

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Orbit Recovery in Dihedral and Projected Multi-Reference Alignment

Permalink

<https://escholarship.org/uc/item/1pk2n3kx>

ISBN

9798193265542

Author

Weicht, Tait Josef

Publication Date

2026-09-08

Peer reviewed|Thesis/dissertation

Orbit Recovery in Dihedral and Projected Multi-Reference Alignment

By

TAIT JOSEF WEICHT
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Applied Mathematics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Albert Fannjiang, Co-Chair

Alexander Wein, Co-Chair

Yunpeng Shi

Committee in Charge

2026

To my mom, who has been my biggest supporter from the beginning.

Contents

Abstract	v
Acknowledgments	vii
Chapter 1. Introduction	1
1.1. Research Arc	1
1.2. Outline	5
1.3. Notational Conventions	7
Chapter 2. Tensor Decomposition and Algebraic Geometry	8
2.1. Tensors and Indexing	8
2.2. The Canonical Polyadic (CP) Decomposition	9
2.3. Symmetric Tensors	12
2.4. Cutting out Varieties of Rank-1 Tensors	13
Chapter 3. Improving Bounds for Finding Rank-1 Matrices in a Subspace	17
3.1. Introduction	17
3.2. Setting and Results	21
3.3. Preliminaries and Perspective	24
3.4. Conjecture and Supporting Results	26
3.5. Extension to Symmetric Matrices	33
3.6. Future Directions	36
Chapter 4. Tensor Formats, Symmetrized Kronecker Products and Symmetric Lifts	37
4.1. Symmetric Tensors in an Irredundant Format	37
4.2. Linear Maps between Symmetric Tensor Spaces	40
4.3. Tensors with Partial Symmetries.	45
4.4. Efficiently Cutting Out a Segre-Veronese Variety	46

Chapter 5. Beyond Frequency Marching: Orbit Recovery in Dihedral and Projected Multireference Alignment	52
5.1. Introduction	52
5.2. Dihedral MRA and MRA with Projection	60
5.3. Method for Recovery	65
5.4. Analysis of the Aided Recovery Procedure	83
5.5. Analysis of the Base Case	95
5.6. Putting it Together: Proofs of Main Results	104
5.7. Efficient Implementation	106
5.8. Future Directions	111
5.9. Deferred Proofs	112
Bibliography	118

Abstract

Methods for structure determination in single particle cryogenic electron microscopy upended imaging sciences by demonstrating that crystallization is not necessary to resolve biological molecules at the atomic level. As imaging techniques have matured, so has the mathematical theory used to analyze these data. In particular, the method of moments and algebraic techniques show promise for the challenges posed by a low signal-to-noise ratio, the unknown orientation of particles in the imaging plane, and different classes of particles in a mixture. Cryo-EM remains a challenging problem approached from this direction and has motivated the study of simpler models capturing its core challenges. In particular, multireference alignment (MRA) is a well-studied model in which one observes copies of a signal corrupted by unknown cyclic shifts and isotropic Gaussian noise.

In this work, I study two variants of multireference alignment, one introducing a step modeling tomographic projection (projected MRA) and another allowing reflections of the signal as well as cyclic shifts to appear (dihedral MRA). For signals of length a power of two, I develop an efficient recursive method for recovering signals from the third moment tensor of the samples. Beyond these specific orbit recovery problems, a subroutine using in my algorithm solves a linear system with a solution constrained to be a symmetric rank-1 matrix. Such problems arise in algebraic compressive sensing. Even when a given linear system is under-determined, the rank-1 constraint provides a “sparsity” assumption which may ensure a unique solution. To prove that my algorithm succeeds, I develop a construction which can be used to certify that specific linear systems with a rank-1 constraint have unique solutions. We show the algorithm succeeds in recovering a generic signal on the condition that a particular symbolic matrix has full column-rank. This condition can be easily checked on the computer in polynomial time, and we give strong evidence the matrices have full column-rank for a couple of dimensions.

The foray into algebraic compressive sensing also reveals a close connection to the problem of finding a set of rank-1 matrices “planted” in a subspace of matrices. An algorithm of Johnston, Lovitz and Vijayaraghavan for finding the intersection of a linear subspace with a conic variety can recover such rank-1 matrices as a particular application. However, the upper bound on generic subspace dimension for which these authors prove their algorithm will succeed proves pessimistic in this specific case. We conjecture that a simple dimension counting bound gives the sharp threshold for the maximum dimension of generic subspace from which the JLV intersection algorithm can successfully recover a set of planted rank-1 matrices. To give evidence for this conjecture, I implemented a method of computer-proof which certifies the sharp boundary in several thousand specific cases. We give an analogous conjecture for the variety of symmetric rank-1

matrices and similar computer-proof of a similar sharp boundary in this case. Moreover, I give a floating point implementation of the JLV intersection algorithm in Python which can compute the intersection of any linear subspace with any Segre-Veronese variety (the varieties of rank-1 tensor with any set of partial symmetries and of any shape). Using this implementation, I give further evidence for our conjecture, showing that the algorithm does indeed recover planted matrices beyond the bound proven in the original paper up to our newly conjectured threshold. As an immediate consequence, we more than double the maximum rank of generic third- and fourth-order tensors (with or without partial symmetry) for which the JLV intersection algorithm will give a provably correct decomposition.

Acknowledgments

A great thank you to my two co-advisors, Dr. Alexander Wein and Dr. Albert Fannjiang. Alex, you pushed me to write more precisely and communicate more clearly, and my work would be in a substantially rougher place without your attention to detail and helpful critiques. Albert, you pushed me to clarify steps and computations carefully when I made leaps and encouraged my curiosity in a variety of mathematical areas. I also appreciate your kindness and understanding as I went through challenges which impacted my ability to focus and my mental well-being.

The research in Chapters 3, 4 and 5 was supported through the both US National Science Foundation grant CCF-1934568 for the UC Davis TETRAPODS Institute of Data Science (UCD4IDS) and the National Science Foundation CAREER Award CCF-2338091 (to Dr. Wein). A sincere thanks to Dr. Naoki Saito, who invited me to be part of UCD4IDS. Thanks also to Dr. Matthias Köppe for helpful consultation and support for work in Chapter 5.

To those in the front office and lecturers in the UC Davis math department, you were some of my biggest supporters. Tina Denena, Marji LeGrand, and Minh Hoang, you each at various times offered me guidance and emotional support. To Dr. Korana Burke and Dr. Thomas Rohit, your advice has been invaluable and you gave me so many opportunities to be an educator. Dr. Janko Gravner and Dr. Edward Tavernetti, it was a pleasure serving COSMOS cluster 6 as a teaching assistant. Dr. Ali Dad-del, it was a joy to join COSMOS cluster 5 and build robots with such great students. Your holistic approach to these students' experience is something I will remember as I seek to teach well.

Next, thank you to friends and colleagues in and out of the math department. To the folks at the Queer in Computational and Applied Math workshops I and II, you made me feel that there was space for me in mathematics. In particular, Dr. Michael Roberts, Dr. Andy Bernoff, and Dr. Julie Blackwood, your interest in my career development and the guidance and resources you offered have immeasurably encouraged me. To my officemates, Mary Claire and Connor, I will miss seeing you in the office, on Russell field for Frisbee, and in Community Park for spikeball. Sharing the ups and downs of graduate school with you helped get me through. Kayleigh and Chris, you are most excellent hosts, and I will forever treasure dissecting messy TV relationship drama with you. Nathan, Jamie, Darian, and Jordan, you've all kept me grounded. Brian, you have been a great housemate, climbing buddy and friend.

Finally, my family has supported me each step of the way, especially my mom, who I can always count on for getting to the heart of any matter and for sound advice.

Introduction

1.1. Research Arc

It was in the early stages of graduate school that I was introduced to the inverse problem of structure determination in single particle cryogenic electron microscopy (cryo-EM). In single particle cryo-EM, many particles are frozen in a thin layer of ice. A transmission electron microscope takes an image of the entire ice layer from which individual particle locations may be determined on the micrograph. From a theoretical perspective, determining a protein or small particle structure from a micrograph is not one problem but has multitude of mathematical approaches each with open questions and undiscovered improvements. Structure determination in single particle cryo-EM is made challenging by several factors:

- (1) often, each individual particle image has a very low signal-to-noise ratio (SNR),
- (2) the orientation of each particle relative to the imaging plane is unknown,
- (3) images recorded are 2D and (tomographic) projections of the true 3D structure, and
- (4) particles may not be identical having gross structural differences of small differences in conformation.

The first factor raises the question, ‘how does the number of samples needed depend on the SNR to ensure a given accuracy of recovery?’ This scaling is known as the *sample complexity* of the problem. Taking preparing samples and imaging them is expensive and takes time; knowing how much information can be squeezed from a dataset is valuable. A separate question but equally important is whether efficient algorithms exist that can achieve a known sample complexity.

The second restrict which methods may be used for recovery since one must either identify the orientation of each particle relative to some canonical orientation or seek methods that avoid the issue. Under the assumption that a particle has no preferred orientation, one can only recover the structure up to a 3D rotation. One could first aim to determine the orientations of particles up to a global rotation, a sub-task called *synchronization*. If one instead treats individual particle orientations as nuisance parameter, structure determination falls in a category of problems termed *orbit recovery* since the structure is only determined up to the action of a group upon the structure, $SO(3)$ in the case of cryo-EM. The third factor

make reconstructing 3D structure more challenging. Under parallel projection one can no longer distinguish an object and its mirror image. For biological molecules, one must use other information to determine a particle’s chirality. On the image level, the Fourier-slice theorem guarantees that a projection in real space correspond to a slice through the origin in the reciprocal (Fourier) space. As a result, every pair of particle images will overlap along a direction in the reciprocal space, called a *common line*. However, both noise and the fourth factor– the possibility of heterogeneity– can confound methods that make use of common lines estimates.

Initially, I was drawn to this problem by the elegant spectral method of determining orientations developed by Singer and Shkolnisky [62] in which the eigenspaces associated to the top three eigenvalues of a matrix contain information from which particle orientations may be determined. While this work determines how the number of images must scale with the error, it does not examine how the SNR affect the probability that common lines are identified correctly. Further, an elegant argument states that when the SNR is too low it should be impossible to correctly guess any relative orientation information about images: even if one had the true particle structure, when noise is high enough one cannot accurately guess which direction an image is taken from, more formally show by Aguerrebere, Delbracio, Bartesaghi and Sapiro [4].

Discussions with Dr. Alexander Wein about cryo-EM and sample complexity led me to *multireference alignment (MRA)*. MRA is a simpler problem than cryo-EM but captures many of salient complexities. In the MRA model, one aims to estimate a vector of real values we call the signal. One has access to copies of the signal with entries cyclically shifted by an unknown number of indices and corrupted by adding independent Gaussian noise. More broadly, extensions of MRA and MRA-like models consider group actions beyond cyclic shifts, different sampling schemes, and mixtures of different signals [2, 3, 6, 14, 18]. The sample complexity of the original MRA model was proven to be of order $1/\text{SNR}^3$ [3, 10, 56]. The work by Perry, Weed, Bandeira, Rigollet and Singer [56] also showed that there is an efficient algorithm to achieve orbit recovery: running simultaneous diagonalizable on the empirical third moment tensor of the sample distribution will recover each cyclic shift of the signal. The basic MRA problem has been solved; however, orbit recovery in cryo-EM and many MRA variants remain open problems.

The work by Perry and co-authors follows a strategy for parameter estimation called the method of moments. The signal may be considered as a set of parameters underlying a multivariate probability distribution from which samples are drawn. For such a multivariate distribution density, computing a sample average estimates the distribution’s mean (a vector), and centering samples then taking their outer product and averaging estimates the distribution’s covariance (a matrix). Taking a k -order outer products of sample vectors and averaging produces an estimate of the distribution’s k -th moment, which is a tensor. In our

context, a tensor is simply a k -way array of numbers. Broadly, tensor decomposition then aims to split a tensor into components in a meaningful way. In the case of MRA, the *canonical polyadic (CP) decomposition* splits the third moment into a sum of symmetric rank-1 tensors and each rank-1 component when factored yield a different cyclic shift of the signal. More generally, one hopes that the methods used to decompose tensors can be turned to extracting parameters of interest from moments.

Determining how to extract the structure of a molecule using the method of moments in single particle cryo-EM continues to be an active area of research [11]. Some directions of study involve taking the model and removing features that constitute challenges e.g. algorithms for orbit recovery of “unprojected” cryo-EM [7, 50]. Alternatively, one can take the MRA model and include additional features e.g. including reflections and cyclic shifts of the signal in the sample distribution (dihedral MRA) [13], modeling a continuous cycle shift operation (continuous MRA) [55], or adding a “projection” step (projected MRA) [7]. In my work here, I studied two modifications to the original MRA problem. In the following I give a chronological overview of the work I did, whereas the order in this dissertation follows.

GOAL 1. Determine how to achieve orbit recovery efficiently (in polynomial time) for MRA with a projection step and dihedral MRA.

Initially, I began by studying a model of projected MRA. In projected MRA, one applies an operation that adds together entries of a vector after a cyclic shift has been applied to simulate the tomographic projection of an object (see Figure 5.1 in Chapter 5). Bandeira, Blum-Smith, Kileel, Niles-Weed, Perry and Wein propose one model MRA with projection with an odd dimensional signal [7, Problem 4.7]. I used the even dimensional formulation that was originally rejected since I realized that if one has a moment from this projected MRA sample distribution, one can extract a sub-tensor which corresponds to a moment for a smaller signal. This observation allows one to recursively reduce the orbit recovery problem to one half the size by extracting sub-tensors.

Under this recursive framework, one can recover half information in the signal. I determined that when this information is known, one can set up an under-determined linear system, but one wherein the solution has the structure of a rank-1 symmetric matrix. Problems like this fall into a broader class of problems called “linear systems with a canonical polyadic decomposition (CPD)-constrained solution” or LS-CPDs. The low rank condition on the solution may be considered an algebraic version of sparsity— as appears in models of compressed sensing. Thus LS-CPDs appear even more broadly in the field of algebraic compressed sensing with related problems include matrix completion and tensor completion [20]. Boussé, Vervliet, Domanov, Debals, and Lathauwer [19, Section 3] give an exact algebraic algorithm for solving such systems in the

rank-1 case. They give a necessary condition to ensure that their algorithm recovers unique solutions and can guarantee uniqueness for linear operator which has a range in “general position” provided that range is large enough with respect to the solution’s dimension (by general position or generic we mean that there may be a measure zero set of exceptional matrices that do not meet the guarantees, see Definition 3.2.1 for a formal statement). However, we have no guarantees that the specific linear systems arising from the projected MRA problem are in general position. In practice, though our implementations of this algorithm repeatedly succeeded in solving randomly generated problems of this type.

GOAL 2. Determine a sufficient condition to verify that a specific class of linear system with a rank-1 structure constraint will have unique solutions.

We determine such a sufficient condition and can then guarantee our recursive algorithm will recover a generic signal, provided a certain symbolic matrix has full column-rank. We offer strong numerical evidence that such these symbolic matrices do have full column-rank for a couple of dimension of problem size.

Concurrently, my advisor and a fellow graduate student, Jeshu Dastidar, began studying an algorithm by Johnston, Lovitz, and Vijayaraghavan [43] for finding the intersection of a conic variety with a linear subspace. If the subspace contains certain “planted” solution belonging to the variety, under certain conditions the algorithm recovers these. The JLV algorithm provided a breakthrough in decomposing fourth-order tensors of rank quadratic in the tensors’ dimension. Despite the significant differences in presentation and context, [19, Algorithm 1] can be consider an equivalent to [43, Algorithm 1] restricted to the varieties of rank-1 tensors and with some additional pre-processing to take affine (rather than linear) sections. Moreover, our numerical work on the rank-1 constrained linear systems arising in from the projected MRA problem suggested that [43, Theorem 2] was a pessimistic bound; the JLV algorithm for intersections could work for substantially larger generic subspaces than currently proven.

GOAL 3. Find sharp bounds on the dimension of generic subspace for which the JLV algorithm succeeds in finding the intersection of varieties of rank-1 matrices.

Work by my collaborators doubled the subspace dimension bound previously known and determined a definitive breaking point at an even higher threshold (see Proposition 3.2.6). However, I provided evidence that, in fact, the boundary is sharp and the JLV algorithm succeeds up until to this boundary. Modifying the original algorithm to perform some computations over a finite field, I can prove our conjecture of a sharp threshold for specific cases or generate evidence against it (in practice the algorithm generates no evidence against the conjecture in several thousand cases tested). I also provide the first numerical implementation

of the JLV algorithm for the family of Segre-Veronese varieties, allowing one to take the intersection of a subspace with any shape of rank-1 tensors having any number of partial symmetries across modes.

Main Contributions.

1. I developed a robust implementation JLV intersection algorithm in Python to find the intersection of any Segre-Veronese variety with a linear subspace.
2. I modified the JLV algorithm to provide computer-proof of specific cases of a conjecture (Conjecture 3.4.1) suggesting a sharp bound on the dimension of the subspace from which the JLV can correctly extraction planted rank-1 matrices.
3. I generalized the conjecture and computer-proof techniques to rank-1 matrices with an enforced symmetry (Section 3.5).
4. I repurposed the machinery used in work on the JLV algorithm to develop a polynomial-time test for when a specific class of linear systems with a rank-1 constrained solution will have a unique solution.
5. Using work on finding solutions to linear systems with rank-1 constrained solutions, I developed an recursive polynomial-time method which can recover the orbit of a signal with dimension a power of two from third moments in both the dihedral and projected MRA models, provided the test described above is passed.
6. I give numerical evidence that the linear systems with rank-1 constrained solutions arising in our orbit recovery problems do have unique solutions for generic signals.
7. I do some preliminary sensitivity analysis by running numerical experiments to see how small floating point errors accumulate through steps in my recursive algorithm.

1.2. Outline

Chapter 2 outlines some key terminology and notation for dealing with tensors, tensor decomposition, and some algebraic geometry needed to understand further work. This material is not novel and we draw from the review of Kolda [46] on tensor decomposition, Moitra’s book [54] detailing applications, and textbooks for some key definitions from algebraic geometry.

In Chapter 3, we present work done with co-authors Jeshu Dastidar and Alexander Wein improving a bound an algorithm of Johnston–Lovitz–Vijayaraghavan which finds the intersection of a conic variety with a linear subspace [24]. This work has been published to arXiv and is in the review process for publication in the SIAM Journal of Applied Algebraic Geometry. The JLV algorithm and work by Johnston, Lovitz,

Vijayaraghavan provided a breakthrough in decomposing fourth-order tensors of rank quadratic in the tensors' dimension. In our work, we show that the original authors' proof of the algorithm's correctness gives a pessimistic upper bound on the rank of a tensor for which it may find a decomposition; instead, we show algorithm can decompose tensors with rank roughly twice as large as previously guaranteed. Further, we gives a negative result that the JLV algorithm will fail somewhat above the positive result. We conjecture that the intersection algorithm actually works up until this provable breaking point; further, we prove the conjecture for several thousand specific cases using an algorithm we wrote. Chapter 3 contains an edited version of [24] which emphasizes this author's contributions consisting of an implementation of the JLV algorithm in Python to find the intersection of any linear subspace with any Segre-Veronese variety, which include varieties of rank-1 matrices. I also programmed modifications to the algorithm so that it can produce computer-aided proof of the conjecture for any fixed set of parameters (or, in theory, give statistical evidence against the conjecture). Further, I extended these computer-aided proofs and numerical experiments to the case of symmetric matrices which our combinatorial arguments do not consider.

Material in Chapters 4 and 5 are adapted from a paper on orbit recovery for dihedral and projected MRA by the author and Alexander Wein [67]. This paper is published on arXiv as a pre-print. All proofs, code and computations were written or run by this author with guidance and critique from Dr. Wein on their structure and substance. The paper has been adapted as two chapters in this dissertation with some new additions providing additional clarification.

In Chapter 4, we give a format for symmetric tensor tensor, a symmetrized Kronecker product and properties of this symmetric product, and examples of how varieties are described in this form which come from [67, Section 2.2, 4.2]. I also include details for an algorithm which finds a basis of the degree-2 polynomials cutting out Segre-Veronese varieties— that is, varieties of rank-1 tensors with some partial symmetries along modes. This subroutine is an important part of my implementation of the JLV algorithm.

In Chapter 5, we add a “projection” step to the MRA model, making the challenge of determining the signal significantly harder. Using a framework of variety-constrained linear algebra, we develop an algorithm which can recover the signal when it has dimension that is a power of two and conditional on a symbolic matrix being full-rank. We conjecture that these symbolic matrices have full column-rank and give strong evidence for this fact by showing that that at we can find at least one random instance of these matrices which has numerical full rank.

1.3. Notational Conventions

This dissertation contains a number of results from rigorous proven statements to numerical experiments. We use the notation of “THEOREM” to denote proven statements. In some cases, the result we aim to prove is equivalent to a particular symbolic matrix being full rank or a polynomial being non-zero. A technique we use in Chapter 3 is performing computations over a finite field to certify these facts and constitutes proof of specific cases, aided by the computer. However, in some cases such computations prove quite slow to run. We can instead perform these in floating-point arithmetic at the cost of numerical precision. Such results (numerical condition numbers and function evaluations) do not constitute proof since there is a chance that the result may be attributed to error. Nonetheless, for such facts verified by floating-point computation, we use the notation “THEOREM*.” In such situations, the use of the word “Theorem” is to convey that the outcome of the floating-point calculations are exceedingly unlikely to be attributed solely to the accumulation of errors due to rounding in floating point arithmetic and the statement has overwhelming evidence for it. This notation is not an original creation, I came across it in a paper in the area of computation algebraic geometry but cannot presently locate this paper in my records or by search.

Tensor Decomposition and Algebraic Geometry

In this Chapter we introduce some of the core concepts developed to support work in Chapters 3 and 5. We introduce the notation we use to work with tensors, the idea of tensor decomposition and tensor rank. We also introduce some key concepts from algebraic geometry relevant to the theory of tensor decomposition.

2.0.1. Preliminary Notation. We use $\mathbb{K}$ to denote either the field $\mathbb{R}$ or $\mathbb{C}$. A statement involving $\mathbb{K}$ will hold when one substitutes either $\mathbb{R}$ or $\mathbb{C}$ for all instances. We will use the shorthand $\mathbb{K}^n$ then to denote an n -dimensional vector space over the field $\mathbb{K}$. We will use the notation $[n]$ to denote the set of whole numbers from 1 to n ,

$$[n] = \{1, 2, \dots, n-1, n\},$$

and the set $[n)$ to denote whole numbers from 0 to $n-1$,

$$[n) = \{0, 1, 2, \dots, n-1\}.$$

We will primarily use elements $[d]$ to index entries of vectors in the standard basis and elements of $[d)$ to index Fourier coefficients of a vector (entries of a vector in the Fourier basis).

2.1. Tensors and Indexing

2.1.1. Tensors and tensor spaces. In this work, we consider a tensor to be a multi-dimensional array of numbers. The *order* of a tensor is the number of indices used to access elements of the tensor. In this context, a vector is an order-1 tensor and a matrix is an order-2 tensor. The shape of a tensor is a tuple of non-zero integers denoting the values that each index may range over.

We use hard brackets to indicate a specific entry of a vector or matrix, e.g. for a vector $x \in \mathbb{K}^n$, $x[j]$ is the j -th entry of x and for a matrix $A \in \mathbb{K}^{m \times n}$, $A[j, k]$ is the entry in the j -th row and k -th column. We also use the placeholder “:” to indicate full range of indices so that e.g. for a matrix A , $A[j, :]$ is the j -th row as a vector and $A[:, k]$ is the k -th column.

We denote a space of k -order tensors with shape $(n_1, n_2, \dots, n_k)$ by

$$\mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \dots \otimes \mathbb{K}^{n_k} = \text{span}(u^1 \otimes u^2 \otimes \dots \otimes u^k : u^1 \in \mathbb{K}^{n_1}, u^2 \in \mathbb{K}^{n_2}, \dots, u^k \in \mathbb{K}^{n_k}).$$

As a short-hand we will also use $\mathbb{K}^{n_1 \times n_2 \times \dots \times n_k} \equiv \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \dots \otimes \mathbb{K}^{n_k}$. The short-hand

$$(\mathbb{K}^n)^{\otimes k} := \underbrace{\mathbb{K}^n \otimes \mathbb{K}^n \otimes \dots \otimes \mathbb{K}^n}_{k \text{ times}}$$

indicates a repeated tensor product and a space of cubic tensors.

For two tensors $S, T \in \mathbb{K}^{n_1 \times \dots \times n_k}$ we define their inner product as

$$(2.1) \quad \langle S, T \rangle := \sum_{j_1=1}^{n_1} \dots \sum_{j_k=1}^{n_k} \overline{S[j_1, \dots, j_k]} T[j_1, \dots, j_k].$$

2.1.2. Flattening and reshaping. If a vector or matrix has extra structure we may index into it using tuples to emphasize this structure. Flattening (or generally, reshaping) using an isomorphism $\mathbb{K}^m \otimes \mathbb{K}^n \cong \mathbb{K}^{mn}$ may be indicated using our indexing system e.g. if $A \in \mathbb{K}^m \otimes \mathbb{K}^n$ and we wish to define a flattening to $v \in \mathbb{K}^{mn}$ then we use

$$v[(i, j)] := A[i, j]$$

where a lexicographical ordering on (i, j) gives a bijection between standard basis vectors of $\mathbb{K}^{mn}$ and $\mathbb{K}^m \otimes \mathbb{K}^n$. In general, the tensor space $\mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \dots \otimes \mathbb{K}^{n_k}$ is isomorphic to the vector space $\mathbb{K}^{n_1 n_2 \dots n_k}$. The Kronecker product of vectors $u^1 \in \mathbb{K}^{n_1}, u^2 \in \mathbb{K}^{n_2}, \dots, u^k \in \mathbb{K}^{n_k}$ is a vector $u^1 \otimes u^2 \otimes \dots \otimes u^k \in \mathbb{K}^{n_1 n_2 \dots n_k}$ defined entry-wise by

$$(u^1 \otimes u^2 \otimes \dots \otimes u^k)[(j_1, j_2, \dots, j_k)] = u^1[j_1] u^2[j_2] \dots u^k[j_k],$$

where our indexing by $(j_1, j_2, \dots, j_k) \in [n_1] \times [n_2] \times \dots \times [n_k]$ implicitly uses a fixed bijection with an element of $[n_1 n_2 \dots n_k]$ e.g. by using the lexicographic order on the tuples. We say that vectors arising from such Kronecker products have a *tensor product structure*.

In the other direction, for $x \in \mathbb{K}^m$ and $y \in \mathbb{K}^n$ we may accomplish the reshaping of $v \in \mathbb{K}^{mn}$ to a matrix $A \in \mathbb{K}^m \otimes \mathbb{K}^n$ via

$$A[i, j] := v[(i, j)].$$

2.2. The Canonical Polyadic (CP) Decomposition

Considering a third-order tensor $T \in \mathbb{K}^{p \times q \times r}$, the *canonical polyadic decomposition* of a tensor T is a way of it as a sum of rank-one matrices e.g.

$$(2.2) \quad T = \sum_{j=1}^R a^{(j)} \otimes b^{(j)} \otimes c^{(j)}$$

where the sets of vectors $\{a^{(j)}\}_{j=1}^R \subset \mathbb{K}^p$, $\{b^{(j)}\}_{j=1}^R \subset \mathbb{K}^q$ and $\{c^{(j)}\}_{j=1}^R \subset \mathbb{K}^r$ are called *factors*. This definition extends in a natural way to higher order tensors. One can also collect the factors into *factor matrices* $A \in \mathbb{K}^{p \times R}$ where $A[:, j] = a^{(j)}$ and B, C are defined analogously. Using factor matrices we may use the notation

$$(2.3) \quad T = \llbracket A, B, C \rrbracket := \sum_{j=1}^R A[:, j] \otimes B[:, j] \otimes C[:, j]$$

to concisely define a tensor by the factors that compose it. Such decompositions can only be unique up to reordering the elements in the summand, i.e. if Π is an $R \times R$ permutation matrix

$$T = \llbracket A, B, C \rrbracket = \llbracket A\Pi, B\Pi, C\Pi \rrbracket,$$

and rescaling of the factors in a way that preserves the total Frobenius norm of each of the rank-1 tensor summands i.e. if D_1, D_2 and D_3 are $R \times R$ diagonal matrices such that $D_1 D_2 D_3 = I$ then

$$T = \llbracket A, B, C \rrbracket = \llbracket AD_1, BD_2, CD_3 \rrbracket.$$

The *rank* of a tensor is defined to be the smallest number of factors needed to write $\mathcal{T}$ as an CP decomposition [40, 48]. A CP decomposition with the minimal number of factors used is often called a *rank decomposition* [46], though in this work we will use CP decomposition to mean one with minimal rank.

For tensor of order three and higher, as the term “canonical” suggests, often under weak assumptions a rank decomposition is unique up to the indeterminacies described above. This does not occur for matrices, indeed if $X \in \mathbb{R}^{m \times n}$ has a factorization $X = AB^\top$ then $U = AW$ and $V = BW$ where W is any real orthogonal matrix gives an alternative decomposition $X = UV^\top$. The indeterminacy described is called “the rotation problem” and poses an interpretability challenge when trying to determine correlated effects in data. The issue of uniqueness is remedied if one requires that factors be orthogonal which gives the singular value decomposition (SVD) of X . In contrast, a third-order tensor need not have orthogonal sets of factors for that decomposition to be unique, see the excellent review by Kolda [46] for a comprehensive overview of what is known about uniqueness. Uniqueness under these relatively weak conditions makes the CP decomposition attractive as a tool used in unsupervised learning, applied to the method of moments, and in the analysis of temporal data, multi-relational data, latent variables, phylogenetic trees, topic models, etc. See [46, 54, 58, 66] for further discussion.

Unfortunately, it is challenging to determine utilize these nice properties of the CP decomposition since, in general, determining the CP decomposition is an NP-hard computational problem [42]. Further, while the

Eckart-Young theorem states that the best rank- R approximation of a matrix (order-2 tensor) in Frobenius norm can be found by truncating the singular value decomposition, no such result holds for higher order tensors with regards to truncating the CP decomposition. In general, trying to find a lower rank approximation to a tensor may even be ill-posed in the sense that Indeed, de Silva and Lim [27] constructs examples of tensors spaces where lower rank approximation is ill-posed for a positive volume of tensors (i.e. the set of these tensor has a non-trivial Lebesgue measure). These challenges holds broadly, but by setting some conditions on the tensors one aims to decompose there are efficient algorithms. We highlight the simultaneous decomposition algorithm (also called Jennrich’s algorithm). This algorithm has been rediscovered multiple times and is often attributed to a working paper of Harshman [39] though the author credits Dr. Robert Jennrich. In actuality, this paper describes the method of alternating least squares (ALS) to estimate factor matrices.

THEOREM 2.2.1 ([49]¹). *Consider a tensor $T \in \mathbb{K}^p \otimes \mathbb{K}^q \otimes \mathbb{K}^r$ such that*

$$T = \sum_{j=1}^R a^j \otimes b^j \otimes c^j$$

*where the sets of vectors $\{a^j\}_{j=1}^R$ and $\{b^j\}_{j=1}^R$ are linearly independent and the final set $\{c^j\}_{j=1}^R$ contains pairwise-linearly independent vectors. Then the decomposition is unique and an efficient algorithm, **Simultaneous Diagonalization (SD)**, exists to find the decomposition.*

The algorithm is relatively simple, requiring as subroutines the ability to compute SVDs and the Moore–Penrose psuedoinverse of a matrix, denoted as A^+ for a matrix A .

Simultaneous Diagonalization (SD) [49]

Input: a tensor T satisfying the conditions given in Theorem 2.2.1.

Output: factor lists $\{a^{(i)}\}_{i=1}^R \subseteq \mathbb{K}^p$, $\{b^{(i)}\}_{i=1}^R \subseteq \mathbb{K}^q$ and $\{c^{(i)}\}_{i=1}^R \subseteq \mathbb{K}^r$ (up to order and rescaling indeterminacies).

- (1) Select vectors u, v from the unit sphere in $\mathbb{K}^r$, uniformly at random.
- (2) Form the matrices $T_u := \sum_{j=1}^r \overline{u[j]} \mathcal{T}[:, :, j]$ and $T_v := \sum_{j=1}^r \overline{v[j]} \mathcal{T}[:, :, j]$.

¹The uniqueness result attributed to Dr. Jennrich starts with the assumption that $\{c^j\}_{j=1}^R$ is also linearly independent so that one implicitly requires T have rank $R \leq \min\{p, q, r\}$ rather than $R \leq \text{median}\{p, q, r\}$. Kolda suggests the paper of Leurgans, Ross and Abel as “an *apropos* single source since the originator of the idea, if there was a single such person, has been lost to time.”

- (3) Compute the eigenvalues and associated eigenvectors of $T_u(T_v)^+$, call these λ_j and $\{a^j\}$, $j \in [R]$, respectively. Compute the eigenvalues and vectors of $((T_u)^+T_v)^\top$, call these μ_j and $\{b^j\}$, $j \in [R]$.
- (4) If possible, pair up a^j with b^k when $\lambda_i = (\mu_k)^{-1}$. Otherwise, report failure.
- (5) Solve for each c^j so that $\mathcal{T} = \sum_{j=1}^R a^j \otimes b^j \otimes c^j$.

We note that the SD algorithm is usually given for decompositions over $\mathbb{R}$ but one can verify that that when matrices are complex, the same arguments for correctness hold (see e.g. the presentation by Moitra [54]). For example, in Step 2 we conjugate the entries of u and v to match the definition of the inner product in (2.1).

However, these conditions on a tensor where simultaneous diagonalization may be used to find a decomposition are quite stringent. For a third-order complex tensor $T \in \mathbb{C}^p \otimes \mathbb{C}^q \otimes \mathbb{C}^r$, the *expected generic rank* is

$$(2.4) \quad R = \left\lceil \frac{pqr}{1 + (p-1) + (q-1) + (r-1)} \right\rceil.$$

If we sample “random” tensor (tensor from any absolutely continuous distribution on the space $\mathbb{C}^p \otimes \mathbb{C}^q \otimes \mathbb{C}^r$), this is the rank one should expect the tensor to have by a simple dimension counting argument.² Further, work by Chiantini, Ottaviani and Vannieuwenhoven [23] showed that in many cases, a generically chosen tensor of rank one less than expected generic rank is *identifiable* i.e. has a unique decomposition. This leaves a significant gap between what tensors simultaneous diagonalization can decompose and the tensors that might be encountered in the wild and has motivated research into more powerful tensor decomposition algorithms.

2.3. Symmetric Tensors

Let S_k be the symmetric group on the elements of $[k]$. A tensor $T \in (\mathbb{K}^n)^{\otimes k}$ is symmetric when for all indices $j_1, j_2, \dots, j_k \in [n]$ and any permutation $\sigma \in S_k$ we have

$$T[j_1, j_2, \dots, j_k] = T[j_{\sigma(1)}, j_{\sigma(2)}, \dots, j_{\sigma(k)}].$$

²There are both exceptions to this for small exceptional values of p, q and r and in the case of unbalanced tensors e.g. ones where $r > pq - (p-1) - (q-1)$ [23].

We denote the space of symmetric tensors $S^k(\mathbb{K}^n) \subseteq (\mathbb{K}^n)^{\otimes k}$. For a tensor without symmetry its symmetric component is given by the projection $\text{Sym}_k : (\mathbb{K}^n)^{\otimes k} \rightarrow (\mathbb{K}^n)^{\otimes k}$ defined index-wise for $j_1, j_2, \dots, j_k \in [n]$ by

$$(2.5) \quad \text{Sym}_k(T)[j_1, j_2, \dots, j_k] = \frac{1}{k!} \sum_{\sigma \in S_k} T[j_{\sigma(1)}, j_{\sigma(2)}, \dots, j_{\sigma(k)}].$$

The image of $(\mathbb{K}^n)^{\otimes k}$ under Sym_k is $S^k(\mathbb{K}^n)$. Symmetrization may be analogously defined for vectors that have appropriate tensor product structure.

More generally, for any subspace $\mathcal{U} \subseteq \mathbb{K}^n$, we can define

$$(2.6) \quad S^k(\mathcal{U}) := \text{span}\{\text{Sym}_k(u^1 \otimes u^2 \otimes \dots \otimes u^k) : u^1, \dots, u^k \in \mathcal{U}\}.$$

2.4. Cutting out Varieties of Rank-1 Tensors

A tensor $T \in \mathbb{K}^{n_1 \times n_2 \times \dots \times n_L}$ has CP rank of 1 if we can find non-zero vectors $x^j \in \mathbb{K}^{n_j}$, $j \in [L]$ such that

$$(2.7) \quad T = x^1 \otimes x^2 \otimes \dots \otimes x^L.$$

To better understand the CP decomposition and advances in tensor decomposition, it is helpful to define what it means for a tensor to have rank at most 1 from the perspective of algebraic geometry. Namely, it is first useful to have procedures which allows one to “detect” when a tensor is rank at most one and more. Treating entries of any vector $x \in \mathbb{K}^n$ as symbolic variables, we will use the notation $\mathbb{K}[x] \equiv \mathbb{K}[x[1], x[2], \dots, x[n]]$ to define the polynomial ring over $\mathbb{K}$ in the variables $x[j]$, $j \in [n]$. We will also define polynomial rings using the symbolic variables of other objects e.g. using $\mathbb{K}[A] \equiv \mathbb{K}[A[j, k] : j \in [m], k \in [n]]$ to indicate a polynomial ring in the variables of an $m \times n$ matrix.

DEFINITION 2.4.1 (Algebraic Variety). *An algebraic variety in $\mathbb{K}^n$ is a subset $\mathcal{X} \subseteq \mathbb{K}^n$ such that for a finite set of polynomials $f_1, \dots, f_p \in \mathbb{K}[x]$,*

$$\mathcal{X} = \{x \in \mathbb{K}^n : f_1(x) = f_2(x) = \dots = f_p(x) = 0\}$$

i.e. $\mathcal{X}$ is the zero set of the polynomials $f_1, \dots, f_p$ or the polynomials $f_1, \dots, f_p$ cut out the variety $\mathcal{X}$.

The set of tensors of the form given in (2.7) form what is called a *Segre variety*.

DEFINITION 2.4.2 (Conic Variety). *A conic variety $\mathcal{X}$ over $\mathbb{K}^n$ is a variety such that if $x \in \mathcal{X}$, for any $c \in \mathbb{K}$, $cx \in \mathcal{X}$.*

A polynomial is homogeneous if all its non-zero terms have the same degree. A variety is conic if and only if it is cut out by homogeneous polynomials.

DEFINITION 2.4.3 (Ideal). *An ideal $\mathcal{I}$ of the polynomial ring $\mathbb{K}[x]$ is a non-empty subset of $\mathbb{K}[x]$ such that*

- (1) *if $f \in \mathbb{K}[x]$ and $g \in \mathcal{I}$, then $fg \in \mathcal{I}$;*
- (2) *if $f, g \in \mathcal{I}$, then $f + g \in \mathcal{I}$.*

We can define the ideal of an algebraic variety $\mathcal{X} \subset \mathbb{K}^n$ by

$$(2.8) \quad \mathcal{I}(\mathcal{X}) := \{f \in \mathbb{K}[x] : f(x) = 0 \text{ for all } x \in \mathcal{X}\}.$$

Next, define $\mathbb{K}[x]_d$ to be the vector space of degree- d polynomials of $\mathbb{K}[x]$ under addition (plus the zero polynomial). If $\mathcal{X}$ is conic then

$$(2.9) \quad \mathcal{I}(\mathcal{X})_d := \{f \in \mathbb{K}[x]_d : f(x) = 0 \text{ for all } x \in \mathcal{X}\}$$

is a subspace of $\mathbb{K}[x]_d$. Next given a homogeneous degree- d polynomial $f \in \mathbb{K}[x]_d$, there is a unique tensor $T^f \in S^d(\mathbb{K}^n)$ such that $f(x) = \langle x^{\otimes d}, T^f \rangle$ where the inner product is defined for two tensors in (2.1).

OBSERVATION 2.4.4. *Notice that since T^f is symmetric, we also have that $f(x) = \langle \text{Sym}_d(x^{\otimes d}), T^f \rangle$ since the components of $x^{\otimes d}$ orthogonal to $S^d(\mathbb{K}^n)$ will not contribute to the inner product. It follows that we can evaluate and represent polynomials more efficiently using our symmetric lift constructions given in Chapter 4.*

In Chapter 3, we will need a polynomial basis for $\mathcal{I}(\mathcal{X}_1)_2$ where $\mathcal{X}_1$ is the variety of rank-at-most-1 $m \times n$ matrices. As a set we can define this via

$$\mathcal{X}_1 = \{u \otimes v : u \in \mathbb{K}^m, v \in \mathbb{K}^n\} \subseteq \mathbb{K}^{m \times n}.$$

Note that canonically the zero matrix is rank-0 which is one reason for the terminology "rank-at-most-1." The other is that it is well known that the variety of matrices of rank at most r are cut out by their $(r+1) \times (r+1)$ minors [38]. However, to ensure that rank is *exactly* r would require that at least one $r \times r$ minor is non-zero which is a semi-algebraic condition.

EXAMPLE 2.4.5 (The variety of rank-at-most-1 matrices). *For a matrix $X \in \mathbb{K}^{m \times n}$ the 2-minors of X are the determinants of each 2×2 submatrix formed by selecting the rows a and c and columns b and d for*

$a, c \in [m]$ and $b, d \in [n]$, i.e.

$$(2.10) \quad M_{abcd}(X) := X[a, b]X[c, d] - X[a, d]X[c, b].$$

Taking the standard basis vector $\vec{e}_{ab} \in \mathbb{K}^m \otimes \mathbb{K}^n$ defined so that $\langle \vec{e}_{ab}, X \rangle = X[a, b]$,

$$(2.11) \quad M_{abcd}(X) = \langle \vec{e}_{ab} \otimes \vec{e}_{cd} - \vec{e}_{ad} \otimes \vec{e}_{cb}, X^{\otimes 2} \rangle.$$

Note that $\vec{e}_{ab} \otimes \vec{e}_{cd} - \vec{e}_{ad} \otimes \vec{e}_{cb} \notin S^2(\mathbb{K}^{m \times n})$ and is hence not unique. However, the vector

$$(2.12) \quad \begin{aligned} E_{abcd} &:= \text{Sym}_2(\vec{e}_{ab} \otimes \vec{e}_{cd} - \vec{e}_{ad} \otimes \vec{e}_{cb}) \\ &= \frac{1}{2}(\vec{e}_{ab} \otimes \vec{e}_{cd} + \vec{e}_{cd} \otimes \vec{e}_{ab}) - \frac{1}{2}(\vec{e}_{ad} \otimes \vec{e}_{cb} + \vec{e}_{cb} \otimes \vec{e}_{ad}) \in S^2(\mathbb{K}^{m \times n}) \end{aligned}$$

obeys $\langle E_{abcd}, X^{\otimes 2} \rangle = 0$ if $X = u \otimes v$ for vectors $u \in \mathbb{K}^m$ and $v \in \mathbb{K}^n$. Hence, E_{abcd} is the unique element of $S^2(\mathbb{K}^{m \times n})$ which effects evaluation of the polynomial M_{abcd} against a symbolic matrix X .

Identifying M_{abcd} with E_{abcd} it is not hard to see the following.

LEMMA 2.4.6. *For the variety of rank-1 $m \times n$ matrices $\mathcal{X}_1 \subseteq \mathbb{K}^{m \times n}$, the set of vectors $\{E_{abcd} : 1 \leq a < c \leq m, 1 \leq b < d \leq n\}$ is a basis for $\mathcal{I}(\mathcal{X}_1)_2$ when considering it to be a subspace of $S^2(\mathbb{K}^{m \times n})$ by identifying monomials with tensor products of standard basis elements.*

PROOF. As previously stated minors of the form given by (2.10) generate the ideal $\mathcal{I}(\mathcal{X}_1)$. From this it follows that $\text{span}\{E_{abcd}\} = \mathcal{I}(\mathcal{X}_1)_2$. Finally, we need merely note that if $(a, b, c, d) \neq (a', b', c', d')$, then $\langle E_{abcd}, E_{a'b'c'd'} \rangle = 0$. It follows that the vectors are linearly independent since they are orthogonal. In fact, this is an orthonormal basis as $\|E_{abcd}\|_2^2 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$. $\square$

Next, a less known fact is that for any tensor $T \in \mathbb{K}^{n_1 \times \cdots \times n_k}$, degree-2 homogeneous polynomials of the form

$$(2.13) \quad T[j_1, j_2, \dots, j_\ell, \dots, j_k]T[j'_1, j'_2, \dots, j'_\ell, \dots, j'_k] - T[j_1, j_2, \dots, j'_\ell, \dots, j_k]T[j'_1, j'_2, \dots, j_\ell, \dots, j'_k],$$

where each $j_i, j'_i \in [n_i]$ for every $i \in [k]$ and $\ell \in [k]$, not only cut out a general Segre variety $\mathcal{X}_1 \subset \mathbb{K}^{n_1 \times \cdots \times n_k}$ but generate the ideal $\mathcal{I}(\mathcal{X}_1)$ [37]. It follows that the span of these polynomials produce $\mathcal{I}(\mathcal{X}_1)_2 \subseteq S^2(\mathbb{K}^{n_1 \times \cdots \times n_k})$. However, generating all polynomials of the form in (2.13) does not produce a linearly independent set in general.

EXAMPLE 2.4.7. Let T be a $2 \times 2 \times 2$ rank-1 tensor with no symmetries along modes. Observe that each of

$$a = T[1, 1, 1]T[2, 2, 2] - T[2, 1, 1]T[1, 2, 2],$$

$$b = T[2, 1, 1]T[1, 2, 2] - T[2, 2, 1]T[1, 1, 2]$$

$$c = T[1, 1, 1]T[2, 2, 2] - T[1, 2, 1]T[2, 1, 2] \text{ and}$$

$$d = T[1, 2, 1]T[2, 1, 2] - T[2, 2, 1]T[1, 1, 2]$$

fits the form given in (2.13) but $a + b = c + d$. So in general, polynomials of the form (2.13) are a spanning set of $\mathcal{I}(\mathcal{X}_1)_2$ but not a basis.

This example does given an insight into what goes wrong: there were two different ways to swap indices along modes to go from the pair $\{(1, 1, 1), (2, 2, 2)\}$ to the pair $\{(2, 2, 1), (1, 1, 2)\}$. One can also show that a similar issue arises when trying to deal with symmetry where using all polynomials of the form in (2.13) produces a linearly dependent set of polynomials. However, we wait to resolve both issues until Chapter 4 where we develop a better system to handle enforced symmetries in tensors.

Improving Bounds for Finding Rank-1 Matrices in a Subspace

This chapter is adapted from the arXiv preprint [24] with co-authors Dastidar and Wein. Common definitions and exposition were primarily written by Dastidar and Wein and have been lightly edited by this author to highlight my work in the project, specifically implementation of the JLV algorithm and adaptation to perform partial computer-proof which are covered in Section 3.4, as well as the extension symmetric matrices in Section 3.5 which prove relevant for Chapter 5’s work on the multi-reference alignment (MRA) models.

3.1. Introduction

In this chapter we consider the basic computational question of finding rank-1 matrices in a given linear subspace $\mathcal{U} \subseteq \mathbb{R}^{m \times n}$ which is based on joint work with Dastidar and Wein [24]. We assume $\mathcal{U}$ is generated as the linear span of s “planted” rank-1 matrices, along with $R - s$ other matrices, where $R = \dim(\mathcal{U})$ and $0 \leq s \leq R$. As input, we receive an arbitrary basis for $\mathcal{U}$, and the goal is to output the s planted matrices, up to permutation and scalar multiple (in the sense of Proposition 3.2.3). Aside from being a fundamental question in its own right, the $s = R$ case is a key subroutine in various tensor decomposition algorithms, as we will discuss in Section 3.1.2. In the $s = 0$ case there are no planted matrices to recover, so here the goal is to algorithmically generate a *certificate* that no rank-1 matrices exist in $\mathcal{U}$.

For a successful solution to this problem, we desire an algorithm that works for a wide range of possible inputs. To circumvent NP-hardness [21], we follow [43] and assume that the matrices generating $\mathcal{U}$ are chosen *generically*, either from the set of rank-1 matrices or the set of all matrices, as appropriate. The formal meaning of “generic” will be stated in Section 3.2, but essentially this means we ask for an algorithm that works for “almost all” inputs, in the sense that the set of inputs where the algorithm fails has measure zero. We desire an algorithm that succeeds, in this sense, for the largest possible range of parameters m, n, R, s . Furthermore, we desire an algorithm with fast runtime.

The work of Johnston–Lovitz–Vijayaraghavan [43] gives a polynomial-time algorithm for the above problem, under the condition $R \leq \frac{1}{4}(m-1)(n-1)$, and in fact for a more general setting where rank-1 matrices can be replaced by some other algebraic variety. We refer to this as the *JLV algorithm*, and it

will be a key focus of our work. The algorithm is based on the first level of the so-called *Nullstellensatz hierarchy* (and the connection to Hilbert’s Nullstellensatz is described in [43, Remark 12]). Higher levels of the hierarchy give rise to algorithms that are potentially more powerful at the expense of increased computational cost. Higher levels have been analyzed only in the fully non-planted case $s = 0$, where they are shown to work even for nearly full-dimensional subspaces $R \leq (1 - \epsilon)mn$ with runtime $(mn)^{f(\epsilon)}$ [28] (see Remark 6.10 therein).

Without constraints on runtime, the *identifiability* threshold is $R = (m - 1)(n - 1)$. By this we mean that for $0 \leq s \leq R \leq (m - 1)(n - 1)$, the only rank-1 matrices in the span of s generic rank-1 matrices along with $R - s$ generic matrices, will be the s planted matrices (and their scalar multiples) [34, Theorem 4.6.14]. Furthermore, the planted matrices can be recovered by a sufficiently high level of the Nullstellensatz hierarchy, but the runtime will not be polynomial in general [34, Remark 4.6.16]. Conversely, if $R > (m - 1)(n - 1)$ then the subspace will contain “spurious” rank-1 matrices in addition to the planted ones (see [38, Definition 18.1(iii)]). In fact, *any* subspace of dimension $R \geq (m - 1)(n - 1) + 2$ must contain infinitely many rank-1 matrices (see Definition 11.2 in [38] or Theorem 3 in [51]).

While we have focused our discussion on $\mathbb{R}$, our main results will also hold over $\mathbb{C}$. Although this does not affect generic inputs, the base field can matter: the matrices $((1, 0), (0, 1))$ and $((0, 1), (-1, 0))$ have a (complex) rank-1 matrix in their $\mathbb{C}$ -span but no rank-1 matrix in their $\mathbb{R}$ -span. We only work with infinite fields because the notion of “generic” will not make sense over finite fields.

3.1.1. Our Contributions. In this work, our objective is to investigate the precise limits of the JLV algorithm, that is, the first Nullstellensatz level. Our main question is: for large m, n , what is the largest constant $c \in [1/4, 1]$ such that the JLV algorithm succeeds on generic inputs with $R \leq (c - o(1))mn$? (We will see that there appears to be no dependence on s , to first order.) Even though higher levels of the hierarchy might tolerate larger R , the computational complexity quickly becomes impractical: the lowest level is $d = 2$ and in general, any constant level $d \geq 2$ needs to compute with subspaces of dimension $\Theta((mn)^d)$. For this reason, we feel it is valuable to understand the exact performance of the first level, especially given the large gap between existing bounds: $1/4 \leq c \leq 1$.

Towards this, we improve the bounds to $1/2 \leq c \leq 1/\sqrt{2}$. (In the special case $s = 0$, the bound $c \geq 1 - 1/\sqrt{2} \approx 0.293$ was known previously [28, Theorem 6.9].) In the arXiv version of this work, my co-authors give a combinatorial proof that for any $0 \leq s \leq R$, if $R \leq \frac{1}{2}(m - 1)(n - 2)$, the JLV algorithm succeeds on generic inputs verifying that $c \geq 1/2$, see [24, Corollary 2.6]. However, by a dimension counting

argument (see [24, Theorem 2.7]), when

$$\binom{R+1}{2} - s > \binom{m}{2} \binom{n}{2}$$

the JLV algorithm will not succeed, giving $c \leq 1/\sqrt{2}$.

Here I emphasize the numerical experiments and a method of computer-aided proof (Section 3.4) I developed which suggest that the true value is $c = 1/\sqrt{2}$ and that the inverse of [24, Theorem 2.7] also holds for generic subspaces. We highlight the open problem of proving this, which we state as Conjecture 3.4.1. The improvement from $c = 1/4 = 0.25$ to, potentially, $c = 1/\sqrt{2} \approx 0.71$ represents a much wider range of values for $\dim(\mathcal{U})$, as a fraction of the ambient dimension mn . Conjecture 3.4.1 posits an exact formula for the threshold, and we give an algorithm which will can give a computer-proof of this conjecture for any particular value of m, n, R and s values or give evidence against the conjecture. Our code is freely available, including an implementation of the JLV algorithm in Python.

3.1.2. Motivation and Applications. Aside from being a fundamental computational question in linear algebra, the problem of finding planted rank-1 matrices in a subspace is also important from a machine learning point of view, as it is a key subroutine in various tensor decomposition algorithms [43, 47].

To illustrate our contributions to tensor decomposition, consider the (CP) decomposition problem for an order-3 tensor $T \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ discussed in Section 2.2: given a tensor with a rank- R decomposition $T = \sum_{\ell=1}^R a^\ell \otimes b^\ell \otimes c^\ell$ where the components a^ℓ, b^ℓ, c^ℓ are generically chosen vectors from $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}, \mathbb{R}^{n_3}$ respectively, and $(a \otimes b \otimes c)_{ijk} := a_i b_j c_k$, the goal is to recover the list of components up to inherent ambiguities (permutation of $[R]$ and scalar multiplication). The order-4 tensor decomposition problem is analogous but for $T \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$.

Building on [43], our improved algorithmic result for finding rank-1 matrices implies improved results for polynomial-time tensor decomposition. Specifically, we can decompose generic order-4 $n \times n \times n \times n$ tensors up to rank $R \leq \frac{1}{2}(n-2)^2$ for any n , allowing roughly twice as many components as the previous result $R \leq \frac{1}{4}(n-1)^2$ [43]. Further my computer-proof of Conjecture 3.4.1 for some values gives the improvement $R \leq \frac{1}{\sqrt{2}}n(n-1) + \frac{1}{2}$ for $2 \leq n \leq 15$ and could in theory be run to attempt certification of this bound for higher values of n . For order-3 $n_1 \times n_2 \times n_3$ tensors that are unbalanced in the sense $n_3 \geq n_1 n_2 / 2$, we similarly obtain a factor-of-2 improvement from $R \leq \frac{1}{4}(n_1-1)(n_2-1)$ to $R \leq \frac{1}{2}(n_1-1)(n_2-2)$. These improvements follow directly by the ideas in [43, Section 6]. Specific cases where my program has certified Conjecture 3.4.1 gives analogous improvements, see Table 3.1 for details.

3.1.3. Related Work. The problem we study was first defined by JLV [43], but earlier precursors considered various special cases and variations. For instance, the case $s = R$ is closely tied with the celebrated *FOOBI* algorithm for order-4 tensor decomposition [25, 26]. In Section III of [26] it is claimed that this algorithm decomposes generic $n \times n \times n \times n$ tensors of rank up to $R \leq \frac{1}{2}n(n+1)$. However, Appendix A of [43] points out an error in the proof and provides a correct proof under the weaker condition $R \leq \frac{1}{4}(n-1)^2$. Our improvement to $R \leq \frac{1}{2}(n-2)^2$ discussed above recovers the original claim, to first order. Some extensions of FOOBI include [41, 52], but these do not improve the rank condition.

Another precursor to JLV is [19], which considers only the special case $s = 1$ but generalizes it in a few directions: the linear subspace can be replaced by an affine space, and the set of rank-1 matrices can be replaced by rank- r matrices or tensors.

JLV was first to consider the full range $0 \leq s \leq R$, and their result also allows the set of rank-1 matrices to be replaced by an arbitrary affine cone.

In the special case $s = 0$, some extensions to JLV have appeared. First, [16] gives an algorithm in the case of *smoothed* (rather than generic) inputs, which amounts to giving a detailed analysis of noise-robustness. Also, [28] analyzes higher levels of the Nullstellensatz hierarchy, showing that they succeed for nearly full-dimensional subspaces. Generalizing these results to $s > 0$ remains an interesting open problem.

There is a natural variation of our problem where $m = n$ and all the matrices are symmetric. The JLV algorithm also applies here, but the bounds on its success are different from the non-symmetric case [43]. We cover this situation in Section 3.5.

Our improvements to order-3 tensor decomposition apply to unbalanced tensors, e.g. of shape roughly $n \times n \times n^2$. For “square” $n \times n \times n$ tensors, the classical *simultaneous diagonalization* method can decompose generic tensors of rank up to $R \leq n$ [49]. Recent advances [44, 57] have culminated in an algorithm approaching rank $2n$ [47]. This method uses JLV as a subroutine, but the factor $1/4$ is not the bottleneck in the analysis, so our result does not imply an improvement here.

Other algebraic methods for tensor decomposition were proposed prior to JLV, including [30, 64] and references therein. To our knowledge, these have not been proven to succeed on generic inputs.

Acknowledgments. We are thankful to Ben Lovitz and Aravindan Vijayaraghavan for helpful discussions. In particular, Ben Lovitz pointed us to many relevant references, e.g., concerning the identifiability question.

3.2. Setting and Results

Let us now formalize the basic computational question of interest. Throughout, we will work over a field $\mathbb{F}$ which can be either $\mathbb{R}$ or $\mathbb{C}$. Throughout this section, $\mathcal{V} = \mathbb{F}^{m \times n}$ is a vector space over the field $\mathbb{F}$ and $\mathcal{U} \subseteq \mathcal{V}$ is an R -dimensional subspace. We always assume $m, n \geq 2$. A *rank-1* matrix $v \in \mathbb{F}^{m \times n}$ is one that can be expressed as $v = xy^\top$ for non-zero $x \in \mathbb{F}^m, y \in \mathbb{F}^n$. Equivalently, a matrix has rank at most 1 if and only if all of its 2×2 minors vanish. So the set of rank-at-most-1 matrices can be expressed as the algebraic variety

$$\mathcal{X}_1 = \{v \in \mathbb{F}^{m \times n} : f(v) = 0 \text{ for all } f \in P\} \subseteq \mathcal{V},$$

where

$$P = \{v_{ab}v_{cd} - v_{ad}v_{cb} : 1 \leq a < c \leq m, 1 \leq b < d \leq n\} \subseteq \mathbb{F}[v_{ij} : 1 \leq i \leq m, 1 \leq j \leq n].$$

We assume $\mathcal{U} = \text{span}\{v^1, \dots, v^s, v^{s+1}, \dots, v^R\}$, where we call $v^1, \dots, v^s \in \mathcal{X}_1$ the *planted* matrices (which will have rank 1) and $v^{s+1}, \dots, v^R \in \mathcal{V}$ the *non-planted* matrices (which will have rank greater than 1). The planted ones ($i \leq s$) can be expressed as $v^i = x^i(y^i)^\top$ for vectors $x^i \in \mathbb{F}^m, y^i \in \mathbb{F}^n$. For the non-planted ones ($i > s$), we write $v^i = z^i$ with $z^i \in \mathbb{F}^{m \times n}$. We will assume the underlying vectors x, y, z are chosen *generically*, in the following sense.

DEFINITION 3.2.1. *A property Q of vectors $w \in \mathbb{F}^k$ is generic (or “holds generically” or “holds for a generic choice of w ”), if there is a nonidentically-zero polynomial p on $\mathbb{F}^k$ such that all vectors $w \in \mathbb{F}^k$ with $p(w) \neq 0$ have the property Q .*

This is a standard notion from algebraic geometry and it is well known that a generic property holds for all but a measure-zero set of points w . In our case, w is the concatenation of all vectors x, y, z , and we will aim to prove that our algorithm succeeds generically. We refer to

$$\mathcal{U} = \text{span}\{v^1, \dots, v^s, v^{s+1}, \dots, v^R\} = \text{span}\{x^1(y^1)^\top, \dots, x^s(y^s)^\top, z^{s+1}, \dots, z^R\} \subseteq \mathbb{F}^{m \times n}$$

as a “*generically chosen planted subspace with parameters* $0 \leq s \leq R$.” Generically, such a subspace has dimension exactly R , as we see below.

LEMMA 3.2.2. *If $\mathcal{U} = \text{span}\{v^1, \dots, v^R\} \subseteq \mathbb{F}^{m \times n}$ is a generically chosen planted subspace with parameters $0 \leq s \leq R$ where $R \leq mn$, then the elements $v^1, \dots, v^R$ are linearly independent.*

This fact appears in [43] and holds more generally for an irreducible algebraic variety in place of the rank-1 matrices.

Before stating our main results, we need to unpack some details of the JLV algorithm.

Toward this, for vectors $u, u' \in \mathcal{V} = \mathbb{F}^{m \times n}$, we will denote the symmetrization of their tensor product as

$$u \vee u' := \text{Sym}_2(u \otimes u') = \frac{1}{2}u \otimes u' + \frac{1}{2}u' \otimes u.$$

For a subspace $\mathcal{U} \subseteq \mathcal{V}$, we denote the space of such tensors by $S^2(\mathcal{U}) = \text{span}\{u \vee u' : u, u' \in \mathcal{U}\}$. If $\mathcal{U}$ has basis $\{u^1, \dots, u^k\}$, then $S^2(\mathcal{U})$ has basis $\{u^i \vee u^j : 1 \leq i \leq j \leq k\}$ and thus has dimension $\binom{k+1}{2}$.

A crucial step in the analysis of JLV is to establish

$$(3.1) \quad S^2(\mathcal{U}) \cap \mathcal{I}(\mathcal{X}_1)_2^\perp = \text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\}$$

As discussed in Example 2.4.5, the vectors E_{abcd} , which we now write $E_{abcd} = \vec{e}_{ab} \vee \vec{e}_{cd} - \vec{e}_{ad} \vee \vec{e}_{cb}$, cut out $\mathcal{X}_1$ in that $\langle E_{abcd}, v^{\otimes 2} \rangle = 0$ if $v \in \mathcal{X}_1$. Further, vectors of this type may be used to form a basis for $\mathcal{I}(\mathcal{X}_1)_2$.

In our context then, the critical step in the analysis of JLV is to establish

$$(\text{Goal}) \quad S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp = \text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\},$$

where (a, b, c, d) ranges over $1 \leq a < c \leq m$ and $1 \leq b < d \leq n$. Observe that we have the containment “ $\supseteq$ ” for free since $(v^i)^{\otimes 2} = v^i \vee v^i \in S^2(\mathcal{U})$ and for $i \leq s$, $\langle (v^i)^{\otimes 2}, E_{abcd} \rangle = 0$ since $v^1, \dots, v^s \in \mathcal{X}_1$, so $(v^i)^{\otimes 2} \in \text{span}\{E_{abcd}\}^\perp$. Establishing the forward inclusion is the main challenge, which is the focus of constructions in Section 3.3. The reason (Goal) is helpful is because the left-hand side is a subspace that the algorithm can compute, which gives us access to the right-hand side. Given the right-hand side, JLV show how to extract the desired solutions $v^1, \dots, v^s$, up to permutation and scalar multiple. This is captured by the following fact.

PROPOSITION 3.2.3 ([43]). *There is a polynomial-time algorithm with the following guarantee. Given as input $\text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\}$ where $v^1, \dots, v^s \in \mathbb{F}^k$ are linearly independent vectors, the algorithm will output some $\hat{v}^1, \dots, \hat{v}^s \in \mathbb{F}^k$ for which there exists a permutation π of $[s]$ and non-zero scalars $\alpha_1, \dots, \alpha_s$ with $\hat{v}^i = \alpha_i v^{\pi(i)}$ for all $i \in [s]$.*

Given a basis $\{P_1, \dots, P_s\}$ for $\text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\}$, the procedure for extracting $v^1, \dots, v^s$ is Step 2 of Algorithm 1 in [43]. The procedure involves decomposing a particular tensor using the simultaneous

diagonalization algorithm, and its correctness is justified in the proof of Corollary 16 in [43]. See also Lemma 3.4 of [47].

In the case $s = 0$, if (Goal) holds, i.e., if the left-hand side is $\{0\}$, this constitutes a certificate that $\mathcal{U}$ contains no rank-1 matrices. This is captured by the following fact.

PROPOSITION 3.2.4 ([43], Observation 11). *If $S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp = \{0\}$ for $1 \leq a < c \leq m$ and $1 \leq b < d \leq n$, then $\mathcal{U} \cap \mathcal{X}_1 = \{0\}$.*

The proof is immediate from our discussion above: if there were a non-zero element $v \in \mathcal{U} \cap \mathcal{X}_1$ then $v^{\otimes 2}$ would belong to both $S^2(\mathcal{U})$ and $\text{span}\{E_{abcd}\}^\perp$.

We have now essentially described the JLV algorithm: given a basis $\{u^1, \dots, u^R\}$ for $\mathcal{U}$, it first computes the left-hand side of (Goal) — making use of the basis $\{u^i \vee u^j : 1 \leq i \leq j \leq R\}$ for $S^2(\mathcal{U})$ — and then applies either Proposition 3.2.3 ($s > 0$) or Proposition 3.2.4 ($s = 0$) as appropriate. We now introduce two results proved in the main arXiv paper.

THEOREM 3.2.5 (Positive result [24]). *If $\mathcal{U} = \text{span}\{v^1, \dots, v^R\} \subseteq \mathbb{F}^{m \times n}$ (recall $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$) is a generically chosen planted subspace with parameters $0 \leq s \leq R$ where $R \leq \frac{1}{2}(m-1)(n-2)$, then (Goal) holds and furthermore the JLV algorithm [43, Algorithm 1] succeeds in recovering $v^1, \dots, v^s$ up to permutation and scalar multiple (in the sense of Proposition 3.2.3).*

PROPOSITION 3.2.6 (Negative result [24]). *If $\mathcal{U} = \text{span}\{v^1, \dots, v^R\} \subseteq \mathbb{F}^{m \times n}$ is a generically chosen planted subspace with parameters $0 \leq s \leq R$ where*

$$(3.2) \quad \binom{m}{2} \binom{n}{2} < \binom{R+1}{2} - s,$$

then (Goal) fails to hold and furthermore the JLV algorithm either outputs “Fail” or outputs a list of matrices that is not equivalent to $v^1, \dots, v^s$ via permutation and scalar multiple.

The negative result rules out perfect recovery, but this does not exclude the possibility of finding a superset of matrices that includes all the planted ones. However, we do not think this possibility is realistic: JLV has no reason to succeed without (Goal) because it will attempt to run a tensor decomposition algorithm on a tensor that has no reason to have low enough rank. Numerical experiments in Section 3.4.4 reveal that indeed, JLV appears to output “junk” unrelated to the planted matrices whenever (Goal) fails.

3.3. Preliminaries and Perspective

Towards proving (Goal), we reformulate it via linear algebra and set up the necessary ingredients for our computer-proof method.

3.3.1. A Linear Algebraic View. We start by considering a matrix towards proving the containment “ $\subseteq$ ” in (Goal).

DEFINITION 3.3.1. *Define a matrix M where the columns are indexed by*

$$(i, j) \in \Omega := \{(i, j) : 1 \leq i < j \leq R \text{ or } s < i = j \leq R\}$$

and the rows are indexed by tuples (a, b, c, d) with $1 \leq a < c \leq m$, $1 \leq b < d \leq n$. We define the entries of M by

$$M_{abcd,ij} := \langle E_{abcd}, v^i \vee v^j \rangle.$$

We will explicitly compute these entries soon. We now have the following lemma that gives a concrete linear algebraic approach to prove (Goal).

LEMMA 3.3.2. *If M has full column rank, then (Goal) holds.*

While not needed for our work, we note that the converse also holds, i.e., (Goal) is equivalent to M having full column rank, as long as $v^1, \dots, v^R$ are linearly independent.

PROOF. We prove the contrapositive. Suppose (Goal) is not met, meaning

$$\text{span}\{v^i \vee v^j\} \cap \text{span}\{E_{abcd}\}^\perp \not\subseteq \text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\}.$$

Then there exist coefficients $\{\alpha_{ij} : 1 \leq i \leq j \leq R\}$ such that

$$(3.3) \quad \sum_{1 \leq i \leq j \leq R} \alpha_{ij} v^i \vee v^j \in \text{span}\{E_{abcd}\}^\perp \setminus \text{span}\{(v^i)^{\otimes 2} : 1 \leq i \leq s\}.$$

There must exist a pair $(i, j) \in \Omega$ such that $\alpha_{ij} \neq 0$, since otherwise the left-hand side of (3.3) would lie in $\text{span}\{(v^i)^{\otimes 2} : 1 \leq i \leq s\}$. So for all $1 \leq a < c \leq m$ and $1 \leq b < d \leq n$, we have

$$\begin{aligned} 0 &= \left\langle E_{abcd}, \sum_{1 \leq i \leq j \leq R} \alpha_{ij} v^i \vee v^j \right\rangle \\ &= \sum_{1 \leq i \leq s} \alpha_{ii} \langle E_{abcd}, v^i \vee v^i \rangle + \sum_{(i,j) \in \Omega} \alpha_{ij} \langle E_{abcd}, v^i \vee v^j \rangle \\ &= \sum_{(i,j) \in \Omega} \alpha_{ij} M_{abcd,ij}. \end{aligned}$$

This shows there exists a non-zero $\alpha = (\alpha_{ij})_{(i,j) \in \Omega}$ such that $M\alpha = 0$. $\square$

LEMMA 3.3.3. *The entries of M are given by*

$$M_{abcd,ij} = \frac{1}{2} \left(v_{ab}^i v_{cd}^j + v_{cd}^i v_{ab}^j - v_{ad}^i v_{cb}^j - v_{cb}^i v_{ad}^j \right)$$

for $1 \leq a < c \leq m$, $1 \leq b < d \leq n$, and $(i, j) \in \Omega$.

PROOF. Using $\langle a \otimes b, c \otimes d \rangle = \langle a, c \rangle \langle b, d \rangle$, compute $\langle a \vee b, c \vee d \rangle = \frac{1}{2} (\langle a, c \rangle \langle b, d \rangle + \langle a, d \rangle \langle b, c \rangle)$. Now

$$\begin{aligned} M_{abcd,ij} &= \langle E_{abcd}, v^i \vee v^j \rangle \\ &= \langle e_{ab} \vee e_{cd} - e_{ad} \vee e_{cb}, v^i \vee v^j \rangle \\ &= \frac{1}{2} \left(\langle e_{ab}, v^i \rangle \langle e_{cd}, v^j \rangle + \langle e_{cd}, v^i \rangle \langle e_{ab}, v^j \rangle - \langle e_{ad}, v^i \rangle \langle e_{cb}, v^j \rangle - \langle e_{cb}, v^i \rangle \langle e_{ad}, v^j \rangle \right) \\ &= \frac{1}{2} \left(v_{ab}^i v_{cd}^j + v_{cd}^i v_{ab}^j - v_{ad}^i v_{cb}^j - v_{cb}^i v_{ad}^j \right), \end{aligned}$$

as desired. $\square$

We have now expressed the entries of M in terms of the v vectors. Recall that the v vectors can in turn be expressed in terms of the underlying x, y, z vectors, namely:

$$v_{\alpha\beta}^\ell := \begin{cases} x_\alpha^\ell y_\beta^\ell & \text{if } 1 \leq \ell \leq s, \\ z_{\alpha\beta}^\ell & \text{if } s < \ell \leq R. \end{cases}$$

This allows us to view M as a matrix of symbolic polynomials in the variables x, y, z .

3.4. Conjecture and Supporting Results

3.4.1. Conjecture on the Exact Threshold. In Proposition 3.2.6, we gave a sufficient condition for the JLV algorithm to fail. Numerical tests suggest that the JLV algorithm works right up until this breaking point, leading to the following conjecture.

CONJECTURE 3.4.1. *If $\mathcal{U} = \text{span}\{v^1, \dots, v^R\} \subseteq \mathbb{R}^{m \times n}$ is a generically chosen planted subspace with parameters $0 \leq s \leq R$ where*

$$(3.4) \quad \binom{m}{2} \binom{n}{2} \geq \binom{R+1}{2} - s,$$

then (Goal) holds, and as a consequence the JLV algorithm succeeds.

The left-hand side of (3.4) counts the number of rows of the matrix M defined in the previous section, and the right-hand side counts the number of columns of M . Therefore, (3.4) is the condition under which M could plausibly have full column rank, which is equivalent to (Goal).

We work over the field $\mathbb{F} = \mathbb{R}$ in this section, but the same conclusions hold over $\mathbb{C}$: since M is a symbolic matrix of polynomials with real coefficients, it has the same rank over $\mathbb{C}$ as over $\mathbb{R}$.

We remark that this conjecture is known to hold in the special case $s = 0$ and $\min\{m, n\} = 2$ [28, Theorem 1.10]. In fact, that result shows that for *any* subspace that avoids the rank-1 matrices (not just a generic one), the Nullstellensatz hierarchy can certify this (which, recall, is the objective when $s = 0$) at level $\min\{m, n\}$ (which has polynomial time if $\min\{m, n\}$ is constant). We will provide computer-aided proof for a finite set of new cases, by very different techniques.

In Section 3.2, we introduced the concept of *genericity* and demonstrated via Lemma 3.3.2 that (Goal) will be a generic property of subspaces of a given type if it holds true. As such, we need only produce a single subspace for which the matrix M defined in Section 3.3 has full column rank. We will term such subspaces produced, *witness* subspaces or just *witnesses* since they “witness” the truth of the conjecture. One can certify a subspace is a witness by computing some maximal minor (i.e., a subdeterminant involving all columns) of the M matrix and verifying it is not zero. We call the value of such a non-zero maximal minor (and the rows of M to which it corresponds) a *certificate* for the proof. In exact arithmetic, producing such a single witness-certificate pair constitutes a computer-aided proof that (Goal) holds generically for a specific choice of parameters m, n, s, R (since this implies that some maximal minor is non-zero as a polynomial). Using this observation we have proven Conjecture 3.4.1 in 3703 specific cases.

THEOREM 3.4.2. *For all choices of $m, n \geq 2$ such that $\frac{1}{\sqrt{2}}mn \leq 60$, Conjecture 3.4.1 holds for every s and R satisfying the given inequality (3.4).*

THEOREM 3.4.3. *For all choices of $m, n \geq 2$ such that $\frac{1}{\sqrt{2}}mn \leq 120$, Conjecture 3.4.1 holds whenever $s \in \{0, R\}$ and the given inequality (3.4) holds.*

The values of m and n for which these theorems apply and additional relevant details appear in Table 3.1. We produce witnesses which are particular subspaces spanned by integer-valued basis vectors and certificates as described above, where the determinants are computed in modular arithmetic to verify non-singularity. Further details are given in Section 3.4.3. Witness-certificate pairs are recorded in our supplemental material.

In addition, we have implemented the JLV algorithm in floating-point arithmetic to verify that it can recover planted rank-1 matrices up to a small error attributable to floating-point precision loss. These tests run quicker than our method for producing rigorous certificates, allowing us to run larger problem sizes and give strong evidence for an additional 3289 cases of Conjecture 3.4.1. These data lead to the theorems below marked with an asterisk (*). Since computations were done in floating-point arithmetic with possible loss of precision, we have not rigorously proven that the M matrix is full-rank; however, we believe the experiments justify the truth of these statements beyond a reasonable doubt.

THEOREM* 3.4.4. *For all choices of $m, n \geq 2$ such that $\frac{1}{\sqrt{2}}mn \leq 80$, Conjecture 3.4.1 holds for every s and R satisfying the given inequality (3.4).*

THEOREM* 3.4.5. *For all choices of $m, n \geq 2$ such that $\frac{1}{\sqrt{2}}mn \leq 165$, Conjecture 3.4.1 holds whenever $s \in \{0, R\}$ and the given inequality (3.4) holds.*

These results are all summarized in Table 3.1, with a more in-depth discussion deferred to Section 3.4.4. We also confirm that for R and s beyond the inequality given in Conjecture 3.4.1, the JLV algorithm fails to recover any planted rank-1 matrices (or fails to certify there are no such matrices in the case $s = 0$). This too is discussed in Section 3.4.4.

3.4.2. Determining Boundary Cases. To describe the evidence we have for Conjecture 3.4.1, we first give methods for determining the maximum subspace dimension R given $m, n \geq 2$ and number of planted rank-1 matrices s where $0 \leq s \leq R$ for which the condition of the conjecture will hold.

$m \backslash n$	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	3	4	6	8	10	11	13	15	16	18	20	22	23	25
4	4	6	9	11	13	16	18	21	23	26	28	31	33	36
5	5	8	11	14	17	21	24	27	30	33	36	40	43	46
6	6	10	13	17	21	25	29	33	37	41	45	48	52	56
7	7	11	16	21	25	30	34	39	43	48	53	57	62	66
8	8	13	18	24	29	34	40	45	50	56	61	66	71	77
9	9	15	21	27	33	39	45	51	57	63	69	75	81	87
10	10	16	23	30	37	43	50	57	64	70	77	84	91	97
11	11	18	26	33	41	48	56	63	70	78	85	93	100	107
12	12	20	28	36	45	53	61	69	77	85	93	101	110	118
13	13	22	31	40	48	57	66	75	84	93	101	110	119	128
14	14	23	33	43	52	62	71	81	91	100	110	119	129	138
15	15	25	36	46	56	66	77	87	97	107	118	128	138	148
16	16	27	38	49	60	71	82	93	104	115	126	137	148	159
17	17	29	40	52	64	76	87	99	111	122	134	146	157	169
18	18	30	43	55	68	80	93	105	117	130	142	154	167	179
19	19	32	45	58	72	85	98	111	124	137	150	163	176	190
20	20	34	48	62	76	89	103	117	131	145	158	172	186	200
21	21	36	50	65	79	94	108	123	137	152	166	181	196	210
22	22	37	53	68	83	99	114	129	144	159	175	190	205	220
23	23	39	55	71	87	103	119	135	151	167	183	199	215	231
24	24	41	58	74	91	108	124	141	158	174	191	208	224	241
25	25	42	60	77	95	112	130	147	164	182	199	216	234	251
26	26	44	62	81	99	117	135	153	171	189	207	225	243	261
27	27	46	65	84	103	121	140	159	178	196	215	234	253	271
28	28	48	67	87	106	126	145	165	184	204	223	243	262	282
29	29	49	70	90	110	131	151	171	191	211	232	252	272	292
30	30	51	72	93	114	135	156	177	198	219	240	261	281	302
31	31	53	75	96	118	140	161	183	205	226	248	269	291	312
32	32	55	77	100	122	144	167	189	211	234	256	278	300	323
33	33	56	80	103	126	149	172	195	218	241	264	287	310	333
34	34	58	82	106	130	154	177	201	225	248	272	296	320	343
35	35	60	85	109	134	158	183	207	231	256	280	305	329	353
36	36	61	87	112	137	163	188	213	238	263	288	313	339	364
37	37	63	89	115	141	167	193	219	245	271	297	322	348	374
38	38	65	92	119	145	172	198	225	252	278	305	331	358	384
39	39	67	94	122	149	176	204	231	258	286	313	340	367	394
40	40	68	97	125	153	181	209	237	265	293	321	349	377	405

TABLE 3.1. The value of $R_{\max}(m, n)$ for selected m and n . Grayed entries are redundant since there is symmetry if we exchange the values of m and n . Cases for which we have proved Conjecture 3.4.1 for all $0 \leq s \leq R$ (Theorem 3.4.2) are marked in blue, and cases where we extended via strong numerical evidence (Theorem* 3.4.4) are marked in purple. Cases where we have proof in the special cases $s = 0$ and $s = R$ (Theorem 3.4.3) are marked in magenta (also including the blue and purple regions), and these likewise extend to orange via floating-point calculation (Theorem* 3.4.5). Not all cases in columns $n = 2$ through 5 are listed due to limited space.

DEFINITION 3.4.6. *The maximum value of R as a function of s (the number of planted rank-1 matrices) for which (3.4) holds for given parameters m and n is*

$$R_{\max}(m, n; s) := \left\lfloor \sqrt{\frac{1}{2}m(m-1)n(n-1) + \frac{1}{4} + 2s - \frac{1}{2}} \right\rfloor.$$

The requirement that $s \leq R$ implicitly determines the domain of $R_{\max}(m, n; s)$. Further, $R_{\max}(m, n; s)$ is an increasing function in s . Its minimum is then attained at $s = 0$ and its maximum is attained for the largest s in the domain of $R_{\max}(m, n; s)$.

DEFINITION 3.4.7. *The maximum value of $R_{\max}(m, n; s)$ on its domain is*

$$R_{\max}(m, n) := \max \{s \in \mathbb{Z}^+ : s \leq R_{\max}(m, n; s)\}.$$

The possible values for s are $\{0, 1, \dots, R_{\max}(m, n)\}$. The maximum value then occurs when $s = R$ and we can solve (3.4) to find the explicit expression

$$R_{\max}(m, n) = \left\lfloor \sqrt{\frac{1}{2}m(m-1)n(n-1) + \frac{1}{4} + \frac{1}{2}} \right\rfloor.$$

REMARK 3.4.8. *Observe that $R_{\max}(m, n)$ is at most one larger than $R_{\max}(m, n; 0)$. Since $R_{\max}(m, n; s)$ is increasing and integer-valued, it has the form*

$$R_{\max}(m, n; s) = \begin{cases} R_{\max}(m, n) - 1 & 0 \leq s < s^* \\ R_{\max}(m, n) & s^* \leq s \leq R_{\max}(m, n) \end{cases}, \quad s \in \{0, 1, 2, \dots, R_{\max}(m, n)\}$$

where $s^* \in \{0, 1, \dots, R_{\max}\}$ is some threshold depending on m and n . Hence, $R_{\max}(m, n; s)$ is almost a constant function; however, if we want to ensure that (3.4) truly a sharp condition, we must account for the maximum subspace dimension being slightly larger when there are enough planted rank-1 matrices.

In summary, for given m and n , the (s, R) pairs satisfying (3.4) are those with $s \in \{0, \dots, R_{\max}(m, n)\}$ and $R \in \{s, \dots, R_{\max}(m, n; s)\}$.

3.4.3. Computer-Aided Proofs. Computing a determinant in exact arithmetic may require impractical amounts of memory. Therefore, to construct our certificates, we will produce a subspace spanned by integer-valued basis vectors such that some maximal minor of the matrix M has non-zero determinant as a polynomial over the finite field $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$ for some prime number p . If the determinant is non-zero over $\mathbb{F}_p$ it is also non-zero over $\mathbb{R}$, hence our exact computations constitute proof of the conjecture for specific problem sizes.

3.4.3.1. *Methods.* Here we detail the constructions and computations we do to algorithmically prove Conjecture 3.4.1 for a finite set of cases. Notice that for planted matrices of the form $v^i = x^i \otimes y^i$, $i = 1, \dots, s$, a necessary condition for identifiability of the planted matrices is that each set of factors $\{x^i\}_{i=1}^s \subset \mathbb{R}^m$ and $\{y^i\}_{i=1}^s \subset \mathbb{R}^n$ must be pairwise linearly independent (i.e., no two vectors are scalar multiples). If, for example, x^i is a scalar multiple of x^j then $\{x^i \otimes y^i + \alpha(x^j \otimes y^j) : \alpha \in \mathbb{R}\}$ will be an infinite family of rank-1 matrices in the subspace $\mathcal{U}$. A set of vectors drawn from an absolutely continuous distribution on $\mathbb{R}^d$ will be pairwise linearly independent with probability one, but drawing vectors from $\mathbb{F}_p^d$ will potentially violate this with high probability, depending on p, d and the method of sampling. Over $\mathbb{F}_p^d$, there are exactly $(p^d - 1)/(p - 1)$ pairwise linearly independent vectors. We chose $p = 997$, which is large enough to ensure that pairwise linearly independent sets exist for all values of m, n, s that we consider.

To produce a set of pairwise linearly independent vectors over $\mathbb{F}_p^d$, we sample a vector uniformly at random from $\mathbb{F}_p^d$, but only add it to the set if it is not the zero vector or a multiple modulo p of a vector in the current set. The process is repeated until the set contains s elements. It follows that $\{x^i\}_{i=1}^s$ is uniformly distributed on $(\mathbb{F}_p^m)^s$ conditional on being pairwise linearly independent. The same procedure is used to generate $\{y^i\}_{i=1}^s$. The remaining matrices, $\{v^i\}_{i=s+1}^R$, we sample uniformly at random from $\mathbb{F}_p^{m \times n}$. We will see that this procedure successfully produces the witnesses we need, with high probability.

To determine whether the subspace constitutes a witness to the truth of Conjecture 3.4.1, we construct the matrix M given at the beginning of Section 3.3 with computations done over $\mathbb{F}_p$. To find a non-zero maximal minor if one exists over $\mathbb{F}_p$, we run Gauss–Jordan elimination in modular arithmetic on M with row pivoting. By recording the locations of rows after elimination and pivoting, we may report which rows of the M matrix correspond to the maximal minor. The product of terms on the diagonal modulo p after elimination gives the value of the minor modulo p and either certifies that the subspace $\mathcal{U} = \text{span}\{v_1, v_2, \dots, v_R\} \subset \mathbb{R}^{m \times n}$ produces an M matrix of full rank if the minor is non-zero, or fails to certify that M has full rank. If certification fails, we resample all matrix factors and matrices and proceed with a new subspace.

3.4.3.2. *Results.* We were able to produce pseudo-random witness subspaces on the first initialization (pseudo-random seed set to 0) in all 3703 cases using a modulus of $p = 997$. These certificates constitute proof of Theorem 3.4.2 and Theorem 3.4.3. We note that it appears to be possible to generate witness subspaces for any p provided there is a pairwise linearly independent set i.e., $(p^{\min\{m,n\}} - 1)/(p - 1) \geq s$, but such subspaces are more abundant for larger p .

3.4.4. Numerical Experiments. Over modular arithmetic we can certify that a subspace $\mathcal{U}$ produces a non-singular M matrix, implying success of the JLV algorithm on generic inputs; however, we also aim to

confirm numerically that the full JLV algorithm works correctly over $\mathbb{R}$ in recovering planted rank-1 matrices. To verify this, we ran a number of computations in floating-point arithmetic and report the results here. The routines employed to do linear algebra in floating-point arithmetic run much quicker than Gauss–Jordan elimination, so we verified numerically that the JLV algorithm succeeds even in situations where we have not produced a computer-aided proof of its correctness. This gives strong evidence for cases of the conjecture indicated in Table 3.1 in purple and orange.

In addition to these checks, to understand the JLV algorithm’s failure modes, we ran the algorithm on random subspaces *beyond* where (Goal) holds but where planted solutions are still identifiable (see Section 3.1). For this, we tested all matrix shapes (m, n) with $2 \leq m \leq n$ and $mn/\sqrt{2} < 60$. For each matrix shape (m, n) we tested cases with parameters s, R where $0 \leq s \leq R_{\max}(m, n) + 1$ and $R = R_{\max}(m, n, s) + 1$ as long as $R \leq (m - 1)(n - 1)$, i.e. one over the conjectured bound but still generically identifiable. These cases represent all choices of s, R that are the smallest that exceed the inequality (3.4) for given m and n .

3.4.4.1. *Methods.* For floating-point calculations we produce planted matrices $v^i = x^i \otimes y^i$, $i = 1, \dots, s$ where components x^i and y^i are generated as i.i.d. standard Gaussian random vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. The additional non-planted matrices v^i , $s < i \leq R$ are generated as i.i.d. standard Gaussian matrices in $\mathbb{R}^{m \times n}$.

To perform a test, we used SVD to compute an orthonormal basis for the random subspace, which is used as input to the JLV algorithm. An orthonormal basis is not required, but generally ensures that the rank-1 solutions are not basis elements and finding them cannot be done by inspecting individual basis elements.

As described in Proposition 3.2.3, the JLV algorithm outputs a set of recovered rank-1 matrices, $\{\hat{v}^i\}_{i=1}^s$. Our measure of success was a greedy matching error for the recovered planted matrices. Treating the recovered rank-1 matrices $\{\hat{v}^i\}_{i=1}^s$ and ground truth rank-1 matrices $\{v^i\}_{i=1}^s$ as vectors in $\mathbb{R}^{mn}$, perfect recovery would mean there is a matching via permutation $\pi : [s] \rightarrow [s]$ such that under the matching vectors are co-linear, i.e.,

$$\frac{|\langle v^i, \hat{v}^{\pi(i)} \rangle|}{\|v^i\|_2 \|\hat{v}^{\pi(i)}\|_2} = 1, \quad \text{for all } i \in [s].$$

To measure deviations from perfect recovery, we perform a greedy matching to find π and then report a worst-case matching error

$$(3.5) \quad w = 1 - \min_{i \in [s]} \frac{|\langle v^i, \hat{v}^{\pi(i)} \rangle|}{\|v^i\|_2 \|\hat{v}^{\pi(i)}\|_2},$$

which gives the largest deviation from co-linearity.

We additionally wanted to understand the failure modes of the JLV algorithm, above the conjectured bound. As noted after Proposition 3.2.3, extracting the planted rank-1 matrices $\hat{v}^1, \dots, \hat{v}^s$ is achieved using simultaneous diagonalization of a particular tensor. When the bound (3.4) does not hold, there is no reason to believe that this tensor will be simultaneously diagonalizable but we wondered if the failure of (Goal) could have a “perturbative” effect on this tensor or if it would be catastrophic to the recovery of the planted rank-1 matrices. To this end we modified the simultaneous diagonalization algorithm given in [43, Section 2.5] so that in Step 3 when a set of eigenvalues cannot be matched, instead of outputting ‘Fail’ we run a greedy matching procedure. Similarly, we replace Step 4 of their algorithm, which requires solving a linear system exactly, by a linear least-squares method. These modifications result in a simultaneous diagonalization algorithm robust to small perturbations of its inputs.

3.4.4.2. *Results.* In the 6992 cases we tested where (3.4) holds, we ran the JLV algorithm on randomly generated subspaces and it successfully recovered the planted matrices with worst-case matching error (as defined in (3.5)) no worse than 4.96×10^{-14} (or correctly determined that there were no planted matrices in the case $s = 0$). For the cases beyond where we have generated certificates using computations over finite fields, these computations give overwhelming evidence (Theorem*s 3.4.4 and 3.4.5) that Conjecture 3.4.1 still holds in these cases. These cases appear as cells colored purple and orange in Table 3.1.

In the 2481 cases we tested where (3.4) does not hold, we found that the JLV algorithm fails. For those cases with planted matrices ($s > 0$), the worst-case matching error was always at least 0.042, corresponding to an angular deviation of roughly 0.29 radians or 16.6 degrees. Furthermore, the *average* value (over the 2481 cases) of the worst-case matching error was 0.759. To understand how the JLV algorithm fails, recall that JLV computes the subspace intersection $S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp$, and we know by Proposition 3.2.6 that when (3.4) fails, this subspace must contain “spurious” vectors in addition to the desired solutions $(v^i)^{\otimes 2}$. We saw that these spurious vectors had a catastrophic effect on the subsequent application of simultaneous diagonalization. Furthermore, we found that

$$\dim(S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp) = \binom{R+1}{2} - \binom{m}{2} \binom{n}{2} \geq s + 1$$

for each case tested, meaning the bound (3.2) found in the proof of Proposition 3.2.6 held with *equality*. For the cases where $s = 0$, we know from Proposition 3.2.6 that $S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp$ must be non-empty, meaning JLV fails to certify, and our experiments confirmed this. Data and the programs written in Python can be found in the following [repository](https://github.com/lemniscate8/linear-intersect-varieties)¹ hosted on GitHub. See README.md for an explanation of the programs and records provided.

¹Full URL is: <https://github.com/lemniscate8/linear-intersect-varieties>

3.5. Extension to Symmetric Matrices

Our work above considers subspaces $\mathcal{U} \subseteq \mathbb{R}^{m \times n}$ of matrices; however, it is also natural to consider subspaces of *symmetric* matrices $\mathcal{U} \subseteq S^2(\mathbb{R}^m) \subset \mathbb{R}^{m \times m}$, and to search for planted symmetric rank-1 matrices in such a subspace. As one motivation, this problem is a subroutine in symmetric tensor decomposition, which arises when applying the method of moments.

First, we note that it is possible to define the space of symmetric rank-1 matrices in an identical manner to rank-1 matrices, as the algebraic variety cut out by the vanishing of all 2-by-2 minors, but where the ambient space is the space of symmetric matrices. The variety is then

$$\mathcal{X}_1^\vee = \{v \in S^2(\mathbb{R}^m) : f(v) = 0 \text{ for all } f \in P^\vee\} \subseteq S^2(\mathbb{R}^m),$$

where we identify the variables $v_{ij} \equiv v_{ji}$ for all $i, j \in [m]$ to define

$$P^\vee = \{v_{ab}v_{cd} - v_{ad}v_{cb} : 1 \leq a \leq b \leq m, 1 \leq c \leq d \leq m\}$$

with $P^\vee \subseteq \mathbb{R}[v_{ij} : 1 \leq i \leq j \leq m]$. Due to this identification, the set $P^\vee$ as written contains polynomials that are linearly dependent; the number of linearly independent polynomials for this variety is

$$\binom{\binom{m+1}{2} + 1}{2} - \binom{m+3}{4} = \frac{1}{12}(m+1)m^2(m-1).$$

Here, the first term is the dimension of the homogeneous degree-2 polynomials in variables v_{ab} , and the subtrahend is the dimension of $\text{span}\{v^{\otimes 2} : v \in \mathcal{X}_1^\vee\} \cong S^4(\mathbb{R}^m)$. See Section 4.4, for a more details on how to find a basis for $P^\vee$.

We construct a planted subspace as $\mathcal{U} = \text{span}\{v^1, \dots, v^s, v^{s+1}, \dots, v^R\} \subseteq S^2(\mathbb{R}^m)$ where $v^1, \dots, v^s \in \mathcal{X}_1^\vee$ are the *planted symmetric matrices* and $v^{s+1}, \dots, v^R \in S^2(\mathbb{R}^m)$ are the *non-planted symmetric matrices*. The planted symmetric matrices may be expressed as $v^i = x^i(x^i)^\top$ for vectors $x^i \in \mathbb{R}^m$ when $i \leq s$. Viewing the entries of the x vectors as symbolic variables, there are $s \cdot m$ variables needed to express the planted symmetric matrices. For the non-planted symmetric matrices, we take $v^i = z^i$ for symbolic z variables with symmetry enforced: $z_{ab}^i = z_{ba}^i$ for all $a, b \in [m]$. Accounting for symmetry, there are $(R-s)\binom{m+1}{2}$ z -variables in total. We concatenate the x and z variables into a single collection of symbolic variables, and assume these are chosen generically in the sense of Definition 3.2.1. We refer to

$$\mathcal{U} = \text{span}\{x^1(x^1)^\top, \dots, x^s(x^s)^\top, z^{s+1}, \dots, z^R\} \subseteq S^2(\mathbb{R}^m)$$

as a “generically chosen planted subspace over the symmetric matrices with parameters $0 \leq s \leq R$.” A generic subspace of this type is *not* a generic instance of our original (non-symmetric) problem, so our combinatorial methods used to prove Theorem 3.2.5 are not immediately transferable to symmetric subspaces.

The JLV algorithm can be applied to the setting of symmetric matrices, and there is a natural analogue of (Goal) that implies its success, namely

$$(\text{Goal}^\vee) \quad S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp = \text{span}\{(v^1)^{\otimes 2}, \dots, (v^s)^{\otimes 2}\},$$

where $1 \leq a < c \leq m$ and $1 \leq b < d \leq m$. From this, we can deduce a similar result to Proposition 3.2.6, a negative result for subspaces over the symmetric matrices.

PROPOSITION 3.5.1. *If $\mathcal{U} = \text{span}\{v^1, \dots, v^R\} \subseteq \mathbb{R}^{m \times m}$ is a generically chosen planted subspace over the symmetric matrices with parameters $0 \leq s \leq R$ where $\frac{1}{12}(m+1)m^2(m-1) < \binom{R+1}{2} - s$, then $(\text{Goal}^\vee)$ fails to hold.*

Since $s \leq R$, the condition above is implied by $\frac{1}{\sqrt{6}}m^2 + 1 < R$. The proof of Proposition 3.5.1 is identical in form to the proof of Theorem 3.2.6 but replacing (Goal) with $(\text{Goal}^\vee)$ and the relevant dimensions with those appropriate to the symmetric case. We can produce a similar bound

$$(3.6) \quad \dim(S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^\perp) \geq \binom{R+1}{2} - \frac{1}{12}(m+1)m^2(m-1).$$

We now introduce a natural conjecture following similar principles to Conjecture 3.4.1.

CONJECTURE 3.5.2. *If $\mathcal{U} \subseteq S^2(\mathbb{R}^m)$ is a generically chosen planted subspace over the symmetric matrices with parameters $0 \leq s \leq R$, where s and R satisfy*

$$(3.7) \quad \frac{1}{12}(m+1)m^2(m-1) \geq \binom{R+1}{2} - s,$$

then $(\text{Goal}^\vee)$ holds, and as a consequence the JLV algorithm succeeds.

For large m , condition (3.7) gives $R \sim \frac{1}{\sqrt{6}}m^2$. To compare, Theorem 2 in [43] guarantees success of the JLV algorithm when $R \leq \frac{1}{6}m(m-1)$. Hence, proof of Conjecture 3.5.2 would indicate that the JLV algorithm succeeds for subspaces considerably larger than previously shown (by a factor of roughly 2.5).

While we do not have a proof of this conjecture, we have analogous results to Theorems 3.4.2 and 3.4.3 that prove the conjecture in specific cases.

m	2	3	4	5	6	7	8	9	10	11	12	13	14
$R_{\max}^{\vee}$	2	4	6	10	15	20	26	33	41	49	59	69	80
m	15	16	17	18	19	20	21	22	23	24	25	26	27
$R_{\max}^{\vee}$	92	104	118	132	147	163	180	197	216	235	255	276	297

TABLE 3.2. The value of $R_{\max}^{\vee}(m)$ for selected m . Cases for which we have proof of Conjecture 3.5.2 for all $0 \leq s \leq R$ are shown in blue and in magenta are cases for which the conjecture is proved only in the specific cases of $s = 0$ and $s = R$, the latter having implications for symmetric tensor decomposition.

THEOREM 3.5.3. *For all $2 \leq m \leq 14$, Conjecture 3.5.2 holds for every s and R satisfying inequality (3.7).*

THEOREM 3.5.4. *For all $2 \leq m \leq 21$, Conjecture 3.5.2 holds when $s = 0$ or $s = R$ and satisfy inequality (3.7).*

Similar to Definition 3.4.6 we give the boundaries for subspace dimension depending on the proportion of planted rank-1 matrices.

DEFINITION 3.5.5. *The maximum subspace dimension R that fits the conditions of Conjecture 3.5.2 is*

$$R_{\max}^{\vee}(m; s) := \left\lfloor \sqrt{\frac{1}{6}(m+1)m^2(m-1) + \frac{1}{4} + 2s - \frac{1}{2}} \right\rfloor.$$

Similarly, we define the maximum of possible value for $R_{\max}^{\vee}(m; s)$ any choice of s .

DEFINITION 3.5.6.

$$R_{\max}^{\vee}(m) := \max \{s \in \mathbb{Z}^+ : s \leq R_{\max}^{\vee}(m; s)\}.$$

Further, by similar reasoning to that in Section 3.4.2, it has the following form:

$$R_{\max}^{\vee}(m) = \left\lfloor \sqrt{\frac{1}{6}(m+1)m^2(m-1) + \frac{1}{4} + \frac{1}{2}} \right\rfloor.$$

3.5.0.1. *Certificates and Numerical Experiments.* We ran similar computations over a finite field to prove Conjecture 3.5.2 for a range of matrix sizes. The main difference was that planted matrices took the form $v^i = x^i \otimes x^i$ for $i \in [s]$. The set $\{x^i\}_{i=1}^s$ was uniform on $(\mathbb{F}_p^m)^s$ conditional on being pairwise linearly independent for exact computations, and $\{x^i\}_{i=1}^s$ were independent standard Gaussian vectors in $\mathbb{R}^m$ for floating-point computations. To produce additional basis elements from the space of symmetric matrices we selected entries in the upper triangular portion of each v^i , $i = s+1, \dots, R$ uniformly from $\mathbb{F}_p$ for exact computations, and i.i.d. Gaussian for floating-point computations. We enforced the required symmetry by assigning entries below the diagonal to the value of their counterparts above the diagonal.

Findings were very similar to computations done in the previous sections. We produced certificates and proved Conjecture 3.5.2 in all 441 cases we attempted (see Table 3.2). Additional numerical tests verified that it was possible to recover the random planted symmetric rank-1 matrices via the JLV algorithm with the worst of the worst-case matching error being 1.89×10^{-15} (i.e. negligible error).

We also ran 414 tests over the bound (3.7), from $m = 2$ to $m = 14$ with parameters $0 \leq s \leq R_{\max}^{\vee}(m) + 1$ and $R = R_{\max}^{\vee}(m) + 1$ provided that $R \leq \binom{m+1}{2} - m$, i.e. dimension is one over the conjectured bound in (3.7) but small enough that planted matrices will be identifiable. Among these tests, the worst-case matching error was always at least 0.078 or an angular deviation of roughly 0.40 radians, indicating recovery was generally unsuccessful when we planted matrices ($s > 0$). Furthermore, the *average* value of the worst-case matching error was 0.776. In addition, we found that

$$\dim(S^2(\mathcal{U}) \cap \text{span}\{E_{abcd}\}^{\perp}) = \binom{R+1}{2} - \frac{1}{12}(m+1)m^2(m-1) \geq s+1$$

for every case tested, suggesting the bound in (3.6) holds with equality for generic subspaces.

3.6. Future Directions

A clear direction for future study to prove both Conjecture 3.4.1 and Conjecture 3.5.2 generally though we note that the combinatorial techniques used to prove Proposition 3.2.5 seem to have reached their limit in verifying $c \geq 1/2$. We also only investigated the JLV algorithm for planted rank-1 matrices and rank-1 symmetric matrices. One could also use a similar dimension counting bound to conjecture a sharp threshold for cutting out third-order tensors or indeed any Segre-Veronese variety (variety of rank-1 tensors with partial symmetries). More challenging would be tightening the subspace dimension bound [43, Theorem 2] further up the Nullstellensatz hierarchy for rank-1 (symmetric) matrices.

Open question: can dimension counting predict sharp bounds on the success of the JLV algorithm when cutting out higher order rank-1 tensors? What about higher up the Nullstellensatz hierarchy?

Tensor Formats, Symmetrized Kronecker Products and Symmetric Lifts

In this Chapter I introduce some of the core concepts developed to support work in Chapter 5. First, I define a symmetrized version of the Kronecker product and prove that it obeys retains many desirable properties had by the usual Kronecker product. I defined symmetrized Kronecker products of vectors and matrices in a uniform way and which avoids storing redundant entries. The product also obeys a symmetrized form of the mixed product rule which aids in analyzing linear functions acting on spaces of symmetric tensors and determining rank of these functions. I also show that with respect to the Deshowitz-Manna ordering on multisets, the symmetric product of upper-triangular matrices remain upper-triangular and that under certain conditions the symmetric product of matrices with orthogonal (or unitary) columns produces a matrix with orthogonal (or unitary) columns. This material is taken from the arXiv preprint [67] and was work done by this author under the guidance of Alexander Wein. See Section 1.2 for further details. Finally, I recast some derivations from Chapter 2 with a better handle on symmetries and find a basis for the degree-2 homogeneous polynomials that cut out rank-1 tensors with arbitrary symmetries across modes which appears as part of my implementation of the JLV algorithm in Python.

4.1. Symmetric Tensors in an Irredundant Format

A symmetric tensor in $S^k(\mathbb{K}^n) \subset (\mathbb{K}^n)^{\otimes k}$ (or the symmetrization of a vector with tensor product structure) may have many redundant entries. We will describe such tensors and linear maps from $S^k(\mathbb{K}^n)$ to another symmetric space in a concise manner by specifying a canonical basis for $S^k(\mathbb{K}^n)$ and identifying elements by their coordinate vectors in this subspace. Since the indices of a k -order symmetric tensor is invariant under the action of S_k we can think of entries such a tensor as being indexed by multisets of k elements. We will disambiguate between sets and multisets reserving the traditional curly braces for sets, e.g. $\{1, 2, 3\}$ is a set, and using the “bag” notation for multisets, e.g. $\{1, 2, 2\}$ is a multiset. We will use the notation $\left(\!\!\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\!\!\right)$ to mean “ n multi-choose k ” where

$$\left(\!\!\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\!\!\right) = \binom{n+k-1}{k}$$

and counts the number of k -element multisets of $[n]$. For a set S , use $\binom{S}{k}$ to denote the set of k -element multi-sets of S . Finally, for a k -element multiset K , we will use $\text{multi}(K)$ to indicate the multiset of multiplicities of the elements in K . For example, $\text{multi}(\{1, 1, 2, 2, 3, 4, 4, 4\}) = \{1, 2, 2, 3\}$. In conjunction with the notation that

$$\binom{n}{m_1, m_2, \dots, m_k} = \frac{n!}{m_1! m_2! \dots m_k!}$$

we may count the number of distinct arrangements of a multiset K into a $|K|$ -tuple via $\binom{|K|}{\text{multi}(K)}$.

For each multiset $\{j_1, j_2, \dots, j_k\}$ of $[n]$, define the vector

$$(4.1) \quad \vec{s}_{j_1 j_2 \dots j_k} := \frac{w_{\{j_1, j_2, \dots, j_k\}}}{k!} \sum_{\sigma \in S_k} \vec{e}_{j_{\sigma(1)}} \otimes \vec{e}_{j_{\sigma(2)}} \otimes \dots \otimes \vec{e}_{j_{\sigma(k)}} \in S^k(\mathbb{K}^n)$$

where

$$(4.2) \quad w_{\{j_1, j_2, \dots, j_k\}} := \sqrt{\binom{k}{\text{multi}(\{j_1, j_2, \dots, j_k\})}}.$$

LEMMA 4.1.1. *The set $\mathcal{S}_{k,n} = \{\vec{s}_{j_1 j_2 \dots j_k} : 1 \leq j_1 \leq \dots \leq j_k \leq n\}$ is an orthonormal basis for $S^k(\mathbb{K}^n)$.*

PROOF. Since there are $\binom{n}{k}$ -many such vectors in $\mathcal{S}_{k,n}$ and $\dim S^k(\mathbb{K}^n) = \binom{n}{k}$, verifying the set is orthonormal imply they are indeed an orthonormal basis for $S^k(\mathbb{K}^n)$. Using the fact that $\langle u^1 \otimes \dots \otimes u^k, v^1 \otimes \dots \otimes v^k \rangle = \prod_{j=1}^k \langle u^j, v^j \rangle$, we may write

$$\langle \vec{s}_{j_1 \dots j_k}, \vec{s}_{j'_1 \dots j'_k} \rangle = \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \prod_{i=1}^k \langle \vec{e}_{j_{\sigma(i)}}, \vec{e}_{j'_{\tau(i)}} \rangle.$$

We may apply a form of re-indexing where we write $\tau = \rho \circ \sigma$ for a unique $\rho \in S_k$ in the inner sum. Reordering the product by σ^{-1} eliminates the dependence on σ so that

$$\langle \vec{s}_{j_1 \dots j_k}, \vec{s}_{j'_1 \dots j'_k} \rangle = \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\rho \in S_k} \prod_{i=1}^k \langle \vec{e}_{j_i}, \vec{e}_{j'_{\rho(i)}} \rangle = \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{k!} \sum_{\rho \in S_k} \prod_{i=1}^k \langle \vec{e}_{j_i}, \vec{e}_{j'_{\rho(i)}} \rangle.$$

For this expression to be non-zero, each inner product must be non-zero. This requires there be a permutation sending the elements of $(j_1, \dots, j_k)$ to $(j'_1, \dots, j'_k)$. This only holds if $\{j_1, \dots, j_k\} = \{j'_1, \dots, j'_k\}$. So when $\{j_1, \dots, j_k\} \neq \{j'_1, \dots, j'_k\}$, $\langle \vec{s}_{j_1 \dots j_k}, \vec{s}_{j'_1 \dots j'_k} \rangle = 0$. Next,

$$\|\vec{s}_{j_1 \dots j_k}\|^2 = \frac{w_{\{j_1, \dots, j_k\}}^2}{k!} \sum_{\rho \in S_k} \prod_{i=1}^k \langle \vec{e}_{j_i}, \vec{e}_{j_{\rho(i)}} \rangle.$$

The sum then counts the number of permutations that keep the sequence $j_1, \dots, j_k$ the same. Since $\frac{w_{\{j_1, \dots, j_k\}}^2}{k!}$ is the reciprocal of this quantity (see (4.2)), $\|\vec{s}_{j_1, \dots, j_k}\|^2 = 1$. $\square$

We will define $(\mathbb{K}^n)^{\otimes k}$ to be the space of coordinate vectors of $S^k(\mathbb{K}^n)$ in the basis $\mathcal{S}_{k,n}$ i.e. the vector of coefficients when writing a symmetric tensor as a sum of basis elements. A vector in $(\mathbb{K}^n)^{\otimes k}$ then has $\binom{n}{k}$ entries which we index using k -element multisets of $[n]$ identifying standard basis vectors by the notation $\vec{e}_{\{j_1, \dots, j_k\}}$, for $j_1, \dots, j_k \in [n]$. One can then identify this space with $\mathbb{K}^{\binom{n}{k}}$ by forgetting our special indexing structure and sending $\vec{e}_M$ to the vector $\vec{e}_j$ where j is the position of M in the some total order on k -element multisets of $[n]$. The symbol “ $\otimes$ ” will be used to indicate these symmetrized products.

DEFINITION 4.1.2. *We define a symmetrized Kronecker product of vectors $u^1, u^2, \dots, u^k \in \mathbb{K}^n$, denoted $u^1 \otimes u^2 \otimes \dots \otimes u^k \in (\mathbb{K}^n)^{\otimes k}$, as the coordinate-vector of $\text{Sym}_k(u^1 \otimes u^2 \otimes \dots \otimes u^k)$ in $\mathcal{S}_{n,k}$. Entry-wise,*

$$(4.3) \quad (u^1 \otimes u^2 \otimes \dots \otimes u^k)[\{j_1, j_2, \dots, j_k\}] := \frac{w_{\{j_1, j_2, \dots, j_k\}}}{k!} \sum_{\sigma \in S_k} u^1[j_{\sigma(1)}] u^2[j_{\sigma(2)}] \dots u^k[j_{\sigma(k)}].$$

We may express the standard basis on $(\mathbb{K}^n)^{\otimes k}$ using the standard basis in $\mathbb{K}^n$ and our symmetrized product as follows

$$(4.4) \quad \vec{e}_{\{j_1, \dots, j_k\}} := w_{\{j_1, \dots, j_k\}}(\vec{e}_{j_1} \otimes \dots \otimes \vec{e}_{j_k}).$$

If $x, y \in (\mathbb{K}^n)^{\otimes k}$ then

$$\langle x, y \rangle = \sum_{1 \leq j_1 \leq \dots \leq j_k \leq n} x[\{j_1, \dots, j_k\}] y[\{j_1, \dots, j_k\}]$$

gives the inner product on $(\mathbb{K}^n)^{\otimes k}$.

The primary advantage of working over $(\mathbb{K}^n)^{\otimes k}$ is that the storing symmetric tensors as elements in $(\mathbb{K}^n)^{\otimes k}$ requires an array of size n^k while de-duplicating entries in this manner requires $\binom{n}{k} \approx \frac{1}{k!} n^k$ entries which represents significant less storage for higher order tensors. Further, in this format, symmetry cannot be degraded by issues of numerical precision since we do not have duplicated entries which can accumulate different errors.

LEMMA 4.1.3. *For vectors $u^1, \dots, u^k, v^1, \dots, v^k \in \mathbb{K}^n$, Sym_k as given in (2.5) and our inner product on $(\mathbb{K}^n)^{\otimes k}$, our symmetrized Kronecker product satisfies*

$$\langle u^1 \otimes u^2 \otimes \dots \otimes u^k, v^1 \otimes v^2 \otimes \dots \otimes v^k \rangle = \langle \text{Sym}_k(u^1 \otimes u^2 \otimes \dots \otimes u^k), \text{Sym}_k(v^1 \otimes v^2 \otimes \dots \otimes v^k) \rangle.$$

PROOF. We use Definition 4.1.2 for the symmetrized Kronecker product to expand,

$$\langle u^1 \otimes \dots \otimes u^k, v^1 \otimes \dots \otimes v^k \rangle = \sum_{1 \leq j_1 \leq \dots \leq j_k \leq n} \frac{w_{\{j_1, \dots, j_k\}}^2}{(k!)^2} \left(\sum_{\sigma \in S_k} u^1[j_{\sigma(1)}] \dots u^k[j_{\sigma(k)}] \right) \left(\sum_{\sigma \in S_k} v^1[j_{\sigma(1)}] \dots v^k[j_{\sigma(k)}] \right).$$

Next, notice that $w_{\{j_1, \dots, j_k\}}^2$ counts the number of unique permutations of the multiset $\{j_1, \dots, j_k\}$. We could instead write the sum indexed by all k -tuples in $[n]^k$ giving

$$\sum_{j_1, j_2, \dots, j_k \in [n]} \overline{\left(\frac{1}{k!} \sum_{\sigma \in S_k} u^1[j_{\sigma(1)}] \cdots u^k[j_{\sigma(k)}] \right)} \left(\frac{1}{k!} \sum_{\sigma \in S_k} v^1[j_{\sigma(1)}] \cdots v^k[j_{\sigma(k)}] \right).$$

Finally, we may rewrite the two sums as entries of Kronecker products

$$\sum_{j_1, j_2, \dots, j_k \in [n]} \overline{\left(\frac{1}{k!} \sum_{\sigma \in S_k} (u^1 \otimes \cdots \otimes u^k)[(j_{\sigma(1)}, \dots, j_{\sigma(k)})] \right)} \left(\frac{1}{k!} \sum_{\sigma \in S_k} (v^1 \otimes \cdots \otimes v^k)[(j_{\sigma(1)}, \dots, j_{\sigma(k)})] \right),$$

which then is the inner product of symmetric projections,

$$\langle \text{Sym}_k(u^1 \otimes u^2 \otimes \cdots \otimes u^k), \text{Sym}_k(v^1 \otimes v^2 \otimes \cdots \otimes v^k) \rangle. \quad \square$$

COROLLARY 4.1.4. *For vectors $u^1, \dots, u^k, v^1, \dots, v^k \in \mathbb{K}^n$, and the inner product on $(\mathbb{K}^n)^{\otimes k}$,*

$$\langle u^1 \otimes u^2 \otimes \cdots \otimes u^k, v^1 \otimes v^2 \otimes \cdots \otimes v^k \rangle = \frac{1}{k!} \sum_{\sigma \in S_k} \prod_{i=1}^k \langle u^i, v^{\sigma(i)} \rangle.$$

PROOF SKETCH. If one applies the Lemma 4.1.3 and uses the same simplifications seen in the proof Lemma 4.1.1 this formula is not hard to check. $\square$

4.2. Linear Maps between Symmetric Tensor Spaces

We may now identify linear operators $A : S^k(\mathbb{K}^m) \rightarrow S^k(\mathbb{K}^n)$ by identifying A as an $\binom{m}{k} \times \binom{n}{k}$ matrix with entries in K . We can construct such operators by defining a symmetric product on matrices.

DEFINITION 4.2.1 (Symmetrized Kronecker products). *For matrices $A^1, A^2, \dots, A^k \in \mathbb{K}^{m \times n}$ define a fully symmetrized Kronecker product of these matrices, denoted $A^1 \circledast A^2 \circledast \cdots \circledast A^k \in \mathbb{K}^{\binom{m}{k} \times \binom{n}{k}}$, entry-wise by*

$$(4.5) \quad (A^1 \circledast A^2 \circledast \cdots \circledast A^k)[\{j_1, j_2, \dots, j_k\}, \{j'_1, j'_2, \dots, j'_k\}] := \frac{w_{\{j_1, j_2, \dots, j_k\}} w_{\{j'_1, j'_2, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \prod_{i=1}^k A^i[j_{\sigma(i)}, j'_{\tau(i)}],$$

where $\{j_1, j_2, \dots, j_k\}$ is a k -element multiset of $[m]$ and $\{j'_1, j'_2, \dots, j'_k\}$ is a k -element multiset of $[n]$.

LEMMA 4.2.2. *The fully symmetrized Kronecker product may be equivalently defined, using Definition 4.1.2 for the symmetric Kronecker product of vectors, either row-wise by*

$$(4.6) \quad (A^1 \circledast A^2 \circledast \cdots \circledast A^k)[\{j_1, j_2, \dots, j_k\}, :] = \frac{w_{\{j_1, j_2, \dots, j_k\}}}{k!} \sum_{\sigma \in S_k} A^1[j_{\sigma(1)}, :] \circledast A^2[j_{\sigma(2)}, :] \circledast \cdots \circledast A^k[j_{\sigma(k)}, :]$$

for any $\{j_1, j_2, \dots, j_k\}$ a k -element multiset of $[m]$, or column-wise by

$$(4.7) \quad (A^1 \otimes A^2 \otimes \dots \otimes A^k)[\cdot, \{j'_1, j'_2, \dots, j'_k\}] = \frac{w_{\{j'_1, j'_2, \dots, j'_k\}}}{k!} \sum_{\tau \in S_k} A^1[\cdot, j'_{\tau(1)}] \otimes A^2[\cdot, j'_{\tau(2)}] \otimes \dots \otimes A^k[\cdot, j'_{\tau(k)}]$$

for any $\{j'_1, j'_2, \dots, j'_k\}$ a k -element multiset of $[n]$.

PROOF. We will expand (4.6) and (4.7) to show that produces the same entry-wise expression as Definition 4.2.1.

Indexing into the summand of (4.6) — reproduced here:

$$(A^1 \otimes A^2 \otimes \dots \otimes A^k)[\{j_1, j_2, \dots, j_k\}, \cdot] = \frac{w_{\{j_1, \dots, j_k\}}}{k!} \sum_{\sigma \in S_k} A^1[j_{\sigma(1)}, \cdot] \otimes A^2[j_{\sigma(2)}, \cdot] \otimes \dots \otimes A^k[j_{\sigma(k)}, \cdot],$$

— by a multiset $\{j'_1, j'_2, \dots, j'_k\}$ of $[n]$ and expanding by Definition (4.3) gives

$$(A^1[j_{\sigma(1)}, \cdot] \otimes A^2[j_{\sigma(2)}, \cdot] \otimes \dots \otimes A^k[j_{\sigma(k)}, \cdot])[\{j'_1, j'_2, \dots, j'_k\}] = \frac{w_{\{j'_1, \dots, j'_k\}}}{k!} \sum_{\tau \in S_k} \prod_{i=1}^K A^i[j_{\sigma(i)}, j'_{\tau(i)}].$$

Substitution into (4.6) and rearranging sums and terms yields

$$(A^1 \otimes A^2 \otimes \dots \otimes A^k)[\{j_1, j_2, \dots, j_k\}, \{j'_1, j'_2, \dots, j'_k\}] = \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \prod_{i=1}^k A^i[j_{\sigma(i)}, j'_{\tau(i)}]$$

which is exactly Definition 4.2.1.

Next, indexing into the summand of the (4.7) — reproduced here:

$$(4.8) \quad (A^1 \otimes A^2 \otimes \dots \otimes A^k)[\cdot, \{j'_1, j'_2, \dots, j'_k\}] = \frac{w_{\{j'_1, \dots, j'_k\}}}{k!} \sum_{\tau \in S_k} A^1[\cdot, j'_{\tau(1)}] \otimes A^2[\cdot, j'_{\tau(2)}] \otimes \dots \otimes A^k[\cdot, j'_{\tau(k)}]$$

— by a multiset $\{j_1, j_2, \dots, j_k\}$ of $[m]$ and expanding by Definition (4.3) is essentially the same process and gives

$$(A^1[\cdot, \sigma(j'_1)] \otimes A^2[\cdot, \sigma(j'_2)] \otimes \dots \otimes A^k[\cdot, \sigma(j'_k)])[\{j_1, j_2, \dots, j_k\}] = \frac{w_{\{j_1, \dots, j_k\}}}{k!} \sum_{\tau \in S_k} \prod_{i=1}^K A^i[j_{\sigma(i)}, j'_{\tau(i)}].$$

Substitution into (4.7), rearranging sums and terms also yields

$$(A^1 \otimes A^2 \otimes \dots \otimes A^k)[\{j_1, j_2, \dots, j_k\}, \{j'_1, j'_2, \dots, j'_k\}] = \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \prod_{i=1}^k A^i[j_{\sigma(i)}, j'_{\tau(i)}]. \quad \square$$

If $A^1 = A^2 = \dots = A^k = A$, then we may call $A^{\otimes k}$ the k -th order symmetric lift of the matrix A . Next, our symmetrized Kronecker product obeys a symmetrized version of the mixed-product rule.

LEMMA 4.2.3 (Symmetrized mixed-product rule). *For matrices $A^1, A^2, \dots, A^k \in \mathbb{K}^{m \times p}$ and matrices $B^1, B^2, \dots, B^k \in \mathbb{K}^{p \times n}$,*

$$(A^1 \circledast A^2 \circledast \dots \circledast A^k)(B^1 \circledast B^2 \circledast \dots \circledast B^k) = \frac{1}{k!} \sum_{\rho \in S_k} (A^1 B^{\rho(1)}) \circledast (A^2 B^{\rho(2)}) \circledast \dots \circledast (A^k B^{\rho(k)}).$$

PROOF. We will show the expressions are equivalent. Let $\{j_1, \dots, j_k\}$ be a multiset of $[m]$ and let $\{j'_1, \dots, j'_k\}$ be a multiset of $[n]$. We may write an entry of matrix multiplication $(A^1 \circledast \dots \circledast A^k)(B^1 \circledast \dots \circledast B^k)$ as an inner product of the conjugate-transpose of a row of $(A^1 \circledast \dots \circledast A^k)$ and a column of $(B^1 \circledast \dots \circledast B^k)$ with expressions given by Lemma 4.2.2:

$$\begin{aligned} & (A^1 \circledast \dots \circledast A^k)(B^1 \circledast \dots \circledast B^k)[\{j_1, \dots, j_k\}, \{j'_1, \dots, j'_k\}] \\ &= \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \left(\sum_{\sigma \in S_k} A^1[j_{\sigma(1)}, :] \circledast \dots \circledast A^k[j_{\sigma(k)}, :] \right) \left(\sum_{\tau \in S_k} B^1[:, j'_{\tau(1)}] \circledast \dots \circledast B^k[:, j'_{\tau(k)}] \right) \\ &= \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} (A^1[j_{\sigma(1)}, :] \circledast \dots \circledast A^k[j_{\sigma(k)}, :]) (B^1[:, j'_{\tau(1)}] \circledast \dots \circledast B^k[:, j'_{\tau(k)}]) \\ &= \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \langle (A^1[j_{\sigma(1)}, :] \circledast \dots \circledast A^k[j_{\sigma(k)}, :])^*, B^1[:, j'_{\tau(1)}] \circledast \dots \circledast B^k[:, j'_{\tau(k)}] \rangle. \end{aligned}$$

Applying Corollary 4.1.4 gives

$$\begin{aligned} &= \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \frac{1}{k!} \sum_{\rho \in S_k} \langle (A^i[j_{\sigma(i)}, :])^*, B^{\rho(i)}[:, j'_{\tau(i)}] \rangle \\ &= \frac{1}{k!} \sum_{\rho \in S_k} \frac{w_{\{j_1, \dots, j_k\}} w_{\{j'_1, \dots, j'_k\}}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} (A^i B^{\rho(i)})[j_{\sigma(i)}, j'_{\tau(i)}] \\ &= \frac{1}{k!} \sum_{\rho \in S_k} \left((A^1 B^{\rho(1)}) \circledast (A^2 B^{\rho(2)}) \circledast \dots \circledast (A^k B^{\rho(k)}) \right) [\{j_1, \dots, j_k\}, \{j'_1, \dots, j'_k\}] \end{aligned}$$

where the last line follows from Definition 4.2.1 of the symmetric matrix product. $\square$

In the case of k -th order symmetric lifts, the symmetrized Kronecker product obeys a the mixed-product rule just like the usual Kronecker product.

COROLLARY 4.2.4. *For $A \in \mathbb{K}^{m \times p}$ and $B \in \mathbb{K}^{p \times n}$*

$$(AB)^{\otimes k} = A^{\otimes k} B^{\otimes k}.$$

This follows directly from Lemma 4.2.3.

Next, we will give another very useful property of symmetric lifts. To do this, where as previously it did not matter the exact order on the multiset used to index our symmetrized products (provide such an order is used consistently), we will now define a particular total order on multisets.

DEFINITION 4.2.5 (Dershowitz–Manna (DM) ordering [29]). *For two multisets M and M' of $[n]$, we say that $M' < M$ when there are two multisets X, Y such that*

- (i) $\emptyset \neq X \subseteq M$,
- (ii) $M' = (M \setminus X) \cup Y$,
- (iii) and for every $x \in X$ there is a $y \in Y$ so that $x > y$.

Next, we say that a matrix $R \in \mathbb{K}^{m \times n}$ is *upper triangular* if $m \geq n$ (*upper trapezoidal* if $m < n$) if $R[j, k] = 0$ for $j \in [m]$, $k \in [n]$ whenever $j > k$.

LEMMA 4.2.6. *When $R \in \mathbb{K}^{m \times n}$ is upper triangular (upper trapezoidal), $R^{\otimes k}$ is upper triangular (upper trapezoidal) with respect to the DM order on its entries in rows and columns i.e.*

$$R^{\otimes k}[\langle j_1, \dots, j_k \rangle, \langle j'_1, \dots, j'_k \rangle] = 0 \quad \text{when} \quad \langle j_1, \dots, j_k \rangle > \langle j'_1, \dots, j'_k \rangle.$$

PROOF. Expanding entries in $R^{\otimes k}$ by Definition 4.2.1 gives

$$(R^{\otimes k})[\langle j_1, j_2, \dots, j_k \rangle, \langle j'_1, j'_2, \dots, j'_k \rangle] = \frac{w_{\langle j_1, \dots, j_k \rangle} w_{\langle j'_1, \dots, j'_k \rangle}}{(k!)^2} \sum_{\sigma \in S_k} \sum_{\tau \in S_k} \prod_{i=1}^k R[j_{\sigma(i)}, j'_{\tau(i)}].$$

We may simplify by re-indexing permutation to get

$$(R^{\otimes k})[\langle j_1, j_2, \dots, j_k \rangle, \langle j'_1, j'_2, \dots, j'_k \rangle] = \frac{w_{\langle j_1, \dots, j_k \rangle} w_{\langle j'_1, \dots, j'_k \rangle}}{k!} \sum_{\sigma \in S_k} \prod_{i=1}^k R[j_i, j'_{\sigma(i)}].$$

Now if $M = \langle j_1, j_2, \dots, j_k \rangle$ and $M' = \langle j'_1, j'_2, \dots, j'_k \rangle$ we will show that $(R^{\otimes k})[M, M'] = 0$ whenever $M > M'$ by contraposition. Suppose that $(R^{\otimes k})[M, M'] \neq 0$ for some k -element multisets M and M' of $[d]$. If this entry is non-zero then there must exists a permutation $\sigma \in S_k$ such that for every $i \in [k]$, $j_i \geq j'_{\sigma(i)}$ so that the product $\prod_{i=1}^k R[j_i, j'_{\sigma(i)}]$ is non-zero. Now set $S = \{i \in [k] : j_i > j'_{\sigma(i)}\}$. If $S = \emptyset$ then $M = M'$. Supposing that S is non-empty, set $X = \langle j_i \in M : i \in S \rangle$ and $Y = \langle j'_{\sigma(i)} \in M' : i \in S \rangle$. We then have that $X \subseteq M$, $X \neq \emptyset$ and $M' = (M \setminus X) \cup Y$ by construction. Further, by our assumption on σ , for any $x \in X$ there is a $y \in Y$ so that $x > y$. By Definition 4.2.5, we have $M < M'$. Since either $M = M'$ or $M < M'$, $M \nprec M'$. It follows that when $M > M'$, $(R^{\otimes k})[M, M'] = 0$ and $R^{\otimes k}$ is upper triangular. $\square$

We also note that lifting orthonormal bases produces orthonormal bases in our de-duplicated format.

LEMMA 4.2.7. For $Q \in \mathbb{K}^{m \times n}$, $m \geq n$, if Q has orthonormal ($\mathbb{K} = \mathbb{R}$) or unitary ($\mathbb{K} = \mathbb{C}$) columns then so does $Q^{\otimes k}$.

PROOF. We will show that $(Q^{\otimes k})^* Q^{\otimes k} = I$ by expanding its entries. By matrix multiplication and the column-wise definition of the symmetric lift given by Lemma 4.2.2,

$$((Q^{\otimes k})^* Q^{\otimes k})[\mathcal{J}_1, \dots, \mathcal{J}_k, \mathcal{J}'_1, \dots, \mathcal{J}'_k] = \langle w_{\mathcal{J}_1, \dots, \mathcal{J}_k} Q^{\otimes k}[:, \mathcal{J}_1, \dots, \mathcal{J}_k], w_{\mathcal{J}'_1, \dots, \mathcal{J}'_k} Q^{\otimes k}[:, \mathcal{J}'_1, \dots, \mathcal{J}'_k] \rangle.$$

Next, using Corollary 4.1.4 this inner product becomes

$$\frac{w_{\mathcal{J}_1, \dots, \mathcal{J}_k} w_{\mathcal{J}'_1, \dots, \mathcal{J}'_k}}{k!} \sum_{\sigma \in S_k} \prod_{i=1}^k \langle Q[:, j_i], Q[:, j'_{\sigma(i)}] \rangle = \frac{w_{\mathcal{J}_1, \dots, \mathcal{J}_k} w_{\mathcal{J}'_1, \dots, \mathcal{J}'_k}}{k!} \sum_{\sigma \in S_k} \prod_{i=1}^k \begin{cases} 1 & j_i = j'_{\sigma(i)}, \\ 0 & \text{otherwise,} \end{cases}$$

where the piecewise expression follows from Q having orthonormal columns. Now this expression can only be non-zero if $\mathcal{J}_1, \dots, \mathcal{J}_k = \mathcal{J}'_1, \dots, \mathcal{J}'_k$, otherwise for any permutation at least one $i \in [k]$ will have $j_i = j'_{\sigma(i)}$.

In the case that $\mathcal{J}_1, \dots, \mathcal{J}_k = \mathcal{J}'_1, \dots, \mathcal{J}'_k$,

$$\frac{w_{\mathcal{J}_1, \dots, \mathcal{J}_k}^2}{k!} \sum_{\sigma \in S_k} \begin{cases} 1 & (j_1, j_2, \dots, j_k) = (j_{\sigma(1)}, j_{\sigma(2)}, \dots, j_{\sigma(k)}), \\ 0 & \text{otherwise.} \end{cases}$$

Now, the sum counts how many permutations of the multiset will remain unchanged if we fix an ordering. But this is exactly $\prod_{m \in \text{multi}(\{j_1, \dots, j_k\})} m!$ which cancels with components of the weight $w_{\mathcal{J}_1, \dots, \mathcal{J}_k}^2$. As a result when $\mathcal{J}_1, \dots, \mathcal{J}_k = \mathcal{J}'_1, \dots, \mathcal{J}'_k$ this simplifies to 1 and

$$((Q^{\otimes k})^* Q^{\otimes k})[\mathcal{J}_1, \dots, \mathcal{J}_k, \mathcal{J}'_1, \dots, \mathcal{J}'_k] = \begin{cases} 1 & \mathcal{J}_1, \dots, \mathcal{J}_k = \mathcal{J}'_1, \dots, \mathcal{J}'_k, \\ 0 & \text{otherwise.} \end{cases}$$

This is the identity matrix of size $\binom{n}{k}$ so $Q^{\otimes k}$ has orthonormal columns. $\square$

Lemma 4.2.7 shows that if we need an orthonormal basis over a symmetric space, we may first find an orthonormal a basis by Gram-Schmidt or QR factorization, then create a symmetric lift rather than running these procedures on the lifted $\binom{m}{k} \times \binom{n}{k}$ matrix, which may be quite large.

LEMMA 4.2.8. If $A \in \mathbb{K}^{m \times n}$ with $m \geq n$ has linearly independent columns, then so does $A^{\otimes k}$.

PROOF. Since A has linearly independent columns, it has a thin-QR decomposition $A = QR$ where Q has orthonormal columns and R is a full-rank upper triangular matrix, i.e. $R[j, j'] = 0$ when $j > j'$. Without loss of generality we may assume R has positive real entries along its diagonal, $R[j, j] > 0$ for $j \in [n]$, since

since A has full rank. By Corollary 4.2.4 we have that $A^{\otimes k} = Q^{\otimes k} R^{\otimes k}$. Lemma 4.2.7 shows that $Q^{\otimes k}$ is orthonormal and Lemma 4.2.6 shows $R^{\otimes k}$ is upper triangular so we need only to show $R^{\otimes k}$ is full-rank. If we examine the diagonal of $R^{\otimes k}$, by Definition 4.2.1

$$(R^{\otimes k})[\langle j_1, \dots, j_k \rangle, \langle j_1, \dots, j_k \rangle] = \frac{w_{\langle j_1, \dots, j_k \rangle}^2}{k!} \sum_{\sigma \in S_k} \prod_{i=1}^k R[j_i, j_{\sigma(i)}].$$

For a given $\sigma \in S_k$, we must have that $j_i \leq j_{\sigma(i)}$ otherwise the product inside the sum above is zero. This can only occur for permutations that keep the sequence $j_1, \dots, j_k$ unchanged. As previously noted the number of such permutations is the reciprocal of $\frac{w_{\langle j_1, \dots, j_k \rangle}^2}{k!}$ so

$$(R^{\otimes k})[\langle j_1, \dots, j_k \rangle, \langle j_1, \dots, j_k \rangle] = \prod_{i=1}^k R[j_i, j_i] > 0$$

since R has positive diagonal elements. Hence, $R^{\otimes k}$ has full rank too. Then $Q^{\otimes k} R^{\otimes k}$ is a QR-decomposition of $A^{\otimes k}$ and it follows that if A has linearly independent columns then so does $A^{\otimes k}$. $\square$

4.3. Tensors with Partial Symmetries.

Next, we say that a tensor in the space $(\mathbb{K}^{n_1})^{\otimes k_1} \otimes (\mathbb{K}^{n_2})^{\otimes k_2} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L}$ has L different partial-symmetries along its modes. The order of a tensor in this space is $k_1 + k_2 + \dots + k_L$. We define the standard basis for this space indexed by an L -tuple of multisets $M_i = \langle j_1^i, j_2^i, \dots, j_{k_i}^i \rangle$ for $i = 1, 2, \dots, L$,

$$\vec{e}_{(M_1, M_2, \dots, M_L)} := \vec{e}_{M_1} \otimes \vec{e}_{M_2} \otimes \dots \otimes \vec{e}_{M_L}.$$

Warning: The symmetrized Kronecker product we define is not associative, for example note that $(w \otimes x) \otimes (y \otimes z) \neq w \otimes x \otimes y \otimes z$ since these objects belong to spaces of order-4 tensors with different sets of symmetries enforced along the modes. Specifically, $(w \otimes x) \otimes (y \otimes z) \in ((\mathbb{K}^n)^{\otimes 2})^{\otimes 2} \cong \mathbb{K}^m$ where $m = \binom{n+1}{2} + 1$ whereas $w \otimes x \otimes y \otimes z \in (\mathbb{K}^n)^{\otimes 4} \cong \mathbb{K}^k$ where $k = \binom{n+3}{4}$.

4.3.1. Matricization from an irredundant format. While the vectors in $\mathbb{K}^{\binom{n}{k}}$ are a convenient way of storing the entries of a symmetric tensor without redundancy, to use the multilinear structure of one, one must convert the tensor back to a matrix. First, if $v \in (\mathbb{K}^n)^{\otimes 2}$ and gives the coordinates of a symmetric matrix in our canonical basis, we define the matricization entry-wise by

$$(4.9) \quad \text{matricize}(v)[j_1, j_2] := \frac{1}{w_{\langle j_1, j_2 \rangle}} v[\langle j_1, j_2 \rangle], \quad j_1, j_2 \in [n].$$

More generally for $\ell \leq k$, $\text{matricize}_\ell : (\mathbb{K}^n)^{\otimes k} \mapsto (\mathbb{K}^n)^{\otimes \ell} \otimes (\mathbb{K}^n)^{\otimes k-\ell}$ reshapes a symmetric tensor $\mathcal{T}$ to a matrix via

$$\text{matricize}_\ell(\mathcal{T})[M, K] := \frac{w_M w_K}{w_{M \cup K}} \mathcal{T}[M \cup K]$$

where M is an ℓ -element multiset of $[n]$ and K is a $(k-\ell)$ -element multiset of $[n]$. This particular matrix is ℓ -th *catalecticant matrix* of v expressed in the coordinate bases $\mathcal{S}_{\ell,n}$ and $\mathcal{S}_{k-\ell,n}$ defined in Lemma 4.1.1 for the row space and column space respectively.

Finally, to convert $T \in (\mathbb{K}^{n_1})^{\otimes k_1} \otimes (\mathbb{K}^{n_2})^{\otimes k_2} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L}$ to one without explicit symmetries in $(\mathbb{K}^{n_1})^{\otimes k_1} \otimes (\mathbb{K}^{n_2})^{\otimes k_2} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L}$, we can define the decompression by

$$(4.10) \quad \text{decompress}(T)[j_1^1, j_2^1, \dots, j_{k_1}^1, \dots, j_1^L, j_2^L, \dots, j_{k_L}^L] := \frac{T[\langle j_1^1, j_2^1, \dots, j_{k_1}^1 \rangle, \dots, \langle j_1^L, j_2^L, \dots, j_{k_L}^L \rangle]}{w_{\langle j_1^1, j_2^1, \dots, j_{k_1}^1 \rangle} \dots w_{\langle j_1^L, j_2^L, \dots, j_{k_L}^L \rangle}}.$$

where for $\ell \in [L]$ and $k \in [k_\ell]$, $j_k^\ell \in [n_\ell]$ are indices.

4.4. Efficiently Cutting Out a Segre-Veronese Variety

In this section, we define the Segre-Veronese varieties. These conic varieties contains the rank-1 tensors with partial symmetries along modes and understanding their algebra is critical for understanding some methods in tensor decomposition. Generating the ideal for this variety is well understood from a theoretical perspective, see[15]. However, in this section we will examine how one might construct a basis of degree two determinant-like polynomials which span the vector space of degree two polynomials cutting out these varieties. Such a construction is advantageous since it may be represented very sparsely.

In Section 2.4, we determined a basis for $\mathcal{I}(\mathcal{X}_1)_2$, defined in (2.10). By Lemma 4.1.3 we can recast (2.11) to an inner product in $(\mathbb{K}^{m \times n})^{\otimes 2}$ since

$$M_{abcd}(X) = \langle \vec{e}_{ab} \otimes \vec{e}_{bc} - \vec{e}_{ad} \otimes \vec{e}_{cb}, X^{\otimes 2} \rangle = \left\langle \frac{\vec{e}_{\langle ab, cd \rangle}}{w_{\langle ab, cd \rangle}} - \frac{\vec{e}_{\langle ad, cb \rangle}}{w_{\langle ad, cb \rangle}}, X^{\otimes 2} \right\rangle.$$

Since $a < c$ and $b < d$, $w_{\langle ab, cd \rangle} = w_{\langle ad, cb \rangle} = \sqrt{2}$ and the vectors associated to each minor in our de-duplicated symmetric tensor format is simply the difference of two standard basis vectors

$$(4.11) \quad F_{abcd} := \frac{1}{\sqrt{2}} (\vec{e}_{\langle ab, cd \rangle} - \vec{e}_{\langle ad, cb \rangle}) \in (\mathbb{K}^{m \times n})^{\otimes 2}.$$

We this we can see that a basis for $\mathcal{I}(\mathcal{X}_1)$ such that each vector is the difference of two standard basis vector, appropriately weighted.

4.4.1. Symmetric Rank-1 Matrices. Next, if we identify $S^2(\mathbb{K}^m)$ with $(\mathbb{K}^m)^{\otimes 2}$ we may define the variety of rank-at-most-1 symmetric matrices as the set

$$\mathcal{X}_1^\vee = \{u^{\otimes 2} : u \in \mathbb{K}^m\} \subseteq (\mathbb{K}^m)^{\otimes 2}.$$

EXAMPLE 4.4.1 (The variety of rank-at-most-1 symmetric matrices). *For any matrix $X \in (\mathbb{K}^m)^{\otimes 2}$ stored in our irredundant format, note that $X[\llbracket a, b \rrbracket] = w_{\llbracket a, b \rrbracket} \text{matricize}(X)[a, b]$. Using the (2.10) the 2-minors of X in this case then have the form*

$$M_{abcd}(\text{matricize}(X)) = \frac{X[\llbracket a, b \rrbracket]}{w_{\llbracket a, b \rrbracket}} \cdot \frac{X[\llbracket c, d \rrbracket]}{w_{\llbracket c, d \rrbracket}} - \frac{X[\llbracket a, d \rrbracket]}{w_{\llbracket a, d \rrbracket}} \cdot \frac{X[\llbracket c, b \rrbracket]}{w_{\llbracket c, b \rrbracket}}$$

for $a, b, c, d \in [m]$. Next, observe that $\langle \vec{e}_{\llbracket a, b \rrbracket}, X \rangle = X[\llbracket a, b \rrbracket]$. Without matricization, we can compute these minors by taking inner products against the vectors

$$(4.12) \quad E_{abcd}^\vee := \frac{1}{w_{\llbracket a, b \rrbracket} w_{\llbracket c, d \rrbracket} w_{\llbracket a, b \rrbracket} w_{\llbracket c, d \rrbracket}} \vec{e}_{\llbracket \llbracket a, b \rrbracket, \llbracket c, d \rrbracket \rrbracket} - \frac{1}{w_{\llbracket a, d \rrbracket} w_{\llbracket c, b \rrbracket} w_{\llbracket a, d \rrbracket} w_{\llbracket c, b \rrbracket}} \vec{e}_{\llbracket \llbracket a, d \rrbracket, \llbracket c, b \rrbracket \rrbracket}$$

so that $\langle E_{abcd}^\vee, X^{\otimes 2} \rangle = 0$ if $X \in (\mathbb{K}^m)^{\otimes 2}$ has the form $X = u^{\otimes 2}$ for some $u \in \mathbb{K}^m$.

Some other facts that are helpful: $\text{span}\{(u^{\otimes 2})^{\otimes 2} : u \in \mathbb{K}^m\} \cong \text{span}\{u^{\otimes 4} : u \in \mathbb{K}^m\} \cong S^4(\mathbb{K}^m)$. One can see this easily using the natural injection $x^{\otimes 2} \mapsto x^{\otimes 2}$ and the associativity of repeated Kronecker products. Next, $\mathcal{I}(\mathcal{X}_1^\vee)^\perp = \text{span}\{(u^{\otimes 2})^{\otimes 2} : u \in \mathbb{K}^n\}$ so $\dim \mathcal{I}(\mathcal{X}_1^\vee)^\perp = \binom{m}{4}$ and

$$\dim \mathcal{I}(\mathcal{X}_1^\vee) = \left(\binom{\binom{m}{2}}{2} \right) - \binom{m}{4} = \frac{1}{12}(m+1)m^2(m-1).$$

LEMMA 4.4.2. *For the variety of symmetric rank-1 $m \times m$ matrices $\mathcal{X}_1^\vee$, the set of vectors $\{E_{abcd}^\vee : 1 \leq a < c \leq m, 1 \leq b < d \leq m \text{ and } a \leq b \text{ and } c \leq d\}$ is a basis for $\mathcal{I}(\mathcal{X}_1)_2$ when considering it to be a subspace of $((\mathbb{K}^m)^{\otimes 2})^{\otimes 2}$ by identifying monomials with the symmetrized Kronecker products of standard basis elements.*

PROOF. First, note the extra conditions that $a \leq b$ and $c \leq d$ as compared to Lemma 2.4.6. We proceed with three main steps. First, we show that $\text{span}(E_{abcd}^\vee) \subseteq I_2(\mathcal{X}_1^\vee)$. Second, we show that the number of elements in the set is equal to $\dim I_2(\mathcal{X}_1^\vee)$. Third, we show that the set of vectors $\{E_{abcd}^\vee\}$ is a linearly independent set. Together these three ensure the desired result.

First, suppose that $x \in \mathbb{K}^m$. Then $\langle \vec{e}_{\llbracket ab \rrbracket}, x^{\otimes 2} \rangle = x^{\otimes 2}[\llbracket a, b \rrbracket] = w_{\llbracket a, b \rrbracket} x[a]x[b]$. It follows that for any $a, b, c, d \in [m]$,

$$\langle E_{abcd}^\vee, (x^{\otimes 2})^{\otimes 2} \rangle = \frac{x^{\otimes 2}[\llbracket a, b \rrbracket] x^{\otimes 2}[\llbracket c, d \rrbracket]}{w_{\llbracket a, b \rrbracket} w_{\llbracket c, d \rrbracket}} - \frac{x^{\otimes 2}[\llbracket a, d \rrbracket] x^{\otimes 2}[\llbracket c, b \rrbracket]}{w_{\llbracket a, d \rrbracket} w_{\llbracket c, b \rrbracket}} = x[a]x[b]x[c]x[d] - x[a]x[d]x[c]x[b] = 0$$

so $E_{abcd}^\vee \in \mathcal{I}(\mathcal{X}_1^\vee)_2$.

Next, we need to count the number of unique values of (a, b, c, d) satisfying $1 \leq a < c \leq m, 1 \leq b < d \leq m$ and $a \leq b$ and $c \leq d$. Define $\text{TR}(a, b) = \{(c, d) : 1 \leq c \leq d \leq m, a < c \text{ and } b < d\}$ to be the set of possible values of c and d given $a \leq b$. This set are the integer coordinates within the trapezoid with corners $(a+1, b+1)$, $(a+1, m)$, (m, m) and $(b+1, b+1)$. This degenerates to a triangle when $b = a$. The mutually exclusive conditions $c < b+1$ or $c \geq b+1$ split this into a rectangle of width $(m-b)$ and height $(b-a)$ and triangle with a base $(m-b)$ integer points. It follows that $|\text{TR}(a, b)| = (m-b)(b-a) + \frac{1}{2}(m-b+1)(m-b)$. The number of $E_{abcd}^\vee$ vectors is then

$$\sum_{b=1}^m \sum_{a=1}^b |\text{TR}(a, b)| = \sum_{b=1}^m \sum_{a=1}^b (m-b)(b-a) + \frac{1}{2}(m-b+1)(m-b) = \frac{1}{12}(m+1)m^2(m-1)$$

which matches $\dim \mathcal{I}(\mathcal{X}_1^\vee)$.

Finally, we will show that these vectors are linearly independent. Note that if $\{a, b, c, d\} \neq \{a', b', c', d'\}$ then $\langle E_{abcd}^\vee, E_{a'b'c'd'}^\vee \rangle = 0$. Hence we may partition $\mathcal{E} := \{E_{abcd}^\vee : 1 \leq a < c \leq m, 1 \leq b < d \leq m \text{ and } a \leq b \text{ and } c \leq d\}$ into

$$\bigsqcup_{\{a,b,c,d\}} \underbrace{\{E_{a'b'c'd'}^\vee \in \mathcal{E} : \{a, b, c, d\} = \{a', b', c', d'\}\}}_{=: \mathcal{E}_{abcd}}$$

Now $E_{abcd}^\vee \in \mathcal{E}_{abcd}$ implies $a < c, b < d, a \leq b$ and $c \leq d$. If $a = b$ or $c = d$, then $E_{abcd}^\vee$ is the only element (a singleton) of $\mathcal{E}_{abcd}$. If $a < b$ and $c < d$ then $E_{acbd}^\vee \in \mathcal{E}_{abcd}$ as well. However, comparing the vector $E_{abcd}^\vee$ to $E_{acbd}^\vee$ in (4.12), one can see that these are not scalar multiples of the same vector. It follows that they are linearly independent and that $\text{span}(\mathcal{E}) = \bigoplus_{\{a,b,c,d\}} \text{span}(\mathcal{E}_{abcd})$. $\square$

4.4.2. Segre-Veronese Varieties. A tensor with L partial symmetries $T \in (\mathbb{K}^{n_1})^{\otimes k_1} \otimes (\mathbb{K}^{n_2})^{\otimes k_2} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L}$, is rank-1 if there are non-zero vectors $u^j \in \mathbb{K}^{n_j}$ for $j \in [L]$ such that

$$T = (u^1)^{\otimes k_1} \otimes (u^2)^{\otimes k_2} \otimes \dots \otimes (u^L)^{\otimes k_L}.$$

The set of such tensor (plus the zero tensor) form a *Segre-Veronese variety* which we denote $\text{SV}_{\mathbf{n}, \mathbf{k}}$ where $\mathbf{n} = (n_1, \dots, n_L)$ and $\mathbf{k} = (k_1, \dots, k_L)$. Our main goal in this section is to show how in compressed format we may cut out $\text{SV}_{\mathbf{n}, \mathbf{k}}$ by degree-2 homogeneous polynomials which may computed via $\langle E_{V, V'}, T^{\otimes 2} \rangle$ where

$$E_{V, V'} := \frac{\vec{e}_V}{w_V} - \frac{\vec{e}_{V'}}{w_{V'}} \in ((\mathbb{K}^{n_1})^{\otimes k_1} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L})^{\otimes 2}$$

and V, V' are somewhat complicated indexing objects and w_V and $w_{V'}$ are weights associated with these objects. Eliding details about these objects for the time being, the power of this analysis arises when we

identify the objects as V and V' as vertices of a graph and $E_{V,V'}$ as an edge. At the end we will have an efficient method for constructing a basis for $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2$ e.g. for use in the JLV algorithm.

We will now dive into the weeds to describe the objects and graph structure resulting. Symbolically, we can denote each entries of T in our irredundant format by $T[M^1, M^2, \dots, M^L]$ where each M^j is a k_j -element multiset of $[n_j]$ for each $j \in [L]$. Bernardi [15, Theorem 3.11] shows that polynomials of the form (2.13) will still generate $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})$. In the following, we represent these polynomials in the particular basis we use for tensors with partial symmetries.

LEMMA 4.4.3. *Set*

$$V = \wr(M_1, \dots, M_L), (N_1, \dots, N_L)\S$$

with of M_j and N_j each being k_j -element multisets of $[n_j]$ for each $j \in [L]$. For any $\ell \in [L]$, if possible pick $x \in M_\ell$ and $y \in N_\ell$ such that $y \neq x$. Define new multisets M'_ℓ, N'_ℓ by swapping x and y : $M'_\ell = (M_\ell \setminus \{x\}) \cup \{y\}$ and $N'_\ell = (N_\ell \setminus \{y\}) \cup \{x\}$. Set $V' = \wr((M_1, \dots, M'_\ell, \dots, M_L), (N_1, \dots, N'_\ell, \dots, N_L))\S$. Further, define a weight for V by

$$(4.13) \quad w_V := w_{\wr(M_1, \dots, M_L), (N_1, \dots, N_L)\S} \prod_{j=1}^L (w_{M_j} w_{N_j})$$

and a weight for V' , denoted $w_{V'}$, by the same principle. It follows that

$$(4.14) \quad E_{V,V'} := \frac{\vec{e}_V}{w_V} - \frac{\vec{e}_{V'}}{w_{V'}} \in ((\mathbb{K}^{n_1})^{\otimes k_1} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L})^{\otimes 2}$$

is a non-zero vector such that $\langle E_{V,V'}, T^{\otimes 2} \rangle = 0$ when $T \in \text{SV}_{\mathbf{n},\mathbf{k}}$.

If V and V' are related in that we can produce V' from V (or vice versa) using the procedure of swapping two elements described in Lemma 4.4.3 then we will say that “ V and V' are related by an exchange.” Ultimately, the methodology for deriving this structure follows very similarly to that of Example 4.4.1 in the broad strokes that we can start with the polynomials from (2.13) which should be understood as a polynomial on entries of the decompression of T when T has symmetries. Using the relationship to the compressed form given in (4.10) one naturally arrives at the structures and weights described.

PROOF OF LEMMA 4.4.3. First, the vector is non-trivial since the condition $x \neq y$ ensures $V_1 \neq V_2$ and so $\vec{e}_{V_1} \neq \vec{e}_{V_2}$ and E_{V_1,V_2} . Second, assuming $T = (u^1)^{\otimes k_1} \otimes (u^2)^{\otimes k_2} \otimes \dots \otimes (u^L)^{\otimes k_L}$, it is tedious but not challenging to determine that $\langle E_{V_1,V_2}, T^{\otimes 2} \rangle = 0$ in an analogous manner to that shown in the start of the proof of Lemma 4.4.2. $\square$

Now define the graph $\mathcal{G}_{\mathbf{n}, \mathbf{k}} = (\mathcal{V}, \mathcal{E})$ with the set of vertices $\mathcal{V}$ being all by two-element multisets of tuples of multisets (see Lemma 4.4.3) and a set of edges $\mathcal{E}$ such that $(V, V') \in \mathcal{E}$ when V and V' are related by an exchange. We can consider the edges of this graph to be undirected since if $(V, V') \in \mathcal{E}$ then $(V', V) \in \mathcal{E}$.

First, some graph terminology:

- A subgraph $\mathcal{F}$ of a graph $\mathcal{G}$ is a graph that contains a subset of the $\mathcal{G}$'s vertex set and edge sets.
- A *path* is a sequence of vertices $V_1, V_2, \dots, V_m$ such that $(V_j, V_{j+1}) \in \mathcal{E}$.
- Two vertices are *connected* if there is a path between them (and $\mathcal{G}$ is connected if every two vertices in $\mathcal{V}$ are connected).
- A *cycle* is a path $V_1, V_2, \dots, V_m$ which also has that $(V_m, V_1) \in \mathcal{E}$.
- A *tree* is a connected graph with no cycles.
- A *forest* is the disjoint union of multiple trees.
- A *leaf* of a tree is a vertex connected to only one vertex.
- A *spanning forest* of $\mathcal{G}$ is subgraph of $\mathcal{G}$ which is a forest and contains every vertex of $\mathcal{G}$.

We have three observations:

OBSERVATION 4.4.4. *For $V = \wr(M_1, \dots, M_L), (N_1, \dots, N_L)\S$ and $V' = \wr(M'_1, \dots, M'_L), (N'_1, \dots, N'_L)\S$ if there is any $\ell \in [L]$ so that $M_\ell \cup N_\ell \neq M'_\ell \cup N'_\ell$ then V and V' cannot be related by an exchange.*

OBSERVATION 4.4.5. *For $V = \wr(M_1, \dots, M_L), (N_1, \dots, N_L)\S$ and $V' = \wr(M'_1, \dots, M'_L), (N'_1, \dots, N'_L)\S$ if $M_\ell \cup N_\ell = M'_\ell \cup N'_\ell$ for each $\ell \in [L]$ then there is a sequence of edges in $\mathcal{G}_{\mathbf{n}, \mathbf{k}}$ from V to V' . Hence $\mathcal{G}_{\mathbf{n}, \mathbf{k}}$ has $\left(\prod_{j=1}^L \binom{n_j}{2k_j}\right)$ -many connected components.*

OBSERVATION 4.4.6. *If $V_1, V_2, \dots, V_m$ form a cycle in $\mathcal{G}_{\mathbf{n}, \mathbf{k}}$ (i.e. $(V_j, V_{j+1}) \in E$ for $j \in [m-1]$ and $(V_m, V_1) \in E$). Then*

$$E_{V_1, V_2} + E_{V_2, V_3} + \dots + E_{V_{m-1}, V_m} + E_{V_m, V_1} = \vec{0}$$

and the set of vectors $\{E_{V_1, V_2}, \dots, E_{V_{m-1}, V_m}, E_{V_m, V_1}\}$ is linearly dependent.

Combining these three observations we have a nice result.

PROPOSITION 4.4.7. *Given $\mathcal{G}_{\mathbf{n}, \mathbf{k}} = (\mathcal{V}, \mathcal{E})$ let $\mathcal{F} = (\mathcal{V}, F)$ be any spanning forest on $\mathcal{G}_{\mathbf{n}, \mathbf{k}}$, then the set $\{E_{V, V'} : (V, V') \in F\}$ is a basis for $\mathcal{I}(\text{SV}_{\mathbf{n}, \mathbf{k}})_2$.*

PROOF. Suppose that $(V, V') \in \mathcal{E}$ but $(V, V') \notin F$. By Observation 4.4.5, there is a path between V and V' in $\mathcal{G}_{\mathbf{n}, \mathbf{k}}$. Further, since the trees in $\mathcal{F}$ span G there is a path with edges in $\mathcal{F}$ between V and V' . So

$E_{V,V'} \in \text{span}\{E_{V_1,V_2} : (V_1, V_2) \in F\}$. It then follows that

$$\text{span}\{E_{V,V'} : (V, V') \in F\} = \{E_{V_1,V_2} : (V, V') \in E\} = \mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2.$$

Next, we need show that this set is a basis for the space. Suppose that there are coefficients $c_{V,V'}$ associated to each edge $(V, V') \in F$ so that

$$(4.15) \quad \sum_{(V,V') \in F} c_{V,V'} E_{V,V'} = \vec{0} \in ((\mathbb{K}^{n_1})^{\otimes k_1} \otimes \dots \otimes (\mathbb{K}^{n_L})^{\otimes k_L})^{\otimes 2}.$$

We will show that every $c_{V,V'} = 0$ from which it follows that $\{E_{V_1,V_2} : (V_1, V_2) \in F\}$ is a basis for $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2$. Now $\mathcal{F}$ must have at least one tree and that tree must have at least one leaf, call this V_ℓ . Since V_ℓ is a leaf of a tree in $\mathcal{F}$, it is only connected to one vertex, call this $V_{\ell'}$. Since V_ℓ is a leaf the only vector in (4.15) with $\vec{e}_{V_\ell}$ in its support is $E_{V_\ell,V_{\ell'}}$. It follows, the only way the sum in (4.15) can be the zero vector is if the coefficient $c_{V_\ell,V_{\ell'}} = 0$. Since the coefficient is zero we could remove V_ℓ from $\mathcal{V}$ and $(V_\ell, V_{\ell'})$ from $\mathcal{F}$ to produce a new smaller spanning forest. If the tree was the single edge $(V_\ell, V_{\ell'})$ we also eliminate $V_{\ell'}$ from the graph. Since the original forest has a finite number of trees with a finite number of vertices we can apply this logic to deduce that every $c_{V,V'}$ for $(V, V') \in \mathcal{E}$ is zero, so $\{E_{V_1,V_2} : (V, V') \in F\}$ is a basis for $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2$. $\square$

In conjunction with Observation 4.4.5, Proposition 4.4.7 gives an algorithm to efficiently find the a basis for $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2$.

Algorithm 1 Find the a basis for $\mathcal{I}(\text{SV}_{\mathbf{n},\mathbf{k}})_2$ where $\mathbf{n} = (n_1, \dots, n_L)$ and $\mathbf{k} = (k_1, \dots, k_L)$.

- 1: Let $F = \emptyset$
 - 2: **for** $(G_1, G_2, \dots, G_L) \in \left(\binom{[n_1]}{2k_1}\right) \times \left(\binom{[n_2]}{2k_2}\right) \times \dots \times \left(\binom{[n_L]}{2k_L}\right)$ **do**
 - 3: Pick any $V_r = \langle (M_1, \dots, M_L), (N_1, \dots, N_L) \rangle$ such that $M_j \cup N_j = G_j$ for each $j \in [L]$.
 - 4: Build a breadth first spanning tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}) \subseteq \mathcal{G}_{\mathbf{n},\mathbf{k}}$ starting from vertex V_r .
 - 5: Add each of $E_{V,V'}$ as defined in Lemma 4.4.3 for $(V, V') \in \mathcal{E}$ to F .
 - 6: **return** F
-

This procedure outlined in Algorithm 1 is implemented in Python and can be found in the following [repository](#)¹ hosted on GitHub. The relevant file is titled `segre_veronese_ideal_generation.py`.

¹Full URL is: <https://github.com/lemniscate8/linear-intersect-varieties>

Beyond Frequency Marching: Orbit Recovery in Dihedral and Projected Multireference Alignment

5.1. Introduction

In the *multireference alignment (MRA)* problem [3, 8, 10, 56], we are given samples of the form

$$(5.1) \quad y_i = R_i \cdot x + \sigma \xi_i \in \mathbb{R}^d, \quad i = 1, \dots, N$$

where

- $x \in \mathbb{R}^d$ is a “signal” we aim to recover;
- R_i is an element of the cyclic group $\mathcal{C}_d = \{R_0, R_1, \dots, R_{d-1}\}$ drawn independently and uniformly at random;
- the cyclic group acts on $\mathbb{R}^d$ via cyclic shifts, so $R_\ell \cdot x$ denotes the vector x with entries cycled to the right by an offset of ℓ ;
- $\xi_i \sim \mathcal{N}(0, I)$ is Gaussian noise;
- $\sigma > 0$ is the noise level, which determines the *signal-to-noise ratio* $\text{SNR} := 1/\sigma^2$.

Given samples $\{y_i\}_{i=1}^N$, the goal is to recover x . However, x is only identifiable up to cyclic shift, that is, we aim to recover the *orbit* of x under the group action of cyclic shifts, $\{R_\ell \cdot x : \ell = 0, 1, \dots, d-1\}$.

MRA is one canonical example of a task within the class of *orbit recovery* problems [7], where the observed samples are affected by some group action. These problems have a variety of practical motivations, including the reconstruction of 3-dimensional molecular structure from *cryogenic electron microscopy (cryo-EM)* images. In cryo-EM, individual particles are frozen in a thin layer of ice through which images are taken by a scanning electron microscope. Particle orientations are random relative to the viewing plane. While random rotations may be considered nuisance parameters due to these rotations, we can only hope to recover a 3D structure up to a 3D rotation i.e., the structure’s orbit under the action of $SO(3)$. Two added complexities are that the images are tomographic projections of a 3D electro-static potential function onto a 2D imaging plane, and low amounts of radiation must be used to avoid damage and motion blur from melting ice, leading to noisy images. The tomographic projection results in additional ambiguity since

particles of different chirality will have the same distribution of projections. With projection, one can only hope to recover structure up to handed-ness, or equivalently, up to the action of $O(3)$. See [35] for a historical narrative of cryo-EM's experimental development and [63] for a more detailed discussion of the mathematical details we mention here.

Collecting useful experimental cryo-EM data requires considerable expertise, but even after successful data collection, the mathematics of structure determination as a non-linear inverse problem is highly non-trivial. Orbit recovery in the mathematical models like that those in [63] is far from solved. The MRA problem provides a toy model which captures some key features of the cryo-EM problem, including a phase transition first observed empirically by Sigworth [60] and later proven [3, 10, 56] in which the reconstruction error scales as $1/\text{SNR}^3$ in the high-noise regime. However, the MRA model simplifies several aspects present in cryo-EM since the action of cyclic shifts is diagonalizable while the action of $SO(3)$ is not. In addition, in cryo-EM the tomographic projection operation has a non-trivial kernel so that each sample only contains partial information about the object even when $\sigma = 0$.

We study two different multi-reference alignment (MRA) variants which aim to introduce these challenges to the MRA model. The first variant is *dihedral* multi-reference alignment (dMRA) [13, 31] where samples are drawn from

$$(5.2) \quad y_i = g_i \cdot x + \sigma \xi_i, \quad i = 1, \dots, N$$

where (5.2) modifies (5.1) in that g_i is an element of dihedral group $\mathcal{D}_d = \{R_0, \dots, R_{d-1}, JR_0, \dots, JR_{d-1}\}$ drawn independently and uniformly at random. Here, J acts by reversing the order of entries in the vector. The distribution of the samples $\{y\}_{i=1}^N$ is invariant if we exchange x for any $g \cdot x$ with $g \in \mathcal{D}_d$ so we can only expect to recover x up to its $\mathcal{D}_d$ -orbit.

The second variant, a version of which was introduced in [7, Section 4.4.1], is *projected* multi-reference alignment (pMRA)

$$(5.3) \quad y_i = \Pi(R_i \cdot x) + \sigma \xi_i, \quad i = 1, \dots, N$$

where we modify (5.1) by fixing $x \in \mathbb{R}^{2d} \cap \text{span}([1, -1, 1, \dots, -1]^\top)^\perp$ and using the projection operation $\Pi : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ given by

$$(5.4) \quad (\Pi x)[j] = x[j] + x[2d - 1 - j], \quad j = 1, 2, \dots, d.$$

We call the application of Π “projection” since it could be considered a discretization of the tomographic projection operation. In this context, we can think of entries of x as giving the density for patches of a thin-shell object. Projection then has the effect of summing two entries when they align on an imaging axis (see Figure 5.1). An effect of the projection operation is that we can only hope to recover x up to

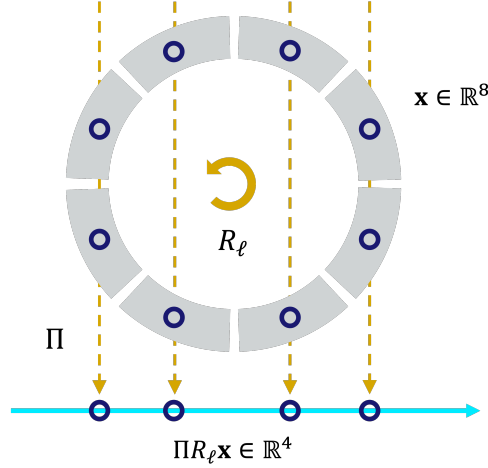


FIGURE 5.1. A visualization of the sampling procedure in the pMRA model as tomographic projection of some thin-shell 2D discretized object.

its $\mathcal{D}_{2d}$ -orbit instead of its $\mathcal{C}_{2d}$ -orbit. In fact, the model is unchanged if our group were dihedral instead of cyclic. The condition that the signal is orthogonal to $(1, -1, 1, \dots, -1)^\top$ ensures the problem is well-posed since the vector $(1, -1, 1, \dots, -1)^\top$ lies in the kernel of $\Pi \circ g$ for every $g \in \mathcal{D}_{2d}$. This vector represents the highest frequency component present in the signal, so excluding this component amounts to a “smoothness” assumption on the signal. Projection has been a key difficulty in extending methods developed for orbit recovery for the MRA model to a model of cryo-EM [50]. Projection has also been shown to cause issues with methods such as maximum likelihood estimation (MLE) when the probability distribution on group actions is misspecified [68].

Method of moments/invariants. In this work, we focus on the high noise regime where σ is much larger than a typical entry of x . The strategy of *synchronization* [9, 61, 62] uses comparisons between samples to label each with a unique group action (up to a global action), after which samples with the same label may be averaged to reduce noise. In the high noise regime, we cannot hope to estimate the group element associated with each sample by pairwise or even n -way comparisons [4]. Instead, one can aim to estimate features of the sample distribution that are invariant under the group action and then use these features to recover the signal’s orbit under the group action. We use *moment tensors* of (the signal part of) the sample distribution which are one class of invariant features (capturing all the polynomial invariants). These

moments are given by

$$(5.5) \quad T_{G,P}^{(r)}(x) := \mathbb{E}_{g \sim \text{Unif}(G)} [(P(g \cdot x))^{\otimes r}]$$

where in the dMRA model $G = \mathcal{D}_d$ and $P = I$, and in the pMRA model $G = \mathcal{C}_{2d}$ and $P = \Pi$. Whenever P is omitted in the subscript, it is assumed to be the identity $P = I$. Each $T_{G,P}^{(r)}(x)$ is an order- r tensor where each entry of the tensor is a homogeneous degree- r G -invariant polynomial of x . Each entry of the r -th order moment may be estimated to an accuracy δ with high probability using $O(\sigma^{2r} \log(d/\delta))$ samples [7, Proposition 6.7] using the empirical moments $\frac{1}{N} \sum_i y_i^{\otimes r}$ and some post-processing steps to remove bias caused by noise. Since the number of samples needed grows exponentially in the order of the tensor we wish to estimate, one might hope that for small R the tensors $\{T_{G,P}^{(r)}(x)\}_{r=1}^R$ *separate* the G -orbits of x , i.e., contain enough information to identify the orbit of x . Unfortunately, even for the original MRA model this does not hold; there exist signals $x \in \mathbb{R}^d$ that will require up to the d -th moment to separate orbits (see [10] for this result in a continuous MRA model and a similar analysis in [56] for the discrete MRA model). However, the signals of this type have Lebesgue measure zero. If we are satisfied with recovering “most” signals, we can instead ask how many moments are needed to uniquely identify the G -orbit for a *generic* signal, meaning we exclude some measure-zero set of “bad” signals (see Definition 5.1.1 for the formal definition). It has been shown in the MRA model that the first and second moment tensors are not sufficient to identify the orbit of x , but the third moment tensor separates the orbits of a generic signals. As a result, the sample complexity of MRA scales like $\Theta(\sigma^6)$ in the high-noise limit (where throughout this discussion we will ignore the logarithmic dependence on other parameters such as dimension and accuracy) [3, 10, 56]. Further, there is a fast algorithm to recover the orbit from the third moment tensor [56]. Beyond MRA, for more general orbit recovery problems, $\Theta(\sigma^{2r})$ samples are both necessary and sufficient, information-theoretically, where r is the number of moments needed to uniquely identify the signal [7].

Both the models we study are strictly harder than the original MRA model in the sense that the invariant tensors in the dMRA and pMRA models have entries which are linear combination of the entries found in the MRA invariant tensors; hence, one obtains less information from these tensors. One therefore needs at least three moments to achieve orbit recovery in both the pMRA and dMRA models. Recently, Edidin and Katz [31] proved that the first three moments do indeed separate generic $\mathcal{D}_d$ -orbits¹ so the sample complexity of dihedral MRA is $\Theta(\sigma^6)$. However, the proof method does not result in a computationally efficient algorithm to recover the orbit. For the pMRA model, even less was known prior to our work:

¹When we say that certain moments “separate generic orbits” we mean that for a generic x , the only vectors y that share the same moments as x are the elements of x ’s orbit.

Bandeira et al. showed that for a similar model [7, Problem 4.7] — in which projection is defined for odd-length signals — the system of polynomials defined by the first three moment tensors will have a finite number of solutions for problems of dimension 3 to 21, but again this does not yield an efficient algorithm for recovery.

In a more general setting that includes cryo-EM, we note that Balanov, Bendory, and Edidin [5] have shown that for an unknown n -dimensional object, the r -th moment of tomographic projection onto an m -dimension subspace determines the r -th moment of the un-projected object provided that $m \geq r$, leading to an algorithm that recovers certain un-projected moments from their projected counterparts. However, this does not handle the important case of $r = 3$ and $m = 2$ in cryo-EM, and does not seem to apply in the models we study here.

Our perspective. We work towards finding a polynomial-time algorithm that can provably recover orbits for generic signals in the dihedral and projected MRA models. In this work we suppose that we have access to exact third order moments $T_{\mathcal{D}_d}^{(3)}(x)$ or $T_{\mathcal{C}_{2d,\Pi}}^{(3)}(x)$, and we assume linear-algebraic computations can be done in exact arithmetic. While questions of robustness and noise stability are pertinent (e.g. [50]), we leave this direction for future work. We insist on using only moments of order ≤ 3 (and in fact the third moment alone will be enough for our purposes) in order to achieve optimal sample complexity in the high-noise limit.

Known algorithms and their limitations. From an algebraic perspective, the entries of $T_{G,P}^{(r)}(x)$ are degree- r homogeneous polynomials in the entries of the vector x , so orbit recovery may be achieved by solving a polynomial system of equations. In principle, the solutions may be found by computing a Gröbner basis with an ordering on the monomials to allow for elimination. In practice, this method is not efficient, and the worst-case time complexity of Gröbner basis computation is doubly exponential in the dimension d . We briefly detail two approaches developed in prior work on the MRA model that motivate our approach.

First, *frequency marching* may be used to achieve orbit recovery in the original MRA model [36, 53] and in *cryogenic electron tomography (cryo-ET)* [7, 50] — i.e., “unprojected” cryo-EM with 3-dimensional observations rather than 2-dimensional projections — using the first three moments of the sample distribution. In both models, the second moment determines the magnitudes of the Fourier or Fourier–Bessel coefficients, also called the *power spectrum*. To recover phases of these coefficients, the third moment of the distribution may be used to set up a linear system of equations modulo 2π for the coefficients’ phases. The phases of coefficients may then be solved for sequentially, in a procedure called *frequency marching* referring to the fact that one initializes a phase for the first Fourier coefficient (corresponding to a low frequency

component of signal) then then uses relationships between phases to determine phases of subsequent Fourier coefficients (higher frequencies). Bendory et al. [12] give a version of frequency marching for the MRA problem and compare its performance to a number of other algorithms. Liu and Moitra [50] modify the frequency marching procedure to achieve $SO(3)$ -orbit recovery in cryo-ET in quasipolynomial time.

However, Edidin and Katz [31] showed that adapting the frequency marching method to dihedral MRA, one can only determine each linear equation needed to march frequencies up to an unknown sign. The ambiguity in sign leads to a combinatorial explosion in which there are exponentially many linear systems, only one of which is guaranteed to have the true x as a solution. While in theory solving the right system recovers the orbit, the runtime of enumerating all systems is impractical.

Second, *tensor decomposition* methods have been shown to achieve orbit recovery in the MRA model [56]. For $x \in \mathbb{R}^d$ the third moment tensor takes the form

$$T_{\mathcal{C}_d}^{(3)}(x) = \sum_{\ell=0}^{d-1} (R_\ell x)^{\otimes 3},$$

which is the sum of d different $d \times d \times d$ rank-1 tensors. Unlike a matrix factorization, the factorization of a tensor into its rank-1 summands (called its *canonical polyadic (CP) decomposition*) does not require the component vectors to be orthogonal for a unique decomposition [48]. Perry et al. [56] showed that *simultaneous diagonalization*, a classical tensor decomposition algorithm, may be used to factor the third moment in a robust way to recover the components $\{R_\ell x\}_{\ell=0}^{d-1}$. Crucially, simultaneous diagonalization requires the rank (number of rank-1 summands to recover) to be no larger than the dimension — in the case of MRA, this holds with equality. However, the third moments arising from the models we study take the form

$$T_{\mathcal{D}_d}^{(3)}(x) = \sum_{g \in \mathcal{D}_d} (g \cdot x)^{\otimes 3} \quad \text{and} \quad T_{\mathcal{C}_{2d}, \Pi}^{(3)}(x) = \sum_{g \in \mathcal{C}_{2d}} (\Pi(g \cdot x))^{\otimes 3}.$$

Each is the sum of $2d$ different $d \times d \times d$ rank-1 tensors, which is more components than simultaneous diagonalization can handle. The best known methods for generic tensor decomposition can reach a rank of $(2 - \epsilon)d$ [47] (see also [22, 45, 57]), which barely falls short of our $2d$. This means we cannot use known tensor methods “out of the box” and must take advantage of the special structure in our tensors to when devising an efficient algorithm.

We also note that methods for *fourth-order* tensor decomposition can handle significantly more components, namely $\Theta(d^2)$ [26, 41, 43, 52]. However, estimating the fourth moment requires sub-optimal sample complexity, so our focus in this work is on methods that only use the third moment and below.

Our contributions. Our main results partially resolve the open problem of orbit recovery in the dihedral and projected MRA models.

We first formally define the algebraic idea of generic vectors.

DEFINITION 5.1.1. *A property Q is said to “hold generically” for vectors $v \in \mathbb{K}^m$ (or to “hold for generic choice of” $v \in \mathbb{K}^m$) if there is a non-zero polynomial p on $\mathbb{K}^m$ such that when $p(v) \neq 0$, v has property Q .*

Next, the results are as follows:

THEOREM 5.1.2 (Dihedral MRA). *Let $d = 2^k$, for $k \geq 2$ and suppose a particular symbolic matrix $M = M(d)$ (see (5.41)) has full column-rank. Then there is a polynomial-time procedure that, for generic $x \in \mathbb{R}^{2d}$, takes as input the exact third moment $T_{\mathcal{D}_{2d}}^{(3)}(x)$ and outputs an element in the $\mathcal{D}_{2d}$ -orbit of x .*

THEOREM 5.1.3 (Projected MRA). *Let $d = 2^k$, for $k \geq 3$, and suppose a particular symbolic matrix $M' = M'(d)$ (see (5.42)) has full column-rank. Then there is a polynomial-time procedure that, for generic $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$, takes as input the exact third moment $T_{\mathcal{C}_{2d}, \Pi}^{(3)}(x)$ and outputs an element in the $\mathcal{D}_{2d}$ -orbit of x .*

For any specific dimension d , one can use a computer to verify that the symbolic matrices $M(d)$ and $M'(d)$ have full column-rank, implying that the algorithm works for generic inputs of that size. This check only needs to be done once and ensures our algorithm then works on generic inputs. Each of these matrices has more rows than columns under the conditions given, and both classes of matrices have an aspect ratio (rows : columns) tending toward 8 : 3 as d goes to infinity. We ran numerical tests giving strong evidence that these matrices have full column-rank numerically for a random initialization in the dimensions $d = 4, 8$ and 16. This gives evidence that the matrices have full column-rank symbolically and that each procedure succeeds in recovering orbits for generic vectors. Further, we provide an implementation of the algorithms in Python and demonstrate their efficacy by recovering signals in $\mathbb{R}^{64}$ and $\mathbb{R}^{64} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ respectively. In both cases the algorithms recover the signal up to cyclic shift and reversal with high accuracy from the moments described. These numerical tests and further details may be found in Section 5.7.1.

The dimension $2d$ must be a power of two since our method uses recursion to extract and solve orbit recovery sub-problems. We first recover the even Fourier coefficients of x by recursively running the algorithm on a sub-tensor. We can then recover the odd Fourier coefficients given the even ones using a different sub-tensor. While this method is reminiscent of frequency marching, we solve for multiple coefficients simultaneously by setting up a *variety-constrained linear system*, a method that synthesizes the work of Boussé et al. [19] on CPD-constrained linear systems and Johnston et al. [43] on efficiently finding the intersection of

a subspace and a conic variety. Unlike frequency marching, our methodology is able to handle the projection operation in the projected MRA model.

Further discussion and related work. A significant body of literature studying orbit recovery exists. We highlight a couple of key problems not addressed in this work that are particularly interesting practically and theoretically, especially with respect to cryo-EM.

In cryo-EM, often we cannot guarantee that all the particles in a sample are identical. Either there may be gross differences, e.g., a protein-complex missing a sub-unit, or conformational differences where proteins flex or bend slightly. Both situations present a form of signal *heterogeneity*. In the simplest form, one assumes a mixture of K different signals. Including heterogeneity in an MRA model as in [56], we assume that instead of one distinct signal x there are K different signals that each appear with some probability. Boumal et al. [18] show that fitting the first three moments, heterogeneity $K > \lceil d/6 \rceil$ causes the problem to be ill-posed but a non-convex optimization approach seems to handle up to (approximately) $K \leq \sqrt{d}$ different signals, when the signals are generated randomly (a more restrictive setting than that of generic signals). Continuous heterogeneity is a considerably different challenge to handle; see [65] for the current state of the art. Heterogeneity is related to the problem of *super-resolution* in MRA [14].

Another reality is that due to boundary effects in the ice layer, particle orientation in cryo-EM may not be uniformly distributed [59]. One can make a similar assumption that some cyclic shifts are more or less likely in the MRA model. If the distribution of group elements is unknown, it must also be estimated. Interestingly, in some orbit recovery models, non-uniformity of the group elements allows for generic recovery using lower order moments than in the uniform case. For example, Abbe et al. [2] give an exact method for recovering the signal and non-uniform distribution of cyclic shifts for the MRA model using the *second* moment rather than the third. Bendory et al. [13] give an information-theoretic proof that the second moment suffices in the dihedral MRA model with a generic distribution on group elements. Similarly for cryo-EM, results by Sharon et al. [59] show that the second moment suffices for accurate estimation, and determine bandlimit conditions on molecular structure and the distribution over viewing angles which ensure well-posedness.

There is also a significant body of research on non-convex methods for orbit recovery, e.g. [18, 32, 33, 68]. Notably, fitting moments by least squares is demonstrated by Edidin and Katz [31] as a practical method for recovery in dihedral MRA, albeit one without guarantees of correctness. In MRA with a non-uniform distribution, Abas et al. [1] develop a generalized method of moments which determines how to re-weight a least-squares fitting of moments to be asymptotically optimal and demonstrate this improves recovery of

the signal and distribution as parameters. Other directions explore maximum likelihood estimation (MLE) as a recovery technique. Xu et al. [68] observe in cryo-EM that misspecifying the probability distribution (e.g., assuming a uniform distribution when viewing angles are non-uniform) does not have a deleterious effect on the likelihood landscape in *unprojected* cryo-EM (cryo-ET) but does result in incorrect estimation when projection is present. Due to this effect, they propose an expectation maximization (EM) algorithm for estimating the signal and viewing-angle distribution jointly. To our knowledge, none of the methods based on non-convex optimization provide a rigorous proof that the global optimum can be found in polynomial time. As a result, this line of work deviates from our primary objective of finding algorithms with provable guarantees.

Concurrent work. Finally, we point out the concurrent and independent work [6], which also studies a variation of the projected MRA problem (with odd-length signals) and gives an algorithm for recovery from the third moment. Their algorithmic approach is rather different from ours: essentially they have adapted the frequency marching method in a way that avoids the combinatorial explosion we mentioned above. Their algorithm appears to also work for dihedral MRA.

Organization. In Section 5.2, we define the dMRA and pMRA models in the Fourier basis and give explicit expressions for the invariants that result. Section 5.3 covers the steps of our algorithm in detail. In Section 5.4, we give sufficient conditions for the recursive step to succeed and give evidence that these conditions hold for generic signals. Since our algorithm is recursive, we discuss how orbit recovery may be achieved for the base case in Section 5.5. In Section 5.7 we give an efficient implementation of the recursive step and details for an implementation in Python of the full algorithm. Finally, in Section 5.8 we discuss future directions.

5.2. Dihedral MRA and MRA with Projection

In this section, we define and compare the two main models of interest in this work: dihedral MRA (dMRA) and MRA with projection (pMRA).

Dihedral MRA. In the dMRA model, we observe N samples according to

$$(5.6) \quad y_i = g_i \cdot x + \sigma \xi_i$$

where x is some fixed signal in $\mathbb{R}^d$, each g_i is a group element of the dihedral group $\mathcal{D}_d$, each ξ_i is independent standard Gaussian noise in $\mathbb{R}^d$, and $\sigma > 0$ is a scalar that determines the signal-to-noise ratio. We will index entries in x starting at zero so that $x = (x[0], x[1], \dots, x[d-1])$.

The dihedral group $\mathcal{D}_d$ may be generated from cyclic shifts and reflections. We will denote these R_ℓ and J respectively with $\mathcal{D}_d = \{R_0, R_1, \dots, R_{d-1}, JR_0, JR_1, \dots, JR_{d-1}\}$. For $\ell \in [d]$, we define the action of $R_\ell \in \mathcal{D}_d$ to be a cycling of entries in x to the right by ℓ positions so that for each index $j \in [d]$,

$$(R_\ell \cdot x)[j] = x[j - \ell \pmod{d}].$$

In general, we will assume that any indexing we do is done cyclically modulo the length of the vector involved so we may omit the modulus. For $J \in \mathcal{D}_d$, we define its action to be reversal of entries in a vector so that for an index $j \in [d]$,

$$(J \cdot x)[j] = x[d - j - 1].$$

Using the samples $\{y_i\}_{i=1}^N$, one aims to recover the signal x up to the action of $\mathcal{D}_d$.

MRA with projection. Our second model is MRA with projection (pMRA). In the pMRA model, we observe N samples according to

$$(5.7) \quad y_i = \Pi(g_i \cdot x) + \sigma \xi_i.$$

Here now x is a fixed signal in $\mathbb{R}^{2d}$, g_i are group elements of the cyclic group $\mathcal{C}_{2d} = \{R_0, R_1, \dots, R_{2d-1}\}$, and again ξ_i is independent standard Gaussian noise (now in $\mathbb{R}^{2d}$) with $\sigma > 0$ a scalar determining the signal-to-noise ratio. We also define the linear operator, $\Pi : \mathbb{C}^{2d} \rightarrow \mathbb{C}^d$, defined entry-wise by for $j \in [d]$ by

$$(5.8) \quad (\Pi x)[j] = x[j] + x[2d - j - 1].$$

This projection operation is a proxy for tomographic projection or the X-ray transform in the sense that if we view the object as being the discretization of some “thin-shell” circular 2D object with $x[j]$ the density in the j -th patch, projection sums two entries when they are align vertically (see Figure 5.1).

In this model, one might aim to recover the signal x up to the action of $\mathcal{D}_{2d}$ from the projections $\{y_i\}_{i=1}^N$. Note that while the group actions are cyclic, we can only hope recover the signal up to action by the dihedral group. Indeed, a signal x and its reflection Jx will produce the same distribution over samples. In this manner, the pMRA model captures the same ambiguity that occurs in tomography; when the orientations in three dimensions are unknown, one can only distinguish an object up to reflection. However, in another regard the model suffers from a small issue. Complete recovery for a signal is impossible since the vector $c = (1, -1, 1, -1, \dots, -1)$ lies in the kernel of ΠR_ℓ for every $\ell \in [d]$. This component is the *Nyquist component* and under the Fourier transform corresponds to the d -th Fourier coefficient of x . Since this component is of highest frequency, the inability to recover this component could be taken as akin to an

optical or electron-imaging system having some limited resolving power. In light of this, we refine the goal: using samples $\{y_i\}_{i=1}^N$, one aims to recover all Fourier coefficients of the signal x except the d -th one, up to the action of $\mathcal{D}_{2d}$.

Advantages of the Fourier Basis. The action of $\mathcal{C}_{2d}$ on $\mathbb{C}^{2d}$ is diagonalized in the Fourier basis and gives the standard representation for the cyclic group. However, $\mathcal{D}_{2d}$ is not a commutative group, the actions of its elements cannot be simultaneously diagonalized, and the matrices giving the action in the Fourier basis are all sparse with non-zero entries only on the diagonal or anti-diagonal.

For a signal $x \in \mathbb{R}^d$, denote the Fourier coefficients of x indexed from zero, by

$$\hat{x}[k] := \sum_{j=0}^{d-1} x[j] \cdot e^{-2\pi i j k/d}$$

where $i = \sqrt{-1}$. Define

$$\hat{x} := (\hat{x}[k] : k = [d]) \in \mathbb{C}^d.$$

Implicitly $\hat{x} = \mathcal{F}_d x$ where $\mathcal{F}_d : \mathbb{C}^{d \times d}$ is a linear transformation giving the change of basis from the standard basis into the Fourier basis in $\mathbb{C}^d$. In addition, since x is a real signal, we have that $\hat{x}$ satisfies $\hat{x}[-k] = \overline{\hat{x}[k]}$ with indices taken modulo d so that they lie in $[d]$. Next, define the following matrices $\hat{R}_{\ell,d} \in \mathbb{C}^{d \times d}$, $\ell \in [d]$, and $\hat{J}_d \in \mathbb{C}^{d \times d}$, where for $v \in \mathbb{C}^d$ with entries indexed by $j \in [d]$,

$$(5.9) \quad (\hat{R}_{\ell,d} \cdot v)[j] = e^{-2\pi i \ell j/(2d)} v[j],$$

$$(5.10) \quad (\hat{J}_d \cdot v)[j] = v[-j],$$

and the matrix $\hat{\Pi}_{2d} \in \mathbb{C}^{d \times 2d}$, where for $v \in \mathbb{C}^{2d}$,

$$(5.11) \quad (\hat{\Pi}_{2d} v)[j] = v[j] + e^{2\pi i j/2d} v[-j].$$

We will omit the subscript dimensions d and $2d$ where it may be inferred from context.

5.2.1. Invariants from the dMRA and pMRA Models. Using the linear operators defined in (5.9)-(5.11) we define the following invariant tensors of a complex vector $\hat{x} \in \mathbb{C}^{2d}$,

$$(5.12) \quad T_{\text{dMRA},2d}^{(r)}(\hat{x}) := \frac{1}{4d} \left(\sum_{\ell=0}^{2d-1} (\hat{R}_{\ell,2d} \cdot \hat{x})^{\otimes r} + \sum_{\ell=0}^{2d-1} (\hat{J} \hat{R}_{\ell,2d} \cdot \hat{x})^{\otimes r} \right),$$

$$(5.13) \quad T_{\text{pMRA},2d}^{(r)}(\hat{x}) := \frac{1}{2d} \left(\sum_{\ell=0}^{2d-1} (\hat{\Pi}_{2d} \hat{R}_{\ell,2d} \cdot \hat{x})^{\otimes r} \right).$$

As suggested by our notation, we claim that these tensors are the same invariants from the dMRA and pMRA models given by (5.5) under a change of basis. First we give the following lemma.

LEMMA 5.2.1. *For any d ,*

$$\hat{R}_\ell = \mathcal{F} R_\ell \mathcal{F}^{-1},$$

for any $\ell \in [d]$,

$$\hat{J} = \mathcal{F} J R_{-1} \mathcal{F}^{-1},$$

and there is a full-rank square matrix C giving a change of basis such that

$$\hat{\Pi}_{2d} = C \Pi_{2d} \mathcal{F}^{-1}.$$

Using this we have the following result.

PROPOSITION 5.2.2. *Let $x \in \mathbb{R}^{2d}$ and let $T_{\mathcal{D}_{2d}}^{(r)}(x)$ and $T_{\mathcal{C}_{2d}, \Pi}^{(r)}(x)$ be r -order invariant tensor of x as defined in (5.5). For $\mathcal{F} \in \mathbb{C}^{2d \times 2d}$ the matrix giving the change of basis from the standard basis into the Fourier basis in $\mathbb{C}^{2d}$,*

$$T_{\text{dMRA}, 2d}^{(r)}(\mathcal{F}x) = \mathcal{F}^{\otimes r} \left(T_{\mathcal{D}_{2d}}^{(r)}(x) \right).$$

For C the change-of-basis matrix in Lemma 5.2.1,

$$T_{\text{pMRA}, 2d}^{(r)}(\mathcal{F}x) = C^{\otimes r} \left(T_{\mathcal{C}_{2d}, \Pi}^{(r)}(x) \right).$$

The subscript $2d$ may be dropped when the dimension is clear from context. See Section 5.9.1 for the proof of Lemma 5.2.1 and Proposition 5.2.2, including the explicit construction of the change-of-basis matrix C . Lemma 5.2.1 is also useful in proving the following fact.

PROPOSITION 5.2.3. *In the pMRA model, two projections $\{\Pi R_{\ell_1} x, \Pi R_{\ell_2} x\}$ from known directions $\ell_1 \neq \ell_2$ are sufficient to recover $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, 1, \dots, -1)^\perp$ when $\gcd(\ell_1 - \ell_2, 2d) = 1$.*

See Section 5.9.1 for the proof. As we mentioned, strategies to label samples provably fail in the high-noise regime; however, Proposition 5.2.3 demonstrates the irony of the pMRA model: the third moment is an overcomplete tensor and we do not have an algorithm to decompose it; we also do not need all the factors from its decomposition, just two! Next, one may wonder why we give our dMRA invariant tensor in (5.12) for a signal of length $2d$ instead of d . By defining both models with a common dimension we may give a relationship between the invariants:

LEMMA 5.2.4. For $\hat{x} \in \mathbb{C}^{2d}$,

$$(\hat{\Pi}_{2d}^{\otimes r})T_{\text{dMRA}, 2d}^{(r)}(\hat{x}) = T_{\text{pMRA}, 2d}^{(r)}(\hat{x}).$$

PROOF. By linearity and Lemma 4.2.4 we may move the projection operator into the summands

$$(\hat{\Pi}^{\otimes r})T_{\text{dMRA}}^{(r)}(\hat{x}) = \frac{1}{4d} \left(\sum_{\ell=0}^{2d-1} (\hat{\Pi}\hat{R}_{\ell}\hat{x})^{\otimes r} + \sum_{\ell=0}^{2d-1} (\hat{\Pi}\hat{J}\hat{R}_{\ell}\hat{x})^{\otimes r} \right).$$

Substituting the relations from Proposition 5.2.1 we have that

$$\hat{\Pi}\hat{J} = C\Pi JR_{-1}\mathcal{F}^{-1} = C\Pi R_{-1}\mathcal{F}^{-1} = C\Pi\mathcal{F}^{-1}\hat{R}_{-1} = \hat{\Pi}\hat{R}_{-1}$$

since $\Pi J = \Pi$. If we re-index the second sum we may combine, yielding the result:

$$(\hat{\Pi}^{\otimes r})T_{\text{dMRA}}^{(r)}(\hat{x}) = \frac{1}{4d} \left(\sum_{\ell=0}^{2d-1} (\hat{\Pi}\hat{R}_{\ell}\hat{x})^{\otimes r} + \sum_{\ell=0}^{2d-1} (\hat{\Pi}\hat{R}_{\ell-1}\hat{x})^{\otimes r} \right) = \frac{1}{2d} \sum_{\ell=0}^{2d-1} (\hat{\Pi}\hat{R}_{\ell}\hat{x})^{\otimes r} = T_{\text{pMRA}}^{(r)}(\hat{x}). \quad \square$$

As a result, the $\mathcal{D}_{2d}$ -invariant polynomials that make up the entries of the $T_{\text{pMRA}}^{(r)}(\hat{x})$ tensor are linear combinations of the $\mathcal{D}_{2d}$ -invariant polynomials in $T_{\text{dMRA}}^{(r)}(\hat{x})$. In other words, the invariant dMRA tensor contains at least as much information as the pMRA tensor. The advantage of our change of basis by the Fourier transform is apparent in computations giving each entry in the invariant tensors for both models.

PROPOSITION 5.2.5. For $\hat{x} \in \mathbb{C}^{2d}$,

$$T_{\text{dMRA}}^{(r)}(\hat{x})[\mathfrak{I}j_1, \dots, j_r] = \frac{w_{\mathfrak{I}j_1, \dots, j_r}}{2} \left(\prod_{k=1}^r \hat{x}[j_k] + \prod_{k=1}^r \hat{x}[-j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r j_k \equiv 0 \pmod{2d} \right\}$$

for $\mathfrak{I}j_1, \dots, j_r$ a multiset of $[2d]$.

For the same signal,

$$T_{\text{pMRA}}^{(r)}(\hat{x})[\mathfrak{I}j_1, \dots, j_r] = \frac{w_{\mathfrak{I}j_1, \dots, j_r}}{2} \sum_{s \in \{-1, 1\}^r} \phi(s, \mathfrak{I}j_1, \dots, j_r) \left(\prod_{k=1}^r \hat{x}[s_k j_k] + \prod_{k=1}^r \hat{x}[-s_k j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r s_k j_k \equiv 0 \pmod{2d} \right\}$$

for $\mathfrak{I}j_1, \dots, j_r$ a multiset of $[d]$ and where ϕ is a complex phase given by

$$\phi(s, \mathfrak{I}j_1, \dots, j_r) := (-1)^{\frac{1}{2d} \sum_{k=1}^r s_k j_k} \exp \left(\frac{\pi i}{2d} \sum_{k=1}^r j_k \right).$$

See Section 5.9.2 for the computations that verify these facts. Observe that each entry of the $T_{\text{dMRA}}^{(r)}(\hat{x})$ tensor is either zero or simply the sum of two monomials up to scale. Further, it is relatively simple to verify these polynomials are invariant under the actions given in (5.9) and (5.10). In contrast, in the standard

basis the entries of $T_{\mathcal{D}_{2d}}^{(r)}(x)$ and $T_{\mathcal{C}_{2d}, \Pi}^{(r)}(x)$ will be the sum of many monomials. We can tie sparsity directly to computational efficiency later and it also eases several future constructions.

5.3. Method for Recovery

In this section, we first define generally the conditions which allow us to extract a subtensor of an invariant tensor that is also an invariant tensor for some model. We demonstrate that this is possible for both the dMRA and pMRA invariant tensors and that the procedure produces smaller dMRA and pMRA invariant tensors such that an orbit recovery problem's dimension may be recursively halved. First, we introduce some notation we use to indicate extracting subtensors.

Restriction and subtensors. Given a vector $x \in \mathbb{K}^n$ and a set of indices $S \subseteq [n]$ we define the restriction of x to S using either the notation $x|_S$ or $x[S] := (x[s] : s \in S) \in \mathbb{K}^{|S|}$. For a matrix $A \in \mathbb{K}^{m \times n}$, $R \subseteq [m]$, and $S \subseteq [n]$ subsets. We will use $A[R, S]$ to denote the submatrix of A with rows in R and columns in S . When $R = [m]$ or $S = [n]$ we will use $A[:, S]$ or $A[R, :]$ respectively.

Similarly, for a tensor $T \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2} \otimes \cdots \otimes \mathbb{R}^{d_r}$ and subsets $S_j \subseteq [d_j]$, $j = 1, \dots, r$ we use the following to denote restriction along multiple modes:

$$T|_{S_1, S_2, \dots, S_r} \text{ or } T[S_1, S_2, \dots, S_r] := [T[s_1, s_2, \dots, s_r] : s_1 \in S_1, s_2 \in S_2, \dots, s_r \in S_r] \in \mathbb{K}^{|S_1| \times \cdots \times |S_r|}$$

is a subtensor of $\mathcal{T}$. For symmetric tensors $\mathcal{T} \in (\mathbb{R}^n)^{\otimes r}$ we may use a single restriction to indicate restriction on all modes by the same subset of indices:

$$T|_S := [T[s_1, s_2, \dots, s_r] : s_1, s_2, \dots, s_r \in S] \in (\mathbb{K}^{|S|})^{\otimes r},$$

or when the subsets $\{S_i\}_{i=1}^r$ are pairwise equal or disjoint, then the subtensor of T defined by

$$T[S_1, S_2, \dots, S_r] := [T[s_1, s_2, \dots, s_r] : s_1 \in S_1, s_2 \in S_2, \dots, s_r \in S_r]$$

is partially symmetric along modes where the subsets S_j and S_k are equal for $j, k \in [r]$.

We quickly give some implications for restriction on matrix-vector products and matrix-matrix products, that are handy.

LEMMA 5.3.1. *Let $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{m \times n}$ be matrices and $v \in \mathbb{K}^n$ be a vector. Let $R \subseteq [m]$, and $S \subseteq [n]$ be subsets of indices. Then*

$$(Av)[R] = A[R, :]v \quad \text{and} \quad (AB)[R, S] = A[R, :]B[:, S].$$

This follows directly from our definition of restriction and matrix multiplication.

5.3.1. Invariant Subtensors.

LEMMA 5.3.2 (Moment subtensors from restriction). *Let G be a finite group with some action on $\mathbb{K}^m$, $h : \mathbb{K}^m \rightarrow \mathbb{K}^n$ be a linear transformation, and $R \subseteq [m]$, $S \subseteq [n]$ both be sets of indices. If $\tilde{G}$ is a finite group with some action on $\mathbb{K}^{|R|}$ and $\tilde{h} : \mathbb{K}^{|R|} \rightarrow \mathbb{K}^{|S|}$ a linear transformation such that for all $v \in \mathbb{K}^m$,*

$$(5.14) \quad \mathcal{I}(h(g \cdot v))|_S : g \in G = k \times \mathcal{I}(\tilde{h}(\tilde{g} \cdot v|_R) : \tilde{g} \in \tilde{G}),$$

where for a multiset M , $k \times M$ indicates a multiset with each item in M repeated k times. Then restricting each mode of a G -invariant tensor to S produces a $\tilde{G}$ -invariant tensor under the projection $\tilde{h}$ up to scale,

$$\left(T_{G,h}^{(r)}(v)\right)|_S = k \frac{|\tilde{G}|}{|G|} T_{\tilde{G},\tilde{h}}^{(r)}(v|_R).$$

PROOF. Proof of Lemma 5.3.2. First,

$$\left[T_{G,h}^{(r)}(v)\right]|_S = \frac{1}{|G|} \sum_{g \in \mathcal{G}} [(h(g \cdot v))^{\otimes r}]|_S = \frac{1}{|G|} \sum_{g \in \mathcal{G}} (h(g \cdot v)|_S)^{\otimes r}$$

by our definition of restriction for a symmetric tensor. Next, applying the multiset condition we have

$$\frac{1}{|G|} \sum_{g \in \mathcal{G}} (h(g \cdot v)|_S)^{\otimes r} = \frac{k}{|G|} \sum_{\tilde{g} \in \tilde{\mathcal{G}}} (\tilde{h}(\tilde{g} \cdot v|_R))^{\otimes r},$$

so with some manipulation,

$$\frac{k}{|G|} \sum_{\tilde{g} \in \tilde{\mathcal{G}}} (\tilde{h}(\tilde{g} \cdot v|_R))^{\otimes r} = \frac{k|\tilde{G}|}{|G|} \cdot \frac{1}{|\tilde{G}|} \sum_{\tilde{g} \in \tilde{\mathcal{G}}} (\tilde{h}(\tilde{g} \cdot v|_R))^{\otimes r} = \frac{k|\tilde{G}|}{|G|} T_{\tilde{G},\tilde{h}}^{(r)}(v|_R),$$

$$\text{and } \left[T_{G,h}^{(r)}(v)\right]|_S = k \frac{|\tilde{G}|}{|G|} T_{\tilde{G},\tilde{h}}^{(r)}(v|_R). \quad \square$$

The property above is useful since when it holds, a particular subtensor is actually an invariant tensor in its own right and — if $k > 1$ — is an invariant over a smaller group.

We now show that the invariant tensors from the dMRA and pMRA models both satisfy Lemma 5.3.2 when restricting to even Fourier coefficients. To denote the even numbers in $[d]$ we use

$$(5.15) \quad E_d = \{j \in [d] : j \equiv 0 \pmod{2}\}.$$

We will omit the subscript and simply write “ E ” if the dimension can be inferred from context. We also use $\hat{\Pi}_{2d} = \hat{\Pi}$ to indicate the projection operation defined in (5.11) and $\hat{\Pi}_{4d}$ an analogous operator defined in the same manner but with the domain $\mathbb{C}^{4d}$.

PROPOSITION 5.3.3. *For $d \geq 1$, the following invariant tensors have invariant subtensors found by restriction:*

(1) For $\hat{x} \in \mathbb{C}^{2d}$,

$$\left(T_{\text{dMRA}, 2d}^{(r)}(\hat{x}) \right) \Big|_{E_{2d}} = T_{\text{dMRA}, d}^{(r)}(\hat{x}|_{E_{2d}}).$$

(2) For $\hat{x} \in \mathbb{C}^{4d}$,

$$\left(T_{\text{pMRA}, 4d}^{(r)}(\hat{x}) \right) \Big|_{E_{2d}} = T_{\text{pMRA}, 2d}^{(r)}(\hat{x}|_{E_{4d}}).$$

PROOF. For each case we show that the multiset condition (5.14) holds so that Lemma 5.3.2 applies. Since we deal with the cyclic group and dihedral group over vector spaces of different sizes we will include the dimension in the subscript on elements, i.e. $\mathcal{C}_d = \{\hat{R}_{\ell, d}\}_{\ell=1}^{d-1}$ where

$$(\hat{R}_{\ell, d} \cdot \hat{x})[j] = e^{-2\pi i j \ell / d} x[j]$$

for $j \in [d]$. Similarly, we will use $\hat{J}_d$ for the reflection element on $\mathbb{C}^d$.

(1) Let $\hat{x} \in \mathbb{C}^{2d}$ be any vector. Given $m \in [d]$ we define $\ell_k = m + dk \in [2d]$ for $k \in \{0, 1\}$. We claim that for each k , $(\hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_{2d}} = \hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}})$ and $(\hat{J}_{2d} \cdot \hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_{2d}} = \hat{J}_d \cdot \hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}})$ so the multiset condition holds. To this end, notice that since $\ell_k = m + kd$,

$$e^{-2\pi i \ell_k / d} = e^{-2\pi i m / d} \implies$$

$$\text{for any } j \in [d], \quad e^{-2\pi i j \ell_k / d} = e^{-2\pi i j m / d} \implies$$

$$e^{-2\pi i j \ell_k / d} \hat{x}[2j] = e^{-2\pi i j m / d} \hat{x}[2j] \implies$$

$$e^{-2\pi i (2j) \ell_k / (2d)} \hat{x}[2j] = e^{-2\pi i j m / d} (\hat{x}|_{E_{2d}})[j] \implies$$

$$(\hat{R}_{\ell_k, 2d} \cdot \hat{x})[2j] = (\hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}}))[j] \implies$$

$$\left((\hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_{2d}} \right)[j] = \left(\hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}}) \right)[j],$$

and so $(\hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_d} = \hat{R}_{m, d} \cdot (\hat{x}|_{E_d})$. Similarly,

$$\begin{aligned}
e^{2\pi i \ell_k / d} &= e^{2\pi i m / d} && \implies \\
\text{for any } j = [d], \quad e^{2\pi i j \ell_k / d} &= e^{2\pi i j m / d} && \implies \\
e^{2\pi i j \ell_k / d} \hat{x}[-2j] &= e^{2\pi i j m / d} \hat{x}[-2j] && \implies \\
e^{-2\pi i (-2j) \ell_k / (2d)} \hat{x}[-2j] &= e^{-2\pi i (-j) m / d} (\hat{x}|_{E_{2d}})[-j] && \implies \\
(\hat{J}_{2d} \cdot \hat{R}_{\ell_k, 2d} \cdot \hat{x})[2j] &= \left(\hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}}) \right) [-j] && \implies \\
\left((\hat{J}_{2d} \cdot \hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_{2d}} \right) [j] &= \left(\hat{J}_d \cdot \hat{R}_{m, d} \cdot (\hat{x}|_{E_{2d}}) \right) [j],
\end{aligned}$$

and so $(\hat{J}_{2d} \cdot \hat{R}_{\ell_k, 2d} \cdot \hat{x})|_{E_d} = \hat{J}_d \cdot \hat{R}_{m, d} \cdot (\hat{x}|_{E_d})$. This completes the association of elements between the sets so that

$$\wr(g \cdot \hat{x})|_{E_d} : g \in \mathcal{D}_{2d} \wr = 2 \times \wr(\tilde{g} \cdot \hat{x}|_{E_d}) : \tilde{g} \in \mathcal{D}_d \wr.$$

(2) Let $\hat{x} \in \mathbb{C}^{4d}$. Set $m \in [2d]$ and let $\ell_k = m + (2d)k \in [4d]$ for $k \in \{0, 1\}$. We aim to show that

$$\left(\hat{\Pi}_{4d}(\hat{R}_{\ell_k, 4d} \cdot \hat{x}) \right) \Big|_{E_{2d}} = \hat{\Pi}_{2d} \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right).$$

Reusing analysis from the first part, we know that $e^{-2\pi i j \ell_k / (2d)} \hat{x}[2j] = e^{-2\pi i j m / (2d)} \hat{x}[2j]$ for all $j \in [2d]$. Rewriting this, we have that $e^{-2\pi i j \ell_k / (2d)} \hat{x}[2j] = e^{-2\pi i (2j) \ell_k / (4d)} \hat{x}[2j] = (\hat{R}_{\ell_k, 4d} \cdot \hat{x})[2j]$ and that $e^{-2\pi i j m / (2d)} \hat{x}[2j] = e^{-2\pi i j m / (2d)} (\hat{x}|_{E_{4d}})[j] = \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right) [j]$ for all $j \in [2d]$.

Now we take linear combinations to produce the d different equations

$$e^{-2\pi i j \ell_k / (2d)} \hat{x}[2j] + e^{2\pi i j / (2d)} e^{2\pi i j \ell_k / (2d)} \hat{x}[-2j] = e^{-2\pi i j m / (2d)} \hat{x}[2j] + e^{2\pi i j / (2d)} e^{2\pi i j m / (2d)} \hat{x}[-2j]$$

with $j \in [d]$. Applying the rewritings gives

$$(\hat{R}_{\ell_k, 4d} \cdot \hat{x})[2j] + \underbrace{e^{2\pi i j / (2d)}}_{= e^{2\pi i (2j) / (4d)}} (\hat{R}_{\ell_k, 4d} \cdot \hat{x})[-2j] = \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right) [j] + e^{2\pi i j / (2d)} \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right) [-j].$$

We then have that by the definition of the $\hat{\Pi}$ operator that

$$\begin{aligned}
\left(\hat{\Pi}_{4d}(\hat{R}_{\ell_k, 4d} \cdot \hat{x}) \right) [2j] &= \left(\hat{\Pi}_{2d} \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right) \right) [j] && \implies \\
\left(\left(\hat{\Pi}_{4d}(\hat{R}_{\ell_k, 4d} \cdot \hat{x}) \right) \Big|_{E_{2d}} \right) [j] &= \left(\hat{\Pi}_{2d} \left(\hat{R}_{m, 2d} \cdot (\hat{x}|_{E_{4d}}) \right) \right) [j].
\end{aligned}$$

Then

$$\wr(\hat{\Pi}_{4d}(g \cdot \hat{x}))|_{E_d} : g \in \mathcal{C}_{4d} \wr = 2 \times \wr(\hat{\Pi}_{2d}(\tilde{g} \cdot (\hat{x}|_{E_{2d}})) : \tilde{g} \in \mathcal{C}_{2d} \wr. \quad \square$$

Proposition 5.3.3 shows that we can recover part of the signal by extracting a subtensor and then solving (recursively) the orbit recovery problem associated to it. Once we recover an element y in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$, our next goal will be to extend y to a vector double its length that lies in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$.

5.3.2. Variety-Constrained Linear Systems. In this section, we introduce a broad class of problems we call *variety-constrained linear systems*. The construction of such systems, an algorithm for solving them, and conditions that guarantee uniqueness of solutions, will allow us to complete the second stage of recovery. We introduced varieties and some core concepts associated with them in Chapter 2 and encourage the reader to review them since they will be relevant throughout this section.

DEFINITION 5.3.4 (Variety-constrained linear system). *We define a variety-constrained linear system (VC-LS) to be*

$$Au = b \quad \text{subject to} \quad u \in \mathcal{X}$$

where $A \in \mathbb{K}^{m \times n}$ is a coefficient matrix, $u \in \mathbb{K}^n$ is the solution vector, $b \in \mathbb{K}^m$ is another vector and $\mathcal{X} \subseteq \mathbb{K}^n$ is an irreducible conic variety.

Notably these systems may have unique solutions even when $m < n$. The condition $u \in \mathcal{X}$ could be considered a form of “sparsity” and as such, algorithms to solve VC-LSs can be considered a form of compressive sensing [20]. Problems of this form include the *canonical polyadic decomposition (CPD)-constrained linear systems* studied in [19], where the authors give an algebraic algorithm when $\mathcal{X}$ is a variety of rank-1 tensors and the constrained system has a single unique solution.

In a different direction, recent work by Johnston, Lovitz, and Vijayaraghavan develops an algorithm [43, Algorithm 1] that can be used to find multiple intersections between a linear subspace $\mathcal{U}$ and a conic variety $\mathcal{X}$, corresponding to a VC-LS with $b = 0$. We observe that the machinery of this intersection algorithm greatly clarifies the general principles in Boussé et al.’s algebraic algorithm and provides a method for dealing with multiple solutions. In Algorithm 2 below, we give a method for solving VC-LSs, synthesizing [19] and [43]. This can potentially handle multiple solutions, although there will only be a single solution in our use cases. We note that the algorithm is not always guaranteed to succeed, and this will be discussed later.

Algorithm 2 Solve a variety-constrained linear system.

Input: Matrix $A \in \mathbb{K}^{m \times n}$, vector $b \in \mathbb{K}^m$, and irreducible conic variety $\mathcal{X} \subseteq \mathbb{K}^n$.

Output: Solution(s) to variety-constrained linear system $Au = b$ subject to $u \in \mathcal{X}$, or “Fail.”

```

1: procedure FINDVARIETYCONSTRAINEDSOLUTIONS( $A, b, \mathcal{X}$ )
2:   Find the subspace  $\mathcal{U} = \{u \in \mathbb{K}^n : Au = \gamma b, \gamma \in \mathbb{K}\}$  of solutions to  $Au = b$  up to scale.
3:   Run the subspace conic-variety intersection algorithm ([43, Algorithm 1]) to find the vectors
       $v_1, v_2, \dots, v_{s'} \in \mathcal{U} \cap \mathcal{X}$ 
      (or report ‘Fail’ if the algorithm fails).
4:   for all  $v_k$  do
5:      $t \leftarrow \langle b, Av_k \rangle / \|b\|_2^2$ 
6:     if  $t \neq 0$  then
7:        $u_k \leftarrow \frac{1}{t} v_k$ 
8:     else
9:       discard  $u_k$ 
10:  return  $\{u_1, u_2, \dots, u_s\}$ 

```

Johnston, Lovitz, and Vijayaraghavan prove that their intersection algorithm succeeds on a certain class of generic inputs with planted solutions, provided the dimension of $\mathcal{U}$ is not too large. In the following sections we will be constructing specialized VC-LS problems, which are not covered by the existing analysis. Towards analyzing these, we next formulate more explicit sufficient conditions for Algorithm 2 to recover a unique solution to a VC-LS. We ask the following:

Given a particular $A \in \mathbb{K}^{m \times n}$ and $u \in \mathcal{X} \subseteq \mathbb{K}^n$, and defining $b := Au$, what are explicit conditions on $A, u, \mathcal{X}$ to guarantee that Algorithm 2 recovers u ?

In this context, we will call $u \in \mathcal{X}$ the *planted solution* to the VC-LS. First, observe that if $b = 0$ then we are finding $\ker A \cap \mathcal{X}$ so the intersection algorithm of JLV may be used without modification and solutions are only unique up to scaling. We will then only consider the case where $u \notin \ker A$ so that $b \neq 0$ and it is possible for a solution to be unique, not just unique up to scale. We will also restrict ourselves here to varieties $\mathcal{X} \subseteq \mathbb{K}^n$ cut out by degree-2 homogeneous polynomials and no linear functions so we may consider $\text{span}\{p_1, p_2, \dots, p_k\} \subseteq (\mathbb{K}^n)^{\otimes 2}$. As discussed in Chapters 2 and 4, varieties of this type include the variety of rank-at-most-1 tensors of any shape and number of symmetries and include the the variety of symmetric matrices of rank-at-most-1 which is our primary use case. We now define a construction which may be used to certify unique recovery. To do this, we first extend our definition for the symmetric product from matrices and vectors to subspaces.

DEFINITION 5.3.5. For subspaces $\mathcal{A}, \mathcal{B} \subseteq \mathbb{K}^n$, define

$$\mathcal{A}^{\otimes k} = \text{span}\{u^{\otimes k} : u \in \mathcal{A}\} \subseteq (\mathbb{K}^n)^{\otimes k}$$

and

$$\mathcal{A} \circledast \mathcal{B} = \text{span}\{u \circledast v : u \in \mathcal{A}, v \in \mathcal{B}\} \subseteq (\mathbb{K}^n)^{\otimes 2}.$$

Next, we generalize a construction used in prior work by the authors with Dastidar [24] to improve a bound the maximum dimension of a subspace in which one can find planted rank-1 matrices.

DEFINITION 5.3.6. *Define a set of matrices denoted by $\mathcal{M}(A, u, \mathcal{X})$ associated to the variety-constrained linear system with planted solution u where*

$$M \in \mathcal{M}(A, u, \mathcal{X})$$

if we may write $M = \Phi U$ for Φ a matrix whose rows form a basis for $\text{span}\{p_1, p_2, \dots, p_k\} \subseteq (\mathbb{K}^n)^{\otimes 2}$ where $p_1, p_2, \dots, p_k$ cut out the irreducible conic variety $\mathcal{X}$, and U a matrix whose columns form a basis for $(\text{span}(u) \circledast \ker A) + (\ker A)^{\otimes 2} \subseteq (\mathbb{K}^n)^{\otimes 2}$.

Note that each matrix in $\mathcal{M}(A, u, \mathcal{X})$ is related by *matrix equivalence* in that if $M, M' \in \mathcal{M}(A, u, \mathcal{X})$ then there exist invertible matrices P and Q such that $P^{-1}M'Q = M$.

PROPOSITION 5.3.7. *If $M \in \mathcal{M}(A, u, \mathcal{X})$ has full column-rank then Algorithm 2 recovers the planted vector u given inputs $A, b, \mathcal{X}$, where $b = Au$.*

Before we prove this, observe that for a representative M to have full column-rank, it must have at least as many rows as columns. This gives a necessary condition that may be used to determine when the kernel of A is too large. If $L = \dim(\ker A)$ and $k = \dim(\text{span}\{p_1, \dots, p_k\})$, then M has k rows and $\left(\binom{L}{2}\right) - 1$ columns. It is then necessary that

$$(5.16) \quad L \leq \sqrt{2(k+1) + \frac{1}{4}} - \frac{1}{2},$$

or else M has fewer rows than columns and necessarily has a non-trivial nullspace. The bound we give in (5.16) is a tightened version of the necessary condition given in [19, Section 3.1.2] which is based off counting the number of entries in second-compound matrices (a matrix containing all 2×2 minors). The use of second-compound matrices is a convenient tool but will produce linearly dependent rows for tensors of order greater than two or with symmetries. See Section 5.4 for more discussion.

In general, the bound (5.16) indicates that the more degree-2 homogeneous polynomials that cut out $\mathcal{X}$, the larger the kernel of A may be while allowing for the possibility of unique recovery. Now we give a proof of the proposition.

PROOF OF PROPOSITION 5.3.7. First, the bulk of our work will be showing that the intersection algorithm in Step 3 will return a single vector v_1 in the direction of u . Then we show that our rescaling produces a vector u_1 such that $Au_1 = b$ ensuring that $u_1 = u$.

We will show by way of contraposition that if $M \in \mathcal{M}(A, u, \mathcal{X})$ has full column-rank then for the subspace $\mathcal{U} = \{u \in \mathbb{K}^n : Au = \gamma b, \gamma \in \mathbb{K}\}$,

$$\text{span}\{p_1, p_2, \dots, p_k\}^\perp \cap \mathcal{U}^{\otimes 2} = \text{span}(u^{\otimes 2}).$$

Suppose that $\text{span}\{p_1, p_2, \dots, p_k\}^\perp \cap \mathcal{U}^{\otimes 2} \neq \text{span}(u^{\otimes 2})$. Now since u was planted, $u \in \mathcal{X}$ and $u \in \mathcal{U}$, $\text{span}\{p_1, p_2, \dots, p_k\}^\perp \cap \mathcal{U}^{\otimes 2} \supseteq \text{span}(u^{\otimes 2})$ so it must be that $\text{span}\{p_1, p_2, \dots, p_k\}^\perp \cap \mathcal{U}^{\otimes 2} \supsetneq \text{span}(u^{\otimes 2})$. Suppose then there is some $w \in (\mathbb{K}^n)^{\otimes 2}$ such that $w \in \text{span}\{p_1, p_2, \dots, p_k\}^\perp$, $w \in \mathcal{U}^{\otimes 2}$ and $w \notin \text{span}(u^{\otimes 2})$. First, if $\{n_1, \dots, n_L\}$ is a basis for $\ker A$, then $\{u, n_1, \dots, n_L\}$ is a basis for $\mathcal{U}$.

Then $\{u^{\otimes 2}\} \cup \{u \otimes n_j : j \in [L]\} \cup \{n_j \otimes n_k : 1 \leq j \leq k \leq L\}$ is a basis for $\mathcal{U}^{\otimes 2}$ and we may write

$$w = c_{\{0,0\}}u^{\otimes 2} + \sum_{j=1}^L c_{\{0,j\}}(u \otimes n_j) + \sum_{1 \leq j \leq k \leq L} c_{\{j,k\}}(n_j \otimes n_k)$$

for coefficients $c_{\{j,k\}} \in \mathbb{K}$, $0 \leq j \leq k \leq L$. Now it must be that $c_{\{j,k\}} \neq 0$ for some $\{j,k\} \neq \{0,0\}$, otherwise $w = c_{\{0,0\}}(u^{\otimes 2})^{\otimes 2} \in \text{span}((u^{\otimes 2})^{\otimes 2})$. Next, observe that if Φ is a matrix with rows a basis for $\text{span}\{p_1, p_2, \dots, p_k\}$ then $\Phi w = \vec{0}$. Since $\Phi u^{\otimes 2} = \vec{0}$ we have

$$\begin{aligned} \vec{0} &= \Phi w = c_{\{0,0\}}\Phi u^{\otimes 2} + \Phi \left(\sum_{j=1}^L c_{\{0,j\}}(u \otimes n_j) + \sum_{1 \leq j \leq k \leq L} c_{\{j,k\}}(n_j \otimes n_k) \right) \\ &= \Phi \left(\sum_{j=1}^L c_{\{0,j\}}(u \otimes n_j) + \sum_{1 \leq j \leq k \leq L} c_{\{j,k\}}(n_j \otimes n_k) \right). \end{aligned}$$

The set $\{u \otimes n_j : j \in [L]\} \cup \{n_j \otimes n_k : 1 \leq j \leq k \leq L\}$ is a basis for the subspace $(\text{span}(u) \otimes \ker A) \cup (\ker A)^{\otimes 2}$. Collecting these vectors as columns of a matrix N and collecting coefficients into a column vector $c = (c_{\{j,k\}} : 0 \leq j \leq k \leq L, \{j,k\} \neq \{0,0\})$, we have

$$\vec{0} = \Phi N c.$$

Observe that $\Phi N \in \mathcal{M}(A, u, \mathcal{X})$ and by the previous observation that at least one $c_{\{j,k\}} \neq 0$, the vector $c \neq \vec{0}$ so ΦN does not have full column-rank. Since column-rank is preserved by matrix equivalence, every matrix $M \in \mathcal{M}(A, u, \mathcal{X})$ lacks full column-rank.

In its second step, the intersection algorithm of JLV runs simultaneous diagonalization on a basis for the subspace $\text{span}\{p_1, p_2, \dots, p_k\}^\perp \cap \mathcal{U}^{\otimes 2}$ which we have shown is equal to $\text{span}(u^{\otimes 2})$ under the assumption M

has full column-rank. However, reshaping any $w \in \text{span}(u^{\otimes 2})$ to an $n \times n \times 1$ tensor and running simultaneous diagonalization in this special case degenerates to the rank-1 factorization of the $n \times n$ matrix which has the form $\alpha uu^\top$ for some $\alpha \in \mathbb{K}$. (We note that when $u \in \mathbb{C}$ this matrix is complex symmetric, not Hermitian.) It follows that Step 3 recovers a vector in the direction of u .

So the intersection algorithm of JLV in step 3 will output a vector v_1 that is a nonzero scalar multiple of u , i.e. $v_1 = \gamma u$ for some $\gamma \in \mathbb{K}$. Finally, we can check that $u_1 = v_1/t$ for $t = \langle b, Au_1 \rangle / \|b\|_2^2$ obeys $Av_1 = b$:

$$Av_1 = \frac{1}{t} Au_1 = \frac{\|b\|_2^2}{\langle b, Au_1 \rangle} Au_1 = \frac{\|b\|_2^2}{\langle b, \gamma Au \rangle} \gamma Au = \frac{\|b\|_2^2}{\langle b, b \rangle} b = b.$$

It follows that $u_1 = u$, meaning Algorithm 2 outputs the correct planted solution. $\square$

5.3.3. Aided Recovery. In this section, we will show how one may set up a variety-constrained linear system using either dMRA or pMRA invariant tensors such that the symmetric lift of the odd Fourier coefficients $(\hat{x}[E^c])^{\otimes 2}$ is the planted solution (or more precisely these coefficients after some group action). First, we note that second order invariants are insufficient to recover the signal in the dMRA model and hence the pMRA model as well.

PROPOSITION 5.3.8 (see [13, 31]). *For $x \in \mathbb{R}^d$, the second order invariant $T_{\text{dMRA}}^{(2)}(\hat{x})$ only determines $|\hat{x}[j]|$ for $j \in [d]$ (the power spectrum of x).*

Proof of this follows almost directly from Proposition 5.2.5. We show an analogous result holds for the pMRA model.

PROPOSITION 5.3.9. *For $x \in \mathbb{R}^{2d}$, the second order invariant $T_{\text{pMRA}}^{(2)}(\hat{x})$ only determines $|\hat{x}[j]|$ for $j \in [2d] \setminus \{d\}$ (the power spectrum of x minus the Nyquist component).*

See Section 5.9.2 for the proof, which follows almost directly from Proposition 5.2.5.

Since in neither model can we recover the phases of the Fourier coefficients from the second order invariants, we can examine third order invariants. Consider for $T = T_{\text{dMRA}}^{(3)}(\hat{x})$ or $T = T_{\text{pMRA}}^{(3)}(\hat{x})$ the following four (partially) symmetric subtensors, recalling that E denotes the even Fourier coefficients:

$$T[E, E, E], \quad T[E^c, E, E], \quad T[E^c, E^c, E], \quad \text{and} \quad T[E^c, E^c, E^c].$$

Since T is symmetric and E and E^c are a partition of $[2d]$, these subtensors represent a disjoint partition of the invariant polynomials that are entries of T . The first subtensor is the subject of Proposition 5.3.3 where we give the conditions under which it is itself an invariant subtensor of the same type as T . One might hope

that the fourth subtensor might yield some useful information about $\hat{x}[E^c]$ in a similar manner but this is not the case. Instead, $T[E^c, E^c, E]$ contains the remaining non-zero invariant polynomials.

PROPOSITION 5.3.10. *For $x \in \mathbb{R}^{2d}$ and $T = T_{\text{dMRA}}^{(3)}(\hat{x})$ or $T = T_{\text{pMRA}}^{(3)}(\hat{x})$,*

$$T[E^c, E, E] = 0 \quad \text{and} \quad T[E^c, E^c, E^c] = 0.$$

PROOF. This follows directly from Proposition 5.2.5 as well. Observe that entries of $T[E, E, E^c]$ are indexed by two even indices and one odd index, so any sum or difference of the three will be odd. Similarly, $T[E^c, E^c, E^c]$ is indexed by three odd indices, so any sum or difference of the three will be odd. Consider the indicator variables in the formulas found in Proposition 5.2.5 for $r = 3$, namely

$$\mathbb{1}\{j_1 + j_2 + j_3 \equiv 0 \pmod{2d}\} \quad \text{and} \quad \mathbb{1}\{s_1 j_1 + s_2 j_2 + s_3 j_3 \equiv 0 \pmod{2d}\}.$$

By the parity argument, the first can never be satisfied, and similarly the second cannot since the sign variables $s_1, s_2, s_3 \in \{-1, 1\}$ do not alter parity. Hence, these subtensors must be identically zero. $\square$

This raises the question:

If we can recover $y \in \mathbb{C}^d$ in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$ from the subtensor $T[E, E, E]$, is the remaining non-zero subtensor $T[E^c, E^c, E]$ enough to extend y to a vector x' in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$?

We will see the answer is yes. Observe that by similar parity arguments to those used in Proposition 5.3.10, the entries of $T[E^c, E^c, E]$ are bihomogeneous polynomials, being degree-2 in the entries of $\hat{x}[E^c]$ and degree-1 in the entries of $\hat{x}[E]$. Following this observation, we may set up a VC-LS problem with a planted solution providing such an extension.

Overarching Strategy. Suppose that $x \in \mathbb{K}^m$ is some signal, G is a finite group with some action on $\mathbb{K}^m$, $h : \mathbb{K}^m \rightarrow \mathbb{K}^n$ is a linear transformation, $R \subseteq [m]$, $S \subseteq [n]$ are sets of indices, $\tilde{G}$ is a finite group with some action on $\mathbb{K}^{|R|}$, $\tilde{h} : \mathbb{K}^{|R|} \rightarrow \mathbb{K}^{|S|}$ is also a linear transformation such that the conditions of Lemma 5.3.2 hold and we may extract an invariant subtensor

$$\left(T_{G,h}^{(3)}(x)\right)[S, S, S] = \frac{k|\tilde{G}|}{|G|} T_{\tilde{G},\tilde{h}}^{(3)}(x[R]).$$

Next, define

$$\mathcal{X}_1^\vee = \{x^{\otimes 2} : x \in \mathbb{K}^{|R^c|}\} \subseteq \left(\mathbb{K}^{|R^c|}\right)^{\otimes 2}$$

to be the variety of symmetric $|R^c| \times |R^c|$ rank-1 matrices.

The following strategy may be applied recursively to the orbit recovery problem:

- (1) Find a representative y of the orbit of $x[R]$ under $\tilde{G}$ by solving the orbit recovery problem associated with $T_{\tilde{G}, \tilde{h}}^{(3)}(x)$.

- (2) Define

$$b := \left(T_{G,h}^{(3)}(x) \right) [S^c, S^c, S] \in \mathbb{K}^{\left(\binom{|S^c|}{2} \right) |S|}$$

to be the vectorization of a particular subtensor of $T_{G,h}^{(3)}(x)$. Next, construct a linear operator using y so that $A := A(y) \in \mathbb{K}^{\left(\binom{|R^c|}{2} \right) \times \left(\binom{|S^c|}{2} \right) |S|}$ and for some x' in the G -orbit of x with component $z = x'[R^c]$,

$$(5.17) \quad A(z^{\otimes 2}) = b.$$

- (3) Solve the variety-constrained linear system constructed in (5.17), given generally as

$$(5.18) \quad Au = b \quad \text{subject to} \quad u \in \mathcal{X}_1^\vee,$$

using Algorithm 2.

We call this procedure *aided recovery* since a specific representative y is used to construct A and aids in the recovery of the rest of the signal. This ensures that only particular representatives of x 's orbit satisfy (5.18). The subsequent portions of this section show how to construct a linear operator as suggested in Step 2 for the tensors $T_{\text{dMRA}}^{(3)}(\hat{x})$ and $T_{\text{pMRA}}^{(3)}(\hat{x})$ with a planted solution as given in (5.17).

Aided Recovery for Dihedral MRA..

DEFINITION 5.3.11. For $y \in \mathbb{C}^d$, define the linear operator $A_{\text{dMRA}}(y) : (\mathbb{C}^d)^{\otimes 2} \rightarrow (\mathbb{C}^d)^{\otimes 2} \otimes \mathbb{C}^d$ column-wise by

$$(5.19) \quad A_{\text{dMRA}}[:, \wr j, k \wr] = \frac{\sqrt{3}}{2} y[-(j+k+1)] (\vec{e}_j \otimes \vec{e}_k \otimes \vec{e}_{-j-k-1} + \vec{e}_{-j-1} \otimes \vec{e}_{-k-1} \otimes \vec{e}_{j+k+1})$$

for $\wr j, k \wr$ a multiset of $[d]$.

For $\hat{x} \in \mathbb{C}^{2d}$, define a vector $b_{\text{dMRA}}(\hat{x}) := b_{\text{dMRA}} \in (\mathbb{C}^d)^{\otimes 2} \otimes \mathbb{C}^d$ given entry-wise by

$$(5.20) \quad b_{\text{dMRA}}[\wr j, k \wr, \wr m \wr] = \left(T_{\text{dMRA}}^{(3)}(\hat{x}) \right) [\wr 2j+1, 2k+1, 2m \wr]$$

where $0 \leq j \leq k < d$ and $0 \leq m < d$.

The vector b_{dMRA} is a particular vectorization of $\left(T_{\text{dMRA}}^{(3)}(\hat{x}) \right) [E^c, E^c, E]$, which is a partially symmetric tensor with symmetry along the first two modes.

PROPOSITION 5.3.12. *Let $x \in \mathbb{R}^{2d}$ with Fourier transform $\hat{x} \in \mathbb{C}^{2d}$, and let $y \in \mathbb{C}^d$ belong to the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$. For $A = A_{\text{dMRA}}(y)$ and $b = b_{\text{dMRA}}(\hat{x})$, there is some $g \in \mathcal{D}_{2d}$ such that $(g \cdot \hat{x})|_E = y$ and the vector $z = (g \cdot \hat{x})|_{E^c}$ satisfies $Az^{\otimes 2} = b$.*

REMARK 5.3.13. *There are generically two $g_1, g_2 \in \mathcal{D}_{2d}$ that meet these conditions and produce vectors z that are negations of each other.*

The crux is that the VC-LS

$$(A_{\text{dMRA}})u = b_{\text{dMRA}} \quad \text{subject to} \quad u \in \mathcal{X}_1^\vee$$

contains a planted solution which may be used to extend y to an element in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$. Applying such extensions recursively gives us our strategy for orbit recovery of signals $x \in \mathbb{R}^{2^{k+1}}$ for $k \geq 2$.

Since much of our indexing is modulo d we will also find it helpful to say that two multisets of the same size, M and M' , are equivalent modulo d denoted by $M \equiv M' \pmod{d}$ when there exists a bijection $\pi : M \rightarrow M'$ such that for every $m \in M$, $m \equiv \pi(m) \pmod{d}$.

PROOF OF PROPOSITION 5.3.12. To begin let $x \in \mathbb{R}^{2d}$, $\hat{x} \in \mathbb{C}^{2d}$ be x 's Fourier coefficients. We first consider the special case $y = \hat{x}[E]$. Using the shorthand $A = A_{\text{dMRA}}(\hat{x}[E])$ and $b = b_{\text{dMRA}}(\hat{x})$, we will show that the choice $g = \text{Id}$, i.e. $z = \hat{x}[E^c]$, satisfies

$$(5.21) \quad Az^{\otimes 2} = b.$$

Then, more generally, we will show that this holds over the orbit of $\hat{x}[E]$ for a corresponding transformation on $\hat{x}[E^c]$.

First, observe that by Definition 4.1.2, for $j', k' \in [d]$,

$$(5.22) \quad (z^{\otimes 2})[\langle j', k' \rangle] = w_{\langle j', k' \rangle} \hat{x}[2j' + 1] \hat{x}[2k' + 1]$$

and by Definition 5.3.11, for $j, k, m \in [d]$ we may find the entries in column of A using indicator functions,

$$(5.23) \quad \begin{aligned} A[\langle j, k \rangle, \langle m \rangle, \langle j', k' \rangle] &= \frac{\sqrt{3}}{2} x[-2(j' + k' + 1)] \times \\ &+ \mathbb{1}\{\langle j, k \rangle \equiv \langle j', k' \rangle \pmod{d}\} \cdot \mathbb{1}\{m \equiv -j' - k' - 1 \pmod{d}\} \\ &+ \mathbb{1}\{\langle -j - 1, -k - 1 \rangle \equiv \langle j', k' \rangle \pmod{d}\} \cdot \mathbb{1}\{m \equiv j' + k' + 1 \pmod{d}\}. \end{aligned}$$

So then entries of $Az^{\otimes 2}$ — the left-hand side of (5.21) — have the form

$$\begin{aligned} (Az^{\otimes 2})[(\wr j, k \S, \wr m \S)] &= A[(\wr j, k \S, \wr m \S), :]z^{\otimes 2} \\ &= \sum_{0 \leq j' \leq k' < d} A[(\wr j, k \S, \wr m \S), \wr j', k' \S](z^{\otimes 2})[\wr j', k' \S]. \end{aligned}$$

Using (5.22) and (5.23), we simplify this to

$$\begin{aligned} (Az^{\otimes 2})[(\wr j, k \S, \wr m \S)] &= \mathbb{1}\{m \equiv -j - k - 1 \pmod{d}\} \cdot \left(\frac{\sqrt{3}w_{\wr j, k \S}}{2} \hat{x}[2j+1]\hat{x}[2k+1]\hat{x}[2(-j-k-1)] \right. \\ &\quad \left. + \frac{\sqrt{3}w_{\wr -j-1, -k-1 \S}}{2} \hat{x}[-2j-1]\hat{x}[-2k-1]\hat{x}[2(j+k+1)] \right). \end{aligned}$$

Now $w_{\wr j, k \S} = w_{\wr -j-1, -k-1 \S}$ since if $j = k$ then $-j - 1 = -k - 1$. We may further simplify to get

$$\begin{aligned} (5.24) \quad (A(\hat{x}[E^c])^{\otimes 2})[(\wr j, k \S, \wr m \S)] &= \frac{\sqrt{3}w_{\wr j, k \S}}{2} \mathbb{1}\{m \equiv -j - k - 1 \pmod{d}\} \times \\ &\quad (\hat{x}[2j+1]\hat{x}[2k+1]\hat{x}[-2(j+k+1)] + \hat{x}[-2j-1]\hat{x}[-2k-1]\hat{x}[2(j+k+1)]). \end{aligned}$$

Next, we examine the entries of b — the right-hand side of (5.21) — using Proposition 5.2.5. We can expand the entries of the third order invariant indexed as

$$\begin{aligned} (5.25) \quad b[(\wr j, k \S, \wr m \S)] &= \left(T_{\text{dMRA}}^{(3)}(\hat{x}) \right) [\wr 2j+1, 2k+1, 2m \S] \\ &= \frac{w_{\wr 2j+1, 2k+1, 2m \S}}{2} \mathbb{1}\{2j+1+2k+1+2m \equiv 0 \pmod{2d}\} \times \\ &\quad (\hat{x}[2j+1]\hat{x}[2k+1]\hat{x}[2m] + \hat{x}[-2j-1]\hat{x}[-2k-1]\hat{x}[-2m]). \end{aligned}$$

The indicator conditions in (5.24) and (5.25) are the same, i.e.,

$$\mathbb{1}\{2j+1+2k+1+2m \equiv 0 \pmod{2d}\} = \mathbb{1}\{m \equiv -j - k - 1 \pmod{d}\}.$$

Next, by parity, $2m \not\equiv 2j+1 \pmod{2d}$ and $2m \not\equiv 2k+1 \pmod{2d}$ for any assignment of indices so we may simplify $w_{\wr 2j+1, 2k+1, 2m \S} = \sqrt{3}w_{\wr 2j+1, 2k+1 \S}$. Similar to our observations before, $w_{\wr 2j+1, 2k+1 \S} = w_{\wr j, k \S}$ so we get

$$\begin{aligned} (5.26) \quad b[(\wr j, k \S, \wr m \S)] &= \frac{\sqrt{3}w_{\wr j, k \S}}{2} \mathbb{1}\{m \equiv -j - k - 1 \pmod{d}\} \times \\ &\quad (\hat{x}[2j+1]\hat{x}[2k+1]\hat{x}[-2(j+k+1)] + \hat{x}[-2j-1]\hat{x}[-2k-1]\hat{x}[2(j+k+1)]). \end{aligned}$$

Comparing (5.24) to (5.26), they are equal and so $A(\hat{x}[E^c])^{\otimes 2} = b$.

Next, suppose that $y = \tilde{g} \cdot \hat{x}[E]$ for $\tilde{g} \in \mathcal{D}_d$ and set $A' = A_{\text{pMRA}}(y)$. The group element $\tilde{g}$ is one of $\{\hat{R}_{\ell,d}, \hat{J}\hat{R}_{\ell,d}\}_{\ell=0}^{d-1}$. If $\tilde{g} = \hat{R}_{\ell,d}$, set $g = \hat{R}_{\ell,2d}$ or $\hat{R}_{\ell+d,2d}$ and if $\tilde{g} = \hat{J}\hat{R}_{\ell,d}$, set $g = \hat{J}\hat{R}_{\ell,2d}$ or $\hat{J}\hat{R}_{\ell+d,2d}$. From the proof of Lemma 5.3.2, one can verify that these choices of group elements will ensure that $(\tilde{g} \cdot \hat{x})|_E = y$. Set $z = (g \cdot \hat{x})|_{E^c}$. We will show that $A'(z^{\otimes 2}) = A(\hat{x}[E^c])^{\otimes 2}$. First, observe that we may factor coefficients out of the columns of A by defining a diagonal matrix $X \in \mathbb{C}^{\binom{d}{2} \times \binom{d}{2}}$,

$$X = \text{diag}(\hat{x}[-2(j+k+1)] : 0 \leq j \leq k < d),$$

and a matrix B such that $BX = A$. Rows and columns of X are indexed by multisets $\{j, k\}$ of $[d]$ under the DM order, and columns of B are

$$B[:, \{j, k\}] = \frac{\sqrt{3}}{2}(\vec{e}_j \otimes \vec{e}_k \otimes \vec{e}_{-j-k-1} + \vec{e}_{-j-1} \otimes \vec{e}_{-k-1} \otimes \vec{e}_{j+k+1}).$$

Observe that in the first case $\tilde{g} = \hat{R}_{\ell,d}$ for some $\ell \in [d]$,

$$A_{\text{dMRA}}(\tilde{g} \cdot (\hat{x}[E])) = BX \underbrace{\text{diag}(e^{2\pi i(2(j+k+1))\ell/(2d)} : 0 \leq j \leq k < d)}_{=: D_1} = AD_1.$$

With a small simplification we have $D_1 = \text{diag}(e^{2\pi i(j+k+1)\ell/d} : 0 \leq j \leq k < d)$. Next, consider that for $\ell' \in [2d]$, if $z = (\hat{R}_{\ell',2d} \cdot \hat{x})|_{E^c}$, then for $j, k \in [d]$,

$$\begin{aligned} z^{\otimes 2}[\{j, k\}] &= w_{\{j, k\}} e^{-2\pi i(2j+1)\ell'/(2d)} \hat{x}[2j+1] e^{-2\pi i(2k+1)\ell'/(2d)} \hat{x}[2k+1] \\ &= e^{-2\pi i(2j+1+2k+1)\ell'/(2d)} w_{\{j, k\}} \hat{x}[2j+1] \hat{x}[2k+1]. \end{aligned}$$

So when $\ell' = \ell$ or $\ell' = \ell + d$,

$$z^{\otimes 2} = \underbrace{\text{diag}(e^{-2\pi i(j+k+1)\ell/d} : 0 \leq j \leq k < d)}_{=: D_2} (x[E^c])^{\otimes 2}.$$

Observe that the phases in D_1 and D_2 are conjugates so $D_1 D_2 = I$ and

$$A' z^{\otimes 2} = AD_1 D_2 (x[E^c])^{\otimes 2} = A(x[E^c])^{\otimes 2} = b.$$

In the second case, we set $\tilde{g} = \hat{J}_d \hat{R}_{\ell,d}$ for some $\ell \in [d]$. Let us consider the effect of the exchange matrix without phase shifting ($\ell = 0$) first. Observe that while $(\hat{J}_{2d} \cdot \hat{x})[E] = \hat{J}_d \cdot (\hat{x}[E])$, $(\hat{J}_{2d} \cdot \hat{x})[E^c] = J_d \cdot (\hat{x}[E^c])$

where

$$J_d[j, j'] = \begin{cases} 1 & j \equiv -j' - 1 \pmod{d} \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the action of reflection around the zero index for a vector of length $2d$ (the action of $\hat{J}_{2d}$) has the effect of reversal (J_d) on the odd coefficients. Then $(J_d \cdot (\hat{x}[E^c]))^{\otimes 2}[\wr j, k] = (\hat{x}[E^c])^{\otimes 2}[\wr -j - 1, -k - 1]$ so we may write

$$J_d^{\otimes 2}[\wr j, k, \wr j', k'] = \begin{cases} 1 & \wr j, k \equiv \wr -j' - 1, -k' - 1 \pmod{d} \\ 0 & \text{otherwise.} \end{cases}$$

Since $J_d^{\otimes 2}$ swaps entries in the vector, it is an involution, i.e. $(J_d^{\otimes 2})^2 = I$. Next notice that

$$J_d^{\otimes 2} X J_d^{\otimes 2} = \text{diag}(\hat{x}[-2((-j-1) + (-k-1) + 1)] : 0 \leq j \leq k < d) = \text{diag}(\hat{x}[2(j+k+1)] : 0 \leq j \leq k < d)$$

so we have that $A_{\text{dMRA}}(\hat{J}_d \cdot \hat{x}[E]) = B J_d^{\otimes 2} X J_d^{\otimes 2}$. Also note that $B[:, \wr j, k] = B[:, \wr -j - 1, -k - 1]$ so $B J_d^{\otimes 2} = B$. Then for any $\ell \in [d]$ with $\tilde{g} = \hat{J}_d \hat{R}_{\ell, d}$,

$$A_{\text{dMRA}}(\tilde{g} \cdot \hat{x}[E]) = B J_d^{\otimes 2} X D_1 J_d^{\otimes 2} = B X D_1 (J_d)^{\otimes 2} = A D_1 (J_d)^{\otimes 2}.$$

Next, consider that for $\ell' = \ell$ or $\ell' = \ell + d$, if $g = \hat{J}_{2d} \hat{R}_{\ell', 2d}$ and $z = (g \cdot \hat{x})|_{E^c}$,

$$z^{\otimes 2} = ((g \cdot \hat{x})[E^c])^{\otimes 2} = J_d^{\otimes 2} D_2 (\hat{x}[E^c])^{\otimes 2}.$$

Combining these two expressions we have

$$A' z^{\otimes 2} = A D_1 J_d^{\otimes 2} J_d^{\otimes 2} D_2 (\hat{x}[E^c])^{\otimes 2} = A D_1 D_2 (\hat{x}[E^c])^{\otimes 2} = A (\hat{x}[E^c])^{\otimes 2} = b. \quad \square$$

Aided Recovery for MRA with Projection. The relationship between the MRA invariants given in Lemma 5.2.4 implies something similar for subtensors of $T_{\text{pMRA}}^{(3)}$. We define the following shorthand for a particular restriction of $\hat{\Pi}^{\otimes 3}$ along each mode of its domain and codomain:

$$(5.27) \quad \hat{\Pi}_{E^c, E^c, E}^{(3)} := (\hat{\Pi}_{2d}[E_d^c, E_{2d}^c])^{\otimes 2} \otimes \hat{\Pi}_{2d}[E_d, E_{2d}].$$

By construction this is a linear transformation from a tensor with partial symmetry in $(\mathbb{C}^d)^{\otimes 2} \otimes \mathbb{C}^d$ to a tensor with partial symmetry in $(\mathbb{C}^{\lfloor (d+1)/2 \rfloor})^{\otimes 2} \otimes \mathbb{C}^{\lfloor d/2 \rfloor}$.

DEFINITION 5.3.14. For $y \in \mathbb{C}^d$, define the linear operator $A_{\text{pMRA}}(y) : (\mathbb{C}^{2d})^{\otimes 2} \rightarrow (\mathbb{C}^{\lfloor d/2 \rfloor})^{\otimes 2} \otimes \mathbb{C}^{\lfloor d/2 \rfloor}$ by

$$(5.28) \quad A_{\text{pMRA}}(y) = \hat{\Pi}_{E^c, E^c, E}^{(3)} A_{\text{dMRA}}(y),$$

with $\hat{\Pi}_{E^c, E^c, E}^{(3)}$ given in (5.27) and A_{dMRA} given in Definition 5.3.11.

For $x \in \mathbb{C}^{2d}$, define a vector $b_{\text{pMRA}}(\hat{x}) := b_{\text{pMRA}} \in (\mathbb{C}^{\lfloor (d+1)/2 \rfloor})^{\otimes 2} \otimes \mathbb{C}^{\lfloor d/2 \rfloor}$ given entry-wise by

$$(5.29) \quad b_{\text{pMRA}}[(j, k), (m)] = \left(T_{\text{pMRA}}^{(3)}(x) \right) [2j+1, 2k+1, 2m]$$

where $0 \leq j \leq k < d$ and $0 \leq m < d$.

PROPOSITION 5.3.15. Let $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ with Fourier transform $\hat{x} \in \mathbb{C}^{2d}$, and let $y \in \mathbb{C}^d$ belong to the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$. For $A = A_{\text{pMRA}}(y)$ and $b = b_{\text{pMRA}}(\hat{x})$, there is some $g \in \mathcal{D}_{2d}$ such that $(g \cdot \hat{x})|_E = y$ and the vector $z = (g \cdot \hat{x})|_{E^c}$ satisfies $Az^{\otimes 2} = b$.

Again, the main point is that

$$(A_{\text{pMRA}})u = b_{\text{pMRA}} \quad \text{subject to} \quad u \in \mathcal{X}_1^\vee$$

also contains a planted solution which may be used to extend y to an element in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$ and so solving this VC-LS can help us achieve orbit recovery.

The proposition follows almost directly from Proposition 5.3.12, though we clarify some of the more complicated steps through smaller lemmas.

LEMMA 5.3.16. For any $x \in \mathbb{C}^{2d}$,

$$b_{\text{pMRA}} = \hat{\Pi}_{E^c, E^c, E}^{(3)} \left(T_{\text{dMRA}}^{(3)}(\hat{x}) \right) [E_{2d}^c, E_{2d}^c, E_{2d}].$$

PROOF. First, observe that

$$b_{\text{pMRA}} = \left(T_{\text{pMRA}}^{(3)}(\hat{x}) \right) [E_d^c, E_d^c, E_d] = \left(\hat{\Pi}^{\otimes 3} T_{\text{dMRA}}^{(3)}(\hat{x}) \right) [E_d^c, E_d^c, E_d]$$

by Lemma 5.2.4. Then by Lemma 5.3.1, we may move the restriction to the rows of $\hat{\Pi}^{\otimes 3}$ yielding

$$b_{\text{pMRA}} = \left((\hat{\Pi}[E_d^c, :])^{\otimes 2} \otimes \hat{\Pi}[E_d, :] \right) T_{\text{dMRA}}^{(3)}(\hat{x}).$$

Finally, notice that for any vector $v \in \mathbb{C}^{2d}$, $\hat{\Pi}[E_d, :]v = \hat{\Pi}[E_d, E_{2d}]v[E_{2d}]$ since by the sparsity structure of $\hat{\Pi}$ given in (5.11), restriction to even rows $\hat{\Pi}[E_d, :]$ eliminates all non-zero entries in odd-indexed columns. The odd columns are all then zero so restriction to non-zero columns has no effect on the matrix-vector product.

For the analogous reasons, $\widehat{\Pi}[E_d^c, :]v = \widehat{\Pi}[E_d^c, E_{2d}^c]v[E_{2d}^c]$. Finally, applying this gives

$$\begin{aligned} \left(T_{\text{pMRA}}^{(3)}(\hat{x})\right)[E_d^c, E_d^c, E_d] &= \left(\widehat{\Pi}[E_d^c, E_{2d}^c]^{\otimes 2} \otimes \widehat{\Pi}[E_d, E_{2d}]\right)\left(T_{\text{dMRA}}^{(3)}(\hat{x})\right)[E_{2d}^c, E_{2d}^c, E_{2d}] \\ &= \widehat{\Pi}_{E^c, E^c, E}^{(3)}\left(T_{\text{dMRA}}^{(3)}(\hat{x})\right)[E_{2d}^c, E_{2d}^c, E_{2d}]. \end{aligned} \quad \square$$

PROOF OF PROPOSITION 5.3.15. Suppose $x \in \mathbb{R}^{2d}$, let $\hat{x} \in \mathbb{C}^{2d}$ be its Fourier transform, and fix some $y \in \mathbb{C}^d$ in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$. Then $y = \tilde{g} \cdot \hat{x}[E]$ for some $\tilde{g} \in \mathcal{D}_d$. Pick $g \in \mathcal{D}_{2d}$ as in Proposition 5.3.15 and set $z = (g \cdot x)[E^c]$. Then we know that z satisfies

$$A_{\text{dMRA}} z^{\otimes 2} = b_{\text{dMRA}}.$$

But then we have

$$\begin{aligned} (\widehat{\Pi}_{E^c, E^c, E}^{(3)} A_{\text{dMRA}}) z^{\otimes 2} &= \widehat{\Pi}_{E^c, E^c, E}^{(3)} b_{\text{dMRA}}, \\ (A_{\text{pMRA}}) z^{\otimes 2} &= b_{\text{pMRA}}, \end{aligned}$$

where the left-hand side follows from the definition of A_{pMRA} and the right-hand side from Lemma 5.3.16. Hence, $z^{\otimes 2}$ is still a planted solution for the variety constrained linear system

$$(A_{\text{pMRA}})u = b_{\text{pMRA}} \quad \text{subject to} \quad u \in \mathcal{X}_1^\vee. \quad \square$$

Careful readers may notice what appears to be an inconsistency: in the pMRA model we cannot recover the Nyquist component $\hat{x}[d]$ when $x \in \mathbb{R}^{2d}$. It appears that we need this component in the construction of A_{pMRA} . We clarify this apparent issue.

LEMMA 5.3.17. *When d is even and $j, k \in [d]$ such that $j + k + 1 \equiv d/2 \equiv -d/2 \pmod{d}$,*

$$A_{\text{pMRA}}[:, \langle j, k \rangle] = \vec{0},$$

so $A_{\text{pMRA}}(y)$ does not depend on the value of $y[d/2]$.

PROOF. When d is even, the matrix $\widehat{\Pi}_{2d}[E_d, E_{2d}]$ used in the construction of $\widehat{\Pi}_{E^c, E^c, E}^{(3)}$ in (5.27) is simply the matrix $\widehat{\Pi}_d$. From (5.11) one can determine that $\widehat{\Pi}_d$ has a zero-column at index $d/2$, so $\widehat{\Pi}_d \vec{e}_{d/2} = \vec{0}$. In Definition 5.3.11, it would appear we need to know $y[d/2]$ to construction a column $A_{\text{dMRA}}[:, \langle j, k \rangle]$ when $j + k + 1 \equiv d/2 \equiv -d/2 \pmod{d}$, which we use in the construction of A_{pMRA} . However, in this case the

column has the form

$$A_{\text{dMRA}}[:, \langle j, k \rangle] = \frac{\sqrt{3}}{2} y[d/2] (\vec{e}_j \otimes \vec{e}_k \otimes \vec{e}_{d/2} + \vec{e}_{-j-1} \otimes \vec{e}_{-k-1} \otimes \vec{e}_{d/2}).$$

Consider the partially symmetric standard basis tensors appearing above. It follows that

$$\begin{aligned} \hat{\Pi}_{E^c, E^c, E}^{(3)}(\vec{e}_j \otimes \vec{e}_k \otimes \vec{e}_{d/2}) &= (\hat{\Pi}_{2d}[E_d^c, E_{2d}^c]^{\otimes 2}(\vec{e}_j \otimes \vec{e}_k)) \otimes (\hat{\Pi}_{2d}[E_d, E_{2d}]\vec{e}_{d/2}) \\ &= (\hat{\Pi}_{2d}[E_d^c, E_{2d}^c]^{\otimes 2}(\vec{e}_j \otimes \vec{e}_k)) \otimes (\hat{\Pi}_d \vec{e}_{d/2}) \\ &= (\hat{\Pi}_{2d}[E_d^c, E_{2d}^c]^{\otimes 2}(\vec{e}_j \otimes \vec{e}_k)) \otimes \vec{0} \\ &= \vec{0}. \end{aligned}$$

Similarly $\hat{\Pi}_{E^c, E^c, E}^{(3)}(\vec{e}_{-j-1} \otimes \vec{e}_{-k-1} \otimes \vec{e}_{d/2}) = \vec{0}$. Thus, the corresponding columns in A_{pMRA} that would appear to use $y[d/2]$ are zero-columns. $\square$

Observe that since A_{pMRA} has some zero-columns it will certainly have a non-trivial nullspace. However, this does not mean the VC-LS is impossible to solve.

5.3.4. Extending Representatives. We have now shown how the subtensor $T[E^c, E^c, E]$ can be used to produce variety-constrained linear systems that contain the remaining information about $\hat{x}$. We now state the full routine for extracting this information, in both the dihedral and projected MRA settings.

Algorithm 3 Aided recovery for dihedral MRA.

Input: A vector y in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$ (even Fourier coeffs. of x) and the tensor $T_{\text{dMRA}}^{(3)}(\hat{x})$ from (5.12).
Output: Either reports an element w in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$ or “Fail” if no element was found.

```

1: procedure EXTENDREPRESENTATIVEDIHEDRALMRA( $y, T$ )
2:    $A \leftarrow A_{\text{dMRA}}(y)$  from Definition 5.3.11
3:    $b \leftarrow T[E^c, E^c, E]$ 
4:    $u \leftarrow \text{FINDVARIETYCONSTRAINEDSOLUTIONS}(A, b, \mathcal{X}_1^\vee)$  (see Algorithm 2)
5:   Reshape  $u$  to a complex symmetric matrix using (4.9):  $U \leftarrow \text{matricize}(u)$ 
6:   Find any  $z$  such that  $zz^\top = U$ 
7:    $w[E] \leftarrow y$ 
8:    $w[E^c] \leftarrow z$ 
9:   return  $w$ 
```

Algorithm 4 Aided recovery for projected MRA.

Input: A vector y in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$ and the tensor $T_{\text{pMRA}}^{(3)}(\hat{x})$ from (5.13).

Output: Either reports an element w in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$ or “Fail” if no element was found.

```

1: procedure EXTENDREPRESENTATIVEPROJECTEDMRA( $y, T$ )
2:    $A \leftarrow A_{\text{pMRA}}(y)$  from Definition 5.3.14
3:    $b \leftarrow T[E^c, E^c, E]$ 
4:    $u \leftarrow \text{FINDVARIETYCONSTRAINEDSOLUTIONS}(A, b, \mathcal{X}_1^\vee)$  (see Algorithm 2)
5:   Reshape  $u$  to a complex symmetric matrix using (4.9):  $U \leftarrow \text{matricize}(u)$ 
6:   Find any  $z$  such that  $zz^\top = U$ 
7:    $w[E] \leftarrow y$ 
8:    $w[E^c] \leftarrow z$ 
9:   return  $w$ 

```

We note that matricizing u produces a complex symmetric matrix, not a Hermitian one. To find z , one can simply take the first column of U scaled by the appropriate complex scalar. There is a unique choice of z , up to negation.

The success of these algorithms hinge on the success of the FINDVARIETYCONSTRAINEDSOLUTIONS subroutine applied in this case. In Section 5.4, we analyze the two sets of equivalent matrices $\mathcal{M}(A, u, \mathcal{X})$ corresponding to these two problems, to determine when these methods succeed.

5.4. Analysis of the Aided Recovery Procedure

In Section 5.3.2 we repurposed the subspace–conic variety intersection algorithm from [43] via our modification in Algorithm 2 to solve VC-LS’s. We now apply our general construction from Definition 5.3.6 for the dMRA and pMRA models.

Since this section and the next concern aided recovery in the context of recursion we will assume that d is even.

Let

$$(5.30) \quad C := \{\langle j, k \rangle : j, k \in [d], j + k + 1 < d\}$$

be a set of multisets. Treated as coordinates in the plane (j, k) with $j \leq k$ for even d these make up one fourth of the square $[d] \times [d]$, so $|C| = d^2/4$. We use the Dershowitz–Manna order to construct a map $\beta' : C \rightarrow [d^2/4]$ which maps a multiset in C to its index when all are arranged in increasing order. Next, we will define

$$-C := \{\langle d-1-j-1, d-1-k \rangle : \langle j, k \rangle \in C\}$$

which we may think of as the reflection of the original multisets-as-coordinates in C over the line $k = d - j$. Define the set

$$L_1 := \{\wr(j, k) : j, k \in [d], j + k + 1 = d\}.$$

These three sets form a partition of the set of all two-element multisets of $[d]$, i.e. $C \sqcup -C \sqcup L_1 = \left(\left(\begin{smallmatrix} [d] \\ 2 \end{smallmatrix}\right)\right)$. Using this partition we extend our original ordering to a map from all multisets $\beta : \left(\left(\begin{smallmatrix} [d] \\ 2 \end{smallmatrix}\right)\right) \rightarrow [d^2/4] \cup \{0\}$ by $\beta(C) = \beta(-C) = \beta'(C)$ and $\beta(L_1) = 0$ for any $j \in [d]$. Finally define a function $\text{sgn} : \left(\left(\begin{smallmatrix} [d] \\ 2 \end{smallmatrix}\right)\right) \rightarrow \{-1, 0, 1\}$ by $\text{sgn}(C) = 1$, $\text{sgn}(-C) = -1$ and $\text{sgn}(L_1) = 0$.

DEFINITION 5.4.1. For $\hat{x} \in \mathbb{C}^{2d}$ the matrix $M_{\text{dMRA}}(\hat{x})$ is defined to have the rows

(5.31)

$$(M_{\text{dMRA}})[(j_1, k_1, j_2, k_2), :] =$$

$$\begin{aligned} & \hat{x}[2j_1 + 1] \cdot \hat{x}[2k_1 + 1] \cdot \hat{x}[2(j_2 + k_2 + 1)] \cdot \text{sgn}(\wr(j_2, k_2)) \quad \vec{e}_0 \otimes \vec{e}_{\beta(\wr(j_2, k_2))} \\ & - \hat{x}[2j_1 + 1] \cdot \hat{x}[2k_2 + 1] \cdot \hat{x}[2(j_2 + k_1 + 1)] \cdot \text{sgn}(\wr(j_2, k_1)) \quad \vec{e}_0 \otimes \vec{e}_{\beta(\wr(j_2, k_1))} \\ & + \hat{x}[2j_2 + 1] \cdot \hat{x}[2k_2 + 1] \cdot \hat{x}[2(j_1 + k_1 + 1)] \cdot \text{sgn}(\wr(j_1, k_1)) \quad \vec{e}_0 \otimes \vec{e}_{\beta(\wr(j_1, k_1))} \\ & - \hat{x}[2j_2 + 1] \cdot \hat{x}[2k_1 + 1] \cdot \hat{x}[2(j_1 + k_2 + 1)] \cdot \text{sgn}(\wr(j_1, k_2)) \quad \vec{e}_0 \otimes \vec{e}_{\beta(\wr(j_1, k_2))} \\ & + \hat{x}[2(j_1 + k_1 + 1)] \cdot \hat{x}[2(j_2 + k_2 + 1)] \cdot \text{sgn}(\wr(j_1, k_1)) \cdot \text{sgn}(\wr(j_2, k_2)) \quad \vec{e}_{\beta(\wr(j_1, k_1))} \otimes \vec{e}_{\beta(\wr(j_2, k_2))} \\ & - \hat{x}[2(j_1 + k_2 + 1)] \cdot \hat{x}[2(j_2 + k_1 + 1)] \cdot \text{sgn}(\wr(j_1, k_2)) \cdot \text{sgn}(\wr(j_2, k_1)) \quad \vec{e}_{\beta(\wr(j_1, k_2))} \otimes \vec{e}_{\beta(\wr(j_2, k_1))} \end{aligned}$$

with $0 \leq j_1 \leq k_1 < d$, $0 \leq j_2 \leq k_2 < d$, $j_1 < j_2$ and $k_1 < k_2$. Indices are taken modulo $2d$ when accessing entries of $\hat{x}$ and the rowspace of M should be considered a subspace of $(\mathbb{C}^{d^2/4+1})^{\otimes 2} \cap \text{span}(\vec{e}_0 \otimes \vec{e}_0)^\perp$.

For $\hat{x} \in \mathbb{C}^{2d}$, $M_{\text{dMRA}}(\hat{x})$ has $\frac{1}{12}(d+1)d^2(d-1)$ rows and $\binom{d^2/4+2}{2} - 1$ columns. As a result, it has more rows than columns when $d \geq 4$. It is also a very sparse matrix, having at most 6 non-zero entries per row.

PROPOSITION 5.4.2. For $\hat{x} \in \mathbb{C}^{2d}$ and any $g \in \mathcal{D}_{2d}$,

$$M_{\text{dMRA}}(g \cdot \hat{x}) \in \mathcal{M}(A_{\text{dMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee)$$

where $\mathcal{M}(A, u, \mathcal{X})$ is the set of equivalent matrices in Definition 5.3.6 for a variety-constrained linear system $Au = b$ with planted solution $u \in \mathcal{X}$.

CONJECTURE 5.4.3. For $d = 2^k$ with $k \geq 2$, $M_{\text{dMRA}}(\hat{x})$ has full column-rank as a symbolic matrix of polynomials where the variables are the $d+1$ real and $d-1$ imaginary components of the Fourier coefficients of $x \in \mathbb{R}^{2d}$.

See Section 5.4.2 for numerical evidence that these matrices have full column-rank symbolically. We will also drop the dependence on $\hat{x}$, replacing it with the parameter d , and writing $M_{\text{dMRA}}(d)$, to emphasize when we are considering the construction as a symbolic matrix.

REMARK 5.4.4 (Symbolic Matrices). *While we define $M_{\text{dMRA}}(\hat{x})$ in terms of Fourier coefficients of x , the signal x is real so $\hat{x}[j] = \overline{\hat{x}[2d-j]}$ for $j \in [2d]$. Due to this conjugate relationship, $\hat{x}[j]$ and $\hat{x}[2d-j]$ cannot be considered distinct complex variables symbolically. Instead we consider symbolic variables $r_0, r_1, \dots, r_d$ and $c_1, c_2, \dots, c_{d-1}$ and define $\hat{x}$ by*

$$\begin{aligned}\Re(\hat{x}) &=: (r_0, r_1, r_2, \dots, r_{d-1}, r_d, r_{d-1}, \dots, r_2, r_1) \quad \text{and} \\ \Im(\hat{x}) &=: (0, c_1, c_2, \dots, c_{d-1}, 0, -c_{d-1}, -c_{d-2}, \dots, -c_2, -c_1).\end{aligned}$$

Now $M_{\text{dMRA}}(\hat{x})$ is a matrix of symbolic polynomials in these underlying symbolic variables. Such a matrix is said to have full column-rank if some maximal minor evaluates to a nonzero polynomial.

If $M_{\text{dMRA}}(d)$ has full column-rank symbolically, this ensure the success of Algorithm 3 in retrieving the odd Fourier coefficients for generic $x \in \mathbb{R}^{2d}$.

THEOREM 5.4.5 (Aided recovery for dihedral MRA). *Fix even $d \geq 4$ and suppose $M_{\text{dMRA}}(d)$ has full column-rank as a symbolic matrix. Then for generic $x \in \mathbb{R}^{2d}$, if y is an arbitrary element in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$, Algorithm 3 runs in polynomial time on inputs y and $T_{\text{dMRA}}^{(3)}(\hat{x})$ and recovers some $\hat{x}'$ in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$.*

Note that the algorithm is guaranteed to succeed for a specific $\hat{x}$ whenever the instance $M_{\text{dMRA}}(\hat{x})$ has full column-rank.

As one can see in Definition 5.3.6, the construction of $M_{\text{dMRA}}(\hat{x})$ requires a basis for the kernel of $A_{\text{dMRA}}(\hat{x}[E])$. To prove Proposition 5.4.2 (and subsequently Theorem 5.4.5), we give an explicit and sparse basis for this subspace. For $y \in \mathbb{C}^d$, define the vectors

$$(5.32) \quad n_{\chi_{j,k}}(y) := y[j+k+1]\vec{e}_j \circledast \vec{e}_k - y[-j-k-1]\vec{e}_{-k-1} \circledast \vec{e}_{-j-1}$$

where indices on the standard basis vectors and indexing entries of y are taken modulo d . Define

$$(5.33) \quad N_{\text{dMRA}}(y) := \{n_m(y) : m \in C\} \subseteq (\mathbb{C}^d)^{\otimes 2}$$

where C is the set defined in (5.30).

LEMMA 5.4.6. *When $y \in \mathbb{C}^d$ with each $y[j] \neq 0$ for $j \in [d]$, $N_{\text{dMRA}}(y)$ is a basis for $\ker A_{\text{dMRA}}(y)$*

The non-zero condition prevents $A_{\text{dMRA}}(y)$ from having zero-columns. Zero-columns would increase the dimension of the nullspace of A_{dMRA} which — as we have noted — does not necessarily break our method unless the bound in (5.16) is exceeded. However, generic signals will have non-zero Fourier coefficients so we do not consider these cases.

As suggested by the lemma, C , β and sgn are helpful construct this specific basis. Note that for any $j \in [d]$, $n_{\lfloor j, d-j-1 \rfloor}(y)$ is $\vec{0}$, hence it is an edge-case we handle along this anti-diagonal and the reason we split our set of multisets into C , $-C$ and L_1 .

PROOF OF LEMMA 5.4.6. For some $y \in \mathbb{C}^d$ set $A = A_{\text{dMRA}}(y)$. Consider the inner product of two columns of A given by Definition 5.3.11:

$$\begin{aligned} \langle A[:, \lfloor j, k \rfloor], A[:, \lfloor j', k' \rfloor] \rangle &= \frac{3}{4} w_{\lfloor j, k \rfloor}^2 y[-(j+k+1)] \overline{y[-(j'+k'+1)]} \\ &\quad \langle (\vec{e}_j \circledast \vec{e}_k \otimes \vec{e}_{-j-k-1} + \vec{e}_{-j-1} \circledast \vec{e}_{-k-1} \otimes \vec{e}_{j+k+1}), \\ &\quad (\vec{e}_{j'} \circledast \vec{e}_{k'} \otimes \vec{e}_{-j'-k'-1} + \vec{e}_{-j'-1} \circledast \vec{e}_{-k'-1} \otimes \vec{e}_{j'+k'+1}) \rangle. \end{aligned}$$

The columns will be orthogonal unless they have overlapping support. Observe that by Corollary 4.1.4,

$$\begin{aligned} \langle \vec{e}_a \circledast \vec{e}_b \otimes \vec{e}_c, \vec{e}_{a'} \circledast \vec{e}_{b'} \otimes \vec{e}_{c'} \rangle &= \langle \vec{e}_a \circledast \vec{e}_b, \vec{e}_{a'} \circledast \vec{e}_{b'} \rangle \langle \vec{e}_c, \vec{e}_{c'} \rangle \\ &= \frac{1}{2} (\langle \vec{e}_a, e_{a'} \rangle \langle \vec{e}_b, e_{b'} \rangle + \langle \vec{e}_a, e_{b'} \rangle \langle \vec{e}_b, e_{a'} \rangle) \langle \vec{e}_c, e_{c'} \rangle \\ &= \frac{|\lfloor a, b \rfloor|}{2} \mathbb{1}(\lfloor a, b \rfloor = \lfloor a', b' \rfloor) \mathbb{1}(c = c'). \end{aligned}$$

By linearity of the inner product and since are indices are taken modulo d we have

$$\begin{aligned} &\langle (\vec{e}_j \circledast \vec{e}_k \otimes \vec{e}_{-j-k-1} + \vec{e}_{-j-1} \circledast \vec{e}_{-k-1} \otimes \vec{e}_{j+k+1}), (\vec{e}_{j'} \circledast \vec{e}_{k'} \otimes \vec{e}_{-j'-k'-1} + \vec{e}_{-j'-1} \circledast \vec{e}_{-k'-1} \otimes \vec{e}_{j'+k'+1}) \rangle \\ &= \left(\frac{|\lfloor j, k \rfloor|}{2} \mathbb{1}(\lfloor j, k \rfloor \equiv \lfloor j', k' \rfloor \pmod{d}) \mathbb{1}(-j-k-1 \equiv -j'-k'-1 \pmod{d}) \right. \\ &\quad + \frac{|\lfloor j, k \rfloor|}{2} \mathbb{1}(\lfloor j, k \rfloor \equiv \lfloor -j'-1, -k'-1 \rfloor \pmod{d}) \mathbb{1}(-j-k-1 \equiv j'+k'+1 \pmod{d}) \\ &\quad + \frac{|\lfloor -j-1, -k-1 \rfloor|}{2} \mathbb{1}(\lfloor -j-1, -k-1 \rfloor \equiv \lfloor j', k' \rfloor \pmod{d}) \mathbb{1}(j+k+1 \equiv -j'-k'-1 \pmod{d}) \\ &\quad + \frac{|\lfloor -j-1, -k-1 \rfloor|}{2} \mathbb{1}(\lfloor -j-1, -k-1 \rfloor \equiv \lfloor -j'-1, -k'-1 \rfloor \pmod{d}) \\ &\quad \left. \cdot \mathbb{1}(-j-k-1 \equiv j'+k'+1 \pmod{d}) \right) \\ &= |\lfloor j, k \rfloor| \mathbb{1}(\lfloor j, k \rfloor \equiv \lfloor j', k' \rfloor \pmod{d}) + |\lfloor j, k \rfloor| \mathbb{1}(\lfloor j, k \rfloor \equiv \lfloor -j'-1, -k'-1 \rfloor \pmod{d}). \end{aligned}$$

Hence, columns are orthogonal unless $\wr j, k \wr \equiv \wr j', k' \wr$ (i.e. they are the same column) or $\wr j, k \wr \equiv \wr -j' - 1, -k' - 1 \wr$. Notice that in the second case, these columns are scalar multiples of each other. We then have that

$$y[j + k + 1] \cdot A[:, \wr j, k \wr] - y[-(j + k + 1)] \cdot A[:, \wr -j - 1, -k - 1 \wr] = \vec{0}.$$

Hence, all null vectors take the form given in (5.32). Further, for $m_1, m_2 \in C$ with $m_1 \neq m_2$, $\langle n_{m_1}, n_{m_2} \rangle = 0$ since the null vectors have non-overlapping support. Hence, $N_{\text{dMRA}}(y)$ is a orthogonal basis for $\ker A$. $\square$

Having determined a sparse basis for the nullspace of $A_{\text{dMRA}}(y)$ in terms of the entries of y we may now show that $M_{\text{dMRA}}(\hat{x})$ lies in the equivalence class of matrices that determine whether the variety-constrained linear system has a unique planted solution recoverable by Algorithm 3.

PROOF OF PROPOSITION 5.4.2. We will proceed in two stages: first we will show that the matrix $M_{\text{dMRA}}(\hat{x}) \in \mathcal{M}(A_{\text{dMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee)$, and then we will show that for $g \in \mathcal{D}_{2d}$, $M_{\text{dMRA}}(g \cdot \hat{x}) = W_g M_{\text{dMRA}}(\hat{x})$ where W_g is some change-of-basis matrix acting on the rows of $M_{\text{dMRA}}(\hat{x})$ depending on g .

Consider a matrix $N \in \mathbb{C}^{\binom{d}{2} \times |C|}$ with rows

$$N[\wr j, k \wr, :] = \text{sgn}(\wr j, k \wr) w_{\wr j, k \wr} \hat{x}[2(j + k + 1)] \vec{e}_{\beta(\wr j, k \wr)}.$$

The columns of N are the null vectors of N_{dMRA} scaled by the weight $w_{\wr j, k \wr}$ so by Lemma 5.4.6, $\text{colspan } N = \ker A_{\text{dMRA}}$. Next, we can expand the solution vector $\hat{x}[E^c]^{\otimes 2}$ by examining its entries,

$$(\hat{x}[E^c]^{\otimes 2})[\wr j, k \wr] = w_{\wr 2j+1, 2k+1 \wr} \hat{x}[2j + 1] \hat{x}[2k + 1] = w_{\wr j, k \wr} \hat{x}[2j + 1] \hat{x}[2k + 1].$$

If we form the $\mathbb{C}^{\binom{d}{2} \times (|C|+1)}$ block matrix

$$U = \begin{bmatrix} (\hat{x}[E^c])^{\otimes 2} & N \end{bmatrix},$$

where columns are indexed from 0 to $|C|$, the rows have the form

$$U[\wr j, k \wr, :] = w_{\wr j, k \wr} (\hat{x}[2j + 1] \hat{x}[2k + 1] \vec{e}_0 + s(\{j, k\}) \hat{x}[2(j + k + 1)] \vec{e}_{\beta(\wr j, k \wr)}).$$

Define the matrix $\tilde{U}$ to be $U^{\otimes 2}$ with the first column dropped. Using Lemma 4.2.2 we expand along the rows but exclude vectors in the direction $\vec{e}_0 \otimes \vec{e}_0$. This gives the three terms

$$\begin{aligned}
(5.34) \quad \tilde{U}[\langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle, :] &= w_{\langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle} w_{\langle j_1, k_1 \rangle} w_{\langle j_2, k_2 \rangle} \times \\
&(\hat{x}[2j_1 + 1] \hat{x}[2k_1 + 1] \hat{x}[2(j_2 + k_2 + 1)] s(\langle j_2, k_2 \rangle) \vec{e}_0 \otimes \vec{e}_{\beta(\langle j_2, k_2 \rangle)} \\
&+ \hat{x}[2j_2 + 1] \hat{x}[2k_2 + 1] \hat{x}[2(j_1 + k_1 + 1)] s(\langle j_1, k_1 \rangle) \vec{e}_0 \otimes \vec{e}_{\beta(\langle j_1, k_1 \rangle)} \\
&+ \hat{x}[2(j_1 + k_1 + 1)] \hat{x}[2(j_2 + k_2 + 1)] s(\langle j_1, k_1 \rangle) s(\langle j_2, k_2 \rangle) \vec{e}_{\beta(\langle j_1, k_1 \rangle)} \otimes \vec{e}_{\beta(\langle j_2, k_2 \rangle)}).
\end{aligned}$$

Finally, define Φ to have rows indexed by the set $\{(j_1, k_1, j_2, k_2) : 0 \leq j_1 \leq k_1 < d, 0 \leq j_2 \leq k_2 < d, j_1 < j_2 \text{ and } k_1 < k_2\}$ defined row-wise by

$$\Phi[(j_1, k_1, j_2, k_2), :] = (E_{j_1 k_1 j_2 k_2}^\vee)^\top$$

where $E_{j_1 k_1 j_2 k_2}^\vee$ is the vector define in (4.12). These vectors are a basis for $\mathcal{I}(\mathcal{X}_1^\vee)_2 \subseteq ((\mathbb{C}^d)^{\otimes 2})^{\otimes 2}$ by Lemma 2.4.6 and since each vector $E_{j_1 k_1 j_2 k_2}^\vee$ is the difference of two rescaled standard basis vectors, left-multiplication by Φ has the effect of taking a linear combination of exactly two rows of $\tilde{U}$. The weights in (5.34) cancel with those in (4.12) for each row which result in the final expression given in Definition 5.4.1. By construction this in $\mathcal{M}(A_{\text{dMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee)$.

Next, we will show that $M_{\text{dMRA}}(\hat{x})$ transforms equivariantly in that for $g \in \mathcal{D}_{2d}$, $M_{\text{dMRA}}(g \cdot \hat{x}) = g \cdot M_{\text{dMRA}}(\hat{x})$ for some group action on the column space. First, consider the action of $\hat{R}_\ell \in \mathcal{D}_{2d}$. Observe that for each monomial in the row, the sum of indices used to access entries is always $2(j_1 + k_1 + j_2 + k_2 + 2)$. Then

$$M_{\text{dMRA}}(\hat{R}_\ell \cdot \hat{x})[\langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle, :] = e^{-2\pi i \ell (j_1 + k_1 + j_2 + k_2 + 2)/d} M_{\text{dMRA}}(\hat{x})[\langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle, :].$$

If we define a matrix $D_\ell := \text{diag}(e^{-2\pi i \ell (j_1 + k_1 + j_2 + k_2 + 2)/d} : \langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle)$, where $j_1, j_2, k_1, k_2 \in [d]$ and $j_1 \leq k_1, j_2 \leq k_2, j_1 < j_2$ and $k_1 < k_2$, then

$$M_{\text{dMRA}}(\hat{R}_\ell \hat{x}) = D_\ell M_{\text{dMRA}}(\hat{x}) \in \mathcal{M}(A_{\text{dMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee).$$

since D_ℓ is an invertible matrix. It remains to show that exchange also causes a simple change of basis. Recall that $(\hat{J} \cdot \hat{x})[j] = \hat{x}[-j]$ for $j \in [2d]$ and with indices take modulo d . Then

$$M_{\text{dMRA}}(\hat{J} \cdot \hat{x})[\langle j_1, k_1 \rangle, \langle j_2, k_2 \rangle, :] = M_{\text{dMRA}}(\hat{x})[\langle d - j_2 - 1, d - k_2 - 1 \rangle, \langle d - j_1 - 1, d - k_1 - 1 \rangle, :].$$

Observe that if $j_1 \leq k_1$, $j_2 \leq k_2$ and $k_1 < k_2$, then the mapping $j \mapsto d - j - 1$ flips each inequality. Thus, the action of $\hat{J}$ simply permutes the rows of M_{dMRA} by swapping many pairs of rows. Hence, $\hat{J}$ induces some change of basis so in general

$$M_{\text{dMRA}}(g \cdot \hat{x}) \in \mathcal{M}(A_{\text{dMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee)$$

for any $g \in \mathcal{D}_{2d}$. □

Next, we will briefly give an elementary result in invariant theory about the decomposition of a symmetric rank-1 matrix.

LEMMA 5.4.7. *If $z \in \mathbb{K}^d$ and $Z = zz^\top$ then Z determines z up to sign.*

PROOF. If Z is the zero matrix then $z = \vec{0}$. Suppose Z is not the zero matrix and both $Z = z_1 z_1^\top$ and $Z = z_2 z_2^\top$ for $z_1, z_2 \in \mathbb{K}^d$ are rank-1 decompositions. Then there exists some $j \in [d]$ so that $z[j] \neq 0$. The j -th column of Z is $Z[:, j] = z_1[j]z_1 = z_2[j]z_2$. Since $z_1[j] \neq 0$ we know $z_1 = \alpha z_2$ for some $\alpha \in \mathbb{K}$. But then $z_1 z_1^\top = Z = \alpha^2 z_2 z_2^\top$ so $1 = \alpha^2$ and $\alpha = \pm 1$. We can then conclude that $z_2 = \pm z_1$. □

We can now prove the correctness of the recovery step in Algorithm 3.

PROOF OF THEOREM 5.4.5. Since y is in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$, by Lemma 5.3.12 the linear system

$$Au = b \quad \text{with} \quad u \in \mathcal{X}_1^\vee$$

with $A = A_{\text{dMRA}}(\hat{y})$ and $b = b_{\text{dMRA}}(\hat{x})$ has a planted solution of the form $u = z^{\otimes 2}$ such that $z = (g \cdot \hat{x})|_{E^c}$ and $(g \cdot \hat{x})|_E = \hat{y}$ for some $g \in \mathcal{D}_{2d}$. Define $M = M_{\text{dMRA}}(\hat{x})$ to be the specific instance of $M_{\text{dMRA}}(d)$ using the coefficients of $\hat{x}$. Since symbolically $M_{\text{dMRA}}(d)$ has full column-rank and $\hat{x}$ is generic, the specific instance M has full column-rank. Then Proposition 5.4.2 and Proposition 5.3.7 together imply the subroutine FINDVARIETYCONSTRAINEDSOLUTIONS returns u . Now $u = z^{\otimes 2}$ so the matricization of u may be written $U = zz^\top$. By Lemma 5.4.7 there is a z' such that $U = z'z'^\top$ and $z' = z$ or $z' = -z$. We will show either choice of sign in the rank-1 decomposition is valid. If the sign is positive and $z' = z$, by assumption z extends y to some $w \in \mathbb{C}^{2d}$ in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$. If $z' = -z$ observe that $\hat{R}_d = \text{diag}((-1)^j : j \in [2d])$ so $(\hat{R}_d g \cdot z)[E^c] = -(g \cdot z)[E^c]$ and $(\hat{R}_d g \cdot z)[E] = (g \cdot z)[E] = y$. Since $\hat{R}_d g \in \mathcal{D}_{2d}$, either sign results in an element in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$. □

5.4.1. Sufficient Conditions for Aided Recovery in the pMRA Model. As in the previous section, we assume d is even throughout. The construction of M_{pMRA} is very similar to that of M_{dMRA} with

a small modification to the column indexing function β . Define the sets of multisets

$$L_2 := \{\langle j, k \rangle : j, k \in [d], j + k + 1 \equiv d/2 \pmod{d}\},$$

$$C' := C \setminus L_2, \text{ and}$$

$$-C' := \{\langle d-1-j, d-1-k \rangle : \langle j, k \rangle \in C'\},$$

where C is defined in (5.30). The accounting is somewhat trickier but $|L_2 \cap C| = \lceil d/4 \rceil$ (see Figure 5.2 for a visualization). We can then determine that $|C' \sqcup L_2| = \lceil d(d+1)/4 \rceil$. Let $\gamma' : C' \cap L_2 \rightarrow \llbracket \lceil d(d+1)/4 \rceil \rrbracket$ give the DM order on $C' \cap L_2$. We extend this using the partition given by four sets now, $\left(\left(\begin{smallmatrix} d \\ 2 \end{smallmatrix}\right)\right) = L_1 \sqcup C' \sqcup -C' \sqcup L_2$, so that $\gamma : \left(\left(\begin{smallmatrix} d \\ 2 \end{smallmatrix}\right)\right) \rightarrow \llbracket \lceil d(d+1)/4 \rceil \rrbracket \cup \{0\}$ is a map such that $\gamma(C') = \gamma(-C') = \gamma'(C)$, $\gamma(L_2) = \gamma'(L_2)$ and $\gamma(L_1) = 0$. We may reuse the original sign function defined at the start of Section 5.4.

Finally, recall that $\hat{x}[d]$ is indeterminate in the pMRA model. However, in the following we will set $\hat{x}[d] = 1$ to avoid dealing with a special case in construction. Lemma 5.4.13 and the proof of Proposition 5.4.9 will clarify why this is used. With these modification, our M -matrix representative for the pMRA model is effectively the same as in Definition 5.4.1.

DEFINITION 5.4.8. For $\hat{x} \in \mathbb{C}^{2d}$ the matrix $M_{\text{pMRA}}(\hat{x})$ has rows

(5.35)

$$\begin{aligned} (M_{\text{pMRA}})[(j_1, k_1, j_2, k_2), :] = & \\ & \hat{x}[2j_1 + 1] \cdot \hat{x}[2k_1 + 1] \cdot \hat{x}[2(j_2 + k_2 + 1)] \cdot \text{sgn}(\langle j_2, k_2 \rangle) \quad \vec{e}_0 \otimes \vec{e}_{\gamma(\langle j_2, k_2 \rangle)} \\ & - \hat{x}[2j_1 + 1] \cdot \hat{x}[2k_2 + 1] \cdot \hat{x}[2(j_2 + k_1 + 1)] \cdot \text{sgn}(\langle j_2, k_1 \rangle) \quad \vec{e}_0 \otimes \vec{e}_{\gamma(\langle j_2, k_1 \rangle)} \\ & + \hat{x}[2j_2 + 1] \cdot \hat{x}[2k_2 + 1] \cdot \hat{x}[2(j_1 + k_1 + 1)] \cdot \text{sgn}(\langle j_1, k_1 \rangle) \quad \vec{e}_0 \otimes \vec{e}_{\gamma(\langle j_1, k_1 \rangle)} \\ & - \hat{x}[2j_2 + 1] \cdot \hat{x}[2k_1 + 1] \cdot \hat{x}[2(j_1 + k_2 + 1)] \cdot \text{sgn}(\langle j_1, k_2 \rangle) \quad \vec{e}_0 \otimes \vec{e}_{\gamma(\langle j_1, k_2 \rangle)} \\ & + \hat{x}[2(j_1 + k_1 + 1)] \cdot \hat{x}[2(j_2 + k_2 + 1)] \cdot \text{sgn}(\langle j_1, k_1 \rangle) \cdot \text{sgn}(\langle j_2, k_2 \rangle) \quad \vec{e}_{\gamma(\langle j_1, k_1 \rangle)} \otimes \vec{e}_{\gamma(\langle j_2, k_2 \rangle)} \\ & - \hat{x}[2(j_1 + k_2 + 1)] \cdot \hat{x}[2(j_2 + k_1 + 1)] \cdot \text{sgn}(\langle j_1, k_2 \rangle) \cdot \text{sgn}(\langle j_2, k_1 \rangle) \quad \vec{e}_{\gamma(\langle j_1, k_2 \rangle)} \otimes \vec{e}_{\gamma(\langle j_2, k_1 \rangle)} \end{aligned}$$

with $0 \leq j_1 \leq k_1 < d$, $0 \leq j_2 \leq k_2 < d$, $j_1 < j_2$ and $k_1 < k_2$. Indices are taken modulo $2d$ when accessing entries of $\hat{x}$ and the rowspace of M_{pMRA} should be considered a subspace of $(\mathbb{C}^{\lceil d(d+1)/4 \rceil + 1})^{\otimes 2} \cap \text{span}(\vec{e}_0 \otimes \vec{e}_0)^\perp$. Further, one should replace $\hat{x}[d]$ with the constant 1 anywhere it appears as the expression above as M_{pMRA} does not depend on $\hat{x}[d]$'s value but the definition requires a non-zero constant in these positions to be correct.

For $\hat{x} \in \mathbb{C}^{2d}$, $M_{\text{pMRA}}(\hat{x})$ has $\frac{1}{12}(d+1)d^2(d-1)$ rows and $\binom{\lfloor d(d+1)/4 \rfloor + 2}{2} - 1$ columns. As a result, it is square when $d = 4$ and has more rows than columns when $d > 4$. It is also a very sparse matrix, having at most 6 non-zero entries per row.

We have an analogue to Proposition 5.4.2.

PROPOSITION 5.4.9. *For $\hat{x} \in \mathbb{C}^{2d}$ and any $g \in \mathcal{D}_{2d}$,*

$$M_{\text{pMRA}}(g \cdot \hat{x}) \in \mathcal{M}(A_{\text{pMRA}}(\hat{x}[E]), (\hat{x}[E^c])^{\otimes 2}, \mathcal{X}_1^\vee)$$

where $\mathcal{M}(A, u, \mathcal{X})$ is the set of equivalent matrices in Definition 5.3.6 for a variety-constrained linear system $Au = b$ with planted solution $u \in \mathcal{X}$.

CONJECTURE 5.4.10. *For $d = 2^k$ with $k \geq 3$, $M_{\text{pMRA}}(\hat{x})$ has full column-rank as a symbolic matrix in the d real and $d - 1$ imaginary components of the Fourier coefficients of $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$.*

See Section 5.4.2 for evidence toward this conjecture. Similar to the notation $M_{\text{dMRA}}(d)$, we will write $M_{\text{pMRA}}(d)$ and drop the dependence on $\hat{x}$ and reference the parameter d when considering this matrix as a symbolical construct.

REMARK 5.4.11. *See Remark 5.4.4 for how these symbolic variables are formally described in the case of the M_{dMRA} matrix in Definition 5.4.1. The condition that $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ results in one less variable since now*

$$\Re(\hat{x}) =: (r_0, r_1, r_2, \dots, r_{d-1}, 0, r_{d-1}, \dots, r_2, r_1)$$

in the underlying symbolic variables. However, in Definition 5.4.8, to ensure that Proposition 5.4.9 holds, the dependence on $\hat{x}[d] = r_d$ should formally be removed by replacing it with the constant 1 (or any non-zero constant).

THEOREM 5.4.12 (Aided recovery for projected MRA). *Fix even $d \geq 4$ and suppose $M_{\text{pMRA}}(d)$ has full column-rank as a symbolic matrix. Then for generic $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$, if y is an arbitrary element in the $\mathcal{D}_d$ -orbit of $\hat{x}[E]$, Algorithm 4 runs in polynomial time on inputs y and $T_{\text{pMRA}}^{(3)}(\hat{x})$ and recovers some $\hat{x}'$ in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$.*

The construction requires a basis for the kernel of $A_{\text{pMRA}}(\hat{x}[E])$, which we now characterize. Define the set of vectors

$$(5.36) \quad N_{\text{pMRA}}(y) := \{n_m(y) : m \in C'\} \cup \{\vec{e}_j \circledast \vec{e}_{d/2-1-j} : j \in [d]\}$$

where $n_m(y)$ is a null-vector of the form define in (5.32) where indices are taken modulo d on the standard basis vectors.

LEMMA 5.4.13. *When d is even and $y \in \mathbb{C}^d$ with each $y[j] \neq 0$ for $j \in [d]$, $N_{\text{pMRA}}(y)$ is a basis for $\ker A_{\text{pMRA}}(y)$.*

Comparing constructions in Lemma 5.4.6 and Lemma 5.4.13, $\ker A_{\text{dMRA}} \subseteq \ker A_{\text{pMRA}}$ which can also be inferred from the definition of A_{pMRA} . In the pMRA model, without the Nyquist component a small set of the dMRA null vectors (specifically the vectors indexed by multisets in the set L_2) to split into two orthogonal standard basis vectors in $(\mathbb{C}^d)^{\otimes 2}$.

PROOF OF LEMMA 5.4.13. Set $A = A_{\text{pMRA}}(y)$ for $y \in \mathbb{C}^d$. By Definition 5.3.14, $A = \hat{\Pi}_{E^c, E^c, E}^{(3)} A_{\text{dMRA}}(y)$, and in Lemma 5.3.17 we determined that A has zero-columns in the $\langle j, k \rangle$ -th position when $j + k + 1 \equiv d/2$. By these observations we may conclude that $\text{span}(N_{\text{pMRA}}(y)) \subseteq \ker A_{\text{pMRA}}$. Our main goal is to show that these are the only such vectors. First, recall that our indices are taken modulo $2d$ or d depending on the dimension of the matrices involved. If we use $j \in \{-d+1, \dots, 0, \dots, d-1, d\}$ to index columns of $\hat{\Pi}$ then observe that the d -th column is $\vec{0}$ and every other column may be written as

$$\hat{\Pi}[:, j] = e^{\pi i(j-|j|)/2d} \vec{e}_{|j|}.$$

For $j, k, \ell \in \{-d/2+1, \dots, d-1, d/2\}$ the $(\langle j, k \rangle, \langle \ell \rangle)$ -th column of $\hat{\Pi}_{E^c, E^c, E}^{(3)}$ is given by

$$\hat{\Pi}_{E^c, E^c, E}^{(3)}[:, (\langle j, k \rangle, \langle \ell \rangle)] = \begin{cases} \phi(j, k) \sqrt{3} w_{\langle j, k \rangle} \left(\vec{e}_{|j+\frac{1}{2}|-\frac{1}{2}} \otimes \vec{e}_{|k+\frac{1}{2}|-\frac{1}{2}} \otimes \vec{e}_{|\ell|} \right) & \ell \not\equiv d/2 \pmod{d} \\ \vec{0} & \ell \equiv d/2 \end{cases}$$

where $\phi(j, k) = \exp\left(\frac{\pi i(2j+1-|2j+1|+2k+1-|2k+1|-2(j+k+1)+2|j+k+1|)}{2d}\right) = \exp\left(\frac{\pi i(2|j+k+1|-|2j+1|-|2k+1|)}{2d}\right)$. It then follows that the columns of A have the form

$$A[:, \langle j, k \rangle] = (\phi(j, k) + \phi(-j-1, -k-1)) \frac{\sqrt{3} w_{\langle j, k \rangle}}{2} \left(\vec{e}_{|j+\frac{1}{2}|-\frac{1}{2}} \otimes \vec{e}_{|k+\frac{1}{2}|-\frac{1}{2}} \otimes \vec{e}_{|j+k+1|} \right)$$

except when $j + k + 1 \equiv d/2$, in which case the phase factors cancel to form a zero-column, producing the second set of null vectors in (5.36). Next, for a column of A with index $\langle j', k' \rangle$ to have the same support as

the $\{j, k\}$ -th column of A , it must be that

$$|j + 1/2| - 1/2 \equiv |j' + 1/2| - 1/2 \pmod{d},$$

$$|j + 1/2| - 1/2 \equiv |j' + 1/2| - 1/2 \pmod{d},$$

$$|j + k + 1| \equiv |j' + k' + 1| \pmod{d}.$$

The first two equations force $j' \equiv j$ or $-j - 1 \pmod{d}$ and $k' = k$ or $-k - 1 \pmod{d}$. We know from Lemma 5.4.6 that when $\{j', k'\} = \{-j - 1, -k - 1\}$, this produces the set of vectors we index by C' in (5.36). For the other two pairs, if $\{j', k'\} = \{-j - 1, k\}$ or if $\{j', k'\} = \{j, k - 1\}$ then the third modular equation requires either $2j \equiv 1 \pmod{d}$ or $2k \equiv 1 \pmod{d}$, both of which are impossible since d is even. Since the columns have non-overlapping support, they are orthogonal and there can be no additional linearly independent null vectors beyond the ones in $N_{\text{PMRA}}(y)$. $\square$

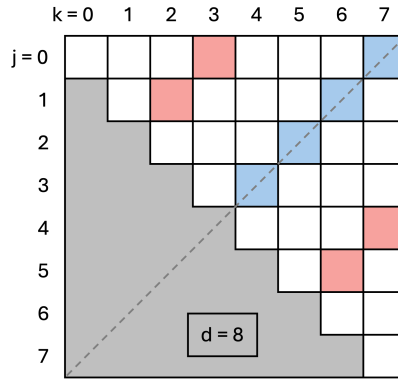


FIGURE 5.2. Illustration of the combinatorics involved in computing nullspace dimensions for $d = 8$. Columns of A_{dMRA} and A_{pMRA} correspond to boxes in the upper triangular region, as they are indexed by pairs $\{j, k\}$ with $0 \leq j \leq k < d$. Boxes that are reflections across the diagonal dashed line correspond to columns that are scalar multiples of each other. Each pair gives an element in the null space. Boxes in blue along the dashed line indicate columns that have a unique support and produce no null vectors. Boxes in red indicate columns that are zero-columns in A_{pMRA} but not A_{dMRA} ; these produce additional null vectors which are tensor products of standard basis elements.

Having verified our choice of basis spans the nullspace of A_{pMRA} , we will now sketch the proof of Proposition 5.4.9, highlighting some subtle differences to Proposition 5.4.2 though the same technique is used.

PROOF SKETCH OF PROPOSITION 5.4.9. In Definition 5.4.8, the function γ serves the same role as β in Definition 5.4.8, being a column-index lookup function to place the vectors of $N_{\text{dMRA}}(\hat{x}[E])$ into a matrix N . The construction of $M_{\text{pMRA}}(\hat{x})$ from this matrix then follows along the same lines as that in Proposition 5.4.2.

However, $M_{\text{pMRA}}(g \cdot \hat{x})$ is no longer row-equivalent to $M_{\text{pMRA}}(\hat{x})$ since we canonically set $\hat{x}[d] = 1$ in our formula. Indeed, when $g \in \{\hat{R}_\ell\}_{\ell=0}^{2d-1} \subseteq \mathcal{D}_{2d}$, multiplying by D_ℓ (defined in the proof of Proposition 5.4.2) may change the sign of $\hat{x}[d]$. However, observe that this coefficient is used to weight one of the basis vectors in $N_{\text{pMRA}}(\hat{x}[E])$ which is just a standard symmetric tensor product of two basis vectors. Multiplying N on the right by a diagonal matrix with a sign in the correct place will fix the sign issue and we can conclude that $M_{\text{pMRA}}(g \cdot \hat{x})$ is still matrix-equivalent to $M_{\text{pMRA}}(\hat{x})$. $\square$

Finally, we can verify that the recursive step succeeds.

PROOF OF THEOREM 5.4.12. This proof follows from identical logic to that in the proof of Theorem 5.4.5 with the replacements $A = A_{\text{pMRA}}$, $b = b_{\text{pMRA}}$, $M = M_{\text{pMRA}}(\hat{x})$ and the reference to Proposition 5.3.12 should be replaced with Proposition 5.3.15. $\square$

5.4.2. Numerical Evidence for Conjectures. For $x \in \mathbb{R}^{2d}$ with Fourier coefficients $\hat{x} = \mathcal{F}x$, one can produce strong numerical evidence that $M_{\text{dMRA}}(\hat{x})$ and $M_{\text{pMRA}}(\hat{x})$ have full column-rank symbolically by picking x at random from some absolutely continuous distribution in $\mathbb{R}^{2d}$ and verifying that the smallest singular value of $M_{\text{dMRA}}(\hat{x})$ or $M_{\text{pMRA}}(\hat{x})$ are each significantly far from zero. Note that demonstrating a *single* choice of x for which M has full column-rank implies that M has full column-rank symbolically, and therefore implies success of aided recovery for generic inputs of this dimension. The condition number of a rectangular matrix M is defined to be

$$(5.37) \quad \kappa(M) = \frac{\sigma_{\max}(M)}{\sigma_{\min}(M)}.$$

Since $\kappa(M) = \kappa(\alpha M)$ for $\alpha \neq 0$, the condition number is a scale-invariant measure of how close M is to having a rank-deficit where if M has linearly-dependent columns, we canonically define $\kappa(M) := \infty$. In the [following repository](#)², we furnish the experiments for the values $d = 4, 8$ and 16 where x is chosen to have standard real Gaussian entries. Out of 500 random trials, we report the smallest condition number obtained for each of $M_{\text{dMRA}}(\hat{x})$ and $M_{\text{pMRA}}(\hat{x})$. We report the smallest condition number since, in practice we need only show that there exists one instantiation of variables $\hat{x}_0, \dots, \hat{x}_{d-1}$ such that $M_{\text{dMRA}}(\hat{x})$ and $M_{\text{pMRA}}(\hat{x})$ have full numerical column-rank. This instance then provides the best possible evidence of these matrices having full column-rank symbolically. The results reported in Table 5.1 support Conjectures 5.4.3 and 5.4.10, respectively. When $d = 32$, we were not able to get our numerical method to converge when computing the smallest singular value for instances of $M_{\text{dMRA}}(\hat{x})$ and $M_{\text{pMRA}}(\hat{x})$. However, we note that the implementation given in Section 5.7.1 and our numerical experiments there give indirect evidence that M_{dMRA} and M_{pMRA}

²Repository is found at <https://github.com/lemniscate8/MRA-variant-orbit-recovery> for those reading in print.

d	$\kappa(M_{\text{dMRA}})$	$\kappa(M_{\text{pMRA}})$
4	6.07	74.96
8	13.36	129.78
16	135.81	1464.38

TABLE 5.1. The smallest condition numbers of both $M_{\text{dMRA}}(\hat{x})$ and $M_{\text{pMRA}}(\hat{x})$ out of 500 trials where each entry of $x \in \mathbb{R}^{2d}$ is chosen independently from a standard normal distribution.

still have full column-rank symbolically in this case. While the smallest sampled condition number grows in the cases we could test, we have no reason to believe that it becomes infinite for any finite d .

5.5. Analysis of the Base Case

The recursive procedure described above can reduce the problem to an orbit recovery problem of constant size. In these cases, the theoretical computational complexity is not an issue, but we do need a proof that the orbit is uniquely identifiable and an explicit procedure to recover it.

5.5.1. Frequency Marching in Dihedral MRA. By the work of Edidin and Katz [31] the question of identifiability has been solved for the dMRA model: namely, the first three moments separate generic orbits.

Counting the number of rows and columns of M_{dMRA} shows that our recursive method cannot work starting from $d = 1$ ($x \in \mathbb{R}^2$), which is unfortunate since this case is trivial to achieve orbit recovery in as the zeroth and first Fourier coefficients are real and can be estimated from the first and second moments, respectively. We then consider the case $d = 2$ ($x \in \mathbb{R}^4$) as our base case.

The technique of frequency marching arises from the observation that the degree-2 cyclic or dihedral invariants give the magnitudes of Fourier coefficients. One can then use the degree-3 invariants to set up a system of equations modulo 2π to solve for the phases of Fourier coefficients. Edidin and Katz considered frequency marching as a strategy for recovery, but primarily considered it as a proof strategy, since it does not give an efficient algorithm. However, since we only care about signals of small dimension for our base case, we demonstrate that frequency marching recovers the orbit of a signal of length four. This example also helps elucidate why this method is inefficient for larger signals (although the concurrent work [6] provides a fix to this).

THEOREM 5.5.1. *When $x \in \mathbb{R}^4$ with $\hat{x}[j] \neq 0$ for $j \in [4]$, the third moment $T_{\text{dMRA}}^{(3)}(\hat{x})$ may be used to recover the $\mathcal{D}_4$ -orbit of $\hat{x}$.*

PROOF OF THEOREM 5.5.1. First, observe that since $x \in \mathbb{R}^4$, $\hat{x}[0]$ and $\hat{x}[2]$ are real and $\hat{x}[1] = \overline{\hat{x}[3]}$. While $T_{\text{dMRA}}^{(1)}(\hat{x})[j] = \hat{x}[0]$ for every $j = 0, 1, 2, 3$, we can also recover this from the third moment via

$$\hat{x}[0] = \sqrt[3]{T_{\text{dMRA}}^{(3)}(\hat{x})[\zeta 0, 0, 0\zeta]/w_{\zeta 0, 0, 0\zeta}}.$$

Similarly, for each $j = 1, 2, 3$,

$$T_{\text{dMRA}}^{(3)}[\zeta 0, j, -j\zeta] = w_{\zeta 0, j, -j\zeta} \hat{x}[0] \hat{x}[j] \hat{x}[-j] = w_{\zeta 0, j, -j\zeta} \hat{x}[0] |\hat{x}[j]|^2$$

so we can recover the magnitudes of Fourier coefficients from the third moment. Using the Fourier coefficient magnitudes we may define the phases of Fourier coefficients to be $\phi \in [0, 2\pi)^4$ where $\hat{x}[j] = |\hat{x}[j]|e^{i\phi[j]}$ for $j = 0, 1, 2, 3$. However, $\phi[0]$ is known, and $\phi[3] \equiv -\phi[1] \pmod{2\pi}$ so we need only determine $\phi[1]$ and $\phi[2]$. Since $\hat{x}[2]$ is real, for some choice of sign $s_1 \in \{-1, 1\}$ we have that $\hat{x}[2] = s_1 |\hat{x}[2]|$ and $\phi[2] = \frac{\pi}{2}(1 - s_1)$. Next, by Proposition 5.2.5,

$$\begin{aligned} T_{\text{dMRA}}^{(3)}[\zeta 1, 1, 2\zeta] &= \frac{w_{\zeta 1, 1, 2\zeta}}{2} (\hat{x}[1]^2 \hat{x}[2] + \hat{x}[3]^2 \hat{x}[2]) \\ &= \frac{w_{\zeta 1, 1, 2\zeta}}{2} |\hat{x}[1]|^2 \hat{x}[2] (e^{2i\phi[1]} + e^{2i\phi[3]}) \\ &= \frac{w_{\zeta 1, 1, 2\zeta}}{2} |\hat{x}[1]|^2 \hat{x}[2] (e^{2i\phi[1]} + e^{-2i\phi[1]}) \\ &= w_{\zeta 1, 1, 2\zeta} |\hat{x}[1]|^2 \hat{x}[2] \cos(2\phi[1]). \end{aligned}$$

Using the expression for $\phi[2]$, we know $\phi[1]$ must satisfy

$$(5.38) \quad \cos(2\phi[1]) = s_1 \frac{T_{\text{dMRA}}^{(3)}[\zeta 1, 1, 2\zeta]}{w_{\zeta 1, 1, 2\zeta} |\hat{x}[1]|^2 |\hat{x}[2]|}.$$

Treating $\phi[1]$ as an unknown value in $[0, 2\pi)$, there are at most four solutions for this value $\phi[1]$. Further, if $\phi[1]$ is a solution, so are $-\phi[1]$ and $\phi[1] + \pi$. It follows that for some choice of sign $s_2 \in \{-1, 1\}$ and $k \in \{0, 1\}$,

$$\phi[1] = \frac{1}{2} s_2 \cos^{-1} \left(s_1 \frac{T_{\text{dMRA}}^{(3)}[\zeta 1, 1, 2\zeta]}{w_{\zeta 1, 1, 2\zeta} |\hat{x}[1]|^2 |\hat{x}[2]|} \right) + k\pi.$$

Finally, for each choice of $s_1, s_2 \in \{-1, 1\}$ and $k \in \{0, 1\}$ the assignments

$$\begin{aligned} \phi[1] &\leftarrow \frac{1}{2} s_2 \cos^{-1} \left(s_1 \frac{T_{\text{dMRA}}^{(3)}[\zeta 1, 1, 2\zeta]}{w_{\zeta 1, 1, 2\zeta} |\hat{x}[1]|^2 |\hat{x}[2]|} \right) + k\pi \\ \phi[2] &\leftarrow \frac{\pi}{2}(1 - s_1) \end{aligned}$$

and $\phi[3] \leftarrow -\phi[1] \pmod{2\pi}$ enumerate (by varying s_1, s_2, k) all 8 elements of a $\mathcal{D}_4$ -orbit for some signal. Since at least one choice of s_1, s_2 and k must produce the true phases for $\hat{x}$, this procedure must recover all elements of the $\mathcal{D}_4$ -orbit of $\hat{x}$. $\square$

Frequency marching succeeds because we only need to use one relationship between the phases $\phi[1]$ and $\phi[2]$ given by (5.38). In general, for $x \in \mathbb{R}^{2d}$, relationships between phases take the form

$$\cos(\phi[j_1] + \phi[j_2] + \phi[j_3]) = \frac{T_{\text{dMRA}}^{(3)}[\imath j_1, j_2, j_3]}{w_{\imath j_1, j_2, j_3} |\hat{x}[j_1]| |\hat{x}[j_2]| |\hat{x}[j_3]|} =: \alpha_{j_1, j_2, j_3}$$

with $j_1, j_2, j_3 \in [2d]$, $j_1 + j_2 + j_3 \equiv 0 \pmod{2d}$ and $\phi[j] = -\phi[2d - j \pmod{2d}]$ for $j \in [2d]$. As noted by Edidin and Katz, we can only determine a linear equation for the phases modulo 2π up to sign, i.e.

$$\phi[j_1] + \phi[j_2] + \phi[j_3] \equiv \pm \cos^{-1}(\alpha_{j_1, j_2, j_3}) \pmod{2d}.$$

In this case $d = 2$, the number of signs was sufficiently small that each assignment of sign variables produced a consistent system of equations, each giving an element of the orbit of $\hat{x}$. For larger d one must resort to an exhaustive search over all choices of signs, most of which will produce inconsistent linear systems. As a result, we use frequency marching in four dimensions as demonstrated in Theorem 5.5.1 as a base case for our recursive method to recover a signal from the dMRA third moment.

Further discussion. Comparing to the frequency marching implementation by Bendory, Boumal, Ma, Zhao and Singer [12, Algorithm 4] for regular MRA, our method differs in that we first estimate $\phi[2]$ and then $\phi[1]$. Their method requires finding the exact value of $\phi[1]$ to begin estimating further frequencies. The method of doing this exactly given at the beginning of [12, Section IV] requires the cyclic invariants and there is not an immediate adaptation wherein we can determine $\phi[1]$ exactly using dihedral invariants. Instead, we observe that when the dimension is $d = 2^k$, a consequence of Lemma 5.3.3 is that one may use the phase $\phi[2^m]$ and the entry $T_{\text{dMRA}}^{(3)}[\imath 2^{m-1}, 2^{m-1}, d - 2^m]$ to recover the phase $\phi[2^{m-1}]$ by solving

$$2\phi[2^{m-1}] + \phi[d - 2^m] \equiv 2\phi[2^{m-1}] - \phi[2^m] \equiv \pm \cos^{-1}(\alpha_{2^{m-1}, 2^{m-1}, 2^m}) \pmod{2d}$$

for any choice of sign and for $m = k - 1, k - 2, \dots, 1$ until $\phi[1]$ is found. Our proof uses this observation for $d = 2^2$ in a single step.

5.5.2. Frequency Marching in Projected MRA. In this section, we will demonstrate that frequency marching can still be used in the projected MRA model as an algorithm for the base case.

First, we will explain why the orbits of signals of small dimension cannot be uniquely identified. Next, we show in Theorem 5.5.2 that for $x \in \mathbb{R}^8 \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ one may narrow down the orbit of $\hat{x}$ to one of four possibilities using a frequency marching-like procedure. Finally, using results of Edidin and Katz [31] but applied to the pMRA model, we show in Theorem 5.5.6 that for $x \in \mathbb{R}^{16} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ the third moment separates generic dihedral orbits.

Unidentifiability for small dimensions. For $d = 2$ ($x \in \mathbb{R}^4 \cap \text{span}(1, -1, 1, -1)^\perp$), examination of the entries of $T_{\text{pMRA}}^{(3)}$ reveals that the phase of $\phi[1]$ (and $\phi[3]$) cannot be determined. This occurs because the only two non-zero entries are $T_{\text{pMRA}}^{(3)}[\zeta 0, 0, 0]$ and $T_{\text{pMRA}}^{(3)}[\zeta 0, 1, 1]$. Each multiset contains a 0 index and as a result these elements are invariants of the form $\hat{x}[0]\hat{x}[j]\hat{x}[-j]$ and contain no information about the phase of $\hat{x}[1]$.

Modified frequency marching. Next, for $d = 4$ ($x \in \mathbb{R}^8 \cap \text{span}(1, -1, \dots, 1, -1)^\perp$), numerical tests we performed indicated that for a randomly generated real signal there were three additional orbits which shared the first three moments with the true signal. We show that this is indeed the case.

THEOREM 5.5.2. *For $x \in \mathbb{R}^8 \cap \text{span}(1, -1, 1, \dots, -1)^\perp$ with $\hat{x}[j] \neq 0$ for all $j \neq 4$, the moment $T_{\text{pMRA}}^{(3)}(\hat{x})$ determines a set of at most four $\mathcal{D}_8$ -orbits, one of which is the orbit of $\hat{x}$.*

PROOF. By similar observations to those in the proof of Theorem 5.5.1 using Proposition 5.2.5, we can see that

$$\begin{aligned} T_{\text{dMRA}}^{(3)}[\zeta 0, 0, 0] &= 8w_{\zeta 0,0,0}\hat{x}[0]^3 \quad \text{and} \\ T_{\text{dMRA}}^{(3)}[\zeta 0, j, j] &= 4w_{\zeta 0,j,j}e^{\pi i(0+j+j)/8}\hat{x}[0]\hat{x}[j]\hat{x}[-j], \end{aligned}$$

so $\hat{x}[0]$ and the Fourier coefficient magnitudes are given by

$$\begin{aligned} \hat{x}[0] &= \sqrt[3]{T_{\text{dMRA}}^{(3)}[\zeta 0, 0, 0]/(8w_{\zeta 0,0,0})} \quad \text{and} \\ |\hat{x}[j]|^2 &= \frac{|T_{\text{dMRA}}^{(3)}[\zeta 0, j, j]|}{4w_{\zeta 0,j,j}\hat{x}[0]}, \quad j = 1, 2, 3. \end{aligned}$$

Next, there are only three non-zero entries of the third moments which do not contain a zero index:

$$\begin{aligned} T_{\text{pMRA}}^{(3)}[\zeta 1, 1, 2] &= e^{i\pi/2}w_{\zeta 1,1,2}(\hat{x}[1]^2\hat{x}[6] + \hat{x}[7]^2\hat{x}[2]), \\ T_{\text{pMRA}}^{(3)}[\zeta 1, 2, 3] &= e^{3i\pi/4}w_{\zeta 1,2,3}(\hat{x}[1]\hat{x}[2]\hat{x}[5] + \hat{x}[7]\hat{x}[6]\hat{x}[3]), \\ T_{\text{pMRA}}^{(3)}[\zeta 2, 3, 3] &= -w_{\zeta 2,2,3}(x[2]\hat{x}[3]^2 + \hat{x}[6]\hat{x}[5]^2). \end{aligned}$$

Using $\hat{x}[j] = e^{i\phi[j]}|\hat{x}[j]|$ with $\phi[j] \in [0, 2\pi)$ and $\phi[8-j] = -\phi[j]$ for $j = 1, 2, 3$, these three entries allow us to determine the following three relationships between phases:

$$\begin{aligned} a_1 &:= \frac{T_{\text{pMRA}}^{(3)}[\zeta 1, 1, 2\zeta]}{2e^{i\pi/2}w_{\zeta 1, 1, 2\zeta}|\hat{x}[1]|^2|\hat{x}[2]|} = \cos(2\phi[1] - \phi[2]), \\ a_2 &:= \frac{T_{\text{pMRA}}^{(3)}[\zeta 1, 2, 3\zeta]}{2e^{3i\pi/4}w_{\zeta 1, 1, 2\zeta}|\hat{x}[1]||\hat{x}[2]||\hat{x}[3]|} = \cos(\phi[1] + \phi[2] - \phi[3]), \\ a_3 &:= -\frac{T_{\text{pMRA}}^{(3)}[\zeta 2, 2, 3\zeta]}{2w_{\zeta 2, 2, 3\zeta}|\hat{x}[2]|^2|\hat{x}[3]|} = \cos(\phi[2] + 2\phi[3]). \end{aligned}$$

Setting $\alpha_j := \arccos(a_j)$ for $j = 1, 2, 3$, this determines a system of three equations up to a sign and multiple of 2π :

$$\begin{aligned} 2\phi[1] - \phi[2] &= \pm\alpha_1 + 2\pi k_1, \\ \phi[1] + \phi[2] - \phi[3] &= \pm\alpha_2 + 2\pi k_2, \\ \phi[2] + 2\phi[3] &= \pm\alpha_3 + 2\pi k_3. \end{aligned}$$

A global change of sign $s_g \in \{-1, 1\}$ for these equations corresponds to the action of $\hat{J}$ on $\hat{x}$ so we need two additional sign variables, $s_1, s_2 \in \{-1, 1\}$ to produce all possible systems. Hence, all solutions are given via the assignments

$$\begin{aligned} \phi[1] &\leftarrow s_g \frac{1}{8} (3(s_1\alpha_1 + 2\pi k_1) + 2(s_2\alpha_2 + 2\pi k_2) + (\alpha_3 + 2\pi k_3)), \\ \phi[2] &\leftarrow s_g \frac{1}{4} (-(s_1\alpha_1 + 2\pi k_1) + 2(s_2\alpha_2 + 2\pi k_2) + (\alpha_3 + 2\pi k_3)), \\ \phi[3] &\leftarrow s_g \frac{1}{8} ((s_1\alpha_1 + 2\pi k_1) - 2(s_2\alpha_2 + 2\pi k_2) + 3(\alpha_3 + 2\pi k_3)), \end{aligned}$$

and $\phi[8-j] \leftarrow -\phi[j] \pmod{2\pi}$ for $j = 1, 2, 3$. Observe that if we fix the sign variables s_1, s_2 and set $s_g = 1$, $k_1 = k_2 = k_3 = 0$ to produce a solution $y(s_1, s_2) \in \mathbb{C}^8$, then each choice of s_g, k_1, k_2 and k_3 produces a solution $z = \hat{J}^{(1-s_g)/2} \hat{R}_{(3k_1+2k_2+k_3)} \cdot y(s_1, s_2)$. It follows that each choice (s_1, s_2) produces an orbit, while choices of (s_g, k_1, k_2, k_3) enumerate the 16 elements of the orbit. Since these are all the possible assignments to produce signals y such that $T_{\text{pMRA}}^{(3)}(y) = T_{\text{pMRA}}^{(3)}(\hat{x})$, at least one choice of s_1 and s_2 must produce the orbit of $\hat{x}$. $\square$

Identifiability for larger signals. Next, for $d = 8$ (where $x \in \mathbb{R}^{16} \cap \text{span}(1, -1, 1, \dots, -1)^\perp$), we will show that a generic signal's orbit is determined by the exact projected MRA third moment. To show this we will need a few constructions and results from Edidin and Katz [31].

Define an $O(2)$ -group action on a real vector $x \in \mathbb{R}^{2d}$ in the Fourier basis generated by the reflection element $\hat{J}$ defined in (5.10) and by a now *continuous* cyclic shift $\hat{R}_\theta$ for $\theta \in [0, 2\pi)$ acting on a signal $\hat{x} \in \mathbb{C}^{2d}$ in the Fourier basis entry-wise by

$$(5.39) \quad (\hat{R}_\theta \cdot \hat{x})[j] = \begin{cases} e^{-ij\theta} \hat{x}[j] & j = 0, 1, \dots, d-1, \\ e^{i(2d-j)\theta} \hat{x}[j] & j = d, d+1, \dots, 2d-1. \end{cases}$$

Observe that when θ a multiple of $\frac{2\pi}{2d}$, the action is that of $\mathcal{D}_{2d}$ as defined in (5.9). Define the *restricted action* of $O(2)$ to be the action on $\hat{x}$ where we have excluded the zeroth Fourier coefficient $\hat{x}[0]$ and Nyquist component $\hat{x}[d]$, i.e., the $O(2)$ -action restricted to a subspace spanned by standard basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{d-1}, \vec{e}_{d+1}, \dots, \vec{e}_{2d-1}$.

REMARK 5.5.3. *Our definition of the action of $O(2)$ in (5.39) differs from the one given by Edidin and Katz in [31, Section 3] since they index the Fourier basis elements using negative indices. In addition, a difference in sign in the phase produces a rotation by θ that is in the opposite direction as Edidin and Katz's definition. These differences in convention will not affect the results.*

Drawing from Edidin and Katz's calculations in [31, Section 3.1.2], in the context of our setup the degree-2 $O(2)$ -invariant polynomials are

$$\hat{x}[j]\hat{x}[-j] = |\hat{x}[j]|^2, \quad 0 \leq j \leq d$$

(with indices taken modulo $2d$), and the degree-3 $O(2)$ -invariant polynomials are

$$\begin{aligned} & \hat{x}[j_1]\hat{x}[j_2]\hat{x}[-j_1-j_2] + \hat{x}[-j_1]\hat{x}[-j_2]\hat{x}[j_1+j_2] \\ &= \hat{x}[j_1]\hat{x}[j_2]\hat{x}[-j_1-j_2] + \overline{\hat{x}[j_1]\hat{x}[j_2]\hat{x}[-j_1-j_2]}, \quad 0 \leq j_1 \leq j_2 \leq d, \quad 0 \leq j_1+j_2 \leq d \end{aligned}$$

(with indices taken modulo $2d$). The invariant polynomials for the restricted $O(2)$ -action exclude the polynomials containing the terms $\hat{x}[0]$ and $\hat{x}[d]$.

THEOREM 5.5.4 (Theorem 3.2 [31]). *When $x \in \mathbb{R}^{2d}$ for $d \neq 4$, the invariants of degree two and three for the restricted action of $O(2)$ separate generic orbits.*

To prove that the invariants from the pMRA model separate generic orbits when $d = 8$ (meaning the signal has length 16) we will first show in Lemma 5.5.5 that we actually have access to the dMRA invariants, except for those involving the term $\hat{x}[d]$. If we had all the invariants, we could simply use Edidin and Katz's results as presented.

LEMMA 5.5.5. Let $x \in \mathbb{R}^{16} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ with Fourier transform $\hat{x}$, and let $\wr j_1, j_2, j_3 \wr$ be a multiset of $[16] \setminus \{8\}$ such that $j_1 + j_2 + j_3 \equiv 0 \pmod{16}$. Then there is a multiset $\wr j'_1, j'_2, j'_3 \wr$ of $[8]$ such that

$$(5.40) \quad T_{\text{pMRA}}^{(3)}(\hat{x})[\wr j'_1, j'_2, j'_3 \wr] = K(\wr j'_1, j'_2, j'_3 \wr) (-1)^{(j_1 + j_2 + j_3)/16} e^{\pi i(j'_1 + j'_2 + j'_3)/16} T_{\text{dMRA}}^{(3)}(\hat{x})[\wr j_1, j_2, j_3 \wr],$$

where $K(\wr j'_1, j'_2, j'_3 \wr)$ is 8 when $\wr j'_1, j'_2, j'_3 \wr = \wr 0, 0, 0 \wr$, 4 when $\wr j'_1, j'_2, j'_3 \wr$ has the form $\wr 0, j, j \wr$, and 2 otherwise.

PROOF. We split into cases based on the number of zero indices in the multi-set $\wr j'_1, j'_2, j'_3 \wr$. Due to the modular condition, non-zero entries in $T_{\text{pMRA}}^{(3)}$ appear when $\wr j'_1, j'_2, j'_3 \wr = \wr 0, 0, 0 \wr$, when $\wr j'_1, j'_2, j'_3 \wr = \wr 0, j, j \wr$ for some $j \in [16] \setminus \{0, 8\}$, or when every entry is non-zero. The first two cases follow by inspection. By the first expression from Proposition 5.2.5, if $\wr j_1, j_2, j_3 \wr$ is a multiset of $[16]$ such that $j_1 + j_2 + j_3 \equiv 0 \pmod{16}$ then

$$T_{\text{dMRA}}^{(3)}(\hat{x})[\wr j_1, j_2, j_3 \wr] = \frac{w_{\wr j_1, j_2, j_3 \wr}}{2} (\hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3]).$$

First, for $\wr j_1, j_2, j_3 \wr = \wr 0, 0, 0 \wr$, using the second expression from Proposition 5.2.5, we may simplify to get

$$T_{\text{pMRA}}^{(3)}(\hat{x})[\wr 0, 0, 0 \wr] = 8w_{\wr 0, 0, 0 \wr} \hat{x}[0]^3 = 8 \cdot T_{\text{dMRA}}^{(3)}(\hat{x})[\wr 0, 0, 0 \wr]$$

and so (5.40) holds for this multiset index. Second, for $j \in \{1, 2, \dots, 7\}$,

$$T_{\text{pMRA}}^{(3)}(\hat{x})[\wr 0, j, j \wr] = 4e^{\pi i(0+j+j)/16} w_{\wr 0, j, j \wr} \hat{x}[0] \hat{x}[j] \hat{x}[-j] = 4e^{\pi i(0+j+j)/16} \cdot T_{\text{dMRA}}^{(3)}(\hat{x})[\wr 0, j, -j \wr]$$

with indices taken modulo 16, so (5.40) holds for these multiset indices. Third, with the bulk of cases remaining, we will show that for any multiset $\wr j_1, j_2, j_3 \wr$ of $[16] \setminus \{0, 8\}$ there is a multiset $\wr j'_1, j'_2, j'_3 \wr$ of $[8] \setminus \{0\}$ so that (5.40) holds. To this end, define a map on indices $f : [16] \setminus \{0, 8\} \rightarrow [8] \setminus \{0\}$ by

$$f(j) = \begin{cases} j & \text{if } j < 8, \\ 16 - j & \text{otherwise.} \end{cases}$$

We extend f to multisets $\wr j_1, j_2, j_3 \wr$ of $[16] \setminus \{0, 8\}$ elementwise by $f(M) = \wr f(m) : m \in M \wr$. For indices $j_1, j_2, j_3 \in [8] \setminus \{0\}$, we may then rewrite the second expression from Proposition 5.2.5 using a summation over the pre-image of f ,

$$\begin{aligned} T_{\text{pMRA}}^{(3)}(\hat{x})[\wr j'_1, j'_2, j'_3 \wr] &= \frac{w_{\wr j'_1, j'_2, j'_3 \wr}}{2} \sum_{\wr j_1, j_2, j_3 \wr \in f^{-1}(\wr j'_1, j'_2, j'_3 \wr)} \left(\phi(\wr j_1, j_2, j_3 \wr, \wr j'_1, j'_2, j'_3 \wr) \right. \\ &\quad \cdot (\hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3]) \mathbb{1} \left\{ \sum_{k=1}^r j_k \equiv 0 \pmod{16} \right\} \Bigg) \end{aligned}$$

where

$$\phi(\wr_{j_1, j_2, j_3}, \wr_{j'_1, j'_2, j'_3}) = (-1)^{(j_1 + j_2 + j_3)/16} e^{\pi i(j'_1 + j'_2 + j'_3)/16}.$$

Defining the set

$$Z = \{\wr_{j_1, j_2, j_3} : j_1, j_2, j_3 \in [16] \setminus \{0, 8\}, j_1 + j_2 + j_3 \equiv 0 \pmod{16}\},$$

we may absorb the indicator function into the sum condition to get

$$T_{\text{pMRA}}^{(3)}(\hat{x})[\wr_{j'_1, j'_2, j'_3}] = \frac{w_{\wr_{j'_1, j'_2, j'_3}}}{2} \sum_{\wr_{j_1, j_2, j_3} \in f^{-1}(\wr_{j'_1, j'_2, j'_3}) \cap Z} \phi(\dots) (\hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3]).$$

However, note that (a) $f^{-1}(\wr_{j'_1, j'_2, j'_3}) \cap Z = (f|_Z)^{-1}(\wr_{j'_1, j'_2, j'_3})$, and (b) the summand is invariant under the substitution $\wr_{j_1, j_2, j_3} \mapsto \wr_{16-j_1, 16-j_2, 16-j_3}$. We will show that, in fact,

$$f|_Z^{-1}(\wr_{j'_1, j'_2, j'_3}) = \{\wr_{j_1, j_2, j_3}, \wr_{16-j_1, 16-j_2, 16-j_3}\}$$

so that summation simplifies to

$$T_{\text{pMRA}}^{(3)}(\hat{x})[\wr_{j'_1, j'_2, j'_3}] = 2(-1)^{(j_1 + j_2 + j_3)/16} e^{\pi i(j'_1 + j'_2 + j'_3)/16} T_{\text{pMRA}}^{(3)}(\hat{x})[\wr_{j_1, j_2, j_3}].$$

First, if $f|_Z(\wr_{j_1, j_2, j_3}) = \wr_{j'_1, j'_2, j'_3}$ then it is simple to see that $f|_Z^{-1}(\wr_{j'_1, j'_2, j'_3}) \supseteq \{\wr_{j_1, j_2, j_3}, \wr_{16-j_1, 16-j_2, 16-j_3}\}$. We will show by contradiction these are the only two multisets in the pre-image. Suppose there is some third multiset $K = \wr_{k_1, k_2, k_3}$ so that $K \in Z$ and $f(K) = \wr_{j'_1, j'_2, j'_3}$. Then by the definition of f , there will be some permutation $\pi \in S_3$ such that either $k_{\pi(i)} = j_i$ or $k_{\pi(i)} = 16 - j_i$ for each $i = 1, 2, 3$. Without loss of generality we may assume that $\pi = \text{id}$. It must be that either $k_1 = 16 - j_1$, $k_2 = j_2$ and $k_{\pi(3)} = j_3$, or $k_1 = j_1$, $k_{\pi(2)} = 16 - j_2$ and $k_3 = 16 - j_3$ for some $\pi \in S_3$. We may assume the first since if $K \in f|_Z^{-1}(\wr_{j'_1, j'_2, j'_3})$ then $\wr_{16-k_1, 16-k_2, 16-k_3} \in f|_Z^{-1}(\wr_{j'_1, j'_2, j'_3})$. But since $K \in Z$, $k_1 + k_2 + k_3 = 16 - j_1 + j_2 + j_3 \equiv 0 \pmod{16}$ and $\{j_1, j_2, j_3\} \in Z$ gives $j_1 + j_2 + j_3 \equiv 0 \pmod{16}$, we find that $(j_1 + j_2 + j_3) - (16 - j_1 + j_2 + j_3) \equiv 0 \pmod{16}$ which implies $2j_1 \equiv 0 \pmod{16}$ but then $j_1 \equiv 0 \pmod{8}$ and $j_1 = 0$ or $j_1 = 8$ so $\wr_{j_1, j_2, j_3} \notin Z$ which contradicts our initial assumptions. $\square$

Lemma 5.5.5 shows that we can recover (most of) the dihedral MRA 3rd moment from the projected MRA 3rd moment. Using this and Theorem 5.5.4 we will verify generic identifiability in the pMRA model in the base case $d = 8$.

THEOREM 5.5.6. *For generic $x \in \mathbb{R}^{16} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$, the moment $T_{\text{pMRA}}^{(3)}(\hat{x})$ uniquely determines the $\mathcal{D}_{16}$ -orbit of x .*

PROOF. Let $x \in \mathbb{R}^{16} \cap \text{span}(1, -1, \dots, 1, -1)^\perp$ be a generic vector. Now suppose that there is a $y \in \mathbb{R}^{16} \cap \text{span}(1, \dots, -1)^\perp$ such that $T_{\text{pMRA}}^{(3)}(\hat{x}) = T_{\text{pMRA}}^{(3)}(\hat{y})$. We will show that y is in the $\mathcal{D}_{16}$ -orbit of x . First, by Lemma 5.5.5, we have that

$$T_{\text{pMRA}}^{(3)}(\hat{x}) = T_{\text{pMRA}}^{(3)}(\hat{y}) \implies T_{\text{dMRA}}^{(3)}(\hat{x})[\imath j_1, j_2, j_3] = T_{\text{dMRA}}^{(3)}(\hat{y})[\imath j_1, j_2, j_3] \text{ for all } j_1, j_2, j_3 \in [16] \setminus \{8\}.$$

Observe that since

$$T_{\text{dMRA}}^{(3)}(\hat{x})[\imath 0, 0, 0] = w_{\imath 0, 0, 0} \hat{x}[0]^3$$

we can infer that $\hat{x}[0] = \hat{y}[0]$, i.e. x and y have the same mean value. Next, for $j \in [16] \setminus \{8\}$ since x is real, $\hat{x}[-j] = \overline{\hat{x}[j]}$ and since

$$T_{\text{dMRA}}^{(3)}(\hat{x})[\imath 0, j, -j] = w_{\imath 0, j, -j} \hat{x}[0] \hat{x}[j] \hat{x}[-j] = w_{\imath 0, j, -j} \hat{x}[0] |\hat{x}[j]|^2,$$

$|\hat{x}[j]| = |\hat{y}[j]|$ for $j \in [8]$, i.e. x and y have the same power spectrum. It remains to show that the phases of $\hat{x}[j]$ and $\hat{y}[j]$ must be related by a dihedral group action. We will first show that $\hat{x}$ and $\hat{y}$ must be in the same $O(2)$ -orbit. Observe that for $j_1, j_2, j_3 \in [16] \setminus \{0, 8\}$ with $j_1 + j_2 - j_3 = 0$,

$$T_{\text{dMRA}}^{(3)}(\hat{x})[\imath j_1, j_2, j_3] = \frac{w_{\imath j_1, j_2, j_3}}{2} (\hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3]),$$

so the $O(2)$ -invariant polynomials for $\hat{x}$ and $\hat{y}$ have the same values. By Theorem 5.5.4, these invariants separate generic $O(2)$ -orbits, so we know that $\hat{x}$ and $\hat{y}$ are in the same $O(2)$ -orbit, i.e. $\hat{y} = \hat{R}_\theta \hat{x}$ or $\hat{y} = \hat{J} \hat{R}_\theta \hat{x}$ for some $\theta \in [0, 2\pi)$. Without loss of generality assume the first occurs. In our final step, we will resolve $\theta = \frac{2\pi k}{16}$ for some $k \in [16]$ using the same argument as in [31, Section 3.3], which implies $\hat{x}$ and $\hat{y}$ are in the same $\mathcal{D}_{16}$ -orbit. Pick $j_1 + j_2 + j_3 = 16$ with $0 < j_1, j_2, j_3 < 8$ (e.g. $j_1 = 3, j_2 = 6$ and $j_3 = 7$). Then

$$\begin{aligned} T_{\text{dMRA}}^{(3)}(\hat{x})[\imath j_1, j_2, j_3] &= T_{\text{dMRA}}^{(3)}(\hat{y})[\imath j_1, j_2, j_3] \\ \implies \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3] &= \hat{y}[j_1] \hat{y}[j_2] \hat{y}[j_3] + \hat{y}[-j_1] \hat{y}[-j_2] \hat{y}[-j_3] \\ \implies \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3] &= e^{-i\theta(j_1+j_2+j_3)} \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + e^{i\theta(j_1+j_2+j_3)} \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3] \\ \implies \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3] &= e^{-16i\theta} \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3] + e^{16i\theta} \hat{x}[-j_1] \hat{x}[-j_2] \hat{x}[-j_3]. \end{aligned}$$

Define $\phi \in [0, 2\pi)$ to be such that $|\hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3]| e^{-i\phi} = \hat{x}[j_1] \hat{x}[j_2] \hat{x}[j_3]$. Then

$$\begin{aligned} e^{-i\phi} + e^{i\phi} &= e^{-i\phi} e^{-16i\theta} + e^{-i\phi} e^{-16i} \\ \implies \cos(\phi) &= \cos(\phi + 16\theta). \end{aligned}$$

Since ϕ is generic, this equality only holds when $16\theta = 2\pi k$ for some $k \in [16)$, so $\theta = \frac{2\pi k}{16}$. Thus, $\hat{y}$ is in the $\mathcal{D}_{16}$ -orbit of $\hat{x}$. Hence, the third order invariants from projected MRA separate generic orbits. $\square$

Algorithmic implications. While unique recovery of an orbit is not possible for signals of length 8, we use frequency marching in 8 dimensions ($d = 4$) as demonstrated in Theorem 5.5.2 as a base case for our recursive method to recover a set of four orbit representatives in $\mathbb{C}^8$ from the dMRA third moment. If each of these is used as the input to Algorithm 4 to recover some larger signal in $\mathbb{C}^{16}$, Theorem 5.5.1 then ensures that any solutions to the variety-constrained linear system belong to the same orbit. Alternatively if any of the other three orbit representatives fail to extend to an orbit representative in $\mathbb{C}^{16}$, they may simply be discarded. See the next section and proof of Theorem 5.1.3 for a rigorous explanation.

5.6. Putting it Together: Proofs of Main Results

First, we define the matrix $M = M(d)$ referenced in Theorem 5.1.2. For $d = 2^k$ and $x \in \mathbb{R}^{2^d}$ we give this matrix as the direct sum of matrices

$$(5.41) \quad M(\hat{x}) := M_{\text{dMRA}}(\hat{x}[K_3]) \oplus M_{\text{dMRA}}(\hat{x}[K_4]) \oplus \cdots \oplus M_{\text{dMRA}}(\hat{x}[K_k]) \oplus M_{\text{dMRA}}(\hat{x}),$$

where $K_m := \{2^{k+1-m}j : j \in [2^m)\}$ so that each $\hat{x}[K_m]$ is a subsequence of 2^m Fourier coefficients of x . We omit the dependence on $\hat{x}$ (as in the statement of Theorem 5.1.2) when considering M to be a symbolic matrix in the $2^k + 1$ real and $2^k - 1$ imaginary components of $\hat{x}$ in the same sense as Remark 5.4.4.

PROOF OF THEOREM 5.1.2. First, observe that the effect of extracting even Fourier coefficients repeatedly has the same effect as indexing by directly by the sets K_m in that $\hat{x}[K_k] = \hat{x}[E_{2 \cdot 2^k}]$ and $(\hat{x}[K_m])[E_{2^m}] = \hat{x}[K_{m-1}]$ where E_d is defined in (5.15). By Proposition 5.3.3, one may extract the subtensor of $T_{\text{dMRA}}(\hat{x})$ which will have the form $T_{\text{dMRA}}^{(3)}(\hat{x}[E_{2 \cdot 2^k}]) = T_{\text{dMRA}}^{(3)}(\hat{x}[K_k])$. Applying this repeatedly one may then extract sequence of sub-tensors of the form

$$T_{\text{dMRA}}^{(3)}(\hat{x}[K_2]), T_{\text{dMRA}}^{(3)}(\hat{x}[K_3]), \dots, T_{\text{dMRA}}^{(3)}(\hat{x}[K_k]), T_{\text{dMRA}}^{(3)}(\hat{x}).$$

For the sub-tensor $T_{\text{dMRA}}^{(3)}(\hat{x}[K_2]) \in (\mathbb{C}^4)^{\otimes 3}$, by Theorem 5.5.1 there is an efficient method for finding a representative y_2 in the $\mathcal{D}_4$ -orbit of $\hat{x}[K_2]$. Since x is generic and under the assumption that the matrix M has full column-rank symbolically, the instance $M(\hat{x})$ has full column-rank and as a result each of

$$M_{\text{dMRA}}(\hat{x}[K_3]), M_{\text{dMRA}}(\hat{x}[K_4]), \dots, M_{\text{dMRA}}(\hat{x}[K_k]), \text{ and } M_{\text{dMRA}}(\hat{x})$$

has full column-rank. Further, we note that for each $g \in \mathcal{D}_{2 \cdot 2^k}$, the instances $M(g \cdot \hat{x})$ will have full column-rank as a consequence of Proposition 5.4.2. Theorem 5.4.5 then guarantees that Algorithm 3 run on inputs y_2 and $T_{\text{dMRA}}^{(3)}(\hat{x}[K_3])$ produces y_3 in the $\mathcal{D}_8$ -orbit of $\hat{x}[K_3]$. At each stage, this process may be repeated to recover y_m the $\mathcal{D}_{2^m}$ -orbit of $\hat{x}[K_m]$ until the final step recovers y in the $\mathcal{D}_{2d}$ -orbit of $\hat{x}$.

Finally, since the method for initializing y_2 as described in the proof of Theorem 5.5.1 takes constant time and Algorithm 3 runs in polynomial time, the recursive procedure as a whole runs in polynomial time. $\square$

Similar to the dihedral MRA case, we define the matrix $M' = M'(d)$ referenced in Theorem 5.1.3. For $d = 2^k$ and $x \in \mathbb{R}^{2d} \cap \text{span}(1, -1, \dots, 1, -1)$, we give this matrix as the direct sum of matrices

$$(5.42) \quad M'(\hat{x}) := M_{\text{pMRA}}(\hat{x}[K_4]) \oplus M_{\text{pMRA}}(\hat{x}[K_5]) \oplus \dots \oplus M_{\text{pMRA}}(\hat{x}[K_k]) \oplus M_{\text{pMRA}}(\hat{x})$$

where K_m is defined as before. We will also omit the dependence on $\hat{x}$ (as in the statement of Theorem 5.1.3) when considering M' to be a symbolic matrix in the 2^k real and $2^k - 1$ imaginary components of $\hat{x}$ in the same sense as in Remark 5.4.11.

PROOF OF THEOREM 5.1.3. In a similar manner to that in the proof of Theorem 5.1.2, by Proposition 5.3.3, we may extract a sequence of sub-tensors of the form

$$T_{\text{pMRA}}^{(3)}(\hat{x}[K_3]), T_{\text{pMRA}}^{(3)}(\hat{x}[K_4]), \dots, T_{\text{pMRA}}^{(3)}(\hat{x}[K_k]), T_{\text{pMRA}}^{(3)}(\hat{x}).$$

For the sub-tensor $T_{\text{dMRA}}^{(3)}(\hat{x}[K_3]) \in (\mathbb{C}^8)^{\otimes 3}$, by Theorem 5.5.2 there is an efficient method for finding a set of four vectors $y_{3,1}, y_{3,2}, y_{3,3}$ and $y_{3,4} \in \mathbb{C}^8$ at least one of which (say $y_{3,1}$ without loss of generality) is in the $\mathcal{D}_4$ -orbit of $\hat{x}[K_2]$. Since x is generic and under the assumption that the matrix M' has full column-rank symbolically, the instance $M(\hat{x})$ has full column-rank and as a result each of

$$M_{\text{pMRA}}(\hat{x}[K_4]), M_{\text{pMRA}}(\hat{x}[K_5]), \dots, M_{\text{pMRA}}(\hat{x}[K_k]), \text{ and } M_{\text{pMRA}}(\hat{x})$$

has full column-rank. Theorem 5.4.12 then guarantees that Algorithm 4 run on inputs $y_{3,1}$ and $T_{\text{dMRA}}^{(3)}(\hat{x}[K_4])$ to produces $y_4 \in \mathbb{C}^{16}$ in the $\mathcal{D}_{16}$ -orbit of $\hat{x}[K_4]$. For the remaining vectors, we have no guarantees that Algorithm 4 will succeed: if it fails on each of $y_{3,j}$ for any $j \in \{2, 3, 4\}$ before succeeding on $y_{3,1}$, there is no issue. However, if Algorithm 4 also succeeds for inputs $y_{3,j}$ and $T_{\text{pMRA}}^{(3)}(\hat{x}[K_4])$ for some $j \in \{2, 3, 4\}$ then it produces some y'_4 such that $b_{\text{pMRA}}(y'_4) = b_{\text{pMRA}}(\hat{x}[K_4])$. But then $T_{\text{pMRA}}(y'_4) = T_{\text{pMRA}}(\hat{x}[K_4])$. Since x is generic, so is the subset of Fourier coefficients $\hat{x}[K_4]$, and by Theorem 5.5.6 we can conclude that y'_4 is also in $\mathcal{D}_{16}$ -orbit of $\hat{x}[K_4]$. Hence, if Algorithm 4 does not output “Fail” for any inputs $y_{3,j}$, $j \in [4]$ and

$T_{\text{pMRA}}^{(3)}(\hat{x}[K_4])$, the output is guaranteed to lie in the $\mathcal{D}_{16}$ -orbit of $\hat{x}[K_4]$. From here the process is then the same as the dihedral case, where Algorithm 4 may be repeatedly applied to recover y_m the $\mathcal{D}_{2^m}$ -orbit of $\hat{x}[K_m]$ until the final step recovers y in the $\mathcal{D}_{2^d}$ -orbit of $\hat{x}$.

Finally, since the method for finding the candidates $y_{3,1}, y_{3,2}, y_{3,3}$ and $y_{3,4}$ given in Theorem 5.5.2 takes constant time and Algorithm 3 runs in polynomial time, the recursive procedure as a whole runs in polynomial time. $\square$

5.7. Efficient Implementation

Our algorithms obscure some computational details in the `FINDVARIETYCONSTRAINEDSOLUTIONS` subroutine (our Algorithm 2) by using Algorithm 1 from [43]. However, we feel the process deserves more coverage since the statement of Algorithm 1 in [43] is given at a high level of abstraction.

We synthesize the application of Boussé et al. [19] to give a method for finding the solution(s) to $Au = b$ subject to $u \in \mathcal{X}_1^\vee$ where $A \in \mathbb{K}^{m \times n}$, $u \in \mathbb{K}^n$, using Algorithm 1 from [43].

Algorithm 2 (detailed procedure).

1. Find any particular solution x_p to $Au = b$ and $\{n_j\}_{j=1}^L$ a basis for the nullspace of A . Concatenate as columns the matrix

$$N = \begin{bmatrix} x_p & n_1 & n_2 & \cdots & n_L \end{bmatrix} \in \mathbb{K}^{n \times (L+1)}.$$

2. Compute a basis $\{c_j\}_{j=1}^s \subseteq (\mathbb{K}^{(L+1)})^{\otimes 2}$ for $\ker(\Phi N^{\otimes 2})$ where Φ has rows a basis for $I(\mathcal{X}_1^\vee)_2$. If $\Phi N^{\otimes 2}$ has full column-rank, report “No solutions.”
3. Reshape each vector to a symmetric matrix $c_j \mapsto C_j \in \mathbb{K}^{(L+1) \times (L+1)}$. Use $\{C_j\}_{j=1}^s$ as slabs of a partially symmetric tensor $\mathcal{T} \in \mathbb{K}^{(L+1) \times (L+1) \times s}$.
4. Run simultaneous diagonalization² on $\mathcal{T}$ so that $\mathcal{T} = \llbracket W, W, V \rrbracket$ or report “Fail” if simultaneous diagonalization fails.
5. Find the normalized weight matrix $\tilde{W} = W \text{diag}(W[1, :])^{-1}$ with all 1’s in the first row.
6. Solution(s) are the columns of $X = U\tilde{W}$.
7. (Optional) Matricize each column of X and compute its symmetric rank-1 decomposition.

In comparison with the original algorithm [43, Algorithm 1], in addition to our specific application which necessitates the normalization step (Step 5.), we note that it is possible to run simultaneous diagonalization

² If $s = 1$, the tensor degenerates to a single matrix $C_1 = \llbracket W, W, V \rrbracket = \llbracket w, w, \lambda \rrbracket = \lambda w w^\top$ for $w \in \mathbb{K}^n$ and $\lambda \in \mathbb{R}^+$ so simultaneous diagonalization becomes a symmetric rank-1 factorization in this special case.

directly on the reshaping of the *coefficient* vectors $\{c_j\}_{j=1}^s$ and save space and time running simultaneous diagonalization on $(L+1) \times (L+1)$ matrices rather than $n \times n$ matrices so that the complexity of this step scales with the dimension of the subspace, not the ambient space. Only in the final step do we change back to the standard basis after which one may compute decompositions of the planted solutions.

Unfortunately, in general $\Phi N^{\otimes 2}$ may be a very large matrix with a small kernel (one-dimensional in the case of unique recovery). In our applications we also can produce a very sparse basis for our nullspace and Φ is also sparse by construction. To take better advantage of these structural properties, we will give an alternative method which converts the problem into a top right singular vector computation.

First, define the set of all possible splittings of a multiset $\wr a, b, c, d \wr$ of $[d]$ into two multisets of two elements as

$$\text{split}(\wr a, b, c, f \wr) = \{\{\wr j, k \wr, \wr \ell, m \wr\} : \wr j, k, \ell, m \wr = \wr a, b, c, f \wr\}.$$

DEFINITION 5.7.1. *Let $\mathcal{Q}$ be the set of vectors*

$$Q_{abcf}^\vee = \sum_{\{\wr j, k \wr, \wr \ell, m \wr\} \in \text{split}(\wr a, b, c, f \wr)} (w_{\wr j, k \wr, \wr \ell, m \wr} w_{\wr j, k \wr} w_{\wr \ell, m \wr}) \vec{e}_{\wr j, k \wr, \wr \ell, m \wr}.$$

PROPOSITION 5.7.2. *$\mathcal{Q}$ is an orthogonal basis for $I(\mathcal{X}_1^\vee)_2^\perp$ which has dimension $\binom{d}{4}$.*

See Section 5.9.3 for the proof. Using this, we may replace Steps 1. and 2. with Steps 1* and 2*:

- 1*. Find $\tilde{x}_p$, the least-squares solution to $Au = b$ (given by $x_p = A^+b$ where A^+ is the psuedoinverse of A) and find $\{\tilde{n}_j\}_{j=1}^L$ an *orthonormal* basis for the nullspace of A . Concatenate as columns the matrix

$$\tilde{N} = \begin{bmatrix} \frac{\tilde{x}_p}{\|\tilde{x}_p\|_2} & \tilde{n}_1 & \tilde{n}_2 & \cdots & \tilde{n}_L \end{bmatrix} \in \mathbb{K}^{d \times (L+1)}.$$

- 2*. Compute a basis $\{c_j\}_{j=1}^s \subseteq (\mathbb{K}^{(L+1)})^{\otimes 2}$ of the *right singular subspace* of $\Phi_\perp \tilde{N}^{\otimes 2}$ associated with the top singular value of 1 where $\Phi_\perp$ has rows an orthonormal basis for $I(\mathcal{X}_1^\vee)_2^\perp$.
If each singular value of $\Phi_\perp \tilde{N}^{\otimes 2}$ is less than 1, report “No solutions.”

Most importantly, we claim that these two methods are equivalent.

PROPOSITION 5.7.3. *The basis produced by steps 1. and 2. and the basis produced by steps 1* and 2* span the same subspace.*

The orthogonal basis $\mathcal{Q}$ can be easily normalized. The advantage of Steps 1* and 2* is that we can find top singular vectors — a method that can be done efficiently using power iteration — instead of finding the

kernel of a massive matrix. By Lemmas 5.4.6 and 5.4.13 bases for the nullspaces needed for the variety-constrained linear systems may be made very sparse (a constant number of entries per row) and orthogonal. This simplifies finding an orthonormal basis in step 1* and the sparsity significantly improves matrix-vector the efficiency of multiplication against $\Phi_{\perp} \tilde{N}^{\otimes 2}$.

THEOREM 5.7.4. *The matrix-vector products $\Phi_{\perp} \tilde{N}^{\otimes 2} v$ and $w^{\top} \Phi_{\perp} \tilde{N}^{\otimes 2}$ may be performed in $O(d^4)$ floating-point operations.*

The number of operations is the same order as the number of rows or columns of $\Phi_{\perp} \tilde{N}^{\otimes 2}$. Without a sparse construction for $\tilde{N}$, a naive matrix-vector multiplication would require $O(d^8)$ operations.

PROOF OF THEOREM 5.7.4. For both the dMRA and pMRA model, $\tilde{N}$ has one dense column $\tilde{x}_p / \|\tilde{x}_p\|_2$. However, by Lemmas 5.4.6 and 5.4.13 the remaining columns have non-overlapping support and contain at most two non-zero entries. As a result, $\tilde{N}$ has at most two non-zero entries per row and d rows. It then follows from Lemma 4.2.2 that $\tilde{N}^{\otimes 2}$ will have at most 4 non-zero entries per row. Next, the rows of $\Phi_{\perp}$ are also sparse, containing at most 3 non-zero entries. The matrix multiplications $\Phi_{\perp} \tilde{N}^{\otimes 2}$ has rows which are linear combinations of the rows of $N^{\otimes 2}$ from which we can infer that each row contains no more than 12 non-zero entries. This gives no more than $23 \binom{d}{4}$ operations by summing scalar products and additions to compute the dot-product of v with every row of $\Phi_{\perp} \tilde{N}^{\otimes 2}$ for the right multiplication.

For the left multiplication, the first column of $\tilde{N}$ is dense with $\binom{d}{2}$ entries. As a result, one column of $\tilde{N}^{\otimes 2}$ contains $O(d^4)$ entries, $O(d^2)$ columns contain $O(d^2)$ entries and $O(d^4)$ columns contain a constant number of entries. Each row of $\Phi_{\perp}$ has a constant number of non-zero entries so performing the multiplication $\Phi_{\perp} \tilde{N}^{\otimes 2}$ required $O(d^4)$ arithmetic operations. Similarly, finding $w^{\top} \Phi_{\perp} \tilde{N}^{\otimes 2}$ will require $O(d^4)$ operations. $\square$

Next, to demonstrate the equivalence of these two methods we first define the principal angles and principal vectors between two subspaces $\mathcal{A}, \mathcal{B} \subseteq \mathbb{K}^n$ where $p \geq \dim(\mathcal{A}) \geq \dim(\mathcal{B}) \geq q$.

DEFINITION 5.7.5. *The principal angles θ_k and principal vectors $u_k \in \mathcal{A}, v_k \in \mathcal{B}$ between two subspaces are recursively defined for $k = 1, 2, \dots, q$ by*

$$\cos(\theta_k) = \max_{u \in \mathcal{A}} \max_{v \in \mathcal{B}} u^* v =: u_k^* v_k, \quad \|u\| = 1, \|v\| = 1$$

under the orthogonality constraints

$$u_j^* u = 0, \quad v_j^* v = 0 \quad j = 1, 2, \dots, k-1.$$

Our simplification arises from an efficient method for finding principles angles by Bjorck and Golub.

THEOREM 5.7.6 ([17, Theorem 1]). Let $Q_{\mathcal{A}}$ and $Q_{\mathcal{B}}$ have columns that are unitary bases for subspaces $\mathcal{A}, \mathcal{B} \subseteq \mathbb{K}^n$. If the $p \times q$ matrix

$$M = Q_{\mathcal{A}}^* Q_{\mathcal{B}}$$

has a singular value decomposition

$$M = X \Sigma Y^*, \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_q)$$

with $X^* X = Y^* Y = Y Y^* = I_q$ and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q$ then

$$\cos(\theta_k) = \sigma_k, \quad u_k = Q_{\mathcal{A}} X[:, k], \quad \text{and} \quad v_k = Q_{\mathcal{B}} Y[:, k], \quad k = 1, 2, \dots, q.$$

Note that the singular values of 1 correspond to principal angles of zero; in other words, we can extract these principal vectors to find a subspace intersection.

PROOF OF PROPOSITION 5.7.3. We will show that both $\ker(\Phi N^{\otimes 2})$ and the right singular subspace of $\Phi_{\perp} \tilde{N}^{\otimes 2}$ associated with the top singular value of 1, may be used to construct a basis for $\text{colspan}(N^{\otimes 2}) \cap I(\mathcal{X}_1^{\vee})_2^{\perp}$.

The first case is simplest. Suppose that $c \in \ker(\Phi N^{\otimes 2})$. Clearly $N^{\otimes 2} c \in \text{colspan}(N^{\otimes 2})$. But $\Phi(N^{\otimes 2} c) = \vec{0}$ so $N^{\otimes 2} c \in \ker(\Phi)$. By construction, $\text{rowspan}(\Phi) = I(\mathcal{X}_1^{\vee})_2$ so $\ker(\Phi) = I(\mathcal{X}_1^{\vee})_2^{\perp}$ and then $N^{\otimes 2} c \in I(\mathcal{X}_1^{\vee})_2^{\perp} \cap \text{colspan}(N^{\otimes 2})$.

Next, for the second case observe by Lemma 4.2.8 since $\tilde{N}$ is an orthonormal basis, so is $\tilde{N}^{\otimes 2}$. By Theorem 5.7.6, the SVD

$$\Phi_{\perp} \tilde{N}^{\otimes 2} = X \Sigma Y^*$$

gives the principal vectors between these subspaces via $u_k = \Phi_{\perp} X[:, k]$ and $v_k = \tilde{N}^{\otimes 2} Y[:, k]$. Suppose that $\sigma_1 = \sigma_2 = \dots = \sigma_s = 1$. Then the first s principal angles are zero and as a consequence of Definition 5.7.5 we have $u_k = v_k$ for $k = 1, 2, \dots, s$. Since $u_k \in I(\mathcal{X}_1^{\vee})_2^{\perp}$ and $v_k \in \text{colspan}(\tilde{N}^{\otimes 2}) = \text{colspan}(N^{\otimes 2})$ it follows that $v_k \in \text{colspan}(N^{\otimes 2}) \cap I(\mathcal{X}_1^{\vee})_2^{\perp}$ for $s = 1, 2, \dots, k$ and these vectors are a basis for the intersection. $\square$

5.7.1. Numerical Experiments with Full Algorithm. We implement Algorithms 3 and 4 using our efficient method for solving variety-constrained linear systems in Python. For our procedure, we generated real signals $x \in \mathbb{R}^{64}$ by selecting entries uniformly at random from a standard normal distribution. The exact third moments given by (5.12) and (5.13) for the dMRA and pMRA model are computed. Our program for reconstruction reads the invariants and applies the Algorithms 3 and 4 recursively. For the base case of the recursion in dihedral MRA, when the recursion reaches a signal length of 4 we use the frequency marching

procedure from Theorem 5.5.1. For the base case in projected MRA, when the recursion reaches a signal length of 8 for projected MRA we use our modified frequency marching procedure from Theorem 5.5.2 and the next step in the recursion determines which of the four orbit representatives extend to a length-16 orbit representative.

To measure any error in recovery, we measure the dihedral-invariant distance

$$(5.43) \quad d(\hat{x}, y) = \min_{g \in \mathcal{D}_{2d}} \|\hat{x} - g \cdot y\|_2$$

between the true set of Fourier coefficients $\hat{x}$ and the vector generated by our procedure y . In projected MRA, we omit the Nyquist component (the 32nd coefficient) of each signal when computing this distance which is equivalent to setting each to zero. In random tests, our algorithm recovers an estimate for the true signal with a reconstruction error close enough to zero to be consistent with loss of precision in floating point computations with a median error of 2.37×10^{-10} for dihedral MRA and 1.21×10^{-11} for projected MRA out of 1000 trials.

Since the algorithm is recursive and floating point calculations are inexact we can also get a sense for how errors propagates through the recursion. We split up the vector of Fourier coefficients $\hat{x} \in \mathbb{C}^2$ based on which coefficients are recovered at each level of recursion. Define the m -th cohort of Fourier coefficients to be the odd coefficients at level m of the recursion, i.e. if $y \in \mathbb{C}^{2^k}$,

$$C_0(y) := (y[0]),$$

$$C_m(y) := (y[2^{k-m}(2j-1)] : j \in [2^{m-1}]), \quad 1 < m \leq k.$$

The errors for cohorts is then $e_0 = |\hat{x}[0] - y[0]|$ and $e_m = d(C_m(\hat{x}), C_m(y))/2^{m-1}$ for $1 < m \leq k$ where we have normalized by the number of elements in the cohort. Table 5.2 gives a median error for each cohort over 1000 samples. While these median errors are small and near machine precision, note that they grow for each cohort of odd Fourier coefficients recovered. We also remark that cohort one is always the Nyquist component so we exclude it from analysis in the pMRA case.

These results can be verified by the Python scripts in the [following repository](#).³ We also furnish unit tests verifying various mathematical calculations done in this paper. See the repository for a detailed description of the contents.

³Repository is found at <https://github.com/lemniscate8/MRA-variant-orbit-recovery> for those reading in print.

Cohort Error	dMRA	pMRA
e_0	2.11×10^{-15}	1.11×10^{-15}
e_1	1.56×10^{-14}	N/A
e_2	2.39×10^{-14}	2.24×10^{-14}
e_3	1.15×10^{-12}	1.88×10^{-14}
e_4	1.76×10^{-12}	8.98×10^{-14}
e_5	3.27×10^{-12}	1.53×10^{-13}
e_6	6.07×10^{-12}	3.45×10^{-13}

TABLE 5.2. Median reconstruction errors for each cohort of odd Fourier coefficients over 1000 random samples.

5.8. Future Directions

Removing the power-of-two assumption. Our methodology exploits the specific structure of both dihedral and projected MRA, relying on a particular recursive structure which requires that the dimension of the signal be a power of two. It would be interesting to adapt our approach to all signal lengths. The concurrent work [5] achieves this using a different style of algorithm (for both dihedral MRA and a variation of projected MRA).

Proving the conjectures. While proof of generic recovery in some dimension d can be verified by checking the rank of certain symbolic matrices, we were not able to prove in general that these matrices always have full column-rank despite their shape approaching an 8 : 3 ratio of rows to columns. Techniques for such a proof would likely lead to improved bounds for finding rank-1 symmetric matrices planted in a subspace like those conjectured in Section 3.5.

Continuous groups and heterogeneity. It would be interesting to explore whether our methods can be applied in *continuous* multi-reference alignment, that is, orbit recovery under the action of an infinite group such as $\text{SO}(2)$ or $\text{O}(2)$, potentially with a projection step. It would additionally be valuable to explore whether our methods can handle a heterogeneous combination of different signals (as described in the introduction) since it remains a difficult impediment for frequency marching-type algorithms.

Other projection operators. With regards to projection operations, a close analysis of Proposition 5.2.5 reveals that most monomials present as entries in $T_{\text{dMRA}}^{(3)}$ also appear in $T_{\text{pMRA}}^{(3)}$. As a result, the effect of projection is mostly negligible for lower order moments in the sense that we have almost the same

information, even though the projection cuts the signal dimension in half. Determining more general conditions on a projection operation so that the sample complexity remains the same as for the “un-projected” version is an interesting direction for future work.

Robustness to noise. Finally, we have given an algorithm in the exact setting where we have abstracted away the estimation of moments from samples. It remains open to give an end-to-end analysis of our method applied to noisy samples, and to bound the sample complexity. This would require, in particular, a robust analysis of the intersection algorithm [43, Algorithm 1]; this currently only exists in a non-planted setting [16]. To illustrate that robustness can be a non-trivial issue, we point to the setting of cryo-ET: while an “exact” algorithm runs in polynomial time [7], a more detailed analysis of the noise tolerance requires *quasi-polynomial* time and sample complexity [50]. Determining whether our method will incur similar costs is an interesting question for further work.

5.9. Deferred Proofs

5.9.1. Group Action and Projection in the Fourier Basis.

PROOF OF PROPOSITION 5.2.1. First, $\mathcal{F} \in \mathbb{C}^{2d \times 2d}$ has entries $\mathcal{F}[j, k] = e^{-2\pi i j k / (2d)}$ and $\mathcal{F}^{-1}[j, k] = \frac{1}{2d} e^{2\pi i j k / (2d)}$.

The elements of the cyclic group on $2d$ elements are generated by R_1 since $R_\ell = (R_1)^\ell$ so it is sufficient to show that $\hat{R}_1 = \mathcal{F} R_1 \mathcal{F}^{-1}$:

$$\begin{aligned}
(\mathcal{F} R_1 \mathcal{F}^{-1})[j, k] &= \mathcal{F}[j, :](R_1 \mathcal{F}^{-1})[:, k] \\
&= \sum_{\ell=0}^{2d-1} e^{-2\pi i j \ell / (2d)} \cdot \frac{1}{2d} e^{2\pi i (\ell-1) k / (2d)} \\
&= e^{-2\pi i k / 2d} \frac{1}{2d} \sum_{\ell=0}^{2d-1} e^{-2\pi i (j-k) \ell / (2d)} \\
&= \begin{cases} e^{-2\pi i k / 2d} & j \equiv k \pmod{2d} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

So $\mathcal{F} R_1 \mathcal{F}^{-1} = \hat{R}_1$ as defined in (5.9) and then $\mathcal{F} R_\ell \mathcal{F}^{-1} = \mathcal{F} (R_1)^\ell \mathcal{F}^{-1} = (\mathcal{F} R_1 \mathcal{F}^{-1})^\ell = (\hat{R}_1)^\ell = \hat{R}_\ell$.

Next, notice that $\hat{J} = JR_{-1}$ since for $x \in \mathbb{K}^{2d}$ with entries $x = (x_0, x_1, \dots, x_{2d-1})$,

$$JR_{-1}x = JR_{-1} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{2d-2} \\ x_{2d-1} \end{bmatrix} = J \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{2d-1} \\ x_0 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_{2d-1} \\ \vdots \\ x_2 \\ x_1 \end{bmatrix} = \hat{J}x.$$

So

$$\begin{aligned} (\mathcal{F}JR_{-1}\mathcal{F}^{-1})[j, k] &= (\mathcal{F}\hat{J}\mathcal{F}^{-1})[j, k] \\ &= (\mathcal{F})[j, :](\hat{J}\mathcal{F}^{-1})[:, k] \\ &= \sum_{\ell=0}^{2d-1} e^{-2\pi i j \ell / (2d)} \cdot \frac{1}{2d} e^{2\pi i (-\ell) k / (2d)} \\ &= \frac{1}{2d} \sum_{\ell=0}^{2d-1} e^{-2\pi i (j+k) \ell / (2d)} \\ &= \begin{cases} 1 & j \equiv -k \pmod{2d} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

This is $\hat{J}$ as given in (5.10) so $\mathcal{F}JR_{-1}\mathcal{F}^{-1} = \hat{J}$.

Finally, we aim to show that there exists a full rank square matrix C such that

$$D\Pi C^{-1} = \hat{\Pi}.$$

Note that the matrix Π has the same rowspan as the matrix $\bar{\Pi} := \Pi^\top \Pi \in \mathbb{R}^{2d \times 2d}$. Inspecting the structure of $\bar{\Pi}$ gives the simplification $\bar{\Pi} = I + J$. Since this matrix is square we may find its conjugation by $\mathcal{F}$,

$$\mathcal{F}\bar{\Pi}\mathcal{F}^{-1} = I + \mathcal{F}J\mathcal{F}^{-1} = I + \mathcal{F}\hat{J}R_1\mathcal{F}^{-1} = I + \hat{J}\mathcal{F}R_1\mathcal{F}^{-1} = I + \hat{J}\hat{R}_1.$$

Elementwise, this gives $(\mathcal{F}\bar{\Pi}\mathcal{F}^{-1}v)[j] = v[j] + e^{-2\pi i (2d-j)/(2d)}v[2d-j]$ for $j = 0, 1, 2, \dots, 2d-1$. With a small algebraic simplification, the rows of this matrix are

$$\mathcal{F}\bar{\Pi}\mathcal{F}^{-1}[j, :] = \vec{e}_j + e^{2\pi i j / (2d)} \vec{e}_{-j},$$

with indices of the standard basis vectors taken modulo $2d$. Notably, $\mathcal{F}\bar{\Pi}\mathcal{F}^{-1}[d, :] = 0$ which is why the Nyquist component is unrecoverable, and $\mathcal{F}\bar{\Pi}\mathcal{F}^{-1}[2d-j, :] = e^{-2\pi i j / (2d)} \mathcal{F}\bar{\Pi}\mathcal{F}^{-1}[2d-j, :]$ so we define $\hat{\Pi}$ to

be the first d rows of $\mathcal{F}\bar{\Pi}\mathcal{F}^{-1}$ (those indexed 0 to $d-1$). To make explicit the change of basis we can write

$$D := \mathcal{F}[[d],:] \Pi^\top \in \mathbb{C}^{d \times d}$$

where it then follows that $D\Pi\mathcal{F}^{-1} = \hat{\Pi}$. (Note $\mathcal{F}[[d],:]$ is the first d rows of the matrix $\mathcal{F}$.) The original projection matrix Π had full row-rank and so does $\hat{\Pi}$ (for both, the rows are an orthogonal set of vectors) so C is a change of basis. $\square$

PROOF OF PROPOSITION 5.2.3. Let $\ell_1, \ell_2 \in [2d]$ be known and suppose that $\gcd(\ell_1 - \ell_2, 2d) = 1$. We will consider the pMRA model in the Fourier basis. Consider the block matrix given by

$$P_{\ell_1, \ell_2} = \begin{bmatrix} \hat{\Pi} \hat{R}_{\ell_1} \\ \hat{\Pi} \hat{R}_{\ell_2} \end{bmatrix} \in \mathbb{C}^{2d \times 2d}.$$

For a signal's Fourier transform $\hat{x}$, $P_{\ell_1, \ell_2} \hat{x}$ produces the concatenation of the projections from directions ℓ_1 and ℓ_2 . Now the kernel of P_{ℓ_1, ℓ_2} contains $\text{span}\{\vec{e}_d\}$ as previously noted but if P_{ℓ_1, ℓ_2} has rank $2d-1$ we may recover the rest of $\hat{x}$. We will demonstrate this by showing that Gaussian elimination produces $2d-1$ pivots. Consider that the rows of P_{ℓ_1, ℓ_2} have the form

$$\begin{aligned} P_{\ell_1, \ell_2}[j, :] &= e^{-2\pi i j \ell_1 / (2d)} \vec{e}_j + e^{2\pi i j (\ell_1 + 1) / (2d)} \vec{e}_{-j}, \\ P_{\ell_1, \ell_2}[j + d, :] &= e^{-2\pi i j \ell_2 / (2d)} \vec{e}_j + e^{2\pi i j (\ell_2 + 1) / (2d)} \vec{e}_{-j}, \end{aligned} \quad \text{for } j \in [d],$$

where indices on are taken modulo $2d$ as needed. For $j = 0$, $P_{\ell_1, \ell_2}[0, :] = P_{\ell_1, \ell_2}[d, :]$ so we get our missing pivot since the row operation $P_{\ell_1, \ell_2}[d, :] \leftarrow P_{\ell_1, \ell_2}[d, :] - P_{\ell_1, \ell_2}[0, :]$ produces a zero row. Next, if we apply the row operations

$$P_{\ell_1, \ell_2}[j + d, :] \leftarrow P_{\ell_1, \ell_2}[j + d, :] - e^{-2\pi i (\ell_2 - \ell_1) / (2d)} P_{\ell_1, \ell_2}[j, :]$$

for each $j = 1, 2, \dots, d-1$ then each row has at most one non-zero entry of the form

$$\begin{aligned} P_{\ell_1, \ell_2}[j + d, :] &= (e^{2\pi i j (\ell_2 + 1) / (2d)} - e^{-2\pi i (\ell_2 - \ell_1) / (2d)} e^{2\pi i j (\ell_1 + 1) / (2d)}) \vec{e}_{-j} \\ &= (e^{2\pi i j (\ell_2 + 1) / (2d)} - e^{-2\pi i (\ell_2 - 2\ell_1 - 1) / (2d)}) \vec{e}_{-j}. \end{aligned}$$

These two phases will only cancel if $-j(\ell_2 + 1) \equiv j(\ell_2 - 2\ell_1 - 1) \pmod{2d}$ for some $j = 1, 2, \dots, d-1$. Simplifying this produces the modular equation $j(\ell_1 - \ell_2) \equiv 0 \pmod{2d}$. Since $\ell_1 - \ell_2$ has no factors in common with $2d$ the modular condition is not satisfied for any row and so each has a non-zero entry in the $(d-j)$ -th column. Thus, the row-echelon form for P_{ℓ_1, ℓ_2} contains pivots in every column except for the d -th one and we may recover all coefficients except the d -th. $\square$

5.9.2. Invariant Tensor Structure.

PROOF OF PROPOSITION 5.2.5. First, for $x \in \mathbb{R}^d$ we will use the structure of the invariant tensor from the original MRA model since we may express the other tensors in terms of this one:

$$T_{\text{MRA}}^{(r)}(\hat{x}) = \frac{1}{d} \sum_{\ell=0}^{d-1} (\hat{R}_\ell \hat{x})^{\otimes r}.$$

We will use Definition 4.1.2 to expand entries in the tensor:

$$\begin{aligned} T_{\text{MRA}}^{(r)}(\hat{x})[\jmath_1, \dots, \jmath_r] &= \frac{1}{d} \sum_{d=\ell}^{d-1} w_{\jmath_1, \dots, \jmath_r} \prod_{k=1}^r (\hat{R}_\ell \hat{x})[j_k] \\ &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{d} \sum_{d=\ell}^{d-1} \prod_{k=1}^r e^{-2\pi i \ell j_k / d} x[j_k] \\ &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{d} \sum_{d=\ell}^{d-1} \exp\left(-\frac{2\pi i \ell}{d} \sum_{k=1}^r j_k\right) \prod_{k=1}^r x[j_k] \\ &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{d} \sum_{d=\ell}^{d-1} \left(\exp\left(-\frac{2\pi i}{d} \sum_{k=1}^r j_k\right) \right)^\ell \prod_{k=1}^r x[j_k] \\ &= w_{\jmath_1, \dots, \jmath_r} \left(\prod_{k=1}^r x[j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r j_k \equiv 0 \pmod{d} \right\}. \end{aligned}$$

The last line follows since the sum is a geometric series with the common ratio a complex root of unity. As a result it only sums to zero when the base is one. This only occurs under the modular condition $\sum_{k=1}^r j_k \equiv 0 \pmod{d}$.

Next, we use the fact that

$$T_{\text{dMRA}}^{(r)}(\hat{x}) = \frac{1}{2} \left(T_{\text{MRA}}^{(r)}(\hat{x}) + T_{\text{MRA}}^{(r)}(\hat{J}\hat{x}) \right),$$

so

$$T_{\text{dMRA}}^{(r)}(\hat{x})[\jmath_1, \dots, \jmath_r] = \frac{w_{\jmath_1, \dots, \jmath_r}}{2} \left(\prod_{k=1}^r x[j_k] + \prod_{k=1}^r x[-j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r j_k \equiv 0 \pmod{d} \right\}.$$

Finally, for the pMRA model with $x \in \mathbb{R}^{2d}$,

$$\begin{aligned} T_{\text{pMRA}}^{(r)}(\hat{x})[\jmath_1, \dots, \jmath_r] &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \prod_{k=1}^r (\hat{\Pi} \hat{R}_\ell \hat{x})[j_k] \\ &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \prod_{k=1}^r (e^{-2\pi i (j_k \ell) / (2d)} x[j_k] + e^{2\pi i (j_k) / (2d)} e^{2\pi i (j_k \ell) / (2d)} x[-j_r]) \\ &= w_{\jmath_1, \dots, \jmath_r} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \prod_{k=1}^r \sum_{s \in \{-1, 1\}} e^{\pi i (1-s)(j_k) / (2d)} e^{-2\pi i (s j_k \ell) / (2d)} x[s j_k]. \end{aligned}$$

We may exchange the product and sum to get a sum over vectors of signs, giving

$$\begin{aligned}
& w_{\{j_1, \dots, j_r\}} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \prod_{k=1}^r \sum_{s \in \{-1, 1\}} e^{\pi i(1-s)(j_k)/(2d)} e^{-2\pi i(s j_k \ell)/(2d)} x[s j_k] \\
&= w_{\{j_1, \dots, j_r\}} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \sum_{s \in \{-1, 1\}^r} \prod_{k=1}^r e^{\pi i(1-s_k)(j_k)/(2d)} e^{-2\pi i(s_k j_k \ell)/(2d)} x[s_k j_k] \\
&= w_{\{j_1, \dots, j_r\}} \frac{1}{2d} \sum_{d=\ell}^{2d-1} \sum_{s \in \{-1, 1\}^r} \exp\left(\frac{\pi i}{2d} \sum_{k=1}^r (1-s_k) j_k\right) \exp\left(\frac{-2\pi i \ell}{2d} \sum_{k=1}^r s_k j_k\right) \prod_{k=1}^r x[s_k j_k] \\
&= w_{\{j_1, \dots, j_r\}} \sum_{s \in \{-1, 1\}^r} \exp\left(\frac{\pi i}{2d} \sum_{k=1}^r (1-s_k) j_k\right) \frac{1}{2d} \sum_{d=\ell}^{2d-1} \left(\exp\left(\frac{-2\pi i}{2d} \sum_{k=1}^r s_k j_k\right) \right)^\ell \prod_{k=1}^r x[s_k j_k] \\
&= w_{\{j_1, \dots, j_r\}} \sum_{s \in \{-1, 1\}^r} \exp\left(\frac{\pi i}{2d} \sum_{k=1}^r j_k\right) (-1)^{\frac{1}{2d} \sum_{k=1}^r s_k j_k} \left(\prod_{k=1}^r x[s_k j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r s_k j_k \equiv 0 \pmod{d} \right\}.
\end{aligned}$$

In the sum, all terms are invariant to negation of the sign vector $s \mapsto -s$ except for the product $\prod_{k=1}^r x[s_k j_k]$ so if we reindex by a negation of the sign vector we simply get

$$w_{\{j_1, \dots, j_r\}} \sum_{s \in \{-1, 1\}^r} \exp\left(\frac{\pi i}{2d} \sum_{k=1}^r j_k\right) (-1)^{\frac{1}{2d} \sum_{k=1}^r s_k j_k} \left(\prod_{k=1}^r x[-s_k j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r -s_k j_k \equiv 0 \pmod{d} \right\}.$$

Averaging these two equivalent equations gives

$$\frac{w_{\{j_1, \dots, j_r\}}}{2} \sum_{s \in \{-1, 1\}^r} \exp\left(\frac{\pi i}{2d} \sum_{k=1}^r j_k\right) (-1)^{\frac{1}{2d} \sum_{k=1}^r s_k j_k} \left(\prod_{k=1}^r x[s_k j_k] + \prod_{k=1}^r x[-s_k j_k] \right) \mathbb{1} \left\{ \sum_{k=1}^r s_k j_k \equiv 0 \pmod{d} \right\}.$$

□

PROOF OF PROPOSITION 5.3.9. Let $x \in \mathbb{R}^{2d}$. By Proposition 5.2.5, for $j_1, j_2 \in [d]$,

(5.44)

$$\begin{aligned}
T_{\text{PMRA}}^{(2)}(\hat{x})[\{j_1, j_2\}] &= \frac{1}{2} (\hat{x}[j_1] \hat{x}[j_2] + \hat{x}[-j_1] \hat{x}[-j_2]) \mathbb{1}\{j_1 + j_2 \equiv 0 \pmod{2d}\} \\
&\quad + \frac{1}{2} e^{\pi i(j_1+j_2)/(2d)} (-1)^{(j_1-j_2)/(2d)} (\hat{x}[j_1] \hat{x}[-j_2] + \hat{x}[-j_1] \hat{x}[j_2]) \mathbb{1}\{j_1 - j_2 \equiv 0 \pmod{2d}\},
\end{aligned}$$

with indexing into $\hat{x}$ taken modulo $2d$. The conditions, $j_1 + j_2 \equiv 0 \pmod{2d}$ and $j_1 - j_2 \equiv 0 \pmod{2d}$ only hold simultaneously when $j_1 = j_2 = 0$. In this case, the entry is a multiple $\hat{x}[0]^2$ and hence one can only determine the value of $|\hat{x}[0]|$. Beyond this, there are no other conditions under which the first indicator variable is non-zero. The second indicator is satisfied along the diagonal, and since x is a real signal $\hat{x}[-j] = \overline{\hat{x}[j]}$, so for $0 < j < d$,

$$T_{\text{PMRA}}^{(2)}(\hat{x})[\{j, j\}] = e^{\pi i j/d} \hat{x}[j] \hat{x}[-j] = e^{\pi i j/d} |\hat{x}[j]|^2.$$

Taking the magnitude of these entries recovers $|x[j]|^2$ for $0 < j < d$ and this is the power spectrum of x excluding the coefficient $\hat{x}[d]$. $\square$

5.9.3. Polynomial Subspaces.

PROOF OF PROPOSITION 5.7.2. As previously noted, $\text{span}(v^{\otimes 2} : v \in \mathcal{X}_1^\vee) \cong (\mathbb{C}^d)^{\otimes 4}$ so $\mathcal{Q}$ is the correct size to be an orthogonal basis for $I(\mathcal{X}_1^\vee)_2^\perp$. First, we will confirm that $\mathcal{Q}$ is a basis by showing it is an orthogonal set of vectors. Notice that since they are standard basis vectors of $((\mathbb{C}^n)^{\otimes 2})^{\otimes 2}$, $\langle (\vec{e}_a \otimes \vec{e}_b) \otimes (\vec{e}_c \otimes \vec{e}_d), (\vec{e}_{a'} \otimes \vec{e}_{b'}) \otimes (\vec{e}_{c'} \otimes \vec{e}_{d'}) \rangle = 0$ whenever $\{\lambda a, b\}, \{\lambda c, d\} \neq \{\lambda a', b'\}, \{\lambda c', d'\}$. When multisets $\{\lambda a, b, c, d\} \neq \{\lambda a', b', c', d'\}$, each produces a different set of two element partitions so $\langle Q_{abcd}^\vee, Q_{a'b'c'd'}^\vee \rangle = 0$.

Next, we will show that for each $E_{abcd}^\vee$ from (4.12) and $Q^\vee \in \mathcal{Q}$, $Q^\vee \perp E_{abcd}^\vee$. Consider that if $\{\lambda a, b, c, d\} \neq \{\lambda a', b', c', d'\}$ then $\langle E_{abcd}^\vee, Q_{abcd}^\vee \rangle = 0$. We need only verify that $\langle E_{abcd}^\vee, Q_{abcd}^\vee \rangle = 0$ and $\langle E_{abcd}^\vee, Q_{acbd}^\vee \rangle = 0$. For the first, expanding terms then canceling weights allows us to write the inner product as a sum of inner products of standard basis vectors in $((\mathbb{C}^n)^{\otimes 2})^{\otimes 2}$,

$$\begin{aligned} \langle E_{abcd}^\vee, Q_{abcd}^\vee \rangle &= \langle \vec{e}_{\{\lambda a, b\}, \{\lambda c, d\}}, \vec{e}_{\{\lambda a, b\}, \{\lambda c, d\}} \rangle \\ &\quad - \langle \vec{e}_{\{\lambda a, d\}, \{\lambda c, b\}}, \vec{e}_{\{\lambda a, b\}, \{\lambda c, d\}} \rangle \\ &\quad + \langle \vec{e}_{\{\lambda a, b\}, \{\lambda c, d\}}, \vec{e}_{\{\lambda a, c\}, \{\lambda b, d\}} \rangle \\ &\quad - \langle \vec{e}_{\{\lambda a, d\}, \{\lambda c, b\}}, \vec{e}_{\{\lambda a, c\}, \{\lambda b, d\}} \rangle \\ &\quad + \langle \vec{e}_{\{\lambda a, b\}, \{\lambda c, d\}}, \vec{e}_{\{\lambda a, d\}, \{\lambda b, c\}} \rangle \\ &\quad - \langle \vec{e}_{\{\lambda a, d\}, \{\lambda c, b\}}, \vec{e}_{\{\lambda a, d\}, \{\lambda b, c\}} \rangle \\ &= 1 - 1 = 0. \end{aligned}$$

The computation for $\langle E_{abcd}^\vee, Q_{acbd}^\vee \rangle = 0$ is effectively identical in structure. So $\mathcal{Q}$ is an orthogonal basis for $I(\mathcal{X}_1^\vee)_2^\perp$. $\square$

Bibliography

- [1] Asaf Abas, Tamir Bendory, and Nir Sharon, *The generalized method of moments for multi-reference alignment*, IEEE Transactions on Signal Processing **70** (2022), 1377–1388 (en). arXiv:2103.02215 [eess].
- [2] Emmanuel Abbe, Tamir Bendory, William Leeb, Joao M. Pereira, Nir Sharon, and Amit Singer, *Multireference Alignment Is Easier With an Aperiodic Translation Distribution*, IEEE Transactions on Information Theory **65** (June 2019), no. 6, 3565–3584.
- [3] Emmanuel Abbe, Joao M. Pereira, and Amit Singer, *Sample complexity of the boolean multireference alignment problem*, 2017 IEEE International Symposium on Information Theory (ISIT), June 2017, pp. 1316–1320.
- [4] Cecilia Aguerrebere, Mauricio Delbracio, Alberto Bartsaghi, and Guillermo Sapiro, *Fundamental Limits in Multi-Image Alignment*, IEEE Transactions on Signal Processing **64** (November 2016), no. 21, 5707–5722 (en).
- [5] Amnon Balanov, Tamir Bendory, and Dan Edidin, *Group-invariant moments under tomographic projections*, arXiv, 2026. Version Number: 1.
- [6] Amnon Balanov, Josh Katz, Tamir Bendory, and Dan Edidin, *Projected multi-reference alignment*, arXiv, 2026. Version Number: 1.
- [7] Afonso S. Bandeira, Ben Blum-Smith, Joe Kileel, Amelia Perry, Jonathan Niles-Weed, and Alexander S. Wein, *Estimation under group actions: recovering orbits from invariants*, Applied and Computational Harmonic Analysis **66** (September 2023), 236–319 (en). arXiv:1712.10163 [math].
- [8] Afonso S. Bandeira, Moses Charikar, Amit Singer, and Andy Zhu, *Multireference alignment using semidefinite programming*, Proceedings of the 5th conference on Innovations in theoretical computer science, January 2014, pp. 459–470 (en).
- [9] Afonso S. Bandeira, Yutong Chen, Roy R. Lederman, and Amit Singer, *Non-unique games over compact groups and orientation estimation in cryo-EM*, Inverse Problems **36** (June 2020), no. 6, 064002.
- [10] Afonso S. Bandeira, Jonathan Niles-Weed, and Philippe Rigollet, *Optimal rates of estimation for multi-reference alignment*, Mathematical Statistics and Learning **2** (February 2020), no. 1, 25–75.
- [11] Tamir Bendory, Alberto Bartsaghi, and Amit Singer, *Single-Particle Cryo-Electron Microscopy: Mathematical Theory, Computational Challenges, and Opportunities*, IEEE Signal Processing Magazine **37** (March 2020), no. 2, 58–76.
- [12] Tamir Bendory, Nicolas Boumal, Chao Ma, Zhizhen Zhao, and Amit Singer, *Bispectrum Inversion with Application to Multireference Alignment*, IEEE Transactions on Signal Processing **66** (February 2018), no. 4, 1037–1050 (en). arXiv:1705.00641 [cs].
- [13] Tamir Bendory, Dan Edidin, William Leeb, and Nir Sharon, *Dihedral Multi-Reference Alignment*, IEEE Transactions on Information Theory **68** (May 2022), no. 5, 3489–3499.
- [14] Tamir Bendory, Ariel Jaffe, William Leeb, Nir Sharon, and Amit Singer, *Super-resolution multi-reference alignment*, Information and Inference: A Journal of the IMA **11** (June 2022), no. 2, 533–555 (en).

- [15] Alessandra Bernardi, *Ideals of varieties parameterized by certain symmetric tensors*, Journal of Pure and Applied Algebra **212** (June 2008), no. 6, 1542–1559 (en).
- [16] Aditya Bhaskara, Eric Evert, Vaidehi Srinivas, and Aravindan Vijayaraghavan, *New Tools for Smoothed Analysis: Least Singular Value Bounds for Random Matrices with Dependent Entries*, Proceedings of the 56th Annual ACM Symposium on Theory of Computing, June 2024, pp. 375–386 (en).
- [17] Ake Björck and Gene H. Golub, *Numerical Methods for Computing Angles Between Linear Subspaces*, Mathematics of Computation **27** (July 1973), no. 123, 579.
- [18] Nicolas Boumal, Tamir Bendory, Roy R. Lederman, and Amit Singer, *Heterogeneous multireference alignment: A single pass approach*, 2018 52nd Annual Conference on Information Sciences and Systems (CISS), March 2018, pp. 1–6.
- [19] M. Boussé, N. Vervliet, I. Domanov, O. Debals, and L. De Lathauwer, *Linear systems with a canonical polyadic decomposition constrained solution: Algorithms and applications: Linear Systems with a CPD Constrained Solution*, Numerical Linear Algebra with Applications **25** (December 2018), no. 6, e2190 (en).
- [20] Paul Breiding, Fulvio Gesmundo, Mateusz Michałek, and Nick Vannieuwenhoven, *Algebraic compressed sensing*, Applied and Computational Harmonic Analysis **65** (July 2023), 374–406 (en).
- [21] Jonathan F. Buss, Gudmund S. Frandsen, and Jeffrey O. Shallit, *The computational complexity of some problems of linear algebra*, Stacs 97, 1997, pp. 451–462.
- [22] Haolin Chen and Luis Rademacher, *Overcomplete Order-3 Tensor Decomposition, Blind Deconvolution, and Gaussian Mixture Models*, SIAM Journal on Mathematics of Data Science **4** (March 2022), no. 1, 336–361 (en).
- [23] Luca Chiantini, Giorgio Ottaviani, and Nick Vannieuwenhoven, *An Algorithm For Generic and Low-Rank Specific Identifiability of Complex Tensors*, SIAM Journal on Matrix Analysis and Applications **35** (January 2014), no. 4, 1265–1287 (en).
- [24] Jeshu Dastidar, Tait Weicht, and Alexander S. Wein, *Improving the Threshold for Finding Rank-1 Matrices in a Subspace*, arXiv, 2025 (en). arXiv:2504.17947 [cs].
- [25] Lieven De Lathauwer, *A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization*, SIAM journal on Matrix Analysis and Applications **28** (2006), no. 3, 642–666.
- [26] Lieven De Lathauwer, Josphine Castaing, and Jean-Francois Cardoso, *Fourth-Order Cumulant-Based Blind Identification of Underdetermined Mixtures*, IEEE Transactions on Signal Processing **55** (June 2007), no. 6, 2965–2973 (en).
- [27] Vin de Silva and Lek-Heng Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, arXiv, 2008 (en). arXiv:math/0607647.
- [28] Harm Derksen, Nathaniel Johnston, and Benjamin Lovitz, *X-arability of mixed quantum states*, arXiv preprint arXiv:2409.18948 (2024).
- [29] Nachum Dershowitz and Zohar Manna, *Proving termination with multiset orderings*, Communications of the ACM **22** (August 1979), no. 8, 465–476 (en).
- [30] Ignat Domanov and Lieven De Lathauwer, *Canonical polyadic decomposition of third-order tensors: Relaxed uniqueness conditions and algebraic algorithm*, Linear Algebra and its Applications **513** (2017), 342–375.
- [31] Dan Edidin and Josh Katz, *The Reflection-Invariant Bispectrum: Signal Recovery in the Dihedral Model*, Journal of Fourier Analysis and Applications **32** (February 2026), no. 1, 1 (en).

- [32] Zhou Fan, Roy R. Lederman, Yi Sun, Tianhao Wang, and Sheng Xu, *Maximum likelihood for high-noise group orbit estimation and single-particle cryo-EM*, The Annals of Statistics **52** (February 2024), no. 1.
- [33] Zhou Fan, Yi Sun, Tianhao Wang, and Yihong Wu, *Likelihood landscape and maximum likelihood estimation for the discrete orbit recovery model*, Communications on Pure and Applied Mathematics **76** (June 2023), no. 6, 1208–1302 (en).
- [34] Hubert Flenner, Liam O’Carroll, and Wolfgang Vogel, *Joins and intersections*, Springer, 1999.
- [35] Joachim Frank, *Generalized single-particle cryo-EM – a historical perspective*, Microscopy **65** (February 2016), no. 1, 3–8 (en).
- [36] Georgios B. Giannakis, *Signal reconstruction from multiple correlations: frequency- and time-domain approaches*, Journal of the Optical Society of America A **6** (May 1989), no. 5, 682 (en).
- [37] Huy Tai Ha, *Box-shaped matrices and the defining ideal of certain blowup surfaces*, arXiv, 2002 (en). arXiv:math/0210251.
- [38] Joe Harris, *Algebraic geometry: a first course*, Vol. 133, Springer Science & Business Media, 1992.
- [39] Richard A Harshman, *Foundations of the PARAFAC procedure: models and conditions for an “explanatory” multimodal factor analysis*, UCLA Working Papers in Phonetics (1970), 1–84 (en).
- [40] Frank L. Hitchcock, *The Expression of a Tensor or a Polyadic as a Sum of Products*, Journal of Mathematics and Physics **6** (1927), no. 1-4, 164–189.
- [41] Samuel B. Hopkins, Tselil Schramm, and Jonathan Shi, *A Robust Spectral Algorithm for Overcomplete Tensor Decomposition*, Proceedings of the Thirty-Second Conference on Learning Theory, June 2019, pp. 1683–1722.
- [42] Johan Håstad, *Tensor rank is NP-complete*, Journal of Algorithms **11** (1990) (1989), 644–654.
- [43] Nathaniel Johnston, Benjamin Lovitz, and Aravindan Vijayaraghavan, *Computing linear sections of varieties: quantum entanglement, tensor decompositions and beyond*, 2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS), November 2023, pp. 1316–1336.
- [44] P. Koiran, *An efficient uniqueness theorem for overcomplete tensor decomposition*, Proceedings of the 2025 annual acm-siam symposium on discrete algorithms (soda), 2025, pp. 1909–1932.
- [45] Pascal Koiran, *An efficient uniqueness theorem for overcomplete tensor decomposition*, arXiv, 2024. Version Number: 2.
- [46] Tamara G. Kolda and Brett W. Bader, *Tensor Decompositions and Applications*, SIAM Review **51** (August 2009), no. 3, 455–500 (en).
- [47] Praveesh K. Kothari, Ankur Moitra, and Alexander S. Wein, *Overcomplete Tensor Decomposition via Koszul-Young Flattenings*, arXiv, 2024. Version Number: 2.
- [48] Joseph B. Kruskal, *Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra and its Applications **18** (1977), no. 2, 95–138 (en).
- [49] S. E. Leurgans, R. T. Ross, and R. B. Abel, *A Decomposition for Three-Way Arrays*, SIAM Journal on Matrix Analysis and Applications **14** (October 1993), no. 4, 1064–1083 (en).
- [50] Allen Liu and Ankur Moitra, *Algorithms from Invariants: Smoothed Analysis of Orbit Recovery over $SO(3)$* , arXiv, 2022 (en). arXiv:2106.02680 [cs, stat].
- [51] Benjamin Lovitz and Nathaniel Johnston, *Entangled subspaces and generic local state discrimination with pre-shared entanglement*, Quantum **6** (2022), 760.
- [52] Tengyu Ma, Jonathan Shi, and David Steurer, *Polynomial-time Tensor Decompositions with Sum-of-Squares*, arXiv, 2016 (en). arXiv:1610.01980 [cs].

- [53] T. Matsuoka and T.J. Ulrych, *Phase estimation using the bispectrum*, Proceedings of the IEEE **72** (1984), no. 10, 1403–1411.
- [54] Ankur Moitra, *Algorithmic Aspects of Machine Learning*, 1st ed., Cambridge University Press, 2018 (en).
- [55] Ankur Moitra and Alexander S. Wein, *Spectral Methods from Tensor Networks*, SIAM Journal on Computing **52** (April 2023), no. 2, STOC19–354–STOC19–384 (en).
- [56] Amelia Perry, Jonathan Weed, Afonso S. Bandeira, Philippe Rigollet, and Amit Singer, *The Sample Complexity of Multireference Alignment*, SIAM Journal on Mathematics of Data Science **1** (January 2019), no. 3, 497–517 (en).
- [57] Elena-Mădălina Persu, *Tensors, sparse problems and conditional hardness*, PhD Thesis, 2018 (en).
- [58] Stephan Rabanser, Oleksandr Shchur, and Stephan Günnemann, *Introduction to tensor decompositions and their applications in machine learning*, arXiv preprint arXiv:1711.10781 (2017).
- [59] Nir Sharon, Joe Kileel, Yuehaw Khoo, Boris Landa, and Amit Singer, *Method of moments for 3D single particle ab initio modeling with non-uniform distribution of viewing angles*, Inverse Problems **36** (April 2020), no. 4, 044003.
- [60] F.J. Sigworth, *A Maximum-Likelihood Approach to Single-Particle Image Refinement*, Journal of Structural Biology **122** (1998), no. 3, 328–339 (en).
- [61] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Applied and Computational Harmonic Analysis **30** (January 2011), no. 1, 20–36 (en).
- [62] A. Singer and Y. Shkolnisky, *Three-Dimensional Structure Determination from Common Lines in Cryo-EM by Eigenvectors and Semidefinite Programming*, SIAM Journal on Imaging Sciences **4** (January 2011), no. 2, 543–572 (en).
- [63] Amit Singer, *Mathematics for cryo-electron microscopy*, Proceedings of the International Congress of Mathematicians (ICM 2018), May 2019, pp. 3995–4014 (en).
- [64] Simon Telen and Nick Vannieuwenhoven, *A normal form algorithm for tensor rank decomposition*, ACM Transactions on Mathematical Software **48** (2022), no. 4, 1–35.
- [65] Bogdan Toader, Fred J. Sigworth, and Roy R. Lederman, *Methods for Cryo-EM Single Particle Reconstruction of Macromolecules Having Continuous Heterogeneity*, Journal of Molecular Biology **435** (May 2023), no. 9, 168020 (en).
- [66] Aravindan Vijayaraghavan, *Efficient tensor decomposition*, arXiv preprint arXiv:2007.15589 (2020).
- [67] Tait Weicht and Alexander S. Wein, *Beyond Frequency Marching: Orbit Recovery in Dihedral and Projected Multireference Alignment*, arXiv, 2026. Version Number: 1.
- [68] Sheng Xu, Anderson Ye Zhang, and Amit Singer, *Misspecified Maximum Likelihood Estimation for Non-Uniform Group Orbit Recovery*, arXiv, 2025 (en). arXiv:2509.22945 [math].