Toeplitz-plus-Hankel Matrix Recovery for Green's Function Computations on General Substrates

Richard Y. Zhang, *Student Member, IEEE,* Jacob K. White, *Fellow, IEEE*
Department of Electrical Engineering & Computer Science, Massachusetts Institute of Technology,
Cambridge MA, 02139, USA

***Abstract*—Rapidly diversifying technology and declining computational costs are popularizing technologically-flexible simulation and verification techniques, even at some cost in performance. This paper investigates a data-driven, sampling-based approach for computing substrate Green's functions, which is more technology-flexible than specialized layered media methods, at the cost of speed and accuracy. Our method is based on assuming that grid-sampled Green's functions can be well-approximated by Toeplitz-plus-Hankel (TPH) matrices, and uses a least-squares procedure to recover a TPH matrix from a small number of samples. We show that sample location is crucial, and that good sample locations are ones that minimize an associated graph-diameter-based condition number estimate. The method's expected effectiveness is demonstrated on noise-polluted samples of layered media Green's functions. More surprisingly, we show that the method is effective even when applied to a substrate geometry that is only mildly planar.**

***Index Terms*—Toeplitz matrices, Hankel matrices, sampling methods, Green's function methods, interconnnections, inductance, multilayer media**

I. Introduction

FOR decades, developers of circuit- and physical-level simulation and verification have focused on increasing capacity and reducing turn-around time. But the accelerating pace of technology diversification, along with the rapidly expanding availability of computing power, is shifting that focus. Flexible generic approaches are gaining popularity, even if slower or less accurate. More mature examples of this shift include: using scripting languages and automatic differentiation for circuit simulator device models, rather than hand-coded C or C++ [1]; and modeling subsystem input-output buffers using automatically-calibrated generic IBIS models, rather than using expert-designed circuit models [2].

Interconnect technology has always been diverse. Wires can be routed on-chip, on-die, on-board, on-connector, in-air, and in-bundles (e.g. Fig. 1). So it is not surprising that the dominant analysis tools, 3-D electromagnetic field solvers, are already technologically-flexible methods. The situation is quite different for on-chip or on-board interconnect. The relevant geometries are far more complicated, and analyzing substrate effects is critical, as they impact delay and cross-talk, and interfere with achieving timing closure [3]–[6]. General 3-D field solvers are prohibitively time-consuming when used to analyze interconnect over substrates, because they must explicitly discretize both the substrate and the wires.

Instead, specialized approaches have been developed that represent the substrate implicitly using a substrate Green's function [7], [8]. If the substrate is a stack of planar layers, as in Fig. 3, then there are a wide variety of techniques for explicitly computing the substrate Green's function in a (block) Toeplitz-plus-Hankel matrix structure [8]–[14]. In turn, this special matrix structure allows the corresponding system of equations to be efficiently solved using the FFT and an iterative method [15]–[19].

The relentless diversification of technology is also impacting substrates, which can now be layered (Fig. 1), finite (Fig. 2, top), tub-like (Fig. 2, middle), finned (Fig. 2, bottom), or one of many other alternatives [4]–[6], [20], [21]. Such general substrates are often only *approximately* planar. In this paper, we describe a data-driven, sampling-based alternative for computing substrate Green's functions, which can be summarized in three steps:

1) Assume that the substrate Green's function has a (block) Toeplitz-plus-Hankel structure;
2) Collect a small number of Green's function samples, via measurements, simulations, or any other means, at *carefully selected* locations;
3) Use least-squares to *recover* a (block) Toeplitz-plus-Hankel representation to the sampled data.

Compared to specialized methods for layered-media Green's functions, the approach is less efficient and less accurate, because analytical expressions for the Toeplitz-plus-Hankel representation of Green's function are explicitly available [8], [22]. But it is also more generic, because it can be expected to work even for approximately planar substrates that do not admit analytical expressions.

Note that the matrix recovery part of the approach can be highly efficient. Since the Toeplitz-plus-Hankel structure has very few degrees-of-freedom, very few samples are needed to uniquely capture the underlying matrix. But a subtler issue, and the central focus for this paper, is determining where the samples should be taken. The entire approach rests on uniquely recovering a matrix given its samples, but poorly chosen sample locations could render this impractical or impossible, if the influence of an underlying degree-of-freedom is insufficiently probed.

In the next section, we provide background on Toeplitz-plus-Hankel matrices, their relationship with translation invariance and planar reflections in Green's functions, and note the data-compression and computational advantages of the representation. We describe the least-squares recovery algorithm in Section III and in Sections IV and V, we describe the graph theory behind selecting the sample locations. We show, given

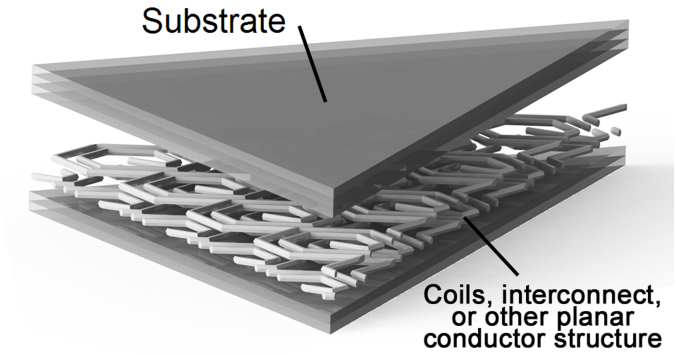


Figure 1: Complex wiring in layered media

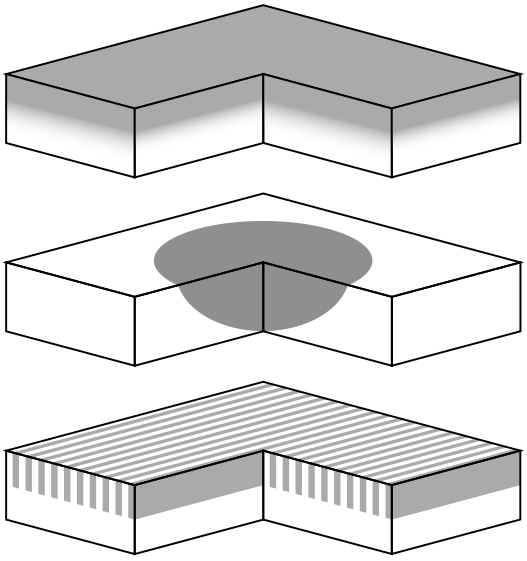


Figure 2: General substrate geometries

well-chosen sample locations, that recovering a Toeplitz-plus-Hankel matrix from sampled data is not only possible, but also reasonably well-conditioned. To illustrate, noise-polluted samples are used to perform the recoveries in Section VI, and the well-conditioning prevents noise from being significantly amplified in the recovered matrix output. Finally, in Section VII, the sampling approach is applied to the electric field Green's function for a general substrate geometry. Results indicate that the Toeplitz-plus-Hankel assumption is surprisingly accurate, even though the domain is only mildly planar. As expected, the assumption degrades when the problem is made fully non-planar.

II. TPH MATRICES

An $n \times n$ matrix, M, is said to be in Toeplitz-plus-Hankel (TPH) form if it can be written as the sum of a Toeplitz (constant diagonal) and a Hankel (constant anti-diagonal) matrix, as in

$$M = T + H \tag{1}$$
$$= \underbrace{\begin{bmatrix} \ddots & t_3 & t_2 & t_1 \\ \ddots & t_4 & t_3 & t_2 \\ \ddots & t_5 & t_4 & t_3 \\ & \ddots & \ddots & \ddots \end{bmatrix}}_{\text{Toeplitz}} + \underbrace{\begin{bmatrix} h_1 & h_2 & h_3 & \iddots \\ h_2 & h_3 & h_4 & \iddots \\ h_3 & h_4 & h_5 & \iddots \\ \iddots & \iddots & \iddots & \end{bmatrix}}_{\text{Hankel}},$$

in which the scalars $t_1, \ldots, t_{2n-1}$ and $h_1, \ldots, h_{2n-1}$ are respectively known as its Toeplitz and Hankel generators.

The definition can also be generalized to block matrices. Suppose that every $t_1, \ldots, t_{2n-1}$ and every $h_1, \ldots, h_{2n-1}$ in (1) are instead defined to be $n \times n$ TPH matrices. Then M would be of size $n^2 \times n^2$, exhibiting the TPH pattern over its matrix blocks, and also within each individual matrix block. In this paper, we refer to such a block matrix as a 2-level TPH matrix, and the original TPH matrix as a 1-level TPH matrix where necessary. Obviously, the definition may be recursively extended to d levels: every matrix block of a d-level matrix is itself a $d-1$-level TPH matrix.

Trivially, every Toeplitz matrix and every Hankel matrix is also a TPH matrix. Likewise, every Toeplitz-block-Toeplitz, Toeplitz-block-Hankel, and Toeplitz-block-Hankel matrix is also a 2-level TPH matrix. The set of d-level TPH matrices is closed under addition: the sum of two d-level TPH matrix yields a third d-level TPH matrix.

A. Application to electromagnetic Green's functions

Fundamental to the analysis of interconnects is the computation of an electromagnetic field quantity (e.g. voltages, flux) given values for a source quantity (e.g. currents, charge) [18], [23]–[25]. For example, in the electric field integral equation (EFIE) over three-dimensions, the $\hat{x}$-directed electric field, E_x, measured at the observation point, $\mathbf{r} = (x, y, z)$, due to $\hat{z}$-direct electric currents, J_z, is given by the integral equation

$$E_x(\mathbf{r}) = \int G_{xz}(\mathbf{r}, \mathbf{r}') J_z(\mathbf{r}')\, d\mathbf{r}'. \tag{2}$$

Similar integral equations can also be reformulated for combinations of various potential fields, charges, and currents [7], [26], [27].

In all cases, the kernel of integration in (2) is known as the Green's function. Serving an analogous role to the impulse response in signals and systems, the Green's function captures the pairwise interaction between point-sources and point-measurements. For example, the value $G_{xz}(\mathbf{r}_1, \mathbf{r}_2)$ corresponds to the $\hat{x}$-directed electric field at $\mathbf{r} = \mathbf{r}_1$, due to a $\hat{z}$-directed current dipole placed at $\mathbf{r}' = \mathbf{r}_2$.

Green's functions that implicitly capture the presence of reflective planar boundaries often yield Toeplitz-plus-Hankel matrices when tabulated over uniform grids. To illustrate, consider the EFIE Green's function from before, which is translation invariant with the form

$$G_{xz}(\mathbf{r}, \mathbf{r}') = \hat{G}_{xz}(x - x', y - y', z - z').$$

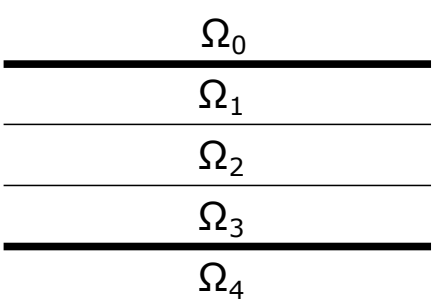


Figure 3: Infinite-extent, planar, piecewise-constant structure that admits closed-form Green's functions.

Suppose a ground plane, modeled as a perfect electric conductor (PEC), were added along the plane $z = 0$. Then by the method of images (see e.g. [27, Ch. 3]), every point-source at $\mathbf{r}' = (x', y', z')$ would be complemented with a mirror point-source at $\mathbf{r}'' = (x', y', -z')$, and their combined response is the new Green's function

$$\begin{aligned} G_{xz}^{PEC}(\mathbf{r}, \mathbf{r}') =& \hat{G}_{xz}(x - x', y - y', z - z') \\ & - \hat{G}_{xz}(x - x', y - y', z + z'). \end{aligned}$$

For every fixed x, x', y, y', tabulating the function G_{xz}^{PEC} over a uniform grid in $\hat{z}$ yields a TPH matrix. Repeating this over different values of x, x', y, y' and stacking the result in turn yields a 3-level TPH matrix that is strictly Toeplitz over its second and third levels.

More formally, for the region above or in-between a multilayer planar structure of infinite extent, as in Fig. 3, each $\hat{p}\hat{q}$-dyad of the associated Green's function will have the form

$$G_{pq}^{ML}(\mathbf{r}, \mathbf{r}') = T_{pq}(\rho, z - z') + H_{pq}(\rho, z + z'), \qquad (3)$$

where $\rho = \sqrt{(x - x')^2 + (y - y')^2}$ is the cylindrical radius [8], [22]. The multilayer planar case is so pervasive that much of the literature on layered- media Green's functions is focused on efficient semianalytical evaluation of T_{pq} and H_{pq} in (3). Then, TPH matrices are easily obtained from (3) by tabulating G_{pq}^{ML} over a uniform grid in $\hat{z}$ and fixed ρ.

B. Computational advantages

The TPH description offers significant data compression when a Green's function is tabulated over a uniform grid. For example, sampling a generic Green's function, $G(\mathbf{r}, \mathbf{r}')$, over a uniform grid of $n \times n \times n$ yields a total of n^6 tabulated values. But if the Green's function is known to be a 3-level TPH matrix, then one can show that there are at most $(4n - 2)^3$ underlying degrees-of-freedom. In the case of layered media (e.g. Fig. 3), these degrees-of-freedom can be directly evaluated and tabulated, using well-established semianalytical methods [8]–[14]. The reduction in formation and storage costs are cubic with respect to n: at $n = 10$, the TPH structure reduces costs by 18 times, but at $n = 100$, this figure grows to more than 15,000 times.

Another aspect of TPH matrices is that they are inexpensive to *apply* to vectors, and hence, inexpensive to *solve* using iterative methods [28]. In the electromagnetics community, it is well-known that the FFT can be used to compute the product of a $n \times n$ Toeplitz matrix and a length-n vector in order $n \log n$ operations [15], [16], [29]. Less well known, but important in this case, is that the FFT, turned "sideways", can also be used to apply a $n \times n$ Hankel matrix to a length-n vector in order $n \log n$ operations. The combination means that if a tabulated Green's function matrix is given as the sum of (block) Toeplitz and Hankel matrices, then it can be applied in order $n \log n$ operations using the FFT [11], [18], [19], [30]–[32].

C. Extraction from data

A $n \times n$ TPH matrix, M, has only $4n - 2$ degrees of freedom, so it is not surprising that all of M can be recovered by sampling a small subset of its entries. Equivalently, it is also possible to infer the T and H matrices from samples of M, though the decomposition is not unique. As discussed in the introduction, a more important issue is determining what samples of M would ensure that that inferring all of M is well-conditioned.

Samples of M should be taken along *entire columns*. To see this, consider tabulating a simple Green's function, $G(z, z')$, via numerical simulations (or hypothetical physical measurements). Every column of M corresponds to the field response measured everywhere on the uniform grid due to a single point-source. Hence, an entire column of M is obtained simultaneously with just single run of numerical simulation program (or a single hypothetical experimental set-up). Moving the point-source along all grid points and repeating the procedure fills the entire matrix. Alternatively, using just a few carefully chosen columns, corresponding to the system response to a few *carefully chosen point-source locations*, all remaining entries of the table may be recovered using the procedure in Section III.

It is worth noting that most electromagnetic Green's functions are singular at $\mathbf{r} = \mathbf{r}'$. In these cases, it is always possible to define a simple analytic function, $G_F(\mathbf{r} - \mathbf{r}')$, that shares the same type of singularity, so that their difference

$$G_S(\mathbf{r}, \mathbf{r}') = G(\mathbf{r}, \mathbf{r}') - G_F(\mathbf{r} - \mathbf{r}')$$

is smooth and easily interpolated from tabulated values [7], [8], [26]. The TPH estimation procedure can then be applied to the smooth remainder G_S.

III. Least-Squares TPH Matrix Recovery

Consider the $n \times n$ Toeplitz-plus-Hankel matrix M, defined

$$M(i, j) = t(i - j - n) + h(i + j - 1) \qquad \forall \{i, j\} \in J, \quad (4)$$

where the set $J = \{\{1, 1\}, \{2, 1\}, \ldots, \{n, n\}\}$ enumerates all matrix element indices in M. Given a subset of the elements of M, whose set of associated indices is denoted by $S \subseteq J$, M can be recovered by solving the minimization problem,

$$\begin{aligned} & \min_{T, H \in \mathbb{R}^{n \times n}} \|T + H - M\|_S, \\ & \text{subject to } T \in \text{Toeplitz}, \ H \in \text{Hankel}, \end{aligned} \qquad (5)$$

where $\|\cdot\|_S$ denotes the matrix Frobenius norm, restricted to elements with indices in S, as in

$$\|X\|_S := \sqrt{\sum_{\{i,j\}\in S} |X(i,j)|^2}.$$

Similar least-squares formulations have appeared in the context of signal processing [33]

In this section, the matrix-norm minimization problem (5) is converted into a (diagonally-scaled) least-squares problem. In the process, we show that in the presence of noise, the efficacy of matrix recovery is determined entirely by the choice of samples. Once well-chosen samples are collected, obtaining the right TPH matrix recovery is relatively routine.

A. Conversion to least-squares problem

The Toeplitz and Hankel matrix constraints in (4) can be summarized by a pair of $\{0,1\}$-matrices E_T and E_H. These matrices relate the $n\times n$ matrices T and H to the $(2n-1)$-long Toeplitz and Hankel vectors t and h, as in

$$E_T t = \mathrm{vec}(T), \qquad E_H h = \mathrm{vec}(H). \tag{6}$$

Note that E_T and E_H each have one nonzero per row.

For the case $S = J$, or equivalently given values for all elements of M, we would have a linear system with n^2 rows and $4n-2$ columns,

$$\begin{bmatrix} E_T & E_H \end{bmatrix} \begin{bmatrix} t \\ h \end{bmatrix} = \mathrm{vec}(M), \\ \iff Ex = b, \tag{7}$$

where $b \in \mathbb{R}^{n^2}$ is the vectorized matrix M, $x \in \mathbb{R}^{4n-2}$ is the vertical concatenation of the t and h generating vectors, and E is the mapping between the generators and the TPH matrix. As can be seen from the $n = 6$ example in Fig. 4, every entry of M corresponds to a row of E with exactly two nonzero values of one, and each column of M corresponds to a block of n rows in E.

The Toeplitz-plus-Hankel decomposition of M is not unique, so E must be singular (see Appendix A). We denote $N \in \mathbb{R}^{n\times c}$ as the c-dimensional right-nullspace matrix of E, such that

$$EN = 0. \tag{8}$$

In addition, since the goal was to reconstruct the full M from a subset of its entries, S is usually a strict subset of J. To represent the sampling, we introduce the restriction operator onto the support of S,

$$R_S(i,i) = \begin{cases} 1 & J(i) \in S, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Multiplication with R_S maps a vector to zero everywhere except within the support of S.

Combining (7)-(9), the matrix-optimization problem in (5) can be reduced to a standard least-squares problem,

$$x_{est} = \underset{x\in\mathbb{R}^n}{\arg\min} \|R_S(Ex-b)\|_2^2 + \|N^T x\|_2^2, \tag{10}$$

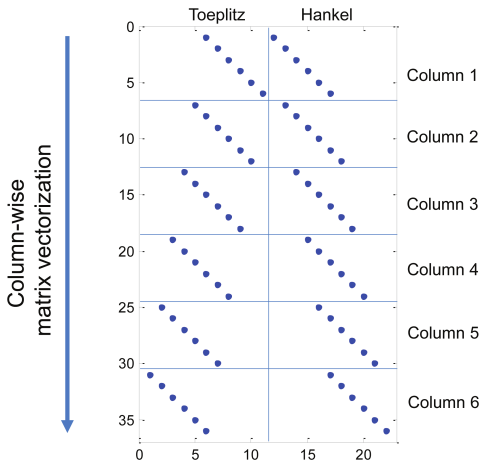


Figure 4: The E matrix, which maps the Toeplitz-plus-Hankel degrees of freedom to elements of the resulting TPH matrix, for a 6×6 example. Note that each dot corresponds to a value of one in the matrix E.

where x_{est} approximates the original generator vector x. An associated approximation to M, denoted M_{est}, can be obtained

$$M_{est} = \mathrm{mat}(Ex_{est}). \tag{11}$$

The minimization problem in (10) is an unconstrained least-squares, with a closed-form solution given by the normal equations,

$$\left[E^T R_S E + NN^T\right] x_{est} = E^T R_S b. \tag{12}$$

B. Error sensitivity & time-complexity

In order to analyze the normal equation error sensitivity, we define κ_S as the condition number restricted to the support of S and orthogonalized to the original null-space. Specifically,

$$\kappa_S(E) = \frac{\lambda_n(E^T R_S E)}{\lambda_{c+1}(E^T R_S E)}, \tag{13}$$

in which $c \triangleq \mathrm{nullity}(E)$ and the eigenvalues are ordered, smallest to largest, as $\lambda_1 \le \cdots \le \lambda_{2n-1}$.

Letting $w \in \mathbb{R}^m$ denote the errors in a given subset of M, and then replacing b with $b + w$ in (12), the reconstruction error $M - M_{est}$ can be related to w using standard matrix perturbation theory [34],

$$\|M - M_{est}\|_F \le \sqrt{\kappa_S(E)}\|w\|_2, \tag{14}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The expression (13) reveals that error sensitivity is driven entirely by the choice of sample locations, S, and not by the underlying values of the matrix.

If we have sampled a good set of entries from M, then we would expect the reconstructed M_{est} to be close to M, even if our samples have errors. Equivalently, a sample set, S, is good if x_{est}, satisfying (12), is reasonably insensitive to errors in b. In other words, a well-chosen S yields a small condition number, κ_S.

On the contrary, poorly chosen samples could cause errors in the data to be magnified. In the worst-case, a sample set could cause the matrix in (12) to become singular. This corresponds to a sample set that "misses" certain degrees-of-freedom in the underlying TPH matrix. In this case, κ_S is unbounded, and the maximum reconstruction error in (14) can be arbitrarily large.

C. Scaled normal equations

The least-squares normal equation (12) is in some sense "artificially" ill-conditioned. The diagonals of $E^T E$ have values that range from 1 to n, and this immediately yields a condition number that grows $O(n)$. It is easy to eliminate this variation by diagonal scaling, sometimes referred to as a Jacobi preconditioning.

Considering defining the diagonal matrix D as

$$D \equiv \mathrm{diag}(E^T R_S E), \tag{15}$$

where the "diag" operator preserves the diagonal entries of its argument, while setting the off-diagonal entries to zero. Using D to define Jacobi-preconditioned versions of E and N,

$$\hat{E} \equiv E D^{-\frac{1}{2}} \qquad \hat{N} \equiv D^{-\frac{1}{2}} N, \tag{16}$$

the scaled normal equation becomes

$$\left[\hat{E}^T R_S \hat{E} + \hat{N}\hat{N}^T\right] \hat{x}_{est} = \hat{E}^T R_S b. \tag{17}$$

Note that diagonal scaling does not interfere with reconstructing M, as easily seen

$$M_{est} = \mathrm{mat}(\hat{E}\hat{x}_{est}). \tag{18}$$

After rescaling, the matrix $\hat{E}^T\hat{E}$ has only 1 along its diagonals. Therefore we would expect the condition number $\kappa_S(\hat{E})$ to be lower than $\kappa_S(E)$.

IV. Optimal Column-based Sample Selection

The results in Section III emphasize the fact that TPH matrix recovery works only when the sample set, S, produces a *well-conditioned* recovery problem. We must balance this objective with taking as few samples as possible. Moreover, driven by the nature of the application, the samples should be collected in a specific pattern, namely as entire columns of the data matrix.

Combined, the problem of picking the best-conditioned columns of the data matrix to sample is a combinatorial eigenvalue optimization

$$S^\star = \underset{S \subseteq K}{\arg\min} \quad |S|\,\kappa_S, \quad \text{subject to } \kappa_S \text{ finite}, \tag{19}$$

such that S be chosen from $K = \{K_1, K_2, \ldots, K_n\}$, where each $K_j = \{\{1,j\},\{2,j\},\ldots,\{n,j\}\}$ denotes an entire column of matrix elements from the data matrix. Note that the minimum number of columns is 4, since there are $2\times(2n-1)$ total degrees of freedom in the TPH matrix (see Section III).

A. Exhaustive search

The optimization in (19) is unstructured and nonconvex, due to its discrete nature and its dependence on the condition number. Regardless, naive exhaustive search can already solve the optimization in *polynomial* time, if the maximum number of columns is fixed to a constant, k. This is because the size of the solution space only grows $\binom{n}{k}$.

But since k is at least 4, the exhaustive search is very expensive in practice. Efficiency may be significantly enhanced by "presolving" the first two columns. More specifically, one can show that every feasible solution to (19) must necessarily satisfy the following statement.

Theorem 1 (Essential columns)**.** *Every feasible selection must necessarily contain the first and last columns, as in*

$$K_1, K_n \in S.$$

Proof: See Section V. ∎

Incorporating this result allows the optimal k-column selection to be provably obtained with a complexity of $O(n^{k-2})$.

B. Four-Column Selection

For a small number of columns, k, it is possible to analyze the structure of the underlying problem. Following this, Section V describes a framework for the optimization based on elements of graph theory. The feasibility of all $k = 4$ column selection is characterized, and a heuristic measure (the graph diameter) is used as a proxy for conditioning. This leads to a heuristic solution to the $k = 4$ column-selection problem.

Theorem 2 (Minimum diameter 4-column set)**.** *Let n be even, then the selection $S = \{K_a, K_b, K_c, K_d\}$, where*

$$a = 1, \qquad b = 1 + \left\lceil \frac{\sqrt{n}}{2} \right\rceil, \tag{20}$$

$$c = b + \frac{n}{2}, \qquad d = n, \tag{21}$$

will yield a 4-column selection that minimizes the TPH graph diameter in the asymptotic limit $n \to \infty$.

Proof: A proof based on the graph formulation in Section V is the subject of an upcoming paper. ∎

Although it is impossible to prove that these choices are optimal, we observe in our numerical experiments that the solutions in Theorem 2 closely match the performance of the optimal solutions found by exhaustive searching.

C. Five-Column Selection

Even the optimal four-column selection becomes inevitably ill-conditioned with growing problem size, n. The condition number can be improved by sampling additional columns, in a sense proving a level of regularization. This paper proposes a simple linear-complexity heuristic for five-column selection:

1) Chose the first four columns based on Theorem 2;
2) Exhaustively search for the fifth column.

There is no strong reason why the optimal five-column selection ought to share columns with a good four-column solution. Nevertheless, in Section VI, we find this procedure to produce good column selections, with an acceptably small optimality gap from the optimal selection.

V. Graph Formulation

The conditioning of the least-squares TPH matrix recovery procedure can be related to properties of a bipartite graph and its associated graph Laplacian. This way, the column-selection optimization in (19) can be posed as a graph sparsification problem on the TPH graph, where the objective is to delete as many edges as possible while preserving the connectivity properties of the original graph.

The general graph sparsification problem is well-studied, admitting many highly effective deterministic [35]–[37] and randomized [38]–[40] algorithms. However, the constraint to only select edges corresponding to entire columns of the data matrix makes our version of the problem quite different; the theory that relies on selecting individual matrix entries does not directly apply. At the same time, the constraint reduces the dimensionality of the problem, making it easier to analyze in closed-form.

A. Reduction to graph sparsification

Since E in (7) is a binary matrix with two nonzeros per row, it can be interpreted as an edgewise incidence matrix for an undirected graph; one that has $4n-2$ vertices and n^2 edges[1]. But we can be more specific, E is an edgewise incidence matrix for a bipartite graph (the edges in a bipartite graph connect one member from each of two disjoint sets of vertices). This follows from the fact that each entry of M is the sum of one value from the Toeplitz vector t and one value from the Hankel vector h. Correspondingly, each row of E has one nonzero in the first $2n-1$ columns (for the Toeplitz value) and one nonzero last $2n-1$ columns (for the Hankel value).

Formally, E is the incidence matrix for the TPH graph defined below. Note that the graph is dense, because there are $O(n^2)$ edges for $O(n)$ vertices, making visualization difficult even for small cases. A highly stylized representation of the graph is shown in Fig. 5.

Definition 3 (The TPH graph)**.** The size-n Toeplitz-plus-Hankel (TPH) graph $\mathcal{G} := \{\{T, H\}, E\}$ is a bipartite graph, whose edges are defined by the algebraic relation

$$\{t_{i-j+n}, h_{i+j-1}\} \in E \qquad 1 \le i, j \le n, \tag{22}$$

over the vertices $T = \{t_1, t_2, \ldots, t_{2n-1}\}$ and $H = \{h_1, h_2, \ldots, h_{2n-1}\}$.

It is reasonable to speak interchangeably of edges in the TPH graph, $\mathcal{G}$, and indices of the TPH matrix, M, since every edge of $\mathcal{G}$ maps uniquely to a single index of M via (22).

Definition 4. Given a set of indices, S, for a TPH matrix, denote $\mathcal{G}_S = \{\{T, H\}, S\}$ as the corresponding edge-induced TPH subgraph, $\mathcal{G}_S \subseteq \mathcal{G}$, i.e. the graph that share the same vertices as the TPH graph, $\mathcal{G}$, but has edges restricted to $e \in S$.

The TPH subgraph $\mathcal{G}_S$ and the S-restricted least-squares recovery problem in Section III are *co-spectral*. To make this statement precise, we provide a brief definition of graph eigenvalues. Given any graph, $\mathcal{G}$, with vertices V and edges E, the Laplacian matrix, L, is given by

$$L(u,v) = \begin{cases} -1 & \{u,v\} \in E, \\ \deg(v) & u = v, \\ 0 & \text{otherwise}, \end{cases} \qquad \forall u, v \in V,$$

and the normalized Laplacian matrix, $\mathcal{L}$, is defined

$$\mathcal{L}(u,v) = \begin{cases} \frac{-1}{\sqrt{\deg(u)\deg(v)}} & \{u,v\} \in E, \\ 1 & u = v, \\ 0 & \text{otherwise}, \end{cases} \qquad \forall u, v \in V,$$

in which the degree of vertex v, denoted $\deg(v)$, gives a count of the number of edges in $\mathcal{G}$ that originate from v. The eigenvalues of $\mathcal{G}$ are defined as the eigenvalues of its (regular) Laplacian matrix, $\lambda_k(\mathcal{G}) \equiv \lambda_k(L)$, while the normalized eigenvalues of $\mathcal{G}$ are those for the normalized Laplacian matrix, $\hat{\lambda}_k(\mathcal{G}) \equiv \lambda_k(\hat{L})$. In this paper, we take the convention $\lambda_1 \le \cdots \le \lambda_{4n-2}$, and $\hat{\lambda}_1 \le \cdots \le \hat{\lambda}_{4n-2}$.

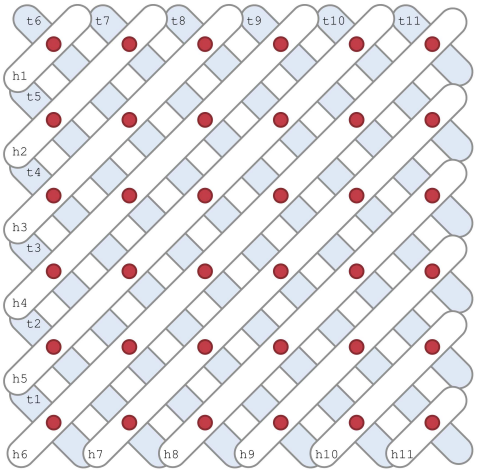


Figure 5: The TPH graph for the $n = 6$ case is depicted in highly stylized form. The 22 vertices (or nodes) in the graph are the elongated ovals labeled $t_1, ..., t_{11}$ and $h_1, ..., h_{11}$. The edges in the graph are black dots, intended to denote "vias" between a T vertex and an H vertex. The figure can also be interpreted as a hypergraph, in which case the elongated ovals are the hyperedges and the black dot vias are the vertices or nodes.

Lemma 5. *The eigenvalues and the normalized eigenvalues of the graph coincide with those of the least-squares systems in (12) and (17):*

$$\lambda_k(E^T R_S E) = \lambda_k(\mathcal{G}_S), \qquad \lambda_k(\hat{E}^T R_S \hat{E}) = \hat{\lambda}_k(\mathcal{G}_S).$$

Proof: Taking E to be the least-squares system in (7), the normal matrix $E^T E$ can be written as the sum of the degree matrix D and adjacency matrix A for $\mathcal{G}$,

$$E^T E = D + A.$$

[1]This paper deviates from the usual graph theory convention of denoting m as the number of edges and n as the number of vertices.

Since the graph $\mathcal{G}$ is bipartite, the matrices $D+A$ and $D-A \equiv L$ are unitarily similar (see [41] for a proof). The spectral equivalence then follows immediately. ∎

The condition number $\kappa_S(\cdot)$, defined in (13), can be expressed via the eigenvalues of $\mathcal{G}_S$ using Lemma 5, as in

$$\kappa_S(E) = \frac{\lambda_n(\mathcal{G}_S)}{\lambda_{c+1}(\mathcal{G}_S)}, \qquad \kappa_S(\hat{E}) = \frac{\hat{\lambda}_n(\mathcal{G}_S)}{\hat{\lambda}_{c+1}(\mathcal{G}_S)}, \tag{23}$$

where $c \triangleq \text{nullity}(E)$ as in (13).

B. Graph connectivity as a metric for feasibility

Returning to the optimization problem (19), we have designated a candidate sample set, S, to be feasible if its condition number, $\kappa_S(E)$, is finite. By our error analysis in (14), only a sample set S with a finite κ_S can successfully recover all elements of a TPH matrix in the presence of noise.

Under the graph formulation, a candidate sample set is feasible if and only if its subgraph is connected.

Lemma 6. *A sample set, S, is feasible in (19) if and only if the subgraph $\mathcal{G}_S$ has exactly 2 connected components.*

Proof: The nullity (i.e. number of connected components) of a graph coincides with the nullity of its Laplacian matrix (see e.g. [42] for a proof). By Lemma 8, the original TPH graph has $\text{nullity}(\mathcal{G}) = 2$. Suppose that $\text{nullity}(\mathcal{G}_S) > 2$. Then $\lambda_{c+1}(\mathcal{G}_S) = 0$, which causes $\kappa_S(E) \to \infty$ in (23). ∎

An immediate application of Lemma 6 is to identify the "essential" elements of M that must be included in order for a sample set S to be feasible. Theorem 1 is one such statement.

Proof of Theorem 1: Columns 1 and n contain edges to the *pendant* vertices, $\{t_1, t_{2n-1}, h_1, h_{2n-1}\}$, each connected to the main graph by a single edge. If these columns are not sampled, then at least one pendant vertex becomes disconnected. By Lemma 6, the resulting sample set is infeasible. ∎

C. The graph diameter as a heuristic for conditioning

A sample set with a small value of κ_S can also be selected under the graph framework. According to the relation in (23), minimizing κ_S is equivalent to maximizing the smallest nonzero normalized eigenvalue, $\hat{\lambda}_{c+1}$; the largest normalized eigenvalue is fixed to $\hat{\lambda}_{4n-2} = 2$ due to $\mathcal{G}$ being bipartite [42, Lemma 1.8].

Regardless, $\hat{\lambda}_{c+1}$ is difficult to optimize using combinatorial arguments, due to its variational, global nature. Instead, Theorem 2 arrives at a selection by using the graph diameter as a proxy for $\hat{\lambda}_{c+1}$. Informally, the diameter is known as the longest "shortest-path" in a graph. It is computed to be the largest number of vertices that must be traversed to travel directly between two vertices on a graph, i.e. without detours, loops or backtracks. For a graph with more than one connected component, we define its diameter as the maximum diameter of its components.

Claim 7. A good column-based sample set S, i.e. one that yields a low condition number $\kappa_S(\hat{E})$ in (23), must also produce a subgraph $\mathcal{G}_S$ with a low diameter.

Proof: Small diameters is a necessary condition for graphs to have large values of $\hat{\lambda}_{c+1}$ [42], [43, Ch.4]. ∎

However, the converse statement is not necessarily true: a subgraph with a low diameter is not obliged to yield a well-conditioned recovery problem. Hence, the graph diameter can only serve as a heuristic to guide the optimization, but cannot be used to prove optimality. Regardless, our result in Theorem 2 was obtained by minimizing the graph diameter, and in practice, we find the heuristic to be highly effective. Our latter numerical results confirm this claim.

VI. Performance Validation

Our results throughout this paper are ultimately rooted in the heuristic of approximating the condition number with the diameter of a certain subgraph. It cannot prove optimality, nor assess the performance of a "typical" feasible solution. To study these properties, we must perform an exhaustive search in the solution space.

A. Comparison to average case

To get a sense of how the four-column choice suggested in Theorem 2 compares against a randomized choice (made while respecting Theorem 1), a cumulative probability distribution is plotted in Fig. 6. The horizontal axis shows the inverse of the the condition number, $1/\kappa_S(\hat{E})$, on a logarithmic scale. Each line plots the percentage of four-column choices that are bettered conditioned than that particular value along the horizontal axis.

The results show that the "average" 4-column choice performs very poorly. Even when the first the columns are correctly selected via Theorem 1, 70% of the choices for the remaining two columns are infeasible, for the problem size of $n = 30$. The portion of infeasible choices increases to 80% for the larger problem size of $n = 70$. The distribution itself also becomes long-tailed as n increases: for $n = 30$, all feasible column choices performs within a factor of 2 of each other, whereas by $n = 70$ the range has expanded to almost 10.

Results show that the four-column choice suggested by Theorem 2 performs very well in practice. In all three problem sizes studied, the optimality gap is sufficiently small to be negligible. The result suggests that the graph diameter is indeed an effective heuristic for $\hat{\lambda}_{c+1}$, for the class of graphs considered in this paper.

B. Optimality gap over a range of problem sizes

Regardless, the choices suggested by Theorem 2 are not optimal for small values of n, so an optimality gap is expected to exist. Figure 7a plots the optimality gap for the four-column selection suggested by Theorem 2, against the global optimum found by exhaustive search. Surprisingly the optimality gap is negligibly small over all problem sizes considered, and does not appear to grow with problem size. This experimental result suggests that Theorem 2 may in fact coincide with the globally optimal four-column solution.

Figure 6b examines the optimality gap of the four-plus-one column selection strategy. Results show a massive improvement in conditioning, even when compared to the globally

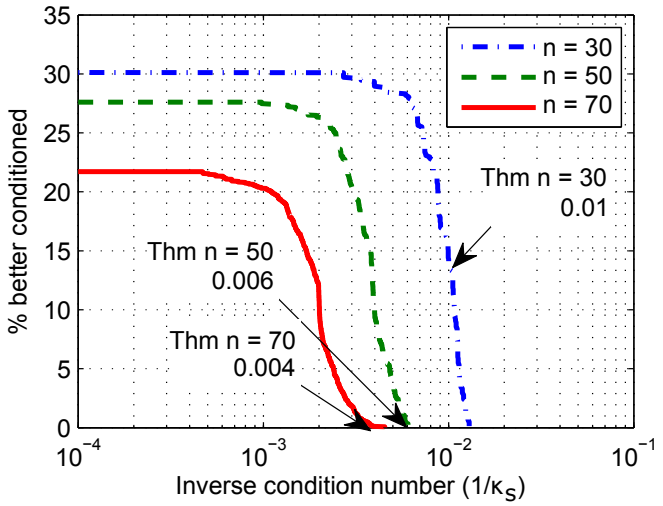


Figure 6: Cumulative probability distribution of the inverse condition number $1/\kappa_S$. Each line plots the percentage of four-column selections that are better conditioned than the value along the horizontal axis.

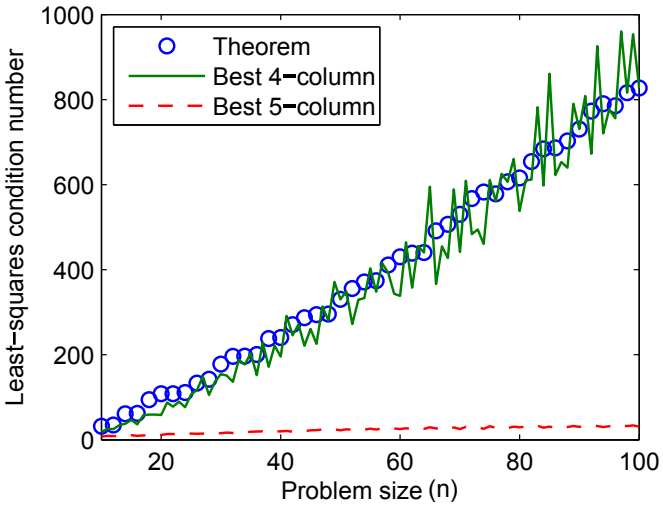
optimal four-column choice. The optimality gap for the four-plus-one strategy is larger in relative terms, and also grows larger with problem size. Regardless, the reduction in the condition number makes it pragmatic choice for real applications.

C. Computational Example

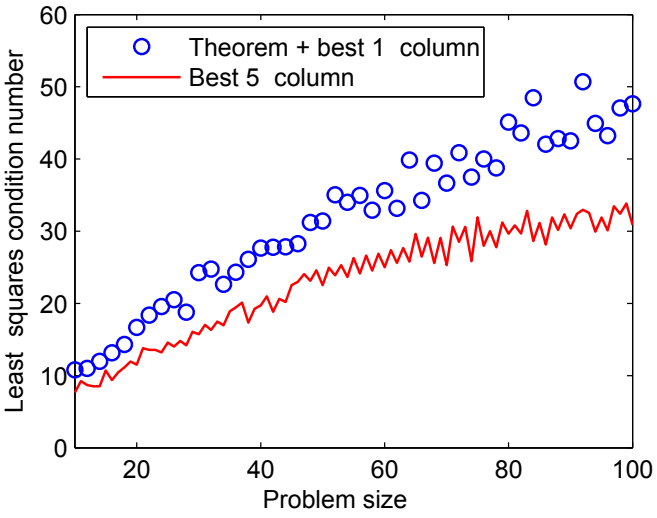
Finally, we illustrate the importance of conditioning in a realistic recovery problem, as well as the effectiveness of the four- and five-column choices in a practical environment. The data for this example is sourced from a realistic TPH matrix in [44], which corresponds to a particular multilayered media Green's function evaluated over a uniform grid.

The data matrix is tainted with 10% additive Gaussian white noise (by Frobenius norm), to result in a signal-to-noise ratio (SNR) of 20 dB. The noise spectrum, in practice, will be determined by the source of the data. Data collected via physical measurements will have errors due to instrument tolerances, while data generated from numerical simulations will have errors due to numerical residuals and modeling errors. Our goal in this subsection is to illustrate the importance of conditioning, without discriminating against a particular data source. For this reason, a flat noise spectrum was used.

We begin by performing the four-column recovery and comparing it against an "arbitrary" choice of four columns. For the example size of $n = 300$, the former is chosen to be the columns $\{1, 10, 160, 300\}$, while the latter is set to $\{1, 75, 225, 300\}$. This latter, "arbitrary" choice was made to simulate a sampling scheme selected via engineering instinct. Since the underlying Green's function is known to be smooth, it is always tempting to sample it using the Gauss-Legendre nodes or Chebyshev nodes, which are optimal for their respective classes of smooth functions. The fourth order Chebyshev nodes over the interval $x \in [-1, 1]$ are $x = \pm\sqrt{2 \pm \sqrt{2}}/2 \approx \{\pm 0.38, \pm 0.92\}$. Stretching this over 300 columns yields approximately the choice described.



(a)

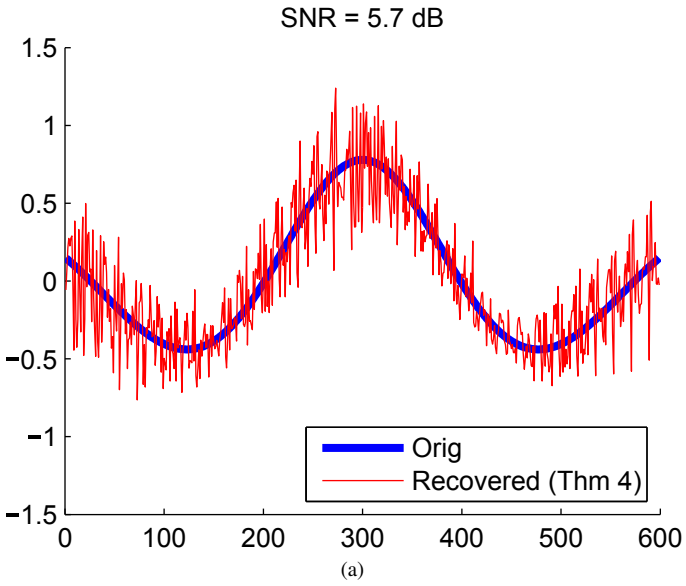


(b)

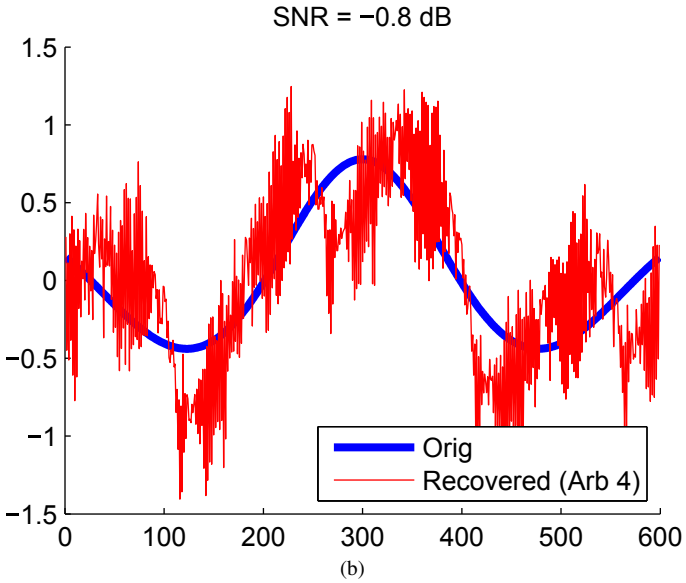
Figure 7: Optimality gap between the approximate minimum diameter solution and the actual solution. (a) four-column set; (b) five-column set.

The recovered Toeplitz vectors for the four-column scheme are shown in Fig. 8; similar levels of noise are also observed in the Hankel vectors. The columns suggested by Theorem 2 yields a reduction of SNR from 20 dB to 5.7 dB, for a noise factor of around 14 dB. The spectrum of the noise is gathered towards high frequencies, such that if the underlying function is known to be smooth, then the SNR can be further improved either via a regularizing assumption, or by post-processing. By comparison, the arbitrary choice has a noise factor around 20.8 dB, and the noise appears to be broad-spectrum. It is difficult to discern the signal from noise at this point, even if assumptions are made about the smoothness of the underlying function.

Next, we perform a five-column recovery. The four columns selected before are kept as-is, and an exhaustive search is performed for the fifth column that would minimize

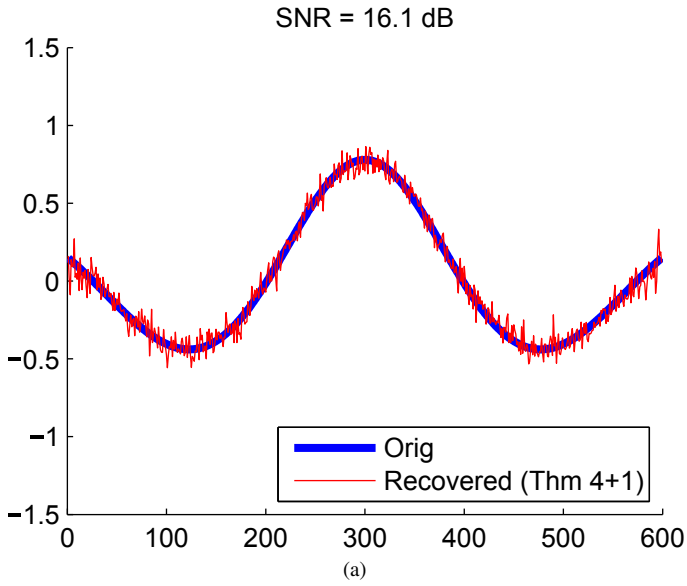


(a)

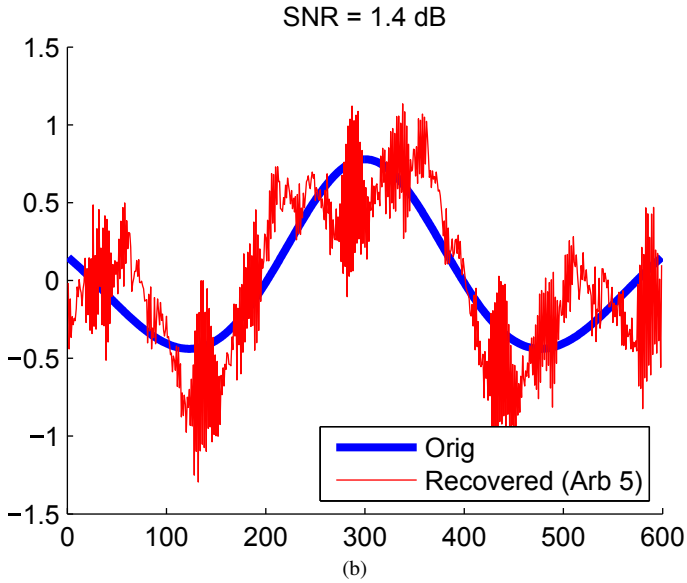


(b)

Figure 8: The amplification of measurement noise in the Toeplitz vector during least-square recovery for $n = 300$ problem. (a) Four columns sampled via the results of this paper; (b) Four column recovered via an arbitrary column choice of $\{1, n/4, 3n/4, n\}$.



(a)



(b)

Figure 9: The amplification of measurement noise in the Toeplitz vector during least-square recovery for $n = 300$ problem. (a) Four columns sampled via the results of this paper, plus one by exhaustive search; (b) Five column recovered via an arbitrary column choice of $\{1, n/4, n/2, 3n/4, n\}$.

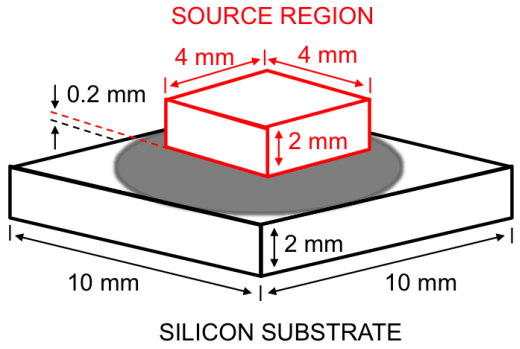
the overall condition number. This results in the choice $\{1, 10, 160, 245, 300\}$. The "arbitrary" choice is also augmented by the most obvious choice: the column in the middle, to yield $\{1, 75, 150, 225, 300\}$.

Results from the four-plus-one recovery scheme are shown in Fig. 9a. Results validate our finding that the condition number dramatically improves with the addition of one more vector. The noise factor improves to 4.9 dB, which is good enough to use directly, without further regularization or post processing.

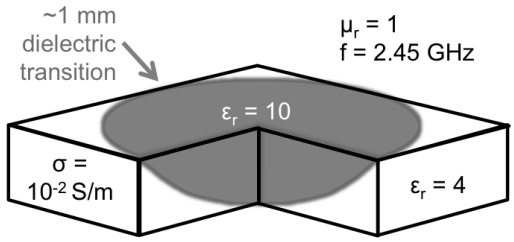
In comparison, the arbitrary five-column scheme shows no significant improvement over the arbitrary four-column scheme. While the noise factor has technically improved, it remains just as difficult as before to discern the signal from the noise.

VII. Recovery of a General Green's Function

Toeplitz-plus-Hankel matrix recovery presents a surprisingly effective data-driven approach for estimating the Green's functions of substrate geometries that satisfy a certain ansatz. Consider the problem of characterizing the electric field response within the box labeled "Source Region" in Fig. 10, due to a Hertzian dipole of any polarity placed anywhere within this same box. Deriving an analytical expression would

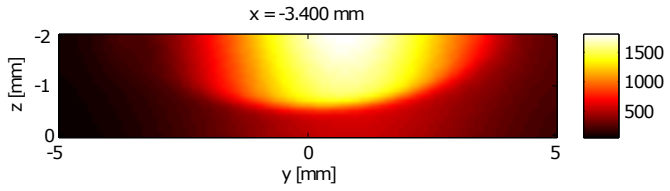


(a)

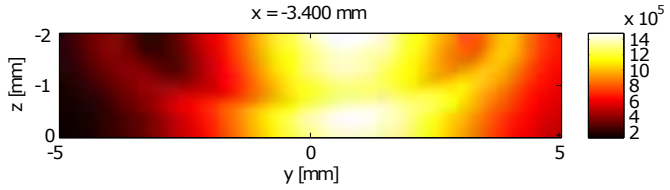


(b)

Figure 10: Ellipsoidal dielectric reflector (with diameter 10 mm and depth 2 mm), embedded within a rectangular silicon substrate: (a) physical dimensions and source layout; (b) electromagnetic properties. The origin $(x, y, z) = (0, 0, 0)$ is set to the center of the top surface of the substrate.



(a)



(b)

Figure 11: Substrate response to an $\hat{x}$-oriented dipole placed at $\mathbf{r}' = (1.67, 0.67, 0.42)$ mm : (a) current density magnitude $|\mathbf{J}(\mathbf{r})|$; (b) electric field magnitude $|\mathbf{E}(\mathbf{r})|$.

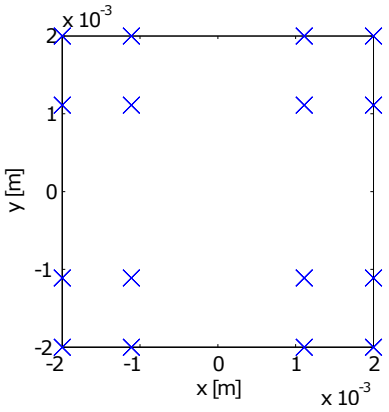


Figure 12: Sampling pattern over source region at $\{-2, -10/9, 10/9, 2\}$ mm.

certainly be difficult. The substrate has dimensions on the same order-of-magnitude as the source region, so edge effects can be significant. Furthermore, the half-ellipsoidal aperture, with diameter of 10 mm and $\epsilon_r = 10$, will experience significant cavity resonance at 2.45 GHz. Finally, the dielectric transitions smoothly into the substrate, but for an analytical expression, this would have to be treated as an abrupt boundary. Figure 11 illustrates a typical response of the substrate; all three effects described above are shown to be significant.

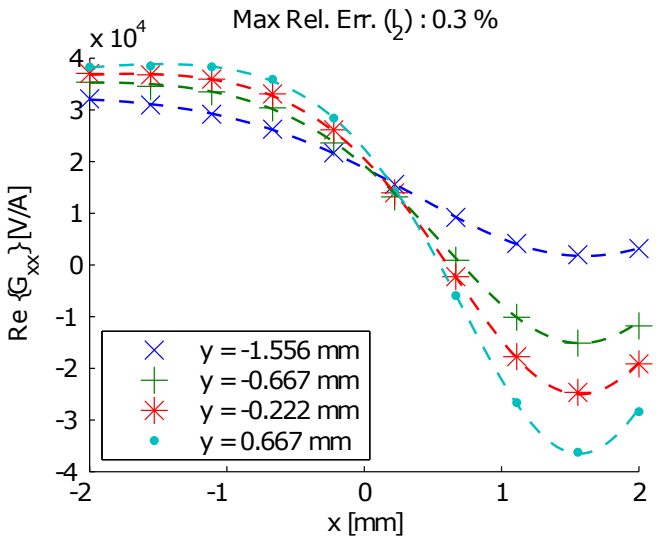
The analytical difficulties motivate a data-driven approach. Discretizing the source region into a uniform grid of $10 \times 10 \times 10$ points, a 1000×1000 matrix may be defined (for each pair of source and field polarity), in which each column contains the electric field at all 1000 grid points due to the influence of a single dipole placed at a corresponding grid point. In turn, the matrix can be populated by performing scattering experiments: exciting the substrate with a dipole and measuring the electric field responses. Data may be collected using real instruments, or via a numerical simulation program. Regardless, naively populating the entire matrix would require 1000 distinct scattering experiments to be performed.

Instead, we make the *ansatz* that the underlying Green's function admits a Toeplitz-plus-Hankel structure in *each of its three dimensions*. More specifically, rearranging the data matrix into a 6-dimensional tensor, we assume the following

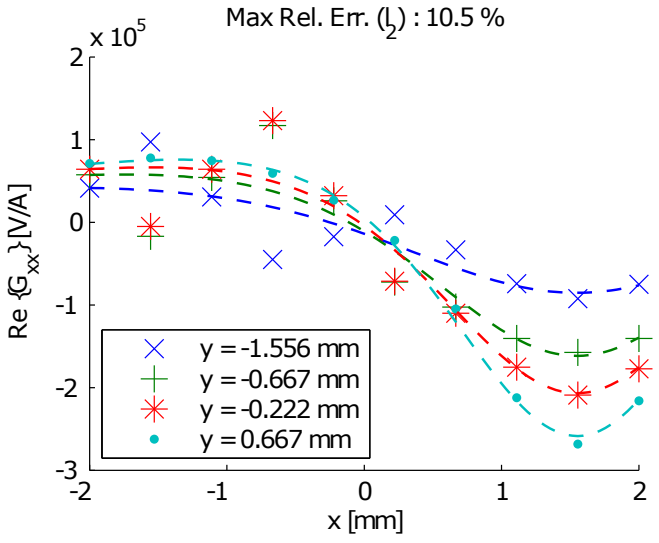
$$G(i, j, k, i', j', k') = \sum_{l=1}^{8} G^{(l)}(i \pm i', j \pm j', k \pm k'), \quad (24)$$

in which (i', j', k') and (i, j, k) are respectively the subscripts to the source and target points on the $10 \times 10 \times 10$ grid.

The ansatz is backed by physical intuition. The half-ellipsoidal cavity is wide and very thin; when it is wider than the source region, its impact is similar to infinite-horizon layered media, which is known to yield the TPH structure along the z direction (see Section II). Moreover, the boundaries of the substrate are rectangular; by separation of variables, one can show that eigenfunctions of a rectangular homogenous medium are TPH along the x and y directions. Taken together, these two observations strongly suggest (and experimental data confirms) that the sampled Green's functions will be well-



(a)



(b)

Figure 13: Electric field response in $\hat{x}$ at $z = 1.978$ mm due to an $\hat{x}$-oriented dipole placed at $\mathbf{r}' = (1.67, 0.67, 0.42)$ mm, with half-ellipsoidal cavity diameter set to: (a) 10 mm; (b) 1 mm. Exact (lines) vs TPH predictions (markers).

approximated by a TPH matrix, but more rigorous analysis is still needed. And more importantly, it remains an open problem to characterize the class of substrates whose Green's function can be accurately approximated by TPH.

Assuming the structure in (24), we proceed to recover the Green's function using the methods in this paper. First, to collect the best-conditioned samples, source locations are chosen at $4^3 = 64$ grid nodes, shown in Fig. 12, corresponding to the choice of $\{1, 3, 8, 10\}$ suggested by Theorem 2 placed along each dimension. For each of the 64 source locations, the scattering problem is solved over all 10^3 points on the grid, using the MARIE package [45], based on the volume integral formulation in [46]. Repeating this process yields a total of 64×10^3 samples.

Table I: TPH recovery of the example Green's function matrix, compared to exhaustively computing every element.

	Exhaustive	TPH Recovery	% Reduction
Data Collection			
Num. scatt. prob.	1,000	64	93.6 %
Num. samples	10^6	6.4×10^3	93.6 %
Comp. time[1]	57,500 s	3,830 s	93.3 %
Storage			
Degrees of freedom[2]	10^6	54,872	94.5%

1 Around 10s to solve each scattering problem, and around 40-50s to evaluate the electric fields using MARIE. Includes the computation of the TPH recovery problem.
2 $54872 = (10 \times 4 - 2)^3$.

To perform the three-dimensional TPH recovery, the least-squares problem (5) in Section III may be modified to enforce the structure in (24), and solved using any generic convex optimization parser, such as YALMIP [47]. Alternatively, it is also possible to perform the recovery one dimension at a time. Doing so immediately draws an equivalence between the conditioning of the three-dimensional problem here and the one-dimensional problem studied in Section III. The details of this latter procedure are left to a separate paper.

The results of the matrix recovery show a remarkable match with elements of the actual matrix: for each of the nine polarization pairs, TPH matrix recovery was able to produce predictions with 0.1-0.3 % relative error in Frobenius norm, despite sampling 6.4% of all matrix entries. Some select predictions of TPH recovery are shown in Fig. 13a. As shown in Table I, the ability to describe the entire matrix using 6.4% of its elements results in a large reduction in computation time and storage requirement.

We emphasize that the effectiveness of the procedure rests solely on the TPH ansatz, described in (24). This is an entirely different setting to most previous work on inference and sparse sensing, in which the underlying matrix is assumed to be low-rank [48]. Indeed, this particular problem is not low-rank. Once the exact 1000×1000 matrices had been computed, their numerical ranks were verified to range between 200 and 600 (for a relative tolerance of 10^{-6}).

But as expected, the method falls apart when the TPH ansatz is rendered invalid. For example, if the half-ellipsoidal cavity is reduced in size to 1 mm diameter and 0.2 mm depth, then the cavity is no longer wide-and-flat, and the scattered field in the source region should not be TPH. Indeed, repeating the above procedure for this modified example results in prediction errors as large as 10% - 80% in Frobenius norm. Some example predictions are shown in Fig. 13b.

VIII. Conclusions

In this paper we showed that a Toeplitz-plus-Hankel (TPH) matrix can be recovered by sampling four or five of its columns, and we showed how to use column sampling in a flexible data-driven approach for characterizing substrate Green's functions, in a form perfectly suited for FFT-based fast interconnect analysis.

The selection of the sampled columns is carefully examined, using a combination of linear algebra and graph techniques,

because a good selection leads to a well-conditioned recovery problem. And if the recovery problem is well-conditioned, then it is possible to accurately recover TPH matrices from noisy measurements or low-accuracy simulation data, as demonstrated in our first example for the recovery of a noise-polluted TPH matrix. Our results also suggests that sampling an additional column (beyond the minimium four) can significantly improve the conditioning of the recovery.

Sampling the columns following the prescription in our Theorem 2 yields a recovery problem that is well-conditioned enough to extract good TPH approximations from mildly non-TPH matrices. In our second example, this was demonstrated by extracting a TPH approximation of the Green's function for a far-from-planar dielectric reflector. The result suggests a practical direction for future work: to characterize the class of substrates that admit TPH approximations of their associated Green's function.

IX. Acknowledgments

The authors are grateful for the financial support of this work provided by the MIT Energy Initiative and the Skolkovo-MIT Initiative in Computational Mathematics.

Appendix

A. Null-space of the TPH decomposition

Proposition 8 (Null-space basis)**.** *The right null-space of E, defined in (7), has dimension 2, and is spanned by spanned by the bases*

$$\begin{bmatrix} \mathbf{1}_{odd} & \mathbf{1}_{even} \\ -\mathbf{1}_{even} & -\mathbf{1}_{odd} \end{bmatrix} \qquad \textit{for } n \textit{ odd,}$$

$$\begin{bmatrix} \mathbf{1}_{odd} & \mathbf{1}_{even} \\ -\mathbf{1}_{odd} & -\mathbf{1}_{even} \end{bmatrix} \qquad \textit{for } n \textit{ even,}$$

where $\mathbf{1}_{odd}$ and $\mathbf{1}_{even}$ define the indicator vectors for odd- and even-indexed coefficients, as in

$$\mathbf{1}_{odd}^T = \overbrace{\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & \cdots \end{bmatrix}}^{2n-1},$$
$$\mathbf{1}_{even}^T = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & \cdots \end{bmatrix}.$$

Proof: See [33], [49] for a proof using linear algebra. ■

References

[1] L. Lemaitre, G. Coram, C. McAndrew, and K. Kundert, "Extensions to Verilog-A to support compact device modeling," in *Proceedings of the 2003 International Workshop on Behavioral Modeling and Simulation, 2003.*, pp. 134–138, IEEE, 2003.

[2] "IBIS Open Forum," 2014.

[3] C. J. Alpert and A. B. Kahng, "Recent directions in netlist partitioning: a survey," *Integration, the VLSI Journal*, vol. 19, pp. 1–81, Aug. 1995.

[4] C.-K. Cheng, J. Lillis, S. Lin, N. H. Chang, *et al.*, *Interconnect analysis and synthesis*. Wiley New York, 2000.

[5] J. Duato, S. Yalamanchili, and L. M. Ni, *Interconnection networks: An engineering approach*. Morgan Kaufmann, 2003.

[6] S. Pasricha and N. Dutt, *On-chip communication architectures: system on chip interconnect*. Morgan Kaufmann, 2010.

[7] W. C. Chew, *Waves and fields in inhomogeneous media*, vol. 522. IEEE press New York, 1995.

[8] K. Michalski and J. Mosig, "Multilayered media Green's functions in integral equation formulations," *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 508–519, Mar. 1997.

[9] Y. Chow, J. Yang, D. Fang, and G. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, pp. 588–592, Mar. 1991.

[10] G. Dural and M. I. Aksun, "Closed-form Green's functions for general sources and stratified media," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 7, pp. 1545–1552, 1995.

[11] A. Niknejad, R. Gharpurey, and R. Meyer, "Numerically stable Green function for modeling and analysis of substrate coupling in integrated circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 4, pp. 305–315, 1998.

[12] J. Zhao, W.-M. Dai, S. Kadur, and D. E. Long, "Efficient three-dimensional extraction based on static and full-wave layered Green's functions," in *Proceedings Design Automation Conference, 1998.*, pp. 224–229, IEEE, 1998.

[13] F. Ling and J. Jin, "Discrete complex image method for Green's functions of general multilayer media," *IEEE Microwave and Guided Wave Letters*, vol. 10, no. 10, pp. 400–402, 2000.

[14] F. Ling, J. Liu, and J. Jin, "Efficient electromagnetic modeling of three-dimensional multilayer microstrip antennas and circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 6, pp. 1628–1635, 2002.

[15] T. Sarkar, E. Arvas, and S. Rao, "Application of FFT and the conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 5, pp. 635–640, 1986.

[16] Y. Zhuang, K.-L. Wu, C. Wu, and J. Litva, "A combined full-wave CG-FFT method for rigorous analysis of large microstrip antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 1, pp. 102–109, 1996.

[17] J. Phillips and J. White, "A precorrected-FFT method for electrostatic analysis of complicated 3-D structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, no. 10, pp. 1059–1072, 1997.

[18] Y. Massoud and J. White, "Simulation and modeling of the effect of substrate conductivity on coupling inductance and circuit crosstalk," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 10, no. 3, pp. 286–291, 2002.

[19] B. J. Rautio, V. Okhmatovski, A. C. Cangellaris, J. C. Rautio, J. K. Lee, *et al.*, "The unified-fft algorithm for fast electromagnetic analysis of planar integrated circuits printed on layered media inside a rectangular enclosure," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 5, pp. 1112–1121, 2014.

[20] J. R. James and P. S. Hall, *Handbook of microstrip antennas*, vol. 28. IET, 1989.

[21] R. Garg, *Microstrip antenna design handbook*. Artech House, 2001.

[22] C.-I. G. Hsu, R. Harrington, K. Michalski, D. Zheng, *et al.*, "Analysis of multiconductor transmission lines of arbitrary cross section in multi-layered uniaxial media," *IEEE Transactions on Microwave Theory and Techniques*, vol. 41, no. 1, pp. 70–78, 1993.

[23] A. E. Ruehli, "Inductance Calculations in a Complex Integrated Circuit Environment," *IBM Journal of Research and Development*, vol. 16, pp. 470–481, Sept. 1972.

[24] A. Ruehli, "Equivalent Circuit Models for Three-Dimensional Multiconductor Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 22, pp. 216–221, Mar. 1974.

[25] M. Kamon, M. Ttsuk, and J. White, "FASTHENRY: a multipole-accelerated 3-D inductance extraction program," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 9, pp. 1750–1758, 1994.

[26] R. F. Harrington and J. L. Harrington, *Field computation by moment methods*. Oxford University Press, 1996.

[27] R. F. Harrington, *Time-Harmonic Electromagnetic Fields (IEEE Press Series on Electromagnetic Wave Theory)*. Wiley-IEEE Press, 2001.

[28] Y. Saad, *Iterative methods for sparse linear systems*. Siam, 2003.

[29] A. E. Yilmaz, J.-M. Jin, and E. Michielssen, "Time domain adaptive integral method for surface integral equations," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 10, pp. 2692–2708, 2004.

[30] J. Phillips, *Rapid solution of potential integral equations in complicated 3-dimensional geometries*. PhD thesis, Massachusetts Institute of Technology, 1997.

[31] A. Niknejad and R. Meyer, "Analysis of eddy-current losses over conductive substrates with applications to monolithic inductors and transformers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 166–176, 2001.

[32] Y. Massoud, S. Majors, J. Kawa, T. Bustami, D. MacMillen, and J. White, "Managing on-chip inductive effects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 10, pp. 789–798, Dec. 2002.

[33] W.-H. Fang and A. Yagle, "Two methods for Toeplitz-plus-Hankel approximation to a data covariance matrix," *IEEE Transactions on Signal Processing*, vol. 40, pp. 1490–1498, June 1992.

[34] N. J. Higham, *Accuracy and stability of numerical algorithms*. Siam, 2002.

[35] S. Boyd, P. Diaconis, and L. Xiao, "Fastest Mixing Markov Chain on a Graph," *SIAM Review*, vol. 46, pp. 667–689, Jan. 2004.

[36] J. Sun, S. Boyd, L. Xiao, and P. Diaconis, "The Fastest Mixing Markov Process on a Graph and a Connection to a Maximum Variance Unfolding Problem," *SIAM Review*, vol. 48, pp. 681–699, Jan. 2006.

[37] J. Batson, D. A. Spielman, and N. Srivastava, "Twice-Ramanujan Sparsifiers," *SIAM Journal on Computing*, vol. 41, pp. 1704–1721, Jan. 2012.

[38] A. A. Benczúr and D. R. Karger, "Approximating st minimum cuts in õ (n 2) time," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 47–55, ACM, 1996.

[39] D. R. Karger and C. Stein, "A new approach to the minimum cut problem," *Journal of the ACM (JACM)*, vol. 43, no. 4, pp. 601–640, 1996.

[40] D. A. Spielman and N. Srivastava, "Graph Sparsification by Effective Resistances," *SIAM Journal on Computing*, vol. 40, pp. 1913–1926, Jan. 2011.

[41] R. Grone, R. Merris, and V. S. Sunder, "The Laplacian spectrum of a graph," *SIAM Journal on Matrix Analysis and Applications*, vol. 11, no. 2, pp. 218–238, 1990.

[42] F. R. K. Chung, *Spectral Graph Theory*. American Mathematical Soc., 1996.

[43] B. Mohar, "Eigenvalues, diameter, and mean distance in graphs," *Graphs and combinatorics*, vol. 7, no. 1, pp. 53–64, 1991.

[44] R. Zhang, J. White, and J. Kassakian, "Fast simulation of complicated 3D structures above lossy magnetic media," *IEEE Transactions on Magnetics*, vol. 50, no. 10, 2014.

[45] J. Villena, A. Polimeridis, L. Wald, E. Adalsteinsson, J. White, and L. Daniel, "MARIE – a MATLAB-based open source software for the fast electromagnetic analysis of MRI systems," in *Proceedings of the 23nd Scientific Meeting of ISMRM, Toronto, Canada*, 2015.

[46] A. Polimeridis, J. Villena, L. Daniel, and J. White, "Stable FFT-JVIE solvers for fast analysis of highly inhomogeneous dielectric objects," *Journal of Computational Physics*, vol. 269, pp. 280–296, 2014.

[47] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *2004 IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289, IEEE, 2004.

[48] E. J. Candes and T. Tao, "The Power of Convex Relaxation: Near-Optimal Matrix Completion," *IEEE Transactions on Information Theory*, vol. 56, pp. 2053–2080, May 2010.

[49] G. Strang and S. MacNamara, "Functions of Difference Matrices Are Toeplitz Plus Hankel," *SIAM Review*, vol. 56, pp. 525–546, Aug. 2014.