BRIDGING AND IMPROVING THEORETICAL AND COMPUTATIONAL ELECTRIC 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 manuscript 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 sub-sampled. 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. Dirichlet-to-Neumann map, albedo operator, inverse problem, matrix completion, sparse, $\mathcal{H}$-matrix, electric impedance tomography, optical tomography

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

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 [35, 27, 36] coupled with the development of highly sophisticated algorithmic pipelines [38, 26, 37] for solving inverse problems have fueled several breakthroughs in a myriad of different

[†]Department of Aerospace Engineering and Engineering Mechanics, The Oden Institute for Computational Engineering and Sciences, UT Austin, Austin, Texas (tanbui@oden.utexas.edu, https://users.oden.utexas.edu/~tanbui/).

[‡]Mathematics Department and Wisconsin Institute for Discovery, UW-Madison, Madison, WI (qinli@math.wisc.edu, http://www.math.wisc.edu/~qinli/).

[§]Mathematics Department, UW-Madison, Madison, WI (lzepeda@math.wisc.edu, https://www.math.wisc.edu/~lzepeda/).

fields such as geophysics, astronomy, biomedical imaging, radar, spectrography, signal processing, communications, among many others [12, 39, 8]. Several of such advances, e.g., magnetic resonance imaging (MRI) [34], computarized tomography (CT) [28], and synthetic aperture radar (SAR) [12], permeates 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 amount of receivers, resulting in a finite number 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: suppose 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. This infinite-dimensional setting is certainly computationally infeasible. Thus in all algorithmic studies, one focuses on designing algorithmic pipelines to perform the reconstruction on the finite-dimensional setting: given a finite number of ItO measurements, how to extract the information to infer the unknown parameters, represented by finite-dimensional vectors? Unlike the theoretical infinite-to-infinite approach, this practical finite-to-finite approach uses a finite amount of data pairs to reconstruct the parameter function characterized by a finite number of degrees of freedom.

It is reasonable to believe that the theory should provide guidance and theoretical guarantees for the algorithms' performance. In practice, however, they have mostly advanced in a disconnected manner. Indeed, 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 amount 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 an optimal experimental design (OED) (see, e.g., [29, 13]) is 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 problem to identify (ideally a small number of) data pairs that are informative about the unknown parameters and then, again 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 finitely small information amounts of data is important for real-world applications where data is

often expensive and potentially cumbersome to obtain.

In this manuscript 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 facilitate more accurate reconstruction compared to using the same amount of otherwise arbitrary data.

As have been discussed, the keys to realize our research program are a) the ability to leverage the structure of the underlying problem to sample only a fraction of data and to complete the missing ones, and b) the availability of theoretical results of infinite dimensional inverse problems. The actual executions are thus problem-specific. We choose electric 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 possesses a $\mathcal{H}$-matrix structure. 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 [31]. Finally, the completed DtN matrix is lifted to the DtN map, allowing us to integrate the theoretical results [2, 33] 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 [23]. Moreover, data-driven approaches have been widely used, in which the most notorious example is the application of compressed sensing to MRI [14], 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. Mathematically, this translates to utilizing the Dirichlet-to-Neumann

(DtN) map to reconstruct the diffusion coefficient in the elliptic equation,

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

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):

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

where n 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 $a(x)$:

$$\Lambda_a : \phi \to d,$$

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

$$(2.2) \qquad \int_{\mathcal{D}} a \nabla u \cdot \nabla v \, \mathrm{d}x = 0, \quad \forall v \in H^1_0(\mathcal{D}).$$

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

$$(2.3) \qquad \langle \Lambda_a \phi, \psi \rangle := \int_{\partial \mathcal{D}} a\partial_n u \psi \, \mathrm{d}x,$$

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

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

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 piece-wise constant and is represented uniquely by the vector a containing its values, and we thus use a and 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 Λ^h_{a} 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 Λ^h_{a} are observed and the rest are unavailable. This is translated in the following hierarchy of increasingly reduced objects

$$(2.4) \qquad \Lambda_{\mathsf{a}} \to \Lambda^h_{\mathsf{a}} \to \Lambda^h_{\mathsf{a}}|_\Omega.$$

Here, again, Λ_a is the DtN map, Λ^h_{a} the DtN matrix (discretization of Λ_{a}, whose size depends on h, the mesh size), and Ω is a subset of matrix indices, indicating where measurements are taken. This reduction process is described in details 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 $\Lambda_{\mathsf{a}}^h|_\Omega$, 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 Λ_{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 $\tilde{\Lambda}_{\mathsf{a}}^h$, $\tilde{\Lambda}_{\mathsf{a}}$ and $\tilde{\mathsf{a}}$ the reconstructed approximations to Λ_{a}^h, Λ_{a}, and a respectively. Ω is judiciously selected, such that

$$(2.5) \qquad\qquad \tilde{\Lambda}_{\mathsf{a}}^h \sim \Lambda_{\mathsf{a}}^h \,,$$

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:

$$(2.6) \qquad\qquad \tilde{\Lambda}_{\mathsf{a}} \sim \Lambda_a.$$

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

$$(2.7) \qquad\qquad \tilde{\Lambda}_{\mathsf{a}}^h \sim \Lambda_{\mathsf{a}}^h \quad \Rightarrow \quad \tilde{\Lambda}_{\mathsf{a}} \sim \Lambda_a \quad \Rightarrow \quad \tilde{\mathsf{a}} \sim a \,.$$

In Section 2.3 we lay out the numerical setup. We recover $\tilde{\Lambda}_{\mathsf{a}}^h$ from the sub-sampled $\Lambda_{\mathsf{a}}^h|_\Omega$, and provide intuition to (2.5) in Section 2.4. Recall from section 1 that the construction of $\tilde{\Lambda}_{\mathsf{a}}^h$ 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 *bringing the gap between theoretical and computational EIT*. Section 2.5 discusses a practical computational algorithm for task II) which aims to approximately reconstruct a from the DtN matrix $\tilde{\Lambda}_{\mathsf{a}}^h$. It is important to point out that—unlike traditional computational inverse problems that uses $\Lambda_{\mathsf{a}}^h|_\Omega$, the incomplete DtN matrix, to reconstruct a—we deploy $\tilde{\Lambda}_{\mathsf{a}}^h$, the completed DtN matrix, to reconstruct a. As shall be shown in section 4, our approach *improves the parameter reconstruction substantially*. Indeed, the reconstructions using $\tilde{\Lambda}_{\mathsf{a}}^h$ and the exact DtN matrix Λ_a^h are visibly identical while the reconstruction directly from $\Lambda_{\mathsf{a}}^h|_\Omega$ 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 $K_j, j = 1, \ldots, N_{\text{el}}$ with Lipschitz boundaries. Denote $\mathcal{D}^h := \cup_{j=1}^{N_{\text{el}}} K_j$, $\overline{\mathcal{D}} = \overline{\mathcal{D}}^h$ the discrete space, and $h = \max_j \operatorname{diam}(K_j)$ the mesh size, we construct the standard linear Lagrange finite element (FE) space

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

as the discrete solution space, and

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

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

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

such that

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

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

$$(2.9) \qquad \int_{\mathcal{D}} \mathsf{a}\,\nabla u^h \cdot \nabla v^h\,\mathrm{d}x = 0, \quad \forall v^h \in V_0^h,$$

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

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

$$(2.10) \qquad \left\langle \Lambda_a^h \phi^h, w^h \right\rangle := \int_{\Omega} a\,\nabla\Phi^h \cdot \nabla W^h\,d\Omega,$$

where W^h is any extension of w^h from $V^h(\partial\Omega)$ to $V^h(\Omega)$ such that $W^h\big|_{\partial\Omega} = w^h$ 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. It can be constructed as follows. Let d be the FEM discretization of d, and S the FEM stiffness matrix, then the numerical solution u^h is

$$\mathsf{u}^h = \left(\mathsf{S}^{ii}\right)^{-1} \cdot \mathsf{S}^{ib} \cdot \phi^h,$$

where

$$\mathsf{S} = \left[\begin{array}{cc} \mathsf{S}^{ii} & \mathsf{S}^{ib} \\ \mathsf{S}^{bi} & \mathsf{S}^{bb} \end{array} \right],$$

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. Furthermore, denote M the map from the discrete solution u^h to the discrete Neumann data on the boundary $\partial\mathcal{D}^h$, the DtN matrix Λ_{a}^h can be formed as

$$(2.11) \qquad \mathsf{d} = \Lambda_{\mathsf{a}}^h \cdot \phi = \mathsf{M} \cdot \left(\mathsf{S}^{ii}\right)^{-1} \mathsf{S}^{ib} \cdot \phi, \quad \text{with} \quad \Lambda_{\mathsf{a}}^h = \mathsf{M} \cdot \left(\mathsf{S}^{ii}\right)^{-1} \mathsf{S}^{ib}.$$

Note that Λ_{a}^h is a square matrix of size $|\partial\mathcal{D}^h| \times |\partial\mathcal{D}^h|$, where $\left|\partial\mathcal{D}^h\right|$ denotes the number of grid points on $\partial\mathcal{D}^h$.

From Λ_{a}^h to $\Lambda_{\mathsf{a}}^h|_\Omega$: In practice, only a small number 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 basis functions ϕ_i. In this case the ijth component of the data matrix d is exactly the ijth entry of the DtN matrix, i.e.,

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

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 $\tilde{\Lambda}_{\mathsf{a}}^h$ 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 $\tilde{\Lambda}_{\mathsf{a}}^h$ 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 low-rank blocks regardless of their size. This decomposition allows us to utilize the matrix completion type methods for the low rank 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.

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 [7]. The goal is to complete the entries in a matrix from a partial knowledge of its entries. To be more specific, let a generic $\mathsf{A} \in \mathbb{R}^{n \times n}$ 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 exist a number of algorithms that aims to reconstruct the entries [30, 20, 15, 9]. We adopt the approach proposed in [11]. Under the assumption that the matrix A is of low rank $(r \ll n)$, we seek to minimize the nuclear norm $\|\mathsf{A}\|_*$, such that the matrix's evaluations at certain locations are fixed by the given data, namely $\mathsf{A}_{ij} = a_{ij}$ for $(i, j) \in \Omega$. The minimization problem now becomes:

$$(2.12) \qquad \min_{\mathsf{A}} \|\mathsf{A}\|_* , \quad \text{s.t.} \quad \mathsf{A}_{ij} = a_{ij} , \quad (i, j) \in \Omega , \quad |\Omega| = m .$$

The objective function $\|\mathsf{A}\|_*$ is the sum of all singular values of A. It can be viewed as the relaxation from ℓ_0-norm of singular values (rank$\{\mathsf{A}\}$) to its ℓ_1-norm [32].

One important advantage of working with (2.12)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 $\|\mathsf{A}\|_*$ coincides with the solution of minimizing rank$\{\mathsf{A}\}$.

We next recall the *decoherent* and *delocalization* conditions [11]. Let

$$\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^\top$$

be the singular value decomposition of A.

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

$$(2.13) \qquad \mu(\mathsf{W}) = n \max_{1 \le i \le n} \|\mathcal{P}_{\mathsf{W}} e_i\|_2 ,$$

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

A1 Decoherent condition: $\max(\mu(U), \mu(V)) \le \mu_0$ for some positive μ_0.
A2 Delocalization condition: The maximum entry of $\sum_{1 \le k \le r} u_k v_k^\top$ 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 (2.12):

THEOREM 2.2 ([11, 31]). *Let* A *be an* $n \times n$ *matrix of rank* r *obeying the decoherent 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*

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

for some $\beta > 2$, *then the minimizer to the problem* (2.12) *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 condition, then with 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 2.4.1 to recover the DtN matrix Λ_{a}^h from its data $\Lambda_{\mathsf{a}}^h|_{\Omega}$. 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 [24], to divide the matrix into sub-blocks, most of which are low-rank. The application of the matrix completion algorithm to these low-rank blocks is then expected to be efficient.

Hierarchical matrix, or commonly referred as $\mathcal{H}$-matrix, is a class of matrices that, upon proper partitioning, have fast decays in singular values in the smaller blocks, leading to the low rank property within these blocks. The concept [17, 18] 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., [5] and references therein). In its original form, one also requires linear computational complexity in finding the partitioning. In our setting, such complexity is irrelevant. We point out that several matrix compression methods can be used to reconstruct low-rank blocks, such as adaptive cross approximation [6], randomized SVD [19], or CUR factorizations [25, 16]. However, they require to sample full rows and columns, or they require matrix-vector multiplications, and thus are not applicable for our context in which we have access to only a subset of entries.

It was shown in [5] that the collection of Green's functions for elliptic equations produces an $\mathcal{H}$-matrix. Recall from the expression for Λ_{a}^h in (2.11) that both M and S^{ib} are sparse matrices. Thus if $\left(\mathsf{S}^{ii} \right)^{-1}$ is an $\mathcal{H}$-matrix, so is Λ_{a}^h.

THEOREM 2.3 (Theorem 4.28 of [5]). *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* $\| \left(\mathsf{S}^{ii} \right)^{-1} - \mathsf{C} \|_2 \leq \epsilon$. *For the accuracy compatible with the finite element 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 finite element method.*

The author of [5] furthermore suggests a way to choose the partition. Indeed $\left(\mathsf{S}^{ii}\right)^{-1}$ 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 written as separable function only if x and y are well-separated, the rank is low only for the blocks of $\left(\mathsf{S}^{ii}\right)^{-1}$ that are not along the diagonal. The same observation was made in [24] which shows that the off-diagonal blocks are of low rank.

Built upon these observation, noting that Λ_{a}^h is an $\mathcal{H}$-matrix, with diagonal blocks, $\Lambda_{\mathsf{a},D}^h$, having the full rank, and the off-diagonal blocks, $\Lambda_{\mathsf{a},O}^h$ being approximately low-rank, we propose to obtain the full data in $\Lambda_{\mathsf{a},D}^h$ but a limited entries in $\Lambda_{\mathsf{a},O}^h$, 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 spot on $\partial\mathcal{D}$, one measures the intensity on that particular spot to fill in the diagonal blocks $\Lambda_{\mathsf{a},D}^h$. 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 $\Lambda_{\mathsf{a},O}^h$. 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 $\Lambda_{\mathsf{a},D}^h$ and $\Lambda_{\mathsf{a},O}^h$.
1. Sample each entry in $\Lambda_{\mathsf{a},D}^h$.
2. For each $\Lambda_{\mathsf{a},O}^h$:
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 (2.12) to reconstruct $\Lambda_{\mathsf{a},O}^h$;
end
Output: Assembled the recovery, denoted by $\tilde{\Lambda}_{\mathsf{a}}^h$.

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

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

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

$$(2.14) \qquad \min_{\mathsf{a}} \|\Lambda_{\mathsf{a}}^h - \tilde{\Lambda}_{\mathsf{a}}^h\|_F^2 + \alpha\|\mathsf{a} - \mathsf{a}_0\|_q^q + \beta R(\mathsf{a}),$$

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. Note that even though the first term may seem benign at first glance, the DtN map, Λ_{a}^h, is highly non-linear in 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 Gauss-Newton method to reconstruct the impedance a.

3. Bridging the gap with matrix completion. To show (2.5) amounts to showing the reconstructed DtN matrix $\tilde{\Lambda}_{\mathsf{a}}^h$ 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 Λ_{a}^h into N blocks with each of size $n \times n$ according to the $\mathcal{H}$-matrix decomposition. 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 (2.12), there exist constants C, c such that if*

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

for some $\beta^i > 2$, then

$$\mathbb{P}(\Lambda_{\mathsf{a}}^h = \tilde{\Lambda}_{\mathsf{a}}^h) \geq 1 - c \sum_{i=1}^{N} n^{-\beta^i}.$$

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

Proof. The proof is a straightforward application of Theorem 2.2.   □

Remark 3.2. We assume that the rank of 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 3.3. The theorem states that the two matrices, the reconstructed and the ground-truth, are exactly the same with high probability. In practice, the data obtained in $\Lambda_{\mathsf{a}}^h|_\Omega$ is often polluted with measurement errors. In [10] the authors discuss the effect of such pollution in the reconstruction.

To quantify (2.6) and (2.7), we will rely on some delicate FEM analysis and Theorem 3.1. To begin, we lift both Λ_{a}^h and $\tilde{\Lambda}_{\mathsf{a}}^h$ matrices to their "corresponding" (or reconstructed) DtN maps $\hat{\Lambda}_{\mathsf{a}}$ and $\tilde{\Lambda}_{\mathsf{a}}$ as follows:

$$(3.1) \qquad \tilde{\Lambda}_{\mathsf{a}} := \left(\Pi^h\right)^* \circ \tilde{\Lambda}_{\mathsf{a}}^h \circ \Pi^h, \quad \hat{\Lambda}_{\mathsf{a}} := \left(\Pi^h\right)^* \circ \Lambda_{\mathsf{a}}^h \circ \Pi^h,$$

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

$$\mathbb{P}\left(\tilde{\Lambda}_{\mathsf{a}} = \hat{\Lambda}_{\mathsf{a}}\right) \geq 1 - c \sum_{i=1}^{N} n^{-\beta^i}.$$

Thus, the following results, except Theorem 3.4 and Theorem 3.6, while being valid deterministically for $\hat{\Lambda}_{\mathsf{a}}$, are valid with high probability for $\tilde{\Lambda}_{\mathsf{a}}$. Let us recall the following well-known result.

THEOREM 3.4 ([2, 33]).   *Let a_1 and a_2 be two piecewise constant functions on $\mathcal{D}$, and a_i are their representation vectors. Denote Λ_{a_i} the corresponding DtN map defined in (2.3). Then a as a function of Λ_{a} is Lipschitz:*

$$(3.2) \qquad \|\mathsf{a}_1 - \mathsf{a}_2\|_\infty \le C \|\Lambda_{\mathsf{a}_1} - \Lambda_{\mathsf{a}_2}\|_{H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)}\,,$$

that is, there is a unique conductivity a for every DtN map Λ_{a}.

Since we assume that a is piecewise constant Theorem 3.4 implies that there exists a unique conductivity $\tilde{\mathsf{a}}$ corresponding to $\hat{\Lambda}_{\mathsf{a}}$ such that

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

The following (whose technical proof is given in the Appendix A) is the justification for (2.6) and (2.7), which, similar to Theorem 3.4, shows the uniqueness the reverse DtN hierarchy in section 2.4.

THEOREM 3.5 (Asymptotic Uniqueness).   *There holds*

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

and thus

$$\lim_{h \to 0} \|\mathsf{a} - \tilde{\mathsf{a}}\|_\infty = 0.$$

Let $\tilde{\mathsf{a}}_1$ and $\tilde{\mathsf{a}}_2$ be the conductivities associated with two reconstructed DtN maps $\hat{\Lambda}_{\mathsf{a}_1}$ and $\hat{\Lambda}_{\mathsf{a}_2}$ corresponding to a_1 and a_2, respectively. Then

$$\|\tilde{\mathsf{a}}_1 - \tilde{\mathsf{a}}_2\|_\infty \le C \left(\left\| \Lambda_{\mathsf{a}_1} - \hat{\Lambda}_{\mathsf{a}_1} \right\|_{H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)} \right.$$

$$\left. + \|\Lambda_{\mathsf{a}_1} - \Lambda_{\mathsf{a}_2}\|_{H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)} + \left\| \Lambda_{\mathsf{a}_2} - \hat{\Lambda}_{\mathsf{a}_2} \right\|_{H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)} \right).$$

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

We also briefly discuss the complexity of the matrix completion.

THEOREM 3.6.   *Denote Λ_{a} a matrix of size $n \times n$ with $n \sim 1/h^d$, and suppose that the matrix can be decomposed in $\mathcal{H}$-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 $|\Omega| = \mathcal{O}\left(r n^{6/5} \log n\right)$ known entries sampled properly, we can reconstruct Λ_{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, 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 $\mathcal{O}\left(r (n^\ell)^{6/5} \log \left((n^\ell)^{6/5}\right)\right)$, 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 $\mathcal{O}\left(2^\ell r(n/2^\ell)^{6/5} \log \left((n/2^\ell)^{6/5}\right)\right)$ 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

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

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

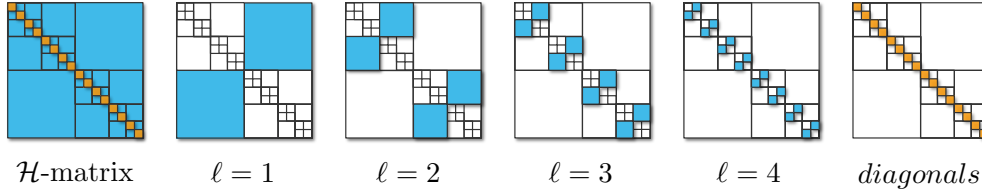where we used the fact that $\sum_{\ell=1}^{n_{levels}} 2^{-\ell/6} = \mathcal{O}(1)$. □

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 $n^{\ell} = n/2^{\ell}$.

Remark 3.7. We have chosen the weak admissibility condition for the sake of simplicity: the theorem follows for other types of partitioning. In particular, as it 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 backend. 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 of low rank. The matrix completion algorithm requires two conditions to be held: the decoherent, and delocalization conditions.

To study the performance of the method, we will first 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 further showcase the reconstruction of the media/conductivity. This final result will be compared with the reconstruction obtained by a complete random sampling of the DtN matrix. The comparison suggests samplings that honor the local low-rank structure of the off-diagonal blocks significantly outperforms a blind random sampling strategy.

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. In subsection 4.3 we demonstrate an extension of the presented method on optical tomography (OT). In OT, 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 low-rank sub-matrices 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 media, 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 low rank, 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 block a) and b) demonstrated in Figure 2(b). As ℓ increases, these blocks have larger sizes accordingly: $n_a = n_h - 1$, and $n_b = (n_h - 1)/2$.

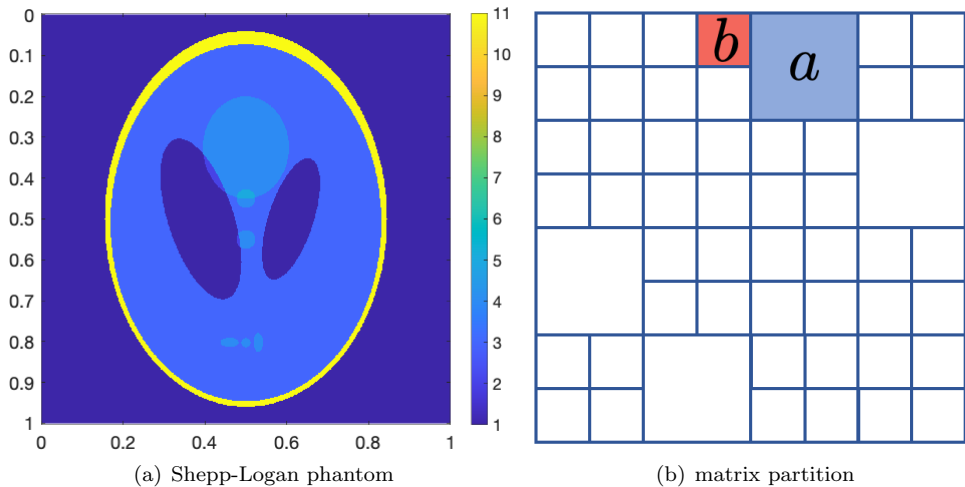(a) Shepp-Logan phantom   (b) matrix partition

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(a) 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 of low rank.
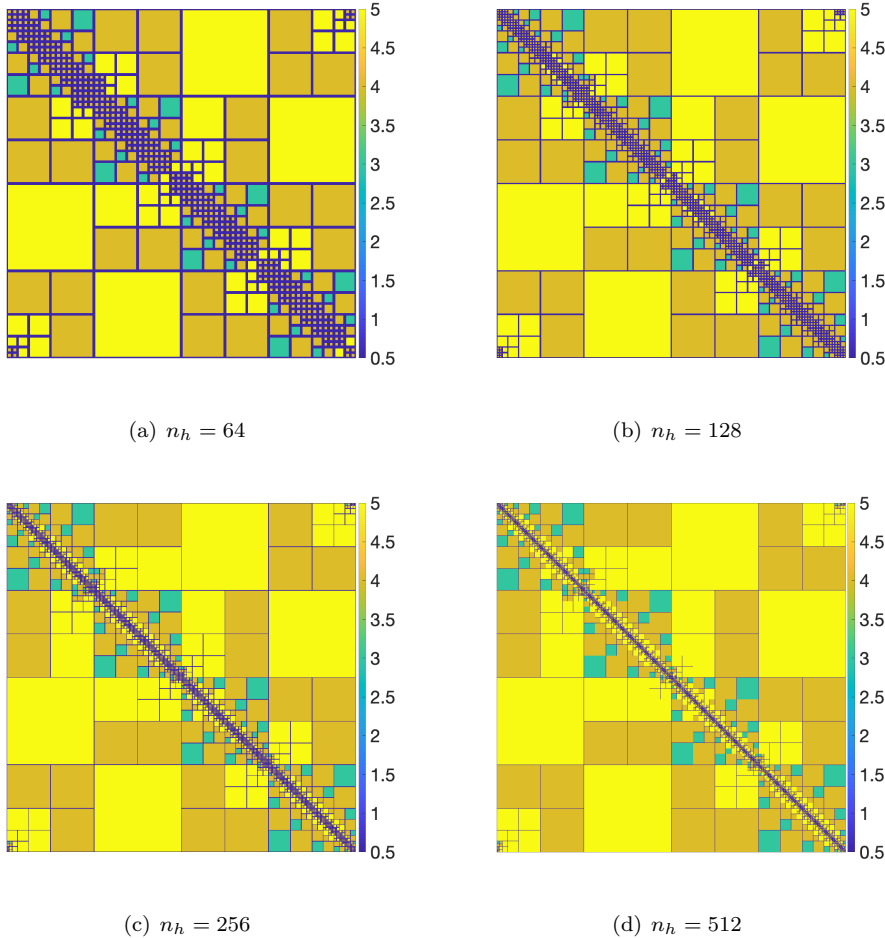

(a) $n_h = 64$

(b) $n_h = 128$

(c) $n_h = 256$

(d) $n_h = 512$

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

To verify the conditions, including the decoherent and delocalization conditions, we plot the coherence indices and the maximum absolute values defined in (2.13) for block 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). These evidence suggest that employing the matrix completion algorithm on off-diagonal blocks will provide satisfying results.

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
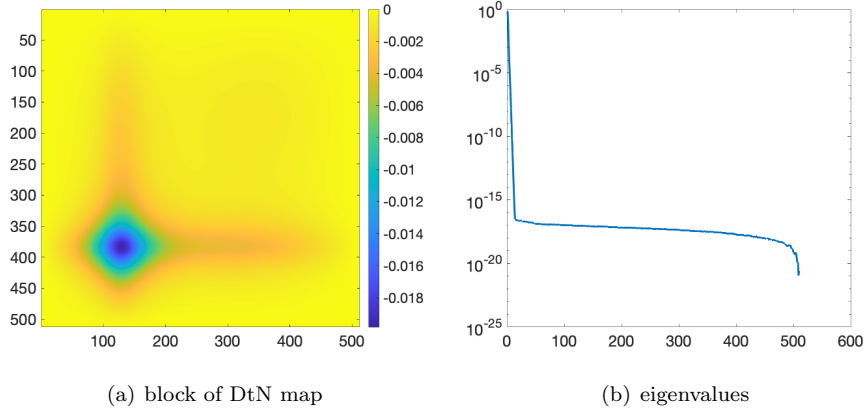
(a) block of DtN map   (b) eigenvalues

Fig. 4: Plot of a typical off-diagonal block of the DtN map, where the color correspond to the value of each entry, along with its eigenvalues.
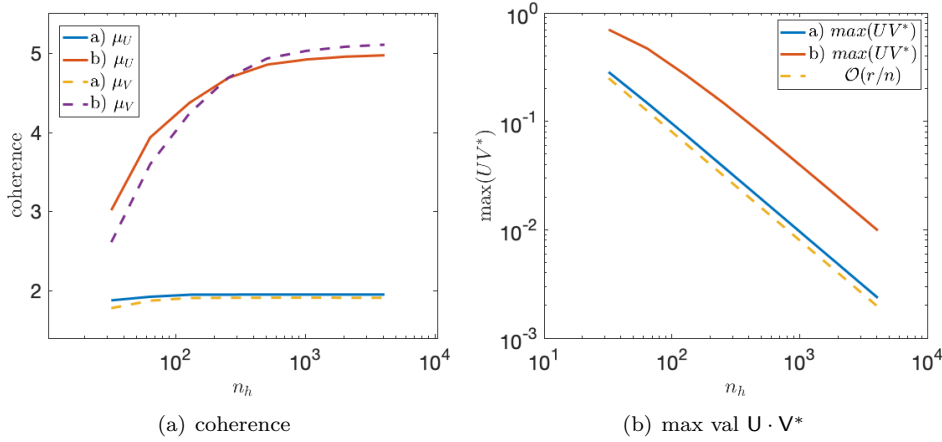

(a) coherence   (b) max val $\mathsf{U} \cdot \mathsf{V}^*$

Fig. 5: Coherence index and maximum absolute value of $\mathsf{U} \cdot \mathsf{V}^*$, for blocks a) and b) at different levels of refinement in the discretization.

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 (2.12). 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 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} error of the ground-truth in Frobenius

(a) original block   (b) mask   (c) reconstruction
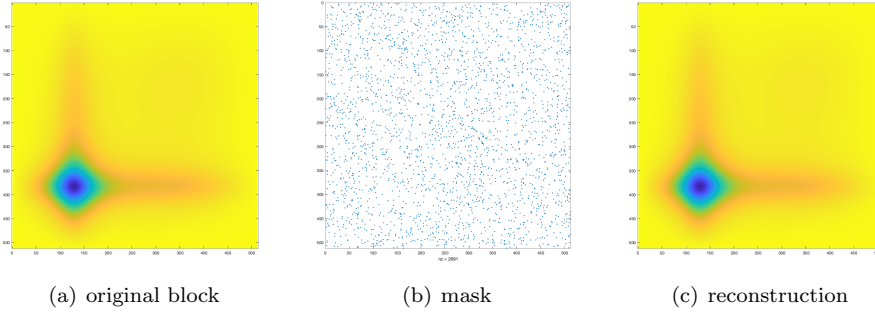
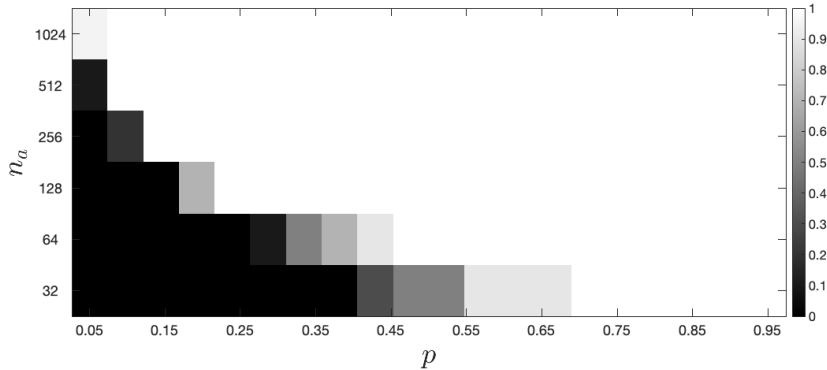Fig. 6: Reconstruction of block a) in the DtN map with with $p = 0.1$.

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.

norm. In Figure 7 we plot the success ratio of reconstructing block a). For low refinement level with coarse discretization, the DtN blocks has small sizes, and the matrix completion algorithm requires a higher percentage of known data for a high success probability of the reconstruction. On refined meshes, 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 (2.14), we use the unconstrained Gauss-Newton method with a constant initial guess. We declare that the optimization algorithm converges if the gradient norm is less than 10^{-9} or the number of iterations exceeds $10,000$. In Figure 9(a) we plot the groundtruth impedance a, and in Figure 9(b) we plot the reconstructed media with the exact DtN matrix Λ_a^h. As can be seen, though the reconstructed media captures the two circular blobs it does not resemble the groundtruth. This is not surprising as the unique reconstruction of the media is guaranteed only with the infinite data limit (i.e. the

full DtN map, instead of the DtN matrix, is known), no discretization, and infinite precision computation. Improved results can be obtained with total variation regularization, for example, to capture sharp edges of the blobs, but this is beyond the scope of this paper. Nevertheless, this reconstruction with the exact DtN matrix provides a benchmark as it is the best case scenario for computation. We now obtain $\Lambda_a^h\big|_\Omega$ by subsampling the exact DtN matrix using the mask in Figure 8 with its design following the criteria in Theorem 3.1. Figure 9(c) shows a reconstruction using the completed matrix $\tilde{\Lambda}_a^h$ obtained from $\Lambda_a^h\big|_\Omega$. We observe that the result is visibly identical with the reconstruction using the exact DtN matrix in Figure 9(b). As a comparison, we also reconstruct the media by solving (2.14) directly with $\Lambda_a^h\big|_\Omega$ instead of the completed DtN matrix $\tilde{\Lambda}_a^h$. As can be seen in Figure 9(d), the reconstruction using $\Lambda_a^h\big|_\Omega$ is not able to capture the two blobs.
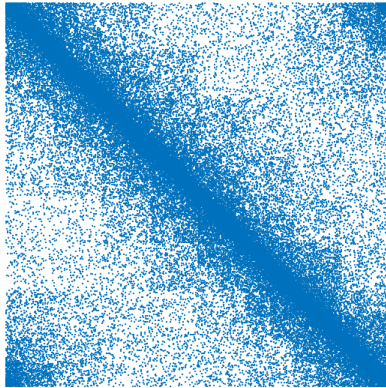

Fig. 8: Mask used to down-sample the full (exact) DtN matrix Λ_a^h. Each blue dot corresponds to a sampled entry, whereas the white color corresponds to not-sampled entries. Note that the number of sampling points is denser close to the diagonal.

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 (RTE) serves as the forward model, and its scattering coefficient—the unknown parameter—reflects the optical property of the media. More specifically, let $f(x,v)$ presents the density of photon particles at location x moving in direction v, then RTE characterizes the dynamics of this distribution function, and in steady state it reads

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

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
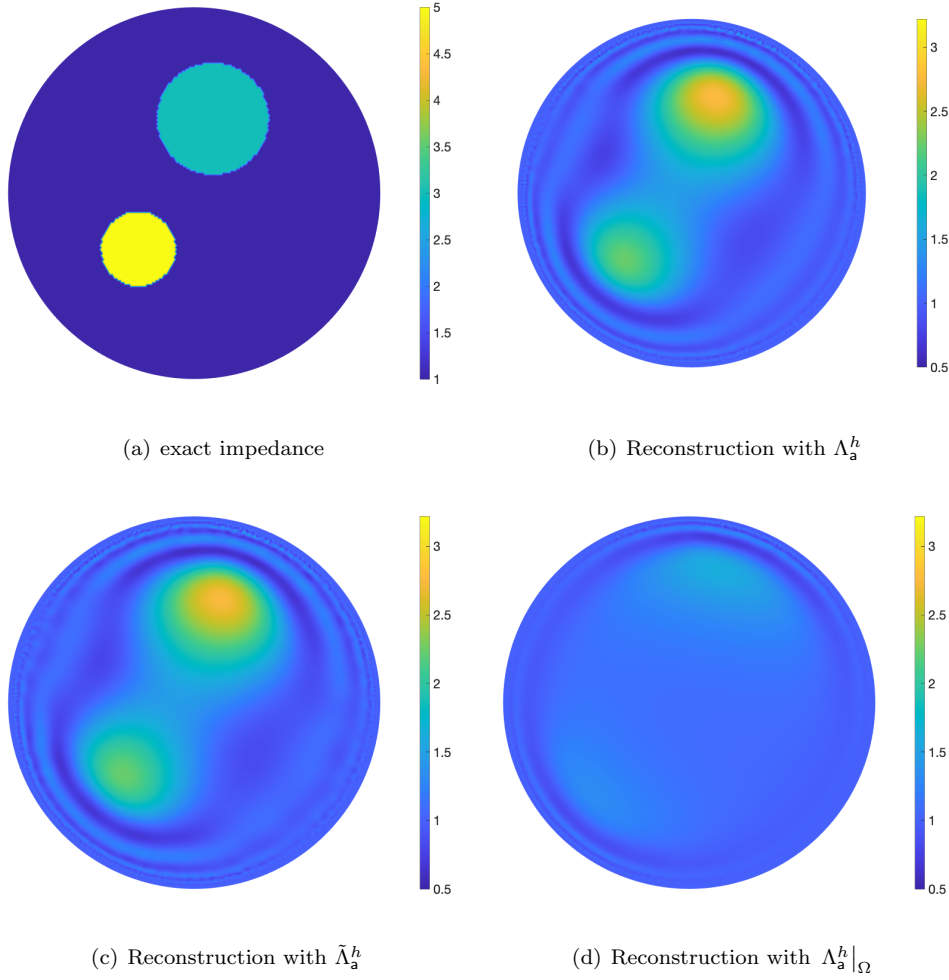
(a) exact impedance

(b) Reconstruction with Λ_{a}^h

(c) Reconstruction with $\tilde{\Lambda}_{\mathsf{a}}^h$

(d) Reconstruction with $\left.\Lambda_{\mathsf{a}}^h\right|_\Omega$

Fig. 9: Reconstructed impedance with the exact DtN matrix Λ_{a}^h, with the completed DtN matrix $\tilde{\Lambda}_{\mathsf{a}}^h$ reconstructed from a subsampled DtN matrix $\left.\Lambda_{\mathsf{a}}^h\right|_\Omega$, and directly with the subsampled DtN matrix.

σ_s. Kn is called the Knudsen number. The "inflow" part of the boundary

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

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 10(a) plots the albedo matrix (discretized albedo operator), along with its eigenvalues in Figure 10(b) in the diffusion regime Kn $\ll$ 1. It is clear that the

operator is approximately low rank. In this case, one would be able to approximate the full operator by solving the optimization problem in (2.12).

In the ballistic regime, $\mathsf{Kn} \sim 1$, then the albedo matrix, plotted in Figure 11(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 11(b), and we plot the ϵ-rank (with $\epsilon = 10^{-6}$) for all the blocks in Figure 11(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 11(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 12 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.
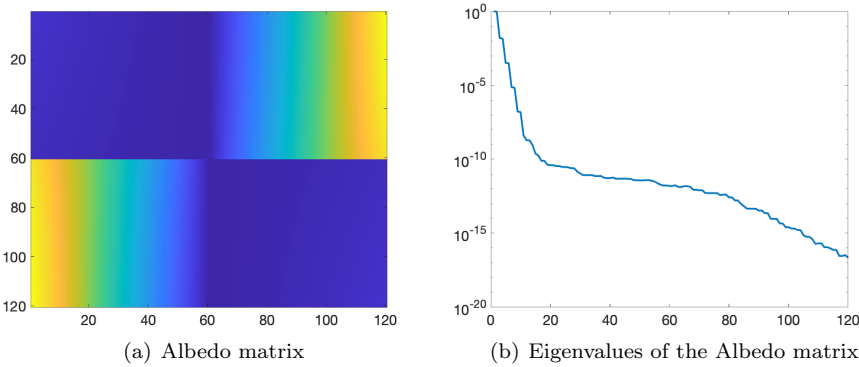

(a) Albedo matrix

(b) Eigenvalues of the Albedo matrix

Fig. 10: Figure 10(a) colors the albedo matrix in the diffusive regime when $\mathsf{Kn} = 2^{-5}$ with the values of its entries. Figure 10(b) plots eigenvalues of the albedo matrix.
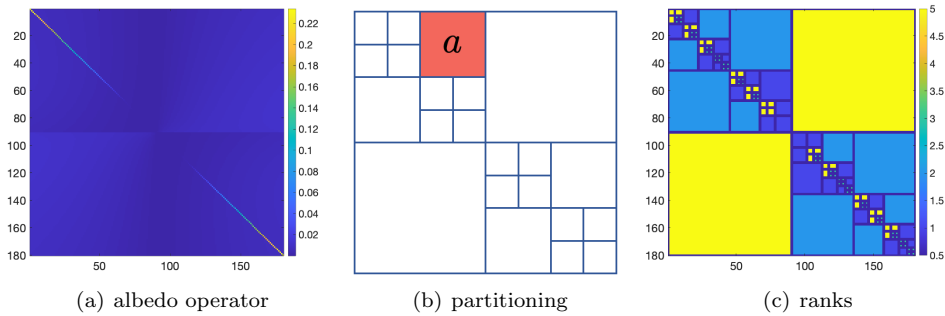

(a) albedo operator

(b) partitioning

(c) ranks

Fig. 11: From left to right, plot of the albedo matrix for $\mathsf{Kn} = 1$, partitioning of the albedo matrix, and the ϵ-rank of each block in the partition. Here each block is colored with its ϵ-rank.
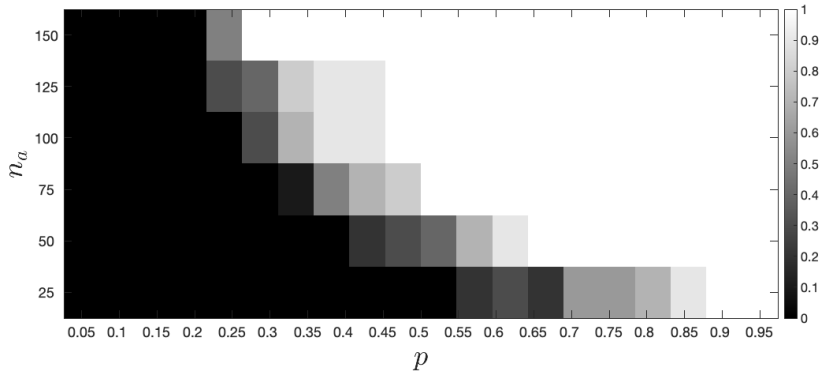
Fig. 12: Success ratio for the reconstruction of block a) in Figure 11(b) of the albedo matrix with $\mathsf{Kn} = 1$ at different refinement level and different value of p. In the plot the color encodes the success ratio.

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 as entries in the input-to-output—Dirichlet-to-Neumann (DtN)—matrix. Since the DtN matrix has the $\mathcal{H}$-matrix structure, we exploit the off-diagonally low-rank property and the matrix completion technique to informatively collect mostly random data points in the matrix, and 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 uses 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 input-to-output 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 3.5. Let us define $\tilde{\Phi}$ as the unique solution of the following problem

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

where again $\phi^h = \Pi^h\phi$. Let us denote $\Lambda_\mathsf{a}^\dagger$ via

$$(A.2) \qquad \left\langle \Lambda_\mathsf{a}^\dagger\phi, \psi \right\rangle := \int_\Omega \mathsf{a}\,\nabla\tilde{\Phi}\cdot\nabla\Psi\,d\Omega,$$

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

LEMMA A.1. *There holds*

$$(A.3) \qquad \left\|\Lambda_a - \Lambda_\mathsf{a}^\dagger\right\|_{H^{1/2}(\partial\Omega)\to H^{-1/2}(\partial\Omega)} \le c\left\|\mathbb{I} - \Pi^h\right\|_{H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)},$$

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

Proof. By definition we have

$$\left|\langle\Lambda_a\phi, \psi\rangle - \langle\Lambda_\mathsf{a}^\dagger\phi, \psi\rangle\right| = \left|\int_\Omega \mathsf{a}\,\nabla\left(\Phi - \tilde{\Phi}\right)\cdot\nabla\Psi\,d\Omega\right|$$

$$\le c\left\|\psi\right\|_{H^{1/2}(\partial\Omega)}\left\|\nabla\left(\Phi - \tilde{\Phi}\right)\right\|_{L^2(\Omega)}$$

$$\le c\left\|\psi\right\|_{H^{1/2}(\partial\Omega)}\left\|\phi\right\|_{H^{1/2}(\partial\Omega)}\left\|\mathbb{I} - \Pi^h\right\|_{H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)},$$

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

Let $\mathbb{P}^h : H^1\left(\Omega\right) \ni \tilde{\Phi} \mapsto \mathbb{P}^h\tilde{\Phi} \in V_{\phi^h}^h\left(\Omega\right)$, where $V_{\phi^h}^h := \left\{v \in V^h\left(\Omega\right) : v|_{\partial\Omega} = \phi^h\right\}$, 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 FEM solution.

LEMMA A.2. *There holds*

$$\left\|\Lambda_\mathsf{a}^\dagger - \hat{\Lambda}_\mathsf{a}\right\|_{H^{1/2}(\partial\Omega)\to H^{-1/2}(\partial\Omega)} \le c\left\|\mathbb{I} - \mathbb{P}^h\right\|_{H^1(\Omega)\to H^1(\Omega)}$$
$$+ c\left\|\mathbb{I} - \Pi^h\right\|_{H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)},$$

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

Proof. We have

$$\left|\left\langle\Lambda_\mathsf{a}^\dagger - \hat{\Lambda}_\mathsf{a}\phi, \psi\right\rangle\right|$$

$$\le \left|\int_\Omega \mathsf{a}\,\nabla\left(\tilde{\Phi} - \mathbb{P}^h\tilde{\Phi}\right)\cdot\nabla\Psi\,d\Omega\right| + \left|\int_\Omega \mathsf{a}\,\nabla\mathbb{P}^h\tilde{\Phi}\cdot\nabla\left(\Psi - \Psi^h\right)\,d\Omega\right|$$

$$\le c\left\|\psi\right\|_{H^{1/2}(\partial\Omega)}\left\|\phi\right\|_{H^{1/2}(\partial\Omega)}\left\|\mathbb{I} - \mathbb{P}^h\right\|_{H^1(\Omega)\to H^1(\Omega)}$$

$$+ c\left\|\phi\right\|_{H^{1/2}(\partial\Omega)}\left(\left\|\Psi - \tilde{\Psi}\right\|_{H^1(\Omega)} + \left\|\tilde{\Psi} - \mathbb{P}^h\tilde{\Psi}\right\|_{H^1(\Omega)}\right)$$

$$\le c\left\|\psi\right\|_{H^{1/2}(\partial\Omega)}\left\|\phi\right\|_{H^{1/2}(\partial\Omega)}\left(\left\|\mathbb{I} - \mathbb{P}^h\right\|_{H^1(\Omega)\to H^1(\Omega)} + \left\|\mathbb{I} - \Pi^h\right\|_{H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)}\right),$$

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

Proof of Theorem 3.5. 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

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

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

$$\left\| \mathbb{I} - \mathbb{P}^h \right\|_{H^1(\Omega) \to H^1(\Omega)} = \sup_{\|\Psi\|_{H^1(\Omega)} \leq 1} \sup_{\|\Phi\|_{H^1(\Omega)} \leq 1} \left\langle \left(\mathbb{I} - \mathbb{P}^h \right) \Psi, \Phi \right\rangle_{H^1(\Omega)}$$
$$= \left\langle \left(\mathbb{I} - \mathbb{P}^h \right) \Psi^*, \Phi^* \right\rangle_{H^1(\Omega)},$$

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 [21, 22] at some Ψ^* and Φ^*. The density of the finite element space $V^h(\Omega)$ in $H^1(\Omega)$ as $h \to 0$ concludes the proof of the first assertion. □

REFERENCES

[1] *Compressed sensing for mri.* https://www.siemens-healthineers.com/en-us/magnetic-resonance-imaging/clinical-specialities/compressed-sensing.

[2] G. ALESSANDRINI AND S. VESSELLA, *Lipschitz stability for the inverse conductivity problem*, Advances in Applied Mathematics, 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 Modeling & Simulation, 9 (2011), pp. 373–406, https://doi.org/10.1137/100791051.

[5] M. BEBENDORF, *Hierarchical Matrices: A Means to Efficiently Solve Elliptic Boundary Value Problems*, Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2008.

[6] M. BEBENDORF, *Adaptive cross approximation of multivariate functions*, Constructive Approximation, 34 (2011), pp. 149–179, https://doi.org/10.1007/s00365-010-9103-x.

[7] J. BENNETT, C. ELKAN, B. LIU, P. SMYTH, AND D. TIKK, *Kdd cup and workshop 2007*, SIGKDD Explor. Newsl., 9 (2007), p. 51–52.

[8] B. BIONDI, *3D Seismic Imaging*, Society of Exploration Geophysicists, 2006, https://doi.org/10.1190/1.9781560801689.

[9] T. BUI-THANH, M. DAMODARAN, AND K. WILLCOX, *Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition*, AIAA Journal, 42 (2004), pp. 1505–1516.

[10] E. J. CANDES AND Y. PLAN, *Matrix completion with noise*, Proceedings of the IEEE, 98 (2010), pp. 925–936.

[11] E. J. CANDES AND B. RECHT, *Exact matrix completion via convex optimization*, Foundations of Computational Mathematics, 9 (2009).

[12] M. CHENEY, *A mathematical tutorial on synthetic aperture radar*, SIAM Rev., 43 (2001), pp. 301–312, https://doi.org/10.1137/S0036144500368859.

[13] H. CHERNOFF, *Sequential Analysis and Optimal Design*, Society for Industrial and Applied Mathematics, 1972, https://doi.org/10.1137/1.9781611970593.

[14] L. FENG, T. BENKERT, K. T. BLOCK, D. K. SODICKSON, R. OTAZO, AND H. CHANDARANA, *Compressed sensing for body mri*, Journal of Magnetic Resonance Imaging, 45 (2017), pp. 966–987.

[15] D. GOLDBERG, D. NICHOLS, B. M. OKI, AND D. TERRY, *Using collaborative filtering to weave an information tapestry*, Communications of the ACM, 35 (1992), pp. 61–70.

[16] S. GOREINOV, E. TYRTYSHNIKOV, AND N. ZAMARASHKIN, *A theory of pseudoskeleton approximations*, Linear Algebra and its Applications, 261 (1997), pp. 1–21.

[17] W. HACKBUSCH, *A sparse matrix arithmetic based on h-matrices. part i: Introduction to h-matrices*, Computing, 62 (1999), pp. 89–108.

[18] W. HACKBUSCH AND B. N. KHOROMSKIJ, *A sparse h -matrix arithmetic. part ii: Application to multi-dimensional problems*, Computing, 64 (2000), p. 21?47.

[19] N. HALKO, P. G. MARTINSSON, AND J. A. TROPP, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Review, 53 (2011), pp. 217–288.

[20] T. HOFMANN, *Latent semantic models for collaborative filtering*, ACM Trans. Inf. Syst., 22 (2004), pp. 89–115.

[21] R. C. JAMES, *Reflexivity and the supremum of linear functionals*, Annals of Mathematics, 66 (1957), pp. 159–169, http://www.jstor.org/stable/1970122.

[22] R. C. JAMES, *Weakly compact sets*, Transactions of the American Mathematical Society, 113 (1964), pp. 129–140, http://www.jstor.org/stable/1994094.

[23] R. LAI AND J. LI, *Solving partial differential equations on manifolds from incomplete interpoint distance*, SIAM Journal on Scientific Computing, 39 (2017), pp. A2231–A2256.

[24] L. LIN, J. LU, AND L. YING, *Fast construction of hierarchical matrix representation from matrix–vector multiplication*, Journal of Computational Physics, 230 (2011), pp. 4071 – 4087.

[25] M. W. MAHONEY AND P. DRINEAS, *CUR matrix decompositions for improved data analysis*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 697–702, https://doi.org/10.1073/pnas.0803205106.

[26] K. MOSEGAARD AND A. TARANTOLA, *Monte Carlo sampling of solutions to inverse problems*, Journal of Geophysical Research, 100 (1995), pp. 12431–12447.

[27] G. NAKAMURA AND G. UHLMANN, *Identification of Lamé parameters by boundary measurements*, American Journal of Mathematics, 115 (1993), pp. 1161–1187.

[28] F. NATTERER, *The Mathematics of Computerized Tomography*, Society for Industrial and Applied Mathematics, 2001, https://doi.org/10.1137/1.9780898719284.

[29] F. PUKELSHEIM, *Optimal Design of Experiments*, Society for Industrial and Applied Mathematics, 2006, https://doi.org/10.1137/1.9780898719109.

[30] A. RAMLATCHAN, M. YANG, Q. LIU, M. LI, J. WANG, AND Y. LI, *A survey of matrix completion methods for recommendation systems*, Big Data Mining and Analytics, 1 (2018), pp. 308–323.

[31] B. RECHT, *A simpler approach to matrix completion*, J. Mach. Learn. Res., 12 (2011), p. 3413?3430.

[32] B. RECHT, M. FAZEL, AND P. A. PARRILO, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review, 52 (2010), pp. 471–501.

[33] L. RONDI, *A remark on a paper by alessandrini and vessella*, Advances in Applied Mathematics, 36 (2006), pp. 67 – 69.

[34] J. F. SCHENCK, *The role of magnetic susceptibility in magnetic resonance imaging: Mri magnetic compatibility of the first and second kinds*, Medical Physics, 23 (1996), pp. 815–850.

[35] J. SYLVESTER AND G. UHLMANN, *A global uniqueness theorem for an inverse boundary value problem*, Annals of Mathematics, 125 (1987), pp. 153–169.

[36] 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, Philadephia, 1990, pp. xx–yy.

[37] 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.

[38] A. TARANTOLA, *Inversion of seismic reflection data in the acoustic approximation*, Geophysics, 49 (1984), pp. 1259–1266, https://doi.org/10.1190/1.1441754.

[39] J. VIRIEUX, A. ASNAASHARI, R. BROSSIER, L. MÉTIVIER, A. RIBODETTI, AND W. ZHOU, *6. An introduction to full waveform inversion*, 2017, pp. R1–1–R1–40, https://doi.org/10.1190/1.9781560803027.entry6.