

Application of optimal control to inversion of self-potential data: theory and synthetic examples

M.S. Malovichko, A.V. Tarasov, N.B. Yavich, K.V. Titov,

Abstract—Last decades, there has been an increased interest in the use of the self-potential (SP) method in hydrogeophysics. In response to this strong interest, we develop a novel approach to the inversion of SP data. Mathematically, the SP inverse problem is the source identification problem for the Poisson equation. Our approach substantially differs from the standard regularization approach, which explicitly includes the forward-problem operator into the cost functional. We formulated the inverse problem as an optimal control problem and then translate it into a variational system. The system is approximated in suitable finite-element spaces giving rise to an algebraic problem with the saddle-point structure. In contrast to the standard approach, which leads to a dense linear system, our method results in a system with a sparse block matrix. It can be efficiently solved by either direct sparse solvers or preconditioned iterative solvers. In this paper, we present the formulation of the problem and its finite-element approximation. We discuss the iterative solution and preconditioning strategies. Our software implementation is based on an industrial finite-element package. We also present a numerical experiment with node-based linear finite elements on tetrahedral grids. Our results suggest that the proposed approach may serve as a rapid and reliable tool for large-scale SP inverse problems. Moreover, the same technique can easily be extended to a wide range of geophysical linear inverse problems, such as inversions of magnetic and gravity data.

Index Terms—inverse problem, self-potential, optimal control, KKT system.

I. INTRODUCTION

THIS paper considers numerical aspects of source identification in the geophysical self-potential (SP) method. Self-potential is one of the oldest geophysical techniques, and is still widely used in mineral exploration, geological mapping and hydrogeology [1]. Last decade, a breakthrough occurred in the use of SP measurements for imaging underground water flows and redox potential [2, 1, 3, 4, 5].

From the mathematical standpoint, the SP inverse problem is the source identification problem for Poisson's equation. Let Ω be a bounded domain in \mathbb{R}^3 with boundary Γ . Let $\Gamma = \Gamma_D \cup \Gamma_N$, where Γ_N represents the air-ground interface, with Γ_D being the rest of the boundary. The electric potential satisfies Poisson's equation,

$$-\operatorname{div}(\sigma \nabla u) = f \quad \text{in } \Omega, \quad (1.1)$$

$$u = g_D \quad \text{on } \Gamma_D, \quad \frac{\partial u}{\partial \nu} = g_N \quad \text{on } \Gamma_N. \quad (1.2)$$

M. Malovichko and N. Yavich are with CDISE, Skolkovo Institute of Science and Technology, Moscow, Russia, and Moscow Institute of Physics and Technology, Dolgoprudny, Russia, e-mail: m.malovichko@skoltech.ru

A.Tarasov and K. Titov are with St.-Petersburg University, Institute of Earth Sciences, St.-Petersburg, Russia

Manuscript received ; revised .

Here u is the scalar electric potential, σ is the electric conductivity, vector ν is a unit outward normal. The right-hand side of (1), f , represents the source term (in Am^{-3}), generated by external currents I , thus $f = \operatorname{div} I$. The external currents arise from gradients of thermodynamic potentials (pressure, chemical potential, and temperature) and are the cross-coupled phenomena (e.g., [6]). We set $g_N = 0$ and $g_D = 0$. Note that so-called "oblique" Dirichlet boundary conditions on Γ_D are beneficial and have gained considerable popularity in geophysics, see [7]. We, however, refrain from it because it would complicate the presentation of our approach. The problem (1) is well-posed; that is, it has a unique and stable solution.

In the SP method, f is unknown. We measure potential in a set of K points using so-called non-polarizing electrodes. Based on observed data, we need to obtain a distribution of current per cubic meter, f . Formally, we set up the following inverse problem.

Given measured data $d \in \mathbb{R}^K$, find f such that

$$\sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i u - d_i)^2 < \epsilon, \quad (2)$$

with f and u satisfying (1).

Here ϵ is the desired misfit, $\{w_i\}$ are positive weights. Typically, the weights represent estimates of noise standard deviations computed from measured data. Each operator \mathcal{Q}_i maps potential u to the value in the i -th electrode. For concreteness, let us define \mathcal{Q}_i as a convolution with the Dirac delta function,

$$\mathcal{Q}_i(u) = \int_{\Omega} u(x) \delta(x - x_i) dV, \quad (3)$$

where x_i is position of i -th electrode.

In geophysics, a common approach to inverse problem (2) is the unconditional minimization of the Tikhonov functional,

$$\underset{f}{\operatorname{minimize}} \quad \frac{1}{2} \|W(\mathcal{G}(f) - d)\|_2^2 + \frac{\alpha}{2} \mathcal{R}(f). \quad (4)$$

Here $\|\cdot\|_2$ is the ℓ^2 norm, $\alpha > 0$ is the regularization parameter, W is a diagonal matrix with weights w_i on its diagonal, \mathcal{G} is the *forward-problem operator* that maps a given f to data vector d , \mathcal{R} is a (frequently quadratic) stabilization functional that imposes constraints on f . Many variants of formulation (4) exists, including non-quadratic \mathcal{R} , various norms in which the data misfit is measured (e.g. $\|\cdot\|_1$), iterative updating of W , etc. Nevertheless, all of them are connected to this basic formulation in one way or another. It is fundamentally important that formulation (4) explicitly includes the forward-problem operator \mathcal{G} .

Formulation (4) forms a basis for linear inversion of geophysical data. We refer to [8] for a comprehensive review. Regarding inversion of SP data, this approach was used in [9, 2, 10, 11, 12, 3, 13]. Similar problems arise in the gravity and the low-susceptibility magnetic inversion, see [14, 15, 16, 17] (among many others). All those examples differ primarily in what quantity is reconstructed (f or I), what are the data (potential u or its derivatives), how the forward problem is formulated (the integral equations, finite differences, etc.) and solved (direct or iterative solvers), and whether σ is constant (gravity and magnetics) or varies in space and, possibly, in time(SP). Note that the gravity migration by [18] in its iterative variant [19] is essentially equivalent to the standard inversion (4). Here, we leave aside methods leading to non-linear inverse problems, [20, 21], and those based on the downward continuation [22].

Although the celebrated minimization (4) is widespread, it posses some drawbacks. The most important one, in our view, is that it leads to a resource-intensive numerical procedure. It requires solving a normal system of linear equations with a large dense matrix, known as the (regularized) Hessian. Direct solvers require $O(n^2)$ storage and $O(n^3)$ operations and quickly become prohibitive for big n , even on modern supercomputers. On the other hand, iterative solvers suffer from slow convergence because the system matrix is badly conditioned. It is hard to improve the convergence because no efficient preconditioning is available for the Hessian. Using the "background" Hessian as a preconditioner is common, but its performance severely degrades as conductivity σ deviates from the background conductivity. The matrix-vector multiplication has to be implemented via the two forward-problem solves for sufficiently big problems. Thus, thousands of solutions to the forward problem are expected on large inverse problems encountered in practice.

The optimal control (OC) framework is an alternative approach to the inverse problem (2). It is a firmly established technique in PDE-constrained optimization [23]. Elements of the OC have been routinely applied in geophysics to compute the gradient and Hessian of the target functional through the adjoint state variable (see [24, 25]), but solely within the standard framework based on the normal system. To the best of our knowledge, the only relevant papers that exploit the OC to derive an alternative to the standard approach are [26, 27].

Within the OC framework, the inverse problem reduces to a variational saddle-point problem followed by approximation in a suitable finite-element (FE) space. The resulting system matrix is three times bigger compared to the Hessian but has a different structure. It is sparse, indefinite, and its blocks are tightly connected to the underlined Poisson equation. The application of modern direct sparse solvers can be very efficient if the size of the problem is not too large. This matrix, in principle, can be preconditioned using the matured theory of the numerical solution of PDEs. Thus, for larger problems, the application of an iterative solver may be beneficial, provided an efficient preconditioner is available. On the other hand, the OC approach has several difficulties, which may explain its low popularity in geophysics. The direct solution of a sparse indefinite system of equations is less stable and more resource-

consuming compared to the positive definite matrices [28, 29]. As for the iterative solution, preconditioning of matrices of this type is not a trivial task.

In this paper, we present an approach to the source reconstruction problem based on the OC formalism. In essence, our approach replaces the normal system of linear equations with a dense matrix with another system with a larger but sparse matrix. This matrix can be solved either directly with a sparse direct solver or iteratively with a carefully designed preconditioner. Although we primarily focus on the general concept, we also discuss the numerical solution. We elaborate on the formulation of the problem, its reduction to a variational system, and a FE discretization. The presented numerical experiments aim to verify the formulas and demonstrate this approach's applicability to realistic 3D problems. We did not set ourselves to analyze various iterative solvers (this part of the work has not been finished yet). Still, we present one option, which seems quite effective, hoping to stimulate further work in this direction. We believe that the OC approach has great potential in geophysics and can become an alternative to the standard minimization (4), at least in some scenarios.

Our paper significantly differs from papers [26, 27]. They solve a coefficient inverse problem of quasi-static electromagnetics, which is non-linear, whereas we solve a linear inverse problem of source reconstruction. They use finite-difference discretization, whereas we employ FEM, which is more geometrically flexible. They discretize differential equations and then find an optimal point (the *discretize-then-optimize* method); we use the *optimize-then-discretize* method, which is a better fit for the final-element framework. In addition, we used a very different approach to preconditioning, compared to those articles, although here we report preliminary results.

The paper is organized as follows. In Section II, we review the standard approach. In Section III, we present the OC approach and derive the variational system. Finite-element discretization is discussed in Section IV. The connection between the two approaches is discussed in Section V. The iterative solution of the arising system of linear equations is covered in Section VI. Section VII is dedicated to numerical experiments. Concluding remarks are given in Section VIII.

II. THE STANDARD APPROACH

Here we review Tikhonov's approach in its classical form, i.e., when the forward operator explicitly enters the target function. Consider the following unconstrained optimization problem,

$$\underset{f}{\text{minimize}} \quad \frac{1}{2} \sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i \mathcal{S} f - d_i)^2 + \frac{\alpha}{2} \int_{\Omega} |\mathcal{F} \nabla f|^2 dV. \quad (5)$$

Here \mathcal{F} is a smooth real-valued function designed to account for the rapid decrease of sensitivity with depth. Operator \mathcal{S} maps a given f to potential u and thus represents the inverse of the differential operator in (1). Since problem (1) is well-posed, \mathcal{S} exists. A scalar $\alpha > 0$ is the regularization parameter, which can be selected in various ways, such as Morozov's discrepancy principle [30, 8]. Looking ahead, any algorithm of

choosing α for the standard method will work for our approach (see Section V).

In geophysical applications, minimization problem (5) is usually discretized first; the optimization is performed on the discrete level. To make this section compatible with the rest of this paper, we employed the finite-element method (FEM) with linear node-based elements in a tetrahedral grid, although many other options are available.

First of all, we define the discrete version of operator \mathcal{S} , that maps a given f to u . To simplify the formulas below, we will assume that $g_N = 0$, which is always true in geophysics. First, (1) translates to the following variational problem,

Find $u \in V$ such that

$$\int_{\Omega} \sigma \nabla u \cdot \nabla v \, dV = \int_{\Omega} f v \, dV, \quad \forall v \in V_0, \quad (6)$$

where

$$V = \{v \in H^1(\Omega) : v = g_D \text{ on } \Gamma_D\}, \quad (7.1)$$

$$V_0 = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}. \quad (7.2)$$

We used the standard notation $H^1(\Omega)$ for the Sobolev space of functions having the square-integrable gradient.

Note that, in FEM textbooks, f is usually specified as $f \in H^{-1}(\Omega)$ (a negative-order Sobolev space dual to $H^1(\Omega)$). Such an f can be highly irregular, but the integral in the right-hand side of (6) exists. However, our final goal is to find an estimation of f through (5), meaning that the second integral in (5) must exist. With this in mind, we assume $f \in H^1(\Omega)$, i.e., much more regular than needed to merely solve the forward problem.

Upon FE discretization of (6), we receive the following system of linear equations,

$$E\tilde{u} = B\tilde{f}. \quad (8)$$

Here \tilde{u} and \tilde{f} consist of degrees of freedom (DOFs) associated with u and f , respectively. Matrix E is the stiffness matrix,

$$E_{ij} = \int_{\Omega} \sigma \nabla \zeta_j \cdot \nabla \zeta_i \, dV, \quad (9)$$

and B is the mass matrix,

$$B_{ij} = \int_{\Omega} \zeta_j \zeta_i \, dV, \quad (10)$$

where ζ_i are the basis functions. The appearance of the mass matrix B in the definition of the load vector, $B\tilde{f}$, is the direct consequence of our choice $f \in H^1(\Omega)$. From (8) we see that $\tilde{u} = E^{-1}B\tilde{f}$; thus, a discrete version of operator \mathcal{S} is $E^{-1}B$.

Problem (5) reduces to the following discrete quadratic program,

$$\underset{f}{\text{minimize}} \quad \frac{1}{2} \left\| W(QE^{-1}B\tilde{f} - d) \right\|_2^2 + \frac{\alpha}{2} \left\| FL^{1/2}\tilde{f} \right\|_2^2, \quad (11)$$

where W is a diagonal matrix of weights, F and Q are discrete representations of \mathcal{F} and $\{\mathcal{Q}_k\}_{k=0}^{K-1}$, respectively. Matrix L is the discrete Laplacian,

$$L_{ij} = \int_{\Omega} \nabla \zeta_j \cdot \nabla \zeta_i \, dV. \quad (12)$$

Problem (11) is known in the optimization community as *the reduced Jacobian* approach. Solution of (11) is given by

$$(J^T W^2 J + \alpha L)\tilde{f} = J^T W^2 d, \quad (13)$$

where

$$J = QS = QE^{-1}B. \quad (14)$$

Matrix J is commonly referred to as *the Jacobian*; $J^T W^2 d$ is known as *the gradient*. The system matrix in (13) is the (regularised) Hessian. Many variants of this basic formulation have been described in the literature. In particular, the norms can be computed in non-Hilbert spaces (e.g., $\|\cdot\|_1$), L may include constraints imposed on the solution, the Jacobian may be computed explicitly or implicitly (in the form of the matrix-vector product) or can be replaced by an approximation, and so on. In all the cases, the Jacobian J is the main ingredient, and major computational challenges relate to it.

It happens pretty often that Jacobian J cannot be stored due to its sheer size. In that case, the only feasible option is an iterative solver, such as the conjugate (CG) gradient iterations. The matrix-vector product includes multiplication by J followed by multiplication by J^T . Thus, two forward simulations per CG iteration are required.

Algebraic problem (13) is badly conditioned. Figure 1 pictures a typical convergence, recorded for a small 2D problem. Note the erratic behavior of the relative misfit and the slow

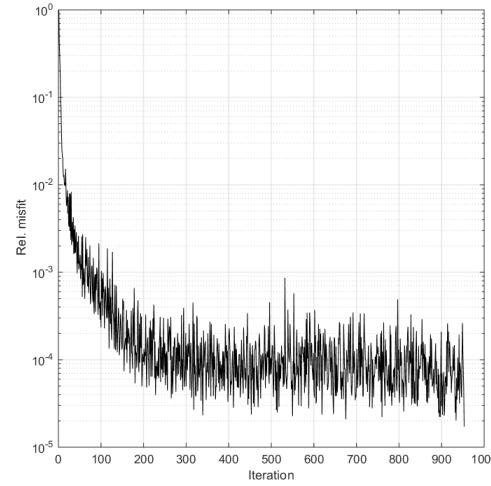


Figure 1: A typical convergence of the conjugate gradients applied to (13). Convergence is almost stalled after the 200th iteration. Convergence can be much worse or much better, depending on the value of α .

progress of the solver after the 200th iteration. This example demonstrates that real-life source reconstruction problems may easily entail a huge computational burden with thousands of forward problem solves. Another observation is that the role of α is twofold: it controls the performance of the iterative solver but also impacts the amount of regularization applied to the solution. Frequently, α is set to a high value solely to reduce computational load, but it may compromise the accuracy of the reconstruction.

III. THE OPTIMAL CONTROL FRAMEWORK

Let us describe our approach to the source identification problem. We formulate (2) as the following linear-quadratic optimal control problem,

$$\underset{f,u}{\text{minimize}} \quad \frac{1}{2} \sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i u - d_i)^2 + \frac{\alpha}{2} \int_{\Omega} |\mathcal{F}\nabla f|^2 dV, \quad (15.1)$$

subject to

$$-\operatorname{div}(\sigma \nabla u) - f = 0, \text{ in } \Omega, \quad (15.2)$$

$$\frac{\partial u}{\partial \nu} = g_N, \text{ on } \Gamma_N, \quad (15.3)$$

$$u = g_D, \text{ on } \Gamma_D. \quad (15.4)$$

Here the electric potential u represents the state variable, the right-hand side f is the control variable, and d (measured data) is the target variable. Operators \mathcal{Q}_i are defined by (3).

Next, we form the Lagrangian function as follows,

$$\begin{aligned} \mathcal{L}(u, f, \lambda) = & \frac{1}{2} \sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i(u) - d_i)^2 + \\ & \frac{\alpha}{2} \int_{\Omega} |\mathcal{F}\nabla f|^2 dV + \int_{\Omega} \lambda (-\operatorname{div}(\sigma \nabla u) - f) dV, \end{aligned} \quad (16)$$

where λ is the adjoint state variable. At this point the Lagrangian includes constraint (15.2), but not (15.3) nor (15.4). In view of the Green-Gauss theorem and constraint (15.3) the following equality holds,

$$\begin{aligned} \int_{\Omega} \lambda \operatorname{div}(\sigma \nabla u) dV = & - \int_{\Omega} \sigma \nabla u \cdot \nabla \lambda dV \\ & + \int_{\Gamma_D} \sigma \frac{\partial u}{\partial \nu} \lambda dS + \int_{\Gamma_N} \sigma g_N \lambda dS \end{aligned} \quad (17)$$

The first surface integral in (17) vanishes if we demand $\lambda = 0$ on Γ_D . In addition, we choose $u = g_D$ on Γ_D to satisfy constraint (15.4). That is, we set $u \in V$, $\lambda \in V_0$ (see definitions in Section II)

Summarizing the discussion above, we set up the following Lagrangian function,

$$\begin{aligned} \mathcal{L}(u, f, \lambda) = & \frac{1}{2} \sum_{i=0}^{K-1} w_i^2 (\mathcal{Q}_i(u) - d_i)^2 + \frac{\alpha}{2} \int_{\Omega} |\mathcal{F}\nabla f|^2 dV \\ & + \int_{\Omega} \sigma \nabla u \cdot \nabla \lambda dV - \int_{\Omega} \lambda f dV - \int_{\Gamma_N} \sigma g_N \lambda dS. \end{aligned} \quad (18)$$

The necessary conditions for a triple (f, λ, u) be the optimal point of \mathcal{L} is that the weak derivatives vanish,

$$\begin{aligned} \mathcal{L}'_f \varphi &= 0, \forall \varphi \in V_0, \\ \mathcal{L}'_\lambda \mu &= 0, \forall \mu \in V_0, \\ \mathcal{L}'_u v &= 0, \forall v \in V_0. \end{aligned} \quad (19)$$

Thus, at the optimal point $(f, \lambda, u) \in V_0 \times V_0 \times V$, the following Euler-Lagrange system must be satisfied,

$$\alpha \int_{\Omega} \mathcal{F}\nabla f \cdot (\mathcal{F}\nabla \varphi) dV - \int_{\Omega} \lambda \varphi dV = 0 \quad \forall \varphi \in V_0, \quad (20.1)$$

$$\int_{\Omega} \sigma \nabla u \cdot \nabla \mu dV - \int_{\Omega} f \mu dV - \int_{\Gamma_N} \sigma g_N \mu dS = 0 \quad \forall \mu \in V_0, \quad (20.2)$$

$$\sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(v) (\mathcal{Q}_i(v) - d_i) + \int_{\Omega} \sigma \nabla \lambda \cdot \nabla v dV = 0, \quad \forall v \in V_0. \quad (20.3)$$

We introduce the following bilinear forms,

$$\mathbb{A}(x, y) = \alpha \int_{\Omega} \mathcal{F}\nabla x \cdot (\mathcal{F}\nabla y) dV, \quad (21.1)$$

$$\mathbb{B}(x, y) = \int_{\Omega} xy dV, \quad (21.2)$$

$$\mathbb{E}(x, y) = \int_{\Omega} \sigma \nabla x \cdot \nabla y dV, \quad (21.3)$$

$$\mathbb{F}(x, y) = \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(x) \mathcal{Q}_i(y), \quad (21.4)$$

$$(21.5)$$

and linear forms,

$$\mathbb{R}(x) = \int_{\Gamma_N} g_N x dS, \quad (22.1)$$

$$\mathbb{S}(x) = \sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_i(x) d_i. \quad (22.2)$$

With these notations system (20) can be rewritten in a compact form,

$$\mathbb{A}(f, \varphi) - \mathbb{B}^t(\lambda, \varphi) = 0, \quad \forall \varphi \in V_0, \quad (23.1)$$

$$-\mathbb{B}(f, \mu) + \mathbb{E}^t(u, \mu) = \mathbb{R}(\mu), \quad \forall \mu \in V_0, \quad (23.2)$$

$$\mathbb{E}(\lambda, v) + \mathbb{F}(u, v) = \mathbb{S}(v), \quad \forall v \in V_0. \quad (23.3)$$

where superscript t means transposition.

IV. DISCRETIZATION

In this section, we discuss discretization of variational system (23) in finite-element spaces.

Domain Ω is divided into a set of tetrahedrons, Ω_h , where h stands for the maximal diameter of tetrahedrons. We utilized the classical nodal-based linear basis functions, which we will denote by ζ_i . Let Z_h be a space spanned by the basis functions. We define

$$\begin{aligned} V_h &= \{v_h \in Z_h : v_h = g_D \text{ on } \Gamma_D\}, \\ V_{h0} &= \{v_h \in Z_h : v_h = 0 \text{ on } \Gamma_D\}. \end{aligned} \quad (24)$$

Thus, system (20) is approximated by the following finite-dimensional system,

$$\begin{aligned} \alpha \int_{\Omega_h} \mathcal{F}\nabla f_h \cdot (\mathcal{F}\nabla \varphi_h) dV - \int_{\Omega_h} \lambda_h \varphi_h dV &= 0, \\ \forall \varphi_h \in V_{h0}, \end{aligned} \quad (25.1)$$

$$\int_{\Omega_h} \sigma \nabla u_h \cdot \nabla \mu_h dV - \int_{\Omega} f_h \mu dV - \int_{\Gamma_N} \sigma g_N \mu_h dS = 0, \quad \forall \mu \in V_{h0}, \quad (25.2)$$

$$\sum_{i=0}^{K-1} w_i^2 \mathcal{Q}_{hi}(v_h) (\mathcal{Q}_{hi}(v_h) - d_i) + \int_{\Omega_h} \sigma \nabla \lambda_h \cdot \nabla v_h dV = 0, \quad \forall v \in V_{h0}. \quad (25.3)$$

Here \mathcal{Q}_{hi} is a finite-dimensional counterpart of \mathcal{Q}_i . Operator \mathcal{Q}_{hi} maps arbitrary function $v_h \in Z_h$ to its value at the position of the i -th electrode, that is, $\mathcal{Q}_{hi} : Z_h \rightarrow \mathbb{R}$. Thus, it can be represented as a weighted sum of basis functions,

$$d_i = \mathcal{Q}_{hi} := \sum_{j=0}^{N-1} q_j \zeta_j(x_i), \quad (26)$$

where q_i are the expansion coefficient, $\zeta_i(x_i)$ are the values of the basis functions at the position of the i -th electrode. Each coefficient q_i depends only on the position of the i -th electrode inside a tetrahedron and the geometry of that tetrahedron. If x_i is located exactly on the j -th grid node then (26) simplifies to $d_i = \zeta_j(x_i)$, so that $\mathcal{Q}_{hi} \in \mathbb{R}^N$ is a row vector containing 1 at the j -th position, and zeros at the other positions.

In the standard fashion, we expand f_h, u_h , and λ_h in functions ζ_i ,

$$f_h = \sum_{i=0}^{N-1} f_i \zeta_i, \quad \lambda_h = \sum_{i=0}^{N-1} \lambda_i \zeta_i, \quad u_h = \sum_{i=0}^{N-1} u_i \zeta_i, \quad (27)$$

where f_i, λ_i , and u_i are the expansion coefficients. Substituting (27) to (25) and recalling that space Z_h is finite-dimensional, we get the following system of linear equations,

$$\sum_{j=0}^{N-1} f_j \alpha \int_{\Omega_h} \mathcal{F} \nabla \zeta_j \cdot \nabla \zeta_i dV - \sum_{j=0}^{N-1} \lambda_i \int_{\Omega_h} \zeta_j \zeta_i dV = 0, \quad i \in [0..N-1], \quad (28.1)$$

$$-\sum_{j=0}^{N-1} f_i \int_{\Omega_h} \zeta_j \zeta_i dV + \sum_{j=0}^{N-1} \int_{\Omega_h} \sigma \nabla \zeta_j \cdot \nabla \zeta_i dV \\ = \int_{\Gamma_N} \sigma g_N \zeta_i dS, \quad i \in [0..N-1], \quad (28.2)$$

$$\sum_{j=0}^{N-1} \left(\sum_{k=0}^{K-1} w_k^2 \mathcal{Q}_{hk}(\zeta_i) \mathcal{Q}_{hk}(\zeta_j) \right) + \sum_{j=0}^{N-1} \lambda_j \int_{\Omega_h} \sigma \nabla \zeta_j \cdot \nabla \zeta_i dV \\ = \sum_{k=0}^{K-1} w_k^2 \mathcal{Q}_{hk}(\zeta_i) d_k, \quad i \in [0..N-1]. \quad (28.3)$$

Now we introduce the shorthand notations,

$$A_{ij} = \alpha \int_{\Omega} \mathcal{F} \nabla \zeta_j \cdot (\mathcal{F} \nabla \zeta_i) dV, \quad (29.1)$$

$$B_{ij} = \int_{\Omega} \zeta_j \zeta_i dV, \quad (29.2)$$

$$E_{ij} = \int_{\Omega} \sigma \nabla \zeta_j \cdot \nabla \zeta_i dV, \quad (29.3)$$

$$F_{ij} = Q^T W^T W Q, \quad (29.4)$$

$$r_i = \int_{\Gamma_N} \sigma g_N \zeta_i dS, \quad (29.5)$$

$$s_i = Q^T W^T W d. \quad (29.6)$$

Thus, we rewrite system (28) in a compact form,

$$\underbrace{\begin{bmatrix} A & -B^T & O \\ -B & O & E^T \\ O & E & F \end{bmatrix}}_{\mathcal{A}} \underbrace{\begin{bmatrix} \tilde{f} \\ \tilde{\lambda} \\ \tilde{u} \end{bmatrix}}_{\xi} = \underbrace{\begin{bmatrix} 0 \\ r \\ s \end{bmatrix}}_b. \quad (30)$$

We exclude the essential DOFs from the system (30). That is, we remove equations that correspond to the expansion coefficients with predefined values, and move corresponding elements of the summations to the right-hand side. Here and below, we assume that (30) is the reduced system with the essential DOFs excluded, so its size is $3n \times 3n$, where $n \leq N$ is the number of non-essential DOFs. Vectors \tilde{f} , $\tilde{\lambda}$, and \tilde{u} contains non-essential values of f_i, λ_i , and u_i , respectively. Matrices A, B, E , and F correspond to bilinear forms $\mathbb{A}, \mathbb{B}, \mathbb{E}$, and \mathbb{F} , respectively. Matrix O is a zero block.

System (30) is known as the Karush-Kuhn-Tucker (KKT) system of linear equations. All blocks in \mathcal{A} are sparse $\mathbb{R}^{n \times n}$ matrices. Block A is a positive-definite stiffness matrix. Note that, if we change the stabilization term in (15) to $\int_{\Omega} |\mathcal{F} f|^2 dV$, then A turns into a mass matrix. Block E is a symmetric positive-definite stiffness matrix. Block F being symmetric positive-semidefinite, $\text{rank}(F) \leq K$, where K is number of data points (we assume that there are no duplicated data). If data are sampled at K different nodal points and $W = I$, then F has an especially simple structure: it is a diagonal matrix with K non-zero entries equal to 1.

V. ON THE EQUIVALENCE OF THE TWO APPROACHES

In this section, we show that the normal system (11) and the KKT system (30) provide exactly the same value of \tilde{f} , but achieve it by solving two different systems of linear equations. In different scenarios, one or the other approach appears to be more beneficial in terms of computing resources.

As in Section II, we assume that $g_N = 0$ to make the presentation easier. Let us rewrite (30) as the three equations,

$$A \tilde{f} - B^T \tilde{\lambda} = 0, \quad (31.1)$$

$$-B \tilde{f} + E^T \tilde{u} = 0, \quad (31.2)$$

$$E \tilde{\lambda} + Q^T W^T W Q \tilde{u} = Q^T W^T W d. \quad (31.3)$$

We eliminate $\tilde{\lambda}$ from the system using (31.3). The result is

$$A \tilde{f} + B^T E^{-1} Q^T W^T W Q \tilde{u} = B^T E^{-1} Q^T W^T W d, \quad (32.1)$$

$$-B\tilde{f} + E^T\tilde{u} = 0, \quad (32.2)$$

Upon eliminating \tilde{u} by using (32.2) we obtain

$$(B^T E^{-1} Q^T W^T W Q E^{-1} B + A) \tilde{f} = B^T E^{-1} Q^T W^T W d. \quad (33)$$

In view of definition (14) and the fact that $\alpha L = A$ we see that (33) is exactly (11).

The fact that the two approaches deliver the same estimation of f means that the questions such as the choice of the regularization parameter, quality of reconstruction, impact of the noise, required density of data sampling, and others have already been addressed in the geophysical literature. It does not mean, however, that the two approaches are identical. The difference between them is in the linear systems, (11) or (30). They require different computing resources in different scenarios. The KKT system is beneficial when the inverse is large and an efficient preconditioner is available to take advantage of the iterative solution of a sparse system.

VI. SOLUTION OF THE SYSTEM OF LINEAR EQUATIONS

System (30) can be solved either directly or iteratively. Advantages of direct sparse solvers include the existence of reliable off-the-shelf packages and a weak dependence (up to a certain limit) of their performance on spectral characteristics of a matrix. Common disadvantages of sparse direct solvers are a long initialization time, high memory consumption, and mediocre parallel scalability. It should be noted that indefinite problems are less stable compared to positive-definite ones and more time-consuming. Also, the performance may vary considerably between various implementations, see review [28], although it represents a snapshot of this field almost 15 years ago. In addition, since matrix $J^*J + \alpha L$ is dense and \mathcal{A} is sparse, troubles with the performance of direct methods will arise for much smaller problems in the former case. All in all, direct sparse solvers are a viable choice for medium-sized problems, where the number of unknowns is hundreds of thousands on a shared-memory machine.

Iterative solvers have to be applied when the problem is large and full factorization of \mathcal{A} is prohibitively costly. Even for modest problems, they may deliver a solution faster than direct solvers. Iterative solvers, especially those based on the short recurrence, require limited memory. The performance of iterative methods critically depends on the preconditioning of a system matrix. Blocks 3×3 systems of linear equations appear in various contexts, e.g. [31, 32]. Preconditioning strategies of such systems are highly dependent on the particular appearance of that matrix and the properties of its blocks. This part of our research has not been finished yet and will be presented in a dedicated paper. We briefly outline some early results here.

The natural choice for (30) is MINRES [33], since \mathcal{A} is symmetric indefinite. MINRES should be accelerated with a positive-definite preconditioner. However, in all the cases we tested, the performance of a preconditioned MINRES was relatively poor. Instead, we use a block-triangular preconditioner, which is among the most efficient accelerators available (see

a review [34]). Specifically, we applied a left block-triangular preconditioner in the following form,

$$\mathcal{P} = \begin{bmatrix} A & O & O \\ -B & -S & O \\ O & E & R \end{bmatrix}, \quad (34)$$

where $S = BA^{-1}B^T$, $R = ES^{-1}E^T + F$. Note that mass matrix B is lumped in the definition of S . Thus, we solve the following preconditioned system,

$$\mathcal{P}^{-1}\mathcal{A}\xi = \mathcal{P}^{-1}b. \quad (35)$$

System (35) is non-symmetric, so a non-symmetric Krylov solver has to be applied. We employed GMRes [33], although BiCGStab, QMR, and some other solvers could also be used. The application of GMRes to (34) leads to some overhead in memory and operations comparing to MINRES applied to the symmetric system. It, however, pays off by the much faster convergence.

In this study, we use preconditioner (34), implemented as three consecutive linear solves with blocks A , S , and R . That is, $x = \mathcal{P}^{-1}y$ is computed via the three consecutive solves,

$$Ax_1 = y_1, \quad (36a)$$

$$Sx_2 = -y_2 - Bx_1, \quad (36b)$$

$$Rx_3 = y_3 - Ex_2, \quad (36c)$$

where the indices at x and y refer to the partitioning of the linear system. We solved systems (36) exactly with the sparse LU factorization.

It should be noted that since we used exact block solvers, GMRes is expected to converge in a single iteration. This is because the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ has all eigenvalues equal to 1. However, in practice, each of the three systems is solved with some round-off error. Accumulation of errors combined with a high condition number of block R leads to an increased number of iterations. The single-iteration convergence can be observed numerically only by explicit factorization of matrix \mathcal{P} , which has no advantage compared to the direct solution of problem (30).

Our implementation is on C++ with the use of MFEM finite-element package [35]. We utilized multi-frontal direct solver UMFPACK [36]. All numerical experiments reported below were performed on a laptop with Intel Core i3-6100 CPU @ 2.70 GHz equipped with 8Gb RAM.

VII. NUMERICAL EXPERIMENTS

We have selected a synthetic example that demonstrates the use of our algorithm with an irregular acquisition system on a geological model with non-flat topography. The computational domain had dimensions $200 \times 200 \times 100$ m. We have created grids, referred to as the small and big one. The small grid contained 150K tets; the big grid contained 376K tets see (Figure 2). On the top of each grid, we specified condition $\frac{\partial u}{\partial v} = 0$, whereas on the rest of the boundary we set up $u = 0$ (see a comment about the "oblique" Dirichlet boundary conditions in Introduction). The number of discrete unknowns was 65K and 167K in the small and big problems, respectively. The source term, f , is depicted in Figure 3(a). It consists

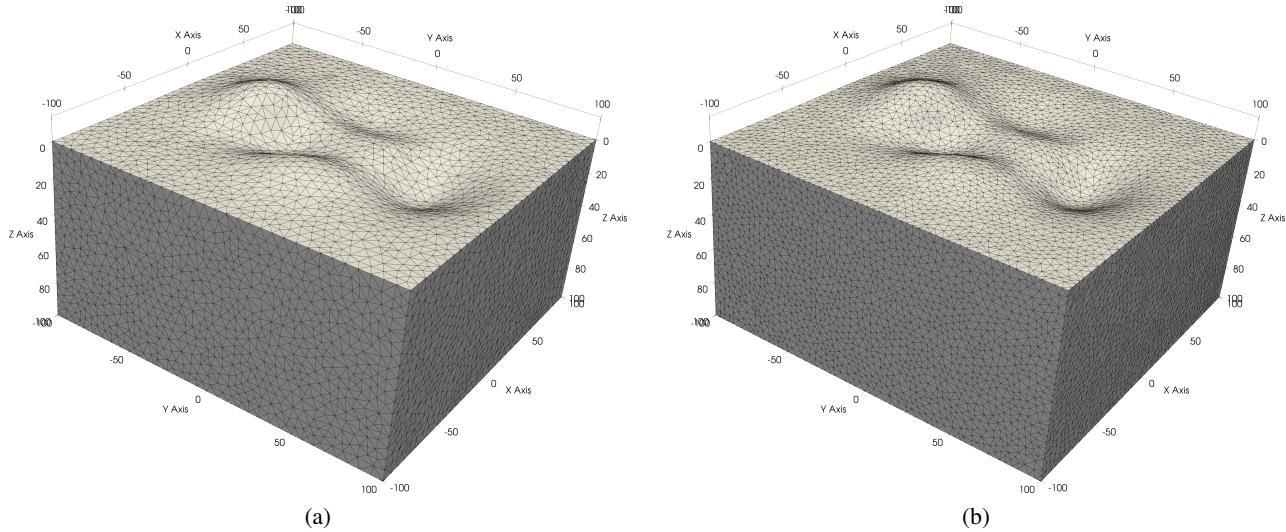


Figure 2: Computational grids used in the numerical experiments. The computational domain has overall physical size of $200 \times 200 \times 100$ m. The top surface represents hills and ditches with the altitude ranging from -13 to 16 m. Panel (a) depicts the small grid containing 150K tetrahedrons. Panel (b) depicts the big grid with 376K tetrahedrons.

of the two smooth distributions of opposite sign centered at $(0, \pm 30, 30)$ m.

The synthetic data were generated using FEM, based on discretization, presented in Section II. The forward solver was implemented in C++ with MFEM package. We used a uniform conductivity, $\sigma = 1$, although our forward and inverse codes support a variable conductivity. Real acquisition grids usually follow more or less regular patterns, but in this test, we randomly scattered 100 receivers in a disc of a radius of 90 m to have even data coverage. The simulated electric potential in electrodes is depicted in Figure 3(b). The synthetic data were noise free.

We solved the two inverse problems with $\alpha = 10^{-5}$. This choice was not optimal in any sense, so we reiterate that any standard algorithm for choosing α [8][Section 2.5] will work equally well for our approach (see Section V). For each model, we compared UMFPACK direct solver and a preconditioned GMRes. Some statistics about the linear system solution is given in Tables II and I. The *Memory* column in the tables refers to the peak memory allocated during the solution of the linear system. These data are based on Linux process high-level statistics. Although these could not be as accurate as a memory profiling tool, they give valuable insight into memory consumption. The computations were performed in the serial mode. Both time and memory consumption were dominated by the LU factorizations either within the direct solver or preconditioner.

Table I: Performance of UMFPACK direct sparse solver on the two test problems.

Problem	Num. unknwns	Time, s	Memory, Gb
Small	65,469	62	2.0
Big	167,052	416	6.3

The convergence of the GMRes solver is given in Figure 4. Surprisingly, the system had to be solved with high accuracy to provide a reasonable source reconstruction matching the one

Table II: Performance of the preconditioned GMRes. The restart parameter was set to 5 and 10 for the small and big problems, respectively.

Problem	Num. unknwns	Total iters	Time, s	Memory, Gb
Small	65,469	14	33	1.3
Big	167,052	24	233	3.9

obtained with the direct solver. We set the solver to reduce the normalized misfit down to 10^{-13} . We attribute this to the fact that at arbitrary iteration n , the three components of the residual, $r_n = \mathcal{A}\xi_n - b$, are of very different scales and get reduced with different rates during iterative updating. This contrasts with the standard approach, where only data misfit is reduced, and the discretized Poisson's equation $Au = f$ is satisfied almost precisely via the Jacobian.

Reconstructed distribution of f were almost identical for all cases, see Figure 5. The positive and negative extrema essentially coincide with the location of the true source and sink. This result is expected because the data are synthetic, the number of measurements is satisfactory, and they are evenly cover the area. At the same time, the amplitude of the reconstructed f is an order of magnitude smaller compared with the true one. We emphasize that this result would be equal to that obtained of standard minimization (13), provided that parameter α , the computational grid, input data, and other parameters are the same and the linear solves are obtained with infinite precision. Thus, it is not a drawback of our approach, but typical behavior of the standard inversion with the "Laplacian" constraints. The main idea behind this calculation is not that our reconstruction is better from the geophysical standpoint (i.e., better resembles the true anomaly), but that it is obtained in a fundamentally different way, which may be better computationally in some scenarios.

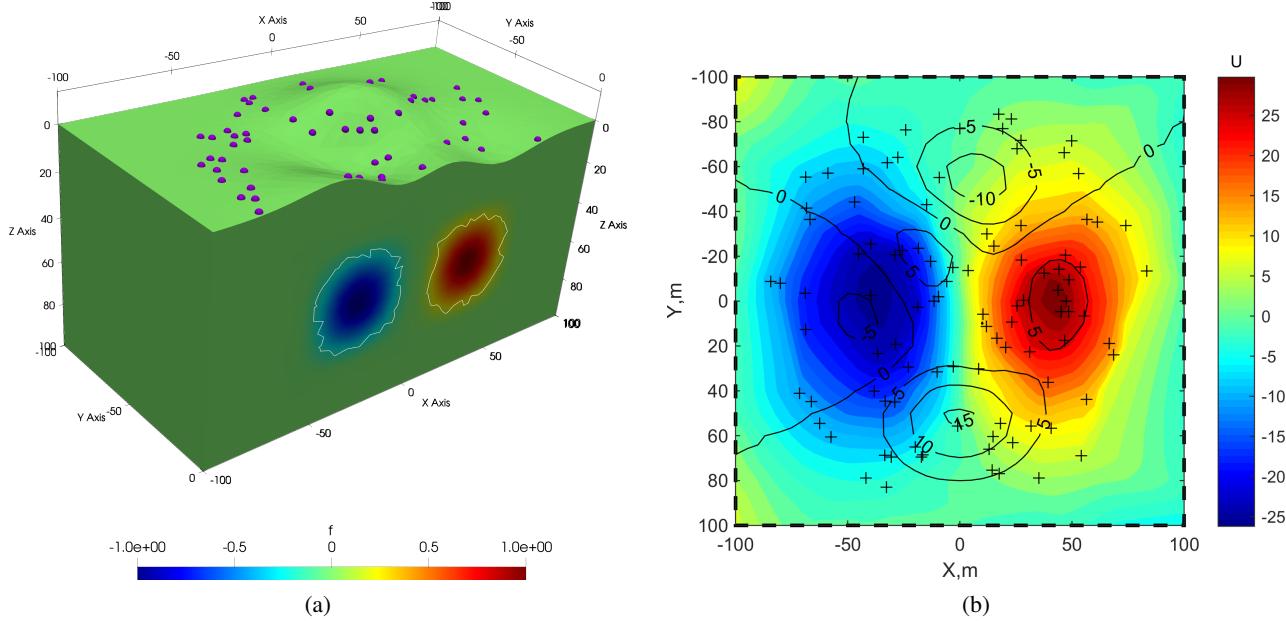


Figure 3: (a) The true source term f , in Am^{-3} , that was used to simulate input data. The two distributions of the opposite signs are placed 60 m apart at a depth of 30 m. The white lines mark level sets at $\pm 0.1 \text{ Am}^{-3}$. (b) The simulated electric potential is shown in color. The crosses mark receiver locations. The topography is shown by the black lines.

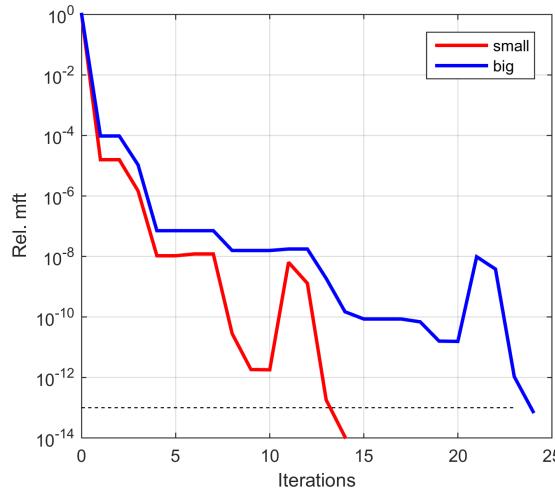


Figure 4: Convergence of the preconditioned GMRes solver at the two test problems. The restart parameter was set to 5 and 10 for the small and big problems, respectively.

VIII. DISCUSSION AND CONCLUSIONS

We have developed a novel approach for source reconstruction in the self-potential method. Derivation of the variational problem and its discrete formulations for the finite-element method are given. The main advantages of this approach are that the resulting system of linear equations is sparse, and its blocks are associated with the PDE. Thus, the system can be solved either by a direct sparse solver (if the problem is medium-size) or by a preconditioned iterative solver (for larger problems). We have proposed a preconditioned iterative method to solve the system of linear equations. The numerical experiments suggested that our approach works reasonably

well. Both the direct and iterative solvers provided almost identical solutions to the inverse problems in a reasonable time. In the cases we tested, the iterative solution was twice as fast as the direct one and consumed 30% less memory. The optimal control framework might be a viable alternative to the standard inversion, especially for large-scale problems. We believe that it will lead to a new generation of efficient inversion algorithms for processing real data.

Let us comment on essential aspects of the proposed approach and future research.

First of all, the study of preconditioners should be enhanced further. We employed the exact block solvers based on the sparse LU factorization. A natural extension would be to study the incomplete LU (ILU) factorization. It constitutes one of the best-known classes of preconditioners. Its application will inevitably slow down the convergence of iterative solvers compared to the exact block solver, but the computational load will be less. The combined effect may be or may not be a reduction of the overall runtime. It is known that the ILU preconditioner is not spectrally equivalent to the original matrix. It, however, would be desirable to have a method whose performance is independent of the grid size. The multigrid should be studied in this context, although it will be a very different approach than the ILU.

Higher-order elements can be employed to allow for coarser grids. Their applicability is justified because f, u , and λ are very smooth in practice. The use of coarse grids may improve the algorithm's performance, although some overhead is inevitable due to an increased number of non-zero elements in the blocks. The numerical implementation will be more involving, but almost certainly, most programming efforts have already been made in industrial FEM packages.

The role of the noise should be investigated further. At first

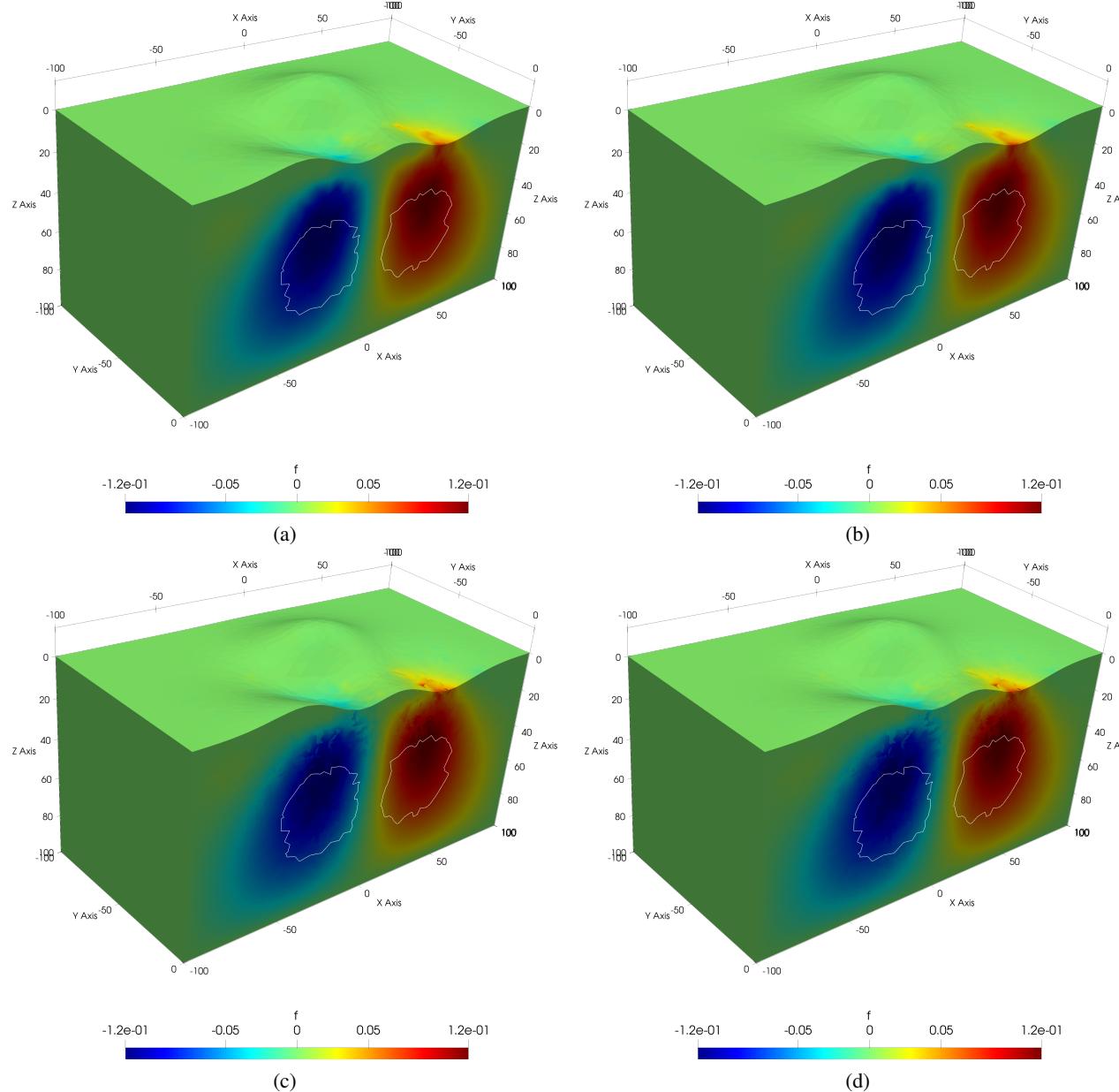


Figure 5: Reconstructed source distributions, f . (a) - The small model, direct solver. (b) - The small model, iterative solver. (c) - The big model, direct solver. (d) - The big model, iterative solver. The white lines mark level sets of the true source function f at ± 0.1 . The slight differences between the small and big grids are due to the difference in spatial sampling.

glance, introducing noise to the data will not give helpful information since both systems (13), and (30) deliver identical \hat{f} , and the data weights W have already appeared in the system matrices. There is, however, a subtle difference when it comes to the iterative solution. In the standard approach, f and u are tightly linked because they both satisfy the discrete Poisson's equation through the Jacobian. Only the data misfit decreases during iterative updating, so the residual norm is a natural criterion for the termination of the iterative process. In our approach, the system matrix consists of three systems. It is unknown beforehand which of these will be satisfied earlier. It may happen that in early iterates, the computed and measured data match each other with high accuracy, whereas,

say, Poisson's equation is still not met. We suspect that it is the reason why the system should be solved with very high accuracy to deliver valuable results. This phenomenon certainly deserves further investigation.

We note that errors in σ strongly influence the reconstruction of f . This phenomenon is well known empirically, for example, [10]. To our knowledge, one of the most systematic studies of this topic, both theoretical and numerical, is [37][Chapter 4]. The common consensus is that accurate reconstruction of σ , based on the inversion of a separate electrical tomography survey, is a precursor to the high-quality retrieval of f .

Finally, we would like to comment on the much higher

accuracy of the reconstruction, reported in [38]. They sought the electric potential on the inaccessible part of the boundary (the borehole casing) given known values of the potential and its normal derivative on the ground surface, and $f = 0$. Such a problem is the Cauchy problem for Poisson's equation. It has a unique theoretical solution but is unstable with respect to the data (see discussion in [39]). This problem is still ill-posed, but the theoretical uniqueness of the solution translates to a higher accuracy comparing to the problem of reconstructing the right-hand side f .

IX. ACKNOWLEDGEMENTS

This study was supported by the Russian Science Foundation, project no. 21-11-00139. The authors acknowledge computational resources granted by Complex for Simulation and Data Processing for Mega-Science Facilities at NRC "Kurchatov Institute", <http://ckp.nrcki.ru/>; and Skoltech CDISE's high-performance computing cluster, Zhores [40].

REFERENCES

- [1] A. Revil and A. Jardani, *The Self-Potential Method: Theory and Applications in Environmental Geosciences*. Cambridge University Press, 2013.
- [2] J. Castermant, C. Mendonça, A. Revil, F. Trolard, G. Bourrié, and N. Linde, "Redox potential distribution inferred from self-potential measurements associated with the corrosion of a buried metallic body," *Geophys. Prospecting*, vol. 56, pp. 269–282, 2008.
- [3] J. B. Rittgers, A. Revil, M. Karaoulis, M. A. Mooney, L. D. Slater, and E. A. Atekwana, "Self-potential signals generated by the corrosion of buried metallic objects with application to contaminant plumes," *Geophysics*, vol. 78, no. 5, pp. EN65–EN82, 2013.
- [4] Y. Bernabé and A. Maineult, "Physics of porous media: Fluid flow through porous media, (second ed.)." in *Treatise on geophysics (2nd ed.)*, G. Schubert, Ed. Oxford: Elsevier, 2015, p. 19–41.
- [5] L. Guerracino and D. Jougnot, "A physically based analytical model to describe effective excess charge for streaming potential generation in water saturated porous media." *J. Geophys. Res.*, vol. 123, p. 52–65, 2018.
- [6] L. Onsager, "Reciprocal relations in irreversible processes. I." *Phys. Rev.*, vol. 37, pp. 405–426, Feb 1931. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.37.405>
- [7] C. Rücker, T. Günther, and K. Spitzer, "Three-dimensional modelling and inversion of dc resistivity data incorporating topography - I. Modeling," *Geophysical Journal International*, vol. 166, pp. 495–505, 2006.
- [8] M. Zhdanov, *Inverse Theory and Applications in Geophysics*. Elsevier, 2015.
- [9] B. J. Minsley, J. Sogade, and F. D. Morgan, "Three-dimensional source inversion of self-potential data," *Journal of Geophysical Research*, vol. 112, p. B2202, 2007.
- [10] A. Jardani, A. Revil, A. Bolève, and J. Dupont, "Three-dimensional inversion of self-potential data used to constrain the pattern of groundwater flow in geothermal fields," *J. Geophys. Res.*, vol. 113, p. B09204, 2008.
- [11] A. Bolève, A. Revil, F. Janod, J. Mattiuzzo, and J.-J. Fry, "Preferential fluid flow pathways in embankment dams imaged by self-potential tomography," *Near Surface Geophysics*, vol. 7, pp. 447–462, 2009.
- [12] S. Ahmed, A. Jardani, A. Revil, and J. Dupont, "SP2DINV: A 2D forward and inverse code for streaming potential problems," *Computers & Geosciences*, vol. 59, pp. 9–16, 2013.
- [13] S. J. Ikard, A. Revil, M. Schmutz, M. Karaoulis, A. Jardani, and M. Mooney, "Characterization of focused seepage through an earthfill dam using geoelectrical methods," *Groundwater*, vol. 52, pp. 952–965, 2014.
- [14] Y. Li and D. Oldenburg, "3-d inversion of magnetic data," *Geophysics*, vol. 61, pp. 394–408, 1996.
- [15] O. Portniaguine and M. Zhdanov, "3-d magnetic inversion with data compression and image focusing," *Geophysics*, vol. 67, no. 5, pp. 1532–1541, 2002.
- [16] Y. Li and D. Oldenburg, "3-d inversion of gravity data," *Geophysics*, vol. 63, pp. 109–119, 1998.
- [17] M. Zhdanov, R. Ellis, and S. Mukherjee, "Three-dimensional regularized focusing inversion of gravity data gradient tensor component data," *Geophysics*, vol. 69, pp. 925–937, 2004.
- [18] M. Zhdanov, X. Liu, G. Wilson, and L. Wan, "Potential field migration for rapid imaging of gravity gradiometry data," *Geophysical prospecting*, vol. 59, pp. 1052–1071, 2011.
- [19] Z. Xu, L. Wan, and M. Zhdanov, "Focusing iterative migration of gravity gradiometry data acquired in the Nordcapp Basin, Barents Sea," *Geophysical prospecting*, vol. 68, no. 7, pp. 2292–2306, 2020.
- [20] D. Zidarov, *Inverse gravimetric problem in geoprospecting and geodesy*. Elsevier Science, 1990.
- [21] A. Camacho, F. Montesinos, and R. Vieira, "Gravity inversion by means of growing bodies," *Geophysics*, vol. 65, no. 1, pp. 95–101, 2000.
- [22] M. Fedi and M. Pilkington, "Understanding imaging methods for potential field data," *Geophysics*, vol. 77, no. 1, pp. G13–G24, 2012.
- [23] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*. Springer, 2009.
- [24] W. L. Rodi, "A technique for improving the accuracy of finite element solutions for magnetotelluric data," *Geophysical Journal International*, vol. 44, no. 2, p. 483–506, 1976.
- [25] R.-E. Plessix, "A review of the adjoint-state method for computing the gradient of a functional with geophysical applications," *Geophysical Journal International*, vol. 167, no. 2, pp. 495–503, 2006.
- [26] E. Haber, U. Ascher, and D. Oldenburg, "On optimization techniques for solving nonlinear inverse problems," *Inverse Problems*, vol. 16, pp. 1263–1280, 2000.
- [27] E. Haber, U. M. Ascher, and D. W. Oldenburg, "Inversion of 3d electromagnetic data in frequency and time domain using an inexact all-at-once approach," *GEOPHYSICS*, vol. 69, no. 5, pp. 1216–1228, 2004.

- [28] N. I. M. Gould, J. A. Scott, and Y. Hu, "A numerical evaluation of sparse direct solvers for the solution of large sparse symmetric linear systems of equations," *ACM Transactions on Mathematical Software*, vol. 33, no. 2, p. art. no. 1236465, 2007.
- [29] T. A. Davis, S. Rajamanickam, and W. M. Sid-Lakhdar, "A survey of direct methods for sparse linear system," *Acta Numerica*, pp. 382–566, 2016.
- [30] T. Bonesky, "Morozov's discrepancy principle and Tikhonov-type functionals," *IOP Inverse Problems*, vol. 25, p. 015015, 2009.
- [31] S. Rhebergen, G. N. Wells, A. J. Wathen, and R. F. Katz, "Three-field block preconditioners for models of coupled magma/mantle dynamics," *SIAM Journal on Scientific Computing*, vol. 37, no. 5, pp. A2270–A2294, 2015.
- [32] 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*, vol. to appear, 2021.
- [33] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. Society for Industrial and Applied Mathematics, 1994.
- [34] F. P. Ali Beik and M. Benzi, "Iterative methods for double saddle point systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 39, no. 2, pp. 902–921, 2018.
- [35] R. Anderson, J. Andrej, A. Barker, J. Bramwell, J.-S. Camier, J. Cerveny, V. Dobrev, Y. Dudouit, A. Fisher, T. 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*, vol. 81, pp. 42–74, 2021.
- [36] T. A. Davis, "Algorithm 832: Umfpack v4.3—an unsymmetric-pattern multifrontal method," *ACM Trans. Math. Softw.*, vol. 30, no. 2, p. 196–199, Jun. 2004. [Online]. Available: <https://doi.org/10.1145/992200.992206>
- [37] B. J. Minsley, "Modeling and inversion of self-potential data," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.
- [38] A. Maineult, "Estimation of the electrical potential distribution along metallic casing from surface self-potential profile," *Journal of Applied Geophysics*, vol. 129, pp. 66–78, 2016.
- [39] 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*, vol. 68, no. 6, pp. 1811–1819, 2021.
- [40] I. Zacharov, R. Arslanov, M. Gunin, D. Stefonishin, A. Bykov, S. Pavlov, O. Panarin, A. Maliutin, S. Rykovyanov, and M. Fedorov, "Zhores—Petaflops supercomputer for data-driven modeling, machine learning and artificial intelligence installed in Skolkovo Institute of Science and Technology," *Open Eng.*, no. 9, pp. 512–520, 2019.



Mikhail S. Malovichko received the B.C. and M.S. degrees in geophysics from St. Petersburg University, St. Petersburg, Russia, and the Ph.D. degree in technical sciences from Lomonosov Moscow State University, Russia, in 2004, 2006, and 2014, respectively. From 2008 to 2014, he was with JSC EMMET, a Russian subsidiary of CGG, analyzing electromagnetic data of marine geophysics and developing data processing software. Since 2014, he is a Research Scientist in the Applied Computational Geophysics Lab at Moscow Institute of Physics and Technology, Dolgoprudny, Russia. In 2018 he joined the Center for Data-Intensive Science and Engineering at Skolkovo Institute of Science and Technology, Moscow, Russia. His research interest includes PDE-based inverse problems, mathematical modeling, imaging methods, optimization, and high-performance computing with applications in medical imaging and geophysics.



Andrey V. Tarasov received the B.S. degree in geology and applied geophysics, the M.S. degree in applied geophysics, and the Ph.D. degree in applied geophysics (the induced polarization method) from St. Petersburg University, Russia, in 1997, 1999, and 2008, respectively. From 1999 to 2005, he was a Research Associate with the Impulse Electromagnetic Methods Laboratory at the Russian State Institute of Exploration Geophysics (VIRG-Rudgeofizika), St. Petersburg, Russia. From 2005 to 2010, he was with ELGEO Ltd., an engineering research company,

developing software for the transient electromagnetic (TEM) and induced polarization (IP) methods. Since 2009, he has been an Associate Professor with the Geophysics Department, St. Petersburg University. His research interests include theory and applications of the induced polarization phenomena and various geophysical methods for mineral exploration.



imaging.

Nikolay B. Yavich received the Ph.D. degree in mathematics from the University of Houston, Houston, TX, USA, in 2009. From 2009 to 2014 he worked in the entities of Fugro and CGG. In 2014 he joined the Applied Computational Geophysics Lab of Moscow Institute of Physics and Technology, Russia. Since 2018, he is with Skolkovo Institute of Science and Technology, Russia. His research interests include large-scale numerical modeling, high-performance computing, fast solvers, inverse problems, computational geophysics, and biomedical



Konstantin V. Titov received the engineer-geophysicist degree from Leningrad Mining Institute, USSR, in 1986, the Ph.D. degree in geophysics from St. Petersburg Mining Institute, Russia, in 1993, and the Habilitation degree in geophysics from St. Petersburg State University, in 2004. He was a Research Engineer and an Assistant Professor at Leningrad (St. Petersburg) Mining Institute from 1986 to 1994 and an Associate Professor from 1994 to 1999. Starting from 1999, he moved to the Russian Institute of Exploration Geophysics, where

he was a Project Leader till 2003. In 2004, he joined St. Petersburg State University as a Professor of Geophysics. Prof. Titov is also the Director of the Department of Geophysics at St. Petersburg State University. His research interests cover geo-electrical methods.