

BRIDGING AND IMPROVING THEORETICAL AND COMPUTATIONAL ELECTRICAL IMPEDANCE TOMOGRAPHY VIA DATA COMPLETION*

TAN BUI-THANH[†], QIN LI[‡], AND LEONARDO ZEPEDA-NÚÑEZ[†]

Abstract. In computational PDE-based inverse problems, a finite amount of data is collected to infer unknown parameters in the PDE. In order to obtain accurate inferences, the collected data must be informative about the unknown parameters. How to decide which data is most informative and how to efficiently sample it is the notoriously challenging task of optimal experimental design (OED). In this context, the best, and often infeasible, scenario is when the full input-to-output (ItO) map, i.e., an infinite amount of data, is available: This is the typical setting in many theoretical inverse problems, which is used to guarantee the unique parameter reconstruction. These two different settings have created a gap between computational and theoretical inverse problems, where finite and infinite amounts of data are used, respectively. In this article we aim to bridge this gap while circumventing the OED task. This is achieved by exploiting the structures of the ItO data from the underlying inverse problem, using the electrical impedance tomography (EIT) problem as an example. To accomplish our goal, we leverage the rank structure of the EIT model and formulate the ItO matrix—the discretized ItO map—as an $\mathcal{H}$ -matrix whose off-diagonal blocks are low rank. This suggests that, when equipped with the matrix completion technique, one can recover the full ItO matrix, with high probability, from a subset of its entries sampled following the rank structure: The data in the diagonal blocks is informative and should be fully sampled, while data in the off-diagonal blocks can be subsampled. This recovered ItO matrix is then utilized to present the full ItO map up to a discretization error, paving the way to connect with the problem in the theoretical setting where the unique reconstruction of parameters is guaranteed. This strategy achieves two goals: (I) *it bridges the gap between the finite- and infinite-dimensional settings for numerical and theoretical inverse problems* and (II) *it improves the quality of computational inverse solutions*. We detail the theory for the EIT model and provide numerical verification to both EIT and optical tomography problems.

Key words. inverse problems, EIT, data completion, matrix completion

AMS subject classifications. 15A83, 35R30, 68Q25

DOI. 10.1137/21M141703X

1. Introduction. Inverse problems—inferring unknown parameters in physical systems from indirect observations—are ubiquitous in engineering and all branches of sciences. The development of a deep theoretical understanding [50, 42, 51] coupled with the development of highly sophisticated algorithmic pipelines [53, 41, 52] for

*Submitted to the journal's Computational Methods in Science and Engineering section May 3, 2021; accepted for publication (in revised form) January 20, 2022; published electronically June 1, 2022.

<https://doi.org/10.1137/21M141703X>

Funding: The work of the first author was partially supported by the National Science Foundation through grants NSF-2108320, NSF-1808576, and NSF-CAREER-1845799; by the Defense Threat Reduction Agency award DTRA-M1802962; by the Department of Energy award DE-SC0018147; by a 2020 ConTex award; and by 2018 UT-Portugal CoLab award. The work of the second author was partially supported by the UW-Madison Data Initiative, the Vilas Young Investigation Award, and the National Science Foundation under grant DMS-1750488. The work of the third author partially supported by the National Science Foundation under grant DMS-2012292. The work of the second and third authors was supported by NSF TRIPODS award 1740707.

[†]Department of Aerospace Engineering and Engineering Mechanics, The Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712-1221 USA (tanbui@oden.utexas.edu, <https://users.odin.utexas.edu/~tanbui/>).

[‡]Mathematics Department and Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI 53705 USA (qinli@math.wisc.edu, <http://www.math.wisc.edu/~qinli/>, lzepeda@math.wisc.edu, <https://www.math.wisc.edu/~lzepeda/>).

solving inverse problems has fueled several breakthroughs in a myriad of different fields such as geophysics, astronomy, biomedical imaging, radar, spectrography, signal processing, and communications, among many others [17, 55, 9]. Several of such advances, e.g., magnetic resonance imaging (MRI) [49], computerized tomography [43], and synthetic aperture radar [17], permeate the modern life, thus making the study of inverse problems a subject of paramount importance at both theoretical and algorithmic levels.

In both theoretical and algorithmic formulations, the object that encodes the accessible knowledge of the unknown parameter is the input-to-output (ItO) map. Although the particular description of this map differs vastly depending on the modeling of the underlying physics, the ItO map generally encodes the impulse response (output) of the medium, or the parameters we seek to reconstruct, from a probing signal (input). At the theoretical level, one assumes that this map is an operator, which maps a functional space of adequate probing signals to another functional space of the corresponding responses. At the practical and numerical level, there is only a finite number of possible probing signals that one can use, and the impulse response can only be sampled by a limited number of receivers, resulting in a finite amount of output data.

Although both theoretical and algorithmic studies seek to shed light on the mechanisms to infer unknown parameters, they are often not consistent with each other. The theoretical study of inverse problems has mainly focused on answering questions on the infinite-dimensional setting: supposing one knows the full ItO map, can the underlying unknown parameter, living in an infinite-dimensional function space, be uniquely and stably reconstructed? In a nutshell, this infinite-to-infinite approach relies on an infinite amount of data to reconstruct the parameter function, which itself has an infinite number of degrees of freedom [31, 54]. This infinite-dimensional setting is certainly computationally infeasible. Thus in algorithmic studies, one often focuses on designing algorithmic pipelines to perform the reconstruction on the finite-dimensional setting: A finite number of ItO measurements are given, and one is to infer the unknown parameters represented by finite-dimensional vectors through either PDE-constrained minimization or Bayesian inference [27, 37, 38].

It is reasonable to believe that the theory should provide guidance and theoretical guarantees for the algorithms' performance. In practice, however, due to the drastically different perspectives taken, as discussed above, they have mostly advanced in a disconnected manner. Indeed, the use of theoretical guarantees is not straightforward: When one translates the problem from the infinite-dimensional setting to a finite-dimensional one, a large amount of information is often lost. For example, from the theoretical perspective, we do not need to quantify the importance of each data pair since they will be all used. In reality, only a finite number of data pairs are practically available; thus we ought to select the ones that best inform the parameter reconstruction. Which data pairs are most informative is typically unknown unless certain optimal experimental design (OED) problems (see, e.g., [44, 19]) are solved. OED is, however, notoriously challenging and computationally expensive.

To bridge the gap between the theoretical study on the infinite-dimensional setting and the numerical study on the finite-dimensional setting and to maximally use the knowledge from theoretical results, it is necessary to understand the structure of the underlying problems so to identify (ideally a small number of) data pairs that are informative about the unknown parameters and then, again by exploiting the structure, to lift the information coded in the finite data pairs to the infinite-dimensional setting where existing theoretical results are applicable. We stress that finding a

finitely small amount of data is important for real-world applications where data is often expensive and potentially cumbersome to obtain.

In this article we initiate a line of work to achieve the goal of bridging the gap between theoretical and computational inverse problems.

In particular, we seek to exploit the structure of the ItO map, and hence the underlying physics of the problem under consideration, to select a subset of *informative* entries in the ItO matrix to complete the missing entries, thus recovering the full ItO matrix. We then take advantage of the completed ItO matrix in two aspects:

- (I) *bridging the gap between the finite- and infinite-dimensional settings, and*
- (II) *improving the quality of computational inverse solution.*

For (I), the completed ItO matrix is lifted to the ItO map in the infinite-dimensional setting where the unique reconstruction of the unknown parameter is guaranteed, which in turn ensures algorithmic convergence. For (II), the completed ItO data matrix is used to reconstruct the parameter through a minimization algorithm. Since completed data contains more information about the parameter than the originally incomplete one, completed data reaches an empirically more accurate inverse solution than its incomplete counterpart. Even without matrix completion, the incomplete but informative data facilitates more accurate reconstruction compared to using the same amount of data selected according to other sampling strategies.

Summary. As has been discussed, the pillars on top of which our research program lies are (a) the ability to leverage the structure of the underlying problem to sample only a fraction of data and to complete the missing data and (b) the availability of theoretical results of infinite-dimensional inverse problems. The actual executions are thus problem-specific. Due to its long history and the wealth of theoretical results we choose electrical impedance tomography (EIT) for this paper. The ItO map in this case is the Dirichlet-to-Neumann (DtN) map and its discretization, the DtN matrix, which can be accurately approximated by an $\mathcal{H}$ -matrix. This allows us to predict the relative importance of entries of the DtN matrix and sample them accordingly following the $\mathcal{H}$ -matrix partitioning. The selected entries are then used to uncover the missing ones through a matrix completion algorithm [46]. The accuracy of matrix completion algorithms depends on a few assumptions. In this case, the algorithms that we chose requires the to-be-reconstructed matrix to be of low rank,¹ delocalized, and decoherent. Upon the correct partitioning according to the guidance of the $\mathcal{H}$ -matrix structure, we have that these assumptions hold true for blocks of the DtN matrix, permitting a good use of the matrix completion algorithm to complete the full DtN matrix. Finally, the completed DtN matrix is lifted to the DtN map, allowing us to integrate the theoretical results [2, 48] on EIT to show the convergence and uniqueness of the reconstructed parameter with a high probability.

Outline. This paper is organized as follows: We present the whole bridging framework including the data selection process, finite element discretization, a matrix completion algorithm, and a generic parameter reconstruction in section 2. Rigorous results justifying the framework for the EIT problem are presented in section 3. Various numerical results, including showing improved inverse solutions using data completion, are presented in section 4 to validate our approach for both EIT and optical tomography problems. Section 5 concludes the paper with future works.

We point out that despite the matrix completion process being almost absent from the inverse problem literature, it was used to solve PDEs in the forward problems [34].

¹This requirement can be relaxed to approximately low rank when the singular values of the matrix decay fast enough.

Moreover, data-driven approaches have been widely used, of which the most notable example is the application of compressed sensing to MRI [20], which already has commercial applications [1].

2. Bridging framework driven by EIT. Throughout the paper we use the Calderón problem as the motivating example. This is considered as a model problem from EIT, in which the voltage is applied on the surface of tissues and the electric intensity is measured on the surface. By changing voltage configurations, many sets of voltage-to-intensity files can be obtained to infer the conductivity of the medium in the tissue [10, 18]. Mathematically, this translates to utilizing the DtN map to reconstruct the diffusion coefficient in the elliptic equation [32, 50],

$$(2.1) \quad \begin{cases} -\nabla \cdot (a(x)\nabla u) = 0, & x \in \mathcal{D} \subset \mathbb{R}^n, \\ u|_{\partial\mathcal{D}} = \phi, \end{cases}$$

where the input ϕ serves as the Dirichlet boundary condition (voltage). The output is also taken on the boundary and is of Neumann type (electric intensity):

$$d = a\partial_\nu u|_{\partial\mathcal{D}},$$

where ν stands for the unit outward normal direction on $\partial\mathcal{D}$. The ItO map from ϕ to d is thus termed the voltage-to-intensity map, or mathematically, the DtN map. This map is parameterized by the medium conductivity $a(x)$:

$$\Lambda_a : \phi \rightarrow d,$$

where the dependence on the conductivity is reflected in the subscript.

Assume $\mathcal{D}$ is a polygonal domain and $\phi \in H^{1/2}(\partial\mathcal{D})$: the weak formulation of (2.1) reads, Find $u \in H^1(\mathcal{D})$ with $u|_{\partial\mathcal{D}} = \phi$ such that

$$(2.2) \quad \int_{\mathcal{D}} a \nabla u \cdot \nabla v \, dx = 0 \quad \forall v \in H_0^1(\mathcal{D}).$$

The DtN map $\Lambda_a\phi$ is defined as the following bilinear form:

$$(2.3) \quad \langle \Lambda_a\phi, \psi \rangle := \int_{\partial\mathcal{D}} a\partial_\nu u \psi \, dx,$$

where u solves (2.2). Using Green's identity and (2.2),

$$\langle \Lambda_a\phi, \psi \rangle := \int_{\mathcal{D}} a \nabla u \cdot \nabla \Psi \, dx.$$

Here $\Psi \in H^1(\mathcal{D})$ can be any extension of ψ such that $\Psi|_{\partial\mathcal{D}} = \psi$. For the rest of the paper, $a(x)$ is assumed to be piecewise constant and is represented uniquely by the vector $\mathbf{a}$ containing its values, and we thus use a and $\mathbf{a}$ interchangeably.

2.1. Sketch of the DtN map discretization hierarchy. In the numerical setup, the solution and the measurements are all discretized and represented by finite-dimensional vectors. The discrete DtN map $\Lambda_{\mathbf{a}}^h$ is therefore a matrix. In this context, the measurements can be viewed as entries in this matrix. Only a small number of measurements are taken in experiments, meaning a small number of the entries in the DtN matrix are available. In other words, a subset of entries Ω of $\Lambda_{\mathbf{a}}^h$ is observed, and

the rest are unavailable. This is translated in the following hierarchy of increasingly reduced objects:

$$(2.4) \quad \Lambda_a \rightarrow \Lambda_a^h \rightarrow \Lambda_a^h|_\Omega.$$

Here, again, Λ_a is the DtN map, Λ_a^h the DtN matrix (discretization of Λ_a , whose size depends on h , the mesh size, and where $\mathbf{a}$ is the discrete version of a in the finite element presentation), and Ω is a subset of matrix indices, indicating where measurements are taken. This reduction process is described in detail in section 2.3.

2.2. Sketch of reversing the DtN map discretization hierarchy. For the reconstruction, we aim to reverse the hierarchy in (2.4). In particular, we start from $\Lambda_a^h|_\Omega$, and by choosing proper data and proper completion algorithms we obtain the full DtN matrix Λ_a^h . This then gets lifted up to represent the DtN map $\Lambda_a = \Lambda_a$ up to a discretization error that depends on h . Due to the involvement of discretization and reconstruction error, the exact recovery of Λ_a^h (and hence Λ_a) is not available. We denote $\tilde{\Lambda}_a^h$, $\tilde{\Lambda}_a$, and $\tilde{a}$ the reconstructed approximations to Λ_a^h , Λ_a , and $\mathbf{a}$, respectively. Ω needs to be judiciously selected such that

$$(2.5) \quad \tilde{\Lambda}_a^h \sim \Lambda_a^h,$$

where $\sim$ means close in some sense (to be defined later). For small h , we lift the matrix back to the map and need to justify

$$(2.6) \quad \tilde{\Lambda}_a \sim \Lambda_a.$$

Finally we seek to establish the closeness of the reconstruction of the medium:

$$(2.7) \quad \tilde{\Lambda}_a^h \sim \Lambda_a^h \quad \Rightarrow \quad \tilde{\Lambda}_a \sim \Lambda_a \quad \Rightarrow \quad \tilde{a} \sim a.$$

In section 2.3 we lay out the numerical setup. We recover $\tilde{\Lambda}_a^h$ from the subsampled $\Lambda_a^h|_\Omega$ and provide intuition to (2.5) in section 2.4. Recall from section 1 that the construction of $\tilde{\Lambda}_a^h$ is twofold: (I) *bridging the gap* and (II) *improving the quality of computational inverse solution*. The proofs for (2.5), (2.6), and (2.7) are given in section 3, which accomplish task (I) of *bridging the gap between theoretical and computational EIT*. Section 2.5 discusses a practical computational algorithm for task (II) which aims to approximately reconstruct $\mathbf{a}$ from the DtN matrix $\tilde{\Lambda}_a^h$. It is important to point out that—unlike traditional algorithms for computational inverse problems that directly use data from $\Lambda_a^h|_\Omega$, the incomplete DtN matrix, to reconstruct $\mathbf{a}$ —we deploy $\tilde{\Lambda}_a^h$, the completed DtN matrix, to reconstruct $\mathbf{a}$. As shall be shown in section 4, our approach *improves the parameter reconstruction substantially*. Indeed, the reconstructions using $\tilde{\Lambda}_a^h$ and the exact DtN matrix Λ_a^h are visibly identical, while the reconstruction directly from $\Lambda_a^h|_\Omega$ is completely off.

2.3. DtN map discretization hierarchy. In what follows we provide details of the DtN map discretization hierarchy (2.4).

From Λ_a to Λ_a^h . Numerically, we first partition the domain $\mathcal{D}$ into N_{el} non-overlapping shape-regular affine elements $K_j, j = 1, \dots, N_{\text{el}}$, with Lipschitz boundaries. Denote $\mathcal{D}^h := \cup_{j=1}^{N_{\text{el}}} K_j$, $\mathcal{D} = \overline{\mathcal{D}^h}$ the discrete space and $h = \max_j \text{diam}(K_j)$ the mesh size. We construct the standard linear Lagrange finite element (FE) space

$$V^h := \left\{ v \in C^0(\mathcal{D}) : v|_{K_j} \in \mathcal{P}^1(K_j) \quad \forall j \right\} \subset C^0(\mathcal{D}) \subset H^1(\mathcal{D})$$

as the discrete solution space and

$$V^h(\partial\mathcal{D}) := \{v \in C^0(\partial\mathcal{D}) : v|_{\partial K \cap \partial\mathcal{D}} \in \mathcal{P}^1(\partial K \cap \partial\mathcal{D})\} = \text{span}\{\phi_i\}$$

as the discrete boundary condition space, where ϕ_i are the linear nodal (Lagrange) basis functions on $\partial\mathcal{D}$. Here $\mathcal{P}^1(K)$ is the space of polynomials of degree at most 1 on K . To project the boundary condition from the continuous level to the discrete one, we define the projection operator

$$(2.8) \quad \Pi^h : H^{1/2}(\partial\mathcal{D}) \ni \phi \mapsto \phi^h := \Pi^h \phi \in V^h(\partial\mathcal{D})$$

such that

$$\|\phi - \phi^h\|_{H^{1/2}(\partial\mathcal{D})} = \|\phi - \Pi^h \phi\|_{H^{1/2}(\partial\mathcal{D})} = \inf_{w^h \in V^h(\partial\mathcal{D})} \|\phi - w^h\|_{H^{1/2}(\partial\mathcal{D})}.$$

The discretization of the weak formulation (2.2) reads, Find $u_h \in V^h$ such that $u^h|_{\partial\mathcal{D}} = \phi^h = \Pi^h \phi$ and

$$(2.9) \quad \int_{\mathcal{D}} \mathbf{a} \nabla u^h \cdot \nabla v^h \, dx = 0 \quad \forall v^h \in V_0^h,$$

where $V_0^h := \{v \in V^h(\mathcal{D}) : v|_{\partial\mathcal{D}} = 0\}$.

Following the standard FE discretization, we define the stiffness matrix $\mathbf{S}$. Denoting $l_k(x)$ the hat functions supported in the domain $\mathcal{D}$ that peaks at the k th nodal point, the stiffness matrix has the form of

$$(2.10) \quad \mathbf{S} = \begin{bmatrix} \mathbf{S}^{ii} & \mathbf{S}^{ib} \\ \mathbf{S}^{bi} & \mathbf{S}^{bb} \end{bmatrix},$$

where i stands for the collection of the indices of the interior degrees of freedom and b is for the degrees of freedom at the boundary. For example, $\mathbf{S}^{ib} = \int_{\mathcal{D}} \mathbf{a} \nabla l_{k_1} \cdot \nabla l_{k_2} \, dx$ for k_1 and k_2 being interior and boundary indices, respectively. As a result, the numerical solution is $u^h = \sum_k \mathbf{u}_k^h l_k(x)$ with

$$\mathbf{u}^h = \begin{bmatrix} -(\mathbf{S}^{ii})^{-1} \cdot \mathbf{S}^{ib} \cdot \phi^h \\ \phi^h \end{bmatrix}.$$

The discretized DtN map Λ_a^h is bilinear on $V^h(\partial\Omega)$: for $\phi^h \in V^h(\partial\Omega)$ and $w^h \in V^h(\partial\Omega)$,

$$(2.11) \quad \langle \Lambda_a^h \phi^h, w^h \rangle := \int_{\Omega} \mathbf{a} \nabla \Phi^h \cdot \nabla W^h \, d\Omega,$$

where W^h is any extension of w^h from $V^h(\partial\Omega)$ to $V^h(\Omega)$ such that $W^h|_{\partial\Omega} = w^h$ and Φ^h is the FE solution obtained from (2.9). The DtN matrix is the matrix representation of Λ_a^h , and we abuse the notation and still call it Λ_a^h .

Using the FE notation as above, denote $\mathbf{M}$ the map from the discrete solution $\mathbf{u}^h$ to the discrete Neumann data on the boundary $\partial\mathcal{D}^h$. The DtN matrix Λ_a^h can be formed as

$$(2.12) \quad \mathbf{d} = \Lambda_a^h \cdot \phi = \mathbf{M} \cdot \begin{bmatrix} -(\mathbf{S}^{ii})^{-1} \mathbf{S}^{ib} \\ \mathbb{I} \end{bmatrix} \cdot \phi \quad \text{with} \quad \Lambda_a^h = \mathbf{M} \cdot \begin{bmatrix} -(\mathbf{S}^{ii})^{-1} \mathbf{S}^{ib} \\ \mathbb{I} \end{bmatrix}.$$

Note that Λ_a^h is a square matrix of size $|\partial\mathcal{D}^h| \times |\partial\mathcal{D}^h|$, where $|\partial\mathcal{D}^h|$ denotes the number of grid points on $\partial\mathcal{D}^h$ and $\mathbb{I}$ is the identity associated to the degrees of freedom at the boundary.

From Λ_a^h to $\Lambda_a^h|_\Omega$. In practice, only a small number of experiments can be conducted, and in each experiment, only a small number of measurements can be taken. For notational convenience, we assume the input ϕ is chosen from the set of basis functions ϕ_i . In this case the ij th component of the data matrix $\mathbf{d}$ is exactly the ij th entry of the DtN matrix, i.e.,

$$d_{ij} \sim \Lambda_{a,ij}^h, \quad (i, j) \in \Omega \subset [1 : |\partial\mathcal{D}^h|]^2.$$

Here we use $\sim$ instead of $=$ to account for potential measuring errors, and Ω , referred to as a mask, is a subset of all indices of Λ_a^h .

2.4. Reverse DtN map discretization hierarchy. As it was argued in section 1, some data pairs are more informative than the others. Choosing the most informative data, or equivalently, selecting the right mask Ω , is of paramount importance in recovering the missing entries in the DtN matrix Λ_a^h . Recall from section 2.2 that, due to errors in the discretization and reconstruction process, we can only obtain an approximation $\tilde{\Lambda}_a^h$ of Λ_a^h . In the following we exploit the structure of the DtN matrix Λ_a^h to determine Ω and employ a matrix completion technique such that $\tilde{\Lambda}_a^h$ is close to Λ_a^h .

To describe the structure of the DtN matrix, we exploit the concept of $\mathcal{H}$ -matrices. With a proper decomposition, the DtN matrix can be partitioned into several blocks that are approximately low rank (ALR). This means many blocks from the big DtN matrix have their singular values decay fast regardless of the resolution and thus can be approximated by a low-rank block. This partitioning allows us to utilize the matrix completion-type methods for the ALR blocks that are not applicable to the full DtN matrix, as it is often of full rank. In the following we briefly review the matrix completion method in section 2.4.1 and evaluate the matrix structure of Λ_a^h in section 2.4.2. The full completion algorithm is presented in Algorithm 2.1.

We stress that in the process of reconstructing the full map, we assume the given data $\Lambda_a^h|_\Omega$ is noiseless, so the error purely comes from the reconstruction process itself. If the collected data in $\Lambda_a^h|_\Omega$ already presents measuring error, stability in the reconstruction process should also be discussed. We leave that to future work.

2.4.1. Matrix completion. Matrix completion has been a popular topic for a decade due to its applications in recommendation systems, including the famous Netflix problem [8]. The goal is to complete the entries in a matrix from a partial knowledge of its entries. To be more specific, let a generic $\mathbf{A} \in \mathbb{R}^{n \times n}$ be the to-be-completed matrix, of which only some of its entries are known. In this setting Ω , with $|\Omega| = m$, is the index set where the entries are known, and a_{ij} the given values with $(i, j) \in \Omega$.

There exists a number of algorithms that aim to reconstruct the entries [45, 28, 22, 11]. We adopt the approach proposed in [14]. Under the assumption that the matrix $\mathbf{A}$ is of low rank ($r \ll n$), we seek to minimize the nuclear norm $\|\mathbf{A}\|_*$ such that the matrix's evaluations at certain locations are fixed by the given data, namely, $\mathbf{A}_{ij} = a_{ij}$ for $(i, j) \in \Omega$. The minimization problem now becomes

$$(2.13) \quad \min_{\mathbf{A}} \|\mathbf{A}\|_* \quad \text{s.t.} \quad \mathbf{A}_{ij} = a_{ij}, \quad (i, j) \in \Omega, \quad |\Omega| = m.$$

The objective function $\|\mathbf{A}\|_*$ is the sum of all singular values of $\mathbf{A}$. It can be viewed as the relaxation from the ℓ_0 -norm of singular values ($\text{rank}\{\mathbf{A}\}$) to its ℓ_1 -norm [47].

One important advantage of working with (2.13) is that it is a convex optimization problem that can be solved efficiently with interior point methods. It is important to point out that this convex relaxation does not come with an accuracy sacrifice: It is found that under very mild conditions—*decoherent* and *delocalization* conditions—the solution of minimizing $\|A\|_*$ coincides with the solution of minimizing $\text{rank}\{A\}$.

We next recall the *decoherent* and *delocalization* conditions [14]. Let

$$A = U\Sigma V^T$$

be the singular value decomposition of A .

DEFINITION 2.1. Let W be a subspace of $\mathbb{R}^n$ of dimension r and $\mathcal{P}_W$ be the orthogonal projection onto W . Then the coherence index of W is defined as

$$(2.14) \quad \mu(W) = n \max_{1 \leq i \leq n} \|\mathcal{P}_W e_i\|_2,$$

where e_i is the i th unit vector of $\mathbb{R}^n$.

A1 *Decoherent condition:* $\max(\mu(U), \mu(V)) \leq \mu_0$ for some positive μ_0 .

A2 *Delocalization condition:* The maximum entry of $\sum_{1 \leq k \leq r} u_k v_k^T$ is bounded from above by $\mu_1 \sqrt{\frac{r}{n^2}}$ for some positive μ_1 .

Let us now state a probabilistic result on the success of (2.13).

THEOREM 2.2 ([14, 46]). Let A be an $n \times n$ matrix of rank r obeying the decoherent and delocalization conditions **A1** and **A2**. Suppose we observe m entries of A with locations sampled uniformly at random. Then there exist constants C, c such that if

$$m \geq C \max \left\{ \mu_1^2, \mu_0^{1/2} \mu_1, \mu_0 n^{1/4} \right\} \beta r (n \log n)$$

for some $\beta > 2$, then the minimizer to the problem (2.13) is unique and equal to A with probability at least $1 - cn^{-\beta}$. For $r \leq \mu_0^{-1} n^{1/5}$ this estimate can be improved to

$$m \geq C \mu_0 \beta r \left(n^{6/5} \log n \right)$$

with the same probability of success.

This theorem suggests that if the to-be-completed matrix A is of low rank satisfying the decoherent and delocalization conditions, then when the number of provided entries m linearly depend on r , the rank, and the entries are sampled uniformly randomly, the matrix can be precisely reconstructed with a high probability.

2.4.2. Structure of Λ_a^h and $\mathcal{H}$ -matrix. We now discuss how we use the matrix completion method discussed in subsection 2.4.1 to recover the DtN matrix Λ_a^h from its data $\Lambda_a^h|_\Omega$. Theorem 2.2, despite providing a general recipe for reconstructing a matrix from its incomplete data, requires the rank r to be significantly smaller than n , the size of the matrix, for the algorithm to be meaningful. However, Λ_a^h is a full-rank matrix, preventing the direct application of the matrix completion algorithm.

It turns out that we can still take advantage of the matrix completion algorithm by exploiting the $\mathcal{H}$ -matrix structure embedded in Λ_a^h . This allows us, through the “peeling” process [35], to divide the matrix into sub-blocks, most of which are ALR. The application of the matrix completion algorithm to these ALR blocks is then expected to be efficient.

A hierarchical matrix, commonly referred as an $\mathcal{H}$ -matrix, is a class of matrices that, upon proper partitioning, have fast decays in singular values in the smaller

blocks, leading to the ALR property within these blocks. The concept [24, 25] was invented initially to divide a given $\mathcal{H}$ -matrix into smaller blocks so that some matrix operations, including matrix-vector multiplication, addition, inverse, Schur complement, and many others, could be significantly sped up (see, e.g., [6, 40, 56, 21] and references therein). In its original form, one also requires linear computational complexity in finding the partitioning. In our setting, such complexity is irrelevant.

Remark 2.3. We point out that there are multiple ways to reconstruct low-rank matrices/blocks of matrices besides the matrix completion algorithm that was described above. Other possible choices are randomized SVD [26], CUR factorizations [39, 23], adaptive cross approximation [5], and some other algorithms that are adaptive in nature [57, 7]. Even though most of these algorithms would be able to reconstruct the DtN map with a lower sample complexity of $\mathcal{O}(n \log n)$ they present two main drawbacks for the applications to EIT: (i) they require the full knowledge of full rows or columns, or matrix-vector multiplication, and (ii) they require an sequential and online sampling for the reconstruction, i.e., the reconstruction algorithm will sequentially sample the columns and rows of the block. Sampling a full row or column is usually infeasible in practice. Indeed, considering the EIT problem where the matrix presents the map from Dirichlet (potential) data to Neumann (flux) data, accessing the full row/columns amounts to taking measurements continuously along the boundary. This is usually considered too expensive and unnecessary in experiments. Experimentally, a much more feasible approach is to take measurement discretely, which amounts to accessing a subset of entries in the matrix. Sampling adaptively, though a useful sampling process, also poses some experimental hassle. In our context, this amounts to running the experiment several times, each time with a slightly different sensor configuration. It is very costly and labor intensive, especially considering that the procedure is applied to a large number of blocks for the DtN matrix. In this case, there is a practical gain of acquiring the data first and then postprocessing it in an offline fashion, which can be greatly parallelized.

We also stress that under the framework of $\mathcal{H}$ matrix, the reconstruction using entries is not completely new. In [36, 16, 33] various ways of data completion strategies have been adopted. The context is different, while the concepts are rather similar.

It was shown in [6] that the collection of Green's functions for elliptic equations produces an $\mathcal{H}$ -matrix. Recall from the expression for Λ_a^h in (2.12) that both M and S^{ib} are sparse matrices. Thus if $(S^{ii})^{-1}$ is an $\mathcal{H}$ -matrix, so is Λ_a^h .

THEOREM 2.4 (Theorem 4.28 of [6]). *Let $\epsilon > 0$ small; then there is an $\mathcal{H}$ -matrix C with local block rank being $k \lesssim (\log n)^2 |\log \epsilon|^{d+1}$ such that $\|(S^{ii})^{-1} - C\|_2 \leq \epsilon$. For the accuracy compatible with the FE method, we take $\epsilon = h^\beta$ and $k \lesssim \log^{d+3} n$. Here n is the size of the matrix, d is the dimension of the problem, h is the mesh size, and β is the accuracy order of the FE method.*

The author of [6] furthermore suggests a way to choose the partition. Indeed $(S^{ii})^{-1}$ is essentially the discrete version of the Green's function $G(x, y)$, and as a function of x parameterized by y it can be approximately written as a summation of separable functions in x and y if the two coordinates are well separated. This reveals that the rank of the approximation only logarithmically depends on the size of the matrix, and since the Green's function can be approximated by a few separable functions only if x and y are well separated, only the blocks of $(S^{ii})^{-1}$ that are not along the diagonal are ALR. The same observation was made in [35] which shows that the off-diagonal blocks are ALR.

Built upon these observations, noting that Λ_a^h is an $\mathcal{H}$ -matrix with diagonal blocks, $\Lambda_{a,D}^h$, having full rank and the off-diagonal blocks, $\Lambda_{a,O}^h$, being ALR, we propose to obtain the full data in $\Lambda_{a,D}^h$ but a small portion of entries in $\Lambda_{a,O}^h$ and then recover the missing entries with the matrix completion method in section 2.4.1. Experimentally, this means for every injected voltage concentrated on one nodal point on $\partial\mathcal{D}$, one measures the intensity on that particular location so to fill the diagonal blocks $\Lambda_{a,D}^h$. One then decreases the density of the detectors as one moves further away along $\partial\mathcal{D}$ and employs the matrix completion algorithm to recover $\Lambda_{a,O}^h$. In Algorithm 2.1, we summarize the whole process of completing Λ_a^h .

Algorithm 2.1 Completing Λ_a^h .

Preparation:

0. Determine the partition, and identify $\Lambda_{a,D}^h$ and $\Lambda_{a,O}^h$.
1. Sample each entry in $\Lambda_{a,D}^h$.
2. For each $\Lambda_{a,O}^h$, do the following:
 - 2.1: Randomly collect $rn^{6/5} \log n$ data points in the block (n : the size of the block).
 - 2.2: Solve the matrix completion problem (2.13) to reconstruct $\Lambda_{a,O}^h$.

end**Output:** Assemble the recovered submatrices, denoted by $\tilde{\Lambda}_a^h$.

Once $\tilde{\Lambda}_a^h$ is formed, it can be lifted to a corresponding DtN map $\tilde{\Lambda}_a$ which in turn corresponds to a unique conductivity $\tilde{\mathbf{a}}$. The analysis on the difference between Λ_a and $\tilde{\Lambda}_a$ (and between $\mathbf{a}$ and $\tilde{\mathbf{a}}$) is presented in section 3.

2.5. Improving inverse solution with matrix completion. With the full map $\tilde{\Lambda}_a^h$ in hand, the reconstruction of the medium $\mathbf{a}$ is now straightforward using classical optimization-based methods. Since this component of the algorithm is rather classical, we briefly review it here.

We consider the reconstructed $\tilde{\Lambda}_a^h$ as the groundtruth data, and we search for the medium impedance $\mathbf{a}$ such that the misfit—with Frobenius norm—between the DtN matrix generated by that $\mathbf{a}$ and the groundtruth data is minimized, i.e.,

$$(2.15) \quad \min_{\mathbf{a}} \|\Lambda_a^h - \tilde{\Lambda}_a^h\|_F^2 + \alpha \|\mathbf{a} - \mathbf{a}_0\|_q^q + \beta R(\mathbf{a}),$$

where the second term is a regularization term taking into account some prior knowledge and the third term is an additional regularization term to enforce desirable properties in the reconstruction. A typical example is a total-variation norm of $\mathbf{a}$ to eliminate its oscillations [15]. Note that even though the first term may seem benign at first glance, the DtN map, Λ_a^h , is highly nonlinear in $\mathbf{a}$. This may imply the existence of many local minima in the objective function landscape, which can greatly tax the capability of standard gradient-based optimization techniques. In this context, both regularization terms can be tuned to attenuate this issue; however, how to tune these methods is outside the scope of this paper. For our numerical results in section 4, we set both regularization parameters α, β to zero, and we will use an off-the-shelf optimization quasi-Newton method [12], with a simple box constraint, to reconstruct the impedance $\mathbf{a}$.

3. Bridging the gap with matrix completion. To show (2.5) amounts to showing the reconstructed DtN matrix $\tilde{\Lambda}_a^h$ is close to the true DtN matrix Λ_a^h . For that we combine the $\mathcal{H}$ -matrix argument and the matrix completion result.

THEOREM 3.1. *Divide the $n \times n$ matrix Λ_a^h into N blocks with each of size $n^i \times n^i$ according to the $\mathcal{H}$ -matrix decomposition with n^i denoting the size of the i th block. Suppose the i th block has rank r^i and obeys the decoherent and delocalization conditions with constants μ_0^i and μ_1^i . Denote m^i the number of observed entries in the i th block with samples chosen uniformly at random. Then to reconstruct Λ_a^h using (2.13), there exist constants C, c such that if*

$$m^i \geq C \max \left\{ (\mu_1^i)^2, \sqrt{\mu_0^i \mu_1^i}, \mu_0^i (n^i)^{1/4} \right\} \beta^i r^i \sum_i (n^i \log n^i), \quad i = 1, \dots, N,$$

for some $\beta^i > 2$, and hence the total sampling number is $m = \sum m^i \sim n^{5/4} N^{3/8}$, then

$$\mathbb{P}(\Lambda_a^h = \tilde{\Lambda}_a^h) \geq 1 - c \sum_{i=1}^N n_i^{-\beta^i}.$$

The sampling sizes can be improved to $m^i \geq C \mu_0^i \beta^i r^i ((n^i)^{6/5} \log n^i)$ if $r^i \leq (n^i)^{1/5} / \mu_0^i$.

Proof. The proof is a straightforward application of Theorem 2.2. To show that the total sampling $m \sim n^{5/4} N^{3/8}$, we note that $\sum_i n_i^2 = n^2$, so the lower bound of $m = \sum m^i \geq C(n^i)^{1/4} \sum_i n^i \log n^i$ becomes, according to Hölder's:

$$C \sum_i (n^i)^{5/4} \log n^i \leq C N^{3/8} \left(\sum_i ((n^i)^{5/4})^{8/5} \right)^{5/8} \sim n^{5/4} N^{3/8}.$$

□

Remark 3.2. Note that the size of N is typically $\log n$, so $N^{3/8}$ barely contributes anything to the sampling complexity.

Remark 3.3. We assume that the rank of the i th block is r^i . One should note that this is only an approximate rank. According to [4], elliptic boundary-to-boundary operators have exponentially decaying singular values, and thus r^i depend on the error tolerance. For a more precise reconstruction, larger r^i may be needed, and it amounts to a higher value of m^i , meaning more data points are needed.

Remark 3.4. The theorem states that the two matrices, the reconstructed and the groundtruth, are exactly the same with high probability. In practice, the data obtained in $\Lambda_a^h|_\Omega$ is often polluted with measurement errors. In [13] the authors discuss the effect of such pollution in the reconstruction.

To quantify (2.6) and (2.7), we will rely on some delicate FE analysis and Theorem 3.1. To begin, we lift both Λ_a^h and $\tilde{\Lambda}_a^h$ matrices to their “corresponding” (or reconstructed) DtN maps $\hat{\Lambda}_a$ and $\tilde{\Lambda}_a$ as follows:

$$(3.1) \quad \tilde{\Lambda}_a := (\Pi^h)^* \circ \tilde{\Lambda}_a^h \circ \Pi^h, \quad \hat{\Lambda}_a := (\Pi^h)^* \circ \Lambda_a^h \circ \Pi^h,$$

where $(\Pi^h)^*$ is the adjoint of Π^h . It is easy to see that both $\tilde{\Lambda}_a$ and $\hat{\Lambda}_a$ are well-defined linear continuous operators from $H^{1/2}(\partial\Omega)$ to $H^{-1/2}(\partial\Omega)$. Theorem 3.1 implies that $\tilde{\Lambda}_a = \hat{\Lambda}_a$ with high probability, i.e.,

$$\mathbb{P}(\tilde{\Lambda}_a = \hat{\Lambda}_a) \geq 1 - c \sum_{i=1}^N n^{-\beta^i}.$$

We now recall the following well-known result.

THEOREM 3.5 ([2, 48]). *Let a_1 and a_2 be two piecewise constant functions on $\mathcal{D}$, and let $\mathbf{a}_i$ be their representation vectors. Denote $\Lambda_{\mathbf{a}_i}$ the corresponding DtN map defined in (2.3). Then $\mathbf{a}$ as a function of $\Lambda_{\mathbf{a}}$ is Lipschitz:*

$$(3.2) \quad \|\mathbf{a}_1 - \mathbf{a}_2\|_{\infty} \leq C \|\Lambda_{\mathbf{a}_1} - \Lambda_{\mathbf{a}_2}\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)};$$

that is, there is a unique conductivity $\mathbf{a}$ for every DtN map $\Lambda_{\mathbf{a}}$.

Since we assume that $\mathbf{a}$ is piecewise constant, Theorem 3.5 implies that there exists a unique conductivity $\tilde{\mathbf{a}}$ corresponding to $\hat{\Lambda}_{\mathbf{a}}$ such that

$$\|\mathbf{a} - \tilde{\mathbf{a}}\|_{\infty} \leq C \left\| \Lambda_{\mathbf{a}} - \hat{\Lambda}_{\mathbf{a}} \right\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)}.$$

Remark 3.6. It is worth noting that the result hides the dependence on the degrees of freedom in $\mathbf{a}$. In particular, in [48], the author showed that the Lipschitz constant C can grow exponentially fast with respect to n , the number of “pieces” used in $\mathbf{a}$, if the medium is accordingly designed. Such study is out of scope of this paper; thus we omit the discussion.

The following (whose technical proof is given in the Appendix A) is the justification for (2.6) and (2.7), which, similar to Theorem 3.5, shows the uniqueness of the reverse DtN hierarchy in section 2.4.

THEOREM 3.7 (asymptotic uniqueness). *There holds*

$$\lim_{h \rightarrow 0} \left\| \Lambda_{\mathbf{a}} - \hat{\Lambda}_{\mathbf{a}} \right\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} = 0,$$

and thus

$$\lim_{h \rightarrow 0} \|\mathbf{a} - \tilde{\mathbf{a}}\|_{\infty} = 0.$$

Let $\tilde{\mathbf{a}}_1$ and $\tilde{\mathbf{a}}_2$ be the conductivities associated with two reconstructed DtN maps $\hat{\Lambda}_{\mathbf{a}_1}$ and $\hat{\Lambda}_{\mathbf{a}_2}$ corresponding to $\mathbf{a}_1$ and $\mathbf{a}_2$, respectively. Then

$$\begin{aligned} \|\tilde{\mathbf{a}}_1 - \tilde{\mathbf{a}}_2\|_{\infty} &\leq C \left(\left\| \Lambda_{\mathbf{a}_1} - \hat{\Lambda}_{\mathbf{a}_1} \right\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} \right. \\ &\quad \left. + \left\| \Lambda_{\mathbf{a}_1} - \Lambda_{\mathbf{a}_2} \right\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} + \left\| \Lambda_{\mathbf{a}_2} - \hat{\Lambda}_{\mathbf{a}_2} \right\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} \right). \end{aligned}$$

That is, if $\mathbf{a}_1 \rightarrow \mathbf{a}_2$, then $\tilde{\mathbf{a}}_1 \rightarrow \tilde{\mathbf{a}}_2$ as $h \rightarrow 0$.

We also briefly discuss the complexity of the matrix completion.

THEOREM 3.8. *Denote $\Lambda_{\mathbf{a}}$ a matrix of size $n \times n$ with $n \sim 1/h^d$, and suppose that the matrix can be decomposed in an $\mathcal{H}$ -matrix with the weak admissibility condition (see Figure 1). Suppose that the decomposition has $L \sim \log n$ levels and that each block has a bounded rank r and satisfies the conditions of Theorem 2.2. Then with $|\Omega| = \mathcal{O}(rn^{6/5} \log n)$ known entries sampled properly, we can reconstruct $\Lambda_{\mathbf{a}}$ with high probability.*

Proof. We consider diagonal and off-diagonal blocks separately. For the partition considered in this theorem, there are n/r diagonal blocks. They are full rank, which requires them to be fully sampled, and thus $\mathcal{O}(rn)$ entries are needed.

For the off-diagonal blocks we need to use randomized sampling. Given that the matrix is partitioned in L levels, we have that at the ℓ th level in the decomposition,

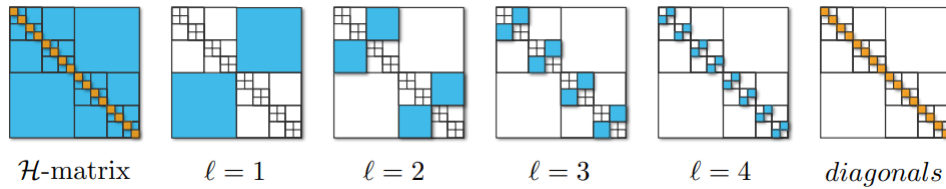


FIG. 1. Sketch of a partition of an $\mathcal{H}$ -matrix in four levels. Blocks in light blue are randomly sampled, and those in orange are fully samples. We show the blocks considered at each level of the partition. As can be seen, we have 2^ℓ number of blocks of size $n^\ell = n/2^\ell$.

each block will have size $n_\ell = n/2^\ell$, and we will have $n/n_\ell = 2^\ell$ of them. Thus following Theorem 2.2, we require $\mathcal{O}(r(n^\ell)^{6/5} \log((n^\ell)^{6/5}))$, or $\mathcal{O}(r(n/2^\ell)^{6/5} \log(n^\ell)^{6/5})$ samples to reconstruct each of the blocks at the ℓ th level in the partition. In summary, we would require $\mathcal{O}(2^\ell r(n/2^\ell)^{6/5} \log((n/2^\ell)^{6/5}))$ to reconstruct, with high-probability, all the blocks at the ℓ th level of the partition.

After adding the number of samples at each level we have that the total number of samples required scales as

$$(3.3) \quad \mathcal{O}\left(\sum_{\ell=1}^{n_{\text{levels}}} 2^\ell r(n/2^\ell)^{6/5} \log(n/2^\ell)^{6/5}\right) = \mathcal{O}\left(rn^{6/5} \log n \sum_{\ell=1}^{n_{\text{levels}}} 2^{-\ell/6}\right)$$

$$(3.4) \quad = \mathcal{O}(rn^{6/5} \log n),$$

where we used the fact that $\sum_{\ell=1}^{n_{\text{levels}}} 2^{-\ell/6} = \mathcal{O}(1)$. $\square$

Remark 3.9. We have chosen the weak admissibility condition for the sake of simplicity: the theorem follows for other types of partitioning. In particular, as will be shown in section 4.1, using a strong admissibility condition with periodic boundary conditions yields similar results.

4. Numerical experiments. We finally present several numerical experiments showcasing the framework introduced above. All the experiments were coded in MATLAB 2019b using CVX to solve the optimization problems with MOSEK [3] as the back-end. The experiments were run on a single-socket workstation running an AMD 2950X processor with 128 GB of RAM.

As the method suggests, we should first run the hierarchical matrix completion algorithm in Algorithm 2.1 to reconstruct the full DtN matrix and then use the completed DtN matrix to reconstruct the media. For an accurate DtN matrix reconstruction, we need to ensure the matrix gets decomposed according to the $\mathcal{H}$ -matrix admission condition, and the matrix completion algorithm is implemented within each off-diagonal block that is ALR. The matrix completion algorithm requires two conditions to be held: the decoherent and delocalization conditions.

Even though these properties may be already familiar for the reader with a background in compressed linear algebra for elliptic operators, we will, for the sake of completeness, demonstrate that the off-diagonal blocks indeed satisfy the decoherent and delocalization conditions. These conditions will be shown to be satisfied independent of the level of numerical refinement. This ensures that the matrix completion algorithm indeed reconstructs the DtN matrix accurately with limited data. With the demonstration of the accurate reconstruction of the DtN matrix, we will further showcase the reconstruction of the medium/conductivity. To show the outperformance of

the proposed method, we also implement and demonstrate the reconstruction using a few other sampling techniques. This includes sampling entries of the DtN matrix using an uniform distribution, sampling columns or rows of the DtN uniformly, and finally, sampling the rows and columns together uniformly. The comparison suggests that the subsampling performs the best when it honors the $\mathcal{H}$ -matrix format and the off-diagonally low-rank structure of the DtN matrix.

In subsection 4.1, we verify the decoherent and the delocalization conditions and present reconstruction of the DtN matrix. In subsection 4.2 we showcase the reconstruction of the conductivity (medium). In subsection 4.3 we demonstrate an extension of the presented method on optical tomography. In optical tomography, the albedo operator maps the incoming light to the outgoing light intensity and is used to reconstruct the scattering coefficient, an optical property of the material.

4.1. Reconstructing the DtN map. As presented in section 2.4 the reconstruction of DtN matrix relies on two key factors: a proper $\mathcal{H}$ -matrix decomposition and the proper use of the matrix completion algorithm in the ALR submatrices that satisfy both decoherent and delocalization conditions. We demonstrate both the $\mathcal{H}$ -matrix decomposition and the final matrix completion results.

We now detail the numerical setup. In $\mathcal{D} = [0, 1]^2$ domain, we choose the Shepp–Logan phantom as the ground truth medium, as plotted in Figure 2(a). On the domain we use the nested grids, with $n_h = 2^\ell + 1$ discrete points per dimension where ℓ is the refinement level. This leads to $n = 2^{\ell+2}$ grid points along the boundary, making a DtN matrix of size $n \times n$. In this DtN matrix, we separate the diagonal and off-diagonal blocks following the strong admissibility condition. It was shown in section 2.4.2 that these off-diagonal blocks are of ALR, and the matrix completion algorithm could potentially bring benefit if the decoherent and the delocalization conditions are satisfied. We choose two representative square blocks to verify these conditions. They are the blocks a and b demonstrated in Figure 2(b). As ℓ increases, these blocks have larger sizes accordingly: $n_a = n_h - 1$, and $n_b = (n_h - 1)/2$.

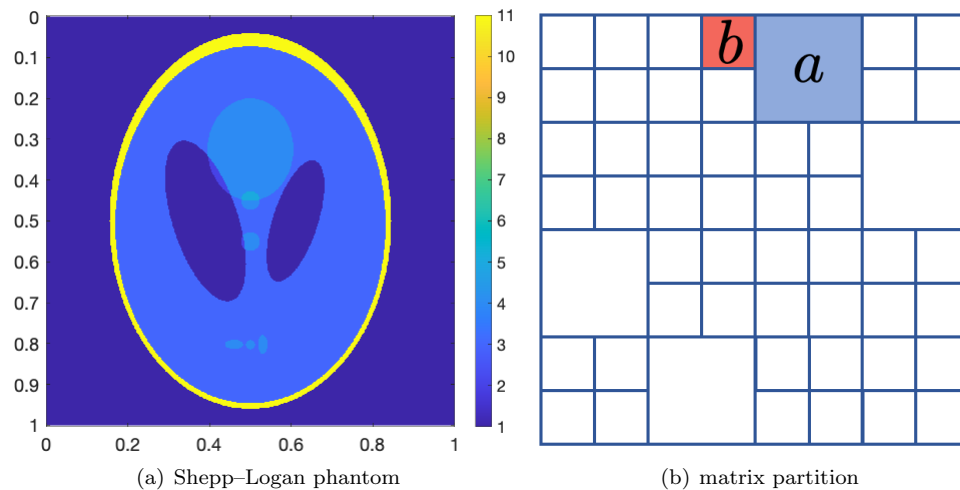


FIG. 2. Figure 2(a) shows the impedance used for the experiments corresponding to the well-known Shepp–Logan phantom, where the color encodes the value of the impedance at each point. Figure 2(b) presents a partition of the matrix together with two blocks used for the numerical experiments.

To show the low-rank structure of the DtN matrix, we plot in Figure 3 different levels of $\mathcal{H}$ -matrix partitioning. At each level of partitioning, we also plot the approximate rank of each block. The rank is evaluated as the number of singular values above $\epsilon = 10^{-6}$. As can be seen uniformly across all refinement levels, the approximate rank of all off-diagonal blocks is smaller than 5. We also plot a typical off-diagonal DtN matrix block and its rank structure, shown in Figure 4. It is clear that these blocks are ALR.

To verify the conditions, including the decoherent and delocalization conditions, we plot the coherence indices and the maximum absolute values defined in (2.14) for blocks a and b . These are shown in Figure 5(a). It is clear that as ℓ increases, the coherence index stays stable for block a and only increases slightly for block b , saturating at a relatively small number quickly. The maximum value in the matrix entry evaluation decreases quickly for both matrix blocks, shown in Figure 5(b). This evidence suggests that employing the matrix completion algorithm on off-diagonal blocks will provide satisfying results.

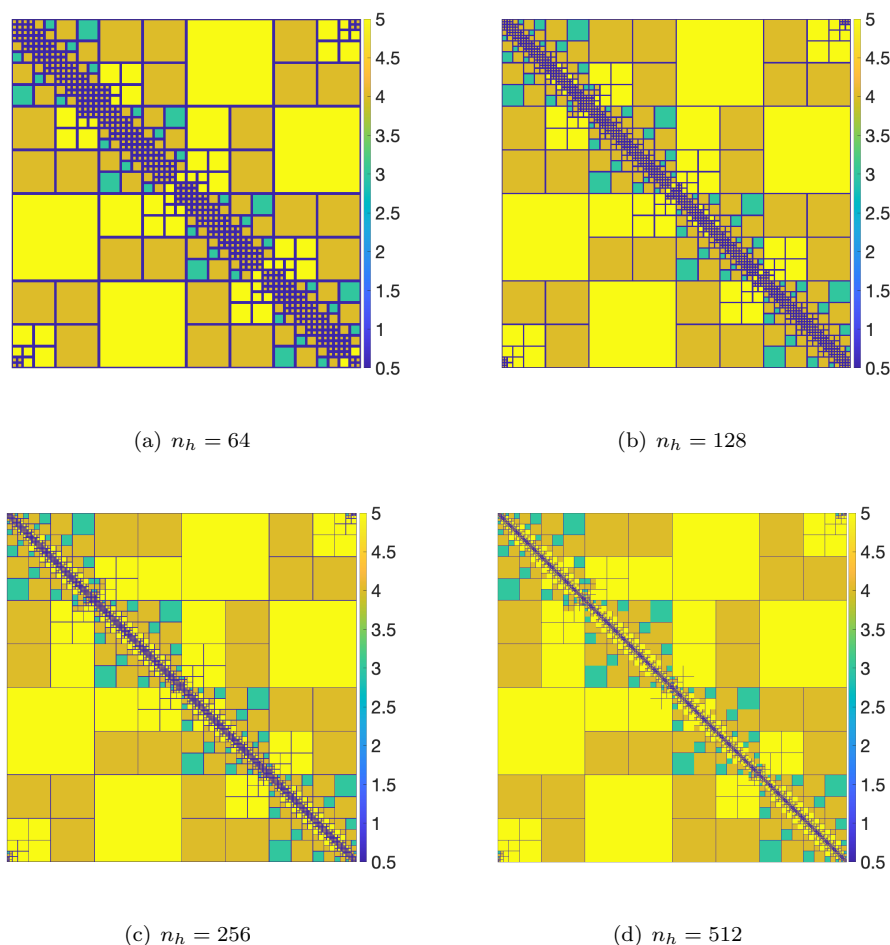


FIG. 3. Partition of the DtN map using different levels of refinements: the blocks are colored with their ϵ -ranks (for $\epsilon = 10^{-6}$).

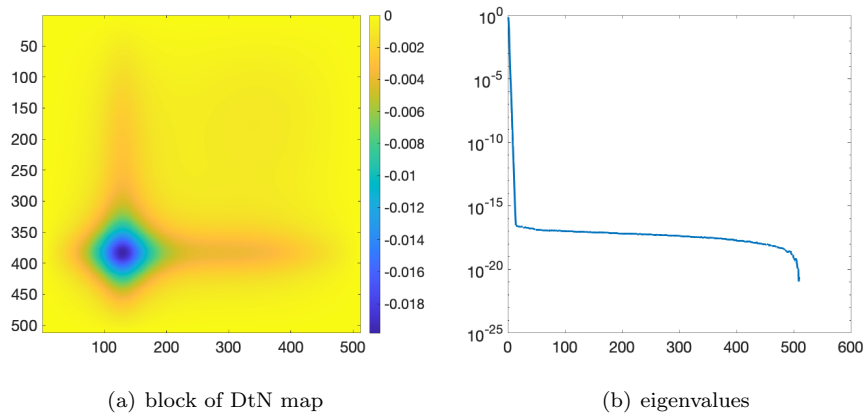


FIG. 4. Plot of a typical off-diagonal block of the DtN map, where the colors correspond to the value of each entry. The singular value of the block decays extremely fast.

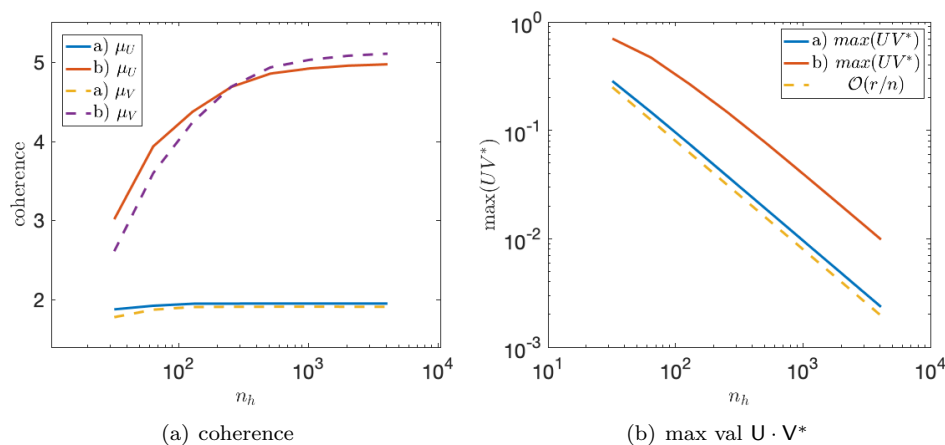


FIG. 5. Coherence index and maximum absolute value of $U \cdot V^*$ for blocks a and b at different levels of refinement in the discretization.

Finally we reconstruct the DtN matrix according to Algorithm 2.1. Since the reconstruction is performed for each block separately, we take the reconstruction of block a as an example. At each level of the refinement, we select entries from block a according to the Bernoulli distribution with parameter p . We then take the values of these entries as the given data to solve the matrix completion optimization problem (2.13). In Figure 6 we plot the original block, the location of the selected entries, and the reconstruction. Clearly, with $p = 0.1$, only ten percent of the data given, we already construct this block with high accuracy.

To quantitatively evaluate the performance of the algorithm, for each predetermined p and refinement level ℓ , we perform the selection and reconstruction process 50 times and document the success ratio as a function of p and ℓ . A successful run is defined as a run where the reconstructed block is within 10^{-4} relative error of the groundtruth in Frobenius norm. In Figure 7 we plot the success ratio of reconstructing block a . For low refinement level with coarse discretization, the DtN blocks have small sizes, and the matrix completion algorithm requires a higher percentage of known data

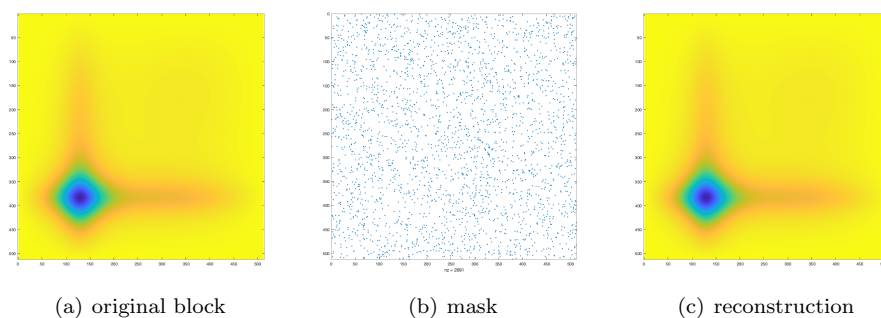


FIG. 6. Reconstruction of block a in the DtN map with $p = 0.1$, meaning the down-sampling rate is 10%.

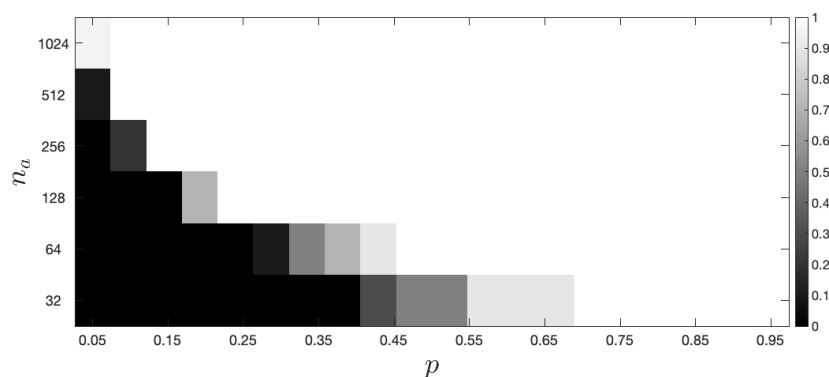


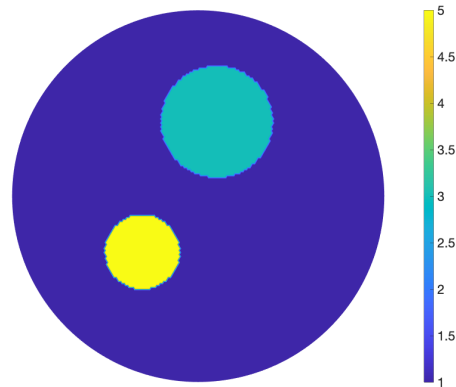
FIG. 7. Success ratio for the reconstruction of block a at different refinements and different density of the sampling mask where $p \sim m/n_a^2$ and m is the number of sampling points. In the plot the color encodes the success ratio, with a lighter color indicating a higher success ratio.

for a high success probability of the reconstruction. On refined meshes with bigger ℓ , small p is sufficient for an accurate reconstruction with high probability. For example, for a matrix of size 512, only up to 5% of the entries are needed to reconstruct the block with a high probability.

4.2. Reconstruction of the media. To solve the minimization problem (2.15), we use the L-BFGS-B method [12] with a constant initial guess.² In order to avoid the inverse crime, we use a different mesh for the forward and adjoint problem in the optimization loop, but we keep that same elements touching the boundary.³ We terminate the optimization process once the gradient norm is less than 10^{-9} or the number of iterations exceeds 10,000. In Figure 8 we plot the groundtruth impedance $\mathbf{a}$, and in Figure 9(a) we plot the reconstructed medium with the exact DtN matrix Λ_a^h . This will be regarded as our reference solution. We note that there is still some error between the reference solution and the exact medium. This is not surprising as the unique reconstruction of the medium is guaranteed only in the noiseless case in the infinite data limit. While we cannot remove the discretization error and only

²In order to regularize the problem, we have a lower bound in the reconstruction to help smooth the oscillations. This seemed to provide better results than using vanilla quasi-Newton methods.

³This simplification allows us to avoid interpolating the data at the boundary.

FIG. 8. *Exact impedance to be reconstructed.*

have the full DtN matrix instead of the DtN map, the reconstruction is bound to show some difference from the groundtruth. Such error can be gradually removed as one refines the mesh to remove the error. Nevertheless, this reconstruction with the exact DtN matrix provides a benchmark, as it is the best possible reconstruction in the current computational framework.

To demonstrate the better performance of the proposed subsampling strategy, we implemented different subsampling strategies and run the same optimization process for reconstructing the medium. In Figure 10 we plot the masks used to subsample the full DtN matrix Λ_a^h with four different strategies, and in Figure 9 we compare their reconstructions of the medium. It is obvious that the reconstructed medium using the completed $\tilde{\Lambda}_a^h$ agrees well with the reference solution, while all other reconstructions fail to capture the medium.

4.3. Optical tomography. We have used the EIT problem to show that data completion can not only bridge the gap between theoretical and computational inverse problems but also help improve computational inverse solutions. While the former depends on the available theories of the inverse problem under consideration, the latter is expected to be valid for all problems. To demonstrate that data completion is also possible for other problems with $\mathcal{H}$ -matrix structure, we now consider an optical tomography problem, where the radiative transfer equation serves as the forward model and its scattering coefficient—the unknown parameter—reflects the optical property of the media. More specifically, let $f(x, v)$ present the density of photon particles at location x moving in direction v ; then the radiative transfer equation characterizes the dynamics of this distribution function, and in steady state it reads

$$v \nabla f = \frac{1}{\text{Kn}} \sigma_s(x) \left[\int_{v'} f dv' - f \right].$$

Here the left-hand side describes the particles moving in direction x with velocity v , and the term on the right suggests the scattering with the intensity characterized by σ_s . Kn is called the Knudsen number. The “inflow” part of the boundary

$$\Gamma_- = \{(x, v) : x \in \partial\mathcal{D}, v \cdot n_x < 0\}$$

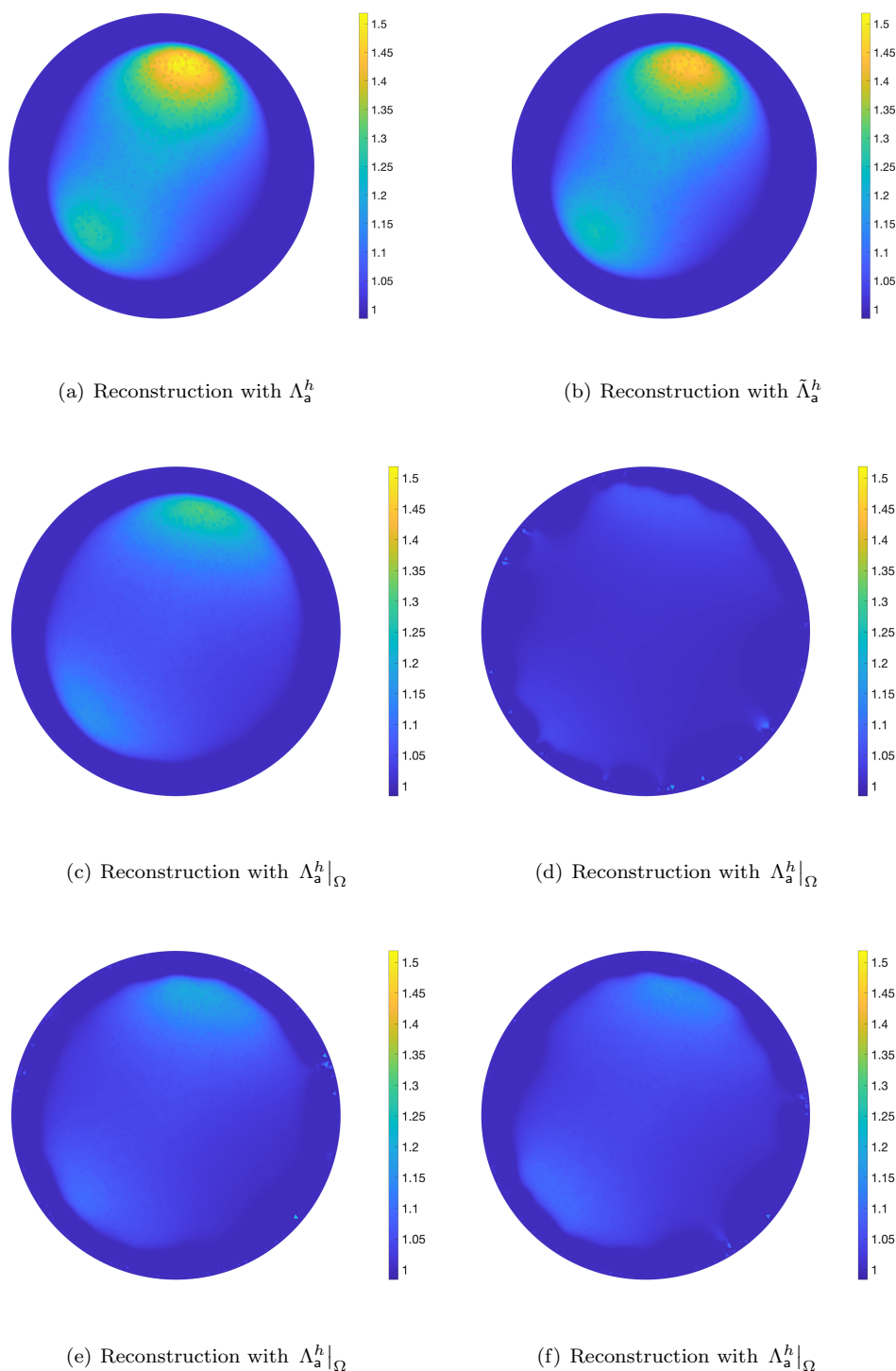
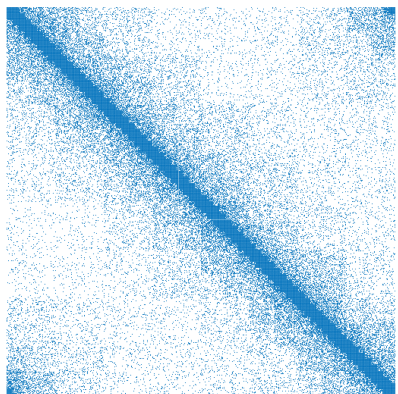
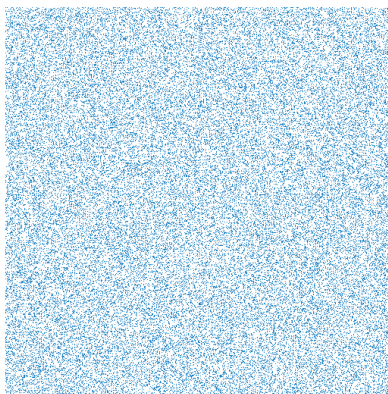


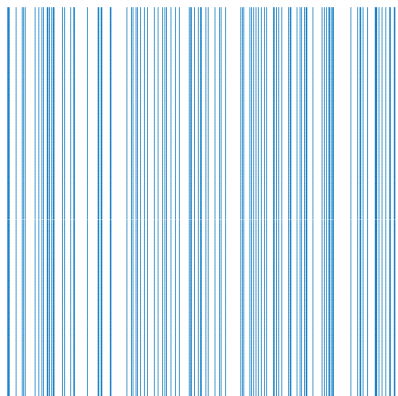
FIG. 9. Reconstructed impedance. In (a) and (b) we show the reference reconstruction and the reconstruction using the DtN matrix $\tilde{\Lambda}_a^h$ completed from a subsampled DtN matrix $\Lambda_a^h|_\Omega$, where entries are drawn according to the mask from Figure 10(a). Subplots (a)–(b), respectively, show the reconstruction from the direct optimization using the subsampled DtN matrix whose entries are drawn according to the four masks shown in Figure 10.



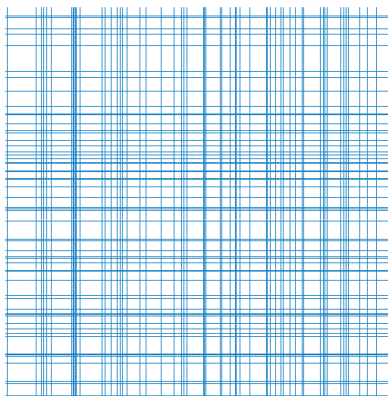
(a) Rank sampling



(b) Uniform sampling



(c) Uniform sampling of columns



(d) Uniform sampling of rows and columns

FIG. 10. Masks used to subsample the full (exact) DtN matrix Λ_a^h . Each blue dot corresponds to a sampled entry, whereas the white color corresponds to locations of the entries not being selected. In all the sampling configurations we extract 20% of the total data in the matrix.

is where lights are shined into the media, and one takes measurement on the “outflow” part of the boundary

$$\Gamma_+ = \{(x, v) : x \in \partial\mathcal{D}, v \cdot n_x > 0\}.$$

The map that directs incoming data to the outgoing data is known as the albedo operator and is used to reconstruct σ_s .

Figure 11(a) plots the albedo matrix (discretized albedo operator), along with its eigenvalues in Figure 11(b) in the diffusion regime $\text{Kn} \ll 1$. It is clear that the operator is ALR. In this case, one would be able to approximate the full operator by solving the optimization problem in (2.13).

In the ballistic regime, $\text{Kn} \sim 1$, then the albedo matrix, plotted in Figure 12(a), is no longer of low rank. However, it is approximately an $\mathcal{H}$ -matrix. A partition of the albedo matrix is shown in Figure 12(b), and we plot the ϵ -rank (with $\epsilon = 10^{-6}$)

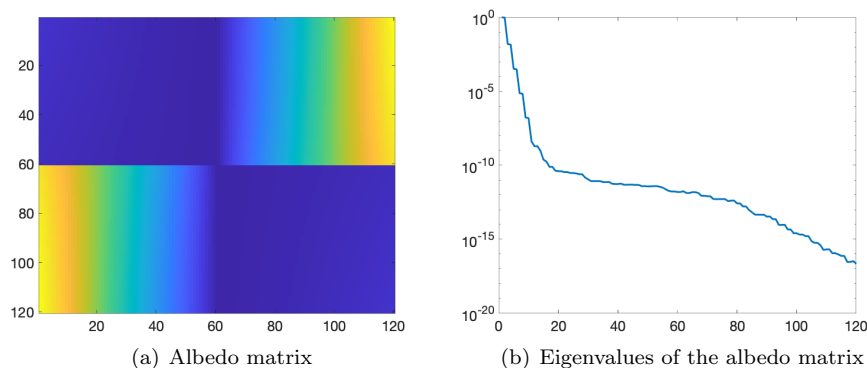


FIG. 11. Panel (a) colors the albedo matrix in the diffusive regime when $\text{Kn} = 2^{-5}$ with the values of its entries. Panel (b) plots eigenvalues of the albedo matrix.

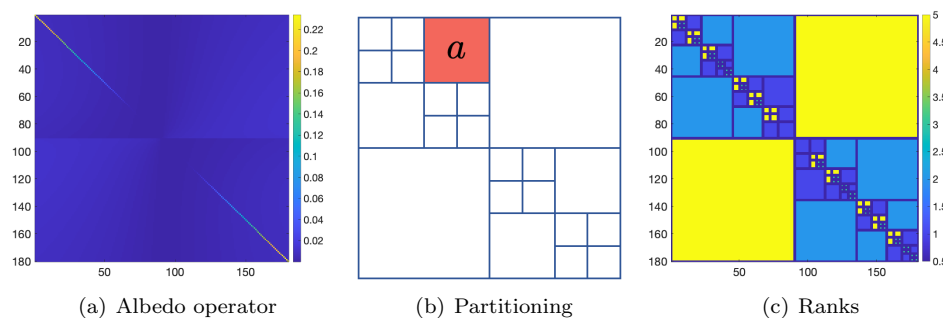


FIG. 12. From left to right, plot of the albedo matrix for $\text{Kn} = 1$, partitioning of the albedo matrix, and the ϵ -rank of each block in the partition. Here each block is colored with its ϵ -rank.

for all the blocks in Figure 12(c). As can be seen, the ϵ -rank is uniformly bounded by 5 in each block. Analogous to the DtN matrix, we present the reconstruction of one typical off-diagonal block in this albedo matrix. For the block shown in Figure 12(b), we select data according to the Bernoulli distribution with parameter p , and the selected entries serve as given data in the matrix completion algorithm. Figure 13 plots the success ratio, computed with 20 experiments for each p and refinement level. Here success means the reconstructed matrix is within 10^{-4} error in Frobenius norm of the groundtruth. It is clear that the chance of successful reconstruction increases as the dimension of the matrix increases, as predicted by the theory. We leave the detailed bridging-the-gap analysis and parameter reconstructions for future work.

5. Conclusions. There is a gap between theoretical and numerical approaches for inverse problems. While in theory, infinite-dimensional datasets, encoded in the ItO map, are available to infer a function living in infinite-dimensional space, in the numerical and experimental settings, both the available data and the reconstructed parameter are finite-dimensional. This mismatch prevents the application of the theory in guiding and improving practical computational inverse solutions: In very rare cases can one assert the unique and stable reconstruction of discretized parameters.

We have presented a framework to bridge this gap using data completion with the EIT problem as the testbed. In particular, we view finitely experimental data points

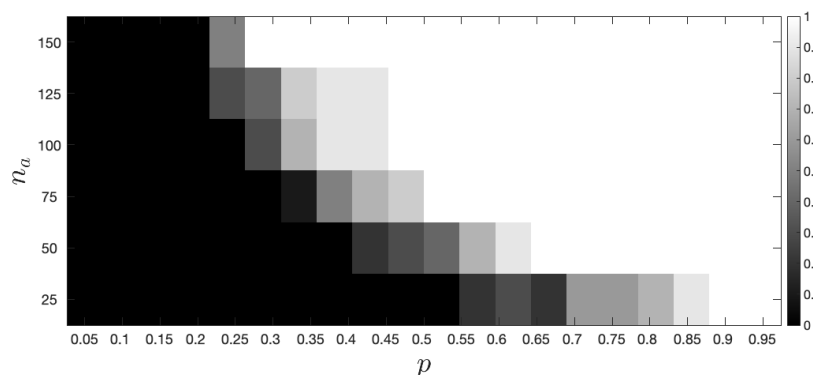


FIG. 13. Success ratio for the reconstruction of block a in Figure 12(b) of the albedo matrix with $\text{Kn} = 1$ at different refinement level and different value of p . In the plot the color encodes the success ratio.

as entries in the ItO (DtN) matrix. Since the DtN matrix has the $\mathcal{H}$ -matrix structure, we exploit the off-diagonally approximately low-rank property and the matrix completion technique to informatively collect mostly random data points in the matrix, and we fill in the unknown entries with a matrix completion method. The goal of matrix completion is twofold: (I) *bridging the gap* and (II) *improving the quality of computational inverse solutions*. For (I), the DtN map is rigorously recovered by lifting the completed DtN matrix, up to discretization error with high probability. This allows us to apply the inverse theory to asymptotically show the unique and stable reconstruction of parameters. For (II) we have numerically demonstrated that—unlike traditional computational inverse problems that use the incomplete DtN matrix—we deploy the completed DtN matrix to reconstruct the unknown parameters. The numerical results have shown that the reconstructions using the completed DtN matrix and the exact DtN matrix are visibly identical, while the reconstruction directly from incomplete DtN matrix is completely off.

We emphasize that the goal of the current paper is to propose a general framework to bridge and improve theoretical and computational inverse problems. For a thorough error analysis, we need a more precise estimate of the decay of the singular values in each block of the ItO matrix. This highly depends on the specific equation encoded in the forward map. This part of error analysis is not yet available in its most precise form in the literature and thus is left for future work.

Appendix A. Proof of Theorem 3.7. Let us define $\tilde{\Phi}$ as the unique solution of the following problem:

$$(A.1) \quad \int_{\Omega} \mathbf{a} \nabla \tilde{\Phi} \cdot \nabla v \, d\Omega = 0 \quad \tilde{\Phi}|_{\partial\Omega} = \phi^h \quad \forall v \in H_0^1(\Omega)$$

where again $\phi^h = \Pi^h \phi$. Let us denote $\Lambda_a^\dagger$ via

$$(A.2) \quad \langle \Lambda_a^\dagger \phi, \psi \rangle := \int_{\Omega} \mathbf{a} \nabla \tilde{\Phi} \cdot \nabla \Psi \, d\Omega,$$

where $\Psi \in H^1(\Omega)$ can be any extension of ψ such that $\Psi|_{\partial\Omega} = \psi$.

LEMMA A.1. *There holds*

$$(A.3) \quad \|\Lambda_a - \Lambda_a^\dagger\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} \leq c \|\mathbb{I} - \Pi^h\|_{H^{1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)},$$

where $\mathbb{I}$ is the identity map and c is a constant independent of the mesh size h .

Proof. By definition we have

$$\begin{aligned} |\langle \Lambda_a \phi, \psi \rangle - \langle \Lambda_a^\dagger \phi, \psi \rangle| &= \left| \int_{\Omega} \mathbf{a} \nabla (\Phi - \tilde{\Phi}) \cdot \nabla \Psi \, d\Omega \right| \\ &\leq c \|\psi\|_{H^{1/2}(\partial\Omega)} \left\| \nabla (\Phi - \tilde{\Phi}) \right\|_{L^2(\Omega)} \\ &\leq c \|\psi\|_{H^{1/2}(\partial\Omega)} \|\phi\|_{H^{1/2}(\partial\Omega)} \|\mathbb{I} - \Pi^h\|_{H^{1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)}, \end{aligned}$$

where we have used the uniform boundedness of $\mathbf{a}$ and definition (A.1). The estimate (A.3) thus follows. $\square$

Let $\mathbb{P}^h : H^1(\Omega) \ni \tilde{\Phi} \mapsto \mathbb{P}^h \tilde{\Phi} \in V_{\phi^h}^h(\Omega)$, where $V_{\phi^h}^h := \{v \in V^h(\Omega) : v|_{\partial\Omega} = \phi^h\}$, be defined as

$$\int_{\Omega} a \nabla \mathbb{P}^h \tilde{\Phi} \cdot \nabla v^h \, d\Omega = \int_{\Omega} a \nabla \tilde{\Phi} \cdot \nabla v^h \, d\Omega \quad \forall v^h \in V_0^h.$$

Note that $\mathbb{P}^h$ is a well-defined linear bounded map and $\Phi^h = \mathbb{P}^h \tilde{\Phi}$, where Φ^h is the FE solution.

LEMMA A.2. *There holds*

$$\begin{aligned} \|\Lambda_a^\dagger - \hat{\Lambda}_a\|_{H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)} &\leq c \|\mathbb{I} - \mathbb{P}^h\|_{H^1(\Omega) \rightarrow H^1(\Omega)} \\ &\quad + c \|\mathbb{I} - \Pi^h\|_{H^{1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)}, \end{aligned}$$

where $\mathbb{I}$ is the identity map and c is a constant independent of the mesh size h .

Proof. We have

$$\begin{aligned} &\left| \langle \Lambda_a^\dagger - \hat{\Lambda}_a \phi, \psi \rangle \right| \\ &\leq \left| \int_{\Omega} \mathbf{a} \nabla (\tilde{\Phi} - \mathbb{P}^h \tilde{\Phi}) \cdot \nabla \Psi \, d\Omega \right| + \left| \int_{\Omega} \mathbf{a} \nabla \mathbb{P}^h \tilde{\Phi} \cdot \nabla (\Psi - \Psi^h) \, d\Omega \right| \\ &\leq c \|\psi\|_{H^{1/2}(\partial\Omega)} \|\phi\|_{H^{1/2}(\partial\Omega)} \|\mathbb{I} - \mathbb{P}^h\|_{H^1(\Omega) \rightarrow H^1(\Omega)} \\ &\quad + c \|\phi\|_{H^{1/2}(\partial\Omega)} \left(\|\Psi - \tilde{\Psi}\|_{H^1(\Omega)} + \|\tilde{\Psi} - \mathbb{P}^h \tilde{\Psi}\|_{H^1(\Omega)} \right) \\ &\leq c \|\psi\|_{H^{1/2}(\partial\Omega)} \|\phi\|_{H^{1/2}(\partial\Omega)} \left(\|\mathbb{I} - \mathbb{P}^h\|_{H^1(\Omega) \rightarrow H^1(\Omega)} + \|\mathbb{I} - \Pi^h\|_{H^{1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)} \right), \end{aligned}$$

where we have defined $\tilde{\Psi}$ as the solution (A.1) with boundary data $\Pi^h \psi$ and taken $\Psi^h = \mathbb{P}^h \tilde{\Psi}$. $\square$

Proof of Theorem 3.7. We provide the proof of the first assertion, as the others are obvious owing to (3.2) and the triangle inequality. From Lemmas A.1–A.2 and the triangle inequality we need to show that

$$\lim_{h \rightarrow 0} \|\mathbb{I} - \mathbb{P}^h\|_{H^1(\Omega) \rightarrow H^1(\Omega)} = 0 \quad \text{and} \quad \lim_{h \rightarrow 0} \|\mathbb{I} - \Pi^h\|_{H^{1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)} = 0.$$

It is sufficient to prove the former, as the proof for the latter is similar. By definition we have

$$\begin{aligned}\|\mathbb{I} - \mathbb{P}^h\|_{H^1(\Omega) \rightarrow H^1(\Omega)} &= \sup_{\|\Psi\|_{H^1(\Omega)} \leq 1} \sup_{\|\Phi\|_{H^1(\Omega)} \leq 1} \langle (\mathbb{I} - \mathbb{P}^h) \Psi, \Phi \rangle_{H^1(\Omega)} \\ &= \langle (\mathbb{I} - \mathbb{P}^h) \Psi^*, \Phi^* \rangle_{H^1(\Omega)},\end{aligned}$$

where $\langle \cdot, \cdot \rangle_{H^1(\Omega)}$ denotes the inner product in $H^1(\Omega)$ and we have used the fact that the suprema are attainable [29, 30] at some Ψ^* and Φ^* . The density of the FE space $V^h(\Omega)$ in $H^1(\Omega)$ as $h \rightarrow 0$ concludes the proof of the first assertion. $\square$

REFERENCES

- [1] *Compressed Sensing for MRI*, <https://www.siemens-healthineers.com/en-us/magnetic-resonance-imaging/clinical-specialties/compressed-sensing>.
- [2] G. ALESSANDRINI AND S. VESSELLA, *Lipschitz stability for the inverse conductivity problem*, Adv. Appl. Math., 35 (2005), pp. 207–241.
- [3] M. APS, *The MOSEK Optimization Toolbox for MATLAB Manual*, version 9.0, 2019.
- [4] I. BABUSKA AND R. LIPTON, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul., 9 (2011), pp. 373–406, <https://doi.org/10.1137/100791051>.
- [5] M. BEBENDORF, *Approximation of boundary element matrices*, Numer. Math., 86 (2000), pp. 565–589.
- [6] M. BEBENDORF, *Hierarchical Matrices: A Means to Efficiently Solve Elliptic Boundary Value Problems*, Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2008.
- [7] M. BEBENDORF, *Adaptive cross approximation of multivariate functions*, Constr. Approx., 34 (2011), pp. 149–179, <https://doi.org/10.1007/s00365-010-9103-x>.
- [8] J. BENNETT, C. ELKAN, B. LIU, P. SMYTH, AND D. TIKK, *KDD cup and workshop 2007*, SIGKDD Explor. Newsl., 9 (2007), p. 51–52, <https://doi.org/10.1145/1345448.1345459>.
- [9] B. BIONDI, *3D Seismic Imaging*, Society of Exploration Geophysicists, Tulsa, OK, 2006, <https://doi.org/10.1190/1.9781560801689>.
- [10] L. BORCEA, *Electrical impedance tomography*, Inverse Problems, 18 (2002), pp. R99–R136, <https://doi.org/10.1088/0266-5611/18/6/201>.
- [11] T. BUI-THANH, M. DAMODARAN, AND K. WILLCOX, *Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition*, AIAA J., 42 (2004), pp. 1505–1516.
- [12] R. H. BYRD, P. LU, J. NOCEDAL, AND C. ZHU, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput., 16 (1995), pp. 1190–1208, <https://doi.org/10.1137/0916069>.
- [13] E. J. CANDÈS AND Y. PLAN, *Matrix completion with noise*, Proc. IEEE, 98 (2010), pp. 925–936.
- [14] E. J. CANDÈS AND B. RECHT, *Exact matrix completion via convex optimization*, Found. Comput. Math., 9 (2009), 717.
- [15] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97, special issue on mathematics and image analysis.
- [16] Z. CHEN, G. ZHOU, AND Q. ZHAO, *Hierarchical factorization strategy for high-order tensor and application to data completion*, IEEE Signal Process. Lett., 28 (2021), pp. 1255–1259, <https://doi.org/10.1109/LSP.2021.3084511>.
- [17] M. CHENEY, *A mathematical tutorial on synthetic aperture radar*, SIAM Rev., 43 (2001), pp. 301–312, <https://doi.org/10.1137/S0036144500368859>.
- [18] M. CHENEY, D. ISAACSON, AND J. C. NEWELL, *Electrical impedance tomography*, SIAM Rev., 41 (1999), pp. 85–101, <https://doi.org/10.1137/S0036144598333613>.
- [19] H. CHERNOFF, *Sequential Analysis and Optimal Design*, SIAM, Philadelphia, 1972, <https://doi.org/10.1137/1.9781611970593>.
- [20] L. FENG, T. BENKERT, K. T. BLOCK, D. K. SODICKSON, R. OTAZO, AND H. CHANDARANA, *Compressed sensing for body MRI*, J. Magn. Resonance Imaging, 45 (2017), pp. 966–987.
- [21] A. GILLMAN AND P.-G. MARTINSSON, *An $o(n)$ algorithm for constructing the solution operator to 2D elliptic boundary value problems in the absence of body loads*, Adv. Comput. Math., 40 (2014), pp. 773–796, <https://doi.org/10.1007/s10444-013-9326-z>.
- [22] D. GOLDBERG, D. NICHOLS, B. M. OKI, AND D. TERRY, *Using collaborative filtering to weave an information tapestry*, Commun. ACM, 35 (1992), pp. 61–70.

- [23] S. GOREINOV, E. TYRTYSHNIKOV, AND N. ZAMARASHKIN, *A theory of pseudoskeleton approximations*, Linear Algebra Appl., 261 (1997), pp. 1–21, [https://doi.org/S0024-3795\(96\)00301-1](https://doi.org/S0024-3795(96)00301-1).
- [24] W. HACKBUSCH, *A sparse matrix arithmetic based on h-matrices. Part I: Introduction to h-matrices*, Computing, 62 (1999), pp. 89–108.
- [25] W. HACKBUSCH AND B. N. KHOROMSKIJ, *A sparse h-matrix arithmetic. Part II: Application to multi-dimensional problems*, Computing, 64 (2000), pp. 21–47.
- [26] N. HALKO, P. G. MARTINSSON, AND J. A. TROPP, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288.
- [27] M. HINZE, R. PINNAU, M. ULBRICH, AND S. ULBRICH, *Optimization with PDE Constraints*, Springer, Dordrecht, 2009, <https://doi.org/10.1007/978-1-4020-8839-1>.
- [28] T. HOFMANN, *Latent semantic models for collaborative filtering*, ACM Trans. Inf. Syst., 22 (2004), pp. 89–115.
- [29] R. C. JAMES, *Reflexivity and the supremum of linear functionals*, Ann. Math., 66 (1957), pp. 159–169, <http://www.jstor.org/stable/1970122>.
- [30] R. C. JAMES, *Weakly compact sets*, Trans. Amer. Math. Soc., 113 (1964), pp. 129–140.
- [31] A. KIRSCH, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, Cham, 2021, <https://doi.org/10.1007/978-1-4419-8474-6>.
- [32] R. KOHN AND M. VOGELIUS, *Determining conductivity by boundary measurements*, Comm. Pure Appl. Math., 37 (1984), pp. 289–298.
- [33] D. KRESSNER, M. STEINLECHNER, AND B. VANDEREYCKEN, *Low-rank tensor completion by riemannian optimization*, BIT, 54 (2014), <https://doi.org/10.1007/s10543-013-0455-z>.
- [34] R. LAI AND J. LI, *Solving partial differential equations on manifolds from incomplete interpoint distance*, SIAM J. Sci. Comput., 39 (2017), pp. A2231–A2256.
- [35] L. LIN, J. LU, AND L. YING, *Fast construction of hierarchical matrix representation from matrix-vector multiplication*, J. Comput. Phys., 230 (2011), pp. 4071–4087.
- [36] Y. LIU, W. SID-LAKHDAR, E. REBROVA, P. GHYSELS, AND X. S. LI, *A parallel hierarchical blocked adaptive cross approximation algorithm*, Int. J. High Performance Comput. Appl., 34 (2020), pp. 394–408, <https://doi.org/10.1177/1094342020918305>.
- [37] B. L.T., G. O., H. M., AND B. VAN BLOEMEN WAANDERS, *Large-Scale PDE-Constrained Optimization: An Introduction*, Springer, Berlin, 2003, https://doi.org/10.1007/978-3-642-55508-4_1.
- [38] M. DASHTI AND A. M. STUART, *The Bayesian approach to inverse problems*, in Handbook of Uncertainty Quantification, Springer, Cham, 2017, pp. 311–428.
- [39] M. W. MAHONEY AND P. DRINEAS, *Cur matrix decompositions for improved data analysis*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 697–702, <https://doi.org/10.1073/pnas.0803205106>.
- [40] P. G. MARTINSSON, *A fast randomized algorithm for computing a hierarchically semiseparable representation of a matrix*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1251–1274.
- [41] K. MOSEGAARD AND A. TARANTOLA, *Monte Carlo sampling of solutions to inverse problems*, J. Geophys. Res., 100 (1995), pp. 12431–12447.
- [42] G. NAKAMURA AND G. UHLMANN, *Identification of Lamé parameters by boundary measurements*, Amer. J. Math., 115 (1993), pp. 1161–1187.
- [43] F. NATTERER, *The Mathematics of Computerized Tomography*, SIAM, Philadelphia, 2001, <https://doi.org/10.1137/1.9780898719284>.
- [44] F. PUKELSHEIM, *Optimal Design of Experiments*, SIAM, Philadelphia, 2006, <https://doi.org/10.1137/1.9780898719109>.
- [45] A. RAMLATCHAN, M. YANG, Q. LIU, M. LI, J. WANG, AND Y. LI, *A survey of matrix completion methods for recommendation systems*, Big Data Min. Analytics, 1 (2018), pp. 308–323.
- [46] B. RECHT, *A simpler approach to matrix completion*, J. Mach. Learn. Res., 12 (2011), p. 3413–3430.
- [47] B. RECHT, M. FAZEL, AND P. A. PARRILO, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev., 52 (2010), pp. 471–501.
- [48] L. RONDÌ, *A remark on a paper by Alessandrini and Vessella*, Adv. Appl. Math., 36 (2006), pp. 67–69.
- [49] J. F. SCHENCK, *The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds*, Med. Phys., 23 (1996), pp. 815–850.
- [50] J. SYLVESTER AND G. UHLMANN, *A global uniqueness theorem for an inverse boundary value problem*, Ann. of Math., 125 (1987), pp. 153–169.
- [51] J. SYLVESTER AND G. UHLMANN, *The Dirichlet to Neumann Map and Applications*, in Inverse Problems in Partial Differential Equations, D. Colton, R. Ewing, and W. Rundell, eds., SIAM, Philadelphia, 1990, pp. xx–yy.

- [52] W. W. SYMES AND J. J. CARAZZONE, *Velocity inversion by differential semblance optimization*, Geophysics, 56 (1991), pp. 654–663, <https://doi.org/10.1190/1.1443082>.
- [53] A. TARANTOLA, *Inversion of seismic reflection data in the acoustic approximation*, Geophysics, 49 (1984), pp. 1259–1266, <https://doi.org/10.1190/1.1441754>.
- [54] G. UHLMANN, *Inside Out: Inverse Problems and Applications*, Math. Sci. Res. Inst. Publ., Cambridge University Press, Cambridge, UK, 2003.
- [55] J. VIRIEUX, A. ASNAASHARI, R. BROSSIER, L. MÉTIVIER, A. RIBODETTI, AND W. ZHOU, *An introduction to full waveform inversion*, in Encyclopedia of Exploration Geophysics, V. Grechka and K. Wapenaar, eds., Society of Exploration Geophysicists, Tulsa, OK, 2017, pp. R1–1–R1–40, <https://doi.org/10.1190/1.9781560803027.entry6>.
- [56] J. XIA, S. CHANDRASEKARAN, M. GU, AND X. S. LI, *Superfast multifrontal method for large structured linear systems of equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1382–1411.
- [57] A. R. ZHANG, Y. LUO, G. RASKUTTI, AND M. YUAN, *Islet: Fast and optimal low-rank tensor regression via importance sketching*, SIAM J. Math. Data Sci., 2 (2020), pp. 444–479, <https://doi.org/10.1137/19M126476X>.