Applied Mathematics and Computation 358 (2019) 436–448

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

# Comparison of statistical inversion with iteratively regularized Gauss Newton method for image reconstruction in electrical impedance tomography

Sanwar Ahmad [a], Thilo Strauss [b], Shyla Kupis [c], Taufiquar Khan [a,*]

[a] Clemson University, School of Mathematical and Statistical Sciences, Clemson, SC 29634, United States
[b] University of Washington, Seattle, WA 98195, United States
[c] Department of Environmental and Earth Sciences, Clemson University, Clemson, SC 29634, United States

ARTICLE INFO

*Keywords:*
Electrical impedance tomography
Statistical inversion
Markov Chain Monte Carlo method
Bayesian inversion
Metropolis–Hastings algorithm

ABSTRACT

In this paper, we investigate image reconstruction from the Electrical Impedance Tomography (EIT) problem using a statistical inversion method based on Bayes' theorem and an Iteratively Regularized Gauss Newton (IRGN) method. We compare the traditional IRGN method with a new Pilot Adaptive Metropolis algorithm that (i) enforces smoothing constraints and (ii) incorporates a sparse prior. The statistical algorithm reduces the reconstruction error in terms of $\ell_2$ and $\ell_1$ norm in comparison to the IRGN method for the synthetic EIT reconstructions presented here. However, there is a trade-off between the reduced computational cost of the deterministic method and the higher resolution of the statistical algorithm. We bridge the gap between these two approaches by using the IRGN method to provide a more informed initial guess to the statistical algorithm. Our coupling procedure improves convergence speed and image resolvability of the proposed statistical algorithm.



## 1. Introduction

The Electrical Impedance Tomography (EIT) problem involves collecting electrical measurements on the smooth boundary $\partial\Omega$ to determine the spatially varying electrical conductivity distribution $\gamma$ within the bounded region $\Omega \subseteq \mathbb{R}^d (d = 2, 3)$. We assume $\gamma$ has a strictly positive, isotropic and bounded conductivity distribution with no current sources inside $\Omega$. We represent the EIT forward problem with the following elliptic differential equation,

$$-\operatorname{div}(\gamma \nabla u) = 0 \quad \text{in } \Omega, \tag{1}$$

where $u \in H^1(\Omega)$ is the electric potential and $\gamma$ is known.

An EIT experiment involves applying an electrical current (Neumann data) $I$ on $\partial\Omega$ and measuring the resulting electrical potential (Dirichlet data) $U$ on $\partial\Omega$. The data collected then provides information on the Neumann-to-Dirichlet (NtD) map. Partial knowledge of the NtD map is used to approximate $\gamma$ from a set of EIT experiments using different input currents [1–3]. EIT can be applied to nondestructive testing [4], the monitoring of oil and gas mixtures in oil pipelines [5], noninvasive medical imaging [6,7], etc. For example, EIT has been used in civil engineering to monitor water infiltration

* Corresponding author.
*E-mail addresses:* sanwara@clemson.edu (S. Ahmad), thilum@gmx.de (T. Strauss), skupis@clemson.edu (S. Kupis), khan@clemson.edu (T. Khan).

https://doi.org/10.1016/j.amc.2019.03.063
0096-3003/

into soil [8–11]. Stacey [12] also studied the EIT's feasibility by monitoring moisture movement in Berea sandstone. Hou and Lynch [13] applied EIT to loaded cementitious composites that were fiber reinforced.

EIT is well known to suffer from a high-degree of nonlinearity and severe ill-posedness [14,15]. Therefore, regularization is required to produce reasonable electrical impedance images. Most reconstruction methods are deterministic, such as the factorization method [17], d-bar method [18], and variational type methods for least-squares fitting. The variational type methods minimize a certain discrepancy functional using Tikhonov-based regularization or an iterative type method of a linearized model or fully nonlinear model such as sparisty cosnstraints [14,15], iteratively regularized Gauss Newton method [16] just to mention a few examples. These analytical methods can be effective in determining specific conductivity distributions, but statistical inversion methods [19] can offer an alternative approach. In [20], Kaipio et al. optimizes the current patterns based on criteria regarding functionals of the posterior covariance matrix. The Bayesian approach has also been used to study the errors from model reduction and partially unknown geometry [21,22]. In [23], Bardsley quantified uncertainty of MCMC-based image reconstruction. Statistically, the sparsity regularization enforces the $\ell_p$ prior on the expansion coefficients for a certain basis like the deterministic approaches for EIT [14,15,24]. In this paper, we apply the Pilot Adaptive Metropolis algorithm proposed in [25–27] to investigate the algorithm's efficacy with synthetic data. We use a TV prior and a combination of a TV prior and a $\ell_1$-type prior to solve a Bayesian inverse problem for which the dimension of the problem is neither too high neither too low (approximately 500 unknowns). In [40], Lassas and Siltanen describe how the solution to a Bayesian inverse problem diverges when the dimension of the problem converges to infinity; however, for the dimension (approximately 500 unknowns) considered in our problem, we found experimentally that our method converges to a solution. Similarly, in [41], Lucka found that, for $\ell_1$-type priors, the efficiency of MCMC methods dramatically decreased when the sparsity or dimension of the problem increased, which we confirmed in numerical experiments were only $\ell_1$ is used as standalone priors. In this study, we find that the our MCMC method converged efficiently when we use a combination of $\ell_1$-type and *TV* priors.

The paper is organized as follows. In Section 2, we describe the complete electrode model (CEM) for the EIT problem. In Section 3, we formulate the inverse problem. We then describe the IRGN method in Section 4. In Section 5, we formulate the statistical inverse problem, and we discuss choices for regularizing the prior density. In Section 6, we discuss the Markov Chain Monte Carlo (MCMC) method and Pilot Adaptive Metropolis algorithm. We compare the inversion results from the statistical approach with the traditional IRGN method in Section 7 to demonstrate the efficacy of our coupled model. Lastly, we conclude with Section 8.

## 2. Complete electrode model

Suppose $L$ electrodes have been fixed around the surface of an object. Current is applied to a subset of these electrodes, and the resulting voltage is measured at all other electrodes. This EIT experiment is repeated several times with different electrode configurations to efficiently characterize the imaged object. The inverse problem then is to reconstruct the conductivity distribution inside the object from a finite set of surface point measurements.

The Eq. (1) is used to solve the electric potential at the electrodes and inside $\Omega$ for some applied current $I$. We obtained (1) by a scaling analysis of Maxwell's equation for electromagnetic fields inside of an object [1,28]. We denote the class of admissible conductivities as

$$Q = \{\gamma \in L^\infty(\Omega) | \gamma(x) \geq \gamma_0 > 0, \text{ for } x \in \Omega \text{ and } \gamma_0 \in \mathbb{R}\}.$$

Currents are applied to electrodes on the surface $\partial\Omega$ of the body. Let $j$ denote the inward unit normal component for the current density that is produced by current on the surface. Then, the Neumann boundary condition is

$$\gamma \frac{\partial u}{\partial n} = j \quad \text{on } \partial\Omega. \tag{2}$$

(1) and (2) are known as the *continuum model*. In practice, the current density $j$ at the electrodes is unknown, but $\int_{E_l} \gamma \frac{\partial u}{\partial n} dS = I_l$ is known, where $n$ is the unit outward normal to $\Omega$, $E_l$ is the surface area of the $l$th electrode and $I_l$ is the current injected into $E_l$. Thus the Neumann condition (2) can be rewritten as

$$\int_{E_l} \gamma \frac{\partial u}{\partial n} dS = I_l \quad \text{for } l = 1, 2, \ldots, L, \tag{3}$$

Furthermore, we know $j = 0$ for the current density on the boundary between the electrodes, such that

$$\gamma \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\Omega / \cup_{l=1}^{L} \overline{E_l}. \tag{4}$$

The potential on the electrodes is considered to be constant: $U_l = \text{constant}$. This property is known as the *shunting effect*, which is represented by

$$u = U_l, \text{ on } E_l, \text{ for } l = 1, 2, \ldots, L. \tag{5}$$

There is also an electro-chemical effect due to the formation of a thin and highly resistive layer between the electrodes and the body. Electrical impedance from this layer, $z_l$, is called the effective contact impedance or surface impedance at $E_l$. This

effect changes (5) to

$$u + z_l \gamma \frac{\partial u}{\partial n} = U_l \quad \text{on } E_l \text{ for } l = 1, 2, \ldots, L. \tag{6}$$

The complete electrode model (CEM) consists of (1), (3), (4) and (6), together with the following conditions.

$$\sum_{l=1}^{L} I_l = 0 \quad \text{(conservation of charge)} \tag{7}$$

$$\sum_{l=1}^{L} U_l = 0 \quad \text{(choice of a ground)} \tag{8}$$

Although the CEM has a unique solution, the accuracy of it is determined by the predicted experimental measurements [29].

For the FEM solution to the forward CEM model, we reformulate the variational form of CEM (see [28] for details) as

$$b((u, U), (v, V)) = f(v, V), \text{ for all } (v, V) \in H^1(\Omega) \oplus \tilde{\mathbb{R}}^L, \tag{9}$$

where $f(v, V) = \sum_{l=1}^{L} I_l V_l$ is a functional that maps $H \mapsto \mathbb{R}$.

Existence and uniqueness of a solution $(u, U) \in H$ can be shown using the Lax-Milgram lemma. In [29], it is shown that $(u, U)$ satisfies the CEM conditions if and only if it satisfies (9).

### 2.1. FEM discretization of CEM

Following the typical FEM approach, let $T = \{T_1, \ldots, T_{|T|}\}$ be the triangulation of $\Omega$, which has $N$ mesh points for the finite dimensional subspace $H_N$ of $H_1(\Omega)$. Any $u_N \in H_N$ is represented by

$$u(x) \approx u_N(x) = \sum_{i=1}^{N} \alpha_i \phi_i(x), \text{ for } \alpha_i \in \mathbb{R},$$

where $\phi_i(x)$ are the basis functions of $H_h$ satisfying $\phi_i(x_k) = \delta_{ik}$ for $i, k = 1, \ldots, N$. Electric potential on the electrodes is given by

$$U^N = \sum_{k=1}^{L-1} \beta_k \nu_k = G\beta,$$

where $\nu_k, k = 1, \ldots, L-1$ compose the basis for $\tilde{\mathbb{R}}^L$ and $G \in \mathbb{R}^{L \times (L-1)}$. We must determine the coefficients $\alpha_i$ and $\beta_k$ in this formulation. Choosing $v = \phi_i$ and $V = \nu_k$ when the set of test functions is of the form $(\phi_1, 0), \ldots, (\phi_N, 0), (\nu_1, 0), \ldots, (\nu_{L-1}, 0)$ results in the following system of equations in matrix form:

$$\hat{A}\theta = f, \tag{10}$$

where $\theta = (\alpha, \beta)^T \in \mathbb{R}^{N+L-1}$, the matrix $\hat{A}$ is the standard FEM system matrix (see [28] for details) and the right hand side $f$ is given by,

$$f(i) = \begin{cases} 0 & \text{for } i = 1, \ldots, N \\ I^T \nu_{i-N} & \text{for } i = N+1, \ldots, N+L-1 \end{cases}. \tag{11}$$

For a fixed current vector $I$ and positive contact impedances $(z_k)_{k=1}^{L}$, we define

$$\mathcal{F} : Q \subset L^\infty(\Omega) \to H, \quad \gamma \mapsto (u, U)$$

to be the forward operator that maps the conductivity $\gamma$ to the solution of the forward problem obtained by FEM described above. [30] showed that this operator is Frèchet differentiable. Recall that Frèchet differentiability of $\mathcal{F}$ in $\gamma$ means that

$$\lim_{||\eta||_\infty \to 0} \frac{||\mathcal{F}(\gamma + \eta) - \mathcal{F}(\gamma) - \mathcal{F}'(\gamma)\eta||_H}{||\eta||_\infty} = 0.$$

The forward problem uses the known conductivity distribution $\gamma$ to find the boundary data associated with a given source. Our goal is to estimate the conductivity distribution $\gamma$ from all pairs of current vectors $I \in \tilde{\mathbb{R}}^L$ and resulting voltage vectors $U \in \tilde{\mathbb{R}}^L$. We minimize the cost functional to recover $\gamma$ since the inverse problem is severely ill-posed from the finite set of measurements.

$$J(\gamma) = ||\mathcal{F}(\gamma) - u_\delta||_2^2$$

where $u_\delta$ approximates the exact data $u$ with the accuracy $\delta$, i.e.,

$$||u - u_\delta|| < \delta. \tag{12}$$

However, regularization is needed to improve the ill-posed problem and instead, we minimize,

$$J_\lambda(\gamma) = ||\mathcal{F}(\gamma) - u_\delta||_2^2 + \lambda R(\gamma - \gamma^*)$$

where $\lambda$ is the regularization parameter, $R(\cdot)$ is the regularization term and $\gamma^*$ is the known background. There are several choices for $R(\cdot)$. The $\ell_p$ regularization $R_{\ell_p}(\gamma - \gamma^*)$ is defined as

$$\mathbf{R}_{\ell_\mathbf{p}}(\gamma - \gamma^*) = ||\gamma - \gamma^*||_\mathbf{p}.$$

where $0 < p \leq 2$ is a constant, [19]. The $\ell_p$ regularization enforces sparsity for $0 < p \leq 1$ and smoothness when $p \geq 2$. Total variation regularization is used for most practical applications to obtain smooth images. Total variation is defined as [42,43],

$$R_{TV}(\gamma) = \sup\left\{\int_\Omega \gamma \ \text{div}\ z dx | z \in C_0^\infty(\Omega), ||z||_\infty \leq 1\right\} = |D\gamma|(\Omega),$$

where $|\cdot|(\Omega)$ is a finite Radon measure. In particular, for $\gamma \in W^{1,1}(\Omega)$, it reduces to the standard notation for TV regularization,

$$R_{TV}(\gamma) = \int_\Omega |\nabla\gamma| dx.$$

The regularization function is represented by a norm for most analytical methods. In this paper we used one of the most successful methods for solving the ill-conditioned problem, $R_{\ell_2}$ known as Tikhonov regularization. The cost functional from Tikhonov regularization is

$$J_\lambda(\gamma) = \frac{1}{2}||\mathcal{F}(\gamma) - u_\delta||_2^2 + \frac{\lambda}{2}||W(\gamma - \gamma^*)||_2^2, \tag{13}$$

where $W$ is weight function and upon discretization becomes a weight matrix. There are several iterative approaches to minimize (13). In this paper, we used a modified iteratively regularized Gauss-Newton (IRGN) method for the minimization, which is described in the next section.

## 3. Iteratively regularized Gauss–Newton method

Suppose $\lambda_k$ is some sequence of regularizing parameters satisfying the conditions

$$\lambda_k \geq \lambda_{k+1} > 0, \quad \sup_{k \in \mathbb{N}\cup\{0\}} \frac{\lambda_k}{\lambda_{k+1}} = \hat{d} < \infty, \quad \lim_{k\to\infty} \lambda_k = 0. \tag{14}$$

Let the unique global minimum of (13) be denoted by $\tilde{\gamma}$. Assume $\tilde{\gamma}$ satisfies the invertibility conditions and the conditions for $\mathcal{F}$ as described in [31]. Then, the unique global minimum of (13) is explicitly given by

$$\tilde{\gamma} = \gamma_k - (\mathcal{F}'(\gamma_k)^T\mathcal{F}'(\gamma_k) + \lambda_k W_2)^{-1}\{\mathcal{F}'(\gamma_k)^T(\mathcal{F}(\gamma_k) - u_\delta) + \lambda_k W_2(\gamma_k - \gamma^*)\}, \tag{15}$$

where $\gamma_k$ is the $k$th approximation to $\gamma$ and $\mathcal{F}'(\gamma_k)$ is the Jacobian matrix at the $k$th iteration, and $W_2 = W^TW$. The above algorithm is generalized further using a line search procedure from Smirnova et al. [31]. A variable step size, $s_k$, is introduced, such that

$$0 < s_k \leq 1. \tag{16}$$

The modified IRGN algorithm is then

$$\gamma_{k+1} = \gamma_k - s_k(\mathcal{F}'(\gamma_k)^T\mathcal{F}'(\gamma_k) + \lambda_k W_2)^{-1}\{\mathcal{F}'(\gamma_k)^T(\mathcal{F}(\gamma_k) - u_\delta) + \lambda_k W_2(\gamma_k - \gamma^*)\}. \tag{17}$$

The line search parameter $s_k$ is chosen to minimize the scalar objective function

$$\Phi(s) = J(\gamma_k + sp_k) \tag{18}$$

where $p_k$ is the search direction, which solves

$$(\mathcal{F}'(\gamma_k)^T\mathcal{F}'(\gamma_k) + \lambda_k W_2)p_k = -\mathcal{F}'(\gamma_k)^T(\mathcal{F}(\gamma_k) - u_\delta) + \lambda_k W_2(\gamma_k - \gamma^*) \tag{19}$$

This step is accomplished through a backtracking strategy until either one of the strong Wolfe conditions,

$$J(\gamma_k + sp_k) \leq J(\gamma_k) + c_1\ s\nabla J(\gamma_k)^T p_k \tag{20}$$

$$|\nabla J(\gamma_k + sp_k)^T p_k| \leq |c_2 \nabla J(\gamma_k)^T p_k|. \tag{21}$$

is satisfied [33], or the maximum number of backtracking steps has been reached. We use the theoretically derived values of $c_1 = 0.0001$ and $c_2 = 0.9$, derived in [33]. Due to the inexact nature of $u_\delta$, we adopt a stopping rule from Bakushinsky and Smirnova [32] to terminate the iterations at the first index $\mathcal{K} = \mathcal{K}(\delta)$, such that

$$||\mathcal{F}(\gamma_\mathcal{K}) - u_\delta||^2 \leq \rho\delta < ||\mathcal{F}(\gamma_k) - u_\delta||^2, \quad 0 \leq k \leq \mathcal{K}, \quad \rho > 1. \tag{22}$$

### 3.1. Derivative of the forward operator

Computing the Jacobian of the forward operator is required for most iterative Newton methods. If the regularity assumptions on the domain and the coefficients are satisfied, the forward operator is differentiable. The operator $\mathcal{F}$ is Frèchet differentiable. It maps $\gamma \in \text{int}(Q)$ to the solution $(u, U) \in H$ of the forward problem with current vector $I$. If $\eta \in L^{\infty}(\Omega) = \{v(x) : \sup_{x\in\Omega} |v(x)| < \infty\}$, such that $\gamma + \eta \in Q$, then the derivative $\mathcal{F}(\gamma)\eta = (w, W) \in H$ satisfies the following variational problem,

$$b((w, W), (v, V)) = -\int_{\Omega} \eta \nabla u \cdot \nabla v dx \tag{23}$$

for all $(v, V) \in H$, where $(u, U) = \mathcal{F}(\gamma)$.

Using (23), the Jacobian has been derived in [30].

## 4. The statistical inverse problem

Inverse problems are typically written in terms of a minimization problem. In this section, we write the inverse problem in terms of the posterior density of the conductivity $\gamma$ given the measurements $U$ on $\partial\Omega$ to obtain the expected value of the conductivity given the surface measurements. This estimate is a reasonable point estimate to the ill-posed inverse problem. In the following, we formulate the finite dimensional posterior density for the conductivity distribution given the finite dimensional measurement on $\partial\Omega$. Let $\delta_L$ represent the $L$ dimensional measurement noise corresponding to the measurements $U_L$. Denote $\gamma_M^*$, $U_L^*$ and $\delta_L^*$ as the random variables for $\gamma_M$, $U_L$ and $\delta_L$, respectively.

### 4.1. The posterior density

For the posterior density $\pi_{\gamma_M^*}(\gamma_M|U_L, \delta_L)$, we first assume the EIT forward operator $\Theta(\gamma_M, \delta_L)$ is well posed, such that $U_L = \Theta(\gamma_M, \delta_L)$, or

$$\pi_{\gamma_M^*}(U_L|\gamma_M, \delta_L) = \Delta(U_L - \Theta(\gamma_M, \delta_L)) = \begin{cases} \infty & \text{if } U_L - \Theta(\gamma_M, \delta_L) = 0, \\ 0 & \text{else,} \end{cases} \tag{24}$$

with the property that $\int_{\mathbb{R}^m} \Delta(x)dx = 1$. $\Delta(\cdot)$ is called the Dirac delta function. The integral over the Dirac delta function represents a distribution function. The joint density of $\gamma_M^*, U_L^*$ and $\delta_L^*$ is

$$\pi_{\gamma_M^*,U_L^*,\delta_L^*}(\gamma_M, U_L, \delta_L) = \pi_{U_L^*}(U_L|\gamma_M, \delta_L)\pi_{\gamma_M^*,\delta_L^*}(\gamma_M, \delta_L) \tag{25}$$

$$= \Delta(U_L - \Theta(\gamma_M, \delta_L))\pi_{\gamma_M^*,\delta_L^*}(\gamma_M, \delta_L), \tag{26}$$

from (24) and $\pi_X(x|Y = y) = \frac{\pi_{X,Y}(x,y)}{\pi_Y(y)}$. Here, the $x$ and $y$ are values of the random variables $X$ and $Y$, respectively. By integrating the measurement noise $\delta_L$ out of Eq. (25), we can then obtain the joint density of $\gamma_M^*$ and $U_L^*$

$$\pi_{\gamma_M^*,U_L^*}(\gamma_M, U_L) = \int_{\mathbb{R}^L} \pi_{\gamma_M^*,U*_L,\delta_L^*}(\gamma_M, U_L, \delta_L)d\delta_L \tag{27}$$

$$= \int_{\mathbb{R}^L} \Delta(U_L - \Theta(\gamma_M, \delta_L))\pi_{\gamma_M^*,\delta_L^*}(\gamma_M, \delta_L)d\delta_L. \tag{28}$$

We assume that the measurement noise $\delta_L^*$ is independent from $\gamma_M^*$, that is $\pi_{\gamma_M^*,\delta_L^*}(\gamma_M, \delta_L) = \pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(\delta_L)$, and that it is additive, that is $U_L = \Theta(\gamma_M) + \delta_L$. Eq. (27) can now be simplified to

$$\pi_{\gamma_M^*,U_L^*}(\gamma_M, U_L) = \int_{\mathbb{R}^L} \Delta(U_L - \Theta(\gamma_M) - \delta_L)\pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(\delta_L)d\delta_L \tag{29}$$

$$= \pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(U_L - \Theta(\gamma_M)). \tag{30}$$

The posterior density of $\gamma_M^*$ given the measurements $g_{\hat{K}}$ is

$$\pi_{\gamma_M^*}(\gamma_M|g_L) = \frac{\pi_{\gamma_M^*,U_L^*}(\gamma_M, U_L)}{\pi_{U_L^*}(U_L)} \tag{31}$$

$$= \frac{\pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(U_L - \Theta(\gamma_M))}{\int_{\mathbb{R}^L} \pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(U_L - \Theta(\gamma_M))d\gamma_M} \tag{32}$$

$$\propto \pi_{\gamma_M^*}(\gamma_M)\pi_{\delta_L^*}(U_L - \Theta(\gamma_M)) \tag{33}$$

from the conditional probability formula. Note that $\pi_{U_L^*}(U_L) = \int_{\mathbb{R}^L} \pi_{\gamma_M^*}(\gamma_M)\ \pi_{\delta_L^*}(U_L - \Theta(\gamma_M))d\gamma_M > 0$ is a constant. The posterior density can be defined by choosing a density for the measurement noise $\delta_L^*$ and a prior density function for $\gamma_M^*$. We assume data with random Gaussian noise,

$$\pi_{\delta_L^*}(U_L - \Theta(\gamma_M)) \propto \exp\left[-\frac{1}{2}(U_L - \Theta(\gamma_M))'C^{-1}(U_L - \Theta(\gamma_M))\right], \tag{34}$$

where $C$ is a positive definite covariance matrix. The density of $\gamma_M^*$ is the prior density because it contains prior knowledge about $\gamma_M^*$. We also assume

$$\pi_{\gamma_M^*}(\gamma_M) \propto \chi_A(\gamma_M)\exp[-\alpha R(\gamma_M)], \tag{35}$$

where $R(\cdot)$ is a regularization function, $\alpha > 0$ is a constant, and $\chi_A(\gamma_M)$ a indicator function with $A = [0, \infty)^n$.

### 4.2. Prior density as a regularization function

In this section, we discuss several choices for the regularization function $R(\cdot)$. We can apply most regularization functions in both the statistical and the analytical settings. However, the statistical setting typically allows to chose from more general regularizations then its analytical counterpart.

#### 4.2.1. The $\ell_p$ prior

The $\ell_p$ regularization $R_{\ell_p}(\gamma_M)$ is defined as

$$R_{\ell_p}(\gamma_M) := \sum_{i=1}^{n} c_i|\gamma_M(i) - \gamma_M^b(i)|^p, \tag{36}$$

where $c_i$ represent the weights, $\gamma_M^b$ is a typical background conductivity and $0 < p \leq 2$ is a constant [19]. $R_{\ell_p}(\gamma_M)$ is a norm if $p \geq 1$ and only defines a metric when $0 < p < 1$. The regularization function is represented by a norm for most analytical methods, but the statistical reconstruction can also handle cases when $0 < p < 1$. In theory, a good choice for the weights $c_i$ are large values at the boundary and exponentially decreasing values towards the center of $\Omega$ because the variance is smaller on the boundary than in the center [34]. The $\ell_p$ regularization enforces sparsity when $0 < p \leq 1$ and enforces smoothness when $p \geq 2$.

#### 4.2.2. The total variation prior

Total variation regularization is used to obtain smooth images, which is needed for most practical applications. The total variation regularization for an infinite dimensional space is defined as $R_{TV_c}(\gamma)$ as previously defined with $\gamma$ as the infinite dimensional version of $\gamma_M$. The discrete analogue for a two-dimensional body of the total variation regularization $R_{TV_c}$ [19] is

$$R_{TV}(\gamma_M) := \sum_{i=1}^{h} l_i|\triangle_i^*\gamma_M|, \tag{37}$$

where $l_i$ is edge length of the $i$th adjacent triangle and

$$\triangle_i^* = (0, 0, \ldots, 0, 1_{a_{(1,i)}}, 0, \ldots, 0, -1_{a_{(2,i)}}, 0, \ldots, 0), \tag{38}$$

where $a = \{a_{(j,i)}\}_{i=1,\ j\in\{1,2\}}^{h}$ is the set of the numbers for all adjacent triangle tuples $(a_{(1,i)}, a_{(2,i)})$.

### 4.3. Combination of the $\ell_p$ and total variation priors

In general, we can choose $\pi_{\gamma_M^*}(\gamma_M)$ to be any kind of prior density on $\Omega$ instead of using the classical analytical choices. We recommend choosing a continuous regularization function for $\gamma_M$ that is integrable over $\Omega$ for proper reconstructions.

The $\ell_p$ and total variation prior can provide better inversion results when combined together [26],

$$R_G(\gamma_M) = \alpha_1 R_{\ell_p}(\gamma_M) + \alpha_2 R_{TV}(\gamma_M), \tag{39}$$

where $\alpha_1, \alpha_2 > 0$ are constants. For proper choice of $\alpha_1, \alpha_2 > 0$, this regularization function can provide smooth images that give the correct background conductivity.

## 5. The Markov Chain Monte Carlo method

We have discussed the posterior density with several meaningful prior densities or regularization functions. Now, we can find the Bayesian estimate for $\gamma_M^*$ based on the measurements $U_L$ using $E(\gamma_M^*|U_L) = \int_{\mathbb{R}^n} \gamma_M \pi_{\gamma_M^*}(\gamma_M|U_L)d\gamma_M$.

There is no direct method for finding the Bayesian estimate $E(\gamma_M^*|U_L)$ because the posterior density $\pi_{\gamma_M^*}(\gamma_M|U_L)$ does not have a closed form. However, we can use the Markov Chain Monte Carlo Method (MCMC) to generate a large set of random

samples $\{\gamma_M^{(i)}\}_{i=B^*}^N$ from the posterior density $\pi_{\gamma_M^*}(\gamma_M|U_L)$, after a burn-in time $B^*$, to approximate the Bayesian estimate from its sample mean,

$$E(\gamma_M^*|g_L) = \int_{\mathbb{R}^d} \gamma_M \pi_{\gamma_M^*}(\gamma_M|U_L) d\gamma_M \approx \frac{1}{N-B^*} \sum_{i=B^*}^{N} \gamma_M^{(i)}. \tag{40}$$

The Gibbs sampler and Metropolis–Hastings algorithm are typically used for randomly sampling the posterior density. The Metropolis–Hastings algorithm is computationally less expensive than the Gibbs sampler for the EIT inverse problem. Hence, here we use a version of the Metropolis–Hastings algorithm to randomly sample from the posterior density. We recommend [35] for a detailed description of the Metropolis–Hastings algorithm.

Consider a Markov chain on the continuous state space $(E, \mathscr{B}, \mathscr{M})$ where $\mathscr{B}$ is a Borel $\sigma$-algebra on $E$ and $\mathscr{M}$ is the normalized Lebesgue measure on $E$. Let $E \subseteq \mathbb{R}^d$ be the support from the target distribution. This means $E$ is the set containing all values a state $x^{(i)}$, of the chain $\{x^{(i)}\}$, can take. Furthermore, let $P(x; A)$ denote a transition kernel for $x \in E$, where $A \in \mathscr{B}$. A transition kernel $P(x; A)$ represents the probability of jumping from a current state $x$ to another state in the set $A$. It is desirable to find a transition kernel $P(\cdot\,;\,\cdot)$ that converges to an invariant distribution $\pi^*$. Here $\pi^*$ is the distribution from the posterior density (31). The transition kernel,

$$P_{MH}(x; A) := \int_A q(x, y)\alpha(x, y)dy + \left[1 - \int_{\mathbb{R}^d} q(x, y)\alpha(x, y)dy\right]\chi_A(x), \tag{41}$$

converges to the invariant distribution $\pi^*$. Here $\chi_A$ is the indicator function over the set $A$ and $q(x, y)$ is a proposal density for generating a new candidate random sample $y$ from a current random sample $x$. For example, $q(x, \cdot)$ can be a multivariate normal density with mean $x$. The acceptance ratio

$$\alpha(x, y) = \begin{cases} \min\left[\frac{P(y|\delta)q(y,x)}{P(x|\delta)q(x,y)}, 1\right], & \text{if } P(x|\delta)q(x, y) > 0 \\ 1, & \text{otherwise,} \end{cases} \tag{42}$$

is the probability of accepting a new random sample $y$.

Selecting the proper proposal density for the Metropolis algorithm is required to randomly sample from the target distribution $\pi^*$ in a reasonable time. This choice is usually complicated because the target density is generally unknown [36–38]. One method of eliminating this problem is using an adaptive Metropolis algorithm by iteratively adapting the proposal density based on all previous samples from the chain. However, adaptive Metropolis algorithms [25,36,38] usually produce a non-Markovian chain that require establishment of the correct ergodic properties.

The approach used in this manuscript is to adapt the proposal distribution a finite amount of times and begin the burn-in time after the last adaption [25,26,39]. This method cannot guarantee that the optimal proposal distribution for the target distribution is obtained after the last adaption. However, the convergence speed respect to the classical Metropolis–Hastings algorithm is usually increased considerably while still maintaining all good properties after the burn-in time. The pilot adaptive Metropolis algorithm (Algorithm 1) still generates a Markov chain after the pilot time.

**Algorithm 1** A Pilot Adaptive Metropolis Algorithm.

```
k = 0;
for i = 1 to N do
    if i ≡ 0 mod m and i ≤ t_r t_0 then
        C_k = Ξ_PAM(C_{k−1});
        k++;
    end if
    Generate y from q_{C_k}(x^{(i−1)}, ·) and u from U(0, 1);
    if u ≤ α(x^{(i−1)}, y) then
        x^{(i)} = y;
    else
        x^{(i)} = x^{(i−1)};
    end if
end for

return {x^{(1)}, x^{(2)}, ..., x^{(N)}}
```

### 5.1. A pilot adaptive metropolis algorithm

This algorithm updates the proposal distribution by changing its covariance matrix so the acceptance ratio of the chain is close to the optimal acceptance ratio $a_o$ of the chain after the last adaption. There is no analytical framework for choosing an optimal acceptance ratio $a_o$ when the target distribution is unknown. Hence, the choice of $a_o$ is usually based on the

result from Gelman et al. [37]. For a normal target and proposal distribution, the optimal acceptance ratio is approximately .45 in the one dimensional case and .234 when the number of dimension converges to infinity [37].

Suppose we perform $t_0$ adaptions, one every $t_r$ iterations, where $1 < t_r t_0 < B^* < N$. Let $c_i$ denote a variable recording if the $i$th iteration of this algorithm has been accepted,

$$c_i := \begin{cases} 1, & i\text{th iteration has been accepted,} \\ 0, & \text{else.} \end{cases} \tag{43}$$

The estimator for the acceptance ratio for the $k$th proposal distribution is $\bar{a}_k = \frac{1}{t_r}\sum_{i=(k-1)t_r+1}^{kt_r} c_i$. Let $1 \gg \epsilon > 0$, where $100\epsilon$ is the percent change per adaption in the entries of the proposal distribution's covariance matrix $C$. The $k$th adaption modifies the current covariance matrix $C_{k-1}$ as

$$C_k = \Xi_{PAM}(C_{k-1}) := \begin{cases} (1+\epsilon)C_{k-1}, & \text{if } \bar{a}_k > a_o, \\ C_{k-1}, & \text{if } \bar{a}_k = a_o, \\ (1-\epsilon)C_{k-1}, & \text{if } \bar{a}_k < a_o. \end{cases} \tag{44}$$

Informally speaking, the algorithm modifies the covariance matrix in the pilot time $t_r t_0$ to converge closely to the optimal acceptance ratio. Then, the standard Metropolis–Hastings algorithm begins with the latest state and proposal distribution of the pilot time.

In Algorithm 1, the pilot adaptive Metropolis algorithm is recapitulated with an arbitrary starting state $x^{(0)} \in E$ and a starting guess for the positive definite covariance matrix $C_0$. The chain only satisfies the Markov property after the last adaption at time $t_r t_0$ in Algorithm 1. The chain usually still moves towards the high probability areas of the target distribution during the pilot time [25–27,35].

## 6. Numerical experiments

We generate synthetic data for our simulations with a known conductivity distribution $\gamma$. We proceed to solve the forward problem using $\gamma$ in the finite element Galerkin method. For each case, we had two different discrete conductivity distributions inside a circular domain $\Omega$. We used $|T| = 4128$ triangles and $N = 2129$ linear basis functions for $u$ and $\gamma$ on $\Omega$. There are $L = 16$ equally spaced electrodes placed along the boundary with each electrode covering a surface area of 5mm. Next, we added random Gaussian noise with a standard deviation that is 1% and 3% of the maximum measurement, respectively. In order to avoid the inverse crime, we reconstructed $\gamma$ on $\Omega$ with a smaller mesh size consisting of 1032 triangles with 549 mesh points.

Our simulations specifications are listed in the following table, which includes the known $\gamma$ with a residual $\delta$ in (22) and a stopping rule for choosing $\rho$ in (12).

The sequence of step lengths $s_k$ is chosen through a backtracking strategy, $s_1, s_1/2, \ldots$ until either the strong Wolf condition from (20) or (21) is satisfied. The maximum number of backtracking steps was set to 16. We also imposed $s_k > (\frac{\hat{d}-1}{\hat{d}})10^{-3}$ to prevent singularity in the numerical computation.

No backtracking is possible if the sequence of regularization parameter $\lambda_k$ decreases too quickly. We define $\lambda_k = \frac{\lambda_1 c}{c+k-1}$ with $\lambda_1 = 1, c = 4,$ to provide us with $\hat{d} = 1.25$.

*Example 1. Single inclusion*

The true conductivity consists of a homogeneous background and one circular inclusion of radius 0.01 mm centered at (0.0242 mm, 0.015 mm). The conductivities of the background and the inclusions are $7 \cdot 10^{-4}$ Ohm and $10^{-8}$ Ohm, respectively, as shown in Fig. 1. The reconstructions with 1% and 3% noise levels are shown in Fig. 2(a) and 2(c) for the IRGN method and Fig. 2(b) and 2(d) for the statistical inversion method.

Inversion results from the IRGN method are smooth from Tikhonov regularization. The inclusion location is effectively captured, but its support is slightly larger than the true inclusion. In particular, its significantly extended towards the center of $\Omega$ when the noise level is higher at 3%. The magnitude of the inclusion conductivity is overestimated by $10^{-4}$ Ohm compared to the true value of $10^{-8}$ Ohm. In contrast, the reconstruction by the statistical inversion method is more localized at the true location with a reasonably homogeneous background. However, with a higher noise level, the background conductivity starts to get distorted. We computed the $\ell_1$ and $\ell_2$ reconstruction errors as $e_1 = \frac{||\gamma_k - \gamma_T||_1}{||\gamma_T||_1}$ and $e_2 = \frac{||\gamma_k - \gamma_T||_2}{||\gamma_T||_2}$, where $\gamma_T$ is the true conductivity distribution.

*Example 2. Double inclusion*

The true conductivity consists of a homogeneous background and double circular inclusions of radius 0.01mm centered at ($\pm$0.036mm, 0mm). The conductivities of the background and the inclusions are $7 \cdot 10^{-4}$ Ohm and $10^{-8}$ Ohm, respectively, as shown in Fig. 3. The reconstructions with 1% and 3% noise levels are shown in Fig. 4(a) and 4(c) for the IRGN method and Fig. 4(b) and 4(d) for the statistical inversion method. We observe that both methods are able to retrieve the inclusions. In IRGN method, the supports of the inclusions are extended towards the center of $\Omega$, and the magnitude of the background conductivity is slightly overestimated. However, the statistical approach provides a more localized solution with a sharper

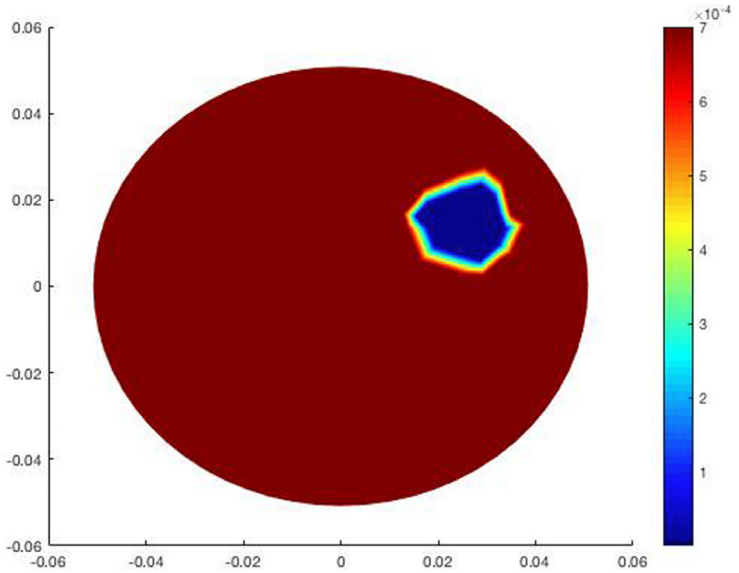


**Fig. 1.** True conductivity distribution with single inclusion.

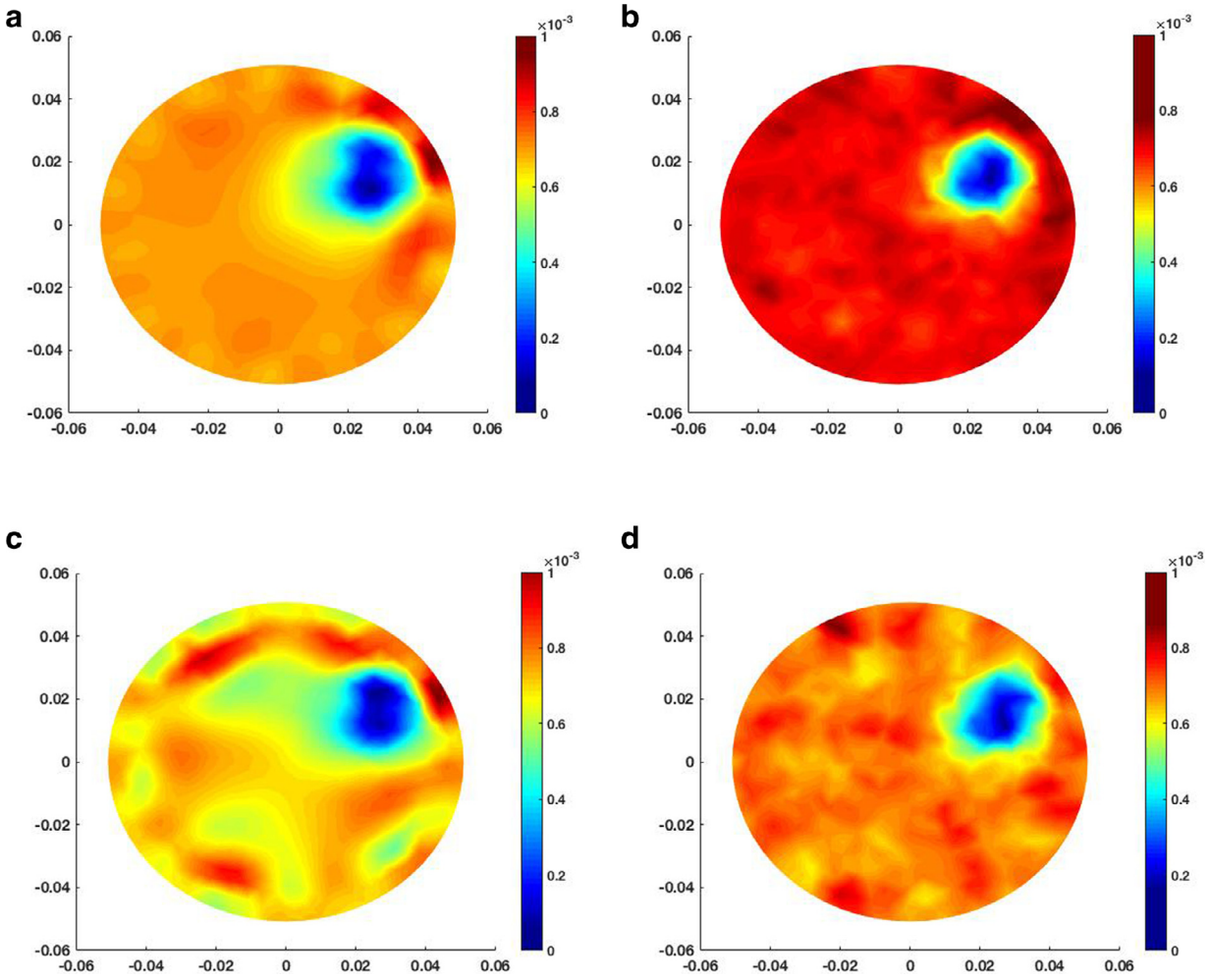


**Fig. 2.** Reconstructions for Example 1: by IRGN method (a) with 1% noise, (c) with 3% noise; and by statistical inversion method (b) with 1% noise, (d) with 3% noise.

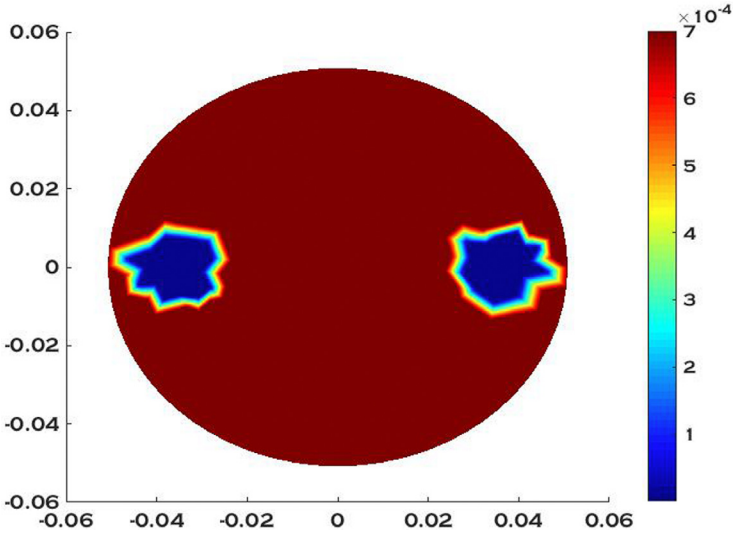


**Fig. 3.** True conductivity distribution with double inclusions.

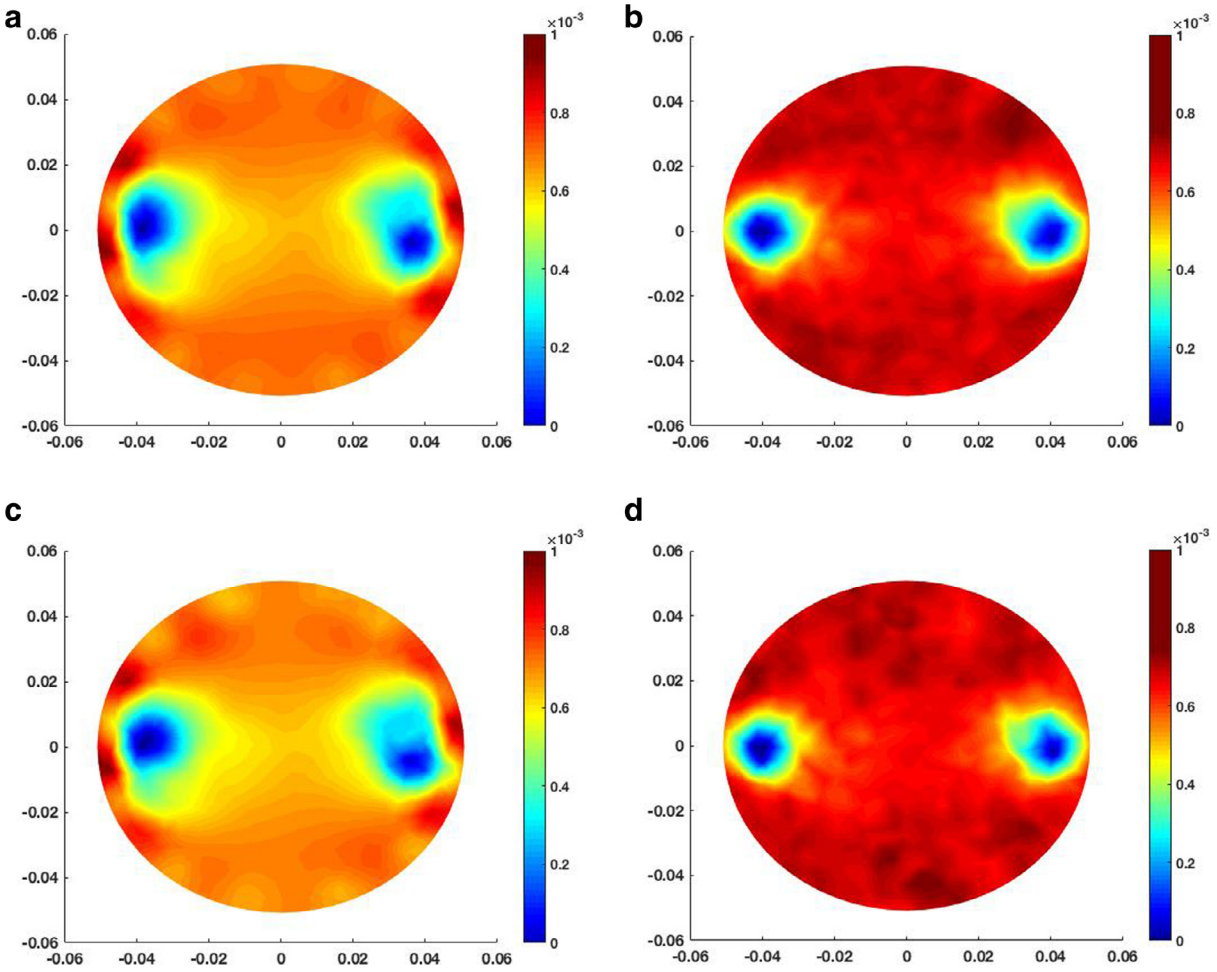


**Fig. 4.** Reconstructions for Example 1: by IRGN method (a) with 1% noise, (c) with 3% noise; and by statistical inversion method (b) with 1% noise, (d) with 3% noise.

**Table 1**
Error comparison between IRGN and statistical inversion method for Example 1.

| | 1% Noise | | 3% Noise | |
|---|---|---|---|---|
| | IRGN | Statistical Inversion | IRGN | Statistical Inversion |
| $\frac{\|\gamma_k-\gamma_T\|_1}{\|\gamma_T\|_1}$ | 0.0677 | 0.0636 | 0.1196 | 0.0911 |
| $\frac{\|\gamma_k-\gamma_T\|_2}{\|\gamma_T\|_2}$ | 0.1187 | 0.1005 | 0.1598 | 0.1279 |

**Table 2**
Error comparison between IRGN and statistical inversion method for Example 2.

| | 1% Noise | | 3% Noise | |
|---|---|---|---|---|
| | IRGN | Statistical Inversion | IRGN | Statistical Inversion |
| $\frac{\|\gamma_k-\gamma_T\|_1}{\|\gamma_T\|_1}$ | 0.1194 | 0.0946 | 0.1225 | 0.1063 |
| $\frac{\|\gamma_k-\gamma_T\|_2}{\|\gamma_T\|_2}$ | 0.1725 | 0.1347 | 0.1775 | 0.1520 |

**Table 3**
Error comparison between IRGN and statistical inversion method for Example 3.

| | 1% Noise | | 3% Noise | |
|---|---|---|---|---|
| | IRGN | Statistical Inversion | IRGN | Statistical Inversion |
| $\frac{\|\gamma_k-\gamma_T\|_1}{\|\gamma_T\|_1}$ | 0.2201 | 0.2016 | 0.2274 | 0.2028 |
| $\frac{\|\gamma_k-\gamma_T\|_2}{\|\gamma_T\|_2}$ | 0.2563 | 0.2370 | 0.2609 | 0.2387 |

**Table 4**
Choice of $\delta$ and $\rho$ for different examples.

| Example | $\delta_{1\%}$ | $\delta_{3\%}$ | $\rho$ |
|---|---|---|---|
| 1 | 0.0328 | 0.0812 | 2 |
| 2 | 0.0590 | 0.1265 | 2 |
| 3 | 0.0616 | 0.1643 | 2 |

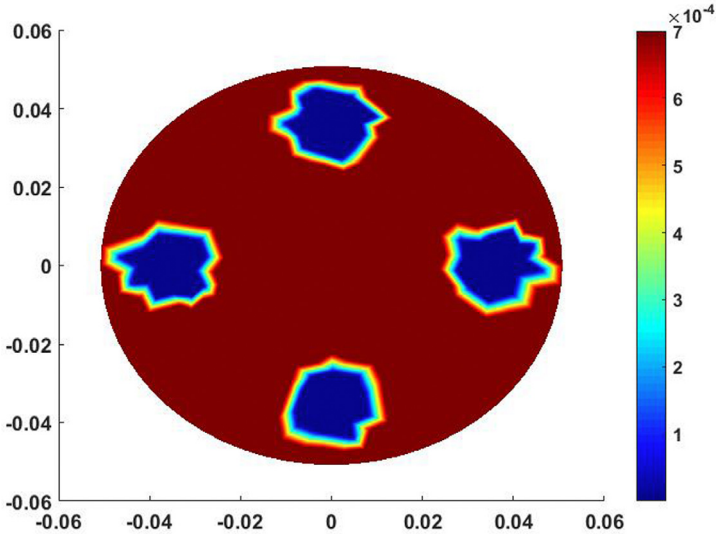


**Fig. 5.** True conductivity distribution with four inclusions.

background. We list the reconstruction errors for $\gamma$ in Table 3, which shows that the errors $e_1$ and $e_2$ are smaller for the statistical inversion method.

*Example 3. Quadruple inclusion*

The true conductivity consists of a homogeneous background and double circular inclusions of radius 0.01 mm centered at ($\pm 0.036$ mm, 0 mm) and (0 mm, $\pm$ 0.036 mm). The conductivities of the background and the inclusions are $7 \cdot 10^{-4}$ Ohm and $10^{-8}$ Ohm, respectively, as shown in Fig. 5. Multiple inclusions are challenging for most numerical algorithms. However,

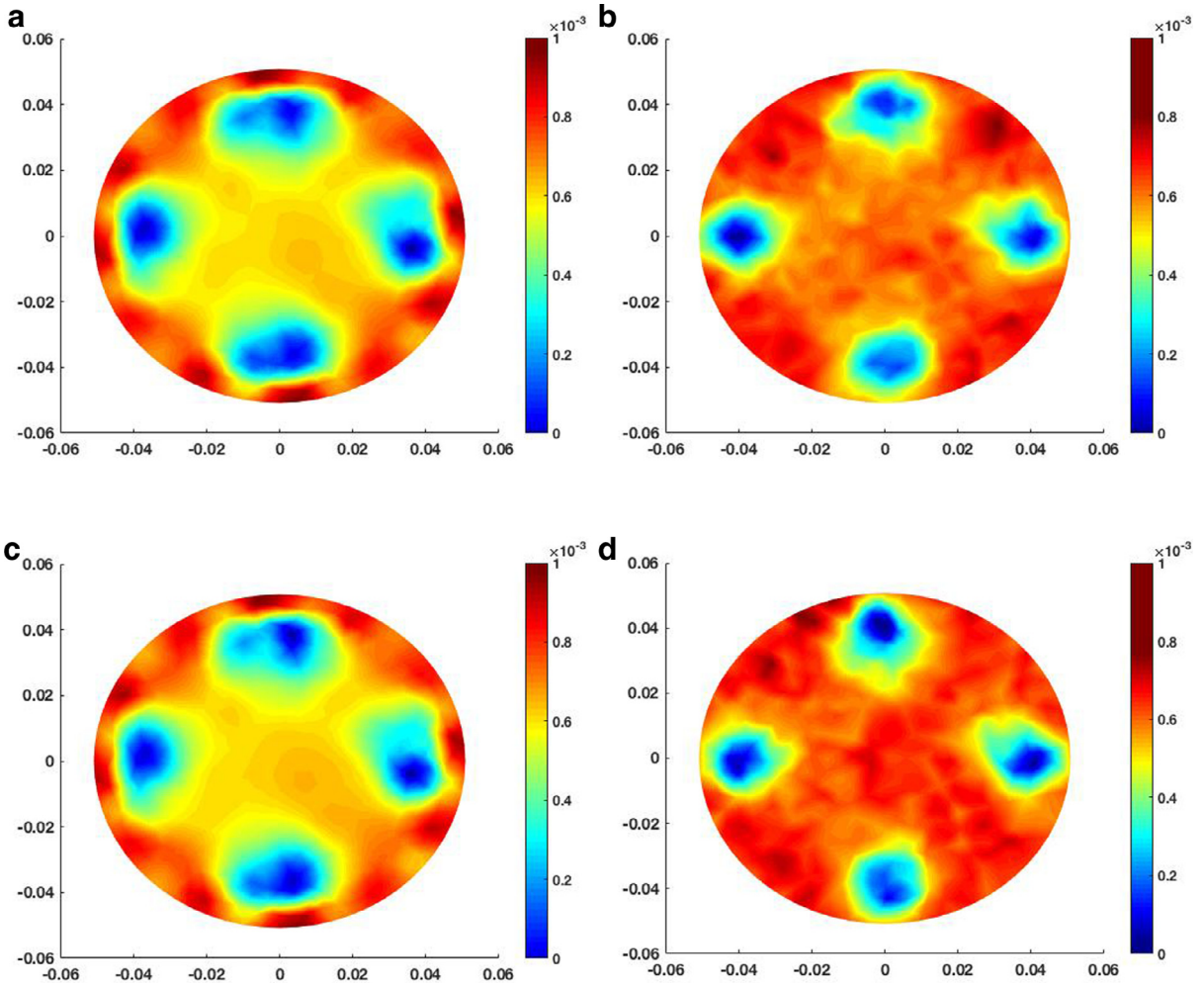


**Fig. 6.** Reconstructions for Example 1: by IRGN method (a) with 1% noise, (c) with 3% noise; and by statistical inversion method (b) with 1% noise, (d) with 3% noise.

both the approaches we used in this paper produce reasonable reconstructions of $\gamma$ for different noise levels. As expected, with higher noise level and more inclusions, their support becomes larger in IRGN method, eventually making them barely observable. On the other hand, the statistical inversion does produce significantly sharper images with more localized inclusions even for a higher noise level.

## 7. Conclusions

We constructed the EIT forward model using the CEM as the Galerkin finite element approximation and then discussed the setup for the linearized inverse problem. We discussed several regularization functions for the statistical inverse problem. We compared these general regularization functions for the statistical inverse problem to the regularization function used in the deterministic setting. We obtained a Bayes' estimate for the electrical conductivity from the Pilot Adaptive Metropolis algorithm that implements a smoothing criteria and enforces sparsity in the prior distribution. We used both the $\ell_p$ and total variation priors for regularization in the Pilot Adaptive Metropolis algorithm. We found that coupling sparsity regularization and smoothness regularization improved performance more than the case when only sparsity constraints or smoothness constraints were used.

The smoothness regularization provided regularity in the solution while the sparsity approach reduced the ill-posed nature of the inverse problem. Both criteria resulted in electrical impedance images with higher resolution. We determined that sparsity and smoothness regularization are needed during EIT inversion for improved image reconstruction. For this study, the statistical algorithm provided better reconstructions compared to IRGN in terms of both $\ell_2$ and $\ell_1$ errors. One drawback to the statistical approach is that it can be computationally expensive to run. For low dimensional cases, the running times were about 40–45 min when we used the deterministic method as an initial guess.

Additional disadvantages of the statistical algorithm and IRGN method are their dependencies on the (1) location of the inclusions and (2) data noise. Reconstruction errors are higher as data noise increases and the location of inclusions move further inside the object. This result is a direct consequence of the ill-posed inverse problem. We avoided this effect by creating inclusions closer to the surface for higher data resolution and using a relatively low noise level. Most experimental noise for EIT does not exceed 10%, but we do not know the noise distribution. Our EIT data is synthetic and produced from a controlled environment. We will continue our work to determine if we can reconstruct high resolution images when the noise distribution is unknown. Our IRGN method and statistical algorithm will be tested further with experimental data. Overall, we found that the statistical approach is a promising and computationally efficient algorithm when it uses the inverted results from the IRGN as an initial starting point.

## References


[1] M. Cheney, D. Isaacson, J.C. Newell, Electrical impedance tomography, SIAM Rev. 41 (1) (1999) 85–101.
[2] L. Borcea, Electrical impedance tomography, Inverse Probl. 18 (6) (2002) R99–R136.
[3] M. Hanke, M. Brühl, Recent progress in electrical impedance tomography, Inverse Probl. 19 (6) (2003) S65–S90.
[4] W. Daily, A. Ramirez, D. LaBrecque, J. Nitao, Electrical resistivity tomography of vadose water movement, Water Resour. Res. 28 (5) (1992) 1429–1442.
[5] O. Isaksen, A.S. Dico, E.A. Hammer, A capacitance-based tomography system for interface measurement in separation vessels, Measur. Sci. Technol. 5 (10) (1994) 1262.
[6] D.S. Holder, Electrical Impedance Tomography: Methods, History and Applications, Institute of Physics Publishing, Bristol, 2005.
[7] R.H. Bayford, Bioimpedance tomography (electrical impedance tomography), Annu. Rev. Biomed. Eng. 8 (2006) 63–91.
[8] W. Daily, A. Ramirez, D. LaBrecque, J. Nitao, Electrical resistivity tomography of vadose water movement, Water Resour. Res. 28 (5) (1992) 1429–1442.
[9] A. Borsic, C. Comina, S. Foti, R. Lancellotta, G. Musso, Imaging heterogeneities with electrical impedance tomography: laboratory results, Géotechnique 55 (7) (2005) 539–547.
[10] C. Comina, R.M. Cosentini, G. Della Vecchia, S. Foti, G. Musso, 3d-electrical resistivity tomography monitoring of salt transport in homogeneous and layered soil samples, Acta Geotech. 6 (4) (2011) 195–203.
[11] R.M. Cosentini, G. Della Vecchia, S. Foti, G. Musso, Estimation of the hydraulic parameters of unsaturated samples by electrical resistivity tomography, Géotechnique 62 (7) (2012) 583–594.
[12] R.W. Stacey, Electrical Impedance Tomography, Department of Energy and by the Department of Petroleum Engineering, Stanford University, 2006.
[13] T.C. Hou, J.P. Lynch, Electrical impedance tomographic methods for sensing strain fields and crack damage in cementitious structures, J. Intell. Mater. Syst. Struct. (2008).
[14] B. Jin, P. Maass, An analysis of electrical impedance tomography with applications to Tikhonov regularization. ESAIM: Control, Optim. Calculus Variat. 18 (04) (2012) 1027–1048.
[15] B. Jin, T. Khan, P. Maass, A reconstruction algorithm for electrical impedance tomography based on sparsity regularization, Int. J. Numer. Methods Eng. 89 (3) (2012) 337–353.
[16] T. Khan, A. Smirnova, 1d inverse problem in diffusion based optical tomography using iteratively regularized gauss-newton algorithm, Appl. Math. Comput. 161 (1) (2005) 149–170.
[17] A. Kirsh, N. Grinberg, The Factorization Method for Inverse Problems, Oxford University Press, 2008.
[18] D. Isaacson, J.L. Mueller, J.C. Newell, S. Siltanen, Reconstructions of chest phantoms by the d-bar method for electrical impedance tomography, IEEE Trans. Med. Imaging 23 (7) (2004) 821–828.
[19] J.P. Kaipio, V. Kolehmainen, E. Somersalo, M. Vauhkonen, Statistical inversion and monte carlo sampling methods in electrical impedance tomography, Inverse Probl. 16 (5) (2000) 1487.
[20] J.P. Kaipio, A. Seppänen, E. Somersalo, H. Haario, Posterior covariance related optimal current patterns in electrical impedance tomography, Inverse Probl. 20 (3) (2004) 919.
[21] A. Nissinen, L.M. Heikkinen, J.P. Kaipio, The Bayesian approximation error approach for electrical impedance tomography experimental results, Measur. Sci. Technol. 19 (1) (2008) 015501.
[22] A. Nissinen, L.M. Heikkinen, V. Kolehmainen, J.P. Kaipio, Compensation of errors due to discretization, domain truncation and unknown contact impedances in electrical impedance tomography, Measur. Sci. Technol. 20 (10) (2009) 105504.
[23] J.M. Bardsley, MCMC-based image reconstruction with uncertainty quantification, SIAM J. Scient. Comput. 34 (3) (2012) A1316–A1332.
[24] B. Jin, P. Maass, Sparsity regularization for parameter identification problems, Inverse Probl. 28 (12) (2012) 123001.
[25] T. Strauss, Statistical Inverse Problems in Electrical Impedance and Diffuse Optical Tomography, Clemson University, 2015. Doctoral Dissertation.
[26] T. Strauss, T. Khan, Statistical inversion in electrical impedance tomography using mixed total variation and non-convex $\ell_p$ regularization prior, J. Inverse Ill-posed Probl. (2015).
[27] T. Strauss, X. Fan, S. Sun, T. Khan, Statistical inversion of absolute permeability in single-phase Darcy flow, Proc. Comput. Sci. 51 (2015) 1188–1197.
[28] P.J. Vauhkonen, Image Reconstruction in Three-Dimensional Electrical Impedance Tomography, Doctoral Dissertation, University of Kuopio, 2004.
[29] E. Somersalo, M. Cheney, D. Isaacson, Existence and uniqueness for electrode models for electric current computed tomography, SIAM J. Appl. Math. 52 (1992) 1023–1040.
[30] A. Lechleiter, A. Rieder, Newton regularizations for impedance tomography: a numerical study, Inverse Probl. 22 (2006) 1967–1987.
[31] A. Smirnova, R.A. Renaut, T. Khan, Convergence and application of a modified iteratively regularized gauss-newton algorithm, Inverse Probl. 23 (2007) 1547–1563.
[32] A.B. Bakushinsky, A. Smirnova, On application of generalized discrepancy principle to iterative methods for nonlinear ill-posed problems, Numer. Func. Anal. Optim. 26 (2005) 35–48.
[33] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, NY, 1999.
[34] V.P. Palamodov, Gabor analysis of the continuum model for impedance tomography, Arkiv för Matematik 40 (1) (2002) 169–187.
[35] S. Chib, E. Greenberg, Understanding the metropolis–hastings algorithm, Am. Stat. 49 (4) (1995) 327–335.
[36] H. Haario, E. Saksman, J. Tamminen, An adaptive metropolis algorithm, Bernoulli 7 (2) (2001) 223–242.
[37] A. Gelman, G. Roberts, W. Gilks, Efficient metropolis jumping rules, Bayesian Stat. 5 (1996) 599–608.
[38] W.R. Gilks, G.O. Roberts, S.K. Sahu, Adaptive Markov chain monte carlo through regeneration, J. Am. Stat. Assoc. 93 (443) (1998) 1045–1054.
[39] A.E. Gelfand, S.K. Sahu, On Markov Chain Monte carlo acceleration, J. Comput. Graph. Stat. 3 (3) (1994) 261–276.
[40] M. Lassas, S. Siltanen, Can one use total variation prior for edge-preserving Bayesian inversion? Inverse Probl. 20 (5) (2004) 1537.
[41] F. Lucka, Fast Markov Chain Monte carlo sampling for sparse Bayesian inference in high-dimensional inverse problems using l1-type priors, Inverse Probl. 28 (12) (2012) 125012.
[42] G. Steidl, Combined first and second order variational approaches for image processing, Jahresbericht der Deutschen Mathematiker-Vereinigung 117 (2) (2015) 133–160.
[43] J.A. Iglesias, G. Mercier, O. Scherzer, A note on convergence of solutions of total variation regularized linear inverse problems, Inverse Probl. 34 (5) (2018) 055011.