THE BAYESIAN FORMULATION OF EIT:
ANALYSIS AND ALGORITHMS

Matthew M. Dunlop and Andrew M. Stuart

Computing & Mathematical Sciences
California Institute of Technology
Pasadena, CA 91125, USA

(Communicated by Samuli Siltanen)

Abstract. We provide a rigorous Bayesian formulation of the EIT problem in an infinite dimensional setting, leading to well-posedness in the Hellinger metric with respect to the data. We focus particularly on the reconstruction of binary fields where the interface between different media is the primary unknown. We consider three different prior models – log-Gaussian, star-shaped and level set. Numerical simulations based on the implementation of MCMC are performed, illustrating the advantages and disadvantages of each type of prior in the reconstruction, in the case where the true conductivity is a binary field, and exhibiting the properties of the resulting posterior distribution.

1. Introduction.

1.1. **Background.** Electrical Impedance Tomography (EIT) is an imaging technique in which the conductivity of a body is inferred from electrode measurements on its surface. Examples include medical imaging, where the methodology is used to non-invasively detect abnormal tissue within a patient, and subsurface imaging where material properties of the subsurface are determined from surface (or occasional interior) measurements of the electrical response; the methodology is often referred to as electrical resistance tomography – ERT – in this context and discussed in more detail below. The concept of EIT appears as early as the late 1970's [15] and ERT the 1930's [28].

A very influential mathematical formulation of the inverse problem associated with EIT dates back to 1980, due to Calderón. He formulated an abstract version of the problem, in which the objective is recovery of the coefficient of a divergence form elliptic PDE from knowledge of its Neumann-to-Dirichlet or Dirichlet-to-Neumann operator. Specifically, in the Dirichlet-to-Neumann case, if $D \subseteq \mathbb{R}^d$ and $g \in H^{1/2}(\partial D)$ is given, let $u \in H^1(D)$ solve

$$\begin{cases} -\nabla \cdot (\sigma \nabla u) = 0 & \text{in } D \\ u = g & \text{on } \partial D. \end{cases}$$

The problem of interest is then to ask does the mapping $\Lambda_\sigma : H^{1/2}(\partial D) \to H^{-1/2}(\partial D)$ given by

$$g \mapsto \sigma \frac{\partial u}{\partial \nu}$$

determine the coefficient σ in D? Physically, g corresponds to boundary voltage measurements, and $\Lambda_\sigma(g)$ corresponds to the current density on the boundary. Much study has been carried out on this problem – some significant results, in the case where all conductivities are in $C^2(\overline{D})$ and $d \geq 3$, concern uniqueness [42], reconstruction [35], stability [2] and partial data [22]. Details of these results are summarised in [38].

In 1996, Nachman proved global uniqueness and provided a reconstruction procedure for the case $d = 2$, involving the use of a scattering transform and solving a D-bar problem [36]. The D-bar equation involved is a differential equation of the form $\overline{\partial} q = f$, where $\overline{\partial}$ denotes the conjugate of the complex derivative and f depends on the scattering transform. A regularised D-bar approach, involving the truncation of the scattering transform, was provided in [23, 24], enabling the recovery of features of discontinuous permeabilities. The regularised D-bar approach is also used in [25], for the case when the data is not of infinite precision. Other work in the area includes joint inference of the shape of the domain and conductivity [26].

For applications, a more physically appropriate model for EIT was provided in 1992 in [40]. This model, referred to as the complete electrode model (CEM), replaces complete boundary potential measurements with measurements of constant potential along electrodes on the boundary, subject to contact impedances. The authors show that predictions from this model agree with experimental measurements up to the measurement precision of 0.1%. For this model they also prove existence and uniqueness of the associated electric potential. It is this model that we shall consider in this paper, and it is outlined in section 2.

When using the CEM, there is a limitation on the number of measurements that can be taken to provide additional information. This is due to the fact that there are only a finite number of electrodes through which current can be injected and voltages read, combined with the linear relationship between current and voltage. The data is therefore finite dimensional in the inverse problem, as distinct from the Calderón problem where knowledge of an infinite dimensional operator is assumed. As a consequence, reconstruction using the CEM often makes use of Tikhonov regularisation; such a regularisation can also be used to account for noise on the data. The paper [13] analyses numerical convergence when an H^1 or TV penalty term is used, with a finite element discretisation of the problem. We will adopt a Bayesian approach to regularisation, and this is discussed below.

A closely related problem to EIT is Electrical Resistivity Tomography (ERT), which concerns subsurface inference from surface potential measurements, see for example [28] which discussed the problem as early as 1933. Physically the main difference between EIT and ERT is that alternating current is typically used for the former, and direct current for the latter. Additionally, due to the scale of ERT, it is a reasonable approximation to model the electrodes as points, rather than using the CEM. Another difference between the two is that the relative contrast between the conductivities of different media are typically higher in subsurface applications than medical applications, which permits the approximation of the problem by a network of resistors in some cases [37]. Nonetheless, the Bayesian theory and MCMC methodology introduced here will be useful for the ERT problem as well as the EIT problem.

Statistical, in particular Bayesian, approaches to EIT inversion have previously been studied, for example in [18, 19, 31]. In [18], the authors prove certain regularity of the forward map associated with the CEM, formulate the Bayesian inverse

problem in terms of the discretised model, and investigate the effect of different priors on reconstruction and behaviour of the posterior. The paper [31] focuses on Whittle-Matèrn priors, using EIT and ERT as examples for numerical simulation. The paper [19] presents a regularised version of the inverse problem, which admits a Bayesian interpretation.

The Bayesian approach to inverse problems, especially in infinite dimensions, is a relatively new technique. Two approaches are typically taken: discretise first and use finite dimensional Bayesian techniques, or apply the Bayesian methodology directly on function space before discretising. The former approach is outlined in [20]. The latter approach for linear problems has been studied in [12, 32, 33, 34]. More recently, this approach has been applied to nonlinear problems [10, 29, 30, 41]. It is this approach that we will be taking in this paper.

1.2. **Our contribution.** The main contributions in our paper are as follows:

(i) This is the first paper to give a rigorous Bayesian formulation of EIT in infinite dimensions.

(ii) This setting leads to proof that the posterior measure is Lipschitz in the data, with respect to the Hellinger metric, for all three priors studied; further stability properties of the posterior with respect to perturbations, such as numerical approximation of the forward model, may be proved similarly.

(iii) We employ a variety of prior models, based on the assumption that the underlying conductivity we wish to recover is binary. We initially look at log-Gaussian priors, before focusing on priors which enforce the binary field property. These binary field priors include both single star-shaped inclusions parametrised by their centre and by a radial function [8], and arbitrary geometric interfaces between the two conductivity values defined via level set functions [17].

(iv) Numerical results using state of the art MCMC demonstrate the importance of the prior choice for accurate reconstruction in the severely underdetermined inverse problems arising in EIT.

1.3. **Organisation of the paper.** In section 2 we describe the forward map associated with the EIT problem, and prove relevant regularity properties. In section 3 we formulate the inverse problem rigorously and describe our three prior models. We then prove existence and well-posedness of the posterior distribution for each of these choices of prior. In section 4 we present results of numerical MCMC simulations to investigate the effect of the choice of prior on the recovery of certain binary conductivity fields. We conclude in section 5.

2. **The forward model.** In subsection 2.1 we describe the complete electrode model for EIT as given in [40]. In subsection 2.2 we give the weak formulation of this model, stating assumptions required for the quoted existence and uniqueness result. Then in subsection 2.3 we define the forward map in terms of this model, and prove that this map is continuous with respect to both uniform convergence and convergence in measure.

2.1. **Problem statement.** Let $D \subseteq \mathbb{R}^d$, $d \leq 3$, be a bounded open set with smooth boundary representing a body, with conductivity $\sigma : \overline{D} \to \mathbb{R}$. A number L of electrodes are attached to the surface of the body. We treat these as subsets $(e_l)_{l=1}^L$ of the boundary ∂D, and assume that they have contact impedances $(z_l)_{l=1}^L \in \mathbb{R}^L$. A current stimulation pattern $(I_l)_{l=1}^L \in \mathbb{R}^L$ is applied to the electrodes. Then the

electric potential v within the body and boundary voltages $(V_l)_{l=1}^L \in \mathbb{R}^L$ on $(e_l)_{l=1}^L$ are modelled by the following partial differential equation.

(1)
$$
\begin{cases}
-\nabla \cdot (\sigma(x)\nabla v(x)) = 0 & x \in D \\[2mm]
\displaystyle\int_{e_l} \sigma \frac{\partial v}{\partial n}\, \mathrm{d}S = I_l & l = 1, \ldots, L \\[2mm]
\sigma(x)\dfrac{\partial v}{\partial n}(x) = 0 & x \in \partial D \setminus \bigcup_{l=1}^L e_l \\[2mm]
v(x) + z_l \sigma(x)\dfrac{\partial v}{\partial n}(x) = V_l & x \in e_l, l = 1, \ldots, L
\end{cases}
$$

This model was first proposed in [40]; a derivation can be found therein. Note that the inputs to this forward model are the conductivity σ, input current $(I_l)_{l=1}^L$ and contact impedances $(z_l)_{l=1}^L$. The solution comprises the function $v : D \to \mathbb{R}$ *and* the vector $V \in \mathbb{R}^L$ of voltages. Also note that solutions to this equation are only defined up to addition of a constant: if (v, V) solves the equation, then so does $(v + c, V + c)$ for any $c \in \mathbb{R}$. This is because it is necessary to choose a reference ground voltage.
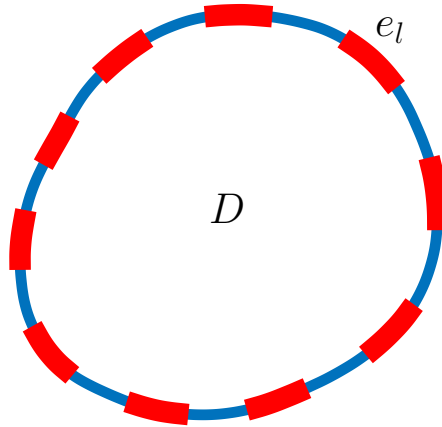

FIGURE 1. An example domain D with electrodes $(e_l)_{l=1}^L$ attached to its boundary.

2.2. **Weak formulation.** We first define the space in which the solution to equation (1) will live. Using the notation of [40] we set

$$\mathbb{H} = H^1(D) \oplus \mathbb{R}^L,$$
$$\|(v, V)\|_{\mathbb{H}}^2 = \|v\|_{H^1(D)}^2 + \|V\|_{\ell^2}^2$$
$$= \|v\|_{L^2(D)}^2 + \|\nabla v\|_{L^2(D)}^2 + \|V\|_{\ell^2}^2.$$

Since solutions are only defined up to addition of a constant, we define the quotient space $(\dot{\mathbb{H}}, \|\cdot\|_{\dot{\mathbb{H}}})$ by

$$\dot{\mathbb{H}} = \mathbb{H}/\mathbb{R},$$
$$\|(v, V)\|_{\dot{\mathbb{H}}} = \inf_{c \in \mathbb{R}} \|(v - c, V - c)\|_{\mathbb{H}}.$$

We will often use the notation $v' = (v, V)$ for brevity. It is more convenient to equip $\dot{\mathbb{H}}$ with an equivalent norm, as stated in the following lemma from [40]:

Lemma 2.1. *Define* $\|v'\|_*$ *by*

$$\|v'\|_*^2 = \|\nabla v\|_{L^2(D)}^2 + \sum_{l=1}^{L} \int_{e_l} |v(x) - V_l|^2 \, \mathrm{d}S.$$

Then $\| \cdot \|_*$ *and* $\| \cdot \|_{\dot{\mathbb{H}}}$ *are equivalent.*

We can now state the weak formulation of the problem as derived in [40]. For this let $w' = (w, W)$.

Proposition 2.2. *Let* $B : \dot{\mathbb{H}} \times \dot{\mathbb{H}} \to \mathbb{R}$ *and* $r : \dot{\mathbb{H}} \to \mathbb{R}$ *be defined by*

$$B(v', w'; \sigma) = \int_D \sigma \nabla v \cdot \nabla w \, \mathrm{d}x + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (v - V_l)(w - W_l) \, \mathrm{d}S,$$

$$r(w') = \sum_{l=1}^{L} I_l W_l.$$

Then if $v' \in \dot{\mathbb{H}}$ *is a strong solution to the problem (1), it satisfies*

$$(2) \qquad\qquad B(v', w'; \sigma) = r(w') \quad \text{for all } w' \in \dot{\mathbb{H}}.$$

We will use the weak formulation (2) to define our forward map for the complete electrode model (1). In order to guarantee a solution to this problem, we make the following assumptions.

Assumptions 2.3. *The conductivity* $\sigma : \overline{D} \to \mathbb{R}$*, the contact impedances* $(z_l)_{l=1}^{L} \in \mathbb{R}^L$ *and the current stimulation pattern* $(I_l)_{l=1}^{L} \in \mathbb{R}^L$ *satisfy*

(i) $\sigma \in L^\infty(\overline{D}; \mathbb{R})$, $\operatorname*{essinf}_{x \in D} \sigma(x) = \sigma_- > 0$;

(ii) $0 < z_- \le z_l \le z_+ < \infty$, $l = 1, \dots, L$;

(iii) $\displaystyle\sum_{l=1}^{L} I_l = 0.$

Under these assumptions, existence of a unique solution to (2) is proved in [40] and stated here for convenience:

Proposition 2.4. *Let Assumptions 2.3 hold, then there is a unique* $[(v, V)] \in \dot{\mathbb{H}}$ *solving (2). We may, without loss of generality, choose the element* $(v, V) \in [(v, V)]$ *of the equivalence class of solutions to be that which satisfies*

$$(3) \qquad\qquad \sum_{l=1}^{L} V_l = 0.$$

Remark 2.5. Assumptions 2.3 (i) and (ii) are to ensure coercivity and boundedness of $B(\cdot, \cdot; \sigma)$. Assumption 2.3 (iii) is necessary for continuity of $r(\cdot)$, and physically may be thought of as a conservation of charge condition. Choosing a solution from the equivalence class corresponds to choosing a reference ground voltage.

2.3. **Continuity of the forward map.** In what follows we will restrict to the set of admissible conductivities, which is defined as follows.

Definition 2.6. A conductivity field $\sigma : \overline{D} \to \mathbb{R}$ is said to be admissible if

(i) there exists $N \in \mathbb{N}$, $\{D_n\}_{n=1}^N$ open disjoint subsets of D for which $\overline{D} = \bigcup_{n=1}^N \overline{D}_j$;

(ii) $\sigma|_{D_j} \in C(\overline{D}_j)$; and

(iii) there exist $\sigma^-, \sigma^+ \in (0, \infty)$ such that $\sigma^- \leq \sigma(x) \leq \sigma^+$ for all $x \in \overline{D}$.

The set of all such conductivities will be denoted $\mathcal{A}(D)$.

Note that any $\sigma \in \mathcal{A}(D)$ will satisfy Assumptions 2.3(i). Assume that the current stimulation pattern $(I_l)_{l=1}^L \in \mathbb{R}^L$ and contact impedances $(z_l)_{l=1}^L \in \mathbb{R}^L$ are known and satisfy Assumptions 2.3(ii)-(iii). Then we may define the solution map $\mathcal{M} : \mathcal{A}(D) \to \mathbb{H}$ to be the unique solution to (2) satisfying (3). The above existence and uniqueness result tells us that this map is well-defined.

In [18] it is shown that $\mathcal{M} : \mathcal{A}(D) \to \mathbb{H}$ is Fréchet differentiable when we equip $\mathcal{A}(D)$ with the supremum norm. Though this is a strong result, this choice of norm is not appropriate for all of the conductivities that we will be considering. We hence establish the following continuity result.

Proposition 2.7. *Fix a current stimulation pattern $(I_l)_{l=1}^L \in \mathbb{R}^L$ and contact impedances $(z_l)_{l=1}^L \in \mathbb{R}^L$ satisfying Assumptions 2.3. Define the solution map $\mathcal{M} : \mathcal{A}(D) \to \mathbb{H}$ as above. Let $\sigma \in \mathcal{A}(D)$ and let $(\sigma_\varepsilon)_{\varepsilon>0} \subseteq \mathcal{A}(D)$ be such that either*

(i) *σ_ε converges to σ uniformly; or*

(ii) *σ_ε converges to σ in (Lebesgue) measure, and there exist $\sigma^-, \sigma^+ \in (0, \infty)$ such that for all $\varepsilon > 0$ and $x \in D$, $\sigma^- \leq \sigma_\varepsilon(x) \leq \sigma^+$.*

Then $\|\mathcal{M}(\sigma_\varepsilon) - \mathcal{M}(\sigma)\|_* \to 0$.

Proof. Define the maps $B : \mathbb{H} \times \mathbb{H} \times \mathcal{A}(D) \to \mathbb{R}$ and $r : \mathbb{H} \to \mathbb{R}$ as in Lemma 2.2, but on $\mathbb{H}$ rather than $\dot{\mathbb{H}}$. Then denoting $u'_\varepsilon = (v_\varepsilon, V^\varepsilon) = \mathcal{M}(\sigma_\varepsilon)$ and $v' = (v, V) = \mathcal{M}(\sigma)$, we have for all $w' \in \mathbb{H}$,

$$B(v'_\varepsilon, w'; \sigma_\varepsilon) = r(w'), \quad B(v', w'; \sigma) = r(w').$$

It follows that

$$0 = B(v'_\varepsilon, w'; \sigma_\varepsilon) - B(v', w'; \sigma_\varepsilon) + B(v', w'; \sigma_\varepsilon) - B(v', w'; \sigma)$$

$$= \int_D \sigma_\varepsilon \nabla(v_\varepsilon - v) \cdot \nabla w \, \mathrm{d}x + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} ((v_\varepsilon - v) - (V_l^\varepsilon - V_l))(w - W_l) \, \mathrm{d}S$$

$$+ \int_D (\sigma_\varepsilon - \sigma) \nabla v \cdot \nabla w \, \mathrm{d}x.$$

Letting $w' = (v_\varepsilon - v, V^\varepsilon - V)$, we see that

$$\int_D \sigma_\varepsilon |\nabla(v_\varepsilon - v)|^2 \, \mathrm{d}x + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} ((v_\varepsilon - v) - (V_l^\varepsilon - V_l))^2 \, \mathrm{d}S$$

$$\leq \int_D |\sigma_\varepsilon - \sigma| |\nabla v \cdot \nabla(v_\varepsilon - v)| \, \mathrm{d}x.$$

In both cases (i) and (ii), we have that $(\sigma_\varepsilon)_{\varepsilon>0}$ is bounded uniformly below by a positive constant. Hence for small enough ε, the left hand side above may be bounded below by $C\|v'_\varepsilon - v'\|^2_*$. We then have by Cauchy-Schwarz

$$\|v'_\varepsilon - v'\|^2_* \leq C \int_D |\sigma_\varepsilon - \sigma||\nabla v \cdot \nabla(v_\varepsilon - v)|\, \mathrm{d}x$$

$$\leq C \left(\int_D |\sigma_\varepsilon - \sigma|^2 |\nabla v|^2\, \mathrm{d}x \right)^{1/2} \cdot \|\nabla(v_\varepsilon - v)\|_{L^2}$$

$$(4) \qquad \leq C \left(\int_D |\sigma_\varepsilon - \sigma|^2 |\nabla v|^2\, \mathrm{d}x \right)^{1/2} \cdot \|v'_\varepsilon - v'\|_*$$

$$(5) \qquad \leq C\|\sigma_\varepsilon - \sigma\|_\infty \|\nabla v\|_{L^2} \|v'_\varepsilon - v'\|_*.$$

If $\sigma_\varepsilon \to \sigma$ uniformly, we deduce from (5) that $\|v'_\varepsilon - v'\|_* \to 0$ and the result follows. If $|\sigma_\varepsilon - \sigma| \to 0$ in measure, then since $|D| < \infty$, it follows that the integrand in (4) tends to zero in measure, see for example Corollary 2.2.6 in [6]. Since σ_ε is assumed to be uniformly bounded, the integrand is dominated by a scalar multiple of the integrable function $|\nabla v|^2$. We claim that this implies that the integrand tends to zero in L^1. Suppose not, and denote the integrand f_ε. Then there exists $\delta > 0$ and a subsequence $(f_{\varepsilon_i})_{i \geq 1}$ such that $\|f_{\varepsilon_i}\|_{L^1} \geq \delta$ for all i. This subsequence still converges to zero in measure, and so admits a further subsequence that converges to zero almost surely. An application of the dominated convergence theorem leads to a contradiction, hence we deduce that f_ε tends to zero in L^1 and the result follows. □

Denote the projection $\Pi : \mathbb{H} \to \mathbb{R}^L$, $(v, V) \mapsto V$. The following lemma shows that the above result still holds if we replace $\mathcal{M}$ by $\Pi \circ \mathcal{M}$.

Corollary 2.8. *Let the assumptions of Proposition 2.7 hold. Then*

$$|\Pi \circ \mathcal{M}(\sigma_\varepsilon) - \Pi \circ \mathcal{M}(\sigma)|_{\ell^2} \to 0.$$

Proof. We show that there exists $C > 0$ such that for all $(v, V) \in \mathbb{H}$ with $\sum_{l=1}^{L} V_l = 0$, $\|(v, V)\|_* \geq C|V|_{\ell^2}$. By the equivalence of $\|\cdot\|_*$ and $\|\cdot\|_{\dot{\mathbb{H}}}$, Lemma 2.1, we have

$$\|(v, V)\|_* \geq C \inf_{c \in \mathbb{R}} (\|v - c\|_{H^1} + |V - c|_{\ell^2}) \geq C \inf_{c \in \mathbb{R}} |V - c|_{\ell^2}$$

The infimum on the right-hand side is attained at

$$c = \frac{1}{L} \sum_{l=1}^{L} V_l = 0.$$

Then by Proposition 2.7, we have

$$0 \leq |\Pi \circ \mathcal{M}(\sigma_\varepsilon) - \Pi \circ \mathcal{M}(\sigma)|_{\ell^2} \leq \|\mathcal{M}(\sigma_\varepsilon) - \mathcal{M}(\sigma)\|_* \to 0$$

□

3. **The inverse problem.** We are interested in the inverse problem of determining the conductivity field from measurements of the voltages $(V_l)_{l=1}^{L}$ on the boundary, for a variety of input currents $(I_l)_{l=1}^{L}$ on the boundary. To this end we introduce the following version of Ohm's law. Observe that the mapping $I \mapsto v'$, taking the current stimulation pattern to the solution of (2), is linear. Then given a conductivity field $\sigma \in \mathcal{A}(D)$, there exists a resistivity matrix $R(\sigma) \in \mathbb{R}^{L \times L}$ such that the boundary voltage measurements $V(\sigma)$ arising from the solution of the forward model are related to I via

$$V(\sigma) = R(\sigma)I$$

By applying several different current stimulation patterns we should be able to infer more about the conductivity σ. Note however that since the mapping $I \mapsto V$ is linear, only linearly independent stimulation patterns will provide more information[1]. Since we have the conservation of charge condition on I, there are at most $L - 1$ linearly independent patterns we can use.

Assume that J linearly independent current patterns $I^{(j)} \in \mathbb{R}^L$, $j = 1, \ldots, J$, $J \leq L - 1$ are applied, and noisy measurements of $V^{(j)} = R(\sigma)I^{(j)}$ are made:

$$y_j = V^{(j)} + \eta_j, \quad \eta_j \sim N(0, \Gamma_0) \text{ iid.}$$

We have

$$y_j = \mathcal{G}_j(\sigma) + \eta_j$$

where $\mathcal{G}_j(\sigma) = R(\sigma)I^{(j)}$. Concatenating these observations, we write

$$y = \mathcal{G}(\sigma) + \eta, \quad \eta \sim N(0, \Gamma)$$

where $\Gamma = \mathrm{diag}(\Gamma_0, \ldots, \Gamma_0)$. The inverse problem is then to recover the conductivity field σ from the data y. This problem is highly ill-posed: the data is finite dimensional, yet we wish to recover a function which, typically, lies in an infinite dimensional space. We take a Bayesian approach by placing a prior distribution on σ. The choice of prior may have significant effect on the resulting posterior distribution, and different choices of prior may be more appropriate depending upon the prior knowledge of the particular experimental set-up under consideration.

In subsection 3.1 we outline three different families of prior models, and show the appropriate regularity of the forward maps arising from them. In subsection 3.2 we describe the likelihood and posterior distribution formally, before rigorously proving that the posterior distribution exists and is Lipschitz with respect to the data in the Hellinger metric.

3.1. **Choices of prior.** In this section we consider three priors, labelled by $i = 1, 2, 3$, defined by functions $F_i : X_i \to \mathcal{A}(D)$ which map draws from prior measures on the Banach spaces X_i to the space of conductivities $\mathcal{A}(D)$. Our prior conductivity distributions will then be the pushforward of the prior measures by these maps F_i. We describe these maps, and establish continuity properties of them needed for the study of the posterior later. In what follows, the space $C_0(\overline{D})$ of continuous functions on $\overline{D}$ will be equipped with the supremum norm $\| \cdot \|_\infty$ and the corresponding Borel σ-algebra.

3.1.1. *Log-Gaussian prior.* We first consider the simple case that the coefficient is given by the exponential of a continuous function. Let $F_1 : C^0(\overline{D}) \to \mathcal{A}(D)$ be defined by $F_1(u) = \exp(u)$. Then it is easily seen that F_1 does indeed map into $\mathcal{A}(D)$. Furthermore, since D is bounded, if $u \in C^0(\overline{D})$ and $(u_\varepsilon)_{\varepsilon>0} \subseteq C^0(\overline{D})$ is a sequence such that $\|u_\varepsilon - u\|_\infty \to 0$, then $\|F_1(u_\varepsilon) - F_1(u)\|_\infty \to 0$.

In this case, we will take our prior measure μ_0 on u to be a non-degenerate Gaussian measure $N(m_0, \mathcal{C}_0)$ on $C^0(\overline{D})$. Note that the push forward of a Gaussian measure by F_1 is a log-Gaussian measure.

---

[1]If there is noise on the measurements, additional linearly dependent observations can be made to effectively reduce the noise level on the original measurements. We can assume that this has been done and scale the noise appropriately.

Example 3.1. Consider the case $D = B(0,1) \subseteq \mathbb{R}^2$. Suppose that u is drawn from a Gaussian measure $\mu_0 = N(0, \mathcal{C})$. Typical samples from $F_1^{\#}(\mu_0)$ are shown below[2]. The covariance $\mathcal{C}$ is chosen such that the samples u almost surely have regularity $u \in C^{\lfloor s \rfloor, s - \lfloor s \rfloor}(D)$ for all $s < t$, where from left to right $t = 2, 1.5, 1, 0.5$ respectively. Here the samples are generated on $[-1,1]^2 \supseteq D$ and then restricted to D, for computational simplicity.
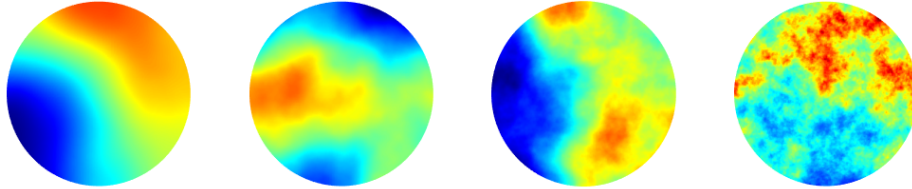
3.1.2. *Star-shaped prior.* We now consider star-shaped inclusions, that is, inclusions parametrised by their centre and a radial function. These were studied in two-dimensions in the paper [8] to parametrise domains for a Bayesian inverse shape scattering problem. In [8] the authors prove well-posedness of the inverse problem in an infinite dimensional setting through the use of shape derivatives and Riesz-Fredholm theory.

Let $D \subseteq \mathbb{R}^d$, and $R_{d-1} = (-\pi, \pi] \times [0, \pi]^{d-2} \subseteq \mathbb{R}^{d-1}$. Let $h : \mathbb{R}^d \to R_{d-1}$ be the continuous function representing the mapping from Cartesian to angular polar coordinates. Define the mapping $A : C_P^0(R_{d-1}) \times D \to \mathcal{B}(D)$ by

$$A(r, x_0) = \left\{ x \in D \mid |x - x_0| \leq r(h(x - x_0)) \right\}$$

where $C_P^0(R_{d-1})$ is the space of continuous periodic functions on R_{d-1}. Then $A(r, x_0)$ describes the set of points in D which lie within the closed surface parametrised in polar coordinates centred at x_0 by

$$\Gamma(\Theta) = (\Theta, r(\Theta)), \quad \Theta \in R_{d-1}.$$

In two dimensions, we have $R_1 = (-\pi, \pi]$ and the mapping $h : \mathbb{R}^2 \to R_1$ is given by

$$h(x, y) = \text{atan2}(y, x) \equiv 2 \arctan\left(\frac{y}{\sqrt{x^2 + y^2} + x} \right)$$

where atan2 is the two-parameter inverse tangent function.

In three dimensions, we have $R_2 = (-\pi, \pi] \times [0, \pi]$ and the mapping $h : \mathbb{R}^3 \to R_2$ is given by

$$h(x, y, z) = \left(\text{atan2}(y, x), \text{arccot}\left(\frac{z}{\sqrt{x^2 + y^2}} \right) \right).$$

Similar expressions for h exist in higher dimensions, though for applications we are only interested in the case $d = 2, 3$.

Define now the map $F_2 : C_P^0(R_{d-1}) \times D \to \mathcal{A}(D)$ by

$$F_2(r, x_0) = u_+ \mathbb{1}_{A(r, x_0)} + u_- \mathbb{1}_{D \setminus A(r, x_0)}$$
$$= (u_+ - u_-) \mathbb{1}_{A(r, x_0)} + u_-,$$

---

[2]Given a measure μ on $(X, \mathcal{X})$ and a measurable map $F : (X, \mathcal{X}) \to (Y, \mathcal{Y})$ between measurable spaces, $F^{\#}(\mu)$ denotes the pushforward of μ by F, i.e. the measure on $(Y, \mathcal{Y})$ given by $F^{\#}(\mu)(A) = \mu(F^{-1}(A))$ for all $A \in \mathcal{Y}$.

where $u_+, u_- > 0$ are the scalar conductivity values. Again it can easily be seen that F_2 does indeed map into $\mathcal{A}(D)$. We claim that this map is continuous in the following sense:

Proposition 3.2. *Define the map* $F_2 : C_P^0(R_{d-1}) \times D \to \mathcal{A}(D)$ *as above. Let* $x_0 \in D$ *and let* $r \in C_P^0(R_{d-1})$ *be Lipschitz continuous.*

(i) *Suppose that* $(r_\varepsilon)_{\varepsilon>0} \subseteq C_P^0(R_{d-1})$ *is a sequence of functions such that* $\|r_\varepsilon - r\|_\infty \to 0$. *Then* $F_2(r_\varepsilon, x_0) \to F_2(r, x_0)$ *in measure*[3].

(ii) *Suppose that* $(x_0^\varepsilon)_{\varepsilon>0} \subseteq D$ *is a sequence of points such that* $|x_0^\varepsilon - x_0| \to 0$. *Then* $F_2(r, x_0^\varepsilon) \to F_2(r, x_0)$ *in measure.*

(iii) *Let* $(r_\varepsilon)_{\varepsilon>0}$, $(x_0^\varepsilon)_{\varepsilon>0}$ *be as above. Then* $F_2(r_\varepsilon, x_0^\varepsilon) \to F_2(r, x_0)$ *in measure.*

Proof. In order to show that a sequence of functions $f_\varepsilon : D \to \mathbb{R}$ converges to $f : D \to \mathbb{R}$ in measure, it suffices to show that there exists a sequence of sets $Z_\varepsilon \subseteq D$ with $|Z_\varepsilon| \to 0$ such that $|f_\varepsilon - f| \leq C\mathbb{1}_{Z_\varepsilon}$. Then for each $\delta > 0$ we have

$$|\{x \in D \,|\, |f_\varepsilon(x) - f(x)| > \delta\}| \leq |\{x \in D \,|\, |f_\varepsilon(x) - f(x)| \neq 0\}| \leq |Z_\varepsilon| \to 0.$$

(i) Fix the centre $x_0 \in D$. Denote $A(r) = A(r, x_0)$. Let $r \in C_P^0(R_{d-1})$ and let $(r_\varepsilon)_{\varepsilon>0} \subseteq C_P^0(R_{d-1})$ be a sequence of functions such that $\|r_\varepsilon - r\|_\infty \to 0$. Then there exists $\gamma(\varepsilon) \to 0$ such that $\|r_\varepsilon - r\|_\infty < \gamma(\varepsilon)$. By definition we then have

$$r(x) - \gamma(\varepsilon) \leq r_\varepsilon(x) \leq r(x) + \gamma(\varepsilon) \quad \text{for all } x \in D \text{ and } \varepsilon > 0.$$

It follows that we have the inclusions

$$A(r - \gamma(\varepsilon)) \subseteq A(r_\varepsilon) \subseteq A(r + \gamma(\varepsilon)),$$
$$A(r - \gamma(\varepsilon)) \subseteq A(r) \subseteq A(r + \gamma(\varepsilon)).$$

Let Δ denote the symmetric difference. We deduce that

$$A(r_\varepsilon)\Delta A(r) \subseteq A(r + \gamma(\varepsilon)) \setminus A(r - \gamma(\varepsilon)).$$

Now the right-hand side is given by

$$A(r + \gamma(\varepsilon)) \setminus A(r - \gamma(\varepsilon))$$
$$= \left\{x \in D \,\big|\, r(h(x - x_0)) - \gamma(\varepsilon) < |x - x_0| \leq r(h(x - x_0)) + \gamma(\varepsilon)\right\}.$$

As $\varepsilon \to 0$, this set decreases to the boundary set

$$\partial A(r) = \left\{x \in D \,\big|\, |x - x_0| = r(h(x - x_0))\right\}.$$

Since the graph of a continuous function has Lebesgue measure zero, we deduce that $|\partial A(r)| = 0$. It follows that

$$\lim_{\varepsilon \to 0} |A(r_\varepsilon)\Delta A(r)| = 0.$$

To conclude, note that

$$|F_2(r_\varepsilon, x_0) - F_2(r, x_0)| \leq |u_+ - u_-||\mathbb{1}_{A(r_\varepsilon)} - \mathbb{1}_{A(r_\varepsilon)}| = C\mathbb{1}_{A(r_\varepsilon)\Delta A(r)}.$$

---

[3]A sequence of functions $(f_\varepsilon)_{\varepsilon>0}$, $f_\varepsilon : D \to \mathbb{R}$, is said to converge in measure to a function $f : D \to \mathbb{R}$ if for all $\delta > 0$, $|\{x \in D \,|\, |f_\varepsilon(x) - f(x)| > \delta\}| \to 0$. Here $|B|$ denotes the Lebesgue measure of a set $B \subseteq \mathbb{R}^d$.

(ii) Let $r \in C_P^0(R_{d-1})$ be Lipschitz continuous. Denote $A(x_0) = A(r, x_0)$. Let $(x_0^\varepsilon) \subseteq D$ be a sequence of points such that $|x_0^\varepsilon - x_0| \to 0$. Note that we may write

$$
\begin{aligned}
A(x_0^\varepsilon) &= \{x \in D \mid |x - x_0^\varepsilon| \le r(h(x - x_0^\varepsilon))\} \\
&= \{x \in \mathbb{R}^d \mid |x - x_0^\varepsilon| \le r(h(x - x_0^\varepsilon))\} \cap D \\
&= ((x_0^\varepsilon - x_0) + \{x \in \mathbb{R}^d \mid |x - x_0| \le r(h(x - x_0))\}) \cap D \\
&=: ((x_0^\varepsilon - x_0) + A(x_0)^*) \cap D.
\end{aligned}
$$

By the distributivity of intersection over symmetric difference, we then have that

$$
\begin{aligned}
A(x_0^\varepsilon)\Delta A(x_0) &= [((x_0^\varepsilon - x_0) + A(x_0)^*) \cap D]\Delta[A(x_0)^* \cap D] \\
&= [((x_0^\varepsilon - x_0) + A(x_0)^*)\Delta A(x_0)^*] \cap D \\
&\subseteq ((x_0^\varepsilon - x_0) + A(x_0)^*)\Delta A(x_0)^*.
\end{aligned}
$$

Therefore, using Theorem 1 from [39], we see that

$$
\begin{aligned}
|A(x_0^\varepsilon)\Delta A(x_0)| &\le |((x_0^\varepsilon - x_0) + A(x_0)^*)\Delta A(x_0)^*| \\
&\le |x_0^\varepsilon - x_0|\mathcal{H}^{d-1}(\partial A(x_0)^*)
\end{aligned}
$$

where $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure. Since we assume that r is Lipschitz, the surface area $\mathcal{H}^{d-1}(\partial A(x_0)^*)$ of the boundary of $A(x_0)^*$ is finite, and so it follows that

$$
\lim_{\varepsilon \to 0} |A(x_0^\varepsilon)\Delta A(x_0)| = 0.
$$

As before, we conclude by noting that

$$
|F_2(r, x_0^\varepsilon) - F_2(r, x_0)| \le |u_+ - u_-||\mathbb{1}_{A(x_0^\varepsilon)} - \mathbb{1}_{A(x_0)}| = C\mathbb{1}_{A(x_0^\varepsilon)\Delta A(x_0)}.
$$

(iii) We have that

$$
\begin{aligned}
|F_2(r_\varepsilon, x_0^\varepsilon) - F_2(r, x_0)| &\le |F_2(r_\varepsilon, x_0^\varepsilon) - F_2(r, x_0^\varepsilon)| + |F_2(r, x_0^\varepsilon) - F_2(r, x_0)| \\
&\le C(\mathbb{1}_{A(r_\varepsilon, x_0^\varepsilon)\Delta A(r, x_0^\varepsilon)} + \mathbb{1}_{A(r, x_0^\varepsilon)\Delta A(r, x_0)}) \\
&\le C\mathbb{1}_{[A(r_\varepsilon, x_0^\varepsilon)\Delta A(r, x_0^\varepsilon)] \cup [A(r, x_0^\varepsilon)\Delta A(r, x_0)]}.
\end{aligned}
$$

Now note that

$$
|A(r_\varepsilon, y_0)\Delta A(r, y_0)| \le |A(r_\varepsilon, y_0)^*\Delta A(r, y_0)^*|.
$$

The right hand-side is independent of y_0 by translation invariance of the Lebesgue measure. By the same argument as part (i) we conclude that it tends to zero. We then have that

$$
\begin{aligned}
|[A(r_\varepsilon, x_0^\varepsilon)\Delta A(r, x_0^\varepsilon)] &\cup [A(r, x_0^\varepsilon)\Delta A(r, x_0)]| \\
&\le |A(r_\varepsilon, x_0^\varepsilon)\Delta A(r, x_0^\varepsilon)| + |A(r, x_0^\varepsilon)\Delta A(r, x_0)| \\
&\le \sup_{y_0 \in D} |A(r_\varepsilon, y_0)\Delta A(r, y_0)| + |A(r, x_0^\varepsilon)\Delta A(r, x_0)|
\end{aligned}
$$

which tends to zero by the discussion above and part (ii).

$\square$

Remark 3.3. Above we assumed that $r : R_{d-1} \to \mathbb{R}$ was Lipschitz continuous. This assumption is only used in the proof of part (ii) of the proposition. If the centre of the star-shaped region is known, this assumption may then be dropped to allow for rougher boundaries.

We need to choose a prior measure μ_0 on (r, x_0). We equip $C_P^0(R_{d-1}) \times D$ with the norm $\|(r, x_0)\| = \max\{\|r\|_\infty, |x_0|\}$ and corresponding Borel σ-algebra. We assume that r and x_0 are independent under the prior so that we may factor $\mu_0 = \sigma_0 \otimes \tau_0$ where σ_0 is a measure on $C_P^0(R_{d-1})$ and τ_0 is a measure on D. We will assume that σ_0 is such that $\sigma_0(B) > 0$ for all balls $B \subseteq C_P^0(R_{d-1})$.

Example 3.4. Consider the case $D = B(0, 1) \subseteq \mathbb{R}^2$. Suppose that r is drawn from a log-Gaussian measure σ_0 on $C_P^0((-\pi, \pi])$, and x_0 is drawn from $\tau_0 = U([-0.5, 0.5]^2)$. Note that $[-0.5, 0.5]^2 \subseteq B(0, 1)$. Typical samples from $F_2^\#(\mu_0)$ are shown below. The covariance of σ_0 is chosen such that the samples r almost surely have regularity $r \in C^{\lfloor s \rfloor, s - \lfloor s \rfloor}((-\pi, \pi])$ for all $s < t$, where from left to right $t = 2.5, 2, 1.5, 1$ respectively.
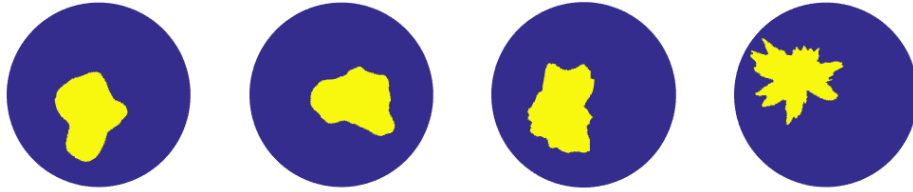
3.1.3. *Level set prior.* We finally consider the case where the inclusions can be described by a single level set function, as in [17]. Let $n \in \mathbb{N}$ and fix constants $-\infty = c_0 < c_1 < \ldots < c_n = \infty$. Given $u : D \to \mathbb{R}$, define $D_i \subseteq D$ by

$$D_i = \{x \in D \mid c_{i-1} \leq u(x) < c_i\}, \quad i = 1, \ldots, n$$

so that $\overline{D} = \bigcup_{i=1}^n \overline{D}_i$ and $D_i \cap D_j = \varnothing$ for $i \neq j$, $i, j \geq 1$. Define also the level sets

$$D_i^0 = \overline{D}_i \cap \overline{D}_{i+1} = \{x \in D \mid u(x) = c_i\}, \quad i = 1, \ldots, n-1.$$

Now given strictly positive functions $f_1, \ldots, f_n \in C^0(\overline{D})$, we define the map $F_3 : C^0(\overline{D}) \to \mathcal{A}(D)$ by

$$F_3(u) = \sum_{i=1}^n f_i \mathbb{1}_{D_i}.$$

Since each f is continuous and strictly positive on a compact set $\overline{D}$, they are uniformly bounded above and below by positive constants, and so F_3 does indeed map into $\mathcal{A}(D)$.

In this paper we are primarily concerned with the case of binary fields, $n = 2$ and f_i constant above, however the theory in proved in the general case. We have the following result regarding continuity of this map, by the same arguments as in [17].

Proposition 3.5. *Define the map $F_3 : C^0(\overline{D}) \to \mathcal{A}(D)$ as above. Let $u \in C^0(\overline{D})$ be such that $|D_i^0| = 0$ for $i = 1, \ldots, n-1$. Suppose that $(u_\varepsilon)_{\varepsilon>0} \subseteq C^0(\overline{D})$ is an approximating sequence of functions so that $\|u_\varepsilon - u\|_\infty \to 0$. Then $F_3(u_\varepsilon) \to F_3(u)$ in measure.*

$F_3(\cdot)$ is continuous at u $F_3(\cdot)$ is discontinuous at u
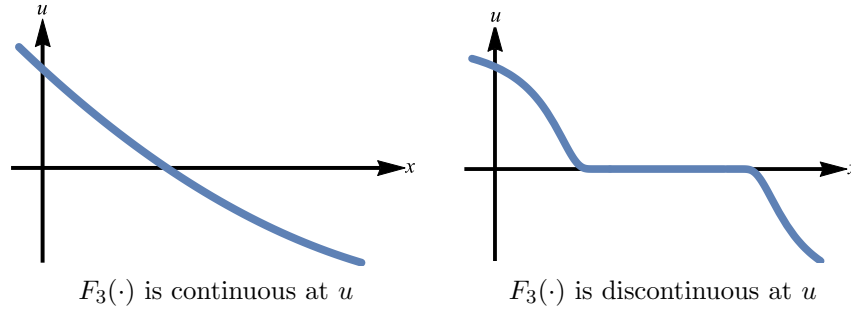
FIGURE 2. The discontinuity of F_3 into $L^p(D)$.

Proof. Denote by $D_{i,\varepsilon}$ and $D_{i,\varepsilon}^0$ the sets as defined above associated with the approximating functions u_ε. We can write

$$F_3(u_\varepsilon) - F_3(u) = \sum_{i=1}^{n}\sum_{j=1}^{n}(f_i - f_j)\mathbb{1}_{D_{i,\varepsilon}\cap D_j} = \sum_{\substack{i,j=1 \\ i\neq j}}^{n}(f_i - f_j)\mathbb{1}_{D_{i,\varepsilon}\cap D_j}.$$

Since $\|u_\varepsilon - u\|_\infty \to 0$, there exists $\gamma(\varepsilon) \to 0$ with $\|u_\varepsilon - u\|_\infty < \gamma(\varepsilon)$. Then we have for all $x \in D$ and $\varepsilon > 0$

$$u(x) - \gamma(\varepsilon) < u_\varepsilon(x) < u(x) + \gamma(\varepsilon).$$

Hence for $|j - i| > 1$ and ε sufficiently small, $D_{i,\varepsilon}\cap D_j = \varnothing$. If $|j - i| = 1$, then

$$D_{i,\varepsilon}\cap D_{i+1} \subseteq \widetilde{D}_{i,\varepsilon} := \{x \in D \,|\, c_i \leq u(x) < c_i + \gamma(\varepsilon)\} \to D_i^0,$$

$$D_{i,\varepsilon}\cap D_{i-1} \subseteq \hat{D}_{i-1,\varepsilon} := \{x \in D \,|\, c_i - \gamma(\varepsilon) \leq u(x) < c_i\} \to \varnothing.$$

By the uniform boundedness of the (f_i), for sufficiently small ε we can then write

$$|F_3(u_\varepsilon) - F_3(u)| \leq \sum_{i=1}^{n-1}|f_i - f_{i+1}|\mathbb{1}_{\widetilde{D}_{i,\varepsilon}} + \sum_{i=2}^{n}|f_i - f_{i-1}|\mathbb{1}_{\hat{D}_{i-1,\varepsilon}}$$

(6)
$$\leq C\mathbb{1}_{Z_\varepsilon}$$

where $Z_\varepsilon \subseteq D$ is given by

$$Z_\varepsilon = \left(\bigcup_{i=1}^{n-1}\widetilde{D}_{i,\varepsilon}\right) \cup \left(\bigcup_{i=2}^{n}\hat{D}_{i-1,\varepsilon}\right) \to \bigcup_{i=1}^{n-1}D_i^0.$$

By the assumption that $|D_i^0| = 0$ for all i, it follows that $|Z_\varepsilon| \to 0$, and so the result follows from the comment at the start of the proof of Proposition 3.2. $\qquad\square$

Note that bound (6) actually above implies the slightly stronger result that, when the c_i-level sets of $u \in X$ have zero measure, F_3 is continuous into $L^p(D)$, $1 \leq p < \infty$, at u. The assumption that the level sets have zero measure is an important one, as illustrated by Figure 2: an arbitrarily small perturbation of u can lead to an order 1 change in $F_3(u)$.

In the Bayesian approach we are taking to this problem, we may choose a prior measure on u such that, almost surely, the Lebesgue measure of the level sets is zero. This is shown to hold for non-degenerate Gaussian measures in [17]. As a result, F_3 will be almost surely continuous under the prior, and this is enough to give the measurability required in Bayes' theorem, as shown in [17].

As in the log-Gaussian case, we take our prior measure μ_0 on u to be a Gaussian measure $N(m_0, \mathcal{C}_0)$ on $C^0(\overline{D})$.

Example 3.6. Consider the case $D = B(0,1) \subseteq \mathbb{R}^2$, $n = 2$, $c_1 = 0$, $f_1 \equiv 1$ and $f_2 \equiv 2$. Suppose that u is drawn from a centred Gaussian measure $\mu_0 = N(0, \mathcal{C})$ on $C^0(\overline{D})$. The covariance $\mathcal{C}$ is chosen such that the samples u almost surely have regularity $u \in C^{\lfloor s \rfloor, s - \lfloor s \rfloor}(D)$ for all $s < t$, where from left to right $t = 4, 3, 2, 1$ respectively. As in the log-Gaussian case, here the samples are generated on $[-1,1]^2 \supseteq D$ and then restricted to D, for computational simplicity.
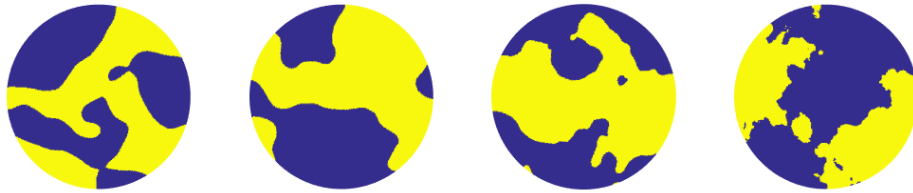
Remark 3.7. In the cases of the star-shaped and level set priors, we have assumed the values that the conductivity takes are known a priori. This may be appropriate in certain situations, for example in medical contexts where conductivities of different types of tissue are known [7]. However it may be preferable to have the flexibility of treating the conductivity values as part of the inference. The theory does not become significantly more involved, as it can be shown that the conclusions of propositions 3.2 and 3.5 still hold when we allow for perturbation of the conductivity values as well. We work with the case of fixed conductivity values for clarity of presentation, but we note that it is possible to place a prior upon these.

3.2. **The likelihood and posterior distribution.** The inverse problem was introduced at the beginning of the section. Now that we have introduced prior distributions, we may provide the Bayesian formulation of the problem.

Let X be a separable Banach space and $F : X \to \mathcal{A}(D)$ a map from the space X where the unknown parameters live to the conductivity space. Choose a set of current stimulation patterns $I^{(j)} \in \mathbb{R}^L$, $j = 1, \dots, J$ and let $\mathcal{M}_j : \mathcal{A}(D) \to \mathbb{H}$ denote the solution map when using stimulation pattern $I^{(j)}$. Recall the projection map $\Pi : \mathbb{H} \to \mathbb{R}^L$ was defined by $\Pi(v, V) = V$.

The data y_j from the jth stimulation pattern is assumed to arise from the map $\mathcal{G}_j : X \to \mathbb{R}^L$, $\mathcal{G}_j = \Pi \circ \mathcal{M}_j \circ F$, via

$$y_j = \mathcal{G}_j(u) + \eta_j, \quad \eta_j \sim N(0, \Gamma_0) \text{ iid.}$$

We concatenate these observations to get data $y \in \mathbb{R}^{JL}$ given by

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathbb{Q}_0 := N(0, \Gamma)$$

where $\Gamma = \text{diag}(\Gamma_0, \dots, \Gamma_0)$ and $\mathcal{G} : X \to \mathbb{R}^{JL}$. This coincides with the setup at the start of the section, with $\sigma = F(u)$.

Assume that $u \sim \mu_0$, where μ_0 is independent of $\mathbb{Q}_0$. From the above, we see that $y|u \sim \mathbb{Q}_u := N(\mathcal{G}(u), \Gamma)$. We use this to find the distribution of $u|y$. First note that

$$\frac{d\mathbb{Q}_u}{d\mathbb{Q}_0}(y) = \exp\left(-\Phi(u; y) + \frac{1}{2}|y|_\Gamma^2\right)$$

where the potential (or negative log-likelihood) $\Phi : X \times Y \to \mathbb{R}$ is given by

$$(7) \qquad \Phi(u; y) = \frac{1}{2}|\mathcal{G}(u) - y|_\Gamma^2 := \frac{1}{2}|\Gamma^{-\frac{1}{2}}(\mathcal{G}(u) - y)|^2.$$

Then under suitable regularity conditions, Bayes' theorem tells us that the distribution μ^y of $u|y$ is as given below:

Theorem 3.8 (Existence and Well-Posedness). *Let $(X, \mathcal{F}, \mu_0)$ denote any of the probability spaces associated with any of the three priors introduced in the previous subsection, and let $\Phi : X \times Y \to \mathbb{R}$ be the potential (7) associated with the corresponding forward map. Then the posterior distribution μ^y of the state u given data y is well-defined. Furthermore, $\mu^y \ll \mu_0$ with Radon-Nikodym derivative*

$$(8) \qquad \frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) = \frac{1}{Z_\mu}\exp(-\Phi(u; y))$$

where for y $\mathbb{Q}_0$-a.s.,

$$Z_\mu := \int_X \exp(-\Phi(u; y))\,\mu_0(\mathrm{d}u) > 0.$$

Additionally, the posterior measure μ^y is locally Lipschitz with respect to y, in the Hellinger distance: for all $y, y' \in Y$ with $\max\{|y|_\Gamma, |y'|_\Gamma\} < \rho$, there exists $C = C(\rho) > 0$ such that

$$d_{\mathrm{Hell}}(\mu^y, \mu^{y'}) \le C|y - y'|_\Gamma.$$

In the proof of the above theorem we will make use of the following version of Bayes' theorem from [10].

Proposition 3.9 (Bayes' theorem). *Define the measure $\nu_0(\mathrm{d}u, \mathrm{d}y) = \mu_0(\mathrm{d}u)\mathbb{Q}_0(\mathrm{d}y)$ on $X \times Y$. Assume that $\Phi : X \times Y \to \mathbb{R}$ is ν_0-measurable and that, for y $\mathbb{Q}$-a.s.*

$$Z_\mu = \int_X \exp(-\Phi(u; y))\,\mu_0(\mathrm{d}u) > 0.$$

Then the conditional distribution of $u|y$ exists and is denoted by μ^y. Furthermore $\mu^y \ll \mu_0$ and, for y $\mathbb{Q}_0$-a.s.,

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0} = \frac{1}{Z_\mu}\exp\left(-\Phi(u; y)\right).$$

We need to verify that the assumptions of this theorem are satisfied. To proceed we first give some regularity properties of the potential Φ:

Proposition 3.10. *Let $(X, \mathcal{F}, \mu_0)$ denote any of the probability spaces associated with the priors introduced in the previous subsection. Then the potential $\Phi : X \times Y \to \mathbb{R}$ associated with the corresponding forward map, given by (7), admits the following properties.*

(i) *There is a continuous $K : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ such that for every $\rho > 0$, $u \in X$ and $y \in Y$ with $|y|_\Gamma < \rho$,*

$$0 \le \Phi(u; y) \le K(\rho, \|u\|_X).$$

In the cases $F = F_2$ and $F = F_3$, K has no dependence on $\|u\|_X$.

(ii) *For any fixed $y \in Y$, $\Phi(\cdot; y) : X \to \mathbb{R}$ is continuous μ_0-almost surely on the probability space $(X, \mathcal{F}, \mu_0)$.*

(iii) There exists $C : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ *such that for every* $y_1, y_2 \in Y$ *with* $\max\{|y_1|_\Gamma, |y_2|_\Gamma\} < \rho$, *and every* $u \in X$,

$$|\Phi(u; y_1) - \Phi(u; y_2)| \leq C(\rho, \|u\|_X)|y_1 - y_2|_\Gamma.$$

Moreover, $C(\rho, \|\cdot\|_X) \in L^2_{\mu_0}(X)$ *for all* $\rho > 0$.

Proof.   (i) From equation (5) in the proof of Proposition 2.7, we see that there exists $C > 0$ such that each $\mathcal{M}_j : \mathcal{A}(D) \to \mathbb{H}$ satisfies

$$\|\mathcal{M}_j(\sigma_1) - \mathcal{M}_j(\sigma_2)\|_* \leq C\|\mathcal{M}_j(\sigma_2)\|_*\|\sigma_1 - \sigma_2\|_\infty$$

for all $\sigma_1, \sigma_2 \in \mathcal{A}(D)$. Taking $\sigma_2 \equiv 1$, say, we deduce that

$$\|\mathcal{M}_j(\sigma_1)\|_* \leq C\|\mathcal{M}_j(1)\|_*\|\sigma_1 - 1\|_\infty + \|\mathcal{M}_j(1)\|_* \leq C(1 + \|\sigma_1\|_\infty).$$

Hence $\|\sigma\|_\infty < \rho$ implies that $\|\mathcal{M}_j(\sigma)\|_* < C(1 + \rho)$. By Corollary 2.8, it follows that $\Pi \circ \mathcal{M}_j : \mathcal{A}(D) \to \mathbb{R}^L$ is bounded on bounded sets with respect to $\|\cdot\|_\infty$ for all j.

In the case $F = F_1$, if $u \in X$ then $\|F(u)\|_\infty \leq e^{\|u\|_X}$. It follows that $|\mathcal{G}(u)|_\Gamma \leq \max_j |\mathcal{G}_j(u)|_\Gamma \leq C(1 + e^{\|u\|_X})$.

Now note that

$$\Phi(u; y) \leq |\mathcal{G}(u)|_\Gamma^2 + |y|_\Gamma^2.$$

Then for any $y \in Y$ with $|y| < \rho$, we may bound

$$\Phi(u; y) \leq C(1 + e^{2\|u\|_X} + \rho^2) =: K(\rho, \|u\|_X).$$

In the cases $F = F_2$ and $F = F_3$, we have that $\|F(u)\|_\infty$ is bounded uniformly over $u \in X$ and so $|\mathcal{G}(u)|_\Gamma \leq \max_j |\mathcal{G}_j(u)|_\Gamma \leq C$. Hence we obtain the bound

$$\Phi(u; y) \leq C(1 + \rho^2) =: K(\rho).$$

(ii) Let $u \sim \mu_0$ and suppose $F : X \to \mathcal{A}(D)$ is such that $\|u_\varepsilon - u\|_X \to 0$ implies that $F(u_\varepsilon) \to F(u)$ either uniformly or in measure. Then Proposition 2.7 tells us that $\mathcal{M}_j \circ F : X \to \mathbb{H}$ is continuous at u for each j. The projection $\Pi : \mathbb{H} \to \mathbb{R}^L$ is continuous, and so $\mathcal{G}_j = \Pi \circ \mathcal{M}_j \circ F$ is continuous at u for each j. In §3.1 it is shown that this is true for $F = F_1$ and $F = F_2$ for any u. For $F = F_3$ it is only true at points u whose level sets have zero measure, however since we are assuming $u \sim \mu_0$, a Gaussian measure, it follows from Proposition 7.2 in [17] that u μ_0-almost surely has this property.

(iii) Let $u \in X$ and $y_1, y_2 \in Y$ with $\max\{|y_1|_\Gamma, |y_2|_\Gamma\} < \rho$. Then we have

$$\begin{aligned}
|\Phi(u; y_1) - \Phi(u; y_2)| &= \frac{1}{2}|\langle y_1 + y_2 - 2\mathcal{G}(u), y_1 - y_2 \rangle_\Gamma| \\
&\leq \frac{1}{2}(|y_1|_\Gamma + |y_2|_\Gamma + 2|\mathcal{G}(u)|_\Gamma)|y_1 - y_2|_\Gamma \\
&\leq (\rho + |\mathcal{G}(u)|_\Gamma)|y_1 - y_2|_\Gamma \\
&=: \tilde{C}(\rho, u)|y_1 - y_2|_\Gamma
\end{aligned}$$

We now consider cases separately based on the prior. In the log-Gaussian case, we may bound

$$\tilde{C}(\rho, u) \leq C(1 + \rho + e^{\|u\|_X}) =: C(\rho, \|u\|_X)$$

using the bound from the proof of part (i). Square-integrability of $C(\rho, \|\cdot\|_X)$ follows since Gaussians have exponential moments.

In the star-shaped and level set prior cases, we have that $|\mathcal{G}(u)|$ is bounded uniformly by a constant. We may hence bound $\tilde{C}(\rho, u)$ above by some

$C(\rho, \|u\|_X) := C(1 + \rho)$ that is independent of u, and so again the square-integrability follows.

$\square$

Proof of Theorem 3.8. Define the product measure $\nu_0(\mathrm{d}u, \mathrm{d}y) = \mu_0(\mathrm{d}u)\mathbb{Q}_0(\mathrm{d}y)$ on $X \times Y$. We showed in Proposition 3.10 that $\Phi(\cdot; y) : X \to \mathbb{R}$ is almost-surely continuous under the prior for all $y \in Y$, and $\Phi(u; \cdot) : Y \to \mathbb{R}$ is locally Lipschitz for all $u \in X$. Together these imply that $\Phi : X \times Y \to \mathbb{R}$ is almost-surely jointly continuous under ν_0. To see this, let $(u, y) \in X \times Y$ and let $(u_n, y_n)_{n \geq 1} \subseteq X \times Y$ be an approximating sequence so that $\|u_n - u\|_X \to 0$ and $|y_n - y|_\Gamma \to 0$. Then we have

$$|\Phi(u_n, y_n) - \Phi(u, y)| \leq |\Phi(u_n, y_n) - \Phi(u_n, y)| + |\Phi(u_n, y) - \Phi(u, y)|.$$

The second term tends to zero μ_0-almost surely by continuity. For the first term, note that the sequences $(\|u_n\|_X)_{n \geq 1}$ and $(|y_n|_\Gamma)_{n \geq 1}$ are bounded, by K and R respectively, say. Then we can use the local Lipschitz property to deduce that

$$|\Phi(u_n, y_n) - \Phi(u_n, y)| \leq C(R, K)|y_n - y|_\Gamma$$

since $C(\cdot, \cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, as defined in the proof of Proposition 3.10, is monotonically increasing in both components. Therefore this term tends to zero, and we obtain the desired continuity. It follows, see for example Lemma 6.1 in [17], that Φ is ν_0-measurable.

For the lower bound on Z_μ, we consider cases separately based on the prior. First we consider the log-Gaussian and level set prior cases so that μ_0 is Gaussian. Let $B \subseteq X$ be any ball. Fix any $\rho > |y|_\Gamma$ and define

$$R = \sup_{u \in B} K(\rho, \|u\|_X)$$

where K is the upper bound from Proposition 3.10(i). This supremum is finite by the continuity of K. Then we have

$$\begin{aligned}
\int_X \exp(-\Phi(u; y)) \, \mu_0(\mathrm{d}u) &\geq \int_B \exp(-\Phi(u; y)) \, \mu_0(\mathrm{d}u) \\
&\geq \int_B \exp(-K(\rho, \|u\|)) \, \mu_0(\mathrm{d}u) \\
&\geq \exp(-R)\mu_0(B).
\end{aligned}$$

Since μ_0 is Gaussian, $\mu_0(B) > 0$ and so $Z_\mu > 0$.

In the star-shaped prior case, proceed as above but take $B = B_1 \times D$ where $B_1 \subseteq C_P^0(R_{d-1})$ is any ball. Then we have

$$\mu_0(B) = (\sigma_0 \times \tau_0)(B_1 \times D) = \sigma_0(B_1)\tau_0(D) > 0$$

by the assumption that σ_0 assigns positive mass to balls, and so again $Z_\mu > 0$. The above hold for all $y \in Y$, and so in particular for y $\mathbb{Q}_0$-almost-surely. We may now apply Bayes' Theorem 3.9 to obtain the existence of μ^y.

The proof of well-posedness is almost identical to that of the analogous result Theorem 2.2 in [17] and is hence omitted.
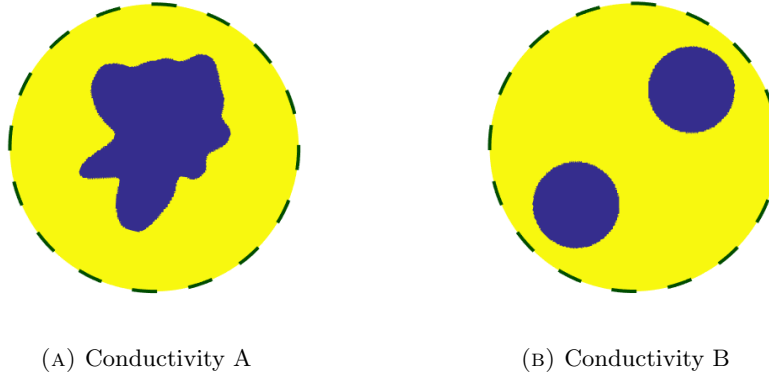
$\square$

(A) Conductivity A   (B) Conductivity B

FIGURE 3. The two binary fields we attempt to recover. Conductivity A is drawn from the star-shaped prior, with $\sigma_0 = h^{\#}\big[N(0.5, 10^9 \cdot (30^2 - \mathcal{A}_D)^{-3})\big]$, $h(z) = (1 + \tanh z)/2$, and $\tau_0 = U([-0.5, 0.5]^2)$. Conductivity B is constructed explicitly, rather than being drawn from a prior.

4. **Numerical experiments.** We investigate the effect of the choice of prior on the recovery of certain binary conductivity fields. The specific fields we consider are shown in Figure 3, where blue represents a conductivity of 1 and yellow a conductivity of 2. Simulations are performed using the EIDORS software [1] to solve the forward model using a first order finite element method; a mesh of 43264 elements is used to create the data and a mesh of 10816 elements is used for simulations in order to avoid an inverse crime [20].

In subsection 4.1 we describe the MCMC sampling algorithm that we will use. In subsection 4.2 we define the parameters we will use for the forward model and the MCMC simulations. We also describe how the data is created, and define our choices of prior distributions. Finally in subsection 4.3 we present the results of the simulation, looking at quality of reconstruction, convergence of the algorithm and some properties of the posterior distribution.

4.1. **Sampling algorithm.** We aim to produce a sequence of samples from μ^y on X, where μ^y is given by (8). We make use of the preconditioned Crank-Nicolson Markov Chain Monte Carlo (pCN-MCMC) method. The pCN-MCMC method is a modification of the standard Random Walk Metropolis (RWM) MCMC method which is well-adapted to Gaussian priors in high dimensions. It was introduced in [5], and its dimension independent properties are analysed and illustrated numerically in [14] and [9] respectively; the pCN nomenclature was introduced in [9]. In the case of the star-shaped prior, we use a Metropolis-within-Gibbs algorithm [43], alternately updating the field with the pCN method above and updating the centre with the standard RWM method.

An advantage of these MCMC methods is that derivatives of the forward map are not needed, only black-box solution of the forward model. However in order to accurately compute some quantity of interest, such as the conditional mean, we may need to produce a very large number of samples and tuning the algorithm to minimise this effect is important. For this reason we compute the effective sample

size from the integrated autocorrelation (neglecting a burn-in period) of a quantity of interest, as in [21].

4.2. **Data and parameters.** We work on a circular domain of radius 1, with 16 equally spaced electrodes on its boundary providing 50% coverage. We take all contact impedances $z_l = 0.01$. We stimulate adjacent electrodes with a current of 0.1, so that the matrix of stimulation patterns $I = (I^{(j)})_{j=1}^{15} \in \mathbb{R}^{16 \times 15}$ is given by

$$I = 0.1 \times \begin{pmatrix} +1 & 0 & \cdots & 0 \\ -1 & +1 & \cdots & 0 \\ 0 & -1 & \ddots & 0 \\ \vdots & \vdots & \ddots & +1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

The conductivity is chosen such that it takes values 1 and 2. We perturb the measurements with white noise $\eta \sim N(0, \gamma^2 I)$, $\gamma = 0.0002$, so that the mean relative error on both sets of data is approximately 10%. The true conductivity fields used to generate the data, henceforth referred to as *Conductivity A* and *Conductivity B*, are shown in Figure 3. In all cases we generate $N = 2.5 \times 10^6$ samples with a burn-in of $k_0 = 5 \times 10^5$ samples.

Our priors on fields will make use of Gaussians with covariances of the form

(9) $$\mathcal{C} = q(\tau^2 - \Delta)^{-\alpha}.$$

These are essentially rescaled Whittle-Matern covariances [31], with τ representing the inverse length scale of the samples, α proportional to their regularity, and q proportional to their amplitude.

In what follows, denote by $\mathcal{A}_N$ the Laplacian with Neumann boundary conditions on $[-1, 1]^2$, restricted to D, so that its domain is given by

$$\mathcal{D}(\mathcal{A}_N) = \left\{ u|_D \ \middle| \ u \in H^2([-1, 1]^2), \frac{\partial u}{\partial n} = 0 \right\}.$$

Defining the Laplacian first on a square and then restricting to D will allow for efficient generation of Gaussian samples via the fast Fourier transform. Note that if we were to consider priors of the form (9) with $\tau = 0$, we should restrict $\mathcal{D}(\mathcal{A}_N)$ further to ensure the invertibility of $\mathcal{A}_N$.

Additionally, denote by $\mathcal{A}_D$ the Laplacian with Dirichlet boundary conditions on $R_1 = (-\pi, \pi]$, so that its domain is given by

$$\mathcal{D}(\mathcal{A}_D) = \left\{ u \in H^2((-\pi, \pi]) \ \middle| \ u(-\pi) = u(\pi) = 0 \right\}.$$

4.2.1. *Gaussian prior.* States are defined on a grid of $2^7 \times 2^7$ points. For both simulations the pCN jump parameter β is taken to be 0.01, with choice of prior

$$\mu_0 = \exp^{\#} \left[N(0.5 \log 2, 10^{16} \cdot (40^2 - \mathcal{A}_N)^{-6}) \right].$$

4.2.2. *Star-shaped prior.* Radial states are defined on a grid of 2^8 points. For Conductivity A, we choose the pCN jump parameter $\beta = 0.03$ and the RWM jump parameter $\delta = 0.01$. For Conductivity B we choose $\beta = 0.01$ and $\delta = 0.005$. For both simulations we use the choice of prior $\mu_0 = \sigma_0 \times \tau_0$, with

$$\sigma_0 = h^{\#} \left[N(0.5, 10^9 \cdot (30^2 - \mathcal{A}_D)^{-3}) \right], \quad \tau_0 = U([-0.5, 0.5]^2),$$

where $h(z) = (1 + \tanh z)/2$.

4.2.3. *Level set prior.* States are defined on a grid of $2^7 \times 2^7$ points. For both simulations the pCN jump parameter β is taken to be 0.005, with choice of prior

$$\mu_0 = N(0, (35^2 - \mathcal{A}_N)^{-5}).$$

4.3. Results.

4.3.1. *Recovery.* Figures 4 and 5 show conductivities arising from the MCMC chains, and Figure 6 shows the values of the misfit Φ at the different sample means. We consider two different types of mean conductivities. First the sample means are calculated in the sample spaces X_i and then pushed forward to the conductivity space by the maps F_i, so that we produce estimates of $F_i(\mathbb{E}(u))$. This preserves the binary nature of the fields in the cases of the star-shaped and level set priors. We also consider the sample means of the pushforwards of the posteriors by F_i, $\mathbb{E}(F_i(u))$, which do not preserve the binary property but provide some visualisation of the interface uncertainty in the posterior.

For Conductivity A, the sample mean $F_2(\mathbb{E}(u))$ arising from the star-shaped prior provides a better reconstruction than the other two prior choices. This is expected, since the true conductivity was drawn from this prior. Whilst the sample means arising from the level set prior are fairly close to the true conductivity (both visually and in terms of Φ), the boundary of the interface appears to have too large a length-scale. Appropriate choice of prior length-scale is a key issue the with the level set method; treating the length-scale hierarchically as another unknown in the problem may be beneficial. On the other hand

The sample means arising from the Gaussian prior fail to recover both the sharp interface and the values of the conductivity, which is reflected in the large values Φ takes. Nonetheless, general qualitative properties of the true conductivity can be seen in the sample means. As with the level set prior, reconstruction could be improved by treating hierarchically the parameters in the prior such as length-scale, marginal variance and mean.

For Conductivity B, the level set prior is most effective in the reconstruction, since a specific number of inclusions isn't fixed a priori as it is for the star-shaped prior. Again the Gaussian prior fails to recover both the sharp interface and the values of the conductivity, however it appears to do a better job than the star-shaped prior at identifying the location and shape of the two inclusions.

In both of the above cases, even though the individual samples coming from using the level set prior contain many small inclusions, these do not show up in the sample means $F_3(\mathbb{E}(u))$, though they are visible in the means $\mathbb{E}(F_3(u))$.

In terms of the values of the misfit as shown in Figure 6, in all but one of the above cases the mean $\mathbb{E}(F_i(u))$ provides a better fit than $F_i(\mathbb{E}(u))$ despite the lack of the binary property. The quality of the recovery cannot be assessed solely on the misfit however; an advantage of the Bayesian approach is that both means are available (along with the full posterior), and whichever is more appropriate could depend on the context.

4.3.2. *Convergence.* In Figure 7, we show the approximate effective sample size (ESS) associated with different quantities of interest. For all choices of prior, these are significantly smaller than the total 2.5×10^6 samples generated. Many more samples may hence be required to produce accurate approximations of the posterior mean.

The chain associated with the star-shaped prior results in the largest ESS, likely because we are only attempting to infer $2^8 + 2$ parameters rather than 2^{14} parameters as in the log-Gaussian and level set cases.

In order to accelerate the convergence of the MCMC we can adjust the jump parameters β and δ. Larger choices of these parameters mean that accepted states will be less correlated with the current state, however the proposed states are less likely to be accepted. The choice $\beta = 1$ in pCN produces proposed states that are independent of the current state, but dependent upon how far the prior is from the posterior, very few or no states may be accepted so that the chain never moves. Similarly, smaller choices of these jump parameters mean that more proposals will be accepted, but the states will be more correlated. A balance hence must be achieved – in our simulations we choose the parameters such that approximately 20-30% of proposals are accepted, though in general the optimal acceptance rate is not known [3].

Alternatively, reconstruction may be accelerated by looking at an approximation of the posterior instead of the exact posterior, for example using the ensemble Kalman filter [16] or a sequential Monte Carlo method [4]. We could also initialise the MCMC chains from EnKF estimates to significantly reduce the burn-in period and hence computational cost. If the derivative of the forward map is available, Hybrid Monte Carlo (HMC) methods could be used to accelerate the convergence [11]. Emulators could also be used to reduce the computational burden of derivative calculation, allowing the use of geometric MCMC methods such as Riemannian Manifold Hamiltonian Monte Carlo (RHMC) and Lagrangian Monte Carlo (LMC) [27].

4.3.3. *Posterior behaviour.* In Figures 8-10 we show kernel density estimates for a number of quantities associated with each posterior distribution. These are calculated using subsequences of the MCMC chains, with lengths of the same order of magnitude as the effective sample sizes, in order to avoid over-fitting. The most regular densities arise in the star-shaped case, with the distribution of all quantities appearing to be very close to uni-modal. More irregularity is seen for the log-Gaussian case, especially in the joint distributions, but they are still very close to uni-modal.

The least regular, more multi-modal densities come from the level set prior. One reason for this is likely the lack of identifiability of the level set function: the forward model only 'sees' the zero level set of the state, and hence cannot distinguish between infinitely many different states. The prior can however distinguish between these states, and will weight them appropriately, which can help explain the shape of the posterior densities.

5. **Conclusions.** The primary contributions of this paper are:

- We have formulated the EIT problem rigorously in the infinite dimensional Bayesian framework.
- We have studied three different prior models, each with their own advantages and disadvantages based on prior knowledge and the nature of the field we are trying to recover.
- With each of these choices of prior we obtain well-posedness of the problem. We can obtain well-posedness using additional prior models, as long as the mapping from the state space to the conductivity space has appropriate regularity.
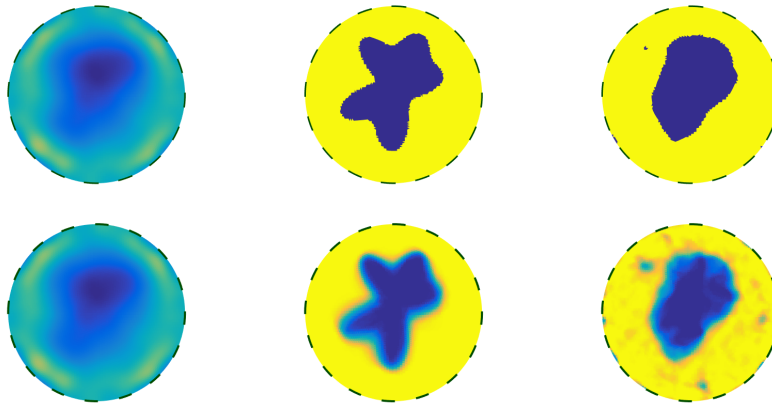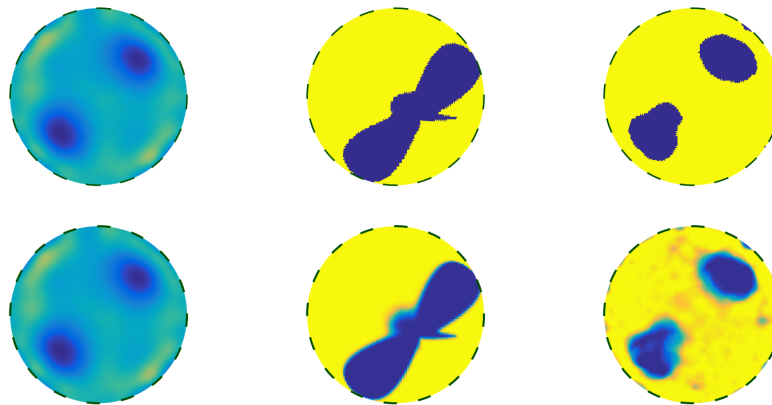
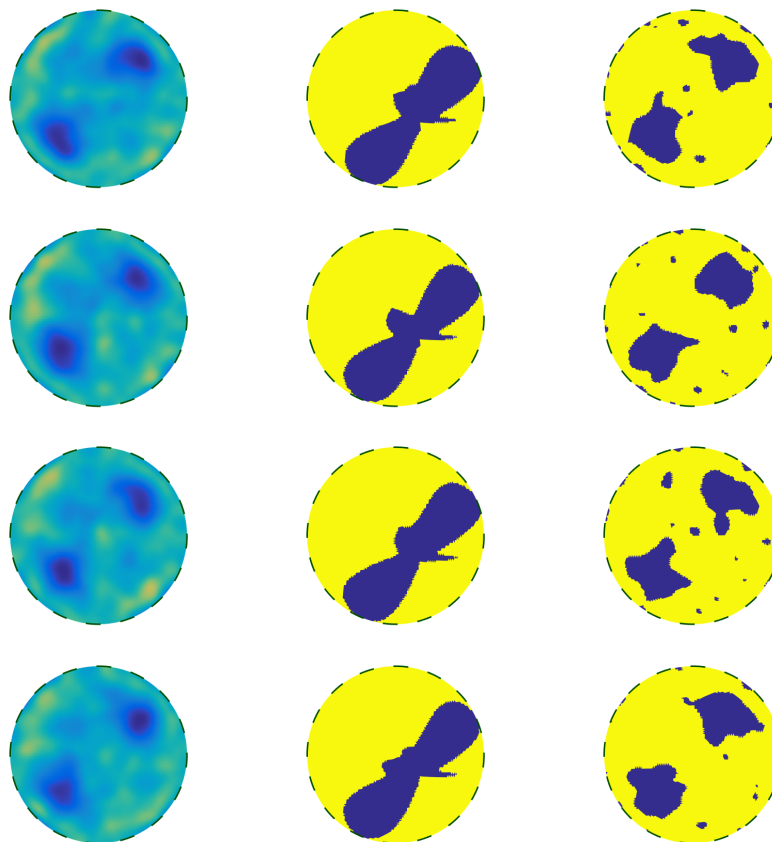(A) (Top) Estimates of $F_i(\mathbb{E}(u))$. (Bottom) Estimates of $\mathbb{E}(F_i(u))$.
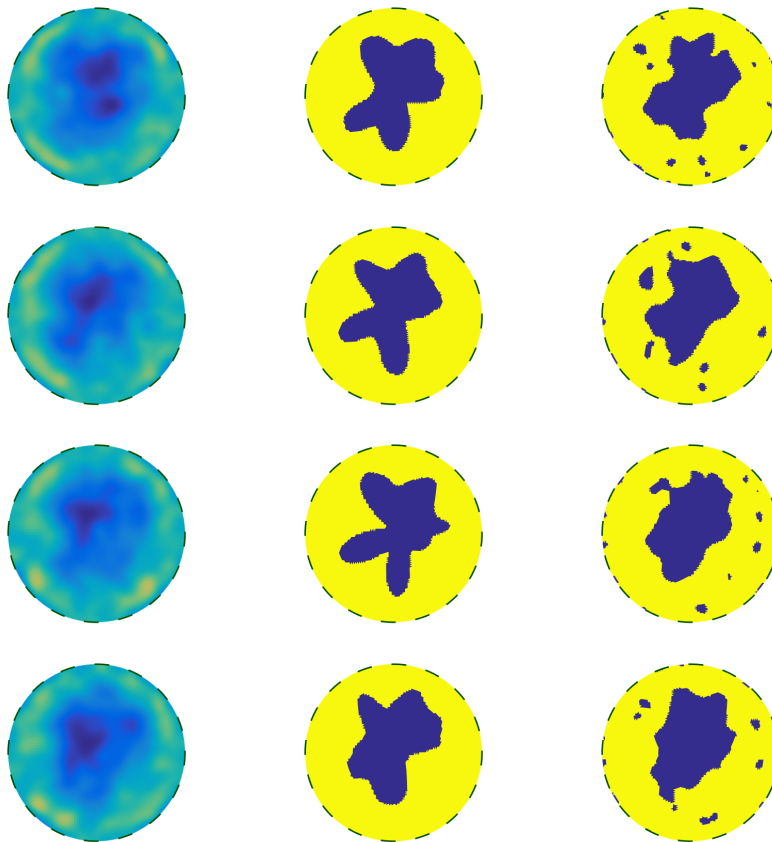
(B) Posterior samples.

FIGURE 4. Recovery of Conductivity A. From left to right, the log-Gaussian, star-shaped and level set priors are used.

(A) (Top) Estimates of $F_i(\mathbb{E}(u))$. (Bottom) Estimates of $\mathbb{E}(F_i(u))$.

(B) Posterior samples.

FIGURE 5. Recovery of Conductivity B. From left to right, the log-Gaussian, star-shaped and level set priors are used.

	Conductivity A	Conductivity B
Log-Gaussian prior	115960	122180
Star-shaped prior	228.84	45013
Level set prior	284.28	193.03

(A) The values of Φ at $F_i(\mathbb{E}(u))$.

	Conductivity A	Conductivity B
Log-Gaussian prior	111960	117950
Star-shaped prior	284.44	41408
Level set prior	137.93	124.42

(B) The values of Φ at $\mathbb{E}(F_i(u)))$.

FIGURE 6. The values of the misfit Φ at the different mean conductivities, for the different conductivities and prior distributions.

Quantity	Estimated ESS
$\hat{u}(0,1)$	40.0
$\hat{u}(0,2)$	90.7
$\hat{u}(1,1)$	35.4
$\hat{u}(1,2)$	44.5
$\hat{u}(1,3)$	36.0
$\hat{u}(2,1)$	101.9
$\hat{u}(2,2)$	37.9
$\hat{u}(2,3)$	89.7

(A) Log-Gaussian prior

Quantity	Estimated ESS
$x_0^{(1)}$	241.7
$x_0^{(2)}$	89.6
$\hat{r}(1)$	101.1
$\hat{r}(2)$	179.4
$\hat{r}(3)$	277.8
$\hat{r}(4)$	214.8
$\hat{r}(5)$	146.7
$\hat{r}(6)$	146.7

(B) Star-shaped prior

Quantity	Estimated ESS
$\hat{u}(0,1)$	26.4
$\hat{u}(0,2)$	28.9
$\hat{u}(1,1)$	27.2
$\hat{u}(1,2)$	23.5
$\hat{u}(1,3)$	23.5
$\hat{u}(2,1)$	26.0
$\hat{u}(2,2)$	26.6
$\hat{u}(2,3)$	24.1

(C) Level set prior

FIGURE 7. (Conductivity B) The estimated effective sample size for each chain, approximated using a variety of quantities, for the different choices of prior. In all cases 2.5×10^6 total MCMC samples are produced, with the initial 5×10^5 discarded as burn-in.
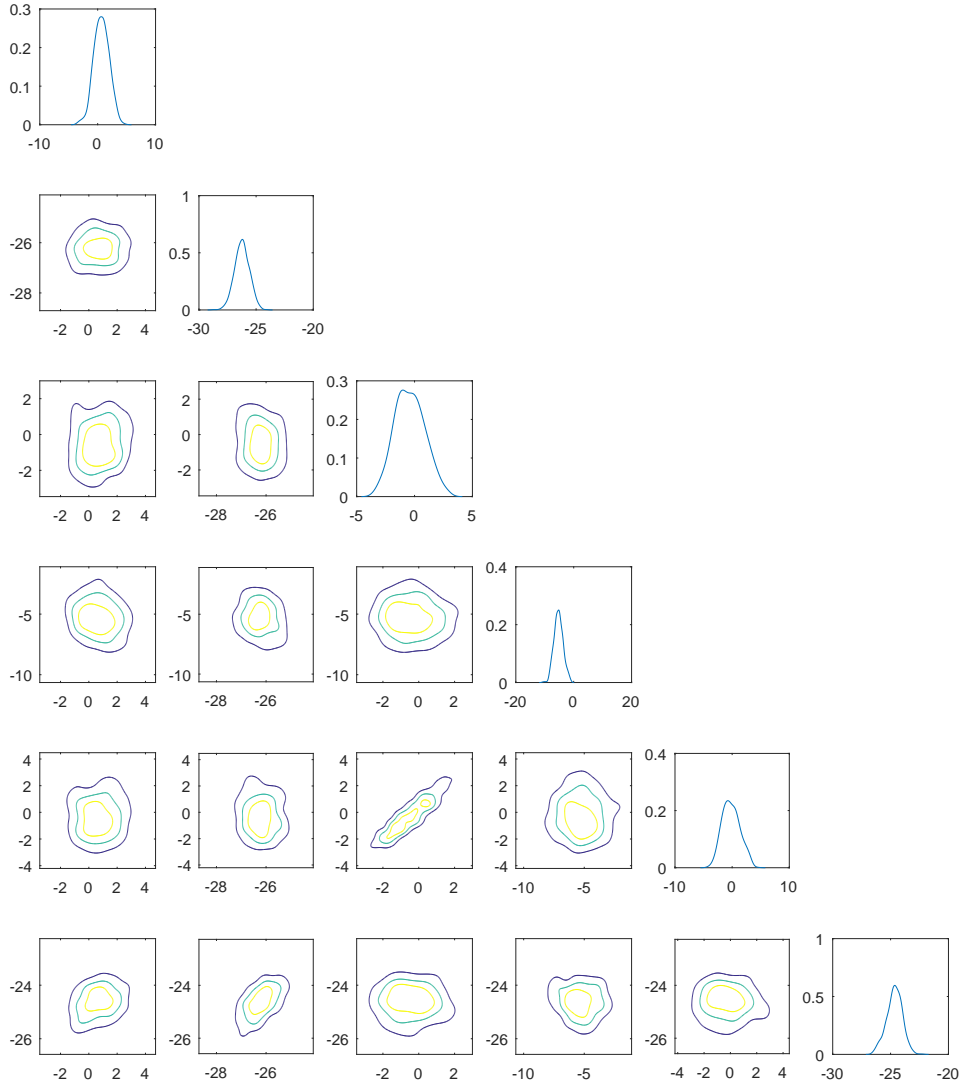
FIGURE 8. (Conductivity B, log-Gaussian prior) Kernel density estimates associated with six Fourier coefficients of u. The diagonal displays the marginal densities of each coefficient, and the off-diagonals the marginal densities of corresponding pairs of coefficients.

- The infinite dimensional formulation of the problem leads to the use of state of the art function space MCMC methods for sampling the posterior distribution.
- Simulations performed using these methods illustrate that the conditional mean provides a reasonable reconstruction of the conductivity, even with fairly significant noise on the measurements. They also illustrate the fact that the
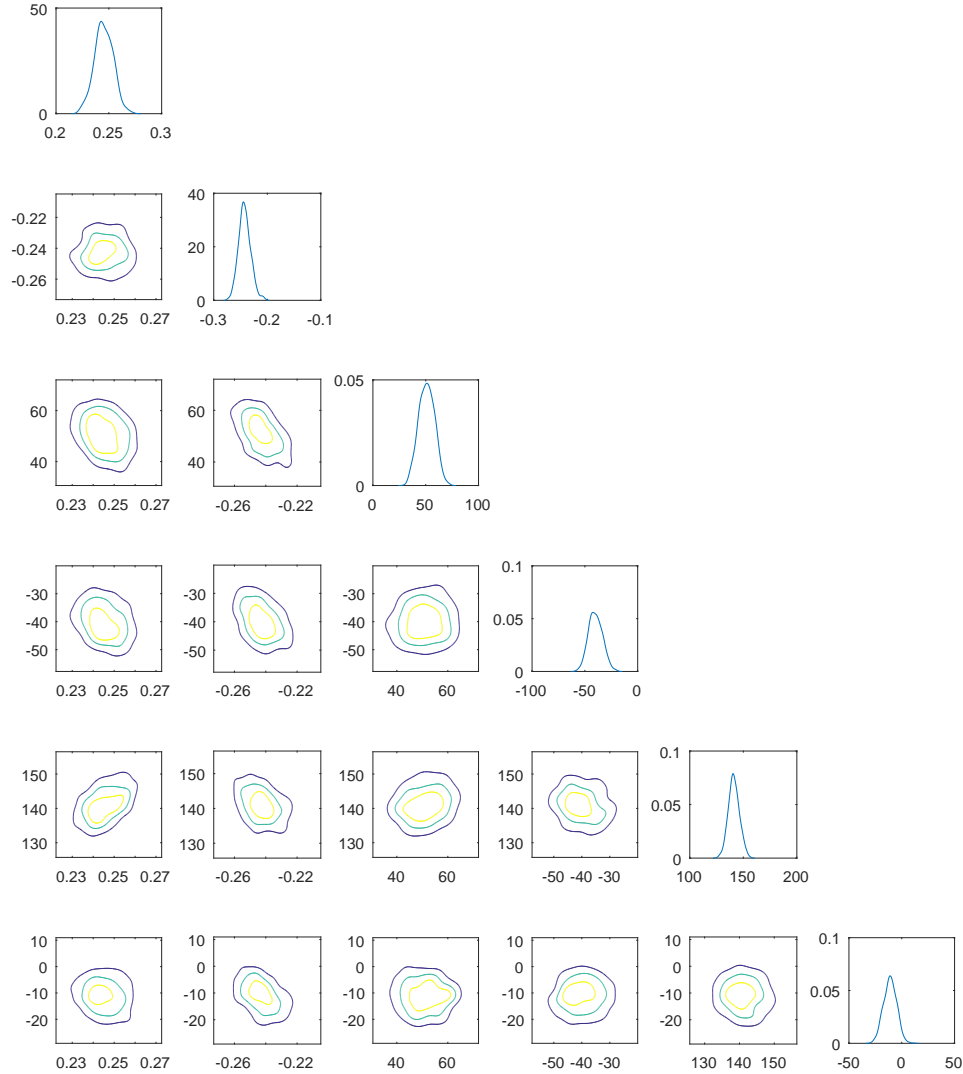
FIGURE 9. (Conductivity B, star-shaped prior) Kernel density estimates associated with the centre $(x_0^{(1)}, x_0^{(2)})$ and four Fourier coefficients of r. The diagonal displays the marginal densities of each quantity, and the off-diagonals the marginal densities of corresponding pairs of quantities.

choice of prior has a significant impact on reconstruction and, in particular, that the geometric priors (star-shaped and level set) can be particularly effective for the (approximately) piecewise constant fields that arise in many applications.

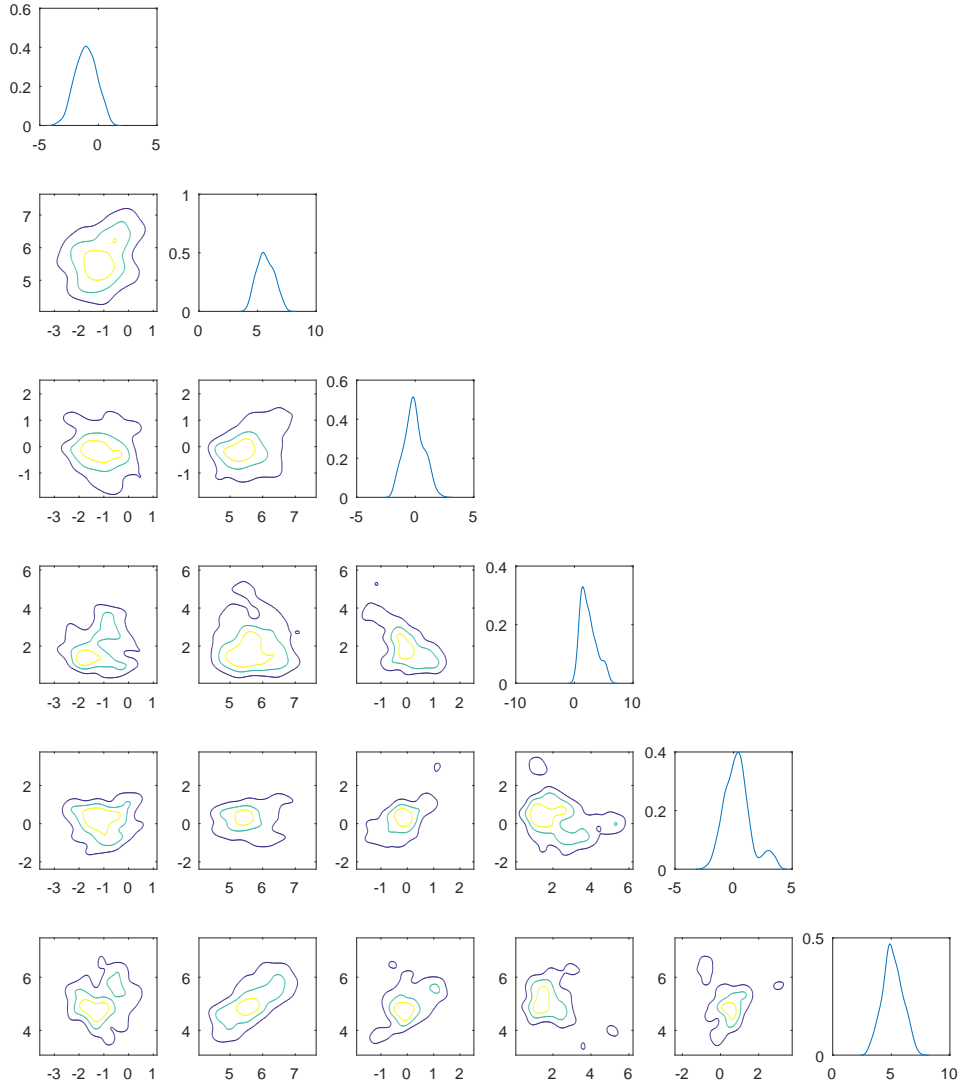Future research directions could include the following:

Figure 10. (Conductivity B, level set prior) Kernel density estimates associated with six Fourier coefficients of u. The diagonal displays the marginal density of each coefficient, and the off-diagonals the marginal densities of corresponding pairs of coefficients. Axes are rescaled by 10^6 for clarity.

- Sampling the exact posterior distribution using MCMC can be computationally expensive. Methods that approximate the posterior may be as effective for calculating quantities such as the conditional mean, with much lower computational load. The relative effectiveness versus cost of different methods could be studied. This could be especially important for simulations in three dimensions, where forward model evaluations are even more expensive.

- When using the level set prior, the length scale of samples could be treated hierarchically as an additional unknown in the problem.
- The star-shaped prior could be extended to describe multiple inclusions, either with the number of inclusions fixed or as an additional unknown.
- The Calderón problem could be considered in place of the CEM. The data could then be either the full Neumann-to-Dirichlet (or Dirichlet-to-Neumann) map, or a finite number of pairs of Dirichlet and Neumann boundary values. In order to perturb the measurements with noise, the former case would require the definition of Gaussian distributions on spaces of operators, and the latter Gaussian distributions on manifolds.

REFERENCES

[1] A. Adler and W. R. B. Lionheart, Uses and abuses of EIDORS: An extensible software base for EIT, *Physiological Measurement*, **27** (2006), S25–S42.

[2] G. Alessandrini, Stable determination of an inclusion by boundary measurements, *Applicable Analysis: An International Journal*, **27** (1988), 153–172.

[3] M. Bédard, Optimal acceptance rates for Metropolis algorithms: Moving beyond 0.234, *Stochastic Processes and their Applications*, **118** (2008), 2198–2222.

[4] A. Beskos, A. Jasra, E. A. Muzaffer and A. M. Stuart, Sequential Monte Carlo methods for Bayesian elliptic inverse problems, *Statistics and Computing*, **25** (2015), 727–737.

[5] A. Beskos, G. O. Roberts, A. M. Stuart and J. Voss, MCMC methods for diffusion bridges, *Stochastics and Dynamics*, **8** (2008), 319–350.

[6] V. I. Bogachev, *Measure Theory Volume I*, Springer, 2007.

[7] L. Borcea, Electrical impedance tomography, *Inverse Problems*, **18** (2002), R99–R136.

[8] T. Bui-Thanh and O. Ghattas, An analysis of infinite dimensional bayesian inverse shape acoustic scattering and its numerical approximation, *SIAM/ASA Journal on Uncertainty Quantification*, **2** (2014), 203–222.

[9] S. L. Cotter, G. O. Roberts, A. M. Stuart and D. White, MCMC methods for functions modifying old algorithms to make them faster, *Statistical Science*, **28** (2013), 424–446.

[10] M. Dashti and A. M. Stuart, The Bayesian Approach to Inverse Problems, In *Handbook of Uncertainty Quantification*, Springer, 2016, 1–118.

[11] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, Hybrid monte carlo, *Physics Letters B*, **195** (1987), 216–222.

[12] J. N. Franklin, Well posed stochastic extensions of ill posed linear problems, *Journal of Mathematical Analysis and Applications*, **31** (1970), 682–716, URL http://adsabs.harvard.edu/abs/1970JIMIA..31..682F.

[13] M. Gehre, B. Jin and X. Lu, An analysis of finite element approximation in electrical impedance tomography, *Inverse Problems*, **30** (2014), 045013, 24pp, URL http://stacks.iop.org/0266-5611/30/i=4/a=045013?key=crossref.284cda5f9645598cb83fccc9d226940a.

[14] M. Hairer, A. M. Stuart and S. J. Vollmer, Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions, *The Annals of Applied Probability*, **24** (2014), 2455–2490.

[15] R. P. Henderson and J. G. Webster, An impedance camera for spatially specific measurements of the thorax, *IEEE Transactions on Bio-Medical Engineering*, **25** (1978), 250–254.

[16] M. A. Iglesias, K. J. H. Law and A. M. Stuart, Ensemble Kalman methods for inverse problems, *Inverse Problems*, **29** (2013), 045001, 20pp, URL http://stacks.iop.org/0266-5611/29/i=4/a=045001?key=crossref.a70b58896eed5588086da7ad3d04954b.

[17] M. A. Iglesias, Y. Lu and A. M. Stuart, A bayesian level set method for geometric inverse problems, *Interfaces Free Bound.*, **18** (2016), 181–217, URL http://arxiv.org/abs/1504.00313.

[18] J. P. Kaipio, V. Kolehmainen, E. Somersalo and M. Vauhkonen, Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography, *Inverse Problems*, **16** (2000), 1487–1522.

[19] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen and E. Somersalo, Inverse problems with structural prior information, *Inverse Problems*, **15** (1999), 713–729.

[20] J. P. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005.

[21] R. E. Kass, Markov Chain Monte Carlo in practice: A roundtable discussion, *The American Statistician*, **52** (1998), 93–100.

[22] C. E. Kenig, J. Sjöstrand and G. Uhlmann, The Calderón problem with partial data on manifolds and applications, *Analysis and PDE*, **6** (2013), 2003–2048.

[23] K. Knudsen, M. Lassas, J. L. Mueller and S. Siltanen, D-bar method for electrical impedance tomography with discontinuous conductivities, *SIAM Journal on Applied Mathematics*, **67** (2007), 893–913.

[24] K. Knudsen, M. Lassas, J. L. Mueller and S. Siltanen, Reconstructions of piecewise constant conductivities by the D-bar method for electrical impedance tomography, *Journal of Physics: Conference Series*, **124** (2008), 012029.

[25] K. Knudsen, M. Lassas, J. L. Mueller and S. Siltanen, Regularized D-bar method for the inverse conductivity problem, *Inverse Problems and Imaging*, **3** (2009), 599–624.

[26] V. Kolehmainen, M. Lassas, P. Ola and S. Siltanen, Recovering boundary shape and conductivity in electrical impedance tomography, *Inverse Problems and Imaging*, **7** (2013), 217–242.

[27] S. Lan, T. Bui-Thanh, M. Christie and M. Girolami, Emulation of Higher-Order Tensors in Manifold Monte Carlo Methods for Bayesian Inverse Problems, *Journal of Computational Physics*, **308** (2016), 81–101.

[28] R. E. Langer, An inverse problem in differential equations, *Bulletin of the American Mathematical Society*, **39** (1933), 814–820.

[29] S. Lasanen, Non-gaussian statistical inverse problems. part I: Posterior distributions, *Inverse Problems & Imaging*, **6** (2012), 215–266.

[30] S. Lasanen, Non-gaussian statistical inverse problems. part II: Posterior convergence for approximated unknowns, *Inverse Problems & Imaging*, **6** (2012), 267–287.

[31] S. Lasanen, J. M. J. Huttunen and L. Roininen, Whittle-Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography, *Inverse Problems and Imaging*, **8** (2014), 561–586, URL http://www.aimsciences.org/journals/displayArticlesnew.jsp?paperID=9912.

[32] M. Lassas, E. Saksman and S. Siltanen, Discretization-invariant Bayesian inversion and Besov space priors, *Inverse Problems and Imaging*, **3** (2009), 87–122.

[33] M. S. Lehtinen, L. Paivarinta and E. Somersalo, Linear inverse problems for generalised random variables, *Inverse Problems*, **5** (1989), 599–612.

[34] A. Mandelbaum, Linear estimators and measurable linear transformations on a Hilbert space, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **65** (1984), 385–397.

[35] A. I. Nachman, Reconstructions from boundary measurements, *Annals of Mathematics*, **128** (1988), 531–576, URL http://cat.inist.fr/?aModele=afficheN&cpsidt=7131706.

[36] A. I. Nachman, Global uniqueness for a two-dimensional inverse boundary value problem, *Annals of Mathematics*, **143** (1996), 71–96.

[37] G. C. Papanicolaou and L. Borcea, Network approximation for transport properties of high contrast materials, *SIAM Journal on Applied Mathematics*, **58** (1998), 501–539.

[38] M. Salo, Calderón problem, Lecture Notes.

[39] D. Schymura, An upper bound on the volume of the symmetric difference of a body and a congruent copy, *Advances in Geometry*, **14** (2014), 287–298.

[40] E. Somersalo, M. Cheney and D. Isaacson, Existence and uniqueness for electrode models for electric current computed tomography, *SIAM Journal on Applied Mathematics*, **52** (1992), 1023–1040.

[41] A. M. Stuart, Inverse problems: A Bayesian perspective, *Acta Numerica*, **19** (2010), 451–559, URL http://webcat.warwick.ac.uk:80/record=b2341209.

[42] J. Sylvester and G. Uhlmann, A global uniqueness theorem for an inverse boundary value, *Annals of Mathematics*, **125** (1987), 153–169.

[43] L. Tierney, Markov chains for exploring posterior distributions, *Annals of Statistics*, **22** (1994), 1701–1728.

E-mail address: mdunlop@caltech.edu
E-mail address: astuart@caltech.edu