Permutations Unlabeled Beyond Sampling Unknown

Ivan Dokmanić ⓘ, *Member, IEEE*

Abstract—A recent unlabeled sampling result by Unnikrishnan, Haghighatshoar, and Vetterli states that with probability one over Gaussian random matrices A with iid entries, any x can be uniquely recovered from an unknown permutation of $y = Ax$ as soon as A has at least twice as many rows as columns. We show that this condition on A implies something much stronger: that an unknown vector x can be recovered from measurements $y = TAx$, when the unknown T belongs to an arbitrary set of invertible, diagonalizable linear transformations $\mathcal{T}$. The set $\mathcal{T}$ can be finite or countably infinite. When it is the set of $m \times m$ permutation matrices, we have the classical unlabeled sampling problem. We show that for almost all A with at least twice as many rows as columns, all x can be recovered either uniquely, or up to a scale depending on $\mathcal{T}$, and that the condition on the size of A is necessary. Our proof is based on vector space geometry. Specializing to permutations, we obtain a simplified proof of the uniqueness result of Unnikrishnan, Haghighatshoar, and Vetterli. In this letter, we are only concerned with uniqueness; stability and algorithms are left for future work.

Index Terms—Sampling, shuffled regression, unlabeled sampling, unknown permutation, unknown transformation.

I. INTRODUCTION

WE "STEAL" a motivating example from [1]: Imagine that you are recording a sound field with a large number of microphones connected to a recording interface. Alas, you forgot to label the cables so you end up with a pile of recordings without knowing which one corresponds to which spatial position. Is there a way to reconstruct the wavefield even without proper labels?

We can model this situation by the following unlabeled sampling problem:

$$y = \Pi A x, \tag{ULS}$$

where $A \in \mathbb{C}^{m \times n}$, $x \in \mathbb{C}^n$, and instead of measuring the usual Ax we get to measure its unknown permutation. If the permutation Π is known (the cables are neatly labeled), (ULS) is simply a linear system.

Many signal processing problems are modeled by (ULS) and related constructions. If the columns of A are samples of harmonic sinusoids, the problem is that of sampling at unknown locations [2], [3]. In simultaneous localization and mapping (SLAM), a robot is sensing an unknown environment, without

"knowing" its own spatial location [4]. If there are a finite number of possible locations and A contains a model of the world as seen through the mobile sensors, then (ULS) models a SLAM scenario. A system similar to (ULS) appears in room geometry reconstruction and microphone positioning by echoes [5]–[7]. A nonlinear instance of sensing with unknown permutations is the unlabeled distance geometry problem where the task is to recover a point set from point-to-point distances, without knowing which distance corresponds to which pair of points [8].

In the context of the Internet of things and fifth-generation communication systems, (ULS) models header-free communication with very short packets [9], [10]. Headers that identify individual nodes are too large compared with the actual payloads, but in many sensing tasks the correct labeling can be inferred from the payload itself. When the nodes are sensors sensing a spatial field which has a subspace representation (for example, an advection-diffusion field [11], [12] or a wavefield [13]), then the problem can be modeled as measuring Ax up to a permutation. Recent work shows that the recovery can be addressed using symmetric polynomials [10], [14].

Further connections exist with tomography with unknown projection angles, an especially relevant topic with the emergence of cryogenic electron microscopy (Cryo-EM) in which we get linear tomographic measurements with unknown angles [15], [16]. Since the Radon transform has a restricted range, the problem can be modeled as (ULS).

Problems of type (ULS) can be split into underdetermined, $m < n$, and (over) determined $(m \geqslant n)$. In the underdetermined case, we need a model for x. When x is sparse, Emiya *et al.* [1] adapt the branch-and-bound technique to efficiently search through all permutations.

We let x be any complex vector and thus study the overdetermined case. In this setting, Unnikrishnan, Haghighatshoar, and Vetterli [17], [18] proved that if A is iid Gaussian, it is possible to recover *every* x uniquely with probability 1 over realizations of A if and only if $m \geqslant 2n$. Their proof involves sophisticated arguments from coding theory. Haghighatshoar and Caire also discuss recovery from an unknown but ordered subset of measurements [19]. Pananjady, Wainwright, and Courtade discuss statistical and computational aspects of unlabeled linear regression [20].

In this letter, we prove the following significant generalization of the above results: Imagine that y was obtained as TAx for some unknown invertible transformation $T \in \mathcal{T}$, where $\mathcal{T}$ is some set of invertible diagonalizable transformations, and A is a known matrix. The set $\mathcal{T}$ can be finite or countably infinite. It can model unknown transfer functions, propagation parameters, and sensing parameters beyond permutations. We show that when $m \geqslant 2n$, for *almost all* matrices A all x can be recovered uniquely or up to a scale. Taking $\mathcal{T}$ to be the set of $m \times m$
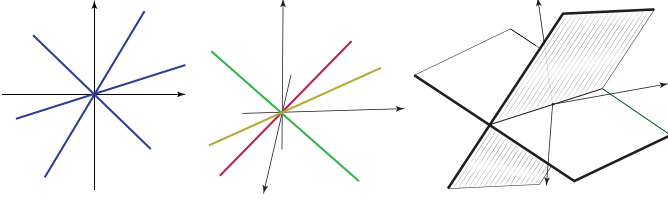
Fig. 1.   Illustration of subspace intersections. Two 1D subspaces (lines through the origin) in 2D generically intersect only at the origin $(1 + 1 \leqslant 2)$; the same holds for two 1D subspaces in 3D $(1 + 1 \leqslant 3)$. Two 2D subspaces (planes through the origin) in 3D generically intersect along a line—a 1D subspace $(2 + 2 > 3)$.

permutation matrices (of cardinality $m!$), we recover the uniqueness result of Unnikrishnan, Haghighatshoar and Vetterli.

Our proof is simple and based on geometric arguments. The gist of it is that random n-dimensional subspaces of $\mathbb{C}^{2n}$ only intersect at the zero vector. On the other hand, if the ambient dimension is smaller than $2n$, then any two n-dimensional subspaces intersect non-trivially; we illustrate this in Fig. 1. The subtleties of the argument depend on the eigenvalues of transformations T; that it can be applied to permutation then follows by studying the eigenvalues of permutation matrices.

In this letter we are only concerned with the question of unique recovery. The important questions of recovery algorithms and their stability are left to future work.

Finally, after our work appeared online, Tsakiris [21] posted a preprint that addresses the phenomenon we describe. Though [21] only deals with finite transformation classes, it extends our results to more general non-invertible transformations via algebraic-geometric arguments.

II. MAIN RESULT

Our main lemma concerns the case of only two transformation matrices $\mathcal{T} = \{I, T\}$, where I is the $m \times m$ identity matrix. We show that x can be recovered from y when y is either Ax or TAx, but we do not know which. The proof relies on studying the size of the intersection of the range of A and the range of TA.

We assume that $T \neq I$ and that $T \in \mathbb{C}^{m \times m}$ has an eigenvalue decomposition $T = \Phi \Lambda \Phi^{-1}$. We will denote by $\bar{\lambda}(T)$ an eigenvalue of T with the largest multiplicity, and denote its multiplicity by $p(T)$. If there are multiple such eigenvalues, and one of them is $\lambda = 1$, we set $\bar{\lambda} = 1$; otherwise we break the tie arbitrarily. Note that the eigenvalues can be complex. Without loss of generality, we order the eigenvalues so that $\lambda_1 = \lambda_2 = \cdots = \lambda_p = \bar{\lambda}$. Since we assume that T is diagonalizable, algebraic and geometric multiplicities coincide.

We denote the Lebesgue measure on $\mathbb{C}^{m \times n}$ by μ and say that a property holds for almost all A when it holds μ-almost everywhere in $\mathbb{C}^{m \times n}$, that is, when it does not hold on $\mathcal{B} \subseteq \mathbb{C}^{m \times n}$ with $\mu(\mathcal{B}) = 0$. Since all the subsequent "almost all" claims in $\mathbb{C}^{m \times n}$ also hold almost everywhere in $\mathbb{R}^{m \times n}$ with respect to the Lebesgue measure on $\mathbb{R}^{m \times n}$, we extend the meaning of almost all to include both cases. We can then state the following:

Lemma 1: Let $T \in \mathbb{C}^{m \times m}$ be an invertible, diagonalizable matrix with an eigenvalue decomposition $T = \Phi \Lambda \Phi^{-1}$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and $A \in \mathbb{C}^{m \times n}$, $m \geqslant 2n$. Then for almost all matrices A, for all y such that $y = Ax = TAz$, we have
- If $p(T) \leqslant m - n$, then $x = z$;
- If $p(T) > m - n$ and $\bar{\lambda}(T) = 1$, then $x = z$;

- If $p(T) > m - n$ and $\bar{\lambda}(T) \neq 1$, then $x = \bar{\lambda}(T)z$.
  Proof: We want to identify conditions on A such that if

$$y = Ax \qquad (1)$$

$$\text{and} \quad y = TAz = \Phi \Lambda \Phi^{-1} Az, \qquad (2)$$

we must have $x = z$. If (1) and (2) hold simultaneously, then

$$TAz \in \mathcal{R}(A).$$

Thus, there exists a vector $y \in \mathcal{R}(A)$ such that also $Ty \in \mathcal{R}(A)$. Our proof hinges on the fact that this situation is very special.

Write $Az = \Phi d$ for some d (this is possible because the columns of Φ form a basis for $\mathbb{C}^m$). Then from (2) we have $y = \Phi \Lambda d$. Since $y \in \mathcal{R}(A)$, it must be that

$$\Lambda d = \Phi^{-1} y \in \mathcal{R}(\widetilde{A}),$$

where we defined the shortcut $\widetilde{A} = \Phi^{-1} A$. From the definition of d we have $d = \Phi^{-1} Az$, so d should also be in the range of $\widetilde{A}$. Another way to write this is as:

$$\left. \begin{array}{l} d \in \mathcal{R}(\widetilde{A}) \\ \Lambda d \in \mathcal{R}(\widetilde{A}) \end{array} \right\} \iff \left\{ \begin{array}{l} Qd = 0 \\ Q \Lambda d = 0 \end{array} \right., \qquad (3)$$

where Q^* is the Hermitian transpose of Q, the columns of $Q^* \in \mathbb{C}^{m \times (m-n)}$ form a basis for the orthocomplement of the range of $\widetilde{A}$ and we used the fact that $\mathcal{R}(\widetilde{A}) = \mathcal{N}(\widetilde{A}^*)^{\perp}$.

Note that (3) is a homogeneous system of $2(m - n)$ equations in m unknowns, so as soon as $2(m - n) < m$, or $m < 2n$, there are inevitably infinitely many solutions regardless of Λ. (Here we assume Q and Λ to be known.) This case is further developed in Proposition 1.

Let rref denote the reduced row echelon form. For A with full column rank (that is, for almost all A), $\widetilde{A}$ also has full column rank, which implies

$$\mathsf{rref}(\widetilde{A}^*) = \left[\ I_{n \times n} \mid [S^*]_{n \times (m-n)}\ \right],$$

with $S \in \mathbb{C}^{(m-n) \times n}$ (for convenience we indicate the block sizes in subscripts). From here we can read out a basis for $\mathcal{N}(\widetilde{A}^*)$ as

$$Q^* = \left[\begin{array}{c} [S^*]_{n \times (m-n)} \\ -I_{(m-n) \times (m-n)} \end{array} \right],$$

with S having full column rank. Setting $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n \times n}$, $\Lambda_2 = \mathrm{diag}(\lambda_{n+1}, \ldots, \lambda_m)$, and partitioning d as $d = [d_1^\top,\ d_2^\top]^\top$ we rewrite (3) as

$$S d_1 - d_2 = 0, \qquad (4)$$

$$S \Lambda_1 d_1 - \Lambda_2 d_2 = 0. \qquad (5)$$

From the first equation we have $d_2 = S d_1$, so that $S \Lambda_1 d_1 - \Lambda_2 S d_1 = 0$, or

$$(S \Lambda_1 - \Lambda_2 S) d_1 = 0. \qquad (6)$$

Let us focus on the top n rows of this equation, with notation illustrated in Fig. 2. In particular, let $\widetilde{S} \in \mathbb{C}^{n \times n}$ denote the top n rows of S, and $\widetilde{\Lambda}_2 \in \mathbb{C}^{n \times n}$ the upper-left $n \times n$ block of Λ_2. From (6) we have that

$$(\widetilde{S} \Lambda_1 - \widetilde{\Lambda}_2 \widetilde{S}) d_1 = 0. \qquad (7)$$
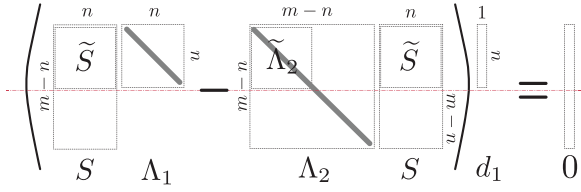
Fig. 2. Dimensions of matrices in (6) and (7).

In order for a nonzero solution d_1 to exist, the system matrix must be singular:

$$q(\widetilde{S}) := \det(\widetilde{S}\Lambda_1 - \widetilde{\Lambda}_2\widetilde{S}) = 0. \qquad (8)$$

The determinant in (8) is a homogeneous multivariate polynomial in the entries of $\widetilde{S}$. This polynomial is either identically zero, or it is zero on a subset of $\mathbb{C}^{n \times n}$ of Lebesgue measure zero. (Similarly, the set of real zeros has measure zero in $\mathbb{R}^{n \times n}$.) Consequently, it is either identically zero or nonzero for almost all A.

We now find sufficient conditions for $q(\widetilde{S})$ not to vanish identically. In particular, we find conditions so that $q(I_n) \neq 0$. Setting $\widetilde{S} = I_n$, we get $\det(\Lambda_1 - \widetilde{\Lambda}_2) = 0$. Recall that $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\Lambda_2 = \mathrm{diag}(\lambda_{n+1}, \ldots, \lambda_m)$, and that the eigenvalue with the largest multiplicity is listed first. Writing $p := p(T)$ for brevity, we have:

1. If $p \leqslant n$, clearly the determinant in (8) cannot be zero. Since the determinant is not identically zero, it is zero only for a set of A of measure zero, hence for almost all A the only solution to (7) is $d_1 = 0$ which implies $y = 0$. Since almost all A have full column rank, $x = z = 0$.

2. If $n < p \leqslant m - n$ we can write (6) as

$$(S\Lambda_1 - \Lambda_2 S)d_1 = (\bar{\lambda}I - \Lambda_2)Sd_1 = 0. \qquad (9)$$

The top $p - n$ rows of $(\bar{\lambda}I - \Lambda_2)$ are zero which implies the same for $(\bar{\lambda}I - \Lambda_2)S$, leaving us with $(m - n) - (p - n) = m - p \geqslant n$ independent nonzero equations "at the bottom". Using an analogous argument as above for the bottom n equations, we again get that $d_1 = 0$ and $x = z$ for almost all A.

3. If $m - n < p$, then $(\bar{\lambda}I - \Lambda_2)S$ has a nontrivial nullspace (we have fewer than n nonzero equations for d_1), and (9) has a nonzero solution.

In the last case, any solution d_1 must be in the nullspace of the bottom $m - p$ rows of S (this can be seen from (9) which also holds for $p > m - n$), so it must be that $d_2 = Sd_1$ is supported only on the top $p - n$ entries (since $(m - p) + (p - n) = m - n$). As a consequence, the vector $d = [d_1^\top, \ d_2^\top]^\top$ is supported on the top $n + (p - n) = p$ entries.

Since $y = \Phi\Lambda d$, this implies that y lives in the eigenspace spanned by the first p eigenvectors corresponding to $\bar{\lambda}$. In summary, if $p > m - n$, then *for all A, $\mathcal{R}(A)$ and $\mathcal{R}(TA)$ intersect on the largest eigenspace corresponding to $\bar{\lambda}(T)$; for almost all A they do not intersect anywhere else.*

Thus for almost all A, all $s \in \mathcal{R}(A) \cap \mathcal{R}(TA)$ are such that $Ts = \bar{\lambda}s$. If $\bar{\lambda} = 1$, we can recover the corresponding x uniquely since the equations $s = Ax$ and $s = TAz$ both have at most one solution, and the solution to the latter is $z = A^\dagger T^{-1}s$, but $T^{-1}s = s$ and $A^\dagger s = x$. Otherwise, if $\bar{\lambda} \neq 1$, we can recover up to a scaling since from $s = TAz$ we have $z = A^\dagger T^{-1}s = \bar{\lambda}^{-1}A^\dagger s = \bar{\lambda}^{-1}x$. ∎

We now show how Lemma 1 implies a similar result for any number of unknown transformations in Theorem 1.

Theorem 1: Let $\mathcal{T} = \{T_k \in \mathbb{C}^{m \times m}\}_{k \in \mathcal{K}}$ be a finite or countably infinite set of invertible diagonalizable transforms, and $A \in \mathbb{C}^{m \times n}$, where $m \geqslant 2n$. Let further $y = TAx$, where $A \in \mathbb{C}^{m \times n}$, $x \in \mathbb{C}^n$, $T \in \mathcal{T}$, and neither x nor T are known. Then

1. If for all $T_1, T_2 \in \mathcal{T}$ we have that $p(T_1^{-1}T_2) \leqslant m - n$ or $\bar{\lambda}(T_1^{-1}T_2) = 1$ then x is uniquely determined by y.

2. If there exist $T_1, T_2 \in \mathcal{T}$ such that $\bar{\lambda}(T_1^{-1}T_2) \neq 1$ and $p(T_1^{-1}T_2) > m - n$, then x is determined up to a scale $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is at most a countable set.

Proof: Denote by $\mathcal{T}_y$ the set of $T \in \mathcal{T}$ for which $y \in \mathcal{R}(TA)$, and by $|\mathcal{T}_y|$ its cardinality. We only need to consider $T \in \mathcal{T}_y$. If $|\mathcal{T}_y| = 1$ we are done.

Suppose $|\mathcal{T}_y| > 1$ and let $T_1, T_2 \in \mathcal{T}_y$. That is, $y = T_1Ax = T_2Az$ for some x and z. Putting $\widetilde{y} = T_1^{-1}y$ we can write

$$\left.\begin{array}{r} y = T_1Ax \\ y = T_2Az \end{array}\right\} \iff \left\{\begin{array}{l} \widetilde{y} = Ax \\ \widetilde{y} = T_1^{-1}T_2Az \end{array}\right. . \qquad (10)$$

1. By Lemma 1, we have that for almost all matrices A, this implies $x = z$, for all y. In other words, the set $\mathcal{B}_{T_1,T_2}$ of "bad" matrices A where it does not hold is of Lebesgue measure zero, $\mu(\mathcal{B}_{T_1,T_2}) = 0$. The set of matrices for which it might fail for any choice of T_1 and T_2 is

$$\mathcal{B} = \bigcup_{\substack{T_1,T_2 \in \mathcal{T}_y \\ T_1 \neq T_2}} \mathcal{B}_{T_1,T_2},$$

but by the subadditivity of measure (and noting that the set of all pairs in $\mathcal{T}$ is countable),

$$\mu(\mathcal{B}) \leqslant \sum_{\substack{T_1,T_2 \in \mathcal{T}_y \\ T_1 \neq T_2}} \mu(\mathcal{B}_{T_1,T_2}) = 0.$$

2. Again using Lemma 1 and reasoning as in **1.**, for a fixed T_1, T_2 and almost all A, we can uniquely recover any x up to a scaling by $\bar{\lambda}(T_1^{-1}T_2)$. Thus the claim of the theorem holds with $\mathcal{A} = \{\bar{\lambda}(T_1^{-1}T_2) : T_1, T_2 \in \mathcal{T}_y, T_1 \neq T_2\}$. ∎

Theorem 1 establishes that under suitable conditions on A, for a rather general class of possible transformations $\mathcal{T}$, x can be recovered from $y = TAx$, where $T \in \mathcal{T}$ is unknown. We now specialize these results to classical unlabeled sensing. We begin with a fact about eigenvalues of permutations.

Lemma 2: For any permutation matrix Π, $\bar{\lambda}(\Pi) = 1$.

Proof: Every $m \times m$ permutation Π can be written as a product of r disjoint cycles $\Pi = C_1 C_2 \cdots C_r$. Since the cycles are disjoint, the sum of the lengths is exactly m.

Denote by $\mathcal{W}_i$ the set of ℓ_i-th roots of unity, where ℓ_i is the length of the ith cycle C_i, $\mathcal{W}_i = \{e^{\mathrm{i}2\pi p/\ell_i} : p \in \{0, 1, \ldots, \ell_i - 1\}\}$. Then the set of all eigenvalues of Π is [22]

$$\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2 \cup \cdots \cup \mathcal{W}_r,$$

and the geometric multiplicity of each eigenvalue is the number of times it appears in sets $\mathcal{W}_i$, $i \in \{1, 2, \ldots, r\}$. Note that every $\mathcal{W}_i$ contains a 1, since $1 = e^{\mathrm{i}2\pi \cdot 0}$. Therefore, the eigenvalue of Π with the largest multiplicity is 1 (there could be other eigenvalues with the same maximal multiplicity). ∎

We will also use a partial converse to Lemma 1 to show that $m \geqslant 2n$ is necessary for permutations.

Proposition 1 (Partial converse): Let $n \leqslant m < 2n$, and assume that there exist $T_1, T_2 \in \mathcal{T}$ such that $T_1^{-1} T_2$ has no eigenspace of dimension larger than $m - n$. Then for almost all matrices A, there exist $x, z \in \mathbb{C}^n$ such that $x \neq z$ and $T_1 A x = T_2 A z$.

Proof: For almost all A the ranges of $T_1 A$ and $T_2 A$ have dimension n. Since the sum of the dimensions of the two range spaces exceeds the dimension of the ambient space, $n + n > m$, they must have a non-trivial intersection:

$$\dim \mathcal{R}(A) \cap \mathcal{R}(TA) \geqslant 2n - m \geqslant 1. \quad (11)$$

Let $s \in S := \mathcal{R}(T_1 A) \cap \mathcal{R}(T_2 A)$. Then there exist x and z such that $s = T_1 A x$ and $s = T_2 A z$.

The unknown x can be recovered only if $x = z$, in which case Ax is an eigenvector of $T_1^{-1} T_2$ with an eigenvalue 1 (recall (10)). This can happen only if $\mathcal{R}(A)$ and the corresponding eigenspace of $T_1^{-1} T_2$ have a nontrivial intersection. But if every eigenspace E_λ of $T_1^{-1} T_2$ has dimension at most $m - n$, then $\dim \mathcal{R}(A) + \dim(E_\lambda) \leqslant m$ and the two intersect only for a set of matrices A of measure zero. Since from (11) a nonzero s does exist, it must be that $x \neq z$. ∎

We can now easily prove the following:

Corollary 1 (Unnikrishnan et al., 2015): If $\mathcal{P} \subset \mathbb{R}^{m \times m}$ is the set of all $m!$ permutation matrices of m elements, then any x can be uniquely recovered from measurements $y = \Pi A x$, where both $\Pi \in \mathcal{P}$ and $x \in \mathbb{C}^n$ are unknown, for almost all matrices $A \in \mathbb{C}^{m \times n}$ with $m \geqslant 2n$. Conversely, if $m < 2n$ then for almost all A there exist $x \neq z$ and permutations $\Pi_1 \neq \Pi_2$ such that $\Pi_1 A x = \Pi_2 A z$.

Proof: Recoverability when $m \geqslant 2n$ is a straightforward consequence of Theorem 1 and Lemma 2, by noting that for any two permutations Π_1 and Π_2, $\Pi_1^{-1} \Pi_2$ is also a permutation. To prove the converse, note that cyclic shift by 1 (which is a permutation) has m distinct eigenvalues $\{ e^{i2k\Pi/m} : k \in \{0, 1, \dots, m-1\} \}$, so that all its eigenspaces have dimension 1. Whenever $m < 2n$, this implies $p(T) = 1 \leqslant m - n$, and the claim follows from Proposition 1. Namely, denoting the cyclic shift by Π_c, for all Π_1, Π_2 such that $\Pi_1^\top \Pi_2 = \Pi_c$ and almost all A, there exist x, z such that $\Pi_1 A x = \Pi_2 A z$ and $x \neq z$. ∎

III. EXTENSION TO ROW-SELECTION MATRICES

In [18] the authors state a more general result that allows row-selection matrices. They prove that instead of a permutation of Ax, one can measure any permuted subset of k entries of Ax and still get unique recovery as long as $k \geqslant 2n$. Although this case is beyond the scope of this letter, we outline an intuition for how our arguments might apply.

The measurement can be written as $y = R\Pi A$, where R is the top k rows of an $m \times m$ identity matrix and Π an unknown permutation. Same as before, it is sufficient to show that for almost all A and two fixed permutations Π_1 and Π_2

$$R\Pi_1 A x = R\Pi_2 A z \quad (12)$$

implies $x = z$. The same R can be used on both sides since different row selections can be absorbed into Π_1 and Π_2. Once that is established an argument parallel to Theorem 1 proves the result.

Suppose that (12) holds. Both $R\Pi_1 A$ and $R\Pi_2 A$ consist of k rows of A in some permuted order. Some rows of A, denote them by A_C, might be present in both $R\Pi_1 A$ and $R\Pi_2 A$, while some, denote them by V, W, appear in only one of them. We can thus rewrite (12) as

$$\Pi_1 \begin{bmatrix} V \\ A_C \end{bmatrix} x = \Pi_2 \begin{bmatrix} W \\ A_C \end{bmatrix} z \implies \begin{bmatrix} V \\ A_C \end{bmatrix} x = \Pi \begin{bmatrix} W \\ A_C \end{bmatrix} z, \quad (13)$$

for some permutations Π_1 and Π_2, where $\Pi = \Pi_1'^\top \Pi_2$. We allow any of the blocks to be empty.

At one extreme where A_C is empty, we can absorb Π in W and ask when it can be that $Vx = Wz$? But $\mathcal{R}(V) \cap \mathcal{R}(W) = \{0\}$ for almost all V, W and hence almost all A (see Fig. 1). The detailed discussion of Lemma 1 is not needed because now W varies independently of V (the number of the degrees of freedom doubles).

At the other extreme where V and W are empty, A_C has at least $2n$ rows. Lemma 1 and Corollary 1 guarantee that for almost all A_C (and hence almost all A), the range of A_C does not intersect the range of ΠA_C unless Π has a large eigenspace, in which case this eigenspace corresponds to $\lambda = 1$. Interpolating between empty V and W and empty A_C, we are adding degrees of freedom and making the two matrices less dependent, which makes range intersections less likely.

IV. CONCLUSION AND FUTURE WORK

We presented a generalization of the classical unlabeled sampling canon. Instead of recovering x from an unknown permutation of $y = Ax$, we showed that it can be recovered from rather general linear transformations of y as long as the set of transformations is at most countably infinite and A has sufficiently many rows. As a byproduct, we get a simple, geometric proof of the uniqueness result for classical permutation-based unlabeled sensing.

The set of transformations $\mathcal{T}$ could model unknown room transfer functions where A takes bandlimited spatial samples of speech. It could model different cameras and projections, or the variety of available sensors in any modality. In the classical unlabeled setting, we can expect the permutation ambiguity to be compounded by other uncertainties which can be modeled by $\mathcal{T}$—unknown filters, offsets, spatially-varying gains, etc.

An interesting line of future work is to relax assumptions on T. The fact that T is diagonalizable or invertible does not seem essential, as long as its nullspace is not too large compared to the range space of A. It also seems plausible that nonlinear T should work. The main practical question is that of stability and polynomial-time recovery algorithms. For the case of permutations, results are beginning to emerge; these will point the way to algorithms for more general transformations.

REFERENCES

[1] V. Emiya, A. Bonnefoy, L. Daudet, and R. Gribonval, "Compressed sensing with unknown sensor permutation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2014, pp. 1040–1044.

[2] A. V. Balakrishnan, "On the problem of time jitter in sampling," *IRE Trans. Inf. Theory*, vol. 8, no. 3, pp. 226–236, 1962.

[3] P. Marziliano and M. Vetterli, "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3462–3471, Dec. 2000.

[4] M. Kreković, I. Dokmanić, and M. Vetterli, "EchoSLAM: Simultaneous localization and mapping with acoustic echoes," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 11–15.

[5] I. Dokmanić, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," *Proc. Nat. Acad. Sci.*, vol. 110, no. 30, pp. 12 186–12 191, Jun. 2013.

[6] I. Dokmanić, L. Daudet, and M. Vetterli, "How to localize ten microphones in one finger snap," in *Proc. Eur. Signal Process. Conf.*, 2014, pp. 2275–2279.

[7] I. Dokmanić, L. Daudet, and M. Vetterli, "From acoustic room reconstruction to slam," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 6345–6349.

[8] S. Huang and I. Dokmanić, "Reconstructing point sets from distance distributions," 2018, arXiv: 1804.02465.

[9] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.

[10] X. Song, H. Choi, and Y. Shi, "Permuted linear model for header-free communication via symmetric polynomials," in *Proc. IEEE Int. Symp. Inf. Theory*, 2018, pp. 661–665.

[11] J. Ranieri, I. Dokmanić, A. Chebira, and M. Vetterli, "Sampling and reconstruction of time-varying atmospheric emissions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Oct. 2012, pp. 3673–3676.

[12] M. Martinez-Camara, I. Dokmanić, J. Ranieri, R. Scheibler, M. Vetterli, and A. Stohl, "The Fukushima inverse problem," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2013, pp. 4330–4334.

[13] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3790–3804, Oct. 2006.

[14] M. C. Tsakiris, L. Peng, A. Conca, L. Kneip, Y. Shi, and H. Choi, "An algebraic-geometric approach to shuffled linear regression," 2018, arXiv: 1810.05440.

[15] S. Basu and Y. Bresler, "Uniqueness of tomography with unknown view angles," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1094–1106, Jun. 2000.

[16] Z. Zhao and A. Singer, "Rotationally invariant image representation for viewing direction classification in Cryo-EM," *J. Structural Biol.*, vol. 186, no. 1, pp. 153–166, 2014.

[17] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, "Unlabeled sensing: Solving a linear system with unordered measurements," in *Proc. IEEE 53rd Annu. Allerton Conf. Commun., Control, Comput.*, 2015, pp. 786–793.

[18] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, "Unlabeled sensing with random linear measurements," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3237–3253, Apr. 2018.

[19] S. Haghighatshoar and G. Caire, "Signal recovery from unlabeled samples," Jan. 2017, arXiv: 1701.08701.

[20] A. Pananjady, M. J. Wainwright, and T. A. Courtade, "Linear regression with an unknown permutation: Statistical and computational limits," Aug. 2016, arXiv:1608.02902.

[21] M. C. Tsakiris, "Eigenspace conditions for homomorphic sensing," 2018, arXiv: 1812.07966.

[22] J. Najnudel and A. Nikeghbali, "The distribution of eigenvalues of randomized permutation matrices," *Ann. Inst. Fourier*, vol. 63, no. 3, pp. 773–838, 2013.