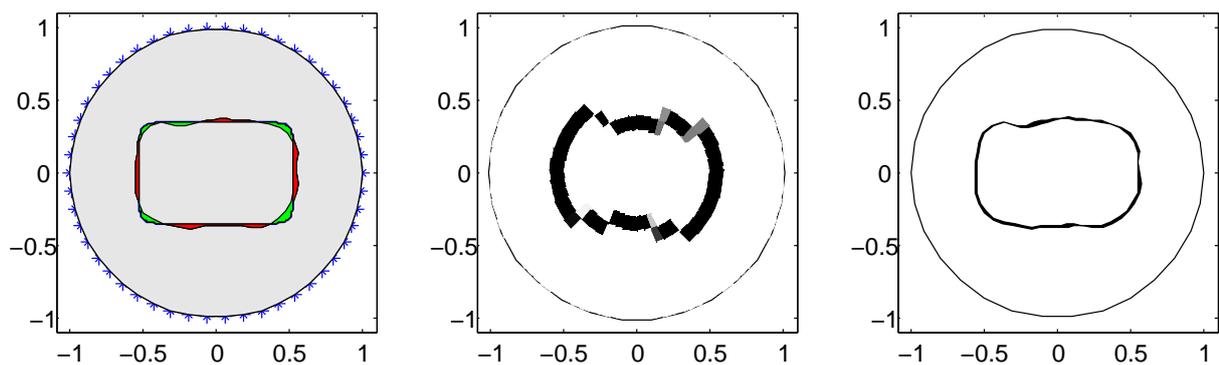


Figure 23: Round-cornered rectangular inclusion with star-shaped model with separate prior information ($\alpha_r = 0.1$, $\alpha_{c_I} = 0.0001$ and $\alpha_{c_E} = 0.0001$) and Laplace prior distributions: estimation errors (left), object boundary histogram (centre) and object boundary credible interval (right).

6 DISCUSSION

This paper has described the Bayesian approach to parameter estimation and the MCMC estimation algorithm, and applied them to the very practical problem of reconstructing the shape of an object from a continuous model EIT data. The MFS provides a simple yet accurate and fast approach to solving the forward problem. It is easy to describe and simple to program. However, it introduces additional, nuisance, parameters which must be estimated along with the variables of interest.

The Bayesian modelling approach gives a rigorous framework for including expert knowledge into the estimation process through prior distributions. Any beliefs regarding the nature of the parameter values, and relationships between the parameters can be incorporated. Also it provides a natural hierarchical structure to describe the dependence between variables which then allows a more intuitive description and interpretation of these relationships. Unfortunately, the prior distributions will contain additional unknown parameters. The framework also allows uncertainty in these parameters to be modelled via hyper-prior distributions. It would be possible to further define hyper-hyper-prior distributions, but this usually does not add anything to the performance, nor even the flexibility, of the model.

A simple MCMC estimation algorithm was developed which allowed all parameters to be well estimated. It is important to emphasise that such algorithms must be designed with care and should be tested widely to have good confidence that they are performing well. The great benefit when using MCMC algorithms is that complex models can be used easily. Also, there is great flexibility in the choice of output. The posterior sample can be used to estimate any summary. For example, posterior marginal distributions can be checked for normality, and where appropriate non-parametric techniques can be used to make inference in place of normal-based methods.

A range of simple examples have been considered and the proposed methods illustrated and developed. In the first set of examples a circular inclusion and a circular object model were used. Although of limited practical use this allowed the focus to be on the estimation of the MFS coefficients. It is clear that estimation can be improved substantially by the inclusion of prior information regarding boundary smoothing and that the two sets of MFS coefficients should be treated separately. These experiments highlight an important point that although maximum likelihood estimation is sometimes possible for such problem, and will produce a good fit to the data, it can leave parameter estimates which are not interpretable. With the inclusion of prior smoothing there is no significant deterioration in the goodness-of-fit but there is a substantial improvement in the interpretability of the MFS coefficients.

In practice it is the star-shaped model which is likely to be of greater use, and this model fitted all data well. It was even possible to perform a fully Bayesian analysis in which prior parameters were also estimated successfully. The bean-shaped and the round-cornered rectangle are challenging shapes to estimate accurately and so this approach can be considered very successful. Although in some examples the improvement over the use of fixed parameters was not always substantial there are examples where it makes a significant difference and hence

leads to a robust approach.

The results clearly indicate that the combined Bayesian/MCMC procedure has worked well, and that the MFS provides a very good and fast approximation to the forward solution. The examples have demonstrated the range of statistical models and prior distributions which can be used and the range of output summaries which are possible using MCMC sampling procedures. Also, the whole approach can easily be generalised making it a feasible approach even for complex modelling problems.

References

- Aykroyd, R. and Cattle, B. (2006). A flexible statistical and efficient computational approach to object location applied to electrical tomography, *Statistics and Computing* **16**: 363–375.
- Aykroyd, R. and Cattle, B. (2007). A flexible statistical and efficient computational approach to object location applied to electrical tomography, *Inverse Problems Sci. Eng.* **15**: 441–461.
- Borman, D., Ingham, D. B., Johansson, B. T. and Lesnic, D. (2009). The method of fundamental solutions for detection of cavities in EIT, *J. Integral Eqns and Appl.* **21**: 381–404.
- Brooks, S., Gelman, A., Jones, G. and Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*, Chapman & Hall/CRC.
- Cowles, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review, *J. Am. Stat. Soc.* **91**: 883–904.
- Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd edn, Chapman & Hall/CRC Texts in Statistical Science.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis* **1**: 515–533.
- Gelman, A., Carlin, J. B., Stern, H. and Rubin, D. (2003). *Bayesian Data Analysis*, second edn, Chapman & Hall/CRC.
- Geyer, C. J. (2011). Introduction to Markov Chain Monte Carlo, in S. Brooks, A. Gelman, G. L. Jones and X.-L. Meng (eds), *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC.
- Haddar, H. and Kress, R. (2005). Conformal mappings and inverse boundary value problems, *Inverse Problems* **21**: 935–953.
- Ivanyshyn, O. (2007). Shape reconstruction of acoustic obstacle from the modulus of the far field pattern, *Inverse Problems and Imaging* **1**: 609–622.
- Ivanyshyn, O. and Kress, R. (2006). Nonlinear integral equations for solving inverse boundary value problems for inclusions and cracks, *J. Integral Eqns and Appl.* **18**: 13–38.

- Karageorghis, A., Lesnic, D. and Marin, L. (2011). The MFS for inverse geometric problems, in L. Marin, L. Munteanu and V. Chiroiu (eds), *Inverse Problems and Computational Mechanics*, Chapter 8, pp. 191–216.
- Karageorghis, A., Lesnic, D. and Marin, L. (2013). A moving pseudo-boundary MFS for void detection, *Numerical Methods for Partial Differential Equations* **29**: 935–960.
- Liu, J. (2008). *Monte Carlo Strategies in Scientific Computing*, Springer.
- Raftery, A. and Lewis, S. (1995). The number of iterations, convergence diagnostics and generic Metropolis algorithms, in W. R. Gilks, S. Richardson and D. J. Spiegelhalter (eds), *Practical Markov Chain Monte Carlo*, Chapman and Hall.
- Roberts, G., Gelman, A. and Gilks, W. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms, *Ann. Appl. Prob.* **7**: 110–120.
- Smyrlis, Y.-S. and Karageorghis, A. (2009). Efficient implementation of the mfs: The three scenarios, *J. Comput. Appl. Math.* **227**: 8392.
- Stuart, A. M. (2010). Inverse problems: A Bayesian perspective, *Acta Numerica* **19**: 451–559.
- Watzenig, D. and Fox, C. (2009). A review of statistical modelling and inference for electrical capacitance tomography, *Meas. Sci. Technol.* **20**: 052002 (22pp).
- West, R., Aykroyd, R., Meng, S. and Williams, R. (2004). Markov chain Monte Carlo techniques and spatial-temporal modelling for medical EIT, *Physiological Measurements* **25**: 181–194.
- West, R., Meng, S., Aykroyd, R. and Williams, R. (2005). Spatial-temporal modelling for electrical impedance imaging of a mixing process, *Rev. Scientific Instruments* **76**: 073703.