PDE-constrained Optimization for Electroencephalographic Source Reconstruction

Article *in* Lobachevskii Journal of Mathematics · May 2024

DOI: 10.1134/S1995080224603266

5 authors, including:

Mikhail Malovichko
Aramco Innovations
40 PUBLICATIONS **170** CITATIONS
SEE PROFILE

Alexandra Razorenova
Skolkovo Institute of Science and Technology
17 PUBLICATIONS **46** CITATIONS
SEE PROFILE

Vasily Golubev
Moscow Institute of Physics and Technology
131 PUBLICATIONS **764** CITATIONS
SEE PROFILE

PDE-constrained Optimization for Electroencephalographic Source Reconstruction

M. S. Malovichko,[1,*] **N. B. Yavich,**[2,**] **A. M. Razorenova,**[2,***]
V. I. Golubev,[1,****] **and N. A. Koshev**[2,3,*****]

(Submitted by A. V. Lapin)

[1] *Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141701 Russia*
[2] *Skolkovo Institute of Science and Technology, Moscow, 121205 Russia*
[3] *Limited liability company "Life improvement by future technologies center", Moscow, 121205 Russia*

Received April 24, 2024; revised May 8, 2024; accepted May 14, 2024

Abstract—This paper introduces a novel numerical method for the inverse problem of electroencephalography (EEG). We pose the inverse EEG problem as an optimal control (OC) problem for Poisson's equation. The optimality conditions lead to a variational system of differential equations. It is discretized directly in finite-element spaces leading to a system of linear equations with a sparse Karush–Kuhn–Tucker matrix. The method uses finite-element discretization and thus can handle MRI-based meshes of almost arbitrary complexity. It extends the well-known mixed quasi-reversibility method (mQRM) in that pointwise noisy data explicitly appear in the formulation making unnecessary tedious interpolation of the noisy data from the electrodes to the scalp surface. The resulting algebraic problem differs considerably from that obtained in the mixed quasi-reversibility, but only slightly larger. The algorithm does not require the formation of the lead-field matrix, which can be beneficial for large matrices. Our tests, both with spherical and MRI-based meshes, demonstrates that the method accurately reconstructs cortical activity.

2010 Mathematical Subject Classification: 86-08, 86A15, 74L05

Keywords and phrases: *optimal control, inverse problems, electroencephalography*

1. INTRODUCTION

Electroencephalography (EEG) has a long history of development and constitute one of the major modalities to study living human brain, see reviews [1–7]. For reliable localization of cortical activations, it is crucial to solving the inverse EEG problem, that is, to map electric potential measured at electrodes to either potential or electric current on the cortex.

Mathematically, this problem is the source reconstruction for Poisson's equation. Depending on the formulation, it can also be considered as the Cauchy problem. The mathematical properties of these problems have been extensively studied, and many approaches have been proposed. We refer to monographs [8–11] as well research papers [12–17] (among many others). Because it is impossible to review all methods proposed for the task, we confine ourselves to those that found applications in the EEG context.

The linear distributed estimators constitute arguably the most popular approach. They are used in practice with great success, e.g., [18–20]. Algorithms in this category, such as MNE-family [20, 21], LORETA-family [22, 23], LAURA [24], FOCUSS [25], WROP [26], and many others, are closely

* E-mail: malovichko.mikhail@gmail.com
** E-mail: n.yavich@skoltech.ru
*** E-mail: alexandra.razorenova@skoltech.ru
**** E-mail: golubev.vi@mipt.ru
***** E-mail: n.koshev@skoltech.ru

connected to the Tikhonov regularization approach. The forward operator is usually based on the boundary-element method (BEM), but the finite-element method (FEM) is catching up; see [27–32].

Methods based on solving the Cauchy problem for Poisson's equation propagate the electric potential and, sometimes, Ohmic current inward the head starting from known values of potential and its normal derivative on the scalp. These methods date back at least to the 1970s, under different names, such as *the current source density*, *surface Laplacian*, *deblurring*, *spatial deconvolution* and others [33–39]. They solve the Cauchy problem for Laplace's or Poisson's equations making severe simplifications to analytically link the potential and its normal derivatives on the scalp with that inside the head. These early techniques have been considered as data-enhancing procedures rather than rigorous solutions to an inverse problem. Today, they are still in use, sometimes as a part of more sophisticated algorithms [40].

The deblurring method by [41–43] was one of the first methods to propagate potential inward an anatomical head model. Remarkably, it employed a FEM on tetrahedral grids, although the BEM would dominate the field for decades. Numerical procedures for potential propagation based on the BEM have been proposed in [44–46]. These approaches are appealing, although they have some limitations rooted in the BEM method, such as the inability to handle surface holes, and a high compute load for large grids.

Thus, despite the variety of mathematical approaches, the majority of source reconstruction methods currently used in EEG community are based on the construction and solution of a normal system of equations with the explicit use of the inverse operator. A remarkable dispatch from this pattern has been made by L. Bourgeois, who proposed the mixed quasi-reversibility method (QRM) in a purely mathematical context [47, 48]. The mixed QRM reduces the Cauchy problem to a variational system. The system is discretized directly in finite-element function spaces. Thus, reconstruction of the potential on the unaccessible boundary (cortex) requires a single linear solve. The FE discretization makes the mixed QRM very flexible, allowing computational domains of almost arbitrary complexity. This approach has been advanced further to account for the non-constant conductivity and adjusted to the EEG source localization in [49, 50].

The mixed QRM has some drawbacks. The main practical disadvantage is the need to convert pointwise measurements on electrodes to a boundary value of potential on the scalp. Accurate interpolation of noisy data to an irregular surface (scalp) is a tedious task of its own. Even more important is that it is unclear how to take proper account of noise characterization, which is typically available by a noise covariance matrix, estimated for a given electrode layout.

In this paper, we study the inverse EEG problem from the optimal control (OC) standpoint. It is a well-established technique for inverse problems associated with PDE [51], but it is rarely applied for the EEG/MEG source analysis. Elements of the OC were used in [52] as a tool to justify the calculation of the lead-field matrix through the adjoint-state technique. In [53], the inverse EEG problem was formulated as a problem of finding the saddle point of the Lagrangian. Those authors did not form the Euler–Lagrange system but approached the saddle point making alternating steps in the state and adjoint-state spaces; thus, this method never constituted an efficient algorithm.

Here we formulate the inverse problem of EEG as an optimization problem with constraints, thus reducing it to the problem of finding an optimal point of Lagrangian. We formulate the Euler–Lagrange variational system and approximate it in the finite-element function spaces. Based on this approach, we proposed a new numerical method for the inverse problem of EEG, which uses pointwise noisy data. This method addresses the critical issue of noise interpolation and, in this respect, can be considered as an "improved" mixed QRM. We assume that the electrical signal is caused by a dipole layer on the surface of the brain, so in our approach eliminates the brain compartment from consideration, making the resulting mesh smaller and avoiding errors associated with interpolating a dipolar right-hand size in finite-element spaces. However, the method can be extended to handle arbitrary oriented dipoles (Section 5, Remark 3). An interesting feature of this approach is that the lead-field matrix is never formed explicitly (Section 5, Remark 2).

The paper is organized as follows. In Section 2, we review the mixed quasi-reversibility method. Section 3 is dedicated to the optimal-control formulation and derivation of the Euler–Lagrange variational system. In Section 4, we verify the derived variational system. The finite-element discretization, the algebraic problem, and the solution of the resulting system of linear equations are considered in Section 5. The FEM convergence is discussed in Section 6. A numerical experiment aimed to verify our numerical method is presented in Section 7. Concluding remarks are given in Section 8.

2. THE MIXED QUASI-REVERSIBILITY METHOD

Here we review the mixed QRM. The QRM can be traced back to [54]. Its mixed formulation within the finite-element framework has been proposed in [47, 48]. It was advanced further in the context of EEG source analysis in [49, 50]. We derive the algorithm with OC formalism. This derivation is similar to one presented in [50], but is adjusted toward the formal Lagrangian approach, used throughout the rest of this paper.

Let $\Omega \in \mathbb{R}^3$ be a part of the head enclosed by the scalp surface, Γ_0, and the cortical surface, Γ_B, that is, $\partial\Omega = \Gamma_0 \cup \Gamma_B$ (Fig. 1). Electric potential on the scalp has zero normal derivative, $\partial u/\partial\nu = 0$. We assume that the potential inside the head is driven by a dipole layer attached to the cortical surface (see Section 5, Remark 3). Let us denote by f the unknown surface density of dipoles.

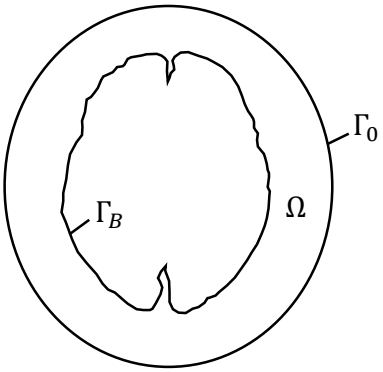


Figure 1. Schematic representation of the head. Γ_0 is the scalp surface, Γ_B is the cortex (can consist of disjoint surfaces, e.g., the two hemispheres). Volume Ω includes different tissues (fat, skull, cerebrospinal fluid, etc.).

Thus, the potential in the head satisfies the following differential equation,

$$-\mathrm{div}(\sigma\nabla u) = 0 \text{ in } \Omega, \tag{1}$$

$$\frac{\partial u}{\partial \nu} = 0 \text{ on } \Gamma_0, \tag{2}$$

$$\sigma\frac{\partial u}{\partial \nu} = f \text{ on } \Gamma_B. \tag{3}$$

The head consists of several compartments of different conductivities. The conductivity is usually considered smooth and bounded inside each compartment with jumps across interfaces

$$0 < \sigma_{min} \le \sigma \le \sigma_{max} < \infty.$$

At the interfaces, the potential is continuous, but the normal Ohmic current experiences jumps,

$$u_i = u_j, \quad \sigma_i\nabla u_i = \sigma_j\nabla u_j \text{ at } \Gamma_{ij}, \tag{4}$$

where Γ_{ij} is the interface between a pair of compartments. To avoid theoretical complications, we assume that the boundaries are smooth, which seems reasonable for EEG studies.

The problem (1)–(3) is understood in the weak sense. Using standard regularity results such as in [55] [Proposition 2.1], we specify $u \in H^1(\Omega)$. Also, note that $\sigma\nabla u \in H_{div}(\Omega)$. Here we used standard notation: $H^1(\Omega)$ is a space of functions having weak square-integrable gradient; $H_{div}(\Omega)$ is a space of functions having weak square-integrable divergence. Thus, the interface conditions (4) not need to be added explicitly to (1)–(3), because they are satisfied by this choice of function spaces.

In the experiments, the potential values u can be measured at some points on the Γ_0, instead of f on Γ_B. So, the problem can be reformulated as follows. We want to reconstruct potential u in the head (including that on the cortex) such that (a) u satisfies Poisson equation, (b) u is smooth almost everywhere, and (c) $u|_{\Gamma_0}$ is equal to some given value g_D. Consider the following inverse problem

$$\underset{u}{\text{minimize}} \quad \frac{\varepsilon}{2}\int_\Omega |\nabla u|^2\,dV, \tag{5}$$

$$-\text{div}(\sigma\nabla u) = 0 \quad \text{in } \Omega, \tag{6}$$

$$\frac{\partial u}{\partial \nu} = 0 \text{ on } \Gamma_0, \tag{7}$$

$$u = g_D \text{ on } \Gamma_0. \tag{8}$$

In (5), $\varepsilon > 0$ is the regularization parameter. Further, the unknown value of f can be calculated as $f = \sigma\frac{\partial u}{\partial \nu}$.

We form the Lagrangian that includes condition (6),

$$\mathcal{L}(u,\lambda) = \frac{\varepsilon}{2}\int_\Omega |\nabla u|^2\,dV - \int_\Omega \lambda\,\text{div}(\sigma\nabla u)\,dV, \tag{9}$$

where λ is the Lagrange multiplier. By applying the Green–Gauss theorem to the second term of (9) and using condition (7), we obtain the following equality

$$\int_\Omega \lambda\,\text{div}(\sigma\nabla u)\,dV = -\int_\Omega \sigma\nabla u\cdot\nabla\lambda\,dV + \int_{\Gamma_B}\sigma\frac{\partial u}{\partial\nu}\lambda\,dV. \tag{10}$$

Let us choose λ such that its boundary values on Γ_B is zero, so, the boundary integral in (10) vanishes:

$$\lambda \in Q := \{v \in H^1(\Omega) : v|_{\Gamma_B} = 0\}. \tag{11}$$

Now, we choose u such that condition (8) is satisfied automatically by this choice

$$u \in P := \{v \in H^1(\Omega) : v|_{\Gamma_0} = g_D\}. \tag{12}$$

Now, we substitute (10) to (9), using definitions (11) and (12), and get

$$\mathcal{L}(u,\lambda) = \frac{\varepsilon}{2}\int_\Omega |\nabla u|^2\,dV + \int_\Omega \sigma\nabla u\cdot\nabla\lambda\,dV. \tag{13}$$

The Lagrangian (13) attains its optimal point whenever $\mathcal{L}'_u = 0$ and $\mathcal{L}'_\lambda = 0$. Thus, we have the following variational problem: find $(u,\lambda) \in P\times Q$, such that

$$\varepsilon\int_\Omega \nabla u\cdot\nabla v\,dV + \int_\Omega \sigma\nabla\lambda\cdot\nabla v\,dV = 0, \quad \forall v \in P_0, \tag{14}$$

$$\int_\Omega \sigma\nabla u\cdot\nabla\mu\,dV = 0, \quad \forall \mu \in Q. \tag{15}$$

Function space P_0 is specified as a set of all elements from $H^1(\Omega)$ having zero boundary value on Γ_0,

$$P_0 := \{v \in H^1(\Omega) : v|_{\Gamma_0} = 0\}.$$

This problem needs to be stabilized (see [47], Remark 3). We add term $-\delta\int_\Omega |\nabla\lambda|^2 dV$ to (13), where $\delta > 0$. The modified variational problem reads: find $(u,\lambda) \in P\times Q$ such that

$$\varepsilon\int_\Omega \nabla u\cdot\nabla v\,dV + \int_\Omega \sigma\nabla\lambda\cdot\nabla v\,dV = 0, \quad \forall v \in P_0, \tag{16}$$

$$\int_\Omega \sigma\nabla u\cdot\nabla\mu\,dV - \delta\int_\Omega \nabla\lambda\cdot\nabla\mu\,dV = 0, \quad \forall \mu \in Q. \tag{17}$$

This is the mixed formulation of the quasi reversibility. Note that, instead of $\int_\Omega |\nabla u|^2 dV$ and $\int_\Omega |\nabla \lambda|^2 dV$ we could have used $\int_\Omega (|u|^2 + |\nabla u|)^2 dV$ and $\int_\Omega (|\lambda|^2 + |\nabla \lambda|)^2 dV$, respectively. This way we would have variational system (19)–(21) from [50].

System (16) and (17) was rigorously studied in [47, 48]. Its numerical discretization and validation in the EEG context were presented in [50].

Note that this algorithm requires as input the boundary value of u on Γ_0, i.e., g_D. In practice, we have discrete data: potential is measured at a small set of electrodes. Interpolation of discrete data onto a whole surface of the grid produces considerable correlated noise. More importantly, it is not clear how to incorporate the noise characterization, typically represented by a noise covariance matrix, into formulation (16) and (17). Finally, formulation (16) and (17) imposes a constraint on the volume distribution of u, although researchers prefer to control the cortical distribution of the potential.

3. THE PDE-CONSTRAINED OPTIMIZATION

In this section, we apply the optimal control formalism to the EEG source reconstruction problem.

The electrical potential is measured in a set of electrodes. Each data point d_i is a value of potential u at the ith electrode. Formally, d_i is the convolution of u with the Dirac δ function

$$d_i = \mathcal{Q}_i(u) := \int_\Omega \delta(x - x_i) u(x)\, dV, \tag{18}$$

where $x_i \in \mathbb{R}^3$ is the position of the ith electrode and $\mathcal{Q}_i$ is the observation operator. Let $d \in \mathbb{R}^K$ be a vector of all data points.

We introduce a diagonal matrix of weights, W, representing estimates of noise standard deviation. Each data point d_i is associated with a weight $w_i > 0$. By assuming that W is diagonal, that is, the noise is uncorrelated, we aim to simplify the formulas below. A dense noise covariance matrix can be readily used here (Section 5, Remark 1).

We set up the following optimization problem

$$\underset{f,u}{\text{minimize}} \quad \Psi(u, f) := \frac{1}{2} \sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i u - d_i)^2 + \frac{\varepsilon}{2} \int_{\Gamma_B} |\nabla_B f|^2\, dS, \tag{19}$$

subject to

$$-\text{div}(\sigma \nabla u) = 0 \quad \text{in } \Omega, \tag{20}$$

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \Gamma_0, \sigma \frac{\partial u}{\partial \nu} = f \quad \text{on} \quad \Gamma_B. \tag{21}$$

The first term in (19) is the misfit between measured data, d_i, and computed data $Q_i u$. The second term is the stabilizing functional with $\varepsilon > 0$ is the regularization parameter. Here ∇_B is the surface gradient defined on Γ_B. Thus, we require f be a weakly smooth function on Γ_B. Constraints (20) and (21) determine the link between the electric potential u and the source function f.

Here we outline the functional framework for u and f. Since data d are "pointwise", the integrals (18) must exist for $\delta \in H^s, s < -3/2$ and any admissible state u. Let us specify the space of admissible states as $U_{ad} := \{u \in H^1(\Omega) \cap \mathcal{C}(\bar{\Omega})\}$. We also specify the space of admissible controls f representing normal currents on cortex (see Section 5, Remark 3)

$$F := \{f \in H^1(\Gamma_B) : f \neq const\}.$$

Next, we derive the variational system for problem (19)–(21). The derivation is based on the Lagrangian formalism aiming to design the proper form of the Euler–Lagrange system. We form the Lagrangian $\mathcal{L}$ as follows

$$\mathcal{L}(u, f, \lambda) = \Psi(u, f) - \int_\Omega \text{div}(\sigma \nabla u)\, \lambda\, dV + \int_{\Gamma_B} \left(\sigma \frac{\partial u}{\partial \nu} - f \right) \lambda\, dS, \tag{22}$$

where λ is the Lagrange multiplier, defined in $\Omega \cup \Gamma_B$. Next, we apply the Green–Gauss theorem to the second term of (22) and use condition $\partial u/\partial \nu = 0$ on Γ_0:

$$\int_\Omega \operatorname{div}(\sigma \nabla u)\, \lambda \, dV = -\int_\Omega \sigma \nabla u \cdot \nabla \lambda \, dV + \int_{\Gamma_B} \sigma \frac{\partial u}{\partial \nu} \lambda \, dS.$$

We have

$$\mathcal{L}(u, f, \lambda) = \Psi(u, f) + \int_\Omega \sigma \nabla u \cdot \nabla \lambda \, dV - \int_{\Gamma_B} f \lambda \, dS, \tag{23}$$

Thus, optimization with constraints (19)–(21) translates to the problem of finding a stationary point of the Lagrangian (23). The necessary conditions for a triple (u, f, λ) be a stationary point are that the derivatives of $\mathcal{L}(u, f, \lambda)$ with respect to u, f, and λ vanish. Since Ψ is convex, the necessary conditions are also sufficient. Thus, we set up the following variational problem: find triple $(f, \lambda, u) \in F \times \Lambda \times U$ which verifies

$$\varepsilon \int_{\Gamma_B} \nabla_B f \cdot \nabla_B \varphi \, dS - \int_{\Gamma_B} \lambda \varphi \, dS = 0, \quad \forall \varphi \in F, \tag{24}$$

$$-\int_{\Gamma_B} f \mu \, dS + \int_\Omega \sigma \nabla u \cdot \nabla \mu \, dV = 0, \quad \forall \mu \in \Lambda, \tag{25}$$

$$\int_\Omega \sigma \nabla \lambda \cdot \nabla v \, dV + \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(u) \mathcal{Q}_i(v) = \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(v) d_i, \quad \forall v \in U, \tag{26}$$

where

$$U := \{v \in U_{ad} : \frac{\partial v}{\partial \nu} = 0 \text{ on } \Gamma_0\}, \quad \Lambda = \{\mu \in H^1(\Omega) : \frac{\partial \mu}{\partial \nu} = 0 \text{ on } \partial\Omega\}.$$

Let us rewrite (24)–(26) as follows: find triple $(f, \lambda, u) \in F \times \Lambda \times U$ which verifies

$$a(f, \varphi) - b^t(\lambda, \varphi) = 0, \quad \forall \varphi \in F, \tag{27}$$

$$-b(f, \mu) + e^t(u, \mu) = 0, \quad \forall \mu \in \Lambda, \tag{28}$$

$$e(\lambda, v) + g(u, v) = R(v), \quad \forall v \in U, \tag{29}$$

where superscript ^{t} stands for transposition, bilinear forms $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $e(\cdot,\cdot)$, $g(\cdot,\cdot)$, and a linear form $R(\cdot)$ are defined as follows

$$a(f, \varphi) = \varepsilon \int_{\Gamma_B} \nabla_B f \cdot \nabla_B \varphi \, dS, \quad b(f, \mu) = \int_{\Gamma_B} f \mu \, dS, \quad e(\lambda, v) = \int_\Omega \sigma \nabla \lambda \cdot \nabla v \, dV,$$

$$g(u, v) = \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(u) \mathcal{Q}_i(v), \quad R(v) = \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(v) d_i.$$

4. SOLVABILITY

In this section, we demonstrate that problem (19)–(21) admits a unique solution. Since the formal Lagrangian method does not constitute a rigorous proof, we prove that (24)–(26) and (19)–(21) are equivalent.

First, let us prove that problem (19)–(21) has a unique optimal control f. The general technique of proving the existence of the optimal control for a linear-quadratic problem is well established and based

on the fact that F is weak lower semi-continuous. Thus, we will check the required conditions. Both U_{ad} and F are convex and closed. We write problem (19)–(21) as follows

$$\underset{f\in F, u\in U_{ad}}{\text{minimize}} \quad \Psi(f,u) := \frac{1}{2}\left\|\mathcal{Q}u - d\right\|_2^2 + \frac{\varepsilon}{2}\left\|\nabla_B f\right\|^2_{L^2(\Gamma_B)}, \tag{30}$$

$$-\mathrm{div}(\sigma\nabla u) = \xi \quad \text{in } \Omega, \tag{31}$$

$$\sigma\frac{\partial u}{\partial \nu} = f \quad \text{on } \Gamma_B, \tag{32}$$

$$\frac{\partial u}{\partial \nu} = g_N \quad \text{on } \Gamma_0. \tag{33}$$

Here operator $\mathcal{Q}: \mathcal{C}(\Omega) \to \mathbb{R}^K$, $\mathcal{Q}v := \{\mathcal{Q}_0(v), .., \mathcal{Q}_{K-1}(v)\}$. In what follows we will use its adjoint $\mathcal{Q}^*: \mathbb{R}^K \to \mathcal{C}(\Omega)^*$, which is $\mathcal{Q}^* z := \sum_{j=0}^{K-1} z_j \delta_j$. Functions g_N and ξ are given (precisely, $\xi = 0$, $g_N = 0$). If f is given as well, than the forward problem (30)–(33) is well-posed. Thus, we define the control-to-target operator, $\mathcal{G} := f \mapsto d : H^1(\Gamma_B) \to \mathbb{R}^K$. Operator $\mathcal{G}$ is linear. Problem (30)–(33) is equivalent to the reduced problem

$$\underset{f\in F}{\text{minimize}} \quad \Phi(f) := \frac{1}{2}\left\|\mathcal{G}f - d\right\|_2^2 + \frac{\varepsilon}{2}\left\|\nabla_B f\right\|^2_{L^2(\Gamma_B)}. \tag{34}$$

The cost function Φ is convex. Thus, we can formulate the following result.

Proposition 1. *For any $\varepsilon > 0$ problem (30)–(33) admits the optimal control $\bar{f}$, and optimal state $\bar{u}$, which are unique.*

Proof. [51][Thm. 1.43]. □

Let us show that problems (19)–(21) and (24)–(26) are equivalent.

Proposition 2. *Control f and state u are the minimizers of (19)–(21) if and only if there exists λ such that (u, f, λ) solve variational system (24)–(26).*

Proof. We will assume $w_i^2 = 1$ (solely to render formulas more compact).

" $\Rightarrow$ ". Let (f, u, λ) be the stationary point of (19)–(21). Let $T : H^1(\Omega) \to L^2(\Gamma_B)$ be the trace operator,

$$Tu := u|_{\Gamma_B}, \quad v \in H^1(\Omega).$$

Its adjoint $T^* : L^2(\Gamma_B) \to H^1(\Omega)^*$ is defined by relation

$$\langle T^* v, u\rangle_{H^1(\Omega)^*, H^1(\Omega)} = \langle v, Tu\rangle_{L^2(\Gamma_B), L^2(\Gamma_B)} = \int_{\Gamma_B} v\, Tu\, dS, \quad v \in L^2(\Gamma_B), u \in H^1(\Omega).$$

From (25) we see that $u \in H^1(\Omega)$ is such that $\partial u/\partial\nu|_{\Gamma_0} = 0$ and

$$\int_\Omega \sigma\nabla u \cdot \nabla\mu\, dV = \int_{\Gamma_B} f\, T\mu\, dS, \quad \forall \mu \in H^1(\Omega).$$

The right-hand side is $\langle T\mu, f\rangle_{L^2(\Gamma_B), L^2(\Gamma_B)} = \langle \mu, T^* f\rangle_{H^1(\Omega), H^1(\Omega)^*}$. Thus, we can rewrite it as $u = ST^* f$, where $S : H^1(\Omega)^* \to H^1(\Omega)$. From (26) it follows that $\lambda \in H^1(\Omega)$, $\partial\lambda/\partial\nu|_{\partial\Omega} = 0$ and

$$\int_\Omega \sigma\nabla\lambda \cdot \nabla v\, dV = \langle d - \mathcal{Q}(u), \mathcal{Q}v\rangle_{\mathbb{R}^K, \mathbb{R}^K}.$$

Since $\langle d - \mathcal{Q}(u), \mathcal{Q}v\rangle_{\mathbb{R}^K, \mathbb{R}^K} = \langle Q^*(d - \mathcal{Q}(u)), v\rangle_{H^1(\Omega)^*, H^1(\Omega)}$, we conclude that

$$\lambda = S^*\mathcal{Q}^*(Qu - d) = S^*\mathcal{Q}^*(QSTf - d),$$

where $S^* : H^1(\Omega)^* \to H^1(\Omega)$. Let us take (24),

$$-\int_{\Gamma_B} T\lambda\,\varphi\,dS + \varepsilon\int_{\Gamma_B}\nabla_B f\cdot\nabla_B\varphi\,dS = 0,\quad \forall\varphi\in H^1(\Gamma_B).$$

By substituting for λ, we get

$$-\int_{\Gamma_B} TS^*\mathcal{Q}^*(Qu-d)\varphi\,dS + \varepsilon\int_{\Gamma_B}\nabla_B f\cdot\nabla_B\varphi\,dS = 0,\quad \forall\varphi\in H^1(\Gamma_B). \tag{35}$$

Since $\nabla_f(Qu-d) = \nabla_f(QST^*f-d) = TS^*\mathcal{Q}^*$ we see that (35) means that the weak gradient of Φ with respect to f is zero. Thus, f minimizes Φ, and (f,u) minimizes (19)–(21).

" $\Leftarrow$ ". Let (u,f) be the minimizer of (19)–(21). It means that $u\in H^1(\Omega)$, $\partial u/\partial\nu|_{\Gamma_0}=0$, and

$$\int_\Omega \nabla u\cdot\nabla v\,dV = \int_{\Gamma_B} fTv\,dS,\quad \forall v\in H^1(\Omega),\ \partial v/\partial\nu|_{\Gamma_0}=0. \tag{36}$$

Thus, we recovered equation (25).

Let us introduce operators S and T with the same definitions that before. Equation (36) can be rewritten as $u = ST^*f$. Thus, (19)–(21) is equivalent to the reduced problem in the form

$$\underset{f}{\text{minimize}}\quad \Phi(f) := \frac{1}{2}\left\|\mathcal{Q}ST^*f-d\right\|_2^2 + \frac{\varepsilon}{2}\left\|\nabla_B f\right\|^2_{L^2(\Gamma_B)}, \tag{37}$$

where $f\in F$. Function f minimizes (37) and so it verifies the following equality

$$\int_{\Gamma_B} TS^*\mathcal{Q}^*(\mathcal{Q}ST^*f-d)\varphi\,dS + \varepsilon\int_{\Gamma_B}\nabla_B f\cdot\nabla_B\varphi\,dS = 0,\quad \forall\varphi\in F. \tag{38}$$

We define

$$\lambda := -S^*\mathcal{Q}^*(\mathcal{Q}u-d), \tag{39}$$

$\lambda\in H^1(\Omega)$, $\partial\lambda/\partial\nu|_{\partial\Omega}=0$. Thus, (38) turns into

$$-\int_{\Gamma_B} T\lambda\varphi\,dS + \varepsilon\int_{\Gamma_B}\nabla_B f\cdot\nabla_B\varphi\,dS = 0,\quad \forall\varphi\in F,$$

and we recovered equation (24).

Finally, from (39) of follows that $(S^*)^{-1}\lambda = -\mathcal{Q}^*(\mathcal{Q}u-d)$, which is in the variational form reads

$$\int_\Omega\nabla\lambda\cdot\nabla v\,dV = -\int_\Omega\mathcal{Q}^*(Qu-d)v\,dV,\quad\forall v\in H^1(\Omega).$$

Since $\int_\Omega\mathcal{Q}^*(Qu-d)v\,dV = \sum_{j=0}^{K-1}(Q_jv-d_j)\,Q_jv$, we recover equation (26). □

5. THE FINITE-ELEMENT DISCRETIZATION

In this section, we formulate a discrete analog of (27)–(29) using FEM. Let us cover domain Ω with a set of tetrahedra, $\Omega_h = \bigcup_i\Omega_i$, where h stands for the maximal diameter. By $\mathcal{T}_h = \bigcup_i\mathcal{T}_i$, we will denote a set of faces of Ω_h on the cortical surface Γ_B. We will assume that partitions Ω_h and $\mathcal{T}_h$ are regular, i.e., their elements will not degenerate as h goes to zero.

By defining the finite-dimensional spaces $F_h\subset H^1(\Gamma_B)$, $U_h\subset H^1(\Omega)$,$\Lambda_h\subset H^1(\Omega)$, we arrive to the following finite-dimensional problem: find triple $(f_h,\lambda_h,u_h)\in F_h\times\Lambda_h\times U_h$ that verifies

$$a_h(f_h,\varphi_h) - b_h^t(\lambda_h,\varphi_h) = 0,\quad\forall\varphi_h\in F_h, \tag{40}$$

$$-b_h(f_h,\mu_h) + e^t(u_h,\mu_h) = 0,\quad\forall\mu_h\in\Lambda_h, \tag{41}$$

$$e(\lambda_h,v_h) + g_h(u_h,v_h) = r_h(v_h),\quad\forall v_h\in U_h, \tag{42}$$

Specifically, in the numerical experiments that follow, we define continuous piecewise-linear basis functions: $\{\alpha_i\}_{i=0}^{N-1}$ on tetrahedral mesh Ω_h and $\{\beta_i\}_{i=0}^{M-1}$ on triangular mesh $\mathcal{T}_h$. That is, we set up the following function spaces

$$U_h = \Lambda_h := \operatorname{span}\{\alpha_0, .., \alpha_{N-1}\}, \quad F_h := \operatorname{span}\{\beta_0, .., \beta_{M-1}\}.$$

The trial functions $u, \lambda \in H^1(\Omega)$ and $f \in H^1(\Gamma_B)$ are approximated by functions $u_h \in U_h$, $\lambda_h \in \Lambda_h$, and $f_h \in F_h$, respectively,

$$u_h = \sum_{i=0}^{N-1} u_i \alpha_i, \quad \lambda_h = \sum_{i=0}^{N-1} \lambda_i \alpha_i, \quad f_h = \sum_{i=0}^{M-1} f_i \beta_i,$$

where u_i, λ_i, and f_i are expansion coefficients (degrees of freedom, DOFs). The test functions v, μ, and φ are replaced by $v_h \in U_h$, $\mu_h \in \Lambda_h$, and $\varphi_h \in F_h$. Thus, infinite-dimensional problem (27)–(29) translates to the following algebraic problem

$$\underbrace{\begin{bmatrix} A & -B^T & O \\ -B & O & E^T \\ O & E & G \end{bmatrix}}_{\mathcal{M}} \underbrace{\begin{bmatrix} \tilde{f} \\ \tilde{\lambda} \\ \tilde{u} \end{bmatrix}}_{\xi} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}}_{b}, \tag{43}$$

where

$$A_{ij} = \varepsilon \int_{\mathcal{T}} \nabla_B \beta_i \cdot \nabla_B \beta_j \, dS, \quad B_{ij} = \int_{\mathcal{T}} \beta_i \alpha_j \, dS, \quad E_{ij} = \int_{\Omega} \sigma \nabla \alpha_i \cdot \nabla \alpha_j \, dV,$$

$G = Q^T W^T W Q$, $\quad r = Q^T W^T W d$, Q is the matrix of observation operator $\mathcal{Q}_h : U_h \to \mathbb{R}^K$ and O's are zero blocks. Vectors $\tilde{f}$, $\tilde{\lambda}$, and $\tilde{u}$ consist of DOFs f_i, λ_i, and u_i, respectively.

Blocks $A \in \mathbb{R}^{M\times M}$ and $E \in \mathbb{R}^{N\times N}$ are sparse positive-definite stiffness matrices. Matrix $B \in \mathbb{R}^{N\times M}$ is a sparse rectangular matrix of full column rank. Matrix $G \in \mathbb{R}^{N\times N}$ is positive semidefinite. Its rank is equal to the number of electrodes minus 1 (we assume that there are no duplicate measurements). System matrix $\mathcal{M}$ is sparse symmetric indefinite. Typically, $N >> M$, so the size of $\mathcal{M}$ is approximately $2N \times 2N$.

Note that a programming implementation of our method requires standard procedures available in existing FEM libraries. In particular, matrices A and E are nothing more than stiffness matrices; matrix B is (a part of) a mass matrix.

Remark 1. *In the presented form, G is sparse. Moreover, if measuring electrodes are located at the mesh nodes, then G will be diagonal. However, one can replace W^TW with a dense noise covariance matrix to account for correlated noise. In this case, block G will be dense.*

Remark 2. *The equivalence between problems (30)–(33) and* (34) *means that the lead-field matrix is implicitly contained in the saddle-point structure of the KKT system. It is especially clear from* (43) *and* (18)*: we see that predicted data d_{prd} are connected to the source $\tilde{f}$ as follows: $d_{prd} = QE^{-T}B\tilde{f}$. Matrix $QE^{-T}B$ is the lead-field matrix, but it is never formed explicitly.*

Remark 3. *Cortex dipoles orthogonal to the cortex represent the most widely used model of the source but not the only one. In specific scenarios, the signal may originate from deeper or shallow sources, and the source may not be orthogonal to the cortex. Extensive literature is devoted to the issues of accurate modeling of sources [27, 30–32], among many others.*

Still, we assume that sources are tied to the cortical cortex and are located on its surface. It makes it possible to identify the sources with the Neumann boundary conditions, approximate with an unknown function on the surface, and include expansion coefficients in the saddle point problem as unknowns. It reduces computations by excluding the brain interior from the problem.

Thus, our method specializes in problems in which the signal originates from cortical dipoles. If this assumption is not met, then an error will arise.

However, our method can be modified to estimate arbitrarily oriented dipoles (we thank the anonymous reviewer for pointing this out to us). As shown in the papers [56–58], one can approximate an arbitrary dipole by Raviart–Thomas basis functions and include unknown coefficients in the problem. The nodal basis functions and the Raviart–Thomas functions are naturally interconnected on the same mesh. In this case, the vector $\tilde{f}$ will contain unknown expansion coefficients, and block B will turn into matrix $\boldsymbol{G}$ from [57].

6. FEM CONVERGENCE

In this section, we demonstrate that finite-dimensional solution (u_h, f_h, λ_h) converges to the infinite-dimensional solution (u, f, λ). We introduce the linear operators $\mathbb{A}_h, \mathbb{B}_h, \mathbb{E}_h$, and $\mathbb{G}_h$ associated with the bilinear forms $a_h(\cdot,\cdot)$, $b_h(\cdot,\cdot)$, $e_h(\cdot,\cdot)$, and $g_h(\cdot,\cdot)$, respectively. By $\langle\cdot,\cdot\rangle$, we will denote the pairing between corresponding function spaces.

The following properties can be verified

$$\beta\,\|\lambda_h\| \le \left(\left\|-\mathbb{B}_h^T\lambda_h\right\|^2 + \|\mathbb{E}_h\lambda_h\|^2\right)^{1/2} \quad \text{(the inf-sup condition)},$$

$$\left(\left\|-\mathbb{B}_h^T\lambda_h\right\|^2 + \|\mathbb{E}_h\lambda_h\|^2\right)^{1/2} \le \beta_2\,\|\lambda\| \quad \text{(continuity)},$$

$$\alpha_1\left(\|f_h\|^2 + \|u_h\|^2\right)^{1/2} \le \langle\mathbb{A}_h f_h, f_h\rangle + \langle\mathbb{G}_h u_h, u_h\rangle \quad \text{(ellipticity)},$$

$$\left(\|\mathbb{A}_h f_h\|^2 + \|\mathbb{G}_h u_h\|^2\right)^{1/2} \le \alpha_2\left(\|f_h\|^2 + \|u_h\|^2\right)^{1/2} \quad \text{(continuity)}.$$

The following result bounds the norm of approximation error $(u-u_h, f-f_h, \lambda-\lambda_h)$ by the norm of the interpolation error. Here we use the method of proving the estimates developed in [59].

Proposition 3. *Let $(f,u,\lambda) \in F\times U\times\Lambda$ be a solution of (27)–(29), and $(f_h,u_h,\lambda)\in F_h\times U_h\times\Lambda_h$ is a solution to (40)–(42). Then, the misfit between the two solutions satisfies to the following estimates*

$$\left(\|f-f_h\|^2 + \|u-u_h\|^2\right)^{1/2} \le \left(1+\frac{\alpha_2}{\alpha_1}\right)\left(1+\frac{\beta_2}{\beta_1}\right)\inf_{\varphi_h,v_h\in F_h\times U_h}\left(\|f-\varphi_h\|^2 + \|u-v_h\|^2\right)^{1/2}. \tag{44}$$

$$\|\lambda-\lambda_h\| \le \left(1+\frac{\beta_2}{\beta_1}\right)\inf_{\mu_h\in\Lambda_h}\|\lambda-\mu_h\|$$
$$+\frac{\alpha_2}{\beta}\left(1+\frac{\alpha_2}{\alpha_1}\right)\left(1+\frac{\beta_2}{\beta_1}\right)\inf_{\varphi_h,v_h\in F_h\times U_h}\left(\|f-\varphi_h\|^2 + \|u-v_h\|^2\right)^{1/2}. \tag{45}$$

Proof. The error $(f-f_h, u-u_h, \lambda-\lambda_h)$ satisfies the following set of equations

$$a_h(f-f_h,\varphi_h) + g_h(u-u_h,v_h) - b_h^t(\lambda-\lambda_h,\varphi_h) + e_h(\lambda-\lambda_h,v_h) = 0, \quad \forall\varphi_h, v_h\in F_h\times U_h, \tag{46}$$

$$-b_h(f-f_h,\mu_h) + e_h^t(u-u_h,\mu_h) = 0, \quad \forall\mu_h\in\Lambda_h. \tag{47}$$

Let us define the space of all functions $(\varphi, v)\in F_h\times U_h$ verifying equation (47)

$$K := \{\varphi\times v\in F_h\times U_h : -b_h(f-\varphi,\mu_h) + e_h^t(u-v,\mu_h) = 0, \quad \forall\mu_h\in\Lambda_h\},$$

Similarly, we define a space of all functions $\mu\in\Lambda_h$ satisfying the transposed equation (47)

$$K^* := \{\mu\in\Lambda_h : -b_h^t(\lambda-\mu,\varphi_h) + e_h(\lambda-\mu,v_h) = 0, \quad \forall\varphi_h, v_h\in F_h\times U_h\}.$$

Now we choose arbitrary $(f_*, v_*) \in K$ and $\lambda_* \in K^*$. Equations (46) and (47) turn into

$$a_h(f_h - f_*, \varphi_h) + g_h(u_h - u_*, v_h) - b_h^t(\lambda_h - \lambda_*, \varphi_h) + e_h(\lambda_h - \lambda_*, v_h) \\ = a_h(f - f_*, \varphi_h) + g_h(u - u_*, v_h) \quad \forall \varphi_h, v_h \in F_h \times U_h, \tag{48}$$

$$-b_h(f_h - f_*, \mu_h) + e_h^t(u_h - u_*, \mu_h) = 0, \quad \forall \mu_h \in \Lambda_h. \tag{49}$$

From the whole set of admissible (φ_h, v_h, μ_h) we select the very special one, $\varphi_h = f_h - \varphi_*$, $v_h = u_h - v_*$, $\mu_h = \lambda_h - \lambda_*$. With this choice, it follows from (48) and (49)

$$a_h(f_h - \varphi_*, f_h - \varphi_*) + g_h(u_h - v_*, u_h - v_*) = a_h(f - \varphi_*, f_h - \varphi_*) + g_h(u - v_*, u_h - v_*).$$

Thus,

$$\|f_h - \varphi_*\|^2 + \|u - v_*\|^2 \le \frac{1}{\alpha_1}\Big(a_h(f_h - \varphi_*, f_h - \varphi_*) + g_h(u_h - v_*, u_h - v_*)\Big) \\ = \frac{1}{\alpha_1}\Big(a_h(f_h - \varphi_*, f - \varphi_*) + g_h(u_h - v_*, u - v_*)\Big) \\ \le \frac{\alpha_2}{\alpha_1}\left(\|f - \varphi_*\|^2 + \|u - v_*\|^2\right)^{1/2}\left(\|f_h - \varphi_*\|^2 + \|u_h - v_*\|^2\right)^{1/2},$$

and

$$\left(\|f_h - \varphi_*\|^2 + \|u_h - v_*\|^2\right)^{1/2} \le \frac{1}{\alpha_1}\frac{\alpha_2}{\alpha_1}\left(\|f - \varphi_*\|^2 + \|u - v_*\|^2\right)^{1/2}.$$

From the triangle inequality it follows

$$\left(\|f - f_h\|^2 + \|u - u_h\|^2\right)^{1/2} \le \left(1 + \frac{\alpha_2}{\alpha_1}\right)\left(\|f - \varphi_*\|^2 + \|u - v_*\|^2\right)^{1/2}.$$

From (48) it follows that

$$-b_h^t(\lambda_h - \lambda_*, \varphi_h) + e_h(\lambda_h - \lambda_*, v_h) = a_h(f - f_h, \varphi_h) + g_h(u - u_h, v_h).$$

Thus,

$$\|\lambda_h - \lambda_*\| \le \frac{1}{\beta_1}\left(\|-\mathbb{B}_h(\lambda_h - \lambda_*)\|^2 + \|\mathbb{E}_h(\lambda_h - \lambda_*)\|^2\right)^{1/2}$$

$$\le \frac{1}{\beta_1}\left(\|-\mathbb{A}_h(f_h - f_*)\|^2 + \|\mathbb{G}_h(u_h - u_*)\|^2\right)^{1/2} \le \alpha_2\left(\|f - f_h\|^2 + \|u - u_h\|^2\right)^{1/2}$$

and, so,

$$\|\lambda - \lambda_h\| \le \|\lambda - \lambda_*\| + \|\lambda_h - \lambda_*\| \le \|\lambda - \lambda_*\| + \frac{\alpha_2}{\beta_1}\left(1 + \frac{\alpha_2}{\alpha_1}\right)\left(\|f - f_*\|^2 + \|u - u_*\|^2\right)^{1/2}.$$

Since (f_*, u_*, μ_*) is arbitrary, we have the following estimates

$$\left(\|f - f_h\|^2 + \|u - u_h\|^2\right)^{1/2} \le \left(1 + \frac{\alpha_2}{\alpha_1}\right)\inf_{\varphi_h \times v_h \in K}\left(\|f - \varphi_h\|^2 + \|u - v_h\|^2\right)^{1/2},$$

$$\|\lambda - \lambda_h\| \le \inf_{\mu_h \in K^*}\|\lambda - \mu_h\| + \frac{\alpha_2}{\beta_1}\left(1 + \frac{\alpha_2}{\alpha_1}\right)\inf_{\varphi_h \times v_h \in K}\left(\|f - \varphi_h\|^2 + \|u - v_h\|^2\right)^{1/2}.$$

Now, proceeding as in [60] [Remark 4.10], we can bound the infimum over K by the infimum over $\mathbb{G}_h \times U_h$

$$\inf_{\varphi_h, v_h \in K}\left(\|f - f_h\|^2 + \|u - u_h\|^2\right)^{1/2} \le \left(1 + \frac{\beta_2}{\beta_1}\right)\inf_{\varphi_h \times v_h \in F_h \times U_h}\left(\|f - \varphi_h\|^2 + \|u - v_h\|^2\right)^{1/2}.$$

Similarly,

$$\inf_{\lambda_h \in K^*} \|\lambda - \lambda_h\| \leq \left(1 + \frac{\beta_2}{\beta_1}\right) \inf_{\lambda_h \in \Lambda_h} \|\lambda - \lambda_h\|.$$

Thus, the estimates (44) and (45) follow. □

This result shows that the norm of approximation error is bounded by the norm of the interpolation error

$$\left(\|f - f_h\|^2 + \|u - u_h\|^2 + \|\lambda - \lambda_h\|^2\right)^{1/2}$$

$$\leq C \inf_{\varphi_h \in F_h, v_h \in U_h, \lambda_h \in \Lambda_h} \left(\|f - \varphi_h\|^2 + \|u - v_h\|^2 + \|\lambda - \mu_h\|^2\right)^{1/2}. \tag{50}$$

We use linear node-based elements on tetrahedral grids for all of f, u, and λ. We apply the classical results on finite-element approximation in Sobolev spaces [61] to (50) to get the following estimate.

Proposition 4. *Let $(f, u, \lambda) \in F \times U \times \Lambda$ be a solution of (27)–(29), and $(f_h, u_h, \lambda) \in F_h \times U_h \times \Lambda_h$ is a solution to (40)–(42). Then, the error satisfies*

$$\left(\|f - f_h\|^2 + \|u - u_h\|^2 + \|\lambda - \lambda_h\|^2\right)^{1/2} \leq Ch \left(|f|^2_{H^2(\Gamma_B)} + |u|^2_{H^2(\Omega)} + |\lambda|^2_{H^2(\Omega)}\right)^{1/2}, \tag{51}$$

where $|\cdot|$ stands for a corresponding seminorm, and C is a constant.

7. NUMERICAL EXPERIMENTS

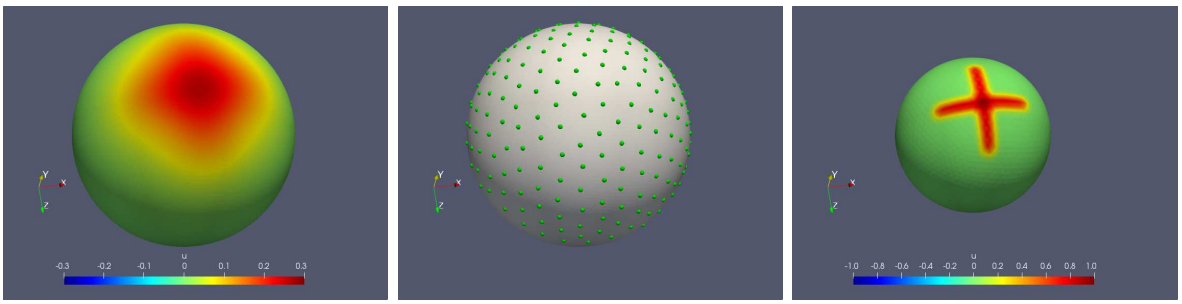
In this section, we demonstrate that the solution of the variational system provides a viable estimate of cortical activity. Our programming implementation is based on C++ finite-element library MFEM [62]. The matrix factorization was performed by a sparse direct solver UMFPACK [63].

In the first series of calculations, we reconstructed the electric potential in a spherical shell with unit conductivity. The inner and outer radii were 0.7 and 1.0, respectively. The shell was split into $86K$ tetrahedra. The potential inside the shell was represented as a finite series of spherical harmonics. The coefficients of the series were selected so that u satisfies Laplace's equation in the shell, condition $\frac{\partial u}{\partial \nu} = 0$ on the outer sphere, and $u = \xi$ on the inner sphere, where ξ is some predefined potential distribution. The method is fully described in [50] [Section IV]; it is similar to the technique in [64], apart from the fact that we solve a boundary-value problem with non-homogeneous boundary conditions and a homogeneous right-hand part. The potential on the inner shell formed a cross of unit amplitude (Fig. 2 c). The potential on the outer shell is depicted in Fig. 2 a. Synthetic EEG data were sampled at 198 electrodes evenly scattered on the hemisphere $y > 0$ (Fig. 2 b).

In the first numerical experiment, we performed source reconstruction for several values of ε from 10^{-12} to 10^{-7}. The data were noise-free. As expected, the smaller ε, the rougher cortical activity is reconstructed, see Fig. 3. We also observe that as ε is getting smaller, the changes in the estimated cortical activity are less visible. In particular, the reconstructed values for $\varepsilon = 10^{-11}$ and $\varepsilon = 10^{-12}$ are very close. The norm of the data residual (the difference between the input and calculated data) is plotted in Fig. 4.

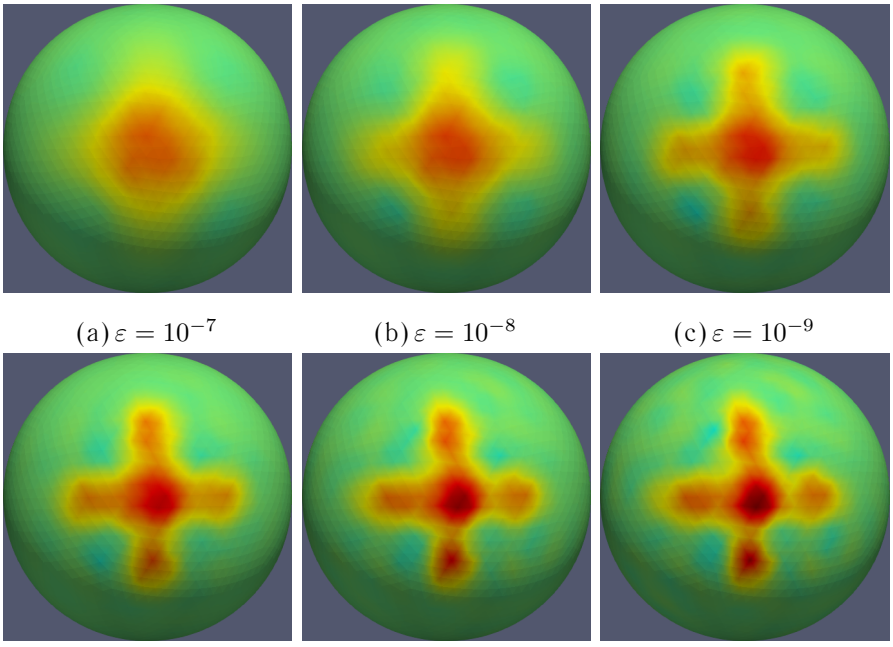
We added noise to the synthetic EEG data in the next numerical experiment. At each electrode position, the noise was Gaussian with a zero mean and the standard deviation equal to 1% of the amplitude of the actual EEG signal at that point. The resulting reconstructions are presented in Fig. 5. The reconstructed distributions remained very similar to those shown before, although the images appeared distorted due to the impact of the noise. To evaluate the data fit, we computed the root-mean-square error (RMSE) as

$$e = \sqrt{\frac{1}{K} \sum_{i=0}^{K-1} \frac{(d_i - \hat{u}_i)^2}{s_i^2}}. \tag{52}$$



(a) (b) (c)

Figure 2. Synthetic EEG data in a spherical shell. (a) Electric potential on the outer sphere (radius 1.0); (b) 198 electrodes distributed on the outer half sphere $y > 0$; (c) Electric potential on the inner sphere (radius 0.7).



(a) $\varepsilon = 10^{-7}$ (b) $\varepsilon = 10^{-8}$ (c) $\varepsilon = 10^{-9}$

(d) $\varepsilon = 10^{-10}$ (e) $\varepsilon = 10^{-11}$ (f) $\varepsilon = 10^{-12}$

Figure 3. Electric potential reconstructed on the inner boundary of the spherical shell for different values of regularization parameter ε. The color scale is the same as in Fig. 2 c. The layout consisted of 198 electrodes evenly distributed on the outer half-sphere centered at the cross. The input data were noise-free.

Here s_i are the estimated values of the standard deviation of the noise at the ith electrode, $\hat{u}_i$ is the final calculated electric potential. Value $e = 1$ means that the data fit is close to optimal: on average, it corresponds to the noise level. The plot of e of ε for the 1% noise is given in Fig. 6. In this sense, the optimal value of regularization is $\varepsilon = 10^{-9}$. Below that level, the data are over-fitted, so, the image of cortical activity on the cortex gets progressively distorted.

We repeated the same calculation with the 5% noise. The reconstructed distribution of u is presented in Fig. 7. The misfit is plotted in Fig. 8. We observe that a higher regularization is needed for the 5% noise compared to the 1% one. The spatial resolution of the reconstructed distributions is lower, as expected.

We applied our algorithm to a realistic head model as proof of the concept. We employed the mesh described in [50]. The model consisted of four compartments to which the following conductivities were assigned: skin 0.33 S/m, skull 0.011 S/m, CSF 1.0 S/m, and brain 0.33 S/m. The mesh contained 587,882 tetrahedra. We set up two active patches in the visual cortex containing 146 unit dipoles. The

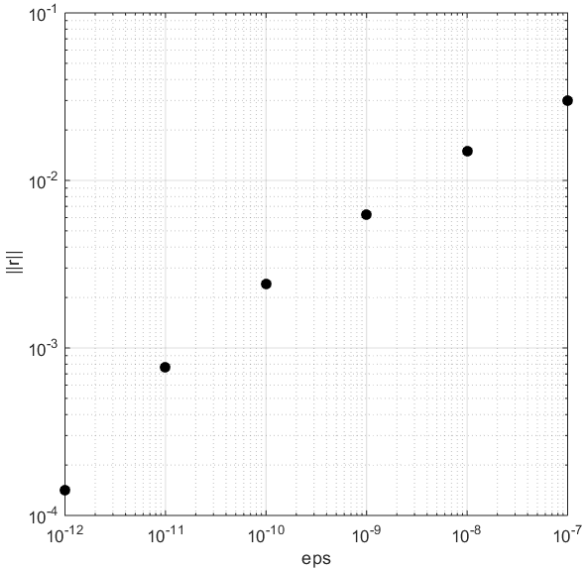
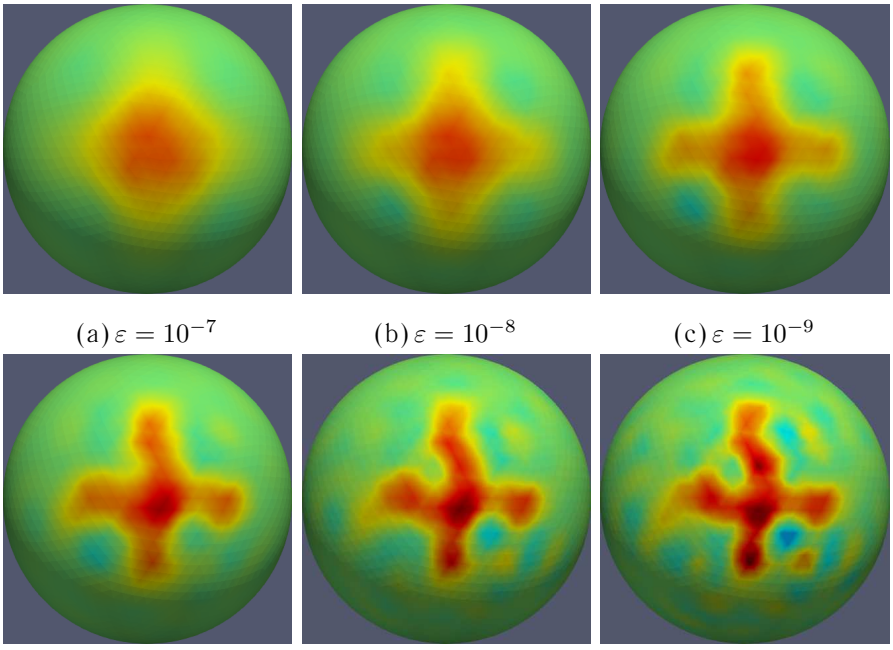


Figure 4. The norm of the data residual $||r||$ as a function of ε.



(a) $\varepsilon = 10^{-7}$ (b) $\varepsilon = 10^{-8}$ (c) $\varepsilon = 10^{-9}$

(d) $\varepsilon = 10^{-10}$ (e) $\varepsilon = 10^{-11}$ (f) $\varepsilon = 10^{-12}$

Figure 5. Electric potential reconstructed on the inner boundary of the spherical shell from the data contaminated with 1% noise. The color scale is the same as in Fig. 2 c.

dipoles were located in the barycenters of the tetrahedra. Orientations of the dipoles were normal to the cortex. The forward problem was solved by FEM using the standard H^1-conforming finite elements of the first order. Synthetic EEG data were obtained by sampling the potential at 175 electrodes extracted from Philips' 256-electrode Geodesic Sensor Net montage. We adjusted electrode locations to make them coincide with some grid vertices. The mesh, brain activations, simulated potential, and electrode locations are depicted in Fig. 9.

The mesh used for reconstruction was created from the same surfaces, but the brain was replaced with a void. We changed meshing parameters to make the mesh differ from the original one, thus avoiding the

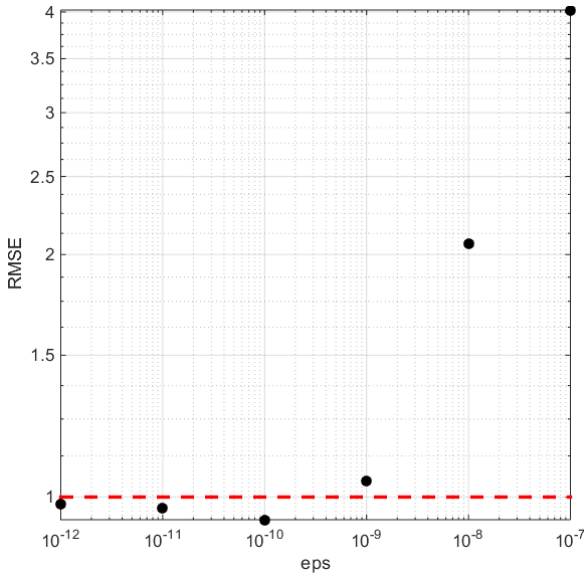


Figure 6. The root-mean-square error e of data contaminated with the 1% noise as a function of ε. The dashed red line marks the value e=1.

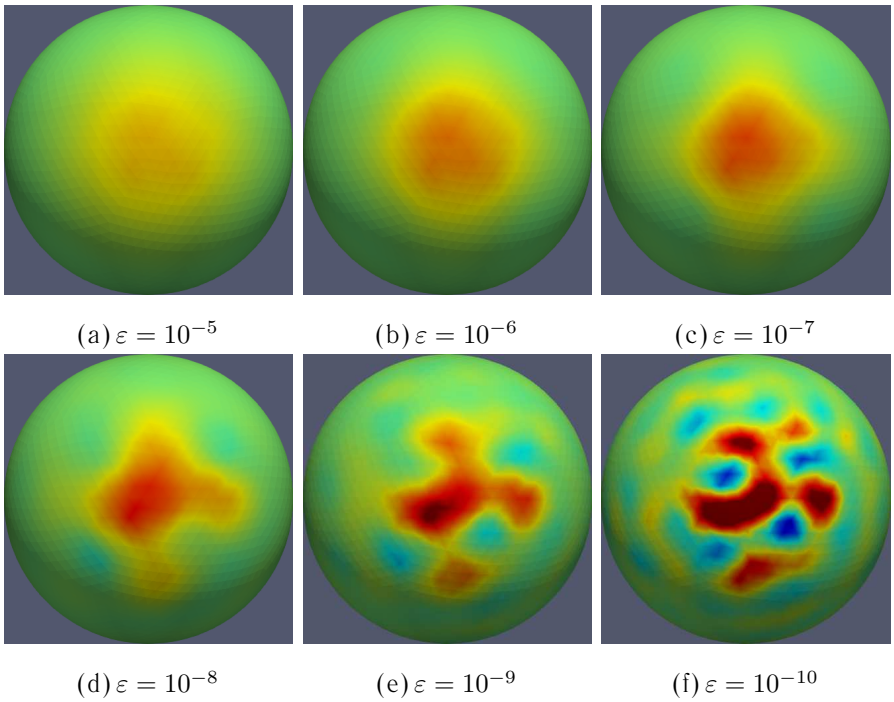


Figure 7. Electric potential reconstructed on the inner boundary of the spherical shell from the data contaminated with 5% noise. The color scale is the same as in Fig. 2 c.

so-called inverse crime. The final grid consisted of 658,513 tetrahedra. Reconstruction was performed using the two values of ε: 10^{-3} and 10^{-4}. The resulting distributions of potential on the cortex are given in Fig. 10.

The reconstructed cortical activity is consistent with the location of the actual sources. Predictably, the reconstruction for bigger ε is more smeared. We have yet to deal with the optimal choice of ε, but generally, the theory of regularization of linear discrete inverse problems is well developed [65, 66].

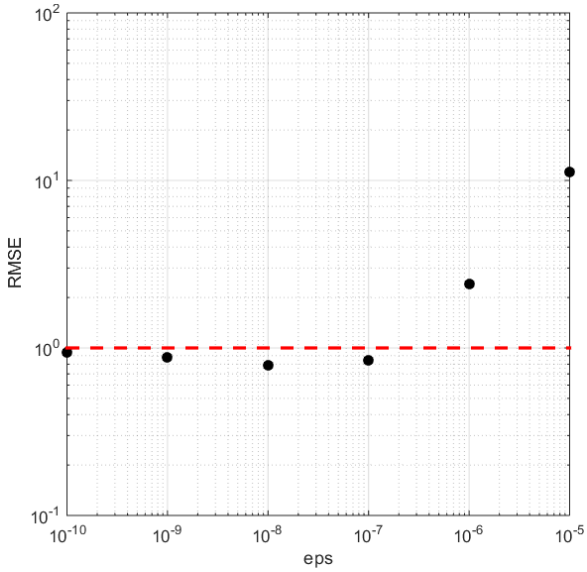


Figure 8. The root-mean-square error e of data contaminated with the 5% noise as a function of ε. The dashed red line marks the value e=1.

Application of Morozov's principle may provide good results [67]. There is also abundant literature targeting specifically EEG/MEG inverse problems, e.g., [68], among many others.

8. CONCLUSIONS

In this article, we applied OC formalism to the inverse EEG problem. The inverse problem is reduced to a system of variational equations giving rise to a system of linear equations. This system is a 3×3 block matrix. It is large symmetric indefinite, and sparse. Thus, the estimation of the EEG sources is performed by a single linear solve. We presented a numerical test demonstrating that this approach provides a consistent estimate of cortical activity. We demonstrated that our algorithm is stable in the presence of noise in the input data.

Our approach fully addresses the critical shortcomings of the mixed QRM: discrete data and noise characteristics are naturally taken into account when solving the inverse problem. In contrast to the standard methods of the MNE/LORETA type, which estimate the power of individual dipoles, our approach reconstructs the density of the dipole layer located on the cortex. On the one hand, this makes our approach more specialized. On the other, it reduces the mesh size and eliminates the need to distinguish between grey and white matter in the model because the area occupied by the brain is excluded from consideration. Also, our approach avoids the problem of simulation of the dipolar right-hand side - a persisting issue with all standard FEM-based methods of MNE/LORETA type. In other respects, our method offers the same flexibility as the FEM-based estimators of MNE/LORETA type, such as the ability to work with anatomical head models of almost arbitrary complexity, non-constant conductivity within head compartments, and holes in the interfaces. The software implementation of this approach does not require special efforts since it uses programming blocks available in standard FEM packages. Finally, our method can be extended to arbitrarily oriented dipoles using Raviart–Thomas basis functions for dipole approximation.

Author Contributions: Conceptualization, software implementation, experimenting, writing, M.M; software implementation, experimenting, N.Y.; experimenting, writing A.R., experimenting, writing, V.G., funding acquisition, administration, N.K.

Funding: This research was supported by the Russian Science Foundation, project no. 21-11-00139.

Conflicts of Interest: The authors declare no conflict of interest.

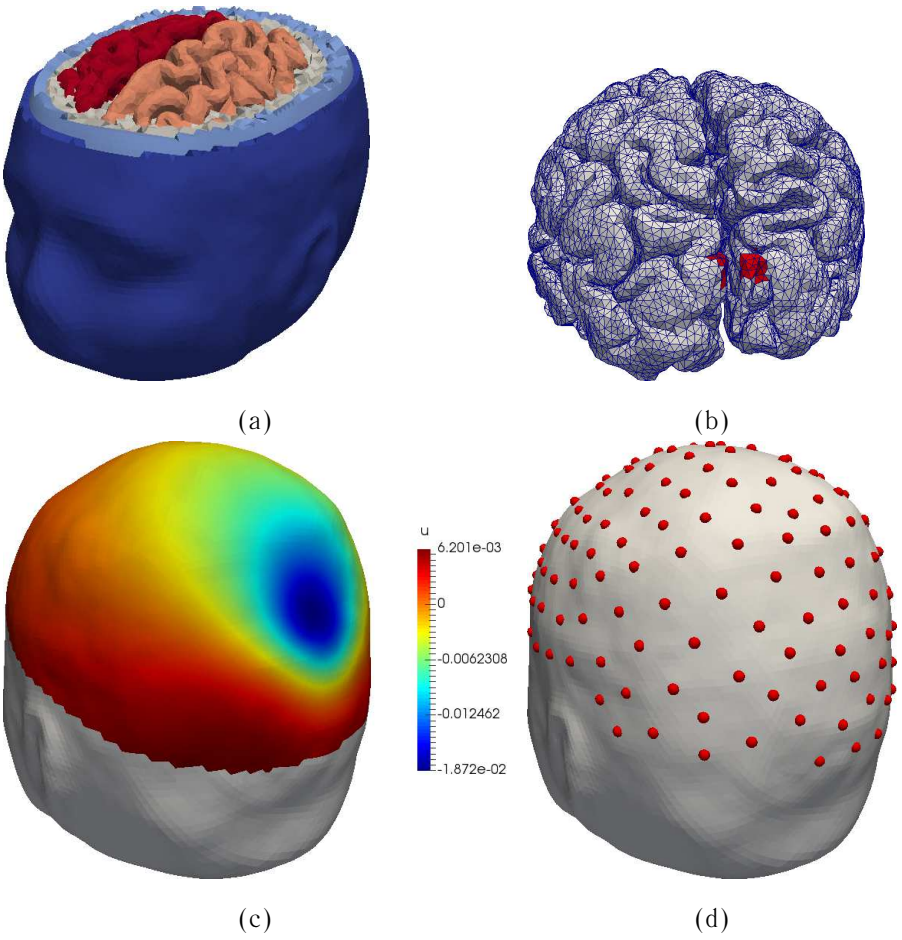


Figure 9. Simulated EEG data used in the numerical experiment. Panel (a). Tetrahedral mesh used for simulating EEG data. Panel (b). Prescribed activations in the visual cortex. Panel (c). Simulated potential on the scalp. Panel (d). 175 electrodes extracted from a 256-electrode montage.

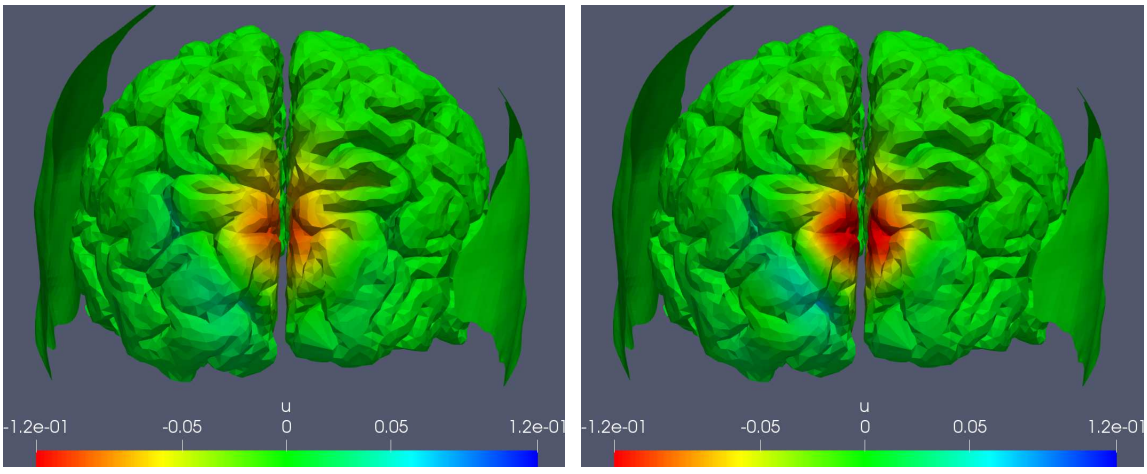


Figure 10. Potential u recovered at the cortex. Regularization parameter was set to $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$ on panels (a) and (b), respectively.

REFERENCES

1. S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping", IEEE Signal Processing Magazine **18** (6), 14–30 (2001).
2. C. M. Michel, M. M. Murray, G. Lantz, S. Gonzales, L. Spinneli, and R. Grave de Perlata, "EEG source imaging", Clinical Neurophysiology **115**, 2195–2222 (2004).
3. L. Soufflet and P. H. Boeijinga, "Linear Inverse Solutions: Simulations from a Realistic Head Model in MEG", Brain Topography **18** (2), 87–99 (2005).
4. R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis", J. of NeuroEngineering and Rehabilitation **6** (25) (2008).
5. R. D. Pascual-Marqui, K. Sekihara, D. Brandeis, and C. M. Michel, "Imaging the electric neuronal generators of EEG/MEG", in Electrical Neuroimaging, Cambridge University Press (2009).
6. C. M. Michel and B. He, *EEG mapping and source imaging,* in Niedermeyer's electroencephalography: basic principles, clinical applications, and related fields, Lippincott Williams & Wilkins (2010).
7. S. P. Ahlfors and M. S. Hamalainen, *MEG and EEG: source estimation*, in Handbook of Neural Activity Measurement, Cambridge University Press (2012).
8. M. M. Lavrentev, V. G. Romanov, and G. P. Shishatskii, *Ill-Posed Problems in Mathematical Physics and Analysis* (Mathematical Society, Providence, 1986).
9. V. Isakov, *Inverse Problems for Partial Differential Equations* (Springer, 2006).
10. A. I. Prilepko, D. G. Orlovsky, and I. A. Vasin, *Methods for solving inverse problems in mathematical physics* (Marcel Dekker, Inc., 2000).
11. M. Klibanov, "Carleman Estimates and Inverse Problems in the Last Two Decades", in Surveys on Solution Methods for Inverse Problems, Springer (2000).
12. M. Klibanov and F. Santosa, "A computational quasi-reversibility method for Cauchy problems for Laplace's equation", SIAM J. Appl. Math. **51** (6), 1653–1675 (1991).
13. V. Kozlov, V. Mazya, and A. Fomin, "An iterative method for solving the Cauchy problem for elliptic equations", Math. Math. Phys. **31** (1), 45–52 (1991).
14. S. I. Kabanikhin and A. L. Karchevsky, "Optimizational method for solving the Cauchy problem for an elliptic equation", J. of Inverse and Ill-posed Problems **3** (1), 21–46 (1995).
15. A. Cimetiere, F. Delvare, M. Jaoua, and F. Pons, "Solution of the Cauchy problem using iterated Tikhonov regularization", Inverse Problems **17**, 553–570 (2001).
16. D. N. Hao, L. T. T. Giang, S. Kabanikhin, and M. Shishlenin, "A finite difference method for the very weak solution to a Cauchy problem for an elliptic equation", J. of Inverse and Ill-posed Problems **26** (6), 835–857 (2018).
17. A. S. Leonov, "Source recovery with a posteriory error estimates in linear partial differential equations", J. of Inverse and Ill-posed Problems **28** (5), 677–692 (2002).
18. A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach", J. of Cognitive Neuroscience **5** (2), 162–176 (1993).
19. M. Fuchs, M. Wagner, T. Kohler, and H. A. Wischmann, "Linear and nonlinear current density reconstructionsLinear and nonlinear current density reconstructions", J. Clin. Neurophysiol. **16**, 267–295 (1999).
20. F. H. Lin, T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hamalainen,"Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates", NeuroImage **31** (1), 160–171 (2006).
21. M. S. Hamalainen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates", Medical & Biological Engineering & Computing **32** (1), 35–42 (1994).
22. R. D. Pascual-Marqui, C. M. Michel, and D. Lehman,"Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain", Int. J. Psychophysiol. **18**, 49–65 (1994).
23. R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details", Methods Find. Exp. Clin. Pharmacol. (Suppl. D) **24**, 5–12 (2002).
24. R. De Peralta-Menendez, M. M. Murray, C. M. Michel, R. Martuzzi, and S. G. Andino, "Electrical neuroimaging based on biophysical constraints", Neuroimage **21**(2), 527–539 (2004).
25. I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm", Electroencephalography and Clinical Neurophysiology **95** (4), 231–251 (1995).
26. R. G. De Peralta-Menendez, O. Hauk, S. Gonzalez-Andino, H. Vogt, and C. Michel, "Linear inverse solutions with optimal resolution kernels applied to electromagnetic tomography", Human Brain Mapping **5** (6), 454–467 (1997).
27. N. Yavich, N. Koshev, M. Malovichko, A. Razorenova, and M. Fedorov, "Conservative Finite Element Modeling of EEG and MEG on Unstructured Grids", IEEE Transactions on Medical Imaging (2021).

28. F. Drechsler, C. H. Wolters, T. Dierkes, H. Si, and L. Grasedyck, "A full subtraction approach for finite element method based source analysis using constrained Delaunay tetrahedralisation", Neuroimage **46** (4), 1055–1065 (2009).
29. D. Gullmar, J. Haueisen, M. Eiselt, F. Giessler, L. Flemming, A. Anwander, T. R. Knosche, C. H. Wolters, M. Dumpelmann, D. S. Tuch, and J. R. Reichenbach, "Influence of anisotropic conductivity on EEG source reconstruction: investigations in a rabbit model", IEEE Transactions on Biomedical Engineering **53** (9), 1841–1850 (2006).
30. C. Engwer, J. Vorwerk, J. Ludewig, and C. H. Wolters, "A Discontinuous Galerkin Method to Solve the EEG Forward Problem Using the Subtraction Approach", SIAM J. on Scientific Computing **39** (1), 138–164 (2017).
31. J. Vorwerk, C. Engwer, S. Pursiainen, and C. H. Wolters, "A Mixed Finite Element Method to Solve the EEG Forward Problem", IEEE transactions on medical imaging **36** (4), 930–941 (2017).
32. S. Pursiainen, J. Vorwerk, and C. H. Wolters, "Electroencephalography (EEG) forward modeling via H(div) finite element sources with focal interpolation", Physics in Medicine and Biology **61** (24), 8502–8520 (2016).
33. C. Nicholson, "Theoretical Analysis of Field Potentials in Anisotropic Ensembles of Neuronal Elements", IEEE Transactions on Biomedical Engineering **20** (4), 278–288 (1973).
34. P. Nicolas and G. Deloche, "Convolution computer processing of the brain electrical image transmission", International J. of Bio-Medical Computing **7** (2), 143–159 (1976).
35. W. J. Freeman, "Use of Spatial Deconvolution to Compensate for Distortion of EEG by Volume Conduction", IEEE Transactions on Biomedical Engineering **27** (8), 421–429 (1980).
36. P. L. Nunez, R. B. Silberstein, P. J. Cadusch, and R. Wijesinghe, "Comparison of high resolution EEG methods having different theoretical bases", Brain Topography **5** (4), 361–364 (1993).
37. R. Srinivasan, P. L. Nunez, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch, "Spatial sampling and filtering of EEG with spline Laplacian to estimate cortical potential", Brain Topography **8** (4), 355–366 (1996).
38. M. Junghofer, T. Elbert, P. Leiderer, P. Berg, and B. Rockstroh, "Mapping EEG-potentials on the surface of the brain: A strategy for uncovering cortical sources", Brain Topography **9** (3), 203–217 (1997).
39. C. E. Tenke and J. Kayser, "Reference-free quantification of EEG spectra: Combining current source density (CSD) and frequency principal components analysis (fPCA)", Clinical Neurophysiology **116** (12), 2826–2846 (2005).
40. D. Haor, R. Shavit, M. Shapiro, and A. B. Geva, "Back-projection cortical potential imaging: theory and results", IEEE Transactions on medical imaging **36** (7), 1583–1595 (2017).
41. A. Gevins, P. Brickett, B. Costales, J. Le, and B. Reutter, "Beyond topographic mapping: toward functional-anatomical imaging with 124-channel EEGs and 3-D MRIs", Brain topography **3** (1), 53–64 (1990).
42. A. S. Gevins and J. Le, "EEEG spatial enhancement method and system", U.S. Patent 5331970 (1992).
43. A. S. Gevins and J. Le, "EEG deblurring method and system for improved spatial detail", U.S. Patent 5568816 (1996).
44. B. He, Y. Wang, and D. Wu, "Estimating cortical potentials from scalp EEGs in a realistically shaped inhomogeneous head model by means of the boundary element method", IEEE Transactions on Biomedical Engineering **46** (10), 1264–1268 (1999).
45. B. He, X. Zhang, J. Lian, H. Sasaki, D. Wu, and V. L. Towle, "Boundary element method-based cortical potential imaging of somatosensory evoked potentials using subjects magnetic resonance images", NeuroImage **16**, 564–576 (2002).
46. M. Clerc and J. Kybic, "Cortical mapping by Laplace–Cauchy transmission using a boundary element method", Inverse Problems **23** (6), 2589-2601 (2007).
47. L. Bourgeois, "A mixed formulation of quasi-reversibility to solve the Cauchy problem for Laplace's equation", Inverse Problems **21** (3), 1087–1104 (2005).
48. L. Bourgeois, "Convergence rates for the quasi-reversibility method to solve the Cauchy problem for Laplace's equation", Inverse Problems **22** (2), 413–430 (2006).
49. N. Koshev, N. Yavich, M. Malovichko, E. Skidchenko, and M. Fedorov, "FEM-based Scalp-to-Cortex EEG data mapping via the solution of the Cauchy problem", J. Inverse Ill-Posed Probl. **28** (4), 517–532 (2020).
50. M. Malovichko, N. Koshev, N. Yavich, A. Razorenova, and M. Fedorov, "Electroencephalographic Source Reconstruction by the Finite-Element Approximation of the Elliptic Cauchy Problem", IEEE Transactions on Biomedical Engineering **68** (6), 1811–1819 (2021).
51. M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints* (Springer, 2009).
52. S. Vallagh, T. Papadopoulo, and M. Clerc, "The adjoint method for general EEG and MEG sensor-based lead field equations", Physics in medicine and biology **54** (1), 135–147 (2009).
53. O. Faugeras, F. Cement, R. Deriche, R. Keriven, T. Papadopoulo, J. Roberts, T. Vieville, F. Devernay, J. Gomes, G. Hermosillo, P. Kornprobst, and D. Lingrand, "The Inverse EEG and MEG Problems: The Adjoint State Approach I: The Continuous Case", in Projet CERMICS, Research Report (1999).
54. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations* (Springer Berlin Heidelberg, 1971).

55. C. Bernardi and R. Verfurth, "Adaptive finite element methods for elliptic equations with non-smooth coefficients", Numer. Math. **85**, 579–608 (2000).
56. I. Tanzer, S. Jarvenpaa, J. Nenonen, and E. Somersalo, "Representation of bioelectric current sources using Whitney elements in the finite element method", Phys. Med. Biol. **50**, 3023–3039 (2005).
57. S. Pursiainen, A. Sorrentino, C. Campi, and M. Piana, "Forward simulation and inverse dipole localization with the lowest order Raviart–Thomas elements for electroencephalography", Inverse problems **27** (4), 1–17 (2011).
58. S. Pursiainen, "Raviart–Thomas-type sources adapted to applied EEG and MEG: implementation and results", Inverse problems **28** (8), 065013 (2012).
59. F. Brezzi, "On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers", ESAIM: Mathematical Modeling and Numerical Analysis **8** (2), 129–151 (1974).
60. D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics* (Cambridge University Press, 2007).
61. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems* (North Holland, 1978).
62. R. Anderson, J. Andrej, A. Barker, J. Bramwell, J. S. Camier, J. Cerveny, V. Dobrev, Y. Dudouit, A. Fisher, Tz. Kolev, W. Pazner, M. Stowell, V. Tomov, I. Akkerman, J. Dahm, D. Medina, and S. Zampini, "MFEM: A Modular Finite Element Methods Library", Computers & Mathematics with Applications **81**, 42–74 (2021).
63. D. A. Timothy, "Algorithm 832: UMFPACK V4.3—an Unsymmetric-Pattern Multifrontal Method", ACM Trans. Math. Softw. **30** (2), 196–199 (2004).
64. J. C. de Munck and M. J. Peters, "A Fast Method to Compute the Potential in the Multisphere Model", IEEE Tran Biomed Eng **40** (11), 1166–1164 (1993).
65. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems* (Society for Industrial and Applied Mathematics, 1998).
66. C. R. Vogel, *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, 2002).
67. T. Bonesky, "Morozovs discrepancy principle and Tikhonov-type functionals", Inverse Problems **25**, 015015 (2009).
68. D. Engemann and A. Gramfort, "Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals", NeuroImage **108**, 328–342 (2014).