



PAPER • OPEN ACCESS

Bayesian model parameter learning in linear inverse problems: application in EEG focal source imaging

To cite this article: Alexandra Koulouri and Ville Rimpiläinen 2025 *Mach. Learn.: Sci. Technol.* **6** 035020

View the [article online](#) for updates and enhancements.

You may also like

- [Eclectic approach as idea of e-educounseling preliminary system model](#)
Jumail, M F Noordin, N M Ibrahim et al.
- [Signal quality in cardiorespiratory monitoring](#)
Gari D Clifford and George B Moody
- [Special issue on applied neurodynamics: from neural dynamics to neural engineering](#)
Hillel J Chiel and Peter J Thomas



PAPER

OPEN ACCESS

RECEIVED
3 January 2025REVISED
28 June 2025ACCEPTED FOR PUBLICATION
8 July 2025PUBLISHED
4 August 2025

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Bayesian model parameter learning in linear inverse problems: application in EEG focal source imaging

Alexandra Koulouri^{1,2} and Ville Rimpiläinen^{3,4,*} ¹ Tampere University, Centre of Mathematics, Faculty of Information Technology and Communication Sciences, Tampere, Finland² University of Bath, Institute for Mathematical Innovation, Bath, United Kingdom³ University of Liverpool, Department of Physics, Liverpool, United Kingdom⁴ University of Bristol, School of Physics, Bristol, United Kingdom

* Author to whom any correspondence should be addressed.

E-mail: a.koulouri84@gmail.com and vrimpila@gmail.com**Keywords:** Bayesian statistics, model uncertainty, parameter learning, physics-informed learning, low-dimensional representation, EEG source imaging, skull conductivity

Abstract

Inverse problems are often described as limited-data problems in which the signal of interest cannot be observed directly. Therefore, a physics-based forward model that relates the signal with the observations is typically needed. Unfortunately, unknown model parameters and imperfect forward models can still undermine the signal recovery. Even though supervised machine learning techniques offer promising avenues to improve the robustness of the solutions, we have to rely on model-based learning when there is no access to ground truth for the training. In this work, we studied a linear inverse problem that included an unknown non-linearly related model parameter and utilized a Bayesian model-based learning approach that allowed reliable signal recovery and subsequently estimation of the unknown model parameter. This approach, often referred to as Bayesian approximation error approach, employed a simplified model of the physics of the problem augmented with an approximation error term that compensated for the simplification. An error subspace was spanned with the help of the eigenvectors of the approximation error covariance matrix which allowed, alongside the primary signal, simultaneous estimation of the induced error. The estimated error and signal were then used to determine the unknown model parameter. For the model parameter estimation, we tested several different approaches: a conditional Gaussian regression, an iterative (model-based) optimization, and a Gaussian process that was modeled with the help of physics-informed learning. In addition, alternating optimization was used as a reference method. As an example application, we focused on the problem of reconstructing brain activity from EEG recordings (a.k.a. EEG source imaging) under the condition that the electrical conductivity of the patient's skull was unknown in the model. Poorly selected conductivity values cause well-documented artifacts in the EEG source imaging results, and the determination of patient-specific head tissue conductivities is a significant technical problem. Our results demonstrated clear improvements in EEG source localization accuracy and provided feasible estimates for the unknown model parameter, skull conductivity.

1. Introduction

1.1. Model uncertainties in inverse problems

In many biomedical imaging, seismic imaging, remote sensing, and tomographic applications, the features of interest cannot be observed directly but must be inferred from other measurable quantities. This is known as an inverse problem [1]. The limited, complicated, or distorted nature of observations, coupled with inherent limitations such as large null spaces in model operators, often make inverse problems ill-posed [2]. To address this, standard regularization techniques are often used [3–5]. Strategies to learn effective regularizers directly from the data are a promising field of research [6–9]. However, reliable and accurate computational

models that relate the features of interest and the observations are also needed. Unfortunately, uncertainties are usually present in these models due to the modeling formulation itself (such as discretization that is required to achieve numerical solutions), imperfect sensing, or lack of knowledge of suitable model parameter values, which can significantly affect the final solution.

In applications like ultrasound imaging and magnetic resonance imaging [10, 11], well-established methods are employed to tackle unknown model parameters. In particular, for the estimation of signals of interest and model parameters, the corresponding bilinear optimization problems can be approached by convexifying the problem using methods such as linearization (lifting) [11–13] or by alternating minimization algorithms [14, 15]. However, optimizing both the primary unknowns and the model parameters can be time consuming, convergence is not always guaranteed or several reinitializations may be required, and approximations for parameters that are *deep* in the model may not always work. Techniques that handle model uncertainties with the help of machine learning (ML) and data-driven approaches [16–18] represent a new direction to accelerate computations and improve estimates in inverse problems [19–21]. However, unlike in many image and pattern recognition problems [22, 23], in inverse problems there is often a limited or no access to ground truth data to train ML algorithms and deep neural networks [10, 23]. In this study, we revisit the application of ML to severely ill-posed inverse problems when no access to ground truth is possible. Specifically, we generate training data stemming from physics-based models and integrate this domain-specific knowledge into the inverse problem formulation and recover both the primary unknown and extract an estimate for the (unknown) model parameter. We approach the problem from a Bayesian perspective, which facilitates the design of optimization algorithms under model uncertainty [24, 25].

As an application, we will focus on imaging brain sources from EEG recordings, due to its importance as a low-cost diagnostic tool and the inherent mathematical challenges and associated model uncertainties. In particular, we study imaging under the influence of erroneous skull conductivity (too high or too low), which has been shown to significantly affect the imaging solution [26–32]. In numerical experiments, we used simulated EEG recordings generated by single dipole sources in the somatosensory cortex of the brain, as similar data sets (EEG and MEG) have previously been found to be suitable for skull conductivity calibration [33–38]. Subsequently, the inferred conductivity value could be used in further EEG studies and in the optimization/planning of transcranial electrical brain stimulation treatment [39–41]. We remark that the developments of this paper are not limited to EEG source imaging but can also be used, for example, in other imaging and deconvolution problems.

1.2. Bayesian approximation error approach and our contributions

In this work, we build on the growing use of Bayesian statistics in ML by providing methodological approaches to efficiently solve linear inverse problems with model parameter uncertainties. In particular, our developments are based on the well-known Bayesian approximation error (BAE) approach [24, 42] in which a *standard* (approximate) observation model is used in the imaging accompanied with an additive error term that takes into account the effects of the approximation. Similarly as in [42, 43], we first derive a posterior distribution which simultaneously predicts the primary unknown signal and a low-order representation of the model-induced error. For the representation, we employ the top-eigenvectors of the modeling error covariance, obtained through Monte Carlo simulations. In addition to this, we derive an analytical expression for the modeling error covariance to investigate the connections between the error, model parameter, and primary signal. The provided insight allowed us to propose other Bayesian strategies to infer the model parameter.

In the application domain, we evolve our previous work, in which we demonstrated that BAE can improve source localization when skull conductivity is unknown [31], and our preliminary study on simultaneous estimation of focal source activity and skull conductivity from EEG recordings [44]. Here, with the help of simulated EEG data that corresponded to focal source activity in the somatosensory brain area, we demonstrate the potential of our approach to reconstruct the underlying focal sources and low-order estimates for the errors induced by unknown skull conductivity. Subsequently, these estimates are used to approximate the skull conductivity.

Our contributions are twofold. First, we propose a Bayesian model-based learning technique for linear inverse problems with unknown model parameters; second, we present different ways to estimate both the primary unknown and the (unknown) model parameter. We demonstrate the approach by imaging the source configuration (primary unknown) of the EEG brain imaging problem and estimating the unknown skull conductivity (model parameter) in the observation model.

Although the BAE approach has been studied in a variety of settings [45–54], its use in the estimation of model parameters has been rather limited [43]. Therefore, in this paper, we concentrate on the model parameter learning aspect of BAE, offer insight on the conditions when it works, and propose two new ways to improve the model parameter estimates. Namely, we first propose to replace the previously used linear

regression, that is based on a conditional Gaussian (CG) approximation [43], with an iterative algorithm, and second, we propose to use a Gaussian process (GP) to infer the unknown model parameter. We show that these new developments are superior to the CG as the boundary data depend non-linearly on the model parameter. Finally, we compare our methodology with other standard approaches dealing with blind inverse problems (in which model parameters are partly or fully unknown) and show that the proposed methodologies, which only rely on precomputed training data and off-line statistical learning, can outperform alternating optimization methods when unknown parameters are deep in the model.

The paper is organized as follows. We first provide mathematical preliminaries in section 2. In section 3, we then present a general overview of the BAE approach and how it can be exploited for model parameter estimation. In section 4, we describe how the BAE and the proposed model parameter estimation techniques can be applied to a model problem in which the source activity of the brain (primary unknown) and the electrical conductivity of the human skull (model parameter) are estimated from EEG recordings. Section 5 describes the implementation of the numerical simulations followed by the results and discussion. Section 6 consists of conclusions and future work. Finally, the appendices provide further details on the theory and analysis and a step-by-step pipeline on how to implement the BAE approach with model parameter estimation.

2. Background

2.1. Bayesian inference in linear inverse problems

We consider the discrete observation model

$$v = A(\sigma)x + e, \quad (1)$$

where $v \in \mathbb{R}^m$ denotes the measurements, m is the number of measurements, $A(\sigma) \in \mathbb{R}^{m \times n}$ ($m \ll n$) is the system matrix (discrete forward model) that depends on a model parameter $\sigma \in \mathbb{R}^{N_\sigma}$, $x \in \mathbb{R}^n$ is the distributed signal (e.g. amplitudes in n locations) that we aim to solve. Finally, e denotes the measurement noise that is modeled as Gaussian $e \sim \mathcal{N}(e_*, \Gamma_e)$ with mean e_* and covariance Γ_e . The model parameter σ is usually (unrealistically) considered known. However, in this paper, we treat σ as unknown, and subsequently estimate it alongside our *primary* unknown x .

In the Bayesian framework, the solution to the inverse problem is the posterior density

$$\pi(x|v) \propto \pi(v|x)\pi(x), \quad (2)$$

where $\pi(v|x)$ is the likelihood and $\pi(x)$ the prior.

In this case, the likelihood can be formulated as

$$\pi(v|x) \propto \exp\left(-\frac{1}{2}(v - A(\sigma)x - e_*)^T \Gamma_e^{-1} (v - A(\sigma)x - e_*)\right). \quad (3)$$

The model parameter σ is usually unknown and we want to obtain information on it, in addition to x . These problems are often called blind inverse problems. Such problems are much harder to solve than standard inverse problems. One approach is to estimate the parameters of the model and the signal by alternating optimizations, but this has certain limitations; namely, convergence is not always reached, and the computations can be time-consuming. Ideally, we would like to marginalize out the unknown parameter σ ; however this is not usually tractable. Hence, instead of the accurate matrix $A(\sigma)$, an approximate (standard) matrix $A_0 = A(\sigma_0)$ with fixed parameters σ_0 is often used in the inversion.

2.2. Bayesian ML

ML in inverse problems relies on training a model with known data, $D = \{(x_s, v_s)\}_{s=1:S}$, in order to predict underlying signals x given a new set of observations, v . In Bayesian ML, model parameter σ is a random variable with distribution $\pi(\sigma|D)$, x and σ are statistically independent and the likelihood is expressed as

$$\pi(v|x) = \int \pi(v|x, e, \sigma) \pi(e) \pi(\sigma|D) d\sigma de. \quad (4)$$

To estimate the primary unknown x , we have to estimate the previous likelihood; however, marginalizing out directly the unknown σ parameter is not always tractable (especially for non-linear parameters). Even if we approximate $\pi(\sigma|D) = \delta(\sigma - \hat{\sigma})$, the question on how we estimate $\hat{\sigma}$ still remains open. Furthermore, in such inverse problems where the access to ground truth distribution for x is (most of the time) impossible, we have to rely on training techniques that employ model-based learning and sampling distributions. However, to guarantee that feasible model parameters have been selected, the training data D (obtained from model distributions) has to be produced under carefully designed simulations.

3. Bayesian approximation error approach

3.1. Overview

To ease the model parameter estimation, we rewrite the observation model (1) with the help of an approximate model $A_0 = A(\sigma_0) \in \mathbb{R}^{m \times n}$, where σ_0 has fixed values, and remove the non-linearity with respect to the unknown model parameter σ as

$$v = A_0 x + \varepsilon + e, \quad (5)$$

where the additive error term $\varepsilon \in \mathbb{R}^m$, referred to as *approximation error* [24, 42], accounts for the simplification. In particular, the approximation error term is related to the unknown parameter as

$$\varepsilon(x, \sigma) = A(\sigma)x - A_0 x. \quad (6)$$

Furthermore, when the variability of the error term ε is relatively low, it can be expressed as a linear combination of few basis functions, for example, as in [42, 43] by using the eigenvalue decomposition of the error covariance, $\Gamma_\varepsilon = \sum_{k=1}^m \lambda_k w_k w_k^T$, where $\lambda_k > 0$ (in descending order) are the eigenvalues and $w_k \in \mathbb{R}^m$ the eigenvectors. Now, we can write the approximation error as

$$\varepsilon(x, \sigma) = \varepsilon_* + \varepsilon' + \varepsilon'' = W^p \alpha + \varepsilon_* + \varepsilon'', \quad (7)$$

where $\varepsilon_* = \mathbb{E}[\varepsilon]$ is the statistical mean of ε and $W^p = [w_1, \dots, w_p] \in \mathbb{R}^{m \times p}$ ($m > p$) are the p top eigenvectors. The term $\varepsilon' = W^p \alpha$ represents the dominant part of the variability in the error, while ε'' is a less informative term described by the remaining $m - p$ eigenvectors with mean $\mathbb{E}[\varepsilon''] = 0$ and covariance $\Gamma_{\varepsilon''} = \sum_{j=p+1}^m \lambda_j w_j w_j^T$. The coefficient vector $\alpha \in \mathbb{R}^p$ represents the low-dimensional error components with mean $\mathbb{E}[\alpha] = 0$ and diagonal covariance $\text{cov}[\alpha] = \text{diag}\{\lambda_1, \dots, \lambda_p\}$. More details on this error decomposition are given in appendix A. Now, by substituting the decomposed error into the observation model (5) we get

$$v = A_0 x + \varepsilon_* + W^p \alpha + \varepsilon'' + e. \quad (8)$$

In practical applications, the basis W^p and corresponding statistics can be learned from training data sampled from the joint model distributions $\pi(x)$ and $\pi(\sigma)$, or computed semi-analytically [42, 43]. Then, we can infer simultaneously x and α based on Bayes' rule, $\pi(x, \alpha | v) \propto \pi(v | x, \alpha) \pi(x, \alpha)$.

Consequently, one way to obtain parameter estimates is by maximizing the conditional probability distribution $\hat{\sigma} := \max_{\sigma} \pi(\sigma | \alpha)$, considering a joint Gaussian for (σ, α) as has been done previously [43]. However, this approach has limitations, since it considers that σ and α are sensibly related through a linear relationship (or at least monotonic relationship), which may not be valid, and the approach also ignores the fact that α depends on the signal x . This approximation can work well when the difference $\sigma - \sigma_0$ is small.

In a more general case, we can consider a mapping from $(\alpha, x) \rightarrow \sigma$. Since the analytical expression of this mapping is often unknown (or too complicated to be derived), we can use non-parametric methods, such as a GP, for the inference of the model parameter σ . Since x and σ are usually independent (this depends on the underlying physics of the problem), the only way x can affect σ in this model is through its interaction with the variable α ; in our case, this kind of interaction occurs because α depends on x through equations (6) and (7). Hence, GP learns a statistical relationship, not a causal one. GP is effectively learning the inverse function of σ from the structure imposed by $A(\sigma)$. Overall, as we shall see later by approximating (6) and obtaining statistical knowledge through training data, we can recover meaningful estimates $\hat{\sigma}$ under certain conditions.

3.2. Bayesian inference of primary signal and approximation error

To jointly estimate both x and α , we need the posterior distribution $\pi(x, \alpha | v)$. To reduce computational complexity, we adopt an approximation in which the variables $(x, \alpha, e, \varepsilon'')$ are treated as mutually Gaussian and uncorrelated⁵ [42, 43]. The resulting approximate likelihood is

$$\begin{aligned} \tilde{\pi}(v | x, \alpha) &\propto \exp\left(-\frac{1}{2}(v - A_0 x - \varepsilon_* - W^p \alpha - e_*)^T\right. \\ &\quad \left. \times (\Gamma_{\varepsilon''} + \Gamma_e)^{-1}(v - A_0 x - \varepsilon_* - W^p \alpha - e_*)\right). \end{aligned} \quad (9)$$

⁵ We note that even though in reality approximation errors usually depend on x , this *technical* approximation often results in very similar inverse solutions [24, 25, 42].

The corresponding posterior density is

$$\tilde{\pi}(x, \alpha | \nu) \propto \exp\left(-\frac{1}{2}\|L_{\varepsilon''+e}(\nu - A_0x - W^p\alpha - \varepsilon_* - e_*)\|_2^2\right) \pi(x, \alpha), \quad (10)$$

where $L_{\varepsilon''+e}$ comes from the Cholesky factorization of $(\Gamma_\varepsilon + \Gamma_e)^{-1}$. Finally, the maximum *a posteriori* (MAP) estimate is obtained by solving:

$$(\hat{x}, \hat{\alpha}) \leftarrow \arg \min_{x, \alpha} \left\{ \frac{1}{2} \|L_{\varepsilon''+e}(A_0x + W^p\alpha + \varepsilon_* + e_* - \nu)\|_2^2 - \log \pi(x, \alpha) \right\}. \quad (11)$$

Next, we can use the point estimates $(\hat{x}, \hat{\alpha})$ to infer σ with the help of learning techniques.

3.3. Inference of the model parameter

3.3.1. Variable interactions

The inference of the model parameter σ from the estimated error (as in [43])

$$\hat{\varepsilon} \approx W^p \hat{\alpha} + \varepsilon_* \quad (12)$$

is feasible if the basis functions W^p primarily capture the error induced by the unknown model parameter σ . However, this is not always the case, and in many situations, the estimated signal \hat{x} must also be considered when estimating σ . Without loss of generality in the following analysis we consider that $\sigma > 0$ is a scalar.

In particular, even though x and σ are independent variables, x can still influence σ through its interaction with α . This implies that we are dealing with an **interaction effect** rather than a direct causal relationship. To better understand how the linear and nonlinear effects of σ and x are incorporated into the basis functions W^p and, consequently, affect the estimation of α (which we use to infer σ), we express the error covariance semi-analytically using a Taylor expansion of the error (6).

Expanding around the mean σ_* , we obtain:

$$\varepsilon(x, \sigma) = A(\sigma_*)x - A_0x + \sum_k J_k(\sigma_*)z_k(\sigma)x, \quad (13)$$

where $J_k = \frac{1}{k!} \frac{\partial^k A(\sigma)}{\partial \sigma^k} \Big|_{\sigma=\sigma_*} \in \mathbb{R}^{m \times n}$ and $z_k(\sigma) = (\sigma - \sigma_*)^k$.

For simplicity, we assume $\sigma_* = \sigma_0$, and the error covariance can be expressed as:

$$\Gamma_\varepsilon = \sum_k J_k \text{cov}[xz_k] J_k^T + \sum_{k \neq j} J_k \text{cov}[xz_k, xz_j^T] J_j^T. \quad (14)$$

Since z_k and x are mutually independent, it follows from [55] that the covariance matrix of their product is given by:

$$\text{cov}[xz_k] = \text{cov}[x] \text{var}[z_k] + \text{cov}[x] \mathbb{E}[z_k]^2 + \text{var}[z_k] \mathbb{E}[x] \mathbb{E}[x]^T. \quad (15)$$

Furthermore, when the mean $\mathbb{E}[x] = 0$ and x has isotropic covariance, i.e. $\text{cov}[x] = \gamma I_n$, we get

$$\Gamma_\varepsilon = \gamma \sum_k J_k \left(\text{var}[z_k] + \mathbb{E}[z_k]^2 \right) J_k^T + \gamma \sum_{k \neq j} J_k \mathbb{E}[z_k z_j^T] J_j^T. \quad (16)$$

This expression for the error covariance suggests that x has primarily a **scaling effect** on the error covariance. Furthermore, we need to determine whether matrix J_k implicitly encodes information about x , potentially arising from the modeling process (e.g. linearization of a nonlinear problem). Overall, we expect the basis functions W^p (obtained via the eigenvalue decomposition of Γ_ε) to capture error contributions depending both from the unknown model parameter σ and the primary signal x^6 .

⁶ If $\sigma \in \mathbb{R}^{N_\sigma}$ where $N_\sigma > 1$, we can obtain a similar expression, i.e. $\Gamma_\varepsilon = \gamma \sum_{k,j} J_k (\text{cov}[z_k, z_j] \otimes I_n) J_j^T$, where $J_k \in \mathbb{R}^{m \times (n \cdot N_\sigma)}$ and \otimes is the Kronecker product.

3.3.2. CG approximation

In the Bayesian framework, we expect to obtain model parameter estimates by maximizing the conditional probability distribution $\hat{\sigma} := \max_{\sigma} \pi(\sigma|\alpha)$, however this is not always tractable. Instead, a simple approach is to approximate $\pi(\sigma|\alpha) \propto \pi(\sigma, \alpha)$ and rely on Monte Carlo simulations [43]. In particular, a joint Gaussian distribution for the pair (α, σ) can be a good approximation for small differences $\sigma - \sigma_*$. In other words, first order Taylor approximation with only the linear term in (13) is sufficient to describe the modeling error. Therefore, given an estimate for the error coefficients $\hat{\alpha}$, the mean of $\pi(\sigma|\hat{\alpha})$ is given by

$$\hat{\sigma}_{*|\alpha} = \sigma_* + \Gamma_{\sigma\alpha} \Gamma_{\alpha}^{-1} \hat{\alpha}, \quad (17)$$

where σ_* denotes the mean value of the postulated model parameter distribution, the cross-covariance $\Gamma_{\sigma\alpha}$ is estimated using samples of $\alpha^{(s)}$ and $\sigma^{(s)}$, and $\Gamma_{\alpha} = \text{diag}\{\lambda_1, \dots, \lambda_p\}$. As we shall see later (in section 4.2.2), in some cases the CG-based parameter estimation can be used to initialize a nonlinear iterative algorithm to improve the model parameter estimates.

3.3.3. GP

In more general cases, where $\alpha = g(x, \sigma)$, we can employ statistical techniques to learn the inverse function, $\sigma = f(\alpha, x)$, in order to predict σ . As explained earlier (in sections 3.1 and 3.3.1), this expression does not mean that x and σ are dependent; it simply means that given α , we may be able to refine our estimate of σ by knowing x , as the value of α depends on both σ and x . To infer σ , we can model f to be distributed as a GP, $f \sim \mathcal{GP}(m, K)$, with mean function $m(\alpha, x) = \mathbb{E}[f(\alpha, x)]$ and covariance (or kernel) function $K_{(\alpha, x)} = \mathbb{E}[(f(\alpha, x) - m(\alpha, x))(f(\alpha', x') - m(\alpha', x'))^T]$ [23]. In practice, based on training input-output data, $\mathcal{D} = \{x^{(s)}, \alpha^{(s)}\}_{s=1}^{N_s}$ and $f_{\mathcal{D}} = \{\sigma^{(s)}\}_{s=1}^{N_s}$ respectively, we can infer a new model parameter value $\hat{\sigma}$ given a new input estimated from (11). Thus, we can predict $\hat{\sigma}$ given $\hat{\mathcal{D}} = (\hat{x}, \hat{\alpha})$ and the training data \mathcal{D} .

The predictive distribution is $\pi(f_{\hat{\mathcal{D}}}|\hat{\mathcal{D}}, \mathcal{D}, f_{\mathcal{D}}) \sim \mathcal{N}(f_{\hat{\mathcal{D}}|\mathcal{D}}, K_{\hat{\mathcal{D}}|\mathcal{D}})$. The conditional mean based on the training set is

$$\hat{f}_{\hat{\mathcal{D}}|\mathcal{D}} = m(\hat{\mathcal{D}}) + K_{\hat{\mathcal{D}}, \mathcal{D}} K_{\mathcal{D}}^{-1} (f_{\mathcal{D}} - m(\mathcal{D})), \quad (18)$$

where $K_{\hat{\mathcal{D}}|\mathcal{D}} = K_{\hat{\mathcal{D}}} - K_{\hat{\mathcal{D}}, \mathcal{D}} K_{\mathcal{D}}^{-1} K_{\mathcal{D}, \hat{\mathcal{D}}}$ is the conditional variance, $K_{\hat{\mathcal{D}}, \mathcal{D}}$ is a vector of covariances between every training case and $(\hat{\alpha}, \hat{x})$, $K_{\mathcal{D}}$ is a matrix with the training data set covariance, and $K_{\hat{\mathcal{D}}}$ is the variance of $(\hat{\alpha}, \hat{x})$. The choice of the mean function and kernel can be determined based on available (sample) data and properties (e.g. physics) of the inverse problem in question.

4. Dipole source imaging and skull conductivity estimation using somatosensory EEG data

In this section, we present a framework for joint estimation of brain source activity and skull conductivity using EEG responses. Building on the previously presented Bayesian approach, we present a modified single-dipole scanning technique to estimate single sources and approximate modeling errors. Then, we describe how the modeling error, which is dependent on the skull conductivity, can be exploited to calibrate the skull conductivity value. To achieve this, we use three strategies, a CG approximation, an iterative algorithm that can refine the CG-result (CG+Iter.), and finally a GP-based regression. In section 5, we demonstrate the proposed approaches with simulated EEG recordings generated by single-dipole sources in the somatosensory cortex of the brain, as this kind of EEG (and MEG) datasets have been shown to be suitable for skull conductivity calibration [33–35, 37, 38].

4.1. Simultaneous approximation error and source estimation using model-based learning and single dipole scanning

In this work, we apply the described Bayesian framework for the reconstruction of the unknown skull conductivity (model parameter) in the EEG source imaging problem alongside the source activity. Here, we consider the distributed source modeling [56, 57] where the primary unknown $x \in \mathbb{R}^{3n}$ is the distributed (electrical) current dipole source configuration (or field) in a source space that consists of n discrete locations, $|x_i| = \sqrt{x_{i1}^2 + x_{i2}^2 + x_{i3}^2}$ is the amplitude of the source at location i , and (x_{i1}, x_{i2}, x_{i3}) are the components of x_i along the coordinate axes. The observation model is linear, as equation (1), where $v \in \mathbb{R}^m$ are the EEG recordings, and the system matrix $A(\sigma) \in \mathbb{R}^{m \times 3n}$ is called the *leadfield* matrix, which depends on the model parameter, electrical conductivity of the skull $\sigma \in \mathbb{R}^+$.

To obtain estimates for the source and the approximation error, we first compute the MAP estimates of the posterior (11). Since, in the numerical experiments (section 5), we use simulated EEG recordings

generated by single dipole sources within a restricted region (namely, the somatosensory cortex of the brain), we employ the dipole scanning solver that is widely used in EEG studies [35, 58–60]. In the dipole scanning algorithm, the main assumption is that only a single dipole source x_i is active at a time.

In Bayesian perspective, the standard dipole scanning algorithm solves $\hat{x}_i = \arg \max_{x_i} \pi(v|x_i)$, where $\pi(v|x_i)$ is the likelihood considering a single source at location i , and the solution is the estimate \hat{x}_i for which the likelihood $\pi(v|\hat{x}_i)$ is maximized.

For the proposed framework, the approximate likelihood of the dipole scanning algorithm is

$$\tilde{\pi}(v|x_i, \alpha_i) \propto \exp \frac{1}{2} \left((v - A_0^i x_i + W_i^p \alpha_i + \varepsilon_{i*})^T \Gamma_{e+\varepsilon_i}^{-1} (v - A_0^i x_i + W_i^p \alpha_i + \varepsilon_{i*}) \right). \quad (19)$$

Here the subscript i denotes a specific dipole location, the superscript in $A^i(\sigma)$ denotes the i_1 , i_2 , and i_3 columns of the leadfield matrix and $\alpha_i \in \mathbb{R}^p$ are the error coefficients encoding the modeling error related to location i . In particular, the error related to location i is $\varepsilon_i(\sigma) = A^i(\sigma)x_i - A_0^i x_i \in \mathbb{R}^m$ and its representation using basis functions is $\varepsilon_i(\sigma) = W_i^p \alpha_i + \varepsilon_{i*} + \varepsilon_{i''}$. The basis functions $W_i^p \in \mathbb{R}^{m \times p}$ are the p top eigenvectors from the eigenvalue decomposition of the approximation error covariance matrix related to a source at location i which is denoted by $\Gamma_{\varepsilon_i} = \mathbb{E}[(\varepsilon_i - \varepsilon_{i*})(\varepsilon_i - \varepsilon_{i*})^T] \in \mathbb{R}^{m \times m}$ where $\varepsilon_{i*} \in \mathbb{R}^m$ is the mean approximation error at location i .

Therefore, we can compute a MAP estimate of the pair $(\hat{x}_i, \hat{\alpha}_i)$ at each location i , and then select as a solution the pair that results in the maximum posterior $\pi(x_i, \alpha_i|v) \propto \pi(v|x_i, \alpha_i)\pi(x_i, \alpha_i)$. In particular, considering uniform prior for $\pi(x_i)$ and $\pi(\alpha_i) \sim \mathcal{N}(0, \Gamma_{\alpha_i})$ where $\Gamma_{\alpha_i} = \text{diag}\{\lambda_{i1}, \dots, \lambda_{ip}\}$, we have the prior $\pi(x_i, \alpha_i) \propto \pi(\alpha_i)$. Then, the dipole scanning algorithm solves for $i = 1, \dots, n$

$$(\hat{x}_i, \hat{\alpha}_i) = \arg \max_{x_i, \alpha_i} \pi(v|x_i, \alpha_i) \pi(\alpha_i). \quad (20)$$

The final solution is

$$(\hat{x}_l, \hat{\alpha}_l) \leftarrow \max_{i=1, \dots, n} \pi(v|\hat{x}_i, \hat{\alpha}_i) \pi(\hat{\alpha}_i), \quad (21)$$

where index l refers to the location that gives the maximum posterior. From (20) and (19) and for the given prior, we have that

$$(\hat{x}_i, \hat{\alpha}_i) = \arg \max_{x_i, \alpha_i} \exp \left(-\frac{1}{2} \left\| \begin{bmatrix} L_{\varepsilon_i''+e} A_0^i & L_{\varepsilon_i''+e} W_i^p \\ \mathbf{0} & L_{\alpha_i} \end{bmatrix} \begin{bmatrix} x_i \\ \alpha_i \end{bmatrix} - \begin{bmatrix} L_{\varepsilon_i''+e} (v - \varepsilon_{i*} - e_*) \\ 0 \end{bmatrix} \right\|_2^2 \right), \quad (22)$$

where $\|\cdot\|_2^2$ is the ℓ_2 -norm, $L_{\varepsilon_i''+e}$ comes from the Cholesky factorization of $\Gamma_{\varepsilon_i''+e}^{-1} = (\Gamma_{\varepsilon_i''} + \Gamma_e)^{-1}$ and $L_{\alpha_i} = \frac{1}{2} \text{diag}\{\lambda_{i1}^{-1/2}, \dots, \lambda_{ip}^{-1/2}\}$ where $\Gamma_{\alpha_i}^{-1} = L_{\alpha_i} \Gamma_{\alpha_i}^T$. Based on the previous, the MAP estimate for location i is

$$\begin{bmatrix} \hat{x}_i \\ \hat{\alpha}_i \end{bmatrix} := \begin{bmatrix} L_{\varepsilon_i''+e} A_0^i & L_{\varepsilon_i''+e} W_i^p \\ \mathbf{0} & L_{\alpha_i} \end{bmatrix}^{-1} \begin{bmatrix} L_{\varepsilon_i''+e} (v - \varepsilon_{i*} - e_*) \\ 0 \end{bmatrix}. \quad (23)$$

Finally, the solution is the pair denoted with subscript l that minimizes the functional

$$(\hat{x}_l, \hat{\alpha}_l) \leftarrow \min_{i=1:n} \left\{ \|L_{\varepsilon_i''+e} (v - A_0^i \hat{x}_i - \varepsilon_{i*} - W_i^p \hat{\alpha}_i - e_*)\|_2^2 + \|L_{\alpha_i} \hat{\alpha}_i\|_2^2 \right\}. \quad (24)$$

Here, the corresponding primary error is given by $\hat{\varepsilon}_l' = W_l^p \hat{\alpha}_l$.

In the following subsections, we analyze the strategies for estimating σ given the pair $(\hat{x}_l, \hat{\alpha}_l)$.

4.2. Skull conductivity estimation

To obtain sensible estimates for the skull conductivity (model parameter), the estimated $\hat{\alpha}$ (and the corresponding modeling error) has to be clearly correlated with σ . To analyze this, we inspect first the covariance matrix of the approximation error at location l that is analytically given by

$$\Gamma_{\varepsilon_l} = \gamma \left(\text{var}(\sigma) J_l^T J_l + \sum_{k,j \geq 2} \text{cov}[z_k, z_j] J_k^T J_j \right), \quad (25)$$

Algorithm 1. Iterative linear skull conductivity estimation.**Initialization:** initialize $\sigma_1 = \sigma_{CG}$ using the CG solution (26).**DO** Update leadfield model, $A^l(\sigma_t)$ Compute Jacobian matrix, $J_{\sigma_t} = \frac{\partial A^l(\sigma_t)}{\partial \sigma_t}$ as in [61] Compute, $b = w_{l1}\hat{\alpha}_l + \varepsilon_{l*} - (A^l(\sigma_t) - A_0^l)\hat{x}_l$ Solve: $\sigma_{t+1} = \min_{\sigma} \|b - J_{\sigma_t}\hat{x}_l(\sigma - \sigma_t)\|_2^2$ **WHILE** $\|\sigma_{t+1} - \sigma_t\|_2 < \epsilon_{\text{tolerance}}$

where $z_k = (\sigma - \sigma_*)^k$, $J_k^l = \frac{1}{k!} \frac{\partial^k A^l(\sigma_*)}{\partial \sigma^k} \in \mathbb{R}^m$ is the k th derivative of the leadfield $A^l(\sigma)$ ⁷ and γ is the dipole variance. Here, without a loss of generality, we have considered that the standard conductivity is equal to the mean conductivity of the training data $\pi(\sigma)$, i.e. $\sigma_0 \approx \sigma_*$, to ease our analysis.

From the previous expression for the error covariance, we observe that the source x_l introduces a scaling effect on Γ_{ε_l} . Because of that and since the error covariance is $\Gamma_{\varepsilon_l} \approx W_l^p \Gamma_{\alpha_l} (W_l^p)^T$, the bases W_l^p are shaped primarily by the variability in the model parameter σ . However, due to the scaling effect of the source x_l in Γ_{ε_l} , one can expect that the error coefficients α_l will be proportional to the amplitude of x_l . Therefore, we can suspect that the methods that employ only the estimated error $\hat{\alpha}_l$ to infer σ , as the CG approximation described below, will not perform as well as the methods that include both, $\hat{\alpha}_l$ and the estimated source \hat{x}_l as inputs.

4.2.1. CG approximation

Given $\hat{\alpha}_l$, the conditional mean of the skull conductivity distribution $\pi(\sigma|\hat{\alpha}_l)$ is

$$\hat{\sigma}_{CG} = \sigma_* + \Gamma_{\sigma\alpha_l} \Gamma_{\alpha_l}^{-1} \hat{\alpha}_l, \quad (26)$$

where σ_* is the mean value of the postulated skull conductivity distribution, $\Gamma_{\sigma\alpha_l}$ is the cross-covariance between σ and α_l (estimated from the samples of $\alpha_l^{(s)}$ and $\sigma^{(s)}$, as will be described in section 5.1.4), and $\Gamma_{\alpha_l} = \text{diag}\{\lambda_{1l}, \dots, \lambda_{pl}\}$ where λ_{pl} are the first p eigenvalues of covariance Γ_{ε_l} .

This Gaussian approximation approach can provide fast results, since it relies on off-line precomputed statistics based on training data. However, in our EEG source imaging case, the approximation error and the model parameter are non-linearly related, and the approximation error depends also on the dipole values. Hence, the CG approximation may not give sufficiently accurate results further away from the linearization point σ_* .

4.2.2. CG with iterations (CG+Iter.)

In this case, it is possible to improve the CG-based model parameter estimates by iteratively updating the linearization point. Algorithm 1 presents an iterative approach that uses the CG result as an initialization and also considers the estimated value of the source.

4.2.3. GP

Another way to estimate model parameters is to employ a GP. In GP, we can predict $\hat{\sigma}$ given $\hat{\mathcal{D}} = (\hat{\alpha}_l, \hat{x}_l)$ as estimated from (24) and the training data \mathcal{D} . Hence, we have

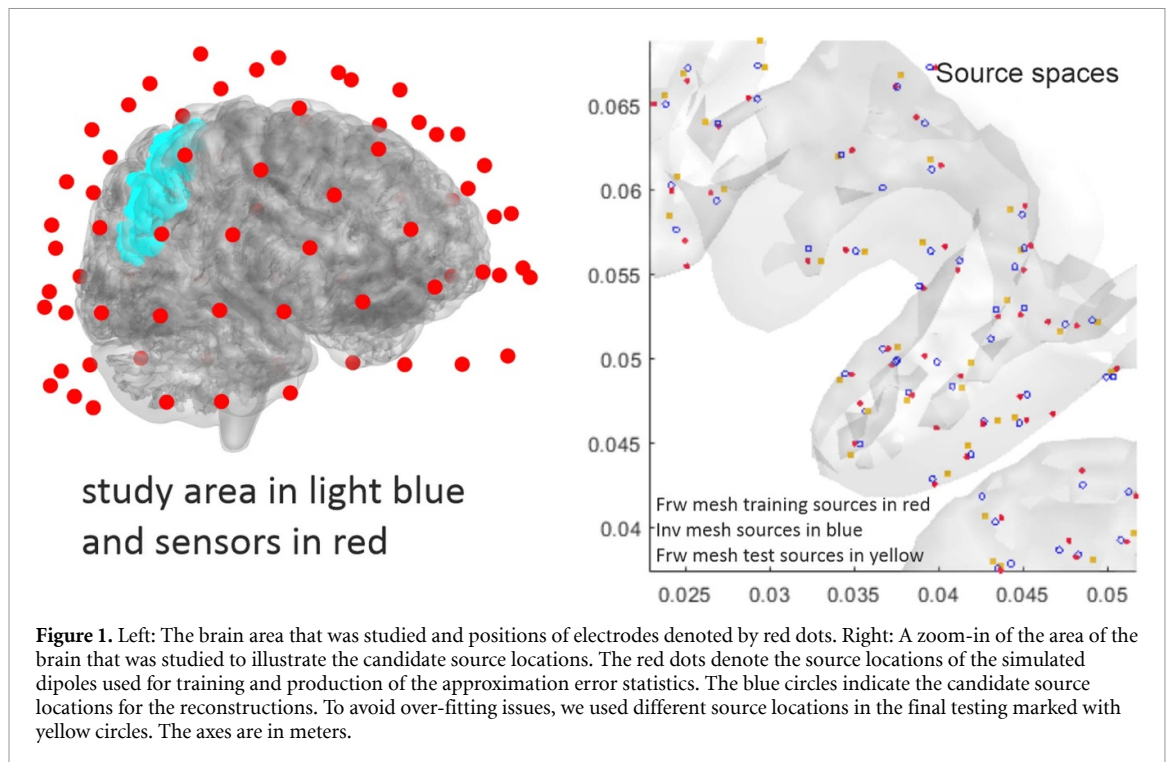
$$\hat{\sigma}_{GP} = m(\hat{\mathcal{D}}) + K_{\hat{\mathcal{D}}, \mathcal{D}} K_{\mathcal{D}}^{-1} (f_{\mathcal{D}} - m(\mathcal{D})). \quad (27)$$

Different kernel K and mean functions $m(\cdot)$ can model complicated relationships between variables, and the way they were chosen in this case is described in section 5.1.6.

5. Implementation, results and discussion

In this section, we use simulated EEG recordings generated by single-dipole sources within the somatosensory cortex of the brain to explore the capability of the proposed approaches to recover single sources and infer skull conductivity values. In section 5.1, we present technical details regarding the meshes used in forward simulations and inversion, dipole source spaces, training data, approximation error sampling, and parameter selection procedures in skull conductivity inference. In addition, we describe the

⁷ In practice, we noticed that the derivatives of order $k, j \geq 2$ were negligible.



used testing data and the different estimates that were computed. Subsequently, in section 5.2, we describe our simulated experiments and present the results under different scenarios to demonstrate the proposed approaches. Finally, we discuss further extensions and applications in section 5.3.

5.1. Implementation details

5.1.1. Meshes

For our study, we built the two used 3D meshes (a fine and a coarse one) with the help of the MRI data of the so-called ernie subject and SimNIBS 4 software⁸ [62]. Four different tissue compartments (scalp, skull, cerebrospinal fluid, and brain) were considered, and 76 electrodes were placed around the head according to the international 10–10 system (see, figure 1). The fine mesh consisted of 2103 623 tetrahedral elements joined in 377 150 nodes, and the coarse mesh of 743 575 tetrahedral elements joined in 136 868 nodes.

5.1.2. Leadfield models

The leadfield matrices used in this study were constructed with the help of custom made software that exploited Finite Element Method with linear basis functions, as in [63].

First, we created $K = 200$ leadfield matrices with skull conductivity samples $\sigma^{(k)}$ drawn from a bounded Gaussian distribution $\pi(\sigma)$ with mean $\sigma_* = 0.0103 \text{ S m}^{-1}$ and standard deviation 0.0035 S m^{-1} . The skull conductivity distribution ranged from 0.0041 S m^{-1} to 0.033 S m^{-1} according to the values reported in [37, 64, 65]. The rest of the tissue electric conductivity values were 0.43 S m^{-1} for the scalp, 1.79 S m^{-1} for cerebrospinal fluid, and 0.33 S m^{-1} for the brain (gray matter and white matter) [66]. We refer to these leadfield matrices as *sample* models, $A(\sigma)$. Out of these, 150 *sample* models were used to estimate the skull conductivity related approximation error statistics (in section 5.1.4) and the remaining 50 *sample* models were used for the training of the GP (in section 5.1.6).

For the final testing, we created two *accurate* leadfield matrices $A(\sigma)$ with the help of the fine mesh and skull conductivity values 0.0061 S m^{-1} and 0.0139 S m^{-1} , that were not included in the sample set. In addition, we created a *standard* model, A_0 , with the help of the coarse mesh and skull conductivity $\sigma_0 = 0.0103 \text{ S m}^{-1}$.

5.1.3. Source spaces

The dipole source space was restricted to an approximately 30 mm thick cross sectional area close to the somatosensory area (as shown in the left image of figure 1). This area was selected because previous studies had found it suitable for model calibration purposes [33–35, 37, 38]. In the right image of figure 1, we show a

⁸ <https://simnibs.github.io/simnibs/build/html/index.html>.

magnification of the region of interest, highlighting the source spaces corresponding to the models considered. In particular, the source space of the standard model (in the coarse mesh) is marked with blue circles in the right image of figure 1, and it uniformly covers the gray matter of the brain with 2.5 mm resolution (distance between neighboring sources). The red dots in the right image of figure 1 indicate the source locations (in the fine mesh) that were used in the estimation of the approximation error statistics, as described in section 5.1.4. The number of points (source locations) in both source spaces was 560. For the final testing, to avoid over-fitting, we produced observations from simulated dipole sources placed in different locations (in the fine mesh, marked with yellow in the right image of figure 1) than the sources that were used to produce the statistics.

5.1.4. Approximation error statistics and training data

The approximation error statistics were estimated from samples. These samples for each location index i were created by evaluating one of the (randomly chosen) *sample* models $A(\sigma^{(k)})$ and the standard model A_0 with a single dipole source $x_i^{(j)}$ as

$$\varepsilon_i^{(s)} = A(\sigma^{(k)})x_i^{(j)} - A_0x_i^{(j)}. \quad (28)$$

A set of $J = 100$ single radial (with respect to the center of the brain) dipole samples $x_i^{(j)}$ with amplitudes drawn from a Rayleigh distribution $\pi(|x_i|) = \text{Rayleigh}(\sqrt{2}\gamma)$ with parameter $\gamma = 1.85$ and mode 2.55, and 150 different sample leadfield models were used to estimate the approximation error statistics at each location i . The superscripts are as follows: $s = j + J(k - 1)$ where $j = 1, \dots, J$, and $k = 1, \dots, K$, and $s = 1, \dots, S$. The mean ε_{i*} and covariance matrix Γ_{ε_i} of the corresponding approximation error at location i were

$$\varepsilon_{i*} = \frac{\sum_{s=1}^S \varepsilon_i^{(s)}}{S} \text{ and } \Gamma_{\varepsilon_i} = \frac{1}{S-1} \sum_{s=1}^S (\varepsilon_i^{(s)} - \varepsilon_{i*}) (\varepsilon_i^{(s)} - \varepsilon_{i*})^T. \quad (29)$$

Based on these, we evaluated samples for $\alpha_i^{(s)}$ and $\varepsilon_i \backslash \backslash^{(s)}$ as $\alpha_i^{(s)} = W_i^T (\varepsilon_i^{(s)} - \varepsilon_{i*})$ and $\varepsilon_i \backslash \backslash^{(s)} = Q_i Q_i^T (\varepsilon_i^{(s)} - \varepsilon_{i*})$, where W_i and Q_i contained the eigenvectors $1, \dots, p$ and $p + 1, \dots, m$ of Γ_{ε_i} respectively (please see appendix A for details). Furthermore, by using these samples we estimated numerically Γ_{α_i} and the cross-covariance $\Gamma_{\sigma\alpha_i}$, i.e. $\Gamma_{\alpha_i} = \frac{1}{S-1} \sum_{s=1}^S \alpha_i^{(s)} (\alpha_i^{(s)})^T$, $\Gamma_{\sigma\alpha_i} = \frac{1}{S-1} \sum_{s=1}^S (\sigma_i^{(s)} - \sigma_{i*}) (\alpha_i^{(s)})^T$ and $\sigma_{i*} = \frac{1}{S} \sum_{s=1}^S \sigma_i^{(s)}$.

5.1.5. Selection of p eigenvectors

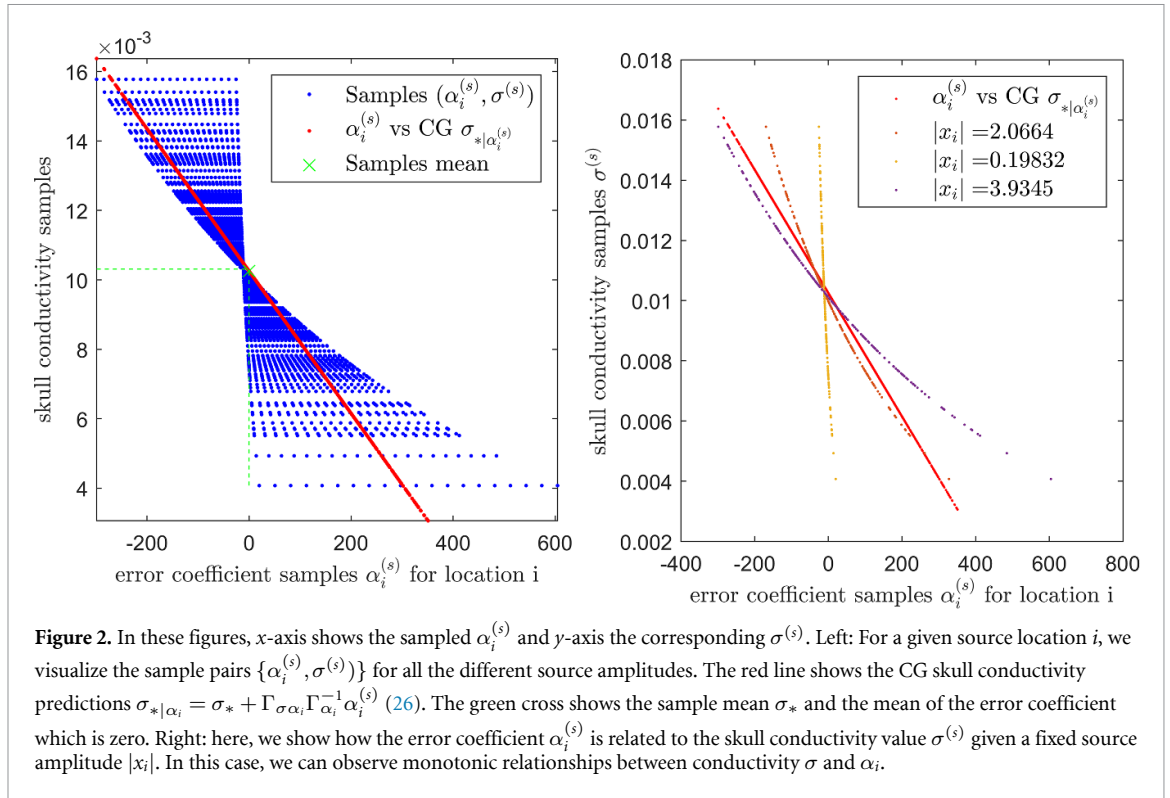
The number of eigenvectors p required to describe the primary error ε' depends on the discrepancies that produce the approximation errors. In general, it is preferable if only few eigenvectors can be used. In the current setup, the source amplitude had only a scaling effect on the covariance matrix Γ_{ε_i} . Furthermore, we noticed that the numerically estimated covariance is well approximated by the first term of equation (25). Then, since we were looking for a single model parameter value (skull conductivity $\sigma > 0$), we set $p = 1$ and use only the eigenvector corresponding to the largest eigenvalue. Thus, we had a single error coefficient $\alpha_i \in \mathbb{R}$ to estimate.

5.1.6. Inference of skull conductivity from training data

In this section, we investigate the relationships between the variables $\{x_i^{(s)}, \alpha_i^{(s)}, \sigma^{(s)}\}_{s=1}^{N_i}$ for a single dipole location i in order to better understand the skull conductivity estimates that could be obtained from CG (26) and GP (27).

5.1.6.1. CG

The red line in figure 2 shows the conditional skull conductivity predictions for different α_i values using $\sigma_{CG} = \sigma_* + \Gamma_{\sigma\alpha_i} \Gamma_{\alpha_i}^{-1} \alpha_i^{(s)}$. In addition, in the left image of figure 2 we have visualized the pairs $\{\alpha_i^{(s)}, \sigma^{(s)}\}$ for different source amplitudes. The horizontal scatter of the error coefficients α_i for a fixed $\sigma^{(s)}$ shows the effect of varying source amplitudes. Furthermore, the green cross shows the sample mean σ_* and the sample mean of the error coefficient. We can observe that for skull conductivity values close to the mean $\sigma_* = 0.0103 \text{ S m}^{-1}$ (which is equal to the standard conductivity σ_0), the conditional conductivity predictions from (26) are accurate. Also, we notice that the red curve (CG conductivity prediction curve) passes through the error coefficients that correspond to the mode of the dipole amplitude distribution $\pi(|x_i|)$. Thus, we can expect feasible conductivity predictions only when the source amplitude is close to the mode. When the underlying skull conductivity is further away from σ_* the predictions deteriorate. A more detailed explanation for this is given in appendix B.



The right image of figure 2 shows the sample pairs $\{\alpha_i^{(s)}, \sigma^{(s)}\}$ for few fixed amplitudes $|x_i|$. We observe monotonic relationships between the conductivity and the error coefficient α_i . Also, we can see that when $\sigma > \sigma_0$, the error coefficient α_i is negative, and when $\sigma < \sigma_0$ the error coefficient α_i is positive. Hence, the sign of the error coefficient could indicate whether the true skull conductivity is lower/higher than σ_0 .

5.1.6.2. GP

To allow more complex relationships (instead of just linear) and to take into account the effect of the source amplitude, we employed a GP to infer the skull conductivity. For this purpose, we needed to define the mean and covariance functions, specified by hyperparameters. In the current problem, the statistics of the underlying (unknown) function $\sigma = f(\alpha_i, x_i), f \sim \mathcal{N}(m(\alpha_i, x_i), K_{\alpha_i, x_i})$, could be approximated based on the expected relationships between the parameters and simulated training data. In particular, the mean function $m(\alpha_i, x_i)$ could be modeled as a polynomial of the following form

$$m(\alpha_i, x_i) = \sum_{0 \leq h \leq H} c_t \left(\frac{\alpha_i + k}{|x_i|} \right)^h, \tag{30}$$

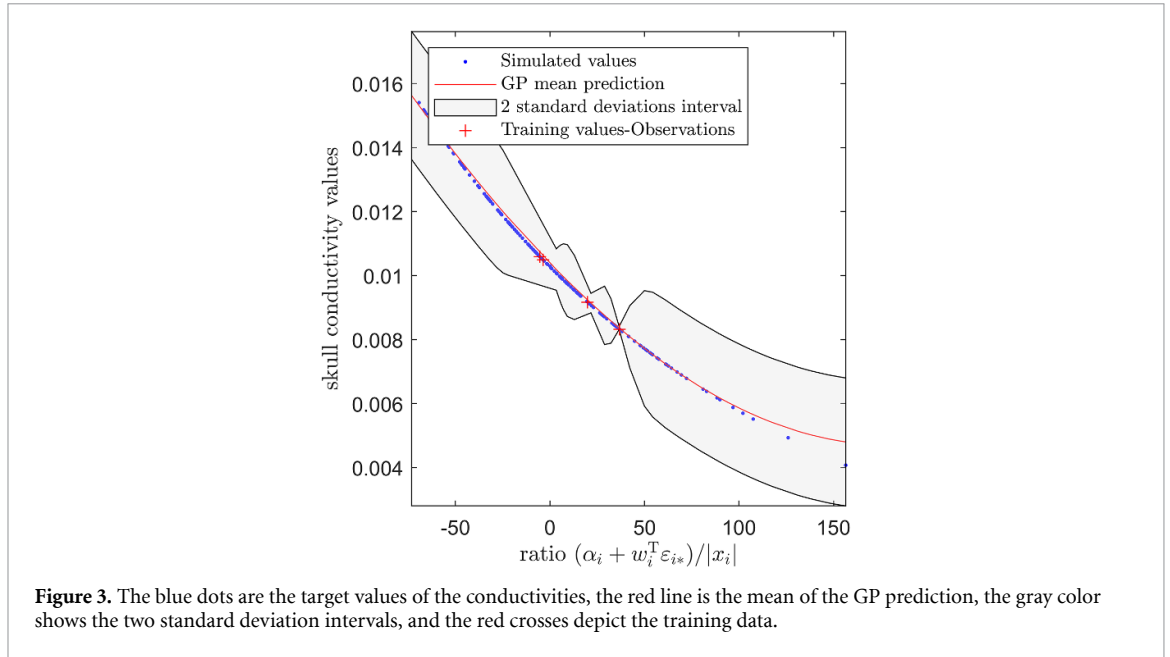
where $k = w_{i1}^T \varepsilon_{i*}$. The elements of the covariance kernel were modeled using the following exponential function

$$k_{\alpha_i, x_i} = s_f^2 \exp \left(-\frac{1}{L^2} \left[\frac{\alpha_i}{|x_i|} - \frac{\alpha_i'}{|x_i'|} \right]^2 \right), \tag{31}$$

where the scaling factor $s_f = 0.001$ has the same order of magnitude as the standard deviation of the conductivity values (vertical scaling) and the length scaling L was set equal to 10 based on the range of $\frac{\alpha_i}{|x_i|}$ values (aiming to avoid under-fitting and over-fitting). The selected mean and covariance functions and their parameters encapsulated our beliefs on the data and our understanding of the related physics. A qualitative explanation for these choices can be found in the appendix C.

In figure 3, the blue dots show the simulated (target) values that were not used in the estimation of the c_t coefficient, the red line shows the mean prediction of the GP regression given by (30), the shaded gray areas represent the (pointwise) two standard deviation limits corresponding to each input, and the red crosses show the observations / training data. Here, we can see that the order $H = 2$ is able to describe the relationships between the three parameters. Higher order H did not improve the fit.

In practice to apply the GP, we divided the artificially generated training data set $\{x_i^{(s)}, \alpha_i^{(s)}, \sigma^{(s)}\}_{s=1}^{N_s}$ into two subsets. The first subset was used to estimate the coefficients c_t through least squares, and the second



subset was used later as input-output training data when solving the problem (27). It is worth noting that the triplets were from additional samples that were not used in the estimation of the approximation error statistics Γ_{ε_i} , ε_{i*} nor W_i in section 5.1.4.

5.1.7. Testing data and different estimates

For the final testing, the EEG recordings were generated using one of the accurate leadfield models (in the fine mesh) as

$$v = A(\sigma_{\text{true}})x + e, \tag{32}$$

where the value of σ_{true} was either 0.00601 S m^{-1} or 0.0139 S m^{-1} that were not included in the training sets, and e denotes the additive (white) measurement noise.

5.1.7.1. Additive measurement noise

Random white noise with zero mean, $e_* = 0$, was added to the observations. In particular, the measurement noise was $e = s_c \bar{e} \in \mathbb{R}^m$ where $\bar{e} \in \mathbb{R}^m$ was sampled from a Gaussian distribution with zero mean and variance one. The scaling parameter $s_c \in \mathbb{R}^+$ was estimated with the help of the chosen signal-to-noise ratio (SNR), defined as $\text{SNR} = 10 \log_{10} \frac{\|v\|_2^2}{\|s_c \bar{e}\|_2^2}$ where $\|\cdot\|_2$ is the ℓ_2 - norm.

In the numerical experiments, two levels of SNR were used, 40 dB and 30 dB. In the following, the measurement noise covariance was $\Gamma_e = s_c^2 I_m$.

5.1.7.2. Single dipole source estimates

Since we studied only single dipole sources, we used the single dipole scanning algorithm [35, 58–60] to estimate the sources.

- When using the *standard* leadfield model (in the coarse mesh) with standard skull conductivity $\sigma_0 = 0.0103 \text{ S m}^{-1}$, we solved the following optimization problem

$$\hat{x}_l \leftarrow \min_{i=1:n} \left\{ \min_{x_i} \|L_e(v - A_0^i x_i)\|_2^2 \right\}, \tag{33}$$

where L_e is a matrix square root of $\Gamma_e^{-1} = L_e^T L_e$. In practice, this functional was minimized for each source space node l considering that the dipoles in other locations were zero, and the solution was the dipole source that minimized the residual.

- When using the proposed Bayesian approximation error algorithm, we found the pair that minimized

$$(\hat{x}_l, \hat{\alpha}_l) \leftarrow \min_{i=1:n} \left\{ \|L_{\varepsilon_i'+e}(v - A_0^i \hat{x}_l - \varepsilon_{i*} - w_{i1} \alpha_l)\|_2^2 + \|L_{\alpha_l} \alpha_l\|_2^2 \right\}, \tag{34}$$

where $(\Gamma_{\varepsilon_i''} + \Gamma_e)^{-1} = L_{\varepsilon_i''+e}^T L_{\varepsilon_i''+e}$, and the estimated $\hat{\alpha}_l$ was subsequently used for the skull conductivity estimation. Note that also this dipole source estimate was computed using the *standard* leadfield model A_0 (i.e. without knowledge of the *true* skull conductivity).

5.1.8. Skull conductivity estimates

Then, the skull conductivity estimation was carried out either by using only the solved $\hat{\alpha}_l$ from the previous step, or the pair $(\hat{\alpha}_l, \hat{x}_l)$.

- CG: the first approach was to use the CG estimation in which the model parameter was solved simply by using the cross-covariance as

$$\hat{\sigma}_{CG} = \sigma_* + \Gamma_{\sigma, \alpha_l} \Gamma_{\alpha_l}^{-1} \hat{\alpha}_l, \quad (35)$$

where Γ_{α_l} and $\Gamma_{\sigma, \alpha_l}$ were computed from the samples $\{\alpha_l^{(s)}, \sigma^{(s)}\}$.

- CG+Iter: to improve the above CG solution, we employed the iterative approach presented in section 4.2.2 (algorithm 1) that took the estimated source also as an input.
- GP: skull conductivity estimation using a GP

$$\hat{\sigma}_{GP} = m(\hat{\alpha}_l, \hat{x}_l) + K_{(\hat{\alpha}_l, \hat{x}_l), \mathcal{D}} K_{\mathcal{D}}^{-1} (f_{\mathcal{D}} - m(\mathcal{D})), \quad (36)$$

where $\mathcal{D} = \{\alpha_l^{(s)}, x_l^{(s)}\}$, $f_{\mathcal{D}} = \{\sigma^{(s)}\}$ and the functions as described in section 5.1.6.2.

5.1.8.1. Reference method: alternating single dipole scanning

We compared our estimates of dipole source and skull conductivity with the estimates of a reference method based on alternating minimization [14, 15]. The corresponding estimates were the following

$$\begin{aligned} \hat{x}_l^{(k+1)} &\leftarrow \min_{i=1:n} \left\{ \min_{x_i} \|L_e \left(v - A(\sigma^{(k)})^i x_i \right)\|_2^2 \right\}, \\ \hat{\sigma}^{(k+1)} &\leftarrow \min_{\sigma} \|L_e \left(v - A(\sigma)^l \hat{x}_l^{(k+1)} \right)\|_2^2, \end{aligned} \quad (37)$$

where $A(\sigma)^l \approx A(\sigma^{(k)})^l + \frac{\partial A^l}{\partial \sigma} |_{\sigma=\sigma^{(k)}} (\sigma - \sigma^{(k)})$ where the derivative was evaluated as in [61].

5.2. Simulated experiments and results

We first study the effects of varying source amplitude on source localization accuracy and skull conductivity estimates. Subsequently, we compared the proposed algorithms with a fixed source amplitude. In all the test cases, we used single radial dipole sources. We randomly picked 88 source locations (that were not used in the training) from the area of the brain shown in figure 1. The EEG measurements v were computed using the *accurate* leadfield model (32) in the fine mesh that had either skull conductivity 0.00601 S m^{-1} or 0.0139 S m^{-1} (that were not included in the training sets).

Random white noise was added to the measurements. In section 5.2.1, a low noise level, $\text{SNR} = 40 \text{ dB}$, was used in order to study the effects of the varying source amplitude on the results. In section 5.2.2, a higher noise level, $\text{SNR} = 30 \text{ dB}$, was used.

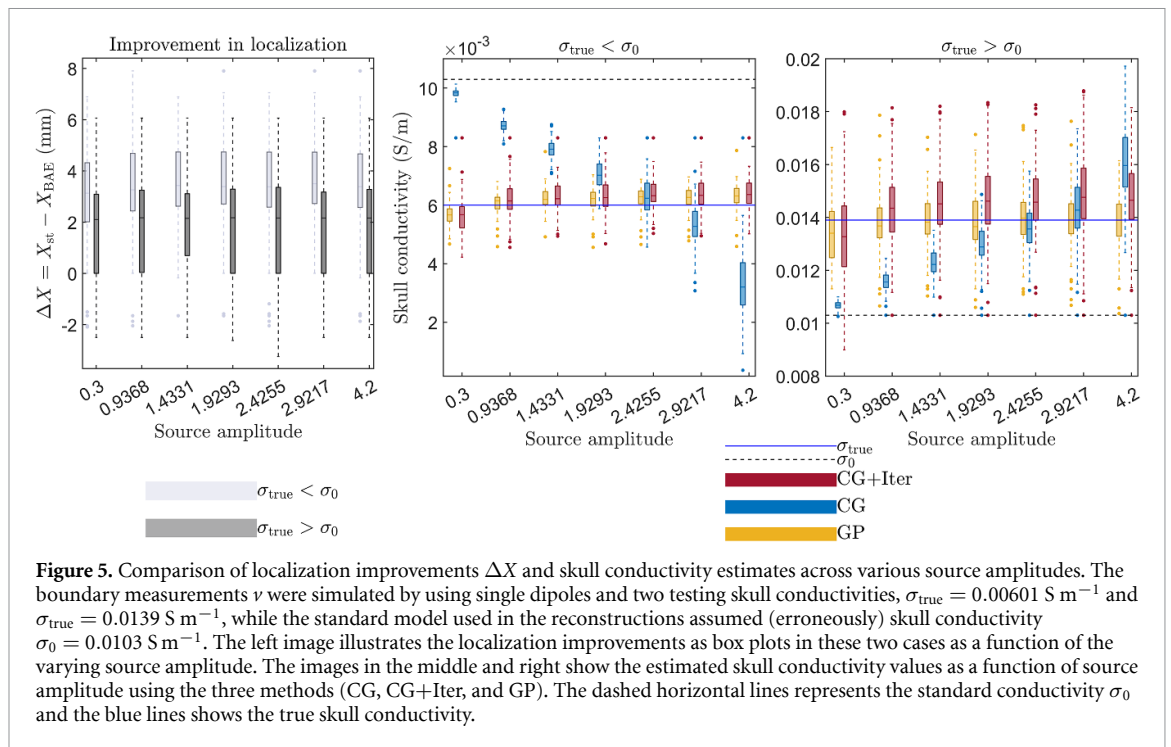
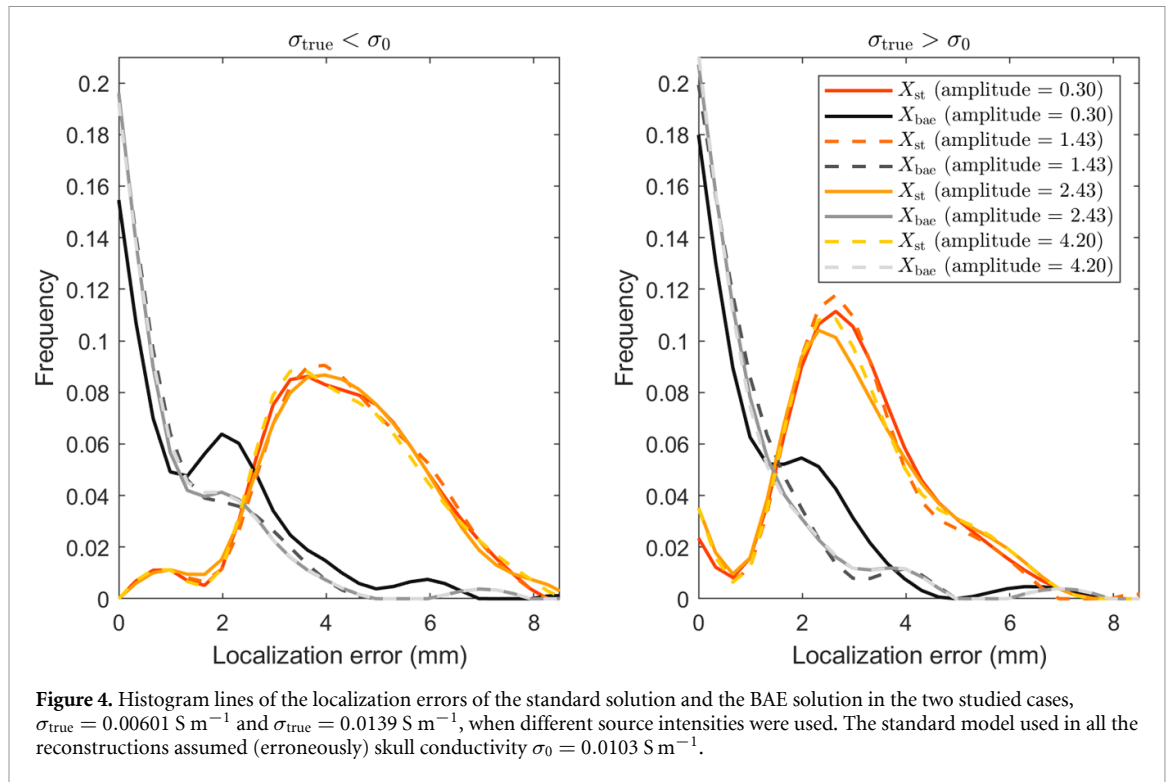
All the dipole source reconstructions were carried out using the *standard* leadfield model (in the coarse mesh) that had fixed skull conductivity $\sigma_0 = 0.0103 \text{ S m}^{-1}$; the alternating dipole scanning obviously updated the leadfield model during the iterations and used the standard leadfield model only in the initial step. For the model parameter (skull conductivity) estimation, CG+Iter also updated the leadfield model in the iterations.

To evaluate the dipole source reconstructions, we calculated localization errors (Euclidean distances in millimeters) between the true and reconstructed source location. In the following, we denote X_{st} , X_{BAE} , and X_{alt} as the localization errors of the standard solution (33), BAE solution (34), and alternating algorithm (37), respectively. Subsequently, we estimated the improvements in source localization ΔX of the algorithms defined as $X_{\text{st}} - X_{\text{BAE}}$ and $X_{\text{st}} - X_{\text{alt}}$.

5.2.1. Effects of varying source amplitude in source localization improvement and skull conductivity estimates

In the first test case, we simulated boundary measurements stemming from 88 radial dipoles (one at a time) with increasing source amplitude (from 0.3 to 4.2, in normalized source amplitude units⁹), and then

⁹ Instead of using the typical single dipole source amplitude values, 10-50 nAm, we used the normalized ones for numerical stability.



computed source reconstructions, using (33) and (34), and skull conductivity estimates using the proposed approaches, CG (35), CG+Iter. (Algorithm 1), and GP (36).

Figure 4 shows the histograms of dipole localization errors, X_{st} and X_{BAE} , of the standard solution (33) and the BAE solution (34), respectively, for the four different source intensities. Each histogram line corresponds to results of 88 simulated sources with a fixed source intensity.

As can be seen, the localization errors of the BAE solutions are much more often smaller than the localization errors of the standard solution in all the cases. Moreover, the localization errors do not seem to have any particularly strong trend with respect to the source intensity.

In figure 5, we show with box plots the source localization improvements ($\Delta X = X_{\text{st}} - X_{\text{BAE}}$, in millimeters) and the estimated skull conductivities in the two test cases as a function of the source amplitude. The

medians are represented by the horizontal lines inside the boxes, the edges of the boxes denote the 25th and 75th percentiles, and the whiskers and outliers indicate the range of values. It is worth emphasizing that the source reconstructions were carried out by first, using the standard leadfield model that assumed (erroneously) skull conductivity value $\sigma_0 = 0.0103 \text{ S m}^{-1}$, and second, using the same standard model accompanied with the statistics of the BAE modeling. Higher ΔX values in the first (left) image of figure 5 indicate greater improvements using the proposed BAE dipole scanning. Light gray refers to the $\sigma_{\text{true}} < \sigma_0$ case and dark gray to the $\sigma_{\text{true}} > \sigma_0$ case. We can see that localization improvements can be achieved using the proposed BAE dipole scanning in both cases, when compared to using the standard dipole scanning. The benefits of BAE are more evident in the case with $\sigma_{\text{true}} = 0.00601 \text{ S m}^{-1} < \sigma_0$ than in the case with $\sigma_{\text{true}} = 0.0139 \text{ S m}^{-1} > \sigma_0$.

The second (middle) and third (right) image in figure 5 show box plots of the estimated skull conductivities in the test cases with $\sigma_{\text{true}} = 0.00601 \text{ S m}^{-1} < \sigma_0$ and $\sigma_{\text{true}} = 0.0139 \text{ S m}^{-1} > \sigma_0$, respectively. The skull conductivities were estimated using the proposed techniques, CG, CG+Iter and GP, across various source amplitudes. The horizontal dashed black line represents the standard conductivity σ_0 , and the solid blue line shows the true conductivity used in the test cases.

Based on these results, the CG-based conductivity estimation performs well only when the source amplitude is close to the mode of the source amplitude distribution $\pi(|x_i|)$ (see appendix B for an explanation). Furthermore, the solution tends to be close to the standard value (σ_0) when the source amplitude is small. This is because for small source amplitudes, the reconstructed error coefficient (α_i) is usually small (due to the scaling effect, see section 4.2) and thus, favors conductivity values close to σ_0 (as was expected based on the red curve in figure 2). Furthermore, we can see that the skull conductivity estimates of CG for increasing dipole amplitude either decrease (in the $\sigma_{\text{true}} < \sigma_0$ case), or increase (in the $\sigma_{\text{true}} > \sigma_0$ case). Therefore, we observe here the interaction effect (explained in sections 3.1 and 3.3.1) based on which a change in the source amplitude can hinder the conductivity estimation with CG, even though, based on the underlying physics, the source amplitude does not depend on the skull conductivity.

As a first improvement to CG, we proposed an iterative algorithm (algorithm 1) that takes as an input the estimated source \hat{x} and uses the CG estimated conductivity as an initialization. As can be seen, the accuracy of the iterative algorithm is much better compared to using only CG, and the solution is much less dependent on the source amplitude. However, the iterative algorithm is computationally much more effortful, e.g. because of the repetitive computations of the Jacobian matrix.

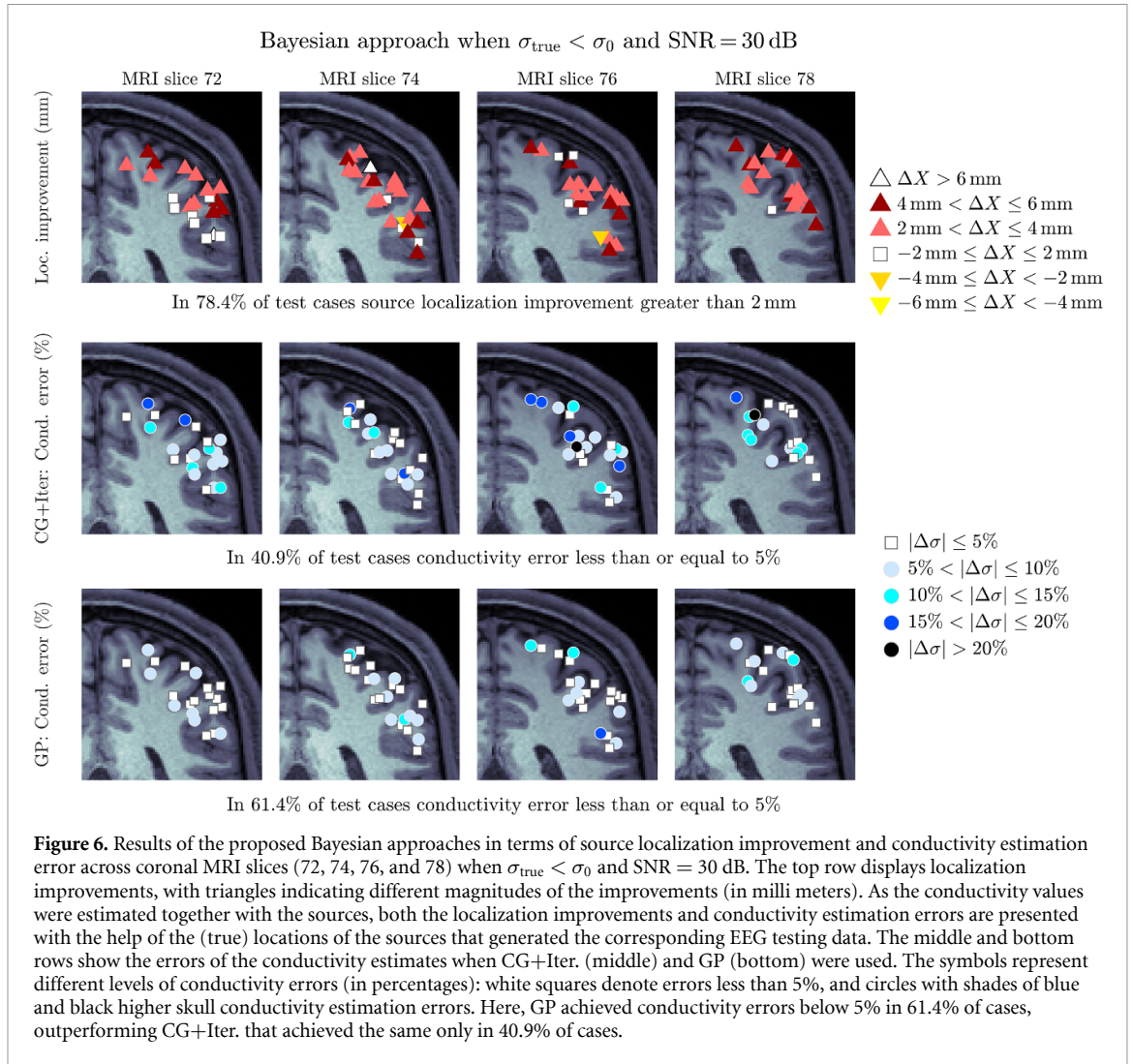
As an alternative to CG+Iter., we proposed to use a GP for the conductivity estimation. GP is computationally much less demanding, the effort being essentially the same as in the non-iterative CG. Based on the results, we can observe that GP performs even better than CG+Iter, with medians closer to the true skull conductivity and shorter whiskers.

5.2.2. Source localization improvements and skull conductivity estimates with the proposed algorithms

In the following two case studies, we fixed the source amplitude to 1.3 (normalized units). We estimated the sources and their localization accuracy, as before. For the model parameter (skull conductivity) estimation, we used only the proposed CG+Iter and GP, as it was shown in the previous section that these two were superior to the simple CG approach. We ran the algorithms for each location with 5 different (measurement) noise realizations at SNR = 30 dB, and we estimated the corresponding localization improvements, $\Delta X = X_{\text{st}} - X_{\text{BAE}}$, and the errors of the conductivity estimates as percentages, $|\Delta\sigma| = 100 \times \frac{|\hat{\sigma} - \sigma_{\text{true}}|}{\sigma_{\text{true}}}$. The presented results are the average values of ΔX and $|\Delta\sigma|$ for each tested location. As a reference method, we computed the source reconstructions and skull conductivity estimates by using the alternating dipole scanning (37).

Figure 6 depicts the results of the proposed Bayesian approach for the $\sigma_{\text{true}} < \sigma_0$ case applied over selected source locations in MRI slices 72, 74, 76, and 78. The top row presents the source localization improvements resulting from the BAE solution with respect to the standard model. The white triangles depict the locations in which the source localization improvements were the highest ($\Delta X > 6 \text{ mm}$), and the red and orange triangles correspond to smaller localization improvements, ranging from 2 mm to 6 mm. In locations marked with white squares, the localization errors were negligibly small ($\leq 1 \text{ mm}$) discretization considered. The yellow (upside down) triangles depict negative localization improvements (between -6 mm and -2 mm), which means that in these cases the *standard* model (with erroneous skull conductivity value) performed better than the proposed BAE approach. The note below the top row gives the percentage of cases in which the localization improvements were greater than 2 mm.

The middle and bottom rows present data related to conductivity errors, comparing the two different methods, CG+Iter. in the middle row and GP in the bottom row. Because the conductivity values were estimated alongside the sources, the conductivity estimation errors are presented with the help of the (true) locations of the sources that generated the corresponding EEG testing data. The symbols (squares and



circles) illustrate different conductivity estimation errors: the white squares indicate the smallest errors ($|\Delta\sigma| \leq 5\%$), and the circles (with different shades of blue and black) indicate higher estimation errors, from 5% to more than 20%.

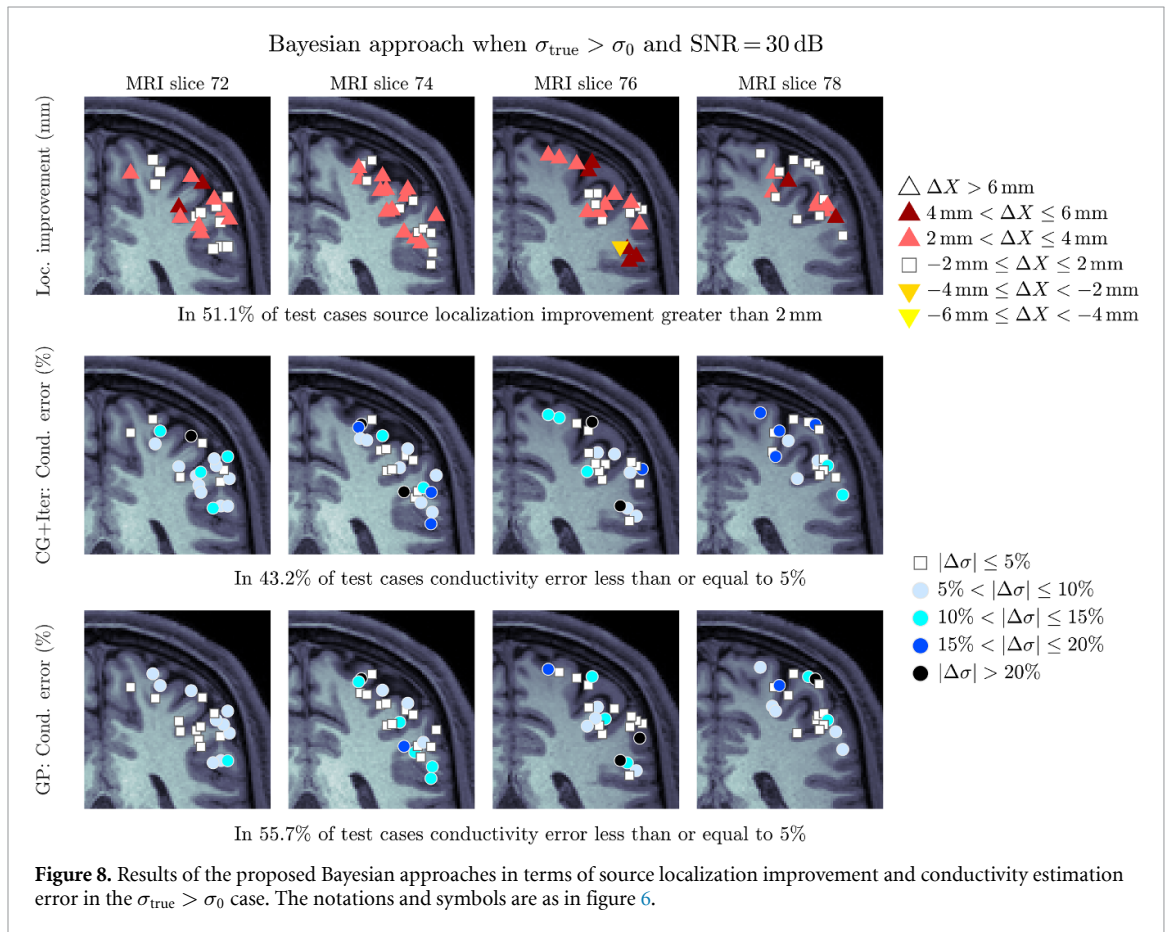
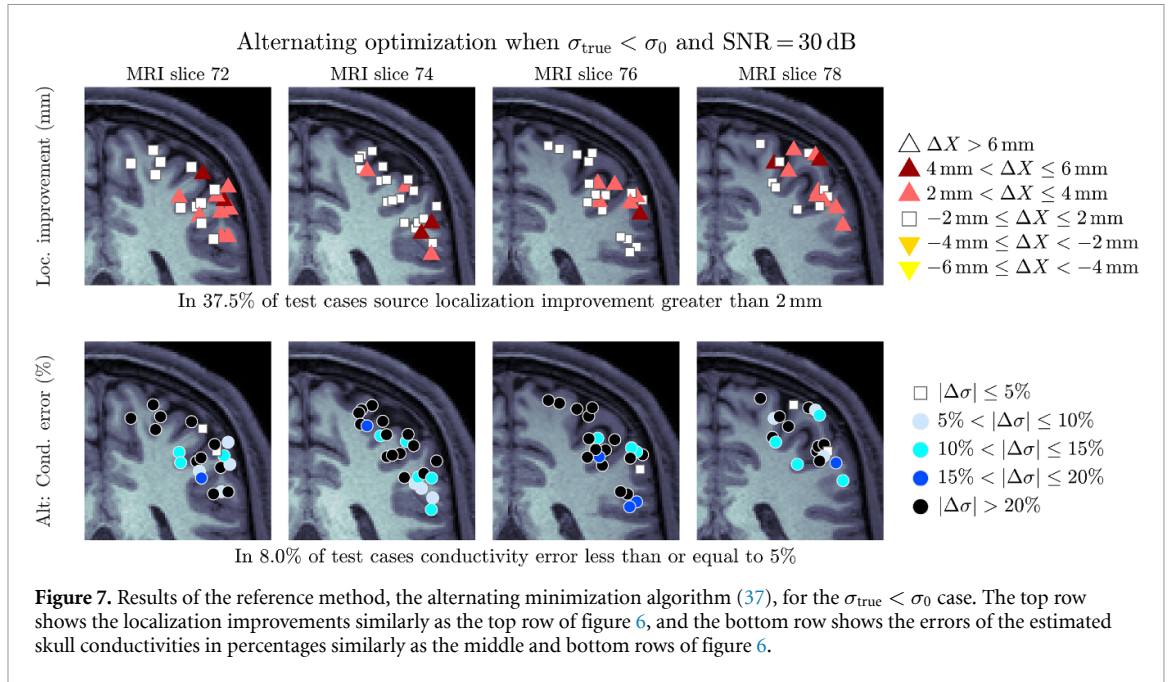
In figure 7, we show the results of the reference method, the alternating dipole scanning set in (37), in the $\sigma_{\text{true}} < \sigma_0$ case. The localization improvements and conductivity estimation errors are presented using the same notations and symbols as in figure 6. Figures 8 and 9 present the corresponding Bayesian and reference results in the $\sigma_{\text{true}} > \sigma_0$ case.

5.3. Discussion

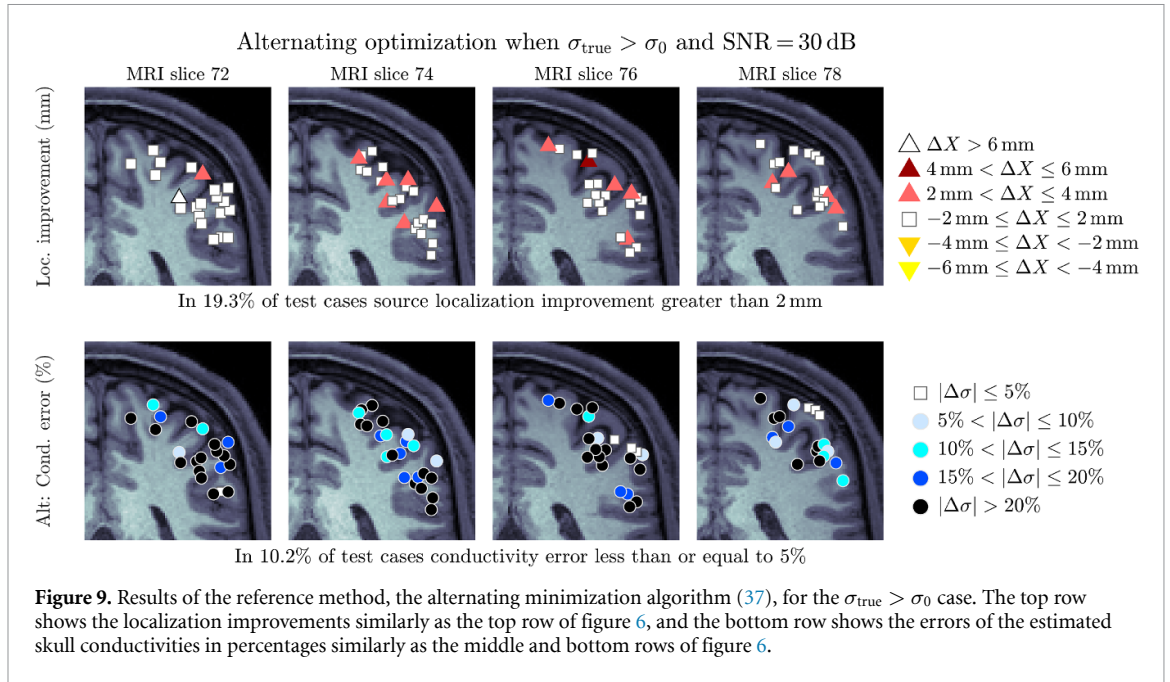
5.3.1. Source estimates

The Bayesian approximation error approach showed substantial improvements in source localization accuracy when compared to both, using the standard model with fixed skull conductivity and the alternating dipole scanning that was used as a reference method. When compared to the standard model, the BAE improved source localization by ≥ 2 mm in 78.4% of the tests in the $\sigma_{\text{true}} < \sigma_0$ case and in 51.1% of the tests in the $\sigma_{\text{true}} > \sigma_0$ case. In the rest of the tests, the performances of the BAE and the standard model were similar; there were only two source locations in which the standard model outperformed the BAE. The alternating dipole scanning performed in most tests equally to the standard model, outperforming the standard model only in 37.5% of the tests in the $\sigma_{\text{true}} < \sigma_0$ case and in 19.3% of the tests in the $\sigma_{\text{true}} > \sigma_0$ case.

The improvements provided by the BAE modeling were particularly clear in the case where the true skull conductivity was lower than σ_0 . This is because the EEG measurements, and thus the approximation errors, are non-linearly proportional to the skull conductivity, roughly $\varepsilon \propto 1/\sigma$. Therefore, the absolute values of the approximation errors in the $\sigma_{\text{true}} < \sigma_0$ case were (on average) higher than in the $\sigma_{\text{true}} > \sigma_0$ case. In the tests where the approximation errors were high (causing poor performance for the standard model), the BAE



approach could effectively alleviate the source localization errors with the help of the embedded model-based training statistics of the approximation errors. The other way around, when the approximation errors were smaller in the $\sigma_{\text{true}} > \sigma_0$ case, the potential improvements in the localization accuracy that could be gained with the help of the BAE approach were more modest. These findings are in line with our previously published results [31].



5.3.2. Model parameter estimates

The CG regression-based model parameter estimation did not work robustly for two main reasons. First, the model parameter cannot be solved using only the low-rank estimator for ε , because the solution additionally depends on the source amplitude (as seen in figures 2 and 5). Second, the model parameter depends non-linearly on ε (and α , as seen in figure 2) which can cause inaccuracies in the estimates particularly when the *true* conductivity values are relatively far from the linearization point, σ_0 .

Thus, the iterative algorithm that used the CG solution as initialization performed better. The estimates were more accurate because the (estimated) source intensity was taken into account and because the problem was iteratively linearized. The trade-off was the computational burden that significantly increased.

GP offered an effective alternative to the above as it did not require iterations. However, the prerequisite was that suitable mean and kernel functions were needed to describe the relationship between the model parameter and ε . In our case, this relationship could be approximated based on the underlying physics (see appendix C). From the tested approaches, GP gave the best model parameter estimates.

The skull conductivity estimates of the alternating optimization (37) were poor. Furthermore, we want to stress that the computation of a new lead field matrix and Jacobian matrix, both required by this algorithm, are highly demanding steps, which makes the alternating optimization unfavorable. In addition, we observed that the alternating optimization was sensitive to the initialization of the conductivity parameter. In some cases, the method did not converge with the initialization σ_0 , and another initialization point close to σ_0 was required.

5.3.3. General remarks, potential and limitations

In this work, by applying the dipole scanning approach that assumes strict single-dipole activity, we successfully removed the dependence of the approximation error on the source location. This was achieved by computing location-specific error estimates; in other words, the error depended solely on the dipole source and the skull conductivity at a fixed location. As a result, the modeling error effects could be described using a single eigenvector, and the computational effort and complexity of the estimation process were reduced considerably in the inversion.

Furthermore, the single-dipole scanning approach employed in our demonstrations has been shown to be equivalent to sLORETA which is a well-established and widely used technique for localizing single smooth source activities [59]. This result enables direct application of the approach to scenarios involving smoother, spatially extended single sources. In particular, visualizing the sLORETA estimate would be equivalent to visualizing a goodness-of-fit distribution defined by $1 - \frac{\|v - A_0^{(i)} \hat{x}^i - w_i \hat{\alpha}_i - \varepsilon_{i*}\|_2}{\|v\|_2}$ for $i = 1 \dots, n$, where $(\hat{x}_i, \hat{\alpha}_i)$ are the estimates obtained from the single-dipole scan.

We note that the design of prior distributions (e.g. Gaussian, Laplace or hierarchical priors) in the BAE posterior (10) of the EEG source imaging problem is not a straightforward task. This due to the fact that the EEG source imaging problem suffers from an inherent limitation, namely the so-called depth-bias [67]

which means that the reconstructed active sources typically locate close to the brain surface, regardless whether the *true* source activity is on the brain surface or deeper in the brain. Unfortunately, this depth-bias problem has not yet been fully resolved; one usually tries to overcome it by using either depth weights and/or hyperparameters in the prior distribution [68–77]. In this work, by considering strictly single-dipole activity, we eliminated the need for these weighting schemes and specially designed priors.

Here, we want to acknowledge that low SNRs can impact both primary unknowns and error estimates, and consequently the accuracy of model parameter inference. However, the specific tolerance to noise depends on properties of the inverse problem, such as the number of measurements and the characteristics of the convolution kernel. In this study, we noticed that with $\text{SNR} = 30$ dB we could still recover meaningful reconstructions. In our previous study [31], we observed that the benefits of BAE were negligible when $\text{SNR} = 20$ dB. As detailed in appendix D, one way to test this would be to compare the noise level of the measurements with the magnitude of the expected approximation errors [42].

One practical observation we made in the current work was that it is preferable to use an average model parameter value (skull conductivity) in the standard model, since it facilitates model parameter inference under Gaussian assumptions and faster convergence when linearized model parameter estimates are later performed. If the model parameter estimation step is not needed, then the choice of the standard model may not always be very critical [54].

Finally, as the BAE approach without model parameter estimation has been used in various (linear and non-linear) problems [45–54], we would expect that in many similar cases as these the joint inference of the primary unknown and error approximation (11) would be tractable. Subsequently, based on the available training data, the properties (physics) of the problem at hand, and the complexity of the approximation error term, different inference strategies for the model parameters could be formulated. In addition to the CG, and the proposed CG+Iter. and GP, others such as polynomial chaos expansions and equivalent techniques could be tested for the model parameter inference.

6. Conclusion and future work

Inverse problems are challenging due to their ill-posed nature, but also because model uncertainties are often present. The lack of access to ground truth data makes it difficult (or impossible) to apply standard and well-known supervised learning techniques for inferring model parameters. Instead, we must rely on meticulously designed simulations to generate sampling distributions for the model parameters, based on evidence, physics, and carefully made assumptions on the problem. This presents a complex challenge, making the task of designing and creating realistic simulations more critical than ever for effective training and testing, and for avoiding such issues as data snooping and overfitting.

In this paper, we employed the Bayesian approximation error approach to solve a blind linear inverse problem with a non-linear (unknown) model parameter in the forward model and tested different approaches to estimate both the primary unknown and the model parameter. To do that, we produced model-based training data that gave us samples of the approximation error. We used these samples to numerically estimate the mean and covariance matrix of the approximation errors, and subsequently a subspace spanned by the top eigenvector of this covariance was used to give a low-rank estimate of the approximation errors in the test cases.

Based on the solved primary unknown and approximation error estimate, we inferred the (unknown) model parameter using three approaches. The first approach was based on CG regression, the second iteratively updated the linearization point, and the third utilized a GP that was modeled with the help of physics-informed learning. In addition, alternating optimization was used as a reference method.

As an application, we studied the EEG source imaging problem in which the (forward) leadfield model contained a non-linear unknown model parameter, the skull conductivity. The utilization of the BAE approach clearly improved the source localization accuracy of the imaging. The proposed CG+Iter. and GP provided superior estimates for the skull conductivity when compared to the CG estimates. While the GP and CG+Iter. yielded comparable results across different source amplitudes, the GP method offers a practical advantage due to its computational efficiency and robustness against local minima, making it a preferable choice in most scenarios.

In the future, the BAE approach with GPs will be tested with experimental EEG data, and the estimated skull conductivities will be compared to results from other (experimental) skull conductivity calibration techniques. In EEG imaging, the BAE approach could be tested with more complicated priors, such as ℓ_1 and group-lasso priors, and in scenarios where several tissue conductivities are recovered simultaneously with the source configuration. However, in these new formulations, the complexity of the modeling error (e.g. existence of higher-order frequencies and thus the requirement to use more eigenvectors to describe it)

should be investigated. Finally, as the proposed developments are not limited to the EEG source imaging, other inverse problems such as kernel estimation in blind deconvolution problems could be studied.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This was supported by the Research Council of Finland: Flagship of Advanced Mathematics for Sensing Imaging and Modeling (359185) and PROFI 6-TAU Imaging Research Platform (336357).

Appendix A. Expressing approximation error via basis functions

Our aim is to express the approximation error term with the help of a small set of basis functions. We begin with the covariance of the approximation error, defined as

$$\Gamma_\varepsilon = \mathbb{E} \left[(\varepsilon - \varepsilon_*) (\varepsilon - \varepsilon_*)^T \right], \quad (38)$$

where ε_* is the mean of the approximation error. Performing eigenvalue decomposition of Γ_ε yields

$$\Gamma_\varepsilon = \sum_{k=1}^m \lambda_k w_k w_k^T = W \Lambda W^T, \quad (39)$$

where $\lambda_k \geq 0$ are the eigenvalues, and $w_k \in \mathbb{R}^m$ are the corresponding orthonormal eigenvectors (principal components) [42, 43].

This decomposition allows us to express the centered approximation error $\varepsilon - \varepsilon_*$ as a linear combination of the eigenvectors:

$$\varepsilon - \varepsilon_* \in \text{span} \{w_1, \dots, w_m\}. \quad (40)$$

In practice, we can write the error as

$$\varepsilon = \varepsilon_* + \varepsilon' + \varepsilon'', \quad (41)$$

where $\varepsilon' = \sum_{k=1}^p \alpha_k w_k$ captures the dominant part of the variation, and $\varepsilon'' = \sum_{j=p+1}^m \beta_j w_j$ is the residual component. Both components are zero-mean:

$$\mathbb{E}[\varepsilon'] = \mathbb{E}[\varepsilon''] = 0. \quad (42)$$

The coefficients α_k and β_j are obtained through inner products:

$$\alpha_k = \langle \varepsilon - \varepsilon_*, w_k \rangle, \quad \beta_j = \langle \varepsilon - \varepsilon_*, w_j \rangle. \quad (43)$$

Thus, the approximation error can be compactly represented using a reduced basis $W^p = [w_1, \dots, w_p] \in \mathbb{R}^{m \times p}$ as

$$\varepsilon' = W^p \alpha, \quad \alpha = (\alpha_1, \dots, \alpha_p)^T \in \mathbb{R}^p. \quad (44)$$

The residual term can be written as

$$\varepsilon'' = Q\beta = QQ^T (\varepsilon - \varepsilon_*), \quad (45)$$

where $Q = [w_{p+1}, \dots, w_m] \in \mathbb{R}^{m \times (m-p)}$ spans the orthogonal complement. The variances of the principal coefficients are

$$\text{var}(\alpha_k) = \lambda_k, \quad \text{for } k = 1, \dots, p. \quad (46)$$

Appendix B. Predicting skull conductivity σ under linear approximations

In this paper, we studied the CG $\pi(\sigma|\alpha_l)$ and the MAP estimates $\sigma_{\text{MAP}} = \max_{\sigma>0} \log \pi(\sigma|\alpha_l)$. We observed that these estimates were accurate only when the true source amplitude at location l was close to the mode of the distribution of the source amplitude $|x_l|$. Here, the source is described by $x_l = |x_l|\vec{n}$, considering a fixed orientation \vec{n} .

In this section, we explain why the estimates performed well only under these circumstances. First, we try to derive an explicit expression for the posterior $\pi(\sigma|\alpha_l)$ using sampling distributions and a linear approximation for the *approximation error*.

We start with

$$\pi(\sigma|\alpha_l) \propto \pi(\sigma)\pi_{|x_l|}(\alpha_l|\sigma) = \pi(\sigma) \int \pi(\alpha_l|\sigma, |x_l|)\pi(|x_l|) d|x_l|.$$

The conditional probability can be approximated as $\pi(\alpha_l|\sigma, |x_l|) \approx \delta(\alpha_l - G_l(\sigma)|x_l| - c)$, where $\delta(\cdot)$ is the Dirac-delta function and G_l is a function of σ .

For small perturbations $\sigma - \sigma_*$, the function $G_l(\sigma)$ can be estimated with the help of the 1st order Taylor series of the *approximation error* around $\sigma_* = \sigma_0$. In particular, we can write

$$\varepsilon_l = w_{l1}\alpha_l + \varepsilon_{l*} = (\sigma - \sigma_*) \frac{\partial A^l(\sigma_*)}{\partial \sigma} \vec{n}|x_l|,$$

where $\frac{\partial A^l(\sigma_*)}{\partial \sigma} \in \mathbb{R}^{m \times 3}$. Now we can multiply both sides with the eigenvector $w_{l1}^T \in \mathbb{R}^m$. This results in

$$\alpha_l + c = (\sigma - \sigma_*) w_{l1}^T \frac{\partial A^l(\sigma_*)}{\partial \sigma} \vec{n}|x_l|,$$

where $c = w_{l1}^T \varepsilon_{l*} \in \mathbb{R}$. Hence, we can define

$$G_l(\sigma) = (\sigma - \sigma_*) w_{l1}^T \frac{\partial A^l(\sigma_*)}{\partial \sigma} \vec{n}.$$

Also, we can write

$$|x_l| = G_l^{-1}(\sigma)(\alpha_l + c) = \frac{k_l^{-1}(\alpha_l + c)}{\sigma - \sigma_*},$$

where $k_l = w_{l1}^T \frac{\partial A^l(\sigma_*)}{\partial \sigma} \vec{n} \in \mathbb{R}$ and where $k_l^{-1}(\alpha_l + c)$ and $\sigma - \sigma_*$ will have the same sign.

When $x_l \in \mathbb{R}^3$ with $x_l \sim \mathcal{N}(0, \gamma^2 I_3)$, the distribution of the amplitude $|x_l|$ is $|x_l| \sim \text{Rayleigh}(\sqrt{2}\gamma)$, i.e. $\pi(|x_l|) = \frac{|x_l|}{2\gamma^2} \exp\left(-\frac{|x_l|^2}{4\gamma^2}\right)$ with $|x_l| > 0$.

Therefore, the posterior becomes

$$\pi(\sigma|\alpha_l) = \pi(\sigma)\pi_{|x_l|}(G_l^{-1}(\sigma)(\alpha_l + c)).$$

Here, the distribution $\pi(\sigma)$ can be approximated as a uniform distribution in the area around the mean σ_* and thus $\pi(\sigma|\alpha_l) \propto \pi_{|x_l|}(G_l^{-1}(\sigma)(\alpha_l + c))$.

Based on the previous, the log of $\pi(\sigma|\alpha_l)$ is

$$\log \pi(\sigma|\alpha_l) \propto \log(G_l^{-1}(\sigma)) - \frac{(G_l^{-1}(\sigma)(\alpha_l + c))^2}{4\gamma^2}.$$

To find the MAP estimate, we solve $\frac{d \log \pi(\sigma|\alpha_l)}{d\sigma} = 0$ which gives

$$\hat{\sigma}_{\text{MAP}} = \sigma_* + k_l^{-1}(\alpha_l + c) \frac{1}{\sqrt{2}\gamma},$$

where $\sqrt{2}\gamma$ is the mode of the source amplitude. Therefore, the skull conductivity is predicted based on the estimated α_l and the *mode* of the source amplitude (instead of the *estimated* amplitude).

Therefore, the CG modeling demonstrates good predictive capabilities only when the source amplitude is close to the mode (as was shown with simulations in section 5.2.1), and more advanced methods that take into account the estimated source amplitude work better in general cases (as was also shown in section 5.2.1).

Appendix C. Choice of the GP function for skull conductivity prediction in the EEG problem

From Ohm's law, we approximate that the electric potential v is proportional to the electric conductivity and a current source as $v \propto \frac{1}{\sigma} |x|$. Furthermore, we approximate that the approximation error (difference of two potentials), $\varepsilon = w_1 \alpha + \varepsilon_*$, is proportional to these variables as

$$\alpha + w_1^T \varepsilon_* \propto \frac{1}{\sigma} |x| - \frac{1}{\sigma_0} |x|. \quad (47)$$

From this, we write

$$\sigma \propto \left(\frac{\alpha + w_1^T \varepsilon_*}{|x|} + \frac{1}{\sigma_0} \right)^{-1}. \quad (48)$$

Now, if we denote $y = \frac{\alpha + w_1^T \varepsilon_*}{|x|}$ and define $\tilde{\sigma}(y) = (y + \frac{1}{\sigma_0})^{-1}$, then by applying Taylor expansion around 0, we get $\tilde{\sigma}(y) = \sum_{i=0}^{\infty} c_i y^i$.

Appendix D. Pipeline for model parameter estimation using the BAE approach

In this section, we describe an example pipeline to set up the BAE approach with combined estimation of primary unknowns and model parameters. We wish to emphasize that the implementation details are problem-specific.

1. Choose an approximate (discrete) observation model with model parameter σ_0 . The corresponding model for the inversion is then of the form: $v = A(\sigma_0)x + W^p \alpha + \nu \in \mathbb{R}^m$, where $\nu = e + \varepsilon_* + \varepsilon''$ includes mean of the (additive) measurement noise e_* , mean of the approximation error term, and the remainder of the approximation error term. Here, the approximation error has been decomposed as $\varepsilon = W^p \alpha + \varepsilon_* + \varepsilon''$. The BAE approach can be used to take into account unknown/uncertain model parameters and numerical (discretization) errors due to model reduction mappings, for example.
2. Measurement noise e : usually in calibration and controlled/repeated experiments the level of noise can be low. If possible, the level of measurement noise should be compared to the approximation error term. For example, one could expect the BAE approach to improve the results when $\|e_*\|^2 + \text{trace} \Gamma_e < \|\varepsilon_*\|^2 + \text{trace} \Gamma_\varepsilon$ [42]. In some cases, when $e_*(k)^2 + \Gamma_e(k, k) < \varepsilon_*(k)^2 + \Gamma_{\varepsilon_*}(k, k)$ holds even for few values of k , the BAE approach might still improve the results [42]. Note that here additive measurement noise has been assumed.
3. Model-based learning:
 - Estimate the parameters of the *model approximation error*, $\varepsilon = A(\sigma)x - A(\sigma_0)x$, as follows:
 - (i) Draw $S + K$ samples: $x^{(s)} \sim \pi(x)$ and $\sigma^{(s)} \sim \pi(\sigma)$.
 - (ii) Estimate *sample models*: $\{A(\sigma^{(1)}), \dots, A(\sigma^{(S+K)})\}$.
 - (iii) Evaluate *approximation error samples*: $\varepsilon^{(s)} = A(\sigma^{(s)})x^{(s)} - A_0 x^{(s)}$.
 - (iv) Estimate the mean ε_* and covariance matrix Γ_ε using the subset $\{x^{(s)}, \sigma^s, \varepsilon^{(s)}\}_{s=1}^S$ as

$$\varepsilon_* = \frac{\sum_{s=1}^S \varepsilon_i^{(s)}}{S} \quad \text{and} \quad \Gamma_\varepsilon = \frac{1}{S-1} \sum_{s=1}^S (\varepsilon^{(s)} - \varepsilon_*) (\varepsilon^{(s)} - \varepsilon_*)^T.$$

- (v) Estimate basis functions W with the help of, for example, eigenvalue decomposition of $\Gamma_\varepsilon = \sum_{k=1}^m \lambda_k w_k w_k^T = W \Lambda W^T$, where $\lambda_1 > \dots > \lambda_m$ and $W = [w_1, \dots, w_m]$.
- (vi) Select the number of basis functions $W^p = [w_1, \dots, w_p]$ needed for model parameter inference based on the complexity of the model *approximation error*. The complexity can be reflected in the distribution of the eigenvalues. The choice of p is problem specific, and the analysis of the (semi)analytical expression of the error covariance can be helpful, as presented in this paper. For example, one could check whether $\varepsilon \propto x\sigma$ and whether any high-order terms are needed to describe the variations caused by the model parameter.
- (vii) Evaluate the samples of $\alpha^{(s)} = (W^p)^T (\varepsilon^{(s)} - \varepsilon_*)$ with mean $\alpha_* = 0$ and covariance $\Gamma_\alpha = \text{diag}(\lambda_1, \dots, \lambda_p)$.
- (viii) The remaining error is $\varepsilon''^{(s)} = Q Q^T (\varepsilon^{(s)} - \varepsilon_*)$, where $Q = [w_{p+1}, \dots, w_m]$ with mean $\mathbb{E}[\varepsilon''] = 0$ and covariance $\Gamma_{\varepsilon''}$.

4. Design the optimization and, if needed, choose appropriate prior for the primary unknown based on the properties of the problem:

$$(\hat{x}, \hat{\alpha}) \leftarrow \min_{x, \alpha} \left\{ \left\| \frac{1}{2} L_{\nu} (A_0 x + W^p \alpha + \nu_* - v) \right\|_2^2 - \ln \pi(x, \alpha) \right\},$$

where $\nu_* = \varepsilon_* + e_*$ and $\Gamma_{\nu} = \Gamma_e + \Gamma_{\varepsilon''}$. For α , one can often use a zero-mean Gaussian smoothness prior.

5. Model parameter σ inference: this step is carried out with the help of the solved $\hat{\alpha}$ from the previous step. In some cases, one may also need the values of \hat{x} for the inference of σ .
- For the following cases, it is often sufficient to utilize the CG approximation, $\hat{\sigma}_{CG} = \sigma_* + \Gamma_{\sigma\alpha} \Gamma_{\alpha}^{-1} \hat{\alpha}$ where $\Gamma_{\sigma\alpha}$ is estimated from samples $\sigma^{(s)}$ and $\alpha^{(s)}$. Feasible model parameter estimates can be expected when
 - (i) Operating close to the linearization point, or in other words, when one can assume that the true value of σ is close to the chosen approximate σ_0 that is used in the standard model and the relationship between σ and α is linear (or almost linear).
 - (ii) The interaction effect of the primary unknown x on σ is only modest.
 - In other cases, the model parameter σ should be inferred using both $(\hat{\alpha}, \hat{x})$. In the following options, the mapping $(\alpha, x) \rightarrow \sigma$ should be smooth and invertible.
 - (i) Option 1: in some cases, such as in this paper, a nonlinear iterative algorithm can be incorporated and initialized with the CG approximation. Iterative algorithms can be time-consuming and computationally effortful. They do not usually require any training data, but then again, they do not benefit from the support of the training data either.
 - (ii) Option 2: a GP can be constructed based on the pre-computed samples $\{x^{(s)}, \sigma^{(s)}, \alpha^{(s)}\}_{s=S+1}^{K+S}$. The mean and covariance functions can be chosen with the help of the underlying physics of the problems and/or the training data. To perform well, the training data should cover a broad range of σ values. GPs can be simple and fast to compute, but require an additional training step.

ORCID iDs

Alexandra Koulouri  0000-0001-7605-3844

Ville Rimpiläinen  0000-0001-8679-7898

References

- [1] Kirsch A 1996 *An Introduction to the Mathematical Theory of Inverse Problems* (Springer)
- [2] Tikhonov A 1943 On the stability of inverse problems *Dokl. Akad. Nauk SSSR* **39** 195–8
- [3] Plato R and Vainikko G 1990 On the regularization of projection methods for solving ill-posed problems *Numer. Math.* **57** 63–79
- [4] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Kluwer Academic Publishers)
- [5] Hansen P C 1998 *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (Society for Industrial and Applied Mathematics)
- [6] Benning M and Burger M 2018 Modern regularization methods for inverse problems *Acta Numer.* **27** 1–111
- [7] Afkham B M, Chung J and Chung M 2021 Learning regularization parameters of inverse problems via deep neural networks *Inverse Problems* **37** 105017
- [8] Mattsson P, Zachariah D and Stoica P 2023 Regularized linear regression via covariance fitting *IEEE Trans. Signal Process.* **71** 1175–83
- [9] Burger M and Kabri S 2013 Learned regularization for inverse problems: Insights from a spectral model (arXiv:2312.09845)
- [10] Michailovich O V and Adam D R 2007 Deconvolution of medical images from microscopic to whole body images *Blind Image Deconvolution: Theory and Applications*, ed P Campisi and K Egiazarian (CRC Press)
- [11] Ni Y and Strohmer T 2024 Auto-calibration and biconvex compressive sensing with applications to parallel MRI
- [12] Ahmed A, Recht B and Romberg J 2014 Blind deconvolution using convex programming *IEEE Trans. Inf. Theory* **60** 1711–32
- [13] Ling S and Strohmer T 2015 Self-calibration and biconvex compressive sensing *Inverse Problems* **31** 115002
- [14] Bolte J, Combettes P L and Pesquet J C 2010 Alternating proximal algorithm for blind image recovery *2010 IEEE Int. Conf. on Image Processing (IEEE)*
- [15] Li X, Ling S, Strohmer T and Wei K 2019 Rapid, robust, and reliable blind deconvolution via nonconvex optimization *Appl. Comput. Harmon. Anal.* **47** 893–934
- [16] Adler J and Öktem O 2017 Solving ill-posed inverse problems using iterative deep neural networks *Inverse Problems* **33** 124007
- [17] Adler J and Öktem O 2018 Learned primal-dual reconstruction *IEEE Trans. Med. Imaging* **37** 1322–32
- [18] Jin K H, McCann M T, Froustey E and Unser M 2017 Deep convolutional neural network for inverse problems in imaging *IEEE Trans. Image Process.* **26** 4509–22
- [19] Arridge S, Maass P, Öktem O and Schönlieb C B 2019 Solving inverse problems using data-driven models *Acta Numer.* **28** 1–174
- [20] Lunz S, Hauptmann A, Tarvainen T, Schönlieb C B and Arridge S 2021 On learned operator correction in inverse problems *SIAM J. Imaging Sci.* **14** 92–127
- [21] Arridge S, Hauptmann A and Korolev Y 2023 Inverse problems with learned forward operators (arXiv:2311.12528)
- [22] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44

- [23] Murphy K P 2022 *Probabilistic Machine Learning* (MIT Press Ltd) (available at: www.ebook.de/de/product/41791478/kevin_p_murphy_probabilistic_machine_learning.html)
- [24] Kaipio J P and Somersalo E 2004 *Statistical and Computational Inverse Problems* (Applied Mathematical Series (Springer))
- [25] Kaipio J and Somersalo E 2007 Statistical inverse problems: Discretization, model reduction and inverse crimes *J. Comput. Appl. Math.* **198** 493–504
- [26] Vanrumste B, Hoey G V, de Walle R V, D’Have M, Lemahieu I and Boon P 2000 Dipole location errors in electroencephalogram source analysis due to volume conductor model errors *Med. Biol. Eng. Comput.* **38** 528–34
- [27] Dannhauer M, Lanfer B, Wolters C and Knösche T 2011 Modeling of the human skull in EEG source analysis *Hum. Brain Mapp.* **32** 1383–99
- [28] Lew S, Sliva D D, Choe M, Grant P E, Okada Y, Wolters C H and Hämäläinen M S 2013 Effects of sutures and fontanels on MEG and EEG source analysis in a realistic infant head model *NeuroImage* **76** 282–93
- [29] Ollikainen J O, Vauhkonen M, Karjalainen P A and Kaipio J P 1999 Effects of local skull inhomogeneities on EEG source estimation *Med. Eng. Phys.* **21** 143–54
- [30] Montes-Restrepo V, van Mierlo P, Strobbe G, Staelens S, Vanderberghe S and Hallez H 2014 Influence of skull modeling approaches on EEG source localization *Brain Topogr.* **27** 95–111
- [31] Rimpiläinen V, Koulouri A, Lucka F, Kaipio J P and Wolters C H 2019 Improved EEG source localization with bayesian uncertainty modelling of unknown skull conductivity *NeuroImage* **188** 252–60
- [32] Vorwerk J, Aydin U, Wolters C H and Butson C R 2019 Influence of head tissue conductivity uncertainties on EEG dipole reconstruction *Front. Neurosci.* **13** 531
- [33] Lew S, Wolters C H, Anwander A, Makeig S and MacLeod R S 2009 Improved EEG source analysis using low-resolution conductivity estimation in a four-compartment finite element head model *Hum. Brain Mapp.* **30** 2862–78
- [34] Papageorgakis C 2017 Patient specific conductivity models: characterization of the skull bones *PhD Thesis* Université Côte d’Azur, France
- [35] Fuchs M, Wagner M, Wischmann H A, Köhler T, Theissen A, Drenckhahn R and Buchner H 1998 Improving source reconstructions by combining bioelectric and biomagnetic data *Electroencephalogr. Clin. Neurophysiol.* **107** 93–111
- [36] Huang M X *et al* 2007 A novel integrated MEG and EEG analysis method for dipolar sources *NeuroImage* **37** 731–48
- [37] Aydin U *et al* 2014 Combining EEG and MEG for the reconstruction of epileptic activity using a calibrated realistic volume conductor model *PLoS One* **9** e93154
- [38] Antonakakis M, Schrader S, Aydin U, Khan A, Gross J, Zervakis M, Rampp S and Wolters C H 2020 Inter-subject variability of skull conductivity and thickness in calibrated realistic head models *NeuroImage* **223** 117353
- [39] Schmidt C, Wagner S, Burger M, Rienen U V and Wolters C H 2015 Impact of uncertain head tissue conductivity in the optimization of transcranial direct current stimulation for an auditory target *J. Neural Eng.* **12** 046028
- [40] Saturnino G B, Thielscher A, Madsen K H, Knösche T R and Weise K 2019 A principled approach to conductivity uncertainty analysis in electric field calculations *NeuroImage* **188** 821–34
- [41] McCann H and Beltrachini L 2021 Does participant’s age impact on tdcS induced fields? Insights from computational simulations *Biomed. Phys. Eng. Express* **7** 045018
- [42] Kaipio J and Kolehmainen V 2013 Approximate marginalization over modeling errors and uncertainties in inverse problems *Bayesian Theory and Applications* ed P Damien, N Polson and D Stephens (Oxford University Press)
- [43] Nissinen A, Kolehmainen V and Kaipio J P 2011 Reconstruction of domain boundary and conductivity in electrical impedance tomography using the approximation error approach *Int. J. Uncertainty Quantif.* **1** 203–22
- [44] Koulouri A and Rimpiläinen V 2020 Simultaneous skull conductivity and focal source imaging from EEG recordings with the help of Bayesian uncertainty modelling *8th European Medical and Biological Engineering Conf.* (Springer) pp 1019–27
- [45] Lipponen A, Seppänen A and Kaipio J P 2011 Nonstationary approximation error approach to imaging of three-dimensional pipe flow: experimental evaluation *Meas. Sci. Technol.* **22** 104013
- [46] Nissinen A, Heikkinen L M, Kolehmainen V and Kaipio J P 2009 Compensation of errors due to discretization, domain truncation and unknown contact impedances in electrical impedance tomography *Meas. Sci. Technol.* **20** 105504
- [47] Arridge S R, Kaipio J P, Kolehmainen V, Schweiger M, Somersalo E, Tarvainen T and Vauhkonen M 2006 Approximation errors and model reduction with an application in optical diffusion tomography *Inverse Problems* **22** 175–95
- [48] Kolehmainen V, Schweiger M, Nissilä I, Tarvainen T, Arridge S R and Kaipio J P 2009 Approximation errors and model reduction in three-dimensional diffuse optical tomography *J. Opt. Soc. Am. A* **26** 2257–68
- [49] Tarvainen T, Kolehmainen V, Pulkkinen A, Vauhkonen M, Schweiger M, Arridge S R and Kaipio J P 2010 An approximation error approach for compensating for modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography *Inverse Problems* **26** 015005
- [50] Koulouri A, Rimpiläinen V, Brookes M and Kaipio J P 2016 Compensation of domain modelling errors in the inverse source problem of the Poisson equation: Application in electroencephalographic imaging *Appl. Num. Math.* **106** 24–36
- [51] Mozumder M, Tarvainen T, Arridge S R, Kaipio J and Kolehmainen V 2013 Compensation of optode sensitivity and position errors in diffuse optical tomography using the approximation error approach *Biomed. Opt. Express* **4** 2015–31
- [52] Alexanderian A, Nicholson R and Petra N 2024 Optimal design of large-scale nonlinear bayesian inverse problems under model uncertainty *Inverse Problems* **40** 095001
- [53] Candiani V, Hyvönen N, Kaipio J P and Kolehmainen V 2021 Approximation error method for imaging the human head by electrical impedance tomography *Inverse Problems* **37** 125008
- [54] Nicholson R, Petra N, Villa U and Kaipio J P 2023 On global normal linear approximations for nonlinear bayesian inverse problems *Inverse Problems* **39** 054001
- [55] Bohrnstedt G W and Goldberger A S 1969 On the exact covariance of products of random variables *J. Am. Stat. Assoc.* **64** 1439–42
- [56] Grech R, Cassar T, Muscat J, Camilleri K, Fabri S, Zervakis M, Xanthopoulos P, Sakkalis V and Vanrumste B 2008 Review on solving the inverse problem in EEG source analysis *J. Neuroeng. Rehabil.* **5** 25
- [57] Michel C M and Brunet D 2019 EEG source imaging: A practical review of the analysis steps *Front. Neurol.* **10** 325
- [58] Rodriguez-Rivera A, Van Veen B and Wakai R 2003 Statistical performance analysis of signal variance-based dipole models for MEG/EEG source localization and detection *IEEE Trans. Biomed. Eng.* **50** 137–49
- [59] Höltershinken M B, Erdbrügger T and Wolters C H 2024 sLORETA is equivalent to single dipole scanning (arXiv:2409.07459)
- [60] Knösche T 1997 Solutions of the neuroelectromagnetic inverse problem: an evaluation study *PhD Thesis* University of Twente, Netherlands
- [61] Vauhkonen M 1997 Electrical impedance tomography and prior information *PhD Thesis* University of Kuopio, Finland

- [62] Puonti O, Van Leemput K, Saturnino G B, Siebner H R, Madsen K H and Thielscher A 2020 Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling *NeuroImage* **219** 117044
- [63] Wolters C H, Grasedyck L and Hackbusch W 2004 Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem *Inverse Problems* **20** 1099–116
- [64] Hoekema R, Wieneke G H, van Veelen C W, van Rijen P C, Huiskamp G J, Ansems J and van Huffelen A C 2003 Measurement of the conductivity of skull, temporarily removed during epilepsy surgery *Brain Topogr.* **16** 29–38
- [65] Homma S, Musha T, Nakajima Y, Okomoto Y, Blom S, Flink R and Hagbarth K E 1995 Conductivity ratios of the scalp-skull-brain head model in estimating equivalent dipole sources in human brain *Neurosci. Res.* **22** 51–55
- [66] Ramon C, Schimpf P H and Haueisen J 2006 Influence of head models on EEG simulations and inverse source localizations *Biomed. Eng. Online* **5** 10
- [67] Badia A E and Duong T H 1998 Some remarks on the problem of source identification from boundary measurements *Inverse Problems* **14** 883–91
- [68] Köhler T, Wagner M, Fuchs M, Wischmann H A, Drenckhahn R and Theissen A 1996 Depth normalization in MEG/EEG current density imaging *Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. (Proc. 18th Annual Int. Conf. IEEE vol 2)* pp 812–3
- [69] Pascual-Marqui R D, Michel C M and Lehmann D 1994 Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain *Int. J. Psychophysiol.* **18** 49–65
- [70] Pascual-Marqui R D 2002 Standardized low resolution brain electromagnetic tomography (sLORETA): technical details *Methods Find Exp. Clin. Pharmacol.* **24** 5–12
- [71] Fuchs M, Wagner M and Wischmann H A 1999 Linear and Nonlinear Current Density Reconstructions *J. Clin. Neurophysiol.* **16** 267–95
- [72] Wagner M, Wischmann H A, Fuchs M, Köhler T and Drenckhahn R 2000 *Current Density Reconstructions Using the L1 Norm* (Springer) pp 393–6
- [73] Palmero-Soler E, Dolan K, Hadamschek V and Tass P A 2007 swLORETA: a novel approach to robust source localization and synchronization tomography *Phys. Med. Biol.* **52** 1783–800
- [74] Buchner H, Knoll G, Fuchs M, Rienäcker A, Beckmann R, Wagner M, Silny J and Pesch J 1997 Inverse localization of electric dipole current sources in finite element models of the human head *Electroencephalogr. Clin. Neurophysiol.* **102** 267–78
- [75] Gramfort A, Luessi M, Larson E, Engemann D A, Strohmeier D, Brodbeck C, Parkkonen L and Hämäläinen M S 2014 MNE software for processing MEG and EEG data *NeuroImage* **86** 446–60
- [76] Koulouri A, Rimpiläinen V, Brookes M and Kaipio J P 2018 *Prior Variances and Depth un-Biased Estimators in EEG Focal Source Imaging Embeb & NBC 2017* (Springer) pp 33–36
- [77] Calvetti D, Pascarella A, Pitolli F, Somersalo E and Vantaggi B 2018 Brain activity mapping from meg data via a hierarchical bayesian algorithm with automatic depth weighting *Brain Topogr.* **32** 363–93