Free Descriptions of Convex Sets

Eitan Levin[†] and Venkat Chandrasekaran[†,‡] [*]

[†] Department of Computing and Mathematical Sciences
[‡] Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125

July 10, 2023

Abstract

Convex sets arising in a variety of applications are well-defined for every relevant dimension. Examples include the simplex and the spectraplex that correspond, respectively, to probability distributions and to quantum states; combinatorial polytopes and their associated relaxations such as the cut polytope and the elliptope in integer programming; and unit balls of commonly-employed regularizers such as the ℓ_p and Schatten norms in inverse problems. Moreover, these sets are often specified using conic descriptions that can be obviously instantiated in any dimension. We develop a systematic framework to study such *free descriptions* of convex sets. We show that free descriptions arise from a recently-identified phenomenon in algebraic topology called *representation stability*, which relates invariants across dimensions in a sequence of group representations. Our framework yields a procedure to obtain parametric families of freely-described convex sets whose structure is adapted to a given application; illustrations are provided via examples that arise in the literature as well as new families that are derived using our procedure. We demonstrate the utility of our framework in two contexts. First, we develop an algorithm for a free analog of the convex regression problem, where a convex function is fit to input-output training data; in our setting, the inputs may be of different dimensions and we seek a convex function that is well-defined for inputs of any dimension (including those that are not in the training set). Second, we prove that many sequences of symmetric conic programs can be solved in constant time, which unifies and strengthens several disparate results in the literature. Our work extensively uses ideas and results from representation stability, and it can be seen as a new point of contact between representation stability and convex geometry via conic descriptions.

Keywords: cone programming, convex optimization, free spectrahedra, graphons, representation theory

Contents

---

[*]Email: {eitanl, venkatc}@caltech.edu

1   Introduction

Convex sets are central objects in numerous areas of the mathematical sciences such as of interest in optimization, statistical inference, control, inverse problems, statistical inference, and information theory. In domains such as inverse problems or integer programming, convex sets are of interest for algorithmic reasons as one seeks optima of various functionals over these sets. In quantum information or in control, the geometric properties of convex sets are of interest as they are used to describe fundamental quantities such as collections of quantum states and channels or of controllers. Convex sets appearing in these areas are often defined in every relevant dimension. For example, unit balls of regularizers in inverse problems (e.g., the ℓ_p or Schatten norms), convex relaxations for intractable graph problems (e.g., elliptope approximations of cut polytopes), controllers (e.g., in dynamical systems in which the system block-diagram structure is agnostic to the dimensions of the states), and families of quantum states arise in every (meaningful) dimension. Thus, convex sets in these and many other areas should be viewed as sequences indexed by dimension. In this paper, we develop a framework to study and derive finitely-parametrized families of sequences of convex sets. Our approach unifies specific cases that have been investigated previously in the literature in the context of control systems and quantum information, and it yields a systematic method to obtain new families in many other contexts.

Our first motivation for this effort is computational and stems from the growing interest in obtaining solution methods for various problems in a data-driven manner. In this paradigm, one identifies procedures given input-output data – inputs representing problem instances and outputs specifying solutions – rather than handcrafting one based on insights about the structure of the problem family. This framework has been fruitfully applied to domains including integer programming, inverse problems, and numerical solvers for PDEs [1, 2, 3, 4, 5, 6]. A fundamental limitation in much of the literature on these topics is that the solution methods learned from data are only applicable in dimensions that are manifest in the provided training data, and extension to inputs of different sizes is handled on a case-by-case basis. In contrast, we wish to learn *algorithms* that should be defined for inputs of any relevant size and constitute a sequence of solution maps, one for each input size. To facilitate numerical search over spaces of algorithms, we seek principled approaches to deriving *finitely-parametrized* such sequences. Our work addresses this challenge for algorithms specified as linear optimization over convex sets, i.e., convex programs.

Our second motivation is mathematical, whereby we wish to facilitate structural understanding of convex sets that can be instantiated in any dimension. Although such sequences of convex sets are ubiquitous, the existence and interplay between the sets in different dimensions is rarely explicitly discussed or exploited in the literature. One body of work in which relations between convex sets in different dimensions play a prominent role is the theory of matrix convexity and free convex algebraic geometry. This theory studies matrix convex sets, which are sequences of convex sets closed under matrix convex combinations that relate the sets in different dimensions [7, 8, 9, 10]. In the present paper we consider sequences of convex sets with more general relations between dimensions, which provides a framework in which to investigate the geometry of sequences of convex sets beyond those that arise in matrix convexity.

A central feature of convex sets is the manner in which they are described. In particular, a canonical way to describe convex sets is via *conic descriptions* in which a set is expressed as an affine section of a convex cone. Conic descriptions have played a central role in modern convex optimization [11]. Indeed, we often classify convex sets based on their conic descriptions—polytopes are affine sections of nonnegative orthants and spectrahedra are affine sections of positive-semidefinite (PSD) cones. Formally, if V, W, U are vector spaces and $K \subseteq U$ is a cone, then a convex subset $C \subseteq V$ can be described using linear maps $A \colon V \to U$, $B \colon W \to U$ and a vector $u \in U$ by

$$C = \{x \in V : \exists y \in W \text{ s.t. } Ax + By + u \in K\}. \tag{Conic}$$

We call the spaces W and U the *description spaces* associated to the conic description. If the cone K is a nonnegative orthant (resp., positive semidefinite cone), then linear optimization over C is a linear (resp., semidefinite) program. The size of the cone K captures the complexity of optimization over C, blending geometry and computation.

An interesting property of conic descriptions of convex sets that commonly arise in practice is that they are often "free" – that is, these descriptions can be obviously instantiated in any desired dimension, thereby yielding a sequence of convex sets.

Example 1.1 (Free descriptions). *The following are simple examples of freely-described sequences of convex sets arising in (quantum) information, graph theory, and the theory of matrix convexity.*

(a) *The simplex in n dimensions is $\Delta^{n-1} = \{x \in \mathbb{R}^n : x \geq 0, \ \mathbb{1}_n^\top x = 1\}$, which is the set of probability distributions over n items. Here $x \geq 0$ denotes an entrywise nonnegative vector x, and $\mathbb{1}_n \in \mathbb{R}^n$ is the vector of all-1's.*

(b) *The spectraplex, or the set of density matrices, of size n is $\mathcal{D}^{n-1} = \{X \in \mathbb{S}^n : X \succeq 0, \ \text{Tr}(X) = 1\}$. It is the set of density matrices describing mixed states in quantum mechanics [12, §2.4]. Here $X \succeq 0$ denotes a symmetric PSD matrix X.*

(c) *The ℓ_2 ball in $\mathbb{R}^n$ is given by $B_{\ell_2}^n = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} 1 & x^\top \\ x & I_n \end{bmatrix} \succeq 0 \right\}.$*

(d) *The elliptope of size n is $\{X \in \mathbb{S}^n : X \succeq 0, \text{diag}(X) = \mathbb{1}_n\}$. It arises in a standard relaxation of the max-cut problem [13].*

(e) *Free spectrahedra of size n are families of sets parametrized by $L_0, \ldots, L_d \in \mathbb{S}^k$ for $k \in \mathbb{N}$ of the form $\left\{ (X_1, \ldots, X_d) \in (\mathbb{S}^n)^d : L_0 \otimes I_n + \sum_{i=1}^d L_i \otimes X_i \succeq 0 \right\}$. They arise in the theory of matrix convexity and free convex algebraic geometry [14, 7].*

(f) *The following is a family of sets parametrized by $\alpha \in \mathbb{R}^7$ that is defined for any n:*

$$\left\{ \alpha_1 \mathrm{diag}(X) + \alpha_2 X \mathbb{1}_n : X \succeq 0, \ (\alpha_3 \mathbb{1}_n^\top X \mathbb{1}_n + \alpha_4 \mathrm{Tr}(X)) \mathbb{1}_n + \alpha_5 \mathrm{diag}(X) + \alpha_6 X \mathbb{1}_n = \alpha_7 \mathbb{1}_n \right\}. \tag{1}$$

All the descriptions in Example 1.1 can be instantiated for every $n \in \mathbb{N}$, yielding infinite sequences of convex sets in different dimensions. Further, Examples 1.1(e)-(f) are finitely-parametrized sequences of sets. The goal of this paper is to develop a systematic framework to study and derive such finitely-parametrized families of sequences of convex sets.

To formalize the notion of free descriptions, we begin by making three observations. First, we note that the sequences of sets in Example 1.1 are described as slices of standard sequences of cones, such as nonnegative orthants and PSD cones; expressing sets in terms of such standard sequences has the benefit that optimization over these sets can be performed using standard off-the-shelf software. Going forward, we shall assume that our cones as well as the vector spaces containing them come from such standard sequences, and in particular, can be instantiated in any dimension. Second, and more importantly, the affine sections of these cones are expressed in terms of vectors and linear maps such as $\mathbb{1}_n \in \mathbb{R}^n$, $I_n \in \mathbb{S}^n$, $\mathrm{diag} \colon \mathbb{S}^n \to \mathbb{R}^n$, which are "free", meaning that they are well-defined in any dimension. Third, finitely-parametrized free descriptions can be obtained by taking linear combinations of any finite collection of such vectors and linear maps, and by viewing the coefficients in the combination as parameters as in Example 1.1(f). Putting these observations together, we seek finite-dimensional spaces of free vectors and linear maps.

The free vectors and linear maps in Example 1.1 are sequences of invariants under sequences of groups, and these invariants are related in a particular way across dimensions. For example, the all-1's vector $\mathbb{1}_n$ of length n is invariant under the group of permutations on n letters acting by permuting coordinates. Further, the all-1's vectors of different lengths are related to each other: extracting the first n entries of $\mathbb{1}_{n+1}$ yields $\mathbb{1}_n$. Similarly, the $n \times n$ identity matrix I_n is invariant under the orthogonal group of size n acting by conjugation, and extracting the top left $n \times n$ submatrix of I_{n+1} yields I_n. Thus, to give a formal definition for free vectors and linear maps, we consider sequences of groups acting on sequences of vector spaces, and we require the spaces in the sequence to be related to each other – specifically, we embed lower-dimensional spaces into higher-dimensional ones and project higher-dimensional spaces onto lower-dimensional ones. Such sequences of group representations are called *consistent sequences*, and they were introduced in [15] in order to study the recently-identified phenomenon of *representation stability*. Representation stability implies that the projections define isomorphisms between the spaces of invariants in the sequence. Consequently, the dimensions of invariants in such a sequence stabilize, and the projections furnish the desired relations between the invariants in different dimensions that constitute free vectors and linear maps. In this way, representation stability yields the desired finite-dimensional spaces of free invariants, and hence finitely-parametrized free descriptions of convex sets.

The combination of symmetry and relations across dimensions that constitute consistent sequences and yield free descriptions arises commonly in applications. Indeed, there are natural embeddings in many applications that relate problem instances of different dimensions. For example, probability distributions over small alphabets can be viewed as distributions over large alphabets by putting zero weight on the additional letters, and small graphs can be viewed as large ones by appending isolated vertices. Further, many application domains exhibit an underlying symmetry, and this structure is inherited by the convex sets appearing in those domains. Polytopes arising in graph theory are invariant under permutations [16], and sets arising in quantum information are invariant under the orthogonal group [17, 18, 19]. In inverse problems, data distributions are often group-invariant, and therefore the unit ball of regularizers used to promote the structure in these distributions are group invariant as well [20, 21].

We use our framework to pursue the two motivating goals we outlined in the beginning of this section. First, we describe a fully algorithmic procedure to obtain parametric families of freely-described convex sets that are adapted to the structure of a given application. Our method only requires a consistent sequence containing the desired convex sets, as well as sequences of description spaces and cones as in (Conic). The former is dictated by the application at hand while the latter is chosen based on the expressivity of the

4

family sought by the user; larger description spaces yield more expressive families but they lead to more computationally expensive optimization over the resulting sets. We apply this procedure to free convex regression: given a consistent sequence $\{V_n\}_{n \in \mathbb{N}}$ and training data $(x_i, y_i) \in V_{n_i} \oplus \mathbb{R}$, we seek a freely-described sequence of convex functions $\{f_n : V_n \to \mathbb{R}\}$ satisfying $f_{n_i}(x_i) \approx y_i$ for all i. An open-source implementation is available at https://github.com/eitangl/anyDimCvxSets. Our procedure learns a function that can be defined on inputs of any size including those that are not part of the training data, thus taking a taking a step towards learning algorithms from data. Second, we obtain structural results for optimization over freely-described convex sets. Specifically, it has been observed in the literature that certain sequences of symmetric semidefinite and relative entropy programs can be solved in constant time [22, 23, 24, 25]. However, the proofs of this fact proceed on a case-by-case basis by showing that the irreducible decompositions of the representations that arise for each program stabilize. We unify the treatment of this phenomenon as well as strengthen existing results in the literature using representation stability, which was developed to explain the ubiquity of the stabilization of irreducible decompositions.

1.1   Our contributions

In Section 2, we formally define free descriptions of sequences of convex sets (Definition 2.14). To do so, we consider consistent sequences (Definition 2.1), which as mentioned above are sequences of groups and representations related by embeddings and projections. Representation stability gives canonical isomorphisms between spaces of invariants in such sequences of representations (Proposition 2.7), allowing us to formally define freely-described elements (Definition 2.12) constituting free descriptions of convex sets, and to derive parametric families of such descriptions. In several problem domains, it is desirable to not only relate the descriptions of the sets across dimensions but also the sets themselves (see Example 2.28). We therefore introduce compatibility conditions relating the convex sets across dimensions (Definition 2.20), and characterize free descriptions of sets satisfying those conditions (Theorem 2.23). This allows us to parametrize only sequences of sets that are compatible across dimensions, and hence extend better to higher dimensions, and fit these sequences to data.

In Section 3, we illustrate our framework with a large number of examples of sequences of convex sets and their descriptions. Some of our examples arise in the literature in application domains including (quantum) information and graph theory, while several others constitute new families that are derived using our procedure and are adapted to each application.

To use our parametric families for learning tasks such as free convex regression described above, we fit a convex set in the largest dimension in which data is available and extend it to a freely-described sequence that can be instantiated in any other dimension. To that end, in Section 4 we study extensions of a convex set in a fixed dimension to a freely-described and compatible sequence, and characterize the existence of such an extension (Theorem 4.15). The *presentation degree* of a consistent sequence plays a fundamental role in our characterization and in the resulting algorithm, and we expose the relevant background theory from the representation stability literature in a manner that is motivated by our computational goals.

In Section 5 we develop an algorithm to computationally parametrize and search over sequences of convex sets (Algorithm 1). Our algorithm only requires a choice of description spaces and cones as inputs from the user. It proceeds by combining recent ideas on computing bases for spaces of invariants with an alternating minimization approach that searches over coefficients in these bases to fit a freely-described convex set to data. We apply our algorithm in Section 5.3 to two stylized free convex regression problems. Specifically, we consider two functions defined on inputs of any size that are not semidefinite representable, namely, the ℓ_π-norm and a (nonnegative and positively homogenous) variant of the quantum entropy; we identify semidefinite approximations entirely from evaluations of these functions on low-dimensional inputs, and are then able to evaluate our approximations on inputs of any size. This task is particularly interesting as neither of the functions can be evaluated exactly using semidefinite programming.

In Section 6, we use our framework to obtain structural results for sequences of invariant conic programs. Specifically, we study sequences of invariant semidefinite and relative entropy programs and give conditions under which they can be solved in constant time. To understand this phenomenon, we formally define constant-sized descriptions for sequences of cones (Definition 6.1) and prove their existence for sequences of symmetric PSD and relative entropy cones, as well as their variants such as sums-of-squares (SOS) and Sums-of-AM/GM-Exponentials (SAGE) cones. Our study of these cones yields Theorems 6.12 and 6.15,

which unify and generalize results in the literature mentioned above.

It is natural to consider limits of freely-described sequences of convex sets, and in particular, to obtain conic descriptions of such limits. In Section 7, we show that free descriptions certifying our compatibility conditions across dimensions often have such limits. Further, these limiting descriptions yield dense subsets of limits of the convex sets in the sequence (Theorem 7.5). Many of our definitions can also be understood in terms of such limits. These results raise the intriguing possibility of studying the complexity of infinite-dimensional conic programs by viewing them as limits of finite-dimensional ones.

We conclude in Section 8 with several directions for future work pertaining to areas such as computational algebra, convex analysis, and statistics. Sections 2.1, 4.1, 6.1 and 7.1 review the needed definitions and results from the representation stability literature, which we present and motivate with a view towards free descriptions.

1.2 Related work

We briefly survey several related areas.

Lifts of convex sets: There is a large literature as well as a systematic framework studying conic descriptions of a fixed convex set, see [26] for a review. In particular, this framework can be applied to study *equivariant lifts* of group-invariant convex sets, which are descriptions of the form (Conic) consisting of group-invariant cones, vectors, and linear maps, see [27] and [26, §4.3]. These are precisely the type of descriptions we consider for each of the convex sets in our sequences. Further, even though this framework is formulated in terms of a fixed convex set, results in this area implicitly concern descriptions of a *sequence* $\{C_n\}$ of convex sets and the complexity of these descriptions in n [28, 27, 29, 30]. Many of the descriptions proposed in this literature can in fact be instantiated in any dimension, and they turn out to be free descriptions according to our definition as we show in Section 3. However, the descriptions of each convex set in the sequence are studied independently of each other in this literature. To our knowledge, however, free descriptions which apply to the entire sequence as well as relations between the sets in the sequence have not been studied systematically. Such study is necessary to obtain parametric families of freely-described sequences of sets, which is significant in the context of of our motivating goal of learning algorithms from data.

Free spectrahedra, noncommutative convex algebraic geometry: A broad research program pursued in several areas involves the study of "matrix" or noncommutative analogues of classical "scalar" or commutative objects. Examples include random matrix theory and free probability [31], studying matrix-valued random variables and their limits as opposed to scalar-valued ones; and noncommutative algebraic geometry [14], studying polynomials in noncommuting variables and their evaluations on matrices as opposed to standard polynomials in commuting variables that are evaluated on scalars [32, 33, 34, 35]. Applying this program to convex sets yields matrix-convex sets and free spectrahedra, the latter being sequences of sets of the form of Example 1.1(e). We refer the reader to [14, 7] for surveys and [36, 37] for some applications. In analogy to the setting of the present paper, results in this area explicitly pertain to sequences of sets which are "freely-described", in the sense that their description can be instantiated in any dimension. For example, free spectrahedra are sequences of sets described by a single linear matrix inequality, and free algebraic varieties are defined by the same noncommutative polynomials instantiated on matrices of any size. Another point of contact with our work is the consideration of relations between the sets in the sequence across dimensions, such as matrix-convex combinations which have been formalized and studied in this literature. Our notion of free descriptions is more general than the ones in this literature however, and it allows us to derive more flexible families of freely-described sets which are adapted to different applications. Further, the relations between sets in different dimensions we consider in this paper are less restrictive than matrix convexity, and again allows us to consider more general families of sets than free spectrahedra (see Section 3.3).

Representation stability: Representation stability arose out of the observation that the decomposition into irreducibles of many sequences of representations stabilize. This phenomenon has been formalized in [15] using consistent sequences, and subsequently studied in [38, 39, 40] from a categorical perspective and in [41, 42, 43] from a limits-based perspective. We relate the categorical and limits-based formalisms to ours in appendix A and Section 7.1, respectively, and refer the reader to [44, 45, 46] for introductions to this area.

Representation stability has been used to study sequences of polyhedral cones and their infinite-dimensional limits [47], as well as sequences of algebraic varieties, their defining equations, and their infinite-dimensional

Symmetric group	$\mathsf{S}_n = \{g \in \mathbb{R}^{n \times n} : g \text{ is a permutation matrix.}\}$
Signed symmetric group	$\mathsf{B}_n = \{g \in \mathbb{R}^{n \times n} : g \text{ is a signed permutation matrix}\}$
Even-signed symmetric group	$\mathsf{D}_n = \{g \in B_n : g \text{ flips evenly-many signed}\}$
Cyclic group	$\mathsf{C}_n = \{s^r : s \in \mathbb{R}^{n \times n} \text{ sends } se_i = e_{i+1 \mod n} \text{ and } r \in [n]\}.$
Orthogonal group	$\mathsf{O}_n = \{g \in \mathbb{R}^{n \times n} : g^\top g = I_n\}$
Space of linear maps	$\mathcal{L}(V, U) = \{A : V \to U \text{ linear}\}; \qquad \mathcal{L}(V) = \mathcal{L}(V, V).$
Direct sum	$V \oplus U = V \times U = \{(v, u) : v \in V, u \in U\}.$
Direct powers	$V^k = V^{\oplus k} = \underbrace{V \oplus \cdots \oplus V}_{k \text{ times}}.$
Tensor product	$V \otimes U = \mathrm{span}\{v \otimes u : v \in V, u \in U\} \cong \mathcal{L}(V, U).$
Tensor power	$V^{\otimes k} = \underbrace{V \otimes \cdots \otimes V}_{k \text{ times}}.$
Symmetric algebra	$\begin{aligned} \mathrm{Sym}^k(V) \ &= \mathrm{span}\{v_1 \cdots v_k : v_i \in V\} \\ &= \{\text{polynomials of degree } = k \text{ on } V\} \\ &= \{\text{symmetric tensors of order } k \text{ over } V\} \\ \mathrm{Sym}^{\leq k}(V) &= \bigoplus_{i=0}^{k} \mathrm{Sym}^i(V). \end{aligned}$
Alternating algebra	$\begin{aligned} \bigwedge^k V &= \mathrm{span}\{v_1 \wedge \cdots \wedge v_k : v_i \in V\} \\ &= \{\text{skew-symmetric tensors of order } k \text{ over } V\}. \end{aligned}$
Symmetric matrices	$\mathbb{S}^n = \{X \in \mathbb{R}^{n \times n} : X^\top = X\} = \mathrm{Sym}^2(\mathbb{R}^n).$
Skew-symmetric matrices	$\mathrm{Skew}(n) = \{X \in \mathbb{R}^{n \times n} : X^\top = -X\} = \bigwedge^2 \mathbb{R}^n.$
Spaces of invariants	$\begin{aligned} V^G &= \{v \in V : g \cdot v = v \text{ for all } g \in G\}, \\ \mathcal{L}(V, U)^G &= \{A \in \mathcal{L}(V, U) : gAg^{-1} = A \text{ for all } g \in G\}. \end{aligned}$

Table 1: Commonly-used groups and vector spaces.

limits [48, 49, 50]. An important distinction between these works and ours is our application of representation stability to *descriptions* of convex sets rather than to their extreme points or rays as in [47]. Thus, we are able to study non-polyhedral sets such as spectrahedra and sets defined by relative entropy programs. Similarly, our study of infinite-dimensional limits in Section 7 focuses on limiting descriptions and not just on limits of the sets themselves.

1.3 Notation and basics

We assume familiarity with the basics of representation theory and convex analysis, and refer the reader to [51, 52] and [53], respectively, for references. We list several standard groups and constructions involving vector spaces in Table 1.

Basics:

- We denote $[n] = \{1, \ldots, n\}$. A *partition* λ of an integer n is a nonincreasing sequence $\lambda = (\lambda_1 \geq \ldots \geq \lambda_k > 0)$ satisfying $|\lambda| := \sum_i \lambda_i = n$.

- For $i \in [n]$, we denote by $e_i \in \mathbb{R}^n$ the ith standard basis vector with a 1 in the ith entry and zero everywhere else, and denote $e_i^{(n)}$ when we wish to emphasize its size.

- If $x \in \mathbb{R}^n$ we denote by $\mathrm{diag}(x) \in \mathbb{S}^n$ the diagonal matrix with x on the diagonal. If $X \in \mathbb{R}^{n \times n}$ we denote by $\mathrm{diag}(X) \in \mathbb{R}^n$ the vector of its diagonal elements.

- All vector spaces in this paper are finite-dimensional real vector spaces equipped with an inner product $\langle \cdot, \cdot \rangle$ unless stated otherwise. We emphasize that some of the inner products we use are nonstandard, hence the transpose of a matrix and the adjoint of the linear operator it represents may differ.

- Given a subspace $W \subseteq V$, denote by $\mathcal{P}_W \colon V \to W$ the orthogonal projection onto W.

- We denote by $\mathbb{R}_+^n$ the cone of entrywise nonnegative vectors in $\mathbb{R}^n$, and by $\mathbb{S}_+^n$ the cone of PSD $n \times n$ matrices. If V is a vector space, denote $\mathrm{Sym}_+^2(V) \cong \mathbb{S}_+^{\dim V}$ the cone of PSD self-adjoint linear maps in $\mathcal{L}(V)$.

Representation theory:

- A (linear) action of a group G on a finite-dimensional vector space V is given by a group homomorphism $\rho \colon G \to \mathrm{GL}(V)$.

  If G is a Lie group, we denote by $\mathrm{D}\rho \colon \mathfrak{g} \to \mathrm{End}(V)$ the induced representation of its Lie algebra $\mathfrak{g}$. Usually ρ is clear from context and we omit it, writing $g \cdot v = \rho(g)v$ for $g \in G$ and $v \in V$ instead.

- All the groups we consider are compact and all group actions are orthogonal, meaning $\langle g \cdot x, g \cdot y \rangle = \langle v, v' \rangle$ for all $x, y \in V$.

- We denote the group ring of G by $\mathbb{R}[G] = \mathrm{span}\{e_g\}_{g \in G}$, where e_g is a basis element indexed by the group element g. This is a ring with multiplication defined by $e_g \cdot e_h = e_{gh}$ for $g, h \in G$ and extended by linearity. Note that a representation of G is the same as a module over the ring $\mathbb{R}[G]$.

- If $H \subseteq G$ is a subgroup, and V is a representation of H, the *induced representation* of G from V is $\mathrm{Ind}_H^G(V) = \mathbb{R}[G] \otimes_{\mathbb{R}[H]} V$. We have $\dim \mathrm{Ind}_H^G(V) = |G/H| \dim V$, and we apply this notion only when H has finite index in G. If $g_1 = \mathrm{id}, g_2, \ldots, g_k$ are coset representatives for G/H, we have

$$\mathrm{Ind}_H^G(V) = \bigoplus_{i=1}^k g_i V, \tag{Ind}$$

together with the following action of G: If $g \in G$ is (uniquely) written as $gg_i = g_j h$ for some $i, j \in [k]$ and $h \in H$, then $g \cdot g_i v = g_j(h \cdot v)$ for any $v \in V$. This construction is independent of the choice of coset representatives.

As vector spaces, we have an isomorphism $\mathrm{Ind}_H^G(V) \cong V^{|G/H|}$. Hence an H-invariant inner product $\langle \cdot, \cdot \rangle$ on V induces a G-invariant inner product on $V^{|G/H|}$ by setting $\langle g_i v, g_j u \rangle = \delta_{i,j} \langle v, u \rangle$ for $v, u \in V$ and $i, j \in [|G/H|]$. Here $\delta_{i,j} = 1$ if $i = j$ and zero otherwise.

We have an isomorphism $(\mathrm{Ind}_H^G V)^H \cong V^H$ sending $v \in V^H$ to $\sum_i g_i v \in (\mathrm{Ind}_H^G V)^H$ and $\widetilde{v} \in (\mathrm{Ind}_H^G V)^H$ to $\mathcal{P}_V \widetilde{v} \in V^H$.

If $H \subseteq H'$ and $G \subseteq G'$ such that $H' \cap G = H$, then we have inclusion $G/H \hookrightarrow G'/H'$ sending $gH \mapsto gH'$, inducing an inclusion $\mathrm{Ind}_H^G V \hookrightarrow \mathrm{Ind}_{H'}^{G'} V$ between induced representations by completing a set of coset representatives for G/H to representatives for G'/H''. Here V is assumed to be an H'-representation.

If V, U are H-representations and $A \colon V \to U$ is an equivariant linear map, we can extend it to a map $\mathrm{Ind}(A) \colon \mathrm{Ind}_H^G V \to \mathrm{Ind}_H^G U$ by defined $\mathrm{Ind}(A)(g_i v) = g_i(Av)$ where g_i is one of the above coset representatives and $v \in V$.

If V is a G-representation and $W \subseteq V$ is an H-subrepresentation, there is a G-equivariant linear map $\mathrm{Ind}_H^G W \to V$ sending $g \otimes w \mapsto g \cdot w$ whose image is precisely $\mathbb{R}[G]W = \mathrm{span}\{g \cdot w\}_{g \in G, w \in W}$.

- The action of G on V induces an action on $V^{\otimes k}$ and $\mathrm{Sym}^k(V)$ by setting $g \cdot v_1 \otimes \cdots \otimes v_k = (gv_1) \otimes \cdots \otimes (gv_k)$ and $g \cdot (v_1 \cdots v_k) = (gv_1) \cdots (gv_k)$ and extending by linearity. If V and U are both representations of G, we have an action of G on $V \otimes U$ by $g \cdot (v \otimes u) = (gv) \otimes (gu)$ and extending by linearity, and on $\mathcal{L}(V, U)$ by $g \cdot A = gAg^{-1}$, making the representations $V \otimes U$ and $\mathcal{L}(V, U)$ isomorphic. Linear maps invariant under this group action are also called *equivariant* or *intertwining*, since they are precisely the linear maps commuting with the group elements.

Convex analysis:

- The *polar* of a convex set $C \subseteq V$ is the convex set
$$C^\circ = \{x \in V : \langle x, y \rangle \le 1 \text{ for all } y \in C\}. \tag{polar}$$

- There are several correspondences between convex sets and functions in convex analysis, which we use to obtain freely-described and compatible functions from sets in Section 2.4. Given a convex subset $C \subseteq V$, its *gauge function* (also called Minkwoski functional) is the convex, nonnegative, and positively-homogeneous function $\gamma_C \colon V \to \mathbb{R}$ defined by
$$\gamma_C(x) = \inf_{t \ge 0} t \quad \text{s.t. } x \in tC. \tag{gauge}$$

Its *support function* is the convex and positively-homogeneous function $h_C \colon V \to \mathbb{R}$ defined by
$$h_C(x) = \sup_{x' \in C} \langle x', x \rangle. \tag{supp}$$

Given a convex subset $E \subseteq V \oplus \mathbb{R}$, we can define a convex function $f_E \colon V \to \mathbb{R}$ by
$$f_E(x) = \inf_{t \in \mathbb{R}} t \quad \text{s.t. } (x, t) \in E. \tag{epi}$$

2   Freely-described and compatible sets

In this section, we formally define freely-described convex sets by generalizing Example 1.1, and explain how to derive parametric families of such sets. To correctly extend a set learned in a fixed dimension to higher dimensions, we also consider compatibility conditions relating sets in different dimensions and derive conditions on free descriptions that certify these conditions. To do so, we use relations between dimensions in the form of embeddings and projections, and their basic properties introduced and studied in the representation stability literature. We begin by reviewing the necessary definitions and results from there.

2.1   Background: Consistent sequences, generation degree, representation stability

We begin by defining the relations between dimensions that we shall use in this paper. A variant of the following definition was made in the seminal paper [15] introducing representation stability.[1]

Definition 2.1 (Consistent sequences). *Fix a family of compact*[2] *groups $\mathscr{G} = \{G_n\}_{n \in \mathbb{N}}$ such that $G_n \subseteq G_{n+1}$. A consistent sequence of $\mathscr{G}$-representations is a sequence $\mathscr{V} = \{(V_n, \varphi_n)\}_{n \in \mathbb{N}}$ satisfying the following properties:*

(a) *V_n is an orthogonal G_n-representation;*

(b) *$\varphi_n \colon V_n \hookrightarrow V_{n+1}$ is a linear G_n-equivariant isometry.*

Unless we want to emphasize the embeddings φ_n, we shall identify V_n with its image inside V_{n+1}. We then write $\mathscr{V} = \{V_n\}$ and take φ_n to be inclusions $V_n \subseteq V_{n+1}$,[3] *see Figure 1.*

As φ_n is an isometry, we have $\varphi_n^* \circ \varphi_n = \mathrm{id}_{V_n}$ so that $\varphi_n^* = \mathcal{P}_{V_n}$ is the orthogonal projector onto V_n in V_{n+1}.

---

[1]The maps φ_n are not assumed to be isometries in the original paper [15] introducing this definition. We make this assumption here to simplify our presentation, and because it holds in all our examples.

[2]Our theory can be extended to reductive groups.

[3]Formally, we obtain such inclusions inside the direct limit of the sequence.

$$\ldots \quad \subseteq \quad G_n \quad \subseteq \quad G_{n+1} \quad \subseteq \quad \ldots$$

$$\circlearrowright \qquad \circlearrowright$$

$$\ldots \quad \overset{\twoheadleftarrow}{\hookrightarrow} \quad V_n \quad \overset{\twoheadleftarrow}{\hookrightarrow} \quad V_{n+1} \quad \overset{\twoheadleftarrow}{\hookrightarrow} \quad \ldots$$

Figure 1: Relations between dimensions.

Example 2.2. *The following are simple examples of consistent sequences arising in optimization. We analyze more sophisticated examples in Section 3.*

(a) *Let $V_n = \mathbb{R}^n$ with the standard inner product, and let $\varphi_n(x) = [x^\top, 0]^\top$ padding a vector with a zero. This is a consistent sequence for many standard sequences of groups, including $G_n = \mathsf{O}_n, \mathsf{B}_n, \mathsf{D}_n, \mathsf{S}_n$ acting by their standard $n \times n$ matrix representations. Here G_n is embedded in G_{n+1} by sending $g \in G_n$ represented as an $n \times n$ matrix to $g \oplus 1$.*

(b) *Let $V_n = \mathbb{S}^n$ with the Frobenius inner product, and let φ_n pad a symmetric matrix by a zero row and column. All the sequences of groups in part (a) act on $\mathbb{S}^n$ by conjugation $g \cdot X = gXg^{-1}$ where $X \in \mathbb{S}^n$ and $g \in G_n$ is represented as a $n \times n$ matrix.*

(c) *Let $V_n = \mathbb{R}^{2^n}$ with the normalized inner product $\langle x, y \rangle = 2^{-n} x^\top y$, and $\varphi_n(x) = x \otimes \mathbb{1}_2$. Here we can take $G_n = \mathsf{C}_{2^n}$ in addition to the standard families of groups in part (a), and embed G_n into G_{n+1} by sending a $2^n \times 2^n$ matrix g to $g \otimes I_2$.*

Remark 2.3. *Our results generalize to the following, more complicated, sequences of group representations. If $\mathcal{N}$ is a strict poset, an $\mathcal{N}$-indexed consistent sequence of $\{G_n\}_{n \in \mathcal{N}}$-representations is a sequence $\{(V_n, \varphi_{N,n})\}_{n < N \in \mathcal{N}}$ of G_n-representations together with embeddings $\varphi_{N,n} \colon V_n \hookrightarrow V_N$ for each $n < N$ such that $\varphi_{N,n}$ is G_n-equivariant, and $\varphi_{M,N} \circ \varphi_{N,n} = \varphi_{M,n}$ whenever $n < N < M$. All our results apply in this setting as well after replacing all occurrences of $n+1$ by $N > n$. This allows us to consider sequences such as $\mathbb{R}^{n_1 \times n_2} \hookrightarrow \mathbb{R}^{m_1 \times m_2}$ where the embedding is padding by zeros, and $\mathbb{R}^n \hookrightarrow \mathbb{R}^{nk}$ for all $n, k \in \mathbb{N}$ by sending $x \mapsto x \otimes \mathbb{1}_k$.*

To handle more complex description spaces, and thereby obtain more expressive families of sets, it is useful to form more complex consistent sequences from simpler ones. The following remark addresses these aims.

Remark 2.4 (Sums, tensors, polynomials)**.** *Fix a family of group $\mathcal{G} = \{G_n\}_{n \in \mathbb{N}}$ such that $G_n \subseteq G_{n+1}$. Suppose $\mathcal{V} = \{(V_n, \varphi_n)\}$ and $\mathcal{U} = \{(U_n, \psi_n)\}$ are consistent sequences of $\mathcal{G}$-representations. Then the following are also consistent sequences of $\mathcal{G}$-representations.*

(Sums) *The direct sum of $\mathcal{V}$ and $\mathcal{U}$ is $\mathcal{V} \oplus \mathcal{U} = \{(V_n \oplus U_n, \varphi_n \oplus \psi_n)\}$.*

If W is a fixed vector space, viewed as a trivial G_n-representation for all n, denote $\mathcal{V} \oplus W = \{(V_n \oplus W, \varphi_n \oplus \mathrm{id}_W)\}$.

(Tensors) *The tensor product of $\mathcal{V}$ and $\mathcal{U}$ is $\mathcal{V} \otimes \mathcal{U} = \{(V_n \otimes U_n, \varphi_n \otimes \psi_n)\}$.*

This is also the sequence of spaces of linear maps $\mathcal{L}(V_n, U_n) \cong V_n \otimes U_n$, where we embed $A_n \colon V_n \to U_n$ to $(\varphi_n \otimes \psi_n) A_n = \psi_n A_n \varphi_n^* \colon V_{n+1} \to U_{n+1}$.

The order-k tensors over $\mathcal{V}$ is $\mathcal{V}^{\otimes k} = \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{k \ times}$.

If W is a fixed vector space, viewed as a trivial G_n-representation for all n, denote $\mathcal{V} \otimes W = \{(V_n \otimes W, \varphi_n \otimes \mathrm{id}_W)\}$.

(Polynomials) *The degree-k polynomials over $\mathcal{V}$ is $\mathrm{Sym}^k \mathcal{V} = \{(\mathrm{Sym}^k V_n, \varphi_n^{\otimes k})\}$, which is also the sequence of order-k symmetric tensors over $\mathcal{V}$. Here we view $\mathrm{Sym}^k V_n \subseteq V^{\otimes k}$ and restrict $\varphi_n^{\otimes k}$ to that subspace.*

The sequence of polynomials of degree at most k is denoted $\mathrm{Sym}^{\leq k} \mathcal{V} = \bigoplus_{j=1}^k \mathrm{Sym}^j \mathcal{V}$.

Similarly, we can form the sequence of kth exterior powers $\bigwedge^k \mathcal{V}$.

10

(Moments) *The sequence of* moment matrices of order k over $\mathscr{V}$ is $\mathrm{Sym}^2(\mathrm{Sym}^{\leq k}\mathscr{V})$. *Its elements can be viewed as symmetric matrices whose rows and columns are indexed by monomials of degree at most k in basis elements for $\mathscr{V}$.*

The group actions on these spaces are given in Section 1.3. In particular, suppose $V_n = \mathbb{R}^n$ with embedding by zero-padding and the action of one of the standard families of groups from Example 2.2(a). Then $V_n^{\otimes k}$ consists of $n \times \cdots \times n$-sized tensors with embeddings by zero padding and $\mathrm{Sym}^k V_n$ consists of homogeneous polynomials of degree k in n variables. In this case, we also have $\mathrm{Sym}^2 V_n = \mathbb{S}^n$ and $\bigwedge^2 V_n = \mathrm{Skew}(n)$. The space $\mathrm{Sym}^2(\mathrm{Sym}^{\leq k} V_n)$ can be viewed as symmetric matrices whose rows and columns are indexed by monomials of degree up to k in n variables. These sequences arises in tensor, sums-of-squares and moment optimization problems [54, 55].

To relate invariants and equivariants across dimensions, we need canonical isomorphisms between the spaces of invariants of a consistent sequence. Proposition 2.7 below shows that the projections $\mathcal{P}_{V_n}$ are such isomorphisms, using the following parameter introduced in [38] to control the complexity of a consistent sequence.

Definition 2.5 (Generation degree)**.** *A consistent sequence $\mathscr{V} = \{V_n\}$ of $\{G_n\}$-representations is generated in degree d if $\mathbb{R}[G_n]V_d = V_n$ for all $n \geq d$. The smallest d for which this holds is called the* generation degree *of the sequence. A subset $\mathcal{S} \subseteq V_d$ is called a* set of generators *for $\mathscr{V}$ if $\mathbb{R}[G_n]\mathcal{S} = V_n$ for all $n \geq d$. A sequence is* finitely-generated *if it is generated in degree d for some $d < \infty$.*

Note that $\mathbb{R}[G_n]V_d = \mathrm{span}\{gV_d\}_{g \in G_n}$, so that $\mathscr{V}$ is generated in degree d if the span of the G_n-orbit of V_d, when embedded in V_n, is all of V_n, for any $n \geq d$. Note also that if $\mathscr{V}$ is generated in degree d then V_d is a set of generators for $\mathscr{V}$.

Example 2.6. *We return to the examples from Example 2.2. Note that the symmetric group S_n is contained in all the groups in Example 2.2(a).*

(a) *The sequence $V_n = \mathbb{R}^n$ with embeddings by zero-padding from Example 2.2(a) is generated in degree 1 for all the standard families of groups listed there. Indeed, any of the canonical basis vectors e_i are obtained from the first one e_1 via the action of S_n.*

(b) *The sequence $V_n = \mathbb{S}^n$ with embeddings by zero-padding from Example 2.2(b) is generated in degree 1 if $G_n = \mathsf{O}_n$, and in degree 2 if $G_n = \mathsf{B}_n, \mathsf{D}_n, \mathsf{S}_n$. Indeed, the first claim follows by the spectral theorem, and the second follows by applying transpositions (Section 1.3) to $e_1 e_1^\top$ and $e_1 e_2^\top + e_2 e_1^\top$.*

(c) *The sequence $V_n = \mathbb{R}^{2^n}$ with embedddings $\varphi_n(x) = x \otimes \mathbb{1}_2$ from Example 2.2(c) is generated in degree 2 if $G_n = \mathsf{O}_{2^n}, \mathsf{B}_{2^n}, \mathsf{D}_{2^n}, \mathsf{S}_{2^n}$ but is not finitely-generated if $G_n = \mathsf{C}_{2^n}$. The first claim follows from the identity $e_1^{(2^n)} = \frac{1}{2}[\mathrm{id} - (1,3) + (2,3)] \cdot e_1^{(2^{n-1})} \otimes \mathbb{1}_2$ where $(i,j) \in \mathsf{S}_{2^n}$ interchanges elements i and j. The last claim follows because the vector $\mathbb{1}_{2^n} \otimes [1,-1]^\top = [1,-1,1,-1,\dots]^\top \in V_{n+1}$ is orthogonal to $\mathbb{R}[G_{n+1}]\varphi_n(V_n)$, for any $n \in \mathbb{N}$.*

As the next proposition shows, finite generation gives us canonical isomorphisms between invariants in different dimensions. We use these isomorphisms to formally define free invariants generalizing the examples in Definition 2.12.

Proposition 2.7. *Suppose $\mathscr{V} = \{(V_n, \varphi_n)\}$ is a consistent sequence of $\{G_n\}$-representations generated in degree d. Then the projections $\varphi_n^* = \mathcal{P}_{V_n} : V_{n+1}^{G_{n+1}} \to V_n^{G_n}$ are injective for all $n \geq d$, and are therefore isomorphisms for all large enough n.*

Proof. First, the map φ_n^* is G_n-equivariant because G_n acts orthogonally, and $G_n \subseteq G_{n+1}$, hence it maps G_{n+1}-invariants in V_{n+1} to G_n-invariants in V_n. Second, suppose $\varphi_n^*(v) = 0$ for some $v \in V_{n+1}^{G_{n+1}}$. For any $u \in V_{n+1}$, write $u = \sum_i g_i \varphi_n(u_i)$ where $u_i \in V_n$ and $g_i \in G_{n+1}$. Because v is G_{n+1}-invariant, we have $\langle v, u \rangle = \langle \varphi_n^*(v), \sum_i u_i \rangle = 0$. As $u \in V_{n+1}$ was arbitrary, we conclude that $v = 0$.

The injectivity of φ_n^* shows that $\dim V_n^{G_n} \geq \dim V_{n+1}^{G_{n+1}}$ for all $n \geq d$, hence the sequence of dimensions $\dim V_n^{G_n}$ eventually stabilizes at which point φ_n^* becomes an isomorphism. $\square$

Proposition 2.7 is stated in the representation stability literature in terms of the adjoints of the projections, viewed as maps between *co*invariants, see [38, §3] for example. More precisely, the projections become isomorphisms starting from the *presentation* degree of the consistent sequence, see Section 4.1 below. Results in representation stability imply that many commonly-encountered consistent sequences are finitely-generated, and yield bounds on the associated generation and presentation degrees, see Theorem 4.11 below for example.

We also define a notion of maps between consistent sequences, which enables us to define embeddings, quotients, and isomorphisms identifying consistent sequences with others.

Definition 2.8 (Morphisms of sequences). *If $\mathscr{V} = \{(V_n, \varphi_n)\}$ and $\mathscr{U} = \{(U_n, \psi_n)\}$ are two consistent sequences of $\{G_n\}$-representations, then a* morphism of consistent sequences $\mathscr{A}: \mathscr{V} \to \mathscr{U}$ *is a collection of linear maps $\mathscr{A} = \{A_n: V_n \to U_n\}$ such that the following hold for each n:*

(a) A_n is G_n-equivariant;

(b) $A_{n+1}\varphi_n = \psi_n A_n$.

If φ_n and ψ_n are inclusions, condition (b) above becomes $A_{n+1}|_{V_n} = A_n$. Note also that the collection of morphisms between two sequences $\mathscr{V} \to \mathscr{U}$ forms a linear space, because if $\{A_n\}$ and $\{A'_n\}$ are morphisms then so is $\{\alpha A_n + \beta A'_n\}$ for any $\alpha, \beta \in \mathbb{R}$.

Example 2.9. *Let $\mathscr{V} = \mathscr{U} = \{\mathbb{S}^n\}$ with the action of $G_n = \mathsf{S}_n$ be the consistent sequence from Example 2.2(b). The morphisms $\mathscr{V} \to \mathscr{U}$ form a 3-dimensional space spanned by $\{A_n^{(1)} X = X\}$, $\{A_n^{(2)} X = \mathrm{diag}(X \mathbb{1}_n)\}_n$, and $\{A_n^{(3)} X = \mathrm{diag}(\mathrm{diag}(X))\}$.*

As the following proposition shows, morphisms yield additional examples of consistent sequences by taking the sequences of their images and kernels. These will play a prominent role in our study of extendability of a fixed convex set to a sequence in Section 4.

Proposition 2.10. *If $\mathscr{V} = \{(V_n, \varphi_n)\}$ and $\mathscr{U} = \{(U_n, \psi_n)\}$ are consistent sequences of $\{G_n\}$-representations and $\mathscr{A} = \{A_n: V_n \to U_n\}$ is a morphism of sequences, then the following are also consistent sequences of $\{G_n\}$-representations.*

(Image) $\mathrm{Im}\mathscr{A} = \{(A_n(V_n), \psi_n)\}$;

(Kernel) $\ker\mathscr{A} = \{(\ker A_n, \varphi_n)\}$.

Proof. As A_n is G_n-equivariant, both its image and its kernel are G_n-representations. The embeddings ψ_n map $A_n(V_n)$ to $A_{n+1}(V_{n+1})$ because for any $x \in V_n$ we have $\psi_n(A_n x) = A_{n+1}(\varphi_n x) \in A_{n+1}(V_{n+1})$. Similarly, the embeddings φ_n map $\ker A_n$ to $\ker A_{n+1}$ because if $A_n x = 0$ then $A_{n+1}(\varphi_n x) = \psi_n(A_{n+1} x) = 0$. The maps φ_n and ψ_n remain G_n-equivariant isometries when restricted to the G_n-subrepresentations $\ker A_n$ and $A_n(V_n)$. $\square$

Furthermore, if $\mathscr{A} = \{A_n: V_n \to U_n\}: \mathscr{V} \to \mathscr{U}$ is a surjection of sequences, meaning it is a morphism of sequences such that each A_n is surjective, and if $\mathscr{V}$ is generated in degree d, then $\mathscr{U}$ is generated in degree d as well as can be seen by considering the images of a generating set for $\mathscr{V}$.

Example 2.11. *In Example 2.9, the image of both $\{A_n^{(2)}\}$ and $\{A_n^{(3)}\}$ is the consistent sequence of diagonal matrices with embeddings by zero-padding, which is isomorphic to Example 2.2(a) with $G_n = \mathsf{S}_n$ via the (iso)morphism of sequences $\{A_n X = \mathrm{diag}(X)\}$. The kernel of $\{A_n^{(2)}\}$ is the consistent sequence of symmetric matrices with zero row and column sums, and the kernel of $\{A_n^{(3)}\}$ is the sequence of zero-diagonal symmetric matrices, both with embeddings by zero-padding.*

Morphisms of sequences have appeared in the representation stability literature as the natural notion of maps between sequences, see [38, Def. 2.1.1] and [39, §3.2]. They also arise when imposing compatibility on convex sets in different dimensions (Theorem 2.23), and have a natural interpretation in terms of limits (Section 7.1).

2.2 Free conic descriptions

In this section, we define and study free descriptions of convex sets, formalizing our observations from Section 1. We begin by defining freely-described vectors and linear maps which generalize the constituents of the free descriptions in Example 1.1.

Definition 2.12 (Freely-described elements). *A freely-described element in a consistent sequence $\mathscr{V} = \{(V_n, \varphi_n)\}$ of $\{G_n\}$-representations is a sequence $\{v_n \in V_n^{G_n}\}$ of invariants satisfying $\varphi_n^*(v_{n+1}) = v_n$ for all n.*

Recall that $\varphi_n^* = \mathcal{P}_{V_n}$, so a freely-described element is a sequence of invariants projecting onto each other. Note that the set of freely-described elements of a given sequence $\mathscr{V}$ is naturally a linear space,[4] because if $\{v_n\}$ and $\{v_n'\}$ are freely-described elements then so is $\{\alpha v_n + \beta v_n'\}$ for any $\alpha, \beta \in \mathbb{R}$. This space can be identified with $V_n^{G_n}$ whenever n is large enough by Proposition 2.7.

Example 2.13. *The following are simple examples of spaces of freely-described elements.*

(a) *Let $\mathscr{V} = \{\mathbb{R}^n\}$ be the consistent sequence from Example 2.2(a) with $G_n = \mathsf{S}_n$. The freely-described elements in $\mathscr{V}$ are $\{\alpha \mathbb{1}_n\}_n$ for $\alpha \in \mathbb{R}$, forming a one-dimensional space.*

(b) *Let $\mathscr{W} = \{\mathbb{S}^n\}$ be the consistent sequence from Example 2.2(b) with $G_n = \mathsf{S}_n$. Then the freely-described elements in $\mathscr{V}$ are $\{\alpha I_n + \beta \mathbb{1}_n \mathbb{1}_n^\top\}_n$ for $\alpha, \beta \in \mathbb{R}$, forming a two-dimensional space.*

(c) *Let $\mathscr{U} = \mathscr{W}^{\otimes 2} = \{\mathcal{L}(\mathbb{S}^n)\}$ where $\mathscr{W}$ is as in (b). Then the freely-described elements in $\mathscr{U}$ are*

$$\left\{ \begin{aligned} A_n X &= \alpha_1 (\mathbb{1}_n^\top X \mathbb{1}_n) \mathbb{1}_n \mathbb{1}_n^\top + \alpha_2 (\mathbb{1}_n^\top X \mathbb{1}_n) I_n + \alpha_3 \mathrm{Tr}(X) \mathbb{1}_n \mathbb{1}_n^\top + \alpha_4 \mathrm{Tr}(X) I_n \\ &\quad + \alpha_5 \left(X \mathbb{1}_n \mathbb{1}_n^\top + \mathbb{1}_n \mathbb{1}_n^\top X \right) + \alpha_6 \left(\mathrm{diag}(X) \mathbb{1}_n^\top + \mathbb{1}_n \mathrm{diag}(X)^\top \right) + \alpha_7 X + \alpha_8 \mathrm{diag}(X \mathbb{1}_n) \\ &\quad + \alpha_9 \mathrm{diag}(\mathrm{diag}(X)) \right\}, \qquad \text{for some } \alpha \in \mathbb{R}^9, \end{aligned} \right.$$

forming a 9-dimensional space.

We now arrive at one of the central definitions of this paper, namely, that of free descriptions of convex sets. These are sequences of conic descriptions consisting of freely-described vectors and linear maps.

Definition 2.14 (Free conic descriptions). *Let $\mathscr{V} = \{V_n\}$, $\mathscr{W} = \{W_n\}$, $\mathscr{U} = \{U_n\}$ be consistent sequences of $\{G_n\}$-representations, and $\mathscr{K} = \{K_n \subseteq U_n\}$ a sequence of convex cones. A sequence of conic descriptions of the form*

$$C_n = \{x \in V_n : \exists y \in W_n \text{ s.t. } A_n x + B_n y + u_n \in K_n\}, \tag{ConicSeq}$$

is called free if $\{A_n\}, \{B_n\}$, and $\{u_n\}$ are freely-described elements of the consistent sequences $\mathscr{V} \otimes \mathscr{U}$, $\mathscr{W} \otimes \mathscr{U}$, and $\mathscr{U}$, respectively.

All the descriptions in Example 1.1 become free when the relevant vector spaces are endowed with natural consistent sequence structure. We consider a few simple examples now, and defer more sophisticated examples to Section 3.

Example 2.15. *Throughout this example, let $\mathscr{V} = \{V_n = \mathbb{R}^n\}$ with embedding by zero-padding and the standard actions of the sequences of groups in Example 2.2(a).*

(Simplex) *Let $G_n = \mathsf{S}_n$ and consider the sequence of simplices in Example 1.1(a). This sequence is given by (ConicSeq) with*

$$\mathscr{U} = \mathscr{V} \oplus \mathbb{R} = \{U_n = \mathbb{R}^{n+1}\}, \quad \mathscr{K} = \{K_n = \mathbb{R}_+^n \oplus \{0\}\}, \quad \mathscr{W} = \{W_n = 0\},$$

---

[4]Formally, the space of freely-described elements is the inverse limit $\varprojlim_n V_n^{G_n}$.

and

$$A_n = \begin{bmatrix} I_n \\ \mathbb{1}_n^\top \end{bmatrix}, \quad B_n = 0, \quad u_n = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Note that $\{A_n\}, \{B_n\}, \{w_n\}$ are all freely-described elements, hence the standard descriptions of simplices is free.

(ℓ_2-ball) *Let $G_n = \mathsf{O}_n$ and consider the ℓ_2-unit balls in Example 1.1(c). These are given by (ConicSeq) with*

$$\mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 1} \mathscr{V}) = \{U_n = \mathbb{S}^{n+1}\}, \quad \mathscr{K} = \{K_n = \mathbb{S}^{n+1}_+\}, \quad \mathscr{W} = \{W_n = 0\},$$

and

$$A_n x = \begin{bmatrix} 0 & x^\top \\ x & 0 \end{bmatrix}, \quad B_n = 0, \quad u_n = I_{n+1}.$$

Again, $\{A_n\}, \{B_n\}, \{u_n\}$ are freely-described elements, hence these descriptions are free.

(ℓ_1-ball) *Let $G_n = \mathsf{B}_n$ and consider the following extended formulation for the unit-ℓ_1 norm ball*

$$C_n = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\} = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n \ \text{s.t.} \ y \geq 0, \ -y \leq x \leq y, \ \langle \mathbb{1}_n, y \rangle \leq 1\}.$$

This is a free description of the form (ConicSeq) with

$$\mathscr{U} = \{U_n = (\mathbb{R}^n)^2 \oplus \mathbb{R}\}, \quad \mathscr{K} = \{K_n = (\mathbb{R}^n_+)^2 \oplus \mathbb{R}_+\}, \quad \mathscr{W} = \mathscr{V}.$$

Here G_n acts on U_n as follows. The subgroup S_n permutes each of the three copies of $\mathbb{R}^n$ in U_n separately, while a sign matrix $\mathrm{diag}(s)$ with $s \in \{\pm 1\}^n$ acts trivially on the first copy and interchanges the coordinates of the second and third copies whose indices lie in $\{i \in [n] : s_i = -1\}$. Define

$$A_n x = (x, -x, 0), \quad B_n y = (y, y, -\mathbb{1}^\top y), \quad u_n = (0, 0, 1).$$

Then $\{A_n\}, \{B_n\}, \{u_n\}$ are freely-described and C_n is given by (ConicSeq).

Remark 2.16 (Other description formats). *While we phrase all our results in terms of descriptions of the form (ConicSeq), other formats for descriptions can be rewritten in that form to fit within our framework. For example, if there are consistent sequences $\mathscr{V} = \{V_n\}, \widetilde{\mathscr{U}} = \{\widetilde{U}_n\}, \widetilde{\mathscr{W}} = \{\widetilde{W}_n\}$ of $\{G_n\}$-representations and cones $\widetilde{\mathscr{K}} = \{\widetilde{K}_n \subseteq \widetilde{U}_n\}$ such that*

$$C_n = \{\widetilde{A}_n z : z \in \widetilde{K}_n, \ \widetilde{B}_n z = w_n\},$$

for $\widetilde{A}_n \in \mathcal{L}(V_n, \widetilde{U}_n)^{G_n}$, $\widetilde{B}_n \in \mathcal{L}(\widetilde{W}_n, \widetilde{U}_n)^{G_n}$ and $w_n \in \widetilde{W}_n^{G_n}$, then C_n is given by (ConicSeq) with

$$\mathscr{W} = \widetilde{\mathscr{U}}, \qquad \mathscr{U} = \widetilde{\mathscr{U}} \oplus \widetilde{\mathscr{W}} \oplus \mathscr{V}, \qquad \mathscr{K} = \widetilde{\mathscr{K}} \oplus 0 \oplus 0,$$

and

$$A_n = \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}, \quad B_n = \begin{bmatrix} I \\ -\widetilde{B}_n \\ -\widetilde{A}_n \end{bmatrix}, \quad u_n = \begin{bmatrix} 0 \\ w_n \\ 0 \end{bmatrix}.$$

Note that $\widetilde{A}_n, \widetilde{B}_n, w_n$ are freely-described if and only if A_n, B_n, u_n are. The description of the simplex in Example 2.15 is an example of such a reformulation.

Remark 2.17. *The convex sets we obtain from Definition 2.14 are often group-invariant themselves. Indeed, standard sequences of cones are often invariant under natural sequences of groups. For example, nonnegative orthants are permutation-invariant, while PSD cones are orthogonally-invariant. It is then easy to check that the convex sets we obtain from Definition 2.14 are all group-invariant. This is not a fundamental restriction of our framework however; to describe non-invariant sets, the different components in a conic description could be required to be invariant under the actions of different sequences of groups.*
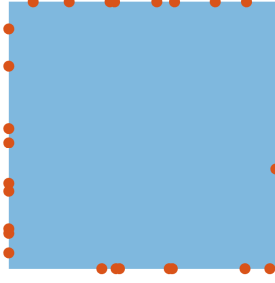
Figure 2: Given points on 2D unit squre, recover unit cube in all dimensions.

Definitions 2.12-2.14 enable us to derive parametric families of free descriptions by finding freely-described bases for the relevant spaces of invariants. Such a basis can be found either manually by giving expressions for basis elements as functions of n and checking Definition 2.12, or computationally as explained in Section 5.1 below.

Example 2.18. *Consider the following illustrative problem. Given points on the boundary of the unit square $[-1,1]^2$ as in Figure 2, recover the sequence $\{[-1,1]^n\}$ of unit cubes for all n. This is a simple task that nevertheless requires parametrizing sequences of convex sets, one in every dimension, and identifying the right parameters fitting the two-dimensional data. We can obtain the desired parametric families from our framework. Indeed, let $\mathcal{V} = \{\mathbb{R}^n\}$ and $G_n = \mathsf{B}_n$ as in Example 2.2(a). For the description spaces, define*

$$\mathscr{U} = \{\mathbb{R}^{2n+1}\} \oplus \{\mathbb{R}^n\}, \qquad \mathscr{K} = \{\mathbb{R}^{2n+1}_+ \oplus 0\}, \qquad \mathscr{W} = \{\mathbb{R}^n\} \oplus \mathbb{R},$$

where $\mathbb{R}^n$ is viewed as a representation of S_n extended to B_n (so signs act trivially), and $\{\mathbb{R}^{2n+1}\}$ is the consistent sequence from the description of the ℓ_1 ball in Example 2.15. We then obtain the following 12-parameter family of freely-described convex sets

$$
\begin{aligned}
C_n = \bigg\{ & x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n, \beta \in \mathbb{R} \ s.t. \ \alpha_1 \begin{bmatrix} I \\ -I \\ 0 \end{bmatrix} x + \begin{bmatrix} \alpha_2 I + \alpha_3 \mathbb{1}_n \mathbb{1}_n^\top & \alpha_4 \mathbb{1}_n \\ \alpha_2 I + \alpha_3 \mathbb{1}_n \mathbb{1}_n^\top & \alpha_4 \mathbb{1}_n \\ \alpha_5 \mathbb{1}_n^\top & \alpha_6 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} + \begin{bmatrix} \alpha_7 \mathbb{1}_n \\ \alpha_7 \mathbb{1}_n \\ \alpha_8 \end{bmatrix} \geq 0, \\
& \begin{bmatrix} \alpha_9 I + \alpha_{10} \mathbb{1}_n \mathbb{1}_n^\top & \alpha_{11} \mathbb{1}_n \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} + \alpha_{12} \mathbb{1}_n = 0 \bigg\}.
\end{aligned}
\tag{2}
$$

This family contains the sequence $\{[-1,1]^n\}$ of unit cubes, which can be written as

$$[-1,1]^n = \left\{ x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n, \beta \in \mathbb{R} \ s.t. \ \begin{bmatrix} I \\ -I \\ 0 \end{bmatrix} x + \begin{bmatrix} I & 0 \\ I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} \geq 0, \ \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} - \mathbb{1}_n = 0 \right\}. \tag{3}$$

To obtain a larger family, we choose the larger description spaces

$$\mathscr{U} = \mathscr{W} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 1}\{\mathbb{R}^{2n+1}\}) = \{\mathbb{S}^{2n+2}\}, \qquad \mathscr{K} = \{\mathbb{S}^{2n+2}_+\}.$$

Then the dimensions of the relevant spaces of invariants are:

$$\dim \mathcal{L}(V_n, U_n)^{G_n} = 4, \quad \dim \mathcal{L}(W_n, U_n)^{G_n} = 108, \quad \dim W_n^{G_n} = 8, \quad \text{for all } n \geq 4. \tag{4}$$

We obtain these dimension counts using the algorithm in Section 5.1, see Example 5.2(c).

2.3 Compatibility across dimensions

The following example shows that a freely-described sequence of sets may not extend correctly to higher dimensions when fitted to data in a fixed dimension.

Example 2.19. *Consider again the problem of fitting the sequence of cubes from Example 2.18. The following freely-described sequence is contained in the family in (2) and fits the two-dimensional data in Figure 2 perfectly*

$$C_n = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} I \\ -I \\ 0 \end{bmatrix} x + \begin{bmatrix} I & 0 \\ I & 0 \\ -\mathbb{1}_n^\top & -1 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \geq 0, \ \begin{bmatrix} I - \mathbb{1}_n \mathbb{1}_n^\top & \mathbb{1}_n \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} = 0 \right\},$$
$$= \frac{3}{2n-1}[-1,1]^n.$$

However, we have $C_n \neq [-1,1]^n$ for all $n \neq 2$, so this sequence incorrectly extends to other dimensions. The issue is that Definition 2.14 of free descriptions only relates descriptions *of convex sets across dimensions, rather than the sets themselves.*

In this section, we define and study relations between the *sets* in a sequence themselves, which we call compatibility conditions. As we will see here and in Section 3, our conditions are satisfied by sequences of sets arising in various applications including the unit cubes, but not by the sequence in Example 2.19. By searching only over sequences of sets satisfying these conditions, we are guaranteed to obtain sets that are not only freely-described, but also correctly related across dimensions.

Definition 2.20 (Compatibility conditions). *Let $\mathcal{V} = \{V_n\}$ be a consistent sequence of $\{G_n\}$-representations. Let $\mathcal{C} = \{C_n \subseteq V_n\}$ be a sequence of convex sets. We say that $\mathcal{C}$ satisfies*

Intersection compatibility *if $C_{n+1} \cap V_n = C_n$;*

Projection compatibility *if $\mathcal{P}_{V_n} C_{n+1} = C_n$.*

Example 2.21. *The following are simple examples of sequences of convex sets satisfying the conditions in Definition 2.20.*

(Simplices) *Let $V_n = \mathbb{R}^n$ and $G_n = \mathsf{S}_n$ as in Example 2.2(a). Consider the sequence of standard simplices Δ^{n-1} (Example 1.1(a)), The sequence $\mathcal{C} = \{\Delta^{n-1}\}$ satisfies intersection compatibility but not projection compatibility.*

(ℓ_p norm balls) *For the same V_n, G_n, the sequence of ℓ_p unit-norm balls $C_n = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$ satisfies intersection and projection compatibility for each $p \in [1, \infty)$.*

(Stability number) *Let $V_n = \mathbb{S}^n$ and $G_n = \mathsf{S}_n$ as in Example 2.2(b), and let $\alpha_n^{-1} : \mathbb{S}^n \to \mathbb{R}$ be the inverse stability number given by $\alpha_n^{-1}(X) = \min_{x \in \Delta^{n-1}} x^\top (X + I_n) x$, see [56]. These are concave functions whose superlevel sets $C_n = \{X \in \mathbb{S}^n : \alpha_n(X)^{-1} \geq 1\}$ are projection-compatible but not intersection-compatible, because if $\alpha_n(X) > 0$ then $\alpha_{n+1}(X) = \alpha_n(X) + 1$.*

We mention that intersection and projection compatibility are dual to each other. The following result is an instance of [53, Cor. 16.3.2].

Proposition 2.22. *Suppose $\{V_n\}$ is a consistent sequence and let $\mathcal{C} = \{C_n \subseteq V_n\}$ be a sequence of convex bodies.[5] Define the polar sequence $\mathcal{C}^\circ = \{C_n^\circ \subseteq V_n\}$ by (polar). Then $\mathcal{C}$ is intersection-compatible if and only if $\mathcal{C}^\circ$ is projection-compatible.*

Example 2.19 shows that a freely-described sequence of sets may not satisfy either compatibility condition. We therefore derive conditions on descriptions that certify our compatibility conditions.

Theorem 2.23. *Let $\{V_n\}, \{W_n\}, \{U_n\}$ be consistent sequences of $\{G_n\}$-representations, let $\{K_n \subseteq U_n\}$ be a sequence of convex cones, and let $\mathcal{C} = \{C_n \subseteq V_n\}$ be described by linear maps $\{A_n : V_n \to U_n\}, \{B_n : W_n \to U_n\}$ and elements $\{u_n \in U_n^{G_n}\}$ as in (ConicSeq). Assume that the cones $\{K_n\}$ are both intersection-compatible and projection-compatible.*

---

[5]A *convex body* is a compact convex set containing the origin in its interior.

(a) *If $\{A_n\}, \{B_n\}, \{B_n^*\}$ are morphisms, $\{u_n\}$ is freely-described, and $u_{n+1} - u_n \in K_{n+1}$ for all n, then $\mathscr{C}$ is intersection-compatible. If, in addition, $\{A_n^*\}$ is a morphism, then $\mathscr{C}$ is also projection-compatible.*

(b) *If $\{A_n\}, \{A_n^*\}\{B_n\}, \{B_n^*\}$ are morphisms, $\{u_n\}$ is freely-described, and $u_{n+1}-u_n \in K_{n+1}+A_{n+1}(V_n^{\perp})+ B_{n+1}(W_n)$, then $\mathscr{C}$ is projection-compatible.*

Proof. (a) First, we show $C_n \subseteq C_{n+1}$. If $x \in C_n$ then there exists $y \in W_n$ satisfying $A_n x + B_n y + u_n \in K_n$. Then

$$A_{n+1}x + B_{n+1}y + u_{n+1} = A_n x + B_n y + u_n + (u_{n+1} - u_n) \in K_{n+1},$$

where we used the facts that $\{A_n\}, \{B_n\}$ are morphisms, that $K_n \subseteq K_{n+1}$ by intersection-compatiblity, and that $u_{n+1} - u_n \in K_{n+1}$. Thus, $x \in C_{n+1}$.

Second, we show $C_{n+1}\cap V_n \subseteq C_n$. If $x \in C_{n+1}\cap V_n$, there exists $y \in W_{n+1}$ satisfying $A_{n+1}x+B_{n+1}y+ u_{n+1} \in K_{n+1}$. Because $\{A_n\}$ is a morphism, we have $A_{n+1}x = A_n x$ and hence $\mathcal{P}_{U_n}A_{n+1}x = A_n x$. Because $\{B_n^*\}$ is a morphism, we have $\mathcal{P}_{U_n}B_{n+1} = B_n \mathcal{P}_{W_n}$, hence $\mathcal{P}_{U_n}B_{n+1}y = B_{n+1}(\mathcal{P}_{W_n}y)$. Finally, we have $\mathcal{P}_{U_n}u_{n+1} = u_n$ because $\{u_n\}$ is freely-described, and $\mathcal{P}_{U_n}K_{n+1} \subseteq K_n$ because $\{K_n\}$ is projection-compatible. Thus, applying $\mathcal{P}_{U_n}$ we obtain $A_n x + B_n(\mathcal{P}_{W_n}y) + u_n \in K_n$, showing that $x \in C_n$. We have shown that $\mathscr{C}$ is intersection-compatible.

Because $\mathscr{C}$ is intersection-compatible, we have $\mathcal{P}_{V_n}C_{n+1} \supseteq C_n$. Conversely, if $x \in C_{n+1}$ then $A_{n+1}x+ B_{n+1}y+u_{n+1} \in K_{n+1}$ for some $y \in W_{n+1}$. If $\{A_n^*\}$ is a morphism, then $\mathcal{P}_{U_n}A_{n+1} = A_n\mathcal{P}_{V_n}$. Applying $\mathcal{P}_{U_n}$ to both sides we obtain $A_n\mathcal{P}_{V_n}x + B_n\mathcal{P}_{W_n}y + u_n \in K_n$ and hence $\mathcal{P}_{V_n}x \in C_n$, showing that $\mathscr{C}$ is projection-compatible.

(b) First, we show $\mathcal{P}_{V_n}C_{n+1} \subseteq C_n$. If $x \in C_{n+1}$ then there is $y \in W_{n+1}$ satisfying $A_{n+1}x+B_{n+1}y+u_{n+1} \in K_{n+1}$. Applying $\mathcal{P}_{U_n}$ to both sides and using the facts that $\{A_n^*\}, \{B_n^*\}$ are morphisms, that $\{u_n\}$ is freely-described, and that $\{K_n\}$ is projection-compatible, we obtain $A_n(\mathcal{P}_{V_n}x)+B_n(\mathcal{P}_{W_n}y)+u_n \in K_n$, showing that $\mathcal{P}_{V_n}x \in C_n$.

Second, we show $C_n \subseteq \mathcal{P}_{V_n}C_{n+1}$. Suppose $A_n x + B_n y + u_n \in K_n$ for $x \in V_n$. Let $x_{\perp} \in V_n^{\perp}$ and $y' \in W_n$ satisfy $u_{n+1} - u_n + A_{n+1}x_{\perp} + B_{n+1}y' \in K_{n+1}$. As $\{A_n\}, \{B_n\}$ are morphisms,

$$A_{n+1}(x + x_{\perp}) + B_{n+1}(y + y') + u_{n+1} = A_n x + B_n y + u_n + (u_{n+1} - u_n + A_{n+1}x_{\perp} + B_{n+1}y') \in K_{n+1},$$

hence $x + x_{\perp} \in C_{n+1}$ and $\mathcal{P}_{V_n}(x + x_{\perp}) = x$. This shows $\mathscr{C}$ is projection-compatible. $\square$

Interestingly, the conditions in Theorem 2.23, which arise naturally when certifying compatibility, actually *imply* that the descriptions are free. We say that a sequence of conic descriptions *certifies compatibility* when it satisfies the hypotheses of Theorem 2.23.

Remark 2.24. *We make a number of remarks about the conditions in Theorem 2.23.*

- *Standard sequences of cones such as nonnegative orthants and PSD cones satisfy both intersection and projection compatibility.*

- *The set of linear maps $\{A_n\}$ and $\{B_n\}$ satisfying the hypotheses of Theorem 2.23(a) form* linear subspaces *of the corresponding spaces of freely-described elements. The set of sequences $\{u_n\}$ satisfying those hypotheses form a* convex cone.

  Similarly, we can parametrize descriptions satisfying the hypotheses of Theorem 2.23(b) by a convex cone, by considering $\{u_n\}$ of the form $u_n = A_n(v_n) + B_n(w_n) + z_n$ where $\{v_n \in V_n^{G_n}\}, \{w_n \in W_n^{G_n}\}$, and $\{z_n \in K_n^{G_n}\}$ are freely-described.

- *Freely-described sets satisfying the hypotheses of Theorem 2.23(b) need not be intersection-compatible, as the elliptope in (12) studied below demonstrates.*

- *Free descriptions certifying compatibility often extend to descriptions of infinite-dimensional limits, see Theorem 7.5 below.*

All the descriptions in Example 2.15 certify the compatibility of the sets they describe since they satisfy the hypotheses of Theorem 2.23(a). We treat more sophisticated examples in Section 3, where we show that many descriptions arising in practice certify compatibility as well, and that we can obtain rich parametric families of such descriptions.

Example 2.25 (Learning the cube). *We can attempt to recover the sequence of unit cubes from the data in Figure 2 by searching only over members of the family (2) certifying intersection and projection compatibility as in Theorem 2.23(a). These form the 7-parameter family*

$$
C_n = \left\{ x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n, \beta \in \mathbb{R} \ s.t. \ \alpha_1 \begin{bmatrix} I \\ -I \\ 0 \end{bmatrix} x + \begin{bmatrix} \alpha_2 I & 0 \\ \alpha_2 I & 0 \\ 0 & \alpha_3 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} + \begin{bmatrix} \alpha_4 \mathbb{1}_n \\ \alpha_4 \mathbb{1}_n \\ \alpha_5 \end{bmatrix} \geq 0, \right.
$$
$$
\left. \begin{bmatrix} \alpha_6 I & 0 \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} + \alpha_7 \mathbb{1}_n = 0 \right\}, \qquad with \ \alpha_4, \ \alpha_7 \geq 0.
$$

The only member of this family fitting the data is the desired sequence of cubes (3).

2.4   Convex functions from sets and their compatibility

Thus far, our theory only concerned convex sets. However, problems such as task 2 in Section 1 of approximating the quantum entropy involve convex *functions*. Fortunately, there are several correspondences between convex sets and functions that allow us to tackle such problems using our theory, see Section 1.3. Our compatibility conditions for sets in Definition 2.20 can be translated via these correspondences to the following conditions on functions. Denote by $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ the extended real line and recall that if $f \colon V \to \overline{\mathbb{R}}$ is a convex function and $A \colon V \to U$ is linear, then $Af \colon U \to \mathbb{R}$ is a convex function defined by $(Af)(x) = \inf_{x' \in A^{-1}(x)} f(x')$, see [53, Thm. 5.7].

Definition 2.26 (Compatibility conditions for functions). *Let $\{V_n\}$ be a consistent sequence of $\{G_n\}$-representations. Let $\mathfrak{f} = \{f_n \colon V_n \to \overline{\mathbb{R}}\}$ be a sequence of convex functions. We say $\mathfrak{f}$ satisfies*

    Intersection compatibility *if $f_{n+1}|_{V_n} = f_n$;*

    Projection compatibility *if $\mathcal{P}_{V_n} f_{n+1} = f_n$.*

The following proposition shows that compatibility of functions can be derived from compatibility of sets using the above correspondences.

Proposition 2.27. *Let $\mathscr{V} = \{V_n\}$ be a consistent sequence of $\{G_n\}$-representations.*

(a) *If a sequence $\{E_n \subseteq V_n \oplus \mathbb{R}\}$ of convex subsets of $\mathscr{V} \oplus \mathbb{R}$ is intersection (resp., projection) compatible, then $\{f_{E_n} \colon V_n \to \mathbb{R}\}$ defined by (epi) is intersection (resp., projection) compatible.*

(b) *If a sequence $\mathscr{C} = \{C_n \subseteq V_n\}$ of convex subsets of $\mathscr{V}$ is intersection (resp., projection) compatible, then $\{\gamma_{C_n}\}$ defined by (gauge) is intersection (resp., projection) compatible.*

(c) *Let $\mathscr{C}$ be as in (b). If $\mathscr{C}$ is projection-compatible, then $\{h_{C_n}\}$ defined by (supp) is intersection-compatible. If $\mathscr{C}$ is intersection-compatible and either all the C_n are compact or $\mathrm{ri}(C_{n+1}) \cap V_n \neq \emptyset$ for all n, then $\{h_{C_n}\}$ is projection-compatible.*

Proof. Part (a) follows from the identities $f_{E_{n+1}}|_{V_n} = f_{E_{n+1} \cap (V_n \oplus \mathbb{R})}$ and $\mathcal{P}_{V_n} f_{E_{n+1}} = f_{\mathcal{P}_{V_n \oplus \mathbb{R}} E_{n+1}}$, as can be verified directly from (epi). Similalrly, part (b) follows from $\gamma_{C_{n+1}}|_{V_n} = \gamma_{C_{n+1} \cap V_n}$ and $\mathcal{P}_{V_n} \gamma_{C_{n+1}} = \gamma_{\mathcal{P}_{V_n} C_{n+1}}$ as can be verified directly from (gauge). For (c), we have $h_{C_{n+1}}|_{V_n} = h_{\mathcal{P}_n C_{n+1}}$ by [53, Cor. 16.3.1], from which the first claim follows. If C_{n+1} is compact, then Sion's minimax theorem gives

$$
(\mathcal{P}_{V_n} h_{C_{n+1}})(x) = \inf_{x' \in \mathcal{P}_{V_n}^{-1}(x)} \sup_{\widetilde{x} \in C_{n+1}} \langle \widetilde{x}, x' \rangle = \sup_{\widetilde{x} \in C_{n+1}} \inf_{x' \in \mathcal{P}_{V_n}^{-1}(x)} \langle \widetilde{x}, x' \rangle = \sup_{\widetilde{x} \in C_{n+1} \cap V_n} \langle \widetilde{x}, x' \rangle = h_{C_{n+1} \cap V_n}.
$$

This identity, which also holds if $\mathrm{ri}(C_{n+1}) \cap V_n \neq \emptyset$ by [53, Cor. 16.3.1], yields the second claim. $\qquad \square$

    The compatibility conditions in Definition 2.26 naturally arise in the context of inverse problems.

Example 2.28 (compatibility in inverse problems). *Consider a consistent sequence $\mathscr{V} = \{V_n\}$ of $\{G_n\}$- representations. A popular approach to recover $x \in V_n$ from $m \in \mathbb{N}$ linear observations takes as input a forward map $A\colon V_n \to \mathbb{R}^m$ and data $y \in \mathbb{R}^m$ and outputs*

$$F_{m,n}(A, y) = \operatorname*{argmin}_{x \in V_n} \, f_n(x) + \lambda \|Ax - y\|_2^2, \tag{5}$$

where $f_n\colon V_n \to \mathbb{R}$ is a convex regularizer promoting desired structure in the solution.

The maps (5) *constitute an algorithm which can clearly be instantiated for any $(A, y) \in \mathcal{L}(V_n, \mathbb{R}^m) \oplus \mathbb{R}^m$ and for any $n, m \in \mathbb{N}$. It is desirable for the different maps in this collection to satisfy*

$$F_{m,n+1}(A\mathcal{P}_n, y) = F_{m,n}(A, y), \tag{6}$$

whenever the corresponding minimizers are unique. Indeed, this says that if the data only depends on the component of $x \in V_{n+1}$ in V_n, then the recovered solution should also lie in V_n to avoid overfitting. When does (6) *hold? If the sequence of regularizers $\mathfrak{f} = \{f_n\}$ is projection-compatible, then*

$$\min_{\widetilde{x} \in V_{n+1}} f_{n+1}(\widetilde{x}) + \lambda \|A\mathcal{P}_n \widetilde{x} - b\|_2^2 = \min_{x \in V_n} \min_{\substack{\widetilde{x} \in V_{n+1} \\ \mathcal{P}_n \widetilde{x} = x}} f_{n+1}(\widetilde{x}) + \lambda \|Ax - y\|_2^2 = \min_{x \in V_n} f_n(x) + \lambda \|Ax - y\|_2^2$$

Moreover, if $x_* = F_{m,n}(A, y)$ minimizes $f_n(x) + \lambda \|Ax - y\|_2^2$ and $\mathfrak{f}$ is intersection-compatible, then $f_n(x_*) + \lambda \|Ax_* - y\|_2^2 = f_{n+1}(x_*) + \lambda \|A\mathcal{P}_n x_* - y\|_2^2$ and hence $x_* = F_{m,n+1}(A\mathcal{P}_n, y)$, showing (6).

3 Examples of freely-described and compatible sets

We consider additional, more sophisticated examples of freely-described sets and functions arising in the literature and the compatibility conditions that they satisfy.

3.1 Regular polygons

The following example illustrates a natural sequence of convex sets that is freely-described but satisfies neither intersection nor projection compatibility. Let $\mathscr{V} = \{(V_n, \varphi_n)\}$ be the consistent sequence $V_n = \mathbb{R}^2$ with $\varphi_n = \operatorname{id}_{\mathbb{R}^2}$ and the standard action of the dihedral group $G_n = \operatorname{Dih}_{2^n}$. Consider the sequence of regular 2^n-gons $\mathscr{C} = \{C_n \subseteq \mathbb{R}^2\}$ defined by

$$C_n = \operatorname{conv}\left\{\begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix}\right\}, \quad \theta_i = \frac{2\pi i}{2^n}, \quad i \in \{0, \ldots, 2^n - 1\}. \tag{7}$$

Because $V_n = V_{n+1}$ while the sets $C_n \neq C_{n+1}$, the sequence $\mathscr{C}$ satisfies neither intersection nor projection compatibility. Nevertheless, it admits the free description

$$C_n = \left\{ x \in \mathbb{R}^2 : \exists y \in \mathbb{R}^{2^n} \text{ s.t. } \begin{bmatrix} -I \\ 0 \\ 0 \end{bmatrix} x + \begin{bmatrix} \cdots & \begin{matrix} \cos(2\pi i/2^n) \\ \sin(2\pi i/2^n) \end{matrix} & \cdots \\ & I_{2^n} & \\ & \mathbb{1}_{2^n}^\top & \end{bmatrix} y + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \in 0 \oplus \mathbb{R}_+^n \oplus 0 \right\},$$

where $\mathscr{W} = \{\mathbb{R}^{2^n}\}_n$ with embeddings $y \mapsto y \otimes [1, 0]^\top$, and $\mathscr{U} = \mathscr{V} \oplus \mathscr{W} \oplus \mathbb{R}$. We put the standard inner products on $\mathbb{R}^{2^n}$. The group permutes the 2^n vertices of C_n, defining a permutation action on $[2^n]$, and it acts on $\mathbb{R}^{2^n}$ by applying these permutations to coordinates.

If $\theta_i = \pi(2i + 1)/2^n$ in (7) instead, the following semidefinite description of C_n given in [29] is also free. Let

$$W_n = (\mathbb{R}^2)^{n-1} \oplus \mathbb{R}, \quad U_n = (\mathbb{S}^3)^{n-1}, \quad K_n = (\mathbb{S}_+^3)^{n-1},$$

where we embed $(X_1, \ldots, X_{n-1}) \in U_n \mapsto (X_1, \ldots, X_{n-1}, 0) \in U_{n+1}$ and $(x^{(1)}, \ldots, x^{(n-2)}, y) \in W_n \mapsto (x^{(1)}, \ldots, x^{(n-2)}, [y, 0]^\top, 0) \in W_{n+1}$. The action of G_n on U_n and W_n is given in [29]. Then it follows from [29, Thm. 3] that $\{C_n\}$ is given by (ConicSeq) with

$$A_n(x) = \left(\begin{bmatrix} 0 & x^\top \\ x & 0 \end{bmatrix}, 0, \ldots, 0 \right),$$

$$B_n(x^{(1)}, \ldots, x^{(n-2)}, y) = \left(\begin{bmatrix} 0 & 0 \\ 0 & \begin{matrix} \frac{x_1^{(1)}}{2} & \frac{x_2^{(1)}}{2} \\ \frac{x_2^{(1)}}{2} & -\frac{x_1^{(1)}}{2} \end{matrix} \end{bmatrix}, \begin{bmatrix} 0 & (x^{(2)})^\top \\ x^{(2)} & \begin{matrix} \frac{x_1^{(3)}}{2} & \frac{x_2^{(3)}}{2} \\ \frac{x_2^{(3)}}{2} & -\frac{x_1^{(3)}}{2} \end{matrix} \end{bmatrix}, \ldots, \right.$$
$$\left. \begin{bmatrix} 0 & (x^{(n-3)})^\top \\ x^{(n-3)} & \begin{matrix} \frac{x_1^{(n-2)}}{2} & \frac{x_2^{(n-2)}}{2} \\ \frac{x_2^{(n-2)}}{2} & -\frac{x_1^{(n-2)}}{2} \end{matrix} \end{bmatrix}, \begin{bmatrix} 0 & (x^{(n-2)})^\top \\ x^{(n-2)} & \begin{matrix} 0 & \frac{y}{2} \\ \frac{y}{2} & 0 \end{matrix} \end{bmatrix} \right),$$

and $u_n = \left(\operatorname{diag}(1, \frac{1}{2}, \frac{1}{2}), \ldots, \operatorname{diag}(1, \frac{1}{2}, \frac{1}{2}) \right)$. It is straightforward to check that all these sequences are freely-described elements of their corresponding consistent sequences.

3.2 Permutahedra, Schur-Horn orbitopes

In this section, we consider permutahedra, which are convex hulls of all permutations of a vector, and their matrix analogous, the Schur-Horn orbitopes, which are convex hulls of all symmetric matrices with a given spectrum. Permutahedra appear as constraint sets in the relaxation of the 2-SUM problem proposed in [57], with applications ranging from gene sequencing to archaeology. These polytopes have also been studied in [58, 59]. Schur-Horn orbitopes were studied in [60], and their efficient description as projected spectrahedra was derived in [61].

Permutahedra: Consider the sequence of standard permutahedra

$$\operatorname{Perm}_n = \operatorname{conv}\left\{ g \cdot [1, 2, \ldots, n]^\top \right\} = \left\{ M[1, 2, \ldots, n]^\top : M \in \mathbb{R}_+^{n \times n}, \ M\mathbb{1}_n = M^\top \mathbb{1}_n = \mathbb{1}_n \right\}, \tag{8}$$

where the second equality follows by the Birkhoff-von Neumann theorem. The sequence $\{\operatorname{Perm}_n\}$, viewed as subsets of the consistent sequence in Example 2.2(a) with $G_n = \mathsf{S}_n$, is neither intersection- nor projection-compatible. Furthermore, their description (8) is not free because the map $M \mapsto M[1, 2, \ldots, n]^\top$ is not G_n-equivariant. The smaller descriptions of these permutahedra in [62, 63] are also not free because they are not equivariant. However, there is a sequence of permutahedra arising naturally from a limiting perspective that is both intersection- and projection-compatible and whose description certifies this compatibility.

Fix $q, m \in \mathbb{N}$ and a vector $\lambda \in \mathbb{R}^q$ with distinct entries, and define $\widetilde{\lambda} = (\lambda_1, \ldots, \lambda_1, \ldots, \lambda_q, \ldots, \lambda_q) \in \mathbb{R}^m$ in which λ_i appears m_i times (so $\sum_i m_i = m$). Let $G_n = \mathsf{S}_{m2^n}$ embedded in G_{n+1} by sending a $m2^n \times m2^n$ matrix $g \in G_n$ to $g \otimes I_2$. Let $\mathscr{V} = \{\mathbb{R}^{m2^n}\}_n$ with embeddings $x \mapsto x \otimes \mathbb{1}_2$, the normalized inner product $\langle x, y \rangle = (m2^n)^{-1} x^\top y$, and the standard action of G_n. We consider convex hulls of all the vectors in $\mathscr{V}$ containing λ_i in m_i/m of its entries, which are given using the Birkhoff-von Neumann theorem by

$$\begin{aligned} \operatorname{Perm}(\lambda)_n &= \operatorname{conv}\{ g \cdot (\widetilde{\lambda} \otimes \mathbb{1}_{m2^n}) \}_{g \in \mathsf{S}_{m2^n}} \\ &= \left\{ M\lambda : M \in \mathbb{R}_+^{m2^n \times q}, \ M\mathbb{1}_q = \mathbb{1}_{m2^n}, \ M^\top \mathbb{1}_{m2^n} = 2^n[m_1, \ldots, m_q]^\top \right\}. \end{aligned} \tag{9}$$

This is an intersection- and projection-compatible sequence of subsets of $\mathscr{V}$. Moreover, the description (9) is free and certifies intersection and projection compatibility. Indeed, let $\mathscr{W} = \mathscr{V}^{\oplus q} = \{\mathbb{R}^{m2^n \times q}\}_n$ and $\mathscr{U} = \mathscr{W} \oplus \mathscr{V}^{\oplus 2} \oplus \mathbb{R}^q$ containing cones $\{\mathbb{R}_+^{m2^n \times q} \oplus 0 \oplus 0\}$. Then (9) is of the form (ConicSeq) with

$$A_n x = (0, -x, 0, 0), \quad B_n M = (M, \ M\lambda, \ M\mathbb{1}_q, \ (m2^n)^{-1} M^\top \mathbb{1}_{m2^n}),$$

and $u_n = (0, 0, -\mathbb{1}_{m2^n}, -\left[\frac{m_1}{m}, \ldots, \frac{m_q}{m}\right]^\top)$. Note that $B_n^*(M, x_1, x_2, \mu) = M + x_1\lambda^\top - x_2\mathbb{1}_q + \mathbb{1}_{m2^n}\mu^\top$ as the adjoint is taken with respect to the normalized inner product. Thus, $\{A_n\}, \{A_n^*\}, \{B_n\}, \{B_n^*\}$ are morphisms and $u_n = u_{n+1}$ under our embedding, so Theorem 2.23(a) applies.

The insight behind this construction is to view elements of $\{V_n\}$ as piecewise-constant functions in $L^2([0, 1])$ such that $v \in V_n$ takes value v_i on $[(i-1)/m2^n, i/m2^n)$ for $i \in [m2^n]$. Then $\{\operatorname{Perm}(\lambda)_n\}$ is the set of piecewise-constant functions that take values λ_i on a subset of $[0, 1]$ of measure m_i/m. See Section 7 for such limit-based interpretations of compatibility, in particular Example 7.6(a). We can extend our consistent sequence to include vectors and matrices of any size by having our consistent sequence be indexed by the poset $\mathbb{N}$ with the divisibility partial order, see Remark 2.3.

Schur-Horn orbitopes: We consider the matrix analogs of the above permutahedra, which are convex hulls of all matrices with a given spectrum. Let $G_n = \mathsf{O}_{m2^n}$ embedding in G_{n+1} by sending a $m2^n \times m2^n$ matrix g to $g \otimes I_2$, let $V_n = \mathbb{S}^{m2^n}$ embedded in V_{n+1} by $X \mapsto X \otimes I_2$, with normalized inner product $\langle X, Y \rangle = (m2^n)^{-1}\mathrm{Tr}(XY)$, and with the action of G_n by conjugation. Consider the sequence of Schur-Horn orbitopes [64, Eq. (19)]

$$\mathrm{SH}(\lambda)_n = \mathrm{conv}\{g \cdot \mathrm{diag}(\lambda \otimes \mathbb{1}_{2^n})\}_{g \in \mathsf{O}_{m2^n}}$$
$$= \left\{ \sum_{i=1}^{q} \lambda_i Y_i : Y_1, \ldots, Y_q \in V_n \text{ s.t. } \sum_{i=1}^{q} Y_i = I, \ Y_i \succeq 0, \ \mathrm{Tr}(Y_i) = m_i 2^n \text{ for } i = 1, \ldots, q \right\}, \quad (10)$$

which is the matrix analog of (9). This is again a free description certifying both intersection- and projection-compatibility. Indeed, let $\mathscr{W} = \mathscr{V}^{\oplus q}$ and $\mathscr{U} = \mathscr{W} \oplus \mathscr{V}^{\oplus 2} \oplus \mathbb{R}^q$ containing the cones $\{K_n = (\mathbb{S}_+^{m2^n})^{\oplus q} \oplus 0 \oplus 0\}$. Then (10) is of the form (ConicSeq) with

$$A_n X = (0, -X, 0, 0), \quad B_n[Y_i]_{i=1}^q = \left([Y_i]_{i=1}^q, \ \sum_i \lambda_i Y_i, \ \sum_i Y_i, \ (m2^n)^{-1}[\mathrm{Tr}(Y_1), \ldots, \mathrm{Tr}(Y_q)] \right),$$

and $u_n = (0, 0, -I_{m2^n}, -[\frac{m_1}{m}, \ldots, \frac{m_q}{m}]^\top)$. Again, note that $B_n^*([Z_i]_i, X_1, X_2, \mu) = [Z_i + \lambda_i X_1 + X_2 + \mu_i I_{m2^n}]_i$ since the adjoint is taken with respect to the above normalized inner product. Then $\{A_n\}, \{A_n^*\}, \{B_n\}, \{B_n^*\}$ are all morphisms and $u_n = u_{n+1}$ under the above embedding, hence Theorem 2.23(a) applies. Once again, we can extend this consistent sequence to include matrices of any size using Remark 2.3, and this sequence has a limit arising in operator algebras, see Example 7.6(b).

Schur-Horn orbitopes are special cases of so-called *spectral polyhedra* studied in [65]. It would be interesting to identify further examples of freely-described sequences of spectral polyhedra arising in applications.

3.3 Free spectrahedra

The family of free spectrahedra of Example 1.1(e) (discussed more extensively in Section 1.2) is precisely the parametric family we obtain from our recipe from Section 2.2 when choosing the appropriate description spaces and imposing compatibility.

Let $\mathscr{V}_0 = \{\mathbb{S}^n\}$ be the sequence from Example 2.2(b) with the action of $G_n = \mathsf{O}_n$ by conjugation. Fix $d, k \in \mathbb{N}$, and let $\mathscr{V} = \mathscr{V}_0^{\oplus d}$, $U_n = \mathbb{S}^k \otimes \mathscr{V}_0$, and $\mathscr{W} = \{W_n = 0\}$.

As the only morphisms $\mathscr{V}_0 \to \mathscr{V}_0$ are multiples of the identity, and elements of $\mathbb{S}^k \otimes \mathbb{S}^n$ are $\dim \mathbb{S}^k$ blocks of $n \times n$ symmetric matrices, we conclude that the morphisms $\mathscr{V} \to \mathscr{U}$ are precisely maps of the form $(X_1, \ldots, X_d) \mapsto \sum_i L_i \otimes X_i$ for some $L_1, \ldots, L_d \in \mathbb{S}^k$. As the only G_n-invariants in $\mathbb{S}^n$ are multiples of I_n, the space of freely-described elements in $\mathscr{U}$ is $\{\{L_0 \otimes I_n\}_n : L_0 \in \mathbb{S}^k\}$, which satisfy $L_0 \otimes (I_{n+1} - I_n) \succeq 0$ if and only if $L_0 \succeq 0$. We conclude that the parametric family of free descriptions of the form (ConicSeq) satisfying Theorem 2.23(a) is

$$(\mathcal{D}_\mathcal{L})_n = \left\{ (X_1, \ldots, X_d) \in (\mathbb{S}^n)^d : L_0 \otimes I_n + \sum_{i=1}^{d} L_i \otimes X_i \succeq 0 \right\}, \quad L_0 \succeq 0.$$

which are free spectrahedra parametrized by $\mathcal{L} = (L_0, \ldots, L_d)$. It is common to assume either $L_0 = I_k$ (the *monic* case) or $L_0 = 0$ (the *homogeneous* case) [7]. As discussed in 1.2, free spectrahedra are fundamental objects in noncommutative free convex and algebraic geometry, see [7, 14] for an introduction. In particular, they satisfy both intersection and projection compatibility, and more generally closure under so-called *matrix-convex combinations*.

3.4 Spectral functions, quantum entropy and its variants

Let $\mathscr{V} = \{V_n = \mathbb{S}^n\}$ with the action of O_n as in Example 2.2(b), and let $\mathscr{V}' = \{V_n' = \mathbb{R}^n\}$ with the action of S_n as in Example 2.2(a). Recall (e.g., [66]) that a convex function $F_n : V_n \to \mathbb{R}$ is O_n-invariant if and only if there exists an S_n-invariant convex function $f_n : V_n' \to \mathbb{R}$ satisfying $F_n(X) = f_n(\lambda(X))$ where $\lambda(X) \in \mathbb{R}^n$ is the vector of eigenvalues of $X \in \mathbb{S}^n$. Furthermore, the sequence $\{F_n : V_n \to \mathbb{R}\}$ is intersection-compatible if and only if the sequence $\{f_n : V_n' \to \mathbb{R}\}$ is.

Examples of such sequences of functions $\mathfrak{F} = \{F_n\}$ and $\mathfrak{f} = \{f_n\}$ arise in (quantum) information theory, where $\mathfrak{F}$ is the quantum analog of classical information-theoretic parameters $\mathfrak{f}$. These are often intersection-compatible as distributions on n states can be viewed as distributions on $n + 1$ states with zero probability on the last state. For example, the negative entropy and relative entropy and their quantum variants are given by

$$h_n(x) = \sum_i x_i \log x_i, \qquad H_n(X) = h_n(\lambda(X)) = \mathrm{Tr}(X \log X),$$

$$D_n(x, y) = \sum_i x_i \log \frac{x_i}{y_i}, \quad S_n(X, Y) = D_n(\lambda(X), \lambda(Y)) = \mathrm{Tr}(X(\log X - \log Y)). \tag{11}$$

Here $\mathrm{dom}(h_n) = \Delta^{n-1}$ and $\mathrm{dom}(D_n) = (\mathbb{R}^n_+)^2$, while $\mathrm{dom}(H_n) = \mathcal{D}^{n-1}$ is the spectraplex from Example 1.1(b) and $\mathrm{dom}(S_n) = (\mathbb{S}^n_+)^2$. We use the standard convention that $0 \log \frac{0}{y} = 0$ even if $y = 0$, and $x \log \frac{x}{0} = \infty$ when $x \neq 0$ [67, §2.3]. Viewing these sequences of functions as defined over $(\mathscr{V}')^{\oplus 2}$ and $\mathscr{V}^{\oplus 2}$, they are intersection-compatible but not projection-compatible (e.g., their domains are not projection-compatible).

Semidefinite approximations: The functions (11) are not semidefinite-representable (i.e., cannot be evaluated using semidefinite programming), though semidefinite approximations of them have been proposed in the literature [68]. We show that these approximations are freely-described, though these descriptions do not certify intersection compatibility in the sense of Theorem 2.23. The family of approximations of [68] to the negative quantum entropy $\{f_{E_n^{(m,k)}}\}$ is parametrized by $m, k \in \mathbb{N}$ and is given by (epi) with

$$
\begin{aligned}
E_n^{(m,k)} = \Bigg\{ & (X, t) \in V_n \oplus \mathbb{R} \, \Big| \, \exists T_0, \ldots, T_m, Z_0, \ldots, Z_k \in \mathbb{S}^n \text{ s.t. } Z_0 = I_n, \ \sum_{j=1}^m w_j T_j = -2^{-k} T_0, \\
& \begin{bmatrix} Z_i & Z_{i+1} \\ Z_{i+1} & X \end{bmatrix} \succeq 0, \text{ for } i = 0, \ldots, k-1, \ \begin{bmatrix} Z_k - X - T_j & -\sqrt{s_j} T_j \\ -\sqrt{s_j} T_j & X - s_j T_j \end{bmatrix} \succeq 0, \\
& \text{for } j = 1, \ldots, m, \ \mathrm{Tr}(T_0) \leq t \Bigg\},
\end{aligned}
$$

where $s, w \in \mathbb{R}^m$ are the nodes and weights for Gauss-Legendre quadrature.

This is a free description of the form (ConicSeq) which is almost, but not quite, an instance of Theorem 2.23(a). Indeed, let $\mathscr{W} = \mathscr{V}^{\oplus(m+k+1)}$ and $\mathscr{U} = \mathscr{V} \oplus (\mathbb{S}^2 \otimes \mathscr{V})^{\oplus(m+k)} \oplus \mathbb{R}$ containing the cones $\{K_n = \{0\} \oplus (\mathbb{S}^2 \otimes \mathbb{S}^n)_+^{\oplus(m+k)} \oplus \mathbb{R}_+\}$. Define

$$A_n(X, t) = \left(0, \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes X \right)^{\oplus k}, \left(\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \otimes X \right)^{\oplus m}, t \right),$$

$$B_n(T_0, \ldots, T_m, Z_1, \ldots, Z_k) = \left(2^{-k} T_0 + \sum_{j=1}^m w_j T_j, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes Z_1, \bigoplus_{i=1}^{k-1} \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes Z_i + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes Z_{i+1} \right), \right.$$

$$\left. \bigoplus_{j=1}^m \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes Z_k - \begin{bmatrix} 1 & \sqrt{s_j} \\ \sqrt{s_j} & s_j \end{bmatrix} \otimes T_j \right), -\mathrm{Tr}(T_0) \right),$$

$$u_n = \left(0, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes I_n, 0, \ldots, 0 \right).$$

Note that $\{A_n\}$ and $\{B_n\}$ are morphisms, that $\{u_n\}$ is a freely-described element of $\mathscr{U}$ satisfying $u_{n+1} - u_n \in K_{n+1}$, but that $\{B_n^*\}$ is *not* a morphism because of the $\mathrm{Tr}(T_0)$ term in B_n. By the proof of Theorem 2.23, this description certifies that $E_n^{(m,k)} \subseteq E_{n+1}^{(m,k)}$ but not $E_{n+1}^{(m,k)} \cap (V_n \oplus \mathbb{R}) \subseteq E_n^{(m,k)}$. We fit a semidefinite approximation to a variant of the quantum entropy from data in Section 5, which is both freely-described and certifies intersection compatibility.

Parametric families: We can use the description spaces of [68] to derive parametric families of freely-described sets. Note that $\mathbb{S}^2 \otimes \mathbb{S}^n \cong (\mathbb{S}^n)^3$ as O_n-representations, and these isomorphisms commute with

zero-padding, so that $\mathscr{U} \cong \mathscr{V}^{1+3(m+k)} \oplus \mathbb{R}$ as consistent sequences. As $\dim(\mathbb{S}^n)^{\mathsf{O}_n} = 1$ and $\dim \mathcal{L}(\mathbb{S}^n)^{\mathsf{O}_n} = 2$, the dimension of invariants parametrizing free descriptions are

$$\dim \mathcal{L}(V_n, U_n)^{\mathsf{O}_n} = 2[1 + 3(m+k)] + 1, \quad \dim \mathcal{L}(W_n, U_n) = 3(m+k+1)[2(m+k)+1],$$
$$\dim U_n^{\mathsf{O}_n} = 2 + 3(m+k).$$

When $m = k = 3$ (the default values in the implementation of [68]), we get

$$\dim \mathcal{L}(V_n, U_n)^{\mathsf{O}_n} = 57, \quad \dim \mathcal{L}(W_n, U_n)^{\mathsf{O}_n} = 570, \quad \dim U_n^{\mathsf{O}_n} = 29.$$

As the only morphisms of sequences $\mathscr{V} \to \mathscr{V}$ are multiples of the identity, and the only morphisms $\mathscr{V} \to \mathbb{R}$ are multiples of the trace, the dimensions of $\{A_n\}, \{B_n\}$ satisfying Theorem 2.23(a) are

$$\dim \Big\{ \{A_n \colon V_n \to U_n\} \text{ morphism} \Big\} = 3(m+k) + 1 = 29,$$
$$\dim \Big\{ \{B_n \colon W_n \to U_n\} : \text{both } \{B_n\} \text{ and } \{B_n^*\} \text{ morphisms} \Big\} = (m+k+1)[3(m+k)+2] = 290.$$

3.5 Graph parameters

Let $\mathscr{V} = \{\mathbb{S}^n\}$ and $G_n = \mathsf{S}_n$ as in Example 2.2(b). Graph parameters are sequences $\{f_n\}$ of G_n-invariant functions $f_n \colon \mathbb{S}^n \to \mathbb{R}$, since those are precisely the functions of (weighted) graphs only depending on the underlying graph rather than on its labelling. Many standard graph parameters are convex (or concave), see [16]. Some graph parameters, such as the max-cut value, are unchanged if we append isolated vertices to the graph, hence they are intersection-compatible. Others, such as the stability number, are nonincreasing when taking induced subgraphs. Moreover, any small graph occurs as an induced subgraph of a larger graph that has the same parameter value. Such parameters are projection-compatible. Standard graph parameters often admit freely-described convex relaxations satisfying the same compatibility conditions, and whose descriptions certify their compatibility. We consider the max-cut and inverse stability number here as examples and derive parametric families of free descriptions from our framework that may be used to fit graph parameters to data.

Max-cut: Computing the max-cut value of a weighted undirected graph reduces to evaluating the support function of the *cut polytope* $\mathcal{C}_n = \mathrm{conv}\{xx^\top : x \in \{\pm 1\}^n\}$. The sequence of cut polytopes $\mathscr{C} = \{\mathcal{C}_n\}$ viewed as subsets of $\mathscr{V}$ is projection-compatible, hence the sequence of their support functions and the max-cut value itself are intersection-compatible by Proposition 2.27(c). Approximation of the max-cut value reduces to approximation of the cut polytopes. A standard outer approximation of the sequence of cut polytopes is the sequence of elliptopes

$$\mathcal{E}_n = \{X \in \mathbb{S}^n : X \succeq 0, \mathrm{diag}(X) = \mathbb{1}_n\}. \tag{12}$$

The sequence $\{\mathcal{E}_n\}$ also satisfies projection compatibility, and the above is a free description certifying this compatibility. Indeed, let $\mathscr{W} = \{0\}$ and $\mathscr{U} = \mathscr{V} \oplus \{\mathbb{R}^n\}$ where the latter sequence is Example 2.2(a), with cones $\mathscr{K} = \{\mathbb{S}_+^n \oplus \{0\}\}$. Then (12) is of the form (ConicSeq) with $A_n X = (X, \mathrm{diag}(X))$, $B_n = 0$, and $u_n = (0, -\mathbb{1}_n)$. Note that $\{A_n\}, \{A_n^*\}$ are morphisms and $u_{n+1} - u_n = (0, -e_{n+1}) = (e_{n+1} e_{n+1}^\top, 0) - A_{n+1} e_{n+1} e_{n+1}^\top$, hence Theorem 2.23(b) applies. Neither the cut polytopes nor the elliptopes is intersection-compatible, as zero-padding a matrix with unit diagonal entries does not yield such a matrix.

Inverse stability number: Computing the inverse stability number reduces to evaluating the support functions of $\mathcal{D}_n = \mathrm{conv}\{xx^\top : x \in \Delta^{n-1}\}$, see [56]. A natural SDP relaxation for this problem is evaluating the support function of

$$\widetilde{\mathcal{D}}_n = \{X \in \mathbb{S}^n : X \succeq 0, X \geq 0, \mathbb{1}_n^\top X \mathbb{1}_n = 1\},$$

where $X \geq 0$ denotes an entrywise nonnegative matrix. Both $\{\mathcal{D}_n\}$ and $\{\widetilde{\mathcal{D}}_n\}$ are intersection-compatible. Moreover, the above description of $\{\widetilde{\mathcal{D}}_n\}$ is free and certifies this compatibility as in Theorem 2.23(b). Indeed, let $\mathscr{W} = \{0\}$ and $\mathscr{U} = \mathscr{V} \oplus \mathbb{R}$ with cones $K_n = (\mathbb{S}_+^n \cap \mathbb{R}_+^{n \times n}) \oplus 0$. Then the above description of $\widetilde{\mathcal{D}}_n$ is of the form (ConicSeq) with $A_n X = (X, \mathbb{1}_n^\top X \mathbb{1}_n)$, $B_n = 0$, and $u_n = (0, -1)$. Note that $\{A_n\}$ is a morphism and $u_n = u_{n+1}$, hence Theorem 2.23(a). Neither $\mathcal{D}_n$ nor $\widetilde{\mathcal{D}}_n$ is intersection-compatible, see Example 2.21.

Parametric families: Convex graph parameters can be obtained from permutation-invariant convex subsets of $\mathscr{V}$, see [16] for examples, hence it is desirable to obtain expressive parametric families of such sets. To that end, let $\mathscr{W} = \mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 2}\{\mathbb{R}^n\}) = \{\mathbb{S}^{\binom{n+1}{2}}\}$. We compute the dimensions of invariants parametrizing free descriptions using the algorithm in Section 5.1, see Example 5.2(b):

$$\dim \mathcal{L}(V_n, U_n)^{G_n} = 93, \quad \dim \mathcal{L}(W_n, U_n)^{G_n} = 1068, \quad \dim W_n^{G_n} = 17, \quad \text{for all } n \geq 8.$$

Using the same algorithm, the dimensions of sequences $\{A_n\}, \{B_n\}$ certifying intersection compatibility as in Theorem 2.23(a) are

$$\begin{aligned}
\dim \Big\{ \{A_n \colon V_n \to U_n\} \text{ morphism} \Big\} &= 19, \\
\dim \Big\{ \{B_n \colon W_n \to U_n\} : \text{both } \{B_n\} \text{ and } \{B_n^*\} \text{ morphisms} \Big\} &= 104.
\end{aligned} \tag{13}$$

3.6 Graphon parameters

A different embedding between graphs arises in the theory of graphons [69], where a weighted graph $X \in \mathbb{S}^{2^n}$ is viewed as a step function $W_X \colon [0,1]^2 \to \mathbb{R}$ defined by $W_X(x,y) = X_{i,j}$ if $(x,y) \in [(i-1)/2^n, i/2^n) \times [(j-1)/2^n, j/2^n)$. See Figure 3. Note that X and $X \otimes \mathbb{1}_2 \mathbb{1}_2^\top \in \mathbb{S}^{2^{n+1}}$ correspond to the same step function, and that the inner product of two such step functions W_X, W_Y in $L^2([0,1]^2)$ equals the normalized Frobenius inner product $\langle X, Y \rangle = 2^{-2n} \mathrm{Tr}(X^\top Y)$. We therefore define the graphon consistent sequence $\mathscr{V} = \{V_n = \mathbb{S}^{2^n}\}$ with embeddings $\varphi_n(X) = X \otimes \mathbb{1}_{2 \times 2}$, the above normalized inner products, and the action of $G_n = \mathsf{S}_{2^n}$ by conjugation. Here G_n is embedded into G_{n+1} by sending a permutation matrix g to $g \otimes I_2$.
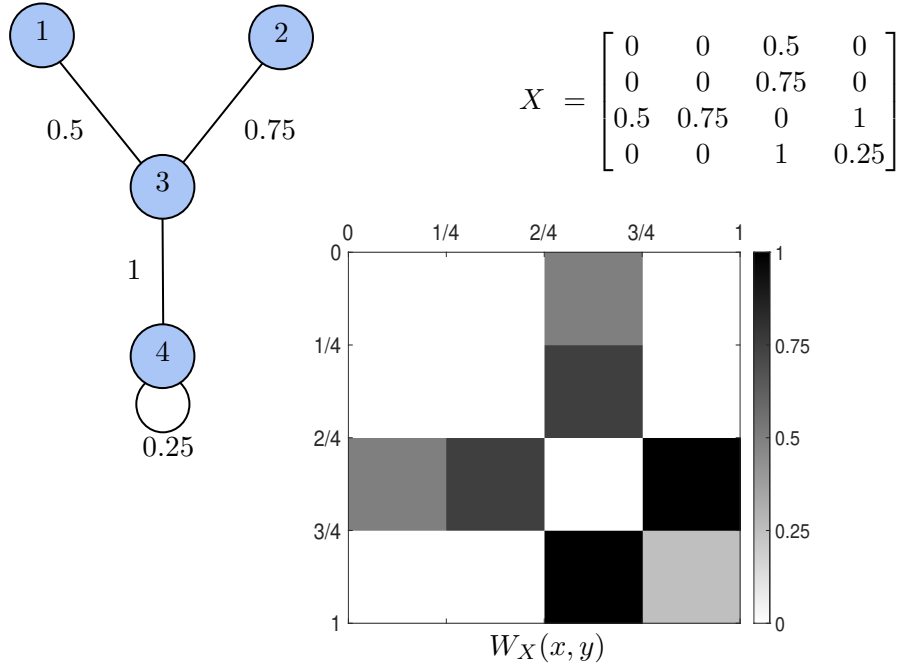

Figure 3: Weighted undirected graph represented as a graph, an adjacency matrix X, and a symmetric function (graphon) W_X on $[0,1]^2$.

The graphon sequence is finitely-generated, as the following computer-assisted proof shows.

Proposition 3.1. *The graphon sequence $\{V_n = \mathbb{S}^{2^n}\}$ is generated in degree 2.*

Proof. Define $E_1^{(n)} = e_1^{(2^n)}(e_1^{(2^n)})^\top$ and $E_2^{(2^n)} = e_1^{(2^n)}(e_2^{(2^n)})^\top + e_2^{(2^n)}(e_1^{(2^n)})^\top$, which span V_n. We verify computationally that $\dim \sum_{i=1}^{2} \mathbb{R}[S_{2^3}]\varphi_2(E_i^{(2)}) = \dim V_3$, see the GitHub repository. Therefore, we can

write

$$E_i^{(3)} = \sum_{j=1}^{2} r_{i,j}\varphi_2(E_j^{(2)}), \quad i \in [2],\ r_{i,j} \in \mathbb{R}[S_{2^3}]. \tag{14}$$

Let $\psi_n(X) = X \oplus 0$ be an embedding of V_n into V_{n+1} by zero-padding, and note that $E_i^{(n+1)} = \psi_n(E_i^{(n)})$ and that ψ_n commutes with a different embedding of S_{2^n} into $\mathsf{S}_{2^{n+1}}$, namely, one that sends $g \mapsto g \oplus I_{2^n}$. Applying ψ_n to (14), we conclude that $E_i^{(n+1)}$ can be written as $\mathbb{R}[S_{2^{n+1}}]$-linear combinations of $\varphi_n(E_i^{(n)})$ for all $n \geq 2$, hence that $\mathbb{R}[S_{2^{n+1}}]\varphi_n(V_n) = V_{n+1}$ for all $n \geq 2$. $\qquad\square$

We can extend our consistent sequence to include symmetric matrices of any size by having our consistent sequence be indexed by the poset $\mathbb{N}$ with the divisibility partial order, see Remark 2.3.

Graphon parameters: A permutation-invariant and intersection-compatible sequence of functions $\mathfrak{f} = \{f_n\colon V_n \to \mathbb{R}\}$ is called a *graphon parameter*, since these are precisely the functions that only depend on graphs via their associated graphons. A family of graphon parameters that plays a central role in the theory of graphons and in extremal combinatorics are graph homomorphism densities [69]. Their convexity is related to weakly-norming graphs and Sidorenko's conjecture, a major open problem in extremal combinatorics [70, 71]. Interestingly, convex graphon parameters that extend continuously to a certain limit of the graphon sequence (Example 7.4(b)) are also projection-compatible by [72, Thm. 3.17].

We get parametric families of graphon parameters by taking the gauge functions of parametric families of intersection-compatible and freely-described convex sets. For example, let $\mathscr{U} = \mathbb{S}^k \otimes \mathscr{V}$ with cones $\mathscr{K} = \{K_n = (\mathbb{S}^k \otimes \mathbb{S}^{2^n})_+\}$ and $\mathscr{W} = \{W_n = 0\}$. Using Example 2.13(c) and Theorem 2.23(a), we get the following parametric family of freely-described and intersection-compatible sets, parametrized by $L_1, \ldots, L_8 \in \mathbb{S}^k$:

$$C_n = \left\{ X \in \mathbb{S}^{2^n} : \frac{\mathbb{1}^\top X \mathbb{1}}{2^{2n}} L_1 \otimes \mathbb{1}\mathbb{1}^\top + \frac{\mathrm{Tr}(X)}{2^n} L_2 \otimes \mathbb{1}\mathbb{1}^\top + L_3 \otimes \frac{1}{n}\left(X\mathbb{1}\mathbb{1}^\top + \mathbb{1}\mathbb{1}^\top X\right) \right.$$

$$\left. + L_4 \otimes \left(\mathrm{diag}(X)\mathbb{1}^\top + \mathbb{1}\mathrm{diag}(X)^\top\right) + L_5 \otimes X + L_7 \otimes \mathbb{1}\mathbb{1}^\top + L_8 \otimes (2^n I_{2^n}) \succeq 0 \right\}, \quad L_8 \succeq 0.$$

Note that all the functions of X appearing in the above description only depend on the associated step function W_X. For example, $\frac{\mathbb{1}^\top X \mathbb{1}}{2^{2n}} = \int_{[0,1]^2} W_X(t,s)\,\mathrm{d}t\,\mathrm{d}s$ and $\frac{\mathrm{Tr}(X)}{2^n} = \int_0^1 W_X(t,t)\,\mathrm{d}t$.

4 Extendability

Theorem 2.23 gives sufficient conditions for a sequence of freely-described convex sets to satisfy compatibility across dimensions. In this section, we consider extending a conically-described set in a fixed dimension to a freely-described sequence of sets satisfying compatibility. As the hypotheses of Theorem 2.23 require sequences of linear maps to be morphisms, we focus on extending a fixed linear map to a morphism of sequences. This question is motivated by our computationaly goals and, to our knowledge, has not been studied in the representation stability literature. We use the results of the present section to computationally parametrize and search over sequences of descriptions satisfying Theorem 2.23 in Section 5. We begin by reviewing additional concepts from the representation stability literature, which we motivate by our new extendability question above.

4.1 Background: Algebraically free sequences, presentation degree

Let $\mathscr{V} = \{V_n\}$, $\mathscr{U} = \{U_n\}$ be consistent sequences of $\{G_n\}$-representations and consider a linear map $A_{n_0} \in \mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$. When can we extend A_{n_0} to a morphism of sequences $\{A_n\}$? We seek conditions on A_{n_0} which are easy to enforce computationally, and we use these conditions in Section 5 to develop an algorithm to parametrize and search over compatible sequences of convex sets. The following proposition gives an equivalent characterization for the existence of such an extension.

Proposition 4.1. *Let $\mathscr{V} = \{V_n\}, \mathscr{U} = \{U_n\}$ be consistent sequences such that $\mathscr{V}$ is generated in degree d, and fix $A_{n_0} \in \mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ for $n_0 \geq d$.*

(a) *There exists an extension $\{A_n \in \mathcal{L}(V_n, U_n)^{G_n}\}_{n < n_0}$ satisfying $A_{n+1}|_{V_n} = A_n$ for all $n < n_0$ if and only if $A_{n_0}(V_j) \subseteq U_j$ for $j \le d$.*

(b) *There exists an extension $\{A_n \in \mathcal{L}(V_n, U_n)^{G_n}\}_{n > n_0}$ satisfying $A_{n+1}|_{V_n} = A_n$ for all $n \ge n_0$ if and only if the following implication holds*

$$\sum_i g_i x_i = 0 \implies \sum_i g_i A_{n_0} x_i = 0, \qquad \text{for all } g_i \in G_n, \ x_i \in V_d, \ n \in \mathbb{N}. \tag{15}$$

If an extension $\{A_n\}$ of A_{n_0} exists, then it is unique.

Proof. (a) If such $\{A_n\}_{n < n_0}$ exists, then it is uniquely given in terms of A_{n_0} by $A_n = A_{n_0}|_{V_n}$. Therefore, we have $A_{n_0}(V_j) = A_j(V_j) \subseteq U_j$ for all $j \le d$. Conversely, suppose $A_{n_0}(V_j) \subseteq U_j$ for $j \le d$. We claim that $A_{n_0}(V_n) \subseteq U_n$ for all $n \ge d$ as well. Indeed, because $\mathcal{V}$ is generated in degree d, we have $A_{n_0}(V_n) = A_{n_0}(\mathbb{R}[G_n]V_d) = \mathbb{R}[G_n]A_{n_0}(V_d) \subseteq \mathbb{R}[G_n]U_d \subseteq U_n$ for $n \ge d$. Defining $A_n = A_{n_0}|_{V_n}$ for each $n < n_0$ yields the desired extension to lower dimensions.

(b) If such $\{A_n\}_{n > n_0}$ exists, it is unique and is explicitly given in terms of A_{n_0} as follows. For any $n > n_0$ and $x \in V_n$, we can write $x = \sum_i g_i x_i$ for some $g_i \in G_n$ and $x_i \in V_d \subseteq V_{n_0}$ by definition of the generation degree. Because $A_n \colon V_n \to U_n$ is G_n-equivariant and satisfies $A_n|_{V_{n_0}} = A_{n_0}$, we have

$$A_n x = A_n \left(\sum_i g_i x_i \right) = \sum_i g_i (A_{n_0} x_i), \tag{16}$$

which expresses A_n in terms of A_{n_0}. The expression (16) shows that (15) is satisfied. Conversely, suppose that (15) is satisfied. For each $n > n_0$ define $A_n \colon V_n \to U_n$ as follows. For any $x \in V_n$, write $x = \sum_i g_i x_i$ for some $g_i \in G_n$ and $x_i \in V_d$, which is possible because $\mathcal{V}$ is generated in degree d, and set $A_n x$ to the right-hand side of (16). This is well-defined because if $x = \sum_i g_i x_i = \sum_j g_j' x_j'$ for $g_i, g_j' \in G_n$ and $x_i, x_j' \in V_d$ then $\sum_i g_i A_{n_0} x_i = \sum_j g_j' A_{n_0} x_j'$ by (15). Moreover, A_n is linear, G_n-equivariant, and extends A_{n_0} by construction, so $\{A_n\}_{n > n_0}$ is the desired extension of A_{n_0}. $\square$

The conditions on A_{n_0} in Proposition 4.1(a) are easy to impose computationally, and we do so in Section 5.1. In contrast, while condition (15) in Proposition 4.1(b) fully characterizes extendability of A_{n_0} to higher dimensions, it is unclear how to impose it computationally. We therefore proceed to study it further. In algebraic terms, elements $x_i \in V_d$ are called the *generators* of $\mathcal{V}$, and expressions of the form $\sum_i g_i x_i = 0$ with $g_i \in G_n$ are called *relations* between those generators over the group G_n. Proposition 4.1(b) shows that A_{n_0} extends to a morphism iff any relation satisfied by the generators of $\mathcal{V}$ is also satisfied by their images under A_{n_0}. We therefore need to understand the relations among the generators in V_d.

We study these relations in two stages. First, we identify two simple types of relations that are satisfied by the images of the generators under A_{n_0} for appropriate $\mathcal{V}, \mathcal{U}$. We then define algebraically free consistent sequences whose generators satisfy only these two types of relations. Second, we express an arbitrary consistent sequence as the quotient of an algebraically free one. The kernel of this quotient morphism is a consistent sequence that encodes all additional relations. To capture the degree starting from which both the generators and the relations between them stabilize, we define the presentation degree of a consistent sequence as the maximum of the generation degree of the sequence itself and that of the above kernel, see Definition 4.9. The presentation degree plays a prominent role in our results in Section 4.2.

We begin carrying out the first stage of the above program and identify two simple types of relations. The first source for relations between generators in V_d arises from relations over G_d. Indeed, if $\sum_i g_i x_i = 0$ for $g_i \in G_d$ and $x_i \in V_d$ then $\sum_i g g_i x_i = 0$ for any $g \in G_n$. Such relations are always satisfied by the images $A_{n_0} x_i$. A second source for such relations arises from subgroups of G_n acting trivially on V_d.

Definition 4.2 (Stabilizing subgroups). *Let $\{V_n\}$ be a consistent sequence of $\{G_n\}$-representations. For any $d \le n$, define the stabilizing subgroups of V_d by*

$$H_{n,d} = \{g \in G_n : g \cdot v = v \text{ for all } v \in V_d\}.$$

Example 4.3. *Different consistent sequences may have the same stabilizing subgroups.*

(a) *If $\mathscr{V} = \{\mathbb{R}^n\}$ with the action of $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a), then $H_{n,d} \cong G_{n-d}$ is the subgroup of (signed) permutations fixing the first d letters. These are also the stabilizing subgroups in Example 2.2(b), and for the sequence $\{\mathbb{R}^{2n+1}\}$ appearing the description of the ℓ_1 norm ball in Example 2.15.*

(b) *If $\mathscr{V} = \{\mathbb{R}^{2^n}\}$ with the action of $G_n = \mathsf{S}_{2^n}$ as in Example 2.2(c), then $H_{n,d}$ is the direct product $(\mathsf{S}_{2^{n-d}})^{2^d}$ of the permutations of each consecutive set of 2^{n-d} coordinates. These are also the stabilizing subgroups for the graphon sequence in Section 3.6. If $G_n = \mathsf{C}_{2^n}$, then $H_{n,d} = \{\mathrm{id}\}$.*

Stabilizing subgroups yield a second source for relations, namely, the relations $(h - \mathrm{id})x = 0$ for all $x \in V_d$ and $h \in H_{n,d}$. Thus, if A_{n_0} extends to a morphism then $A_{n_0}(V_d) \subseteq \bigcap_{n \geq d} U_n^{H_{n,d}}$. Rather than attempt to enforce these constraints computationally, we make a standard simplifying assumption from the representation stability literature. Specifically, we assume that U_d is fixed by $H_{n,d}$ (or a subgroup of it, see below) for all $n \geq d$, in which case there is a simple sufficient condition for the above constraints that we can impose computationally.

Definition 4.4 ($\mathscr{V}$-modules). *Let $\mathscr{V} = \{V_n\}$ and $\mathscr{U} = \{U_n\}$ be consistent sequences of $\{G_n\}$-representations, and let $\{H_{n,d}\}_{d \leq n}$ be the stabilizing subgroups of $\mathscr{V}$ as in Definition 4.2. We say that $\mathscr{U}$ is a $\mathscr{V}$-module if $U_d \subseteq U_n^{H_{n,d}}$ for all $d \leq n$.*

This terminology comes from a categorical approach to representation stability, see Appendix A. Sequences constructed from $\mathscr{V}$-modules as in Remark 2.4 and Proposition 2.10 remain $\mathscr{V}$-modules.

Example 4.5. *$\mathrm{Sym}^2(\mathrm{Sym}^{\leq d}(\bigwedge^k \mathscr{V}))$ is a $\mathscr{V}$-module for any consistent sequence $\mathscr{V}$. This sequence arises in Theorem 6.12 below.*

If $\mathscr{U}$ is a $\mathscr{V}$-module, then imposing $A_{n_0}(V_d) \subseteq U_d$ is sufficient to guarantee $A_{n_0}(V_d) \subseteq \bigcap_{n \geq d} U_n^{H_{n,d}}$ since U_d is contained in the right-hand side of this inclusion. Imposing this sufficient condition can done computationally, see Section 5.1. This concludes the first stage of our program.

To go beyond the above two simple types of relations satisfied by the generators, we begin by defining algebraically free[6] consistent sequences, whose generators do not satisfy any additional types of relations. We then write any consistent sequence as the image under a morphism of sequences of an algebraically free one. The kernel of this morphism precisely captures the relations satisfied by the generators.

Definition 4.6 (Induction, algebraically free sequences). *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations, and for $d \leq n$ let $H_{n,d} \subseteq G_n$ be its stabilizing subgroups.*

(a) *Fix $d \in \mathbb{N}$ and a G_d-representation W. The associated $\mathscr{V}$-induction sequence is defined by*

$$\mathrm{Ind}_{G_d}(W) = \left\{ \mathrm{Ind}_{G_d H_{n,d}}^{G_n} W \right\}_n,$$

where the induced representation is taken to be 0 when $n < d$. This is a $\mathscr{V}$-module.

Here $G_d H_{n,d} = \{gh : g \in G_d, \ h \in H_{n,d}\}$ is the subgroup generated by G_d and $H_{n,d}$ inside G_n.

(b) *A consistent sequence $\mathscr{F}$ is an algebraically free $\mathscr{V}$-module if it is a direct sum of $\mathscr{V}$-induction sequences. The sequence $\mathscr{V}$ itself is algebraically free if it is an algebraically free $\mathscr{V}$-module.*

Definition 4.6 uses the fact that $ghg^{-1} \in H_{n,d}$ for any $g \in G_d$ and $h \in H_{n,d}$ by Definition 4.2, implying that the subgroup generated by G_d and $H_{n,d}$ inside G_n is just the product $G_d H_{n,d}$. To see that $\mathrm{Ind}_{G_d}(W)$ is indeed a consistent sequence of $\{G_n\}$-representations, note that there is a natural inclusion $\mathrm{Ind}_{G_d H_{n,d}}^{G_n} W_d \hookrightarrow \mathrm{Ind}_{G_d H_{n+1,d}}^{G_{n+1}} W_d$ described in Section 1.3 induced by the inclusions $G_d H_{n,d} \subseteq G_d H_{n+1,d}$ and $G_n \subseteq G_{n+1}$ and the fact that $G_d H_{n+1,d} \cap G_n = G_d H_{n,d}$, which follows by Definition 4.2. The inner product and group action on the induced representation with which it becomes a consistent sequence are also described in Section 1.3.

---

[6] Freeness in Definition 4.6 is meant in the algebraic sense of being generated by generators with no nontrivial relations between them, in contrast to Definition 2.14 where it is meant in the sense of dimension-free descriptions.

Example 4.7. *The following are examples of algebraically free sequences.*

    (a) *The sequence $\mathscr{V} = \{V_n = \mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n,$ or S_n as in Example 2.2(a) is the $\mathscr{V}$-induction sequence $\mathscr{V} = \mathrm{Ind}_{G_1} V_1$.*

    (b) *Let $\mathscr{V}$ be as in (a). Then $\mathscr{U} = \{U_n = \mathbb{S}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n,$ or S_n as in Example 2.2(b) is the algebraically free $\mathscr{V}$-module $\mathscr{U} = \bigoplus_{j=1}^2 \mathrm{Ind}_{G_j} W_j$ where $W_1 = U_1$ and $W_2 = \mathbb{R}(e_1 e_2^\top + e_2 e_1^\top) \subseteq U_2$.*

    Any consistent sequence is the image under a morphism of sequences of an algebraically free sequence.

Proposition 4.8. *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations and let $\mathscr{U} = \{U_n\}$ be a $\mathscr{V}$-module. Then $\mathscr{U}$ is generated in degree d if and only if there exists an algebraically free $\mathscr{V}$-module $\mathscr{F}$ generated in degree d and a surjective morphism of sequences $\mathscr{F} \to \mathscr{U}$.*

Proof. If $\mathscr{F} = \{F_n\}$ is a $\mathscr{V}$-module and $\{A_n \colon F_n \to U_n\}$ is a surjective morphism, then for any $n \geq d$ we have $U_n = A_n(F_n) = A_n(\mathbb{R}[G_n]F_d) = \mathbb{R}[G_n]A_n(F_d) = \mathbb{R}[G_n]A_d(F_d) = \mathbb{R}[G_n]U_d$, where we used the fact that $\mathscr{F}$ is generated in degree d; the equivariance of A_n; the fact that $A_n|_{F_d} = A_d$ since $\{A_n\}$ is a morphism; and the surjectivity of A_d. This shows $\mathscr{U}$ is generated in degree d.

    Conversely, if $\mathscr{U}$ is generated in degree d, define the algebraically free $\mathscr{V}$-module $\mathscr{F} = \bigoplus_{i=1}^d \mathrm{Ind}_{G_i}(U_i)$ and consider the morphism $\mathscr{F} \to \mathscr{U}$ defined by $g \otimes u \mapsto g \cdot u$ for each $g \in G_n$, $u \in U_i$, and $i \in [d]$ (see Section 1.3). The image of this morphism in U_n is precisely $\sum_{i=1}^{\min\{d,n\}} \mathbb{R}[G_n]U_i$, hence it is surjective for all n. Finally, $\mathrm{Ind}_{G_i}(U_i)$ is generated in degree i, so $\mathscr{F}$ is generated in degree d. $\qquad\square$

    The kernel of the morphism in Proposition 4.8 precisely encodes all the additional relations beyond the two simple types above satisfied by the generators of $\mathscr{V}$. The generation degree of this kernel then captures the point at which relations stabilize. We therefore define the presentation degree as the maximum of the generation degree of $\mathscr{V}$ and that of this kernel, which captures stabilization of the generators as well as of the relations between them.

Definition 4.9 (Relation and presentation degrees)**.** *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations. We say that a $\mathscr{V}$-module $\mathscr{U}$ is generated in degree d, related in degree r, and presented in degree $k = \max\{d, r\}$ if there exists an algebraically free $\mathscr{V}$-module $\mathscr{F}$ generated in degree d, and a surjective morphism of sequences $\mathscr{F} \to \mathscr{U}$ whose kernel is generated in degree r. The smallest k for which this holds is called the presentation degree of $\mathscr{U}$.*

    Note that the presentation degree is at least as large as the generation degree (cf. Definition 2.5). The presentation degree enables us to strengthen Proposition 2.7 and to quantify more precisely when the projections there become isomorphisms.

Proposition 4.10. *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations and $\mathscr{U}$ be a $\mathscr{V}$-module presented in degree k. Then the maps $\mathcal{P}_{U_n} \colon U_{n+1}^{G_{n+1}} \to U_n^{G_n}$ are isomorphisms for all $n \geq k$.*

Proof. As $\mathscr{U}$ is presented in degree k, there exists an algebraically free $\mathscr{V}$-module $\mathscr{F} = \{F_n\}$ and a surjective morphism $\mathscr{F} \to \mathscr{U}$ such that both its kernel $\mathscr{K} = \{K_n\}$ and $\mathscr{F}$ itself are generated in degree k. Because each map $F_n \to U_n$ is a G_n-equivariant surjection with kernel K_n, its restriction to invariants $F_n^{G_n} \to U_n^{G_n}$ is surjective with kernel $K_n^{G_n}$.

    As $\mathscr{F}$ is an algebraically free $\mathscr{V}$-module, there exist integers d_j and G_{d_j}-representations W_{d_j} satisfying $\mathscr{F} = \bigoplus_j \mathrm{Ind}_{G_{d_j}} W_{d_j}$. Such $\mathscr{F}$ has generation degree $\max_j d_j \leq k$. Therefore, letting $\{H_{n,d}\}$ be the stabilizing subgroups of $\mathscr{V}$, we have for $n \geq k$

$$F_n^{G_n} = \bigoplus_j \left(\mathrm{Ind}_{G_{d_j} H_{n,d_j}}^{G_n} (W_{d_j}) \right)^{G_n} \cong \bigoplus_j W_{d_j}^{G_{d_j}},$$

see Section 1.3. Thus, $\dim F_n^{G_n}$ is constant for $n \geq k$. Moreover, by Proposition 2.7 and the fact that $\mathscr{K}$ and $\mathscr{U}$ are generated in degree k, we have $\dim K_n^{G_n} \geq \dim K_{n+1}^{G_{n+1}}$ and similarly $\dim U_n^{G_n} \geq \dim U_{n+1}^{G_{n+1}}$ for all $n \geq k$.

    By the rank-nullity theorem, we have $\dim U_n^{G_n} = \dim F_n^{G_n} - \dim K_n^{G_n}$. As $\dim F_n^{G_n}$ is constant while both $\dim U_n^{G_n}$ and $\dim K_n^{G_n}$ are nonincreasing for $n \geq k$, we conclude that they are all constant for $n \geq k$. To conclude, note that $\mathcal{P}_{U_n}$ is injective for all $n \geq k$ by Proposition 2.7. $\qquad\square$

The presentation degree allows us to ensure condition (15) is satisfied and hence to extend a fixed linear map to a morphism of sequences in Theorem 4.14, thus answering the question posed in the beginning of this section. Therefore, the presentation degree appears in our extendability result for convex sets (Theorem 4.15), and in our algorithm for computationally parametrizing such sets (Section 5.1). Because of the importance of the presentation degree to our framework, we give a calculus for these degrees. This calculus allows us to bound the presentation degrees of complicated sequences constructed from simpler sequences with known presentation degrees. The following theorem is a consequence of results in [38, 39, 40] concerning calculus for generation degrees. We combine these results to obtain the following analogous calculus for presentation degrees, whose proof is given in Appendix A.

Theorem 4.11 (Calculus for generation and presentation degrees). *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations and let $\mathscr{W}, \mathscr{U}$ be $\mathscr{V}$-modules generated in degrees d_W, d_U and presented in degrees k_W, k_U, respectively. Then*

(Sums) *$\mathscr{W} \oplus \mathscr{U}$ is generated in degree $\max\{d_W, d_U\}$ and presented in degree $\max\{k_W, k_U\}$.*

(Images and kernels) *If $\mathscr{A} : \mathscr{W} \to \mathscr{U}$ and $\mathscr{A}^*$ are morphisms, then $\operatorname{im}\mathscr{A}$ and $\ker\mathscr{A}$ are generated in degree d_W and presented in degree k_W.*

If $\mathscr{V} = \{V_n = \mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a), we further have

(Tensors) *$\mathscr{W} \otimes \mathscr{U}$ is generated in degree $d_W + d_U$ and presented in degree $\max\{k_W + d_U, k_U + d_W\}$.*

(Sym and $\bigwedge$) *$\operatorname{Sym}^\ell \mathscr{W}$, $\bigwedge^\ell \mathscr{W}$ are generated in degree ℓd_W and presented in degree $(\ell - 1)d_W + k_W$.*

Proof. This follows from Theorem A.13 in the appendix. $\square$

Example 4.12. *Let $\mathscr{V} = \{\mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a), which is algebraically free and generated in degree 1. Therefore,*

(a) *$\mathscr{V}^{\otimes k}, \operatorname{Sym}^k \mathscr{V}, \operatorname{Sym}^{\leq k} \mathscr{V}, \bigwedge^k \mathscr{V}$ are generated and presented in degree k, $\operatorname{Sym}^2(\operatorname{Sym}^{\leq k} \mathscr{V})$ is generated and presented in degree $2k$, and $\operatorname{Sym}^\ell(\bigwedge^k \mathscr{V})$ is generated and presented in degree $k\ell$.*

(b) *$\mathscr{V} \otimes \operatorname{Sym}^2(\operatorname{Sym}^{\leq k} \mathscr{V}) = \{\mathcal{L}(\mathbb{R}^n, \mathbb{S}^{\binom{n+k}{k}})\}$ is generated and presented in degree $2(k+1)$ and $[\operatorname{Sym}^2(\operatorname{Sym}^{\leq k} \mathscr{V})]^{\otimes 2} = \{\mathcal{L}(\mathbb{S}^{\binom{k+n}{k}})\}$ is generated and presented in degree $4k$.*

It would be interesting to extend the calculus of Theorem 4.11 to other sequences of groups, such as $G_n = \mathsf{O}_n$, as well as to modules over the consistent sequence of graphons from Section 3.6. It would also be interesting to algorithmically compute presentation degrees, such as by extending the computer-assisted proof of Proposition 3.1.

Remark 4.13 ($\mathscr{V}$-modules vs. stabilizing subgroups). *All of the statements in this section assume a "base" consistent sequence $\mathscr{V}$. Note however that these statements only depend on the stabilizing subgroups $\{H_{n,d}\}$ of $\mathscr{V}$. In fact, any sequence of subgroups $\{H_{n,d} \subseteq G_n\}_{d \leq n}$ satisfying $H_{n+1,d} \supseteq H_{n,d}$, $H_{n,d+1} \subseteq H_{n,d}$, and $H_{n+1,d} \cap G_n = H_{n,d}$ for all $d \leq n$ arise as stabilizing subgroups of a consistent sequence of $\{G_n\}$-representations.*

The stabilizing subgroups play a central role because they determine the sets of embeddings $\{g\varphi_{n-1} \cdots \varphi_d\}_{g \in G_n} \cong G_n/H_{n,d}$ of V_d into V_n, and the combinatorics of these embeddings yields Theorem 4.11. See the proof of [38, Prop. 2.3.6] and Appendix A for example. We formulate our results in terms of $\mathscr{V}$-modules rather than the subgroups $\{H_{n,d}\}$ directly because the sequences we use are often constructed from the same base sequence using Remark 2.4, easing the application of our results, see Section 5 for example.

4.2 Extension to a freely-described and compatible sequence

Let $\mathscr{V} = \{V_n\}, \mathscr{U} = \{U_n\}, \mathscr{W} = \{W_n\}$ be consistent sequences of $\{G_n\}$-representations, and let $\{K_n \subseteq U_n\}$ be convex cones satisfying both intersection and projection compatibility. In this section, we fix n_0 and consider a convex subset $C_{n_0} \subseteq V_{n_0}$ described as in (ConicSeq). We want to extend C_{n_0} with its fixed description to a freely-described and compatible sequence.

If n_0 exceeds the presentation degrees of $\mathscr{W}, \mathscr{U}, \mathscr{V} \otimes \mathscr{U}$ and $\mathscr{W} \otimes \mathscr{U}$, then Proposition 4.10 shows that we can uniquely extend all the invariants in the description of C_{n_0} to freely-described elements, and hence extend C_{n_0} to a freely-described sequence. How can we ensure that this unique extension satisfies our compatibility conditions (Definition 2.20)? More specifically, when do the extensions of the above invariants to freely-described elements satisfy the conditions of Theorem 2.23?

We begin by using the presentation degree to simplify our characterization in Proposition 4.1 for when a fixed equivariant linear map extends to a morphism. The following theorem answers the motivating question we presented in Section 4.

Theorem 4.14. *Let $\mathscr{V}_0$ be a consistent sequence of $\{G_n\}$-representations and let $\mathscr{V} = \{V_n\}, \mathscr{U} = \{U_n\}$ be two $\mathscr{V}_0$-modules. Assume $\mathscr{V}$ is generated in degree d and presented in degree k, and fix $n_0 \geq k$. If $A_{n_0} \in \mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ satisfies $A_{n_0}(V_j) \subseteq U_j$ for $j \leq d$, then A_{n_0} extends to a morphism of sequences $\mathscr{V} \to \mathscr{U}$.*

Proof. Let $\{H_{n,d}\}$ be the stabilizing subgroups of $\mathscr{V}_0$. Suppose first that $\mathscr{V} = \mathscr{F} = \bigoplus_j \operatorname{Ind}_{G_{d_j}} W_{d_j}$ is free. Note that it is generated in degree $\max_j d_j \leq d$.

Let $A_{d_j} = A_{n_0}|_{W_{d_j}}$ and fix $n \geq d_j$. Because $\mathscr{U}$ is a $\mathscr{V}_0$-module, we have $U_{d_j} \subseteq U_n^{H_{n,d_j}}$, so we can view U_{d_j} as a representation of $G_{d_j} H_{n,d_j}$ on which H_{n,d_j} acts trivially. As $A_{d_j}(W_{d_j}) \subseteq U_{d_j}$ and is $G_{d_j} H_{n,d_j}$-equivariant, the following composition defines an equivariant map

$$A_{n,j} \colon \operatorname{Ind}_{G_{d_j} H_{n,d_j}}^{G_n}(W_{d_j}) \xrightarrow{\operatorname{Ind}(A_{d_j})} \operatorname{Ind}_{G_{d_j} H_{n,d_j}}^{G_n} U_{d_j} \xrightarrow{g \otimes u \mapsto g \cdot u} U_n,$$

where the induced map $\operatorname{Ind}(A_{d_j})$ was defined in Section 1.3. Note that $A_{n_0,j} = A_{n_0}|_{\operatorname{Ind}_{G_{d_j}}(W_{d_j})_{n_0}}$, since $A_{n_0,j}(g \otimes w) = g \cdot A_{n_0}w$ for all $g \in G_n$ and $w \in W_{d_j}$. Also, $\{A_{n,j}\}$ defines a morphism $\operatorname{Ind}_{G_{d_j}}(W_{d_j}) \to \mathscr{U}$. Therefore, the desired extension of A_{n_0} to a morphism of sequences $\{A_n\}$ is given by $A_n = \bigoplus_j A_{n,j} \colon V_n \to U_n$.

Now suppose $\mathscr{F}$ is an algebraically free $\mathscr{V}$-module as above with a surjection $\mathscr{F} \to \mathscr{V}$ whose kernel $\mathscr{K} = \{K_n\}$ is generated in degree k. Define the composition

$$\widetilde{A}_{n_0} \colon F_{n_0} \to V_{n_0} \xrightarrow{A_{n_0}} U_{n_0},$$

which satisfies $\widetilde{A}_{n_0}(F_j) \subseteq U_j$ for all $j \leq d$ by assumption and $\widetilde{A}_{n_0}(K_{n_0}) = 0$ by its definition. By the previous paragraph, it extends to a morphism $\{\widetilde{A}_n \colon F_n \to U_n\}$. Because $\mathscr{K}$ is generated in degree k and $n_0 \geq k$, we have $K_n = \mathbb{R}[G_n]K_{n_0}$. Because $\widetilde{A}_n$ is equivariant, we have $\widetilde{A}_n(K_n) = 0$. Therefore, $\widetilde{A}_n$ can be factored as $F_n \to F_n/K_n = V_n \xrightarrow{A_n} U_n$, where the maps A_n in this factorization give the desired extension of A_{n_0} to a morphism $\mathscr{V} \to \mathscr{U}$. $\square$

Comparing Theorem 4.14 with Proposition 4.1, we see that condition (15) there is satisfied merely by choosing n_0 exceeding the presentation degree. To satisfy the conditions in Theorem 2.23, we also use Theorem 4.14 to ensure $\{A_n^*\}$ defines a morphism. To that end, note that $A_{n_0}^*(U_j) \subseteq V_j$ if and only if $A_{n_0}(V_j^\perp) \subseteq U_j^\perp$, where orthogonal complements are taken inside V_{n_0} and U_{n_0}. We can now give conditions guaranteeing extendability of a convex set to a freely-described and compatible sequence. We use the following theorem to computationally parametrize and search over such sequences of sets in Section 5.

Theorem 4.15 (Parametrizing freely-described and compatible sequences)**.** *Let $\mathscr{V}_0$ be a consistent sequence of $\{G_n\}$ representations and let $\mathscr{V} = \{V_n\}$, $\mathscr{W} = \{W_n\}$, and $\mathscr{U} = \{U_n\}$ be $\mathscr{V}_0$-modules generated in degrees d_V, d_U, d_W, respectively, and presented in degree k. Let $\{K_n \subseteq U_n\}$ be a sequence of convex cones that is both intersection- and projection-compatible. Fix $n_0 \geq k$.*

Let $A_{n_0} \in \mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$, $B_{n_0} \in \mathcal{L}(W_{n_0}, U_{n_0})^{G_{n_0}}$ satisfying

$$A_{n_0}(V_i) \subseteq U_i \text{ for } i \leq d_V, \quad B_{n_0}(W_i) \subseteq U_i \text{ for } i \leq d_W, \quad B_{n_0}(W_i^\perp) \subseteq U_i^\perp \text{ for } i \leq d_U. \tag{17}$$

Also let $u_{n_0} \in U_{n_0}^{G_{n_0}}$.

30

Then there are unique extensions of A_{n_0} and B_{n_0} to morphisms of sequences $\{A_n\}\colon \mathscr{V} \to \mathscr{U}$ and $\{B_n\}\colon \mathscr{W} \to \mathscr{U}$ such that $\{B_n^*\}$ is a morphism of sequences $\mathscr{U} \to \mathscr{W}$ as well. Furthermore, there is a unique extension of u_{n_0} to a freely-described element $\{u_n\}$. Let $\mathscr{C} = \{C_n\}$ be the freely-described sequence given by (ConicSeq).

(a) *If $u_{n+1} - u_n \in K_{n+1}$ for all n, then $\mathscr{C}$ is intersection-compatible. If we also have $A_{n_0}(V_i^{\perp}) \subseteq U_i^{\perp}$ for $i \leq d_U$, then it is also projection-compatible.*

(b) *If $u_{n+1} - u_n \in K_{n+1} + A_{n+1}(V_n^{\perp}) + B_{n+1}(U_{n+1})$ for all n, then $\mathscr{C}$ is projection-compatible.*

Proof. This is a combination of Theorem 4.14 and Theorem 2.23. $\qquad\square$

In Section 5, we use Theorem 4.15 to computationally parametrize and search over free descriptions certifying compatibility. We do so by fixing n_0 as in the theorem and finding a basis for the space of linear maps A_{n_0}, B_{n_0} satisfying (17), and searching over the coefficients in this basis. In contrast, we lack a good description for the collection of u_{n_0} satisfying the hypotheses of Theorem 4.15, such as the convex cone of u_{n_0} satisfying the hypotheses in part (a). In the numerical examples in Section 5, we fix a freely-described element $\{u_n\}$ satisfying $u_{n+1} - u_n \in K_{n+1}$ in advance and only search over linear maps A_{n_0}, B_{n_0}. In those examples, we have $\mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq k}\mathscr{V})$ for some consistent sequence $\mathscr{V}$ and $K_n = \mathrm{Sym}_+^2(\mathrm{Sym}^{\leq k}V_n)$ are the corresponding PSD cones, in which case the identities $\{u_n = \mathrm{id}|_{\mathrm{Sym}^{\leq k}V_n}\}$ form such a sequence. Here we view $\mathrm{Sym}^2(\mathrm{Sym}^{\leq k}V_n)$ as self-adjoint endomorphisms of $\mathrm{Sym}^{\leq k}V_n$, and the cone K_n as those endomorphisms which are positive-semidefinite.

Example 4.16. *Let $G_n = \mathsf{B}_n, \mathsf{D}_n,$ or S_n, and let $\mathscr{V}_0 = \{\mathbb{R}^n\}$ as in Example 2.2(a). Let $\mathscr{V} = \mathrm{Sym}^2\mathscr{V}_0 = \{\mathbb{S}^n\}$ and choose description spaces $\mathscr{W} = \mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq k}\mathscr{V}_0) = \{\mathbb{S}^{\binom{n+k}{k}}\}$ with the corresponding PSD cones $\mathscr{K} = \mathrm{Sym}_+^2(\mathrm{Sym}^{\leq k}\mathscr{V}_0)$.*

Theorem 4.11 shows that the presentation degrees of $\mathscr{V}$ and $\mathscr{W} = \mathscr{U}$ are 2 and $2k$, respectively, and their generation degrees are equal to their presentation degrees. Fix $n_0 \geq 2k$, and suppose $C_{n_0} \subseteq \mathbb{S}^{n_0}$ is given by (ConicSeq) *with $A_{n_0}(V_j) \subseteq U_j$ for $j \leq 2$, $B_{n_0}(W_\ell) \subseteq U_\ell$ and $B_{n_0}(W_\ell^{\perp}) \subseteq U_\ell^{\perp}$ for $\ell \leq 2k$, and $u_{n_0} = I_{\binom{n+k}{k}}$. By Theorem 4.15, the unique extension of C_{n_0} to a freely-described sequence $\mathscr{C}$ is intersection-compatible. If, furthermore, $A_{n_0}(V_j^{\perp}) \subseteq U_j^{\perp}$ for $j \leq 2k$ then $\mathscr{C}$ is also projection-compatible.*

5 Computationally parametrizing and searching over descriptions

In this section, we explain how to computationally parametrize free descriptions of convex sets, and search over them in the context of certain regression problems such as the one from Section 1. Our Matlab implementation of the following algorithms can be found at

<div align="center">

https://github.com/eitangl/anyDimCvxSets.

</div>

5.1 Computationally parametrizing descriptions

Suppose we seek a parametric family of convex subsets $\{C_n \subseteq V_n\}$ of a consistent sequence $\mathscr{V} = \{V_n\}$ of $\mathscr{G} = \{G_n\}$-representations. Both $\mathscr{V}$ and $\mathscr{G}$ are usually dictated by the application at hand and the symmetries it exhibits. We then choose description spaces $\mathscr{U} = \{U_n\}, \mathscr{W} = \{W_n\}$ and convex cones $\mathscr{K} = \{K_n \subseteq U_n\}$. These are usually constructed from $\mathscr{V}$ using Remark 2.4, and are therefore $\mathscr{V}$-modules. There is a trade-off in choosing description spaces; large description spaces yield more expressive families, but optimization over the convex sets in these families becomes more expensive. In this section, we explain how to parametrize free descriptions with and without certifying compatibility as in Theorem 2.23. Our procedure is summarized in Algorithm 1.

Fix $n_0 \in \mathbb{N}$. We now explain how to find bases for equivariant linear maps A_{n_0} and B_{n_0} that extend to freely-described elements of $\mathscr{V} \otimes \mathscr{U}$ and $\mathscr{W} \otimes \mathscr{U}$, and to morphisms $\mathscr{V} \to \mathscr{U}$ and $\mathscr{W} \to \mathscr{U}$, respectively. We also specify how large n_0 needs to be for our procedure to work. This elaborates steps 5, 8 and 11 in Algorithm 1.

Algorithm 1 Learn a freely-described (and possibly compatible) sequence of convex sets.

1: **Input:** Consistent sequences $\mathcal{V}, \mathcal{W}, \mathcal{U}$, cones $\mathcal{K}$, and freely-described $\{u_n \in U_n^{G_n}\}$ satisfying $u_{n+1} - u_n \in K_n$.
2: **Output:** Freely-described $\{A_n\}, \{B_n\}$.
3: **if** compatibility not required **then**
4:     Fix $n_0 \geq$ presentation degrees of $\mathcal{V} \otimes \mathcal{U}$ and $\mathcal{W} \otimes \mathcal{U}$.
5:     Find bases for $\mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ and $\mathcal{L}(W_{n_0}, U_{n_0})^{G_{n_0}}$.
6: **else**
7:     Fix $n_0 \geq$ presentation degrees of $\mathcal{V}, \mathcal{W}, \mathcal{U}$.
8:     Find bases for subspaces of $\mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ and $\mathcal{L}(W_{n_0}, U_{n_0})^{G_{n_0}}$ satisfying the hypotheses of Theorem 4.15.
9: **end if**
10: Search over coefficients in above basis to learn A_{n_0}, B_{n_0} at level n_0.
11: For any $n > n_0$, find unique equivariant A_n, B_n projecting onto A_{n_0}, B_{n_0}.

Step 5: Computing basis for equivariant maps. We explain how to compute a basis for invariants in a fixed vector space, then instantiate the algorithm to perform step 5. If V is a representation of a group G, a vector $v \in V$ is G-invariant iff $g \cdot v = v$ for all $g \in G$, which can be rewritten as $v \in \ker(g - I)$ for all $g \in G$. Thus, finding a basis for invariants in a fixed vector space reduces to finding a basis for the kernel of a matrix, though this matrix may be very large or even infinite. We can dramatically reduce the size of this matrix by only considering discrete and continuous generators of G [73]. Formally,

Theorem 5.1 ([73, Thm. 1])**.** *Let G be a real Lie group with finitely-many connected components acting on a vector space V via the homomorphism $\rho: G \to \mathrm{GL}(V)$. Let $\{H_i\}$ be a basis for the Lie algebra $\mathfrak{g}$ of G and $\{h_j\}$ be a finite collection of discrete generators. Then*

$$v \in V^G \iff \mathrm{D}\rho(H_i)v = 0 \text{ and } (\rho(h_j) - \mathrm{id}_V) \cdot v = 0 \quad \text{for all } i, j.$$

Sets of Lie algebra bases and discrete generators for various standard groups are given in [73]. For example, $G = \mathsf{S}_n$ is generated by two elements, namely, the transposition $(1, 2)$ and the n-cycle $(1, \ldots, n)$, reducing the number of group elements that must be considered from the naïve $n!$ to two. For $G = \mathsf{O}_n$, a basis for the Lie algebra $\mathfrak{g} = \mathrm{Skew}(n)$ is given by $E_{i,j} = e_i e_j^\top - e_j e_i^\top$ for $i < j$, and only a single discrete generator such as $h_1 = \mathrm{diag}(-1, 1, \ldots, 1)$ is needed, for a total of $\binom{n}{2} + 1$ elements.

As equivariant linear maps are precisely the invariants in the space of linear maps, Theorem 5.1 allows us to obtain a basis for equivariant maps between fixed vector spaces. Explicitly, if $\rho_V: G_{n_0} \to \mathrm{GL}(V_{n_0})$ and $\rho_U: G_{n_0} \to \mathrm{GL}(U_{n_0})$ are the group homomorphisms defining the actions of G_{n_0} on V_{n_0}, U_{n_0}, then $A_{n_0} \in \mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ if and only if

$$\mathrm{D}\rho_U(H_i)A_{n_0} - A_{n_0}\mathrm{D}\rho_V(H_i) = 0, \quad \rho_U(h_j)A_{n_0} - A_{n_0}\rho_V(h_j) = 0, \quad \text{for all } i, j. \tag{18}$$

The equations (18) express the space $\mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ as the kernel of a matrix, which is often very large and sparse. A basis for the kernel of such a matrix can be computed using its LU decomposition as in [74, 75], or using the algorithm of [73, §5]. A basis for the space $\mathcal{L}(W_{n_0}, U_{n_0})^{G_{n_0}}$ is obtained analogously.

Step 8: Computing a basis for extendable equivariant linear maps. By Theorem 4.15, to find a basis for maps A_{n_0} extending to a morphism we need to find a basis for equivariant linear maps satisfying $A_{n_0}(V_i) \subseteq U_i$ for $i \leq d_V$ where d_V is the generation degree of $\mathcal{V}$. This is again a linear condition on A_{n_0}. Indeed, defining $\varphi_{n_0,i} = \varphi_{n_0-1} \cdots \varphi_i: V_i \hookrightarrow V_{n_0}$ if $i < n_0$ and $\varphi_{n_0,n_0} = \mathrm{id}_{V_{n_0}}$, and similarly for $\psi_{n_0,i}$, we have

$$A_{n_0}(V_i) \subseteq U_i \iff (I - \mathcal{P}_{U_i})A_{n_0}|_{V_i} = 0 \iff (I - \psi_{n_0,i}\psi_{n_0,i}^*)A_{n_0}\varphi_{n_0,i} = 0. \tag{19}$$

The subspace of $\mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ satisfying the hypotheses of Theorem 4.15 is thus again the kernel of a matrix obtained by combining (18) and (19). To also impose $A_{n_0}(V_i^\perp) \subseteq U_i^\perp$ for $i \leq d_U$ where d_U is the

32

generation degree of $\mathscr{U}$, so that $A_{n_0}^*$ extends to a morphism, note that

$$A_{n_0}(V_i^\perp) \subseteq U_i^\perp \iff \mathcal{P}_{U_i} A_{n_0}|_{V_i^\perp} = 0 \iff \psi_{n_0,i}^* A_{n_0}(I - \varphi_{n_0,i}\varphi_{n_0,i}^*) = 0, \tag{20}$$

hence the corresponding subspace of $\mathcal{L}(V_{n_0}, U_{n_0})^{G_{n_0}}$ is the kernel of the matrix obtained by combining (18)-(20). The subspace of $\mathcal{L}(W_{n_0}, U_{n_0})^{G_{n_0}}$ satisfying the hypotheses of Theorem 4.15 is again the kernel of a matrix and its basis is computed similarly.

Step 11: Extending to higher dimensions. Given equivariant A_{n_0}, B_{n_0}, we wish to extend them to freely-described elements. These extensions are morphisms if the hypotheses of Theorem 4.14 are satisfied. We do so computationally by solving a linear system for each $n > n_0$ to which we wish to extend. Specifically, the map $A_n \in \mathcal{L}(V_n, U_n)$ extending a fixed A_{n_0} is the unique solution to the linear system (18) (with n_0 replaced by n) and $\psi_{n,n_0}^* A_n \varphi_{n,n_0} = A_{n_0}$. If A_{n_0} is equivariant and n_0 exceeds the presentation degree of $\mathscr{V} \otimes \mathscr{U}$, or if A_{n_0} also satisfies (19) and n_0 exceeds the presentation degree of $\mathscr{V}$, then this system has a unique solution for any $n > n_0$ by Proposition A.11 and Theorem 4.14. This linear system is typically large and sparse, and we solve it using LSQR [76]. The extension of B_{n_0} is handled similarly, except that n_0 needs to exceed the presentation degrees of both $\mathscr{W}$ and $\mathscr{U}$ to guarantee that both B_{n_0} and $B_{n_0}^*$ extend to morphisms.

Example 5.2 (Dimension counts). *We use the above algorithm to obtain dimension counts for the spaces of linear maps $\{A_n\}, \{B_n\}$ and $\{u_n\}$ parametrizing free descriptions. See the functions* `compute_dims_a`, `compute_dims_b`, *and* `compute_dims_c` *on GitHub for the code computing the dimensions in the three examples below.*

(a) *Let $\mathscr{V} = \{\mathbb{R}^n\}$ with the action of $G_n = \mathsf{S}_n$ as in Example 2.2(a), and let $\mathscr{W} = \mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 2}\mathscr{V})$. Then $\mathscr{V}, \mathscr{U}, \mathscr{V} \otimes \mathscr{U}, \mathscr{W} \otimes \mathscr{U}$ are all $\mathscr{V}$-modules and are presented in degrees $1, 4, 5, 8$, respectively, by Theorem 4.11.*

The dimensions of invariants parametrizing free descriptions were given in (4) as $\dim \mathcal{L}(V_n, U_n)^{G_n} = 39$, $\dim \mathcal{L}(W_n, U_n)^{G_n} = 1068$, and $\dim U_n^{G_n} = 17$ for $n \geq 8$. The dimensions of linear maps $\{A_n\}$ and $\{B_n\}$ satisfying Theorem 2.23 are

$$\dim \Big\{ \{A_n \colon V_n \to U_n\} \text{ morphism} \Big\} = 6,$$

$$\dim \Big\{ \{B_n \colon W_n \to U_n\} : \text{both } \{B_n\} \text{ and } \{B_n^*\} \text{ morphisms} \Big\} = 104.$$

If we further require $\{A_n^*\}$ to be a morphism, the dimension of $\{A_n\}$ decreases to 5.

(b) *Let $\mathscr{V} = \{\mathbb{S}^n\}$ with the action of $G_n = \mathsf{S}_n$ as in Example 2.2(a), and let $\mathscr{W}, \mathscr{U}$ be as in (a). Then $\mathscr{V}, \mathscr{U}, \mathscr{V} \otimes \mathscr{U}, \mathscr{W} \otimes \mathscr{U}$ are all $\mathscr{V}$-modules (by Example 4.3(a)) and are presented in degrees $2, 4, 6, 8$, respectively.*

The dimension of invariant $\{A_n\}$ in this case is $\dim \mathcal{L}(V_n, U_n)^{G_n} = 93$, and the dimensions of $\dim \mathcal{L}(W_n, U_n)^{G_n}$ and $\dim W_n^{G_n}$ are as in (a). The dimensions of linear maps $\{A_n\}$ and $\{B_n\}$ satisfying Theorem 2.23 are 19 and 104, respectively, see (13). If we further require $\{A_n^*\}$ to be a morphism, the dimension of $\{A_n\}$ decreases to 12.

(c) *Let $\mathscr{V} = \{\mathbb{R}^n\}$ with the action of $G_n = \mathsf{B}_n$ as in Example 2.2(a), let $\mathscr{V}' = \{\mathbb{R}^{2n+1}\}$ with the action of $G_n = \mathsf{B}_n$ as in the description of the ℓ_1 ball in Example 2.15, and let $\mathscr{W} = \mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 1}\mathscr{V}')$. Then $\mathscr{V}, \mathscr{U}, \mathscr{V} \otimes \mathscr{U}, \mathscr{W} \otimes \mathscr{U}$ are all $\mathscr{V}$-modules (by Example 4.3(a)) and are presented in degrees $1, 2, 3, 4$, respectively.*

The dimensions of invariants in this case are

$$\dim \mathcal{L}(V_n, U_n)^{G_n} = 4, \quad \dim \mathcal{L}(W_n, U_n)^{G_n} = 108, \quad \dim U_n^{G_n} = 8, \quad \text{for all } n \geq 4.$$

The dimensions of linear maps $\{A_n\}$ and $\{B_n\}$ satisfying Theorem 2.23 are

$$\dim \Big\{ \{A_n \colon V_n \to U_n\} \text{ morphism} \Big\} = 3,$$

$$\dim \Big\{ \{B_n \colon W_n \to U_n\} : \text{both } \{B_n\} \text{ and } \{B_n^*\} \text{ morphisms} \Big\} = 37. \tag{21}$$

If we further require $\{A_n^*\}$ to be a morphism, the dimension does not decrease in this case.

5.2 Searching over descriptions: Convex regression

We give an algorithm to search over the parametric families of descriptions obtained in Section 5.1 in the context of the following regression problem, which includes our tasks from Section 1. Let $\{V_n\}$ be a consistent sequence of $\{G_n\}$-representations and let $\mathfrak{f}^{\mathrm{true}} = \{f_n^{\mathrm{true}} \colon V_n \to \mathbb{R}\}$ be an unknown sequence of functions we wish to estimate given some of its values. Assume the following conditions are satisfied by $\mathfrak{f}^{\mathrm{true}}$:

- f_n^{true} is G_n-invariant for all n,
- $\mathfrak{f}^{\mathrm{true}}$ is intersection-compatible (possibly also projection-compatible),
- f_n^{true} is a gauge function, i.e., a nonnegative positively-homogeneous convex function.

We are given evaluation data $\{(x_i, f_{n_0}^{\mathrm{true}}(x_i))\}_{i=1}^D \subseteq V_{n_0} \times \mathbb{R}$ to which we fit a freely-described and intersection-compatible sequence of gauge functions of the form

$$f_n(x; A_n, B_n, \lambda) = \inf_{\substack{t \geq 0 \\ y \in W_n}} t + \lambda\|y\| \quad \text{s.t. } A_n x + B_n y + t u_n \in K_n \tag{P}$$

$$= \sup_{z \in U_n} -\langle z, A_n x \rangle \quad \text{s.t. } \|B_n^* z\|_* \leq \lambda, \ \langle z, C_n \rangle \leq 1, \ z \in K_n^*, \tag{D}$$

where we take $u_n \in \mathrm{int}(K_n)$ so that Slater's condition is satisfied, and the values of the primal (P) and dual (D) programs above are equal. If $\mathfrak{f}^\dagger$ is also projection-compatible, then we wish the fitted sequence $\{f_n\}$ to be projection-compatible as well. Here $\|\cdot\|$ is an arbitrary intersection- and projection-compatible norm on $\mathscr{U}$, and $\|\cdot\|_*$ is its dual norm. In all experiments below, we use $\mathscr{U} = \mathrm{Sym}^2(\mathrm{Sym}^{\leq k}\mathscr{V}')$ for some $\mathscr{V}$-module $\mathscr{V}'$ and some k with the corresponding PSD cones $\mathscr{K} = \mathrm{Sym}_+^2(\mathrm{Sym}^{\leq k}\mathscr{V}')$ and Frobenius norm. We then choose $\{u_n \in \mathrm{int}(K_n)\}$ to be the sequence of identity matrices. The case $\lambda = 0$ gives the gauge function (gauge) of (ConicSeq), which satisfies intersection and projection compatibility under the conditions in Theorem 2.23(a). It is easy to show that the above $\mathfrak{f}$ is intersection and projection compatible under the same conditions.

To fit $A_{n_0}, B_{n_0}, \lambda$ to the given data, we consider the optimization problem.

$$\min_{\substack{\varepsilon \in \mathbb{R}_+^D \\ A_{n_0}, B_{n_0}, \lambda \geq \lambda_{\min} \\ \{(t_i, y_i)\}, \{z_i\}}} \|\varepsilon\| \qquad \text{s.t.} \tag{Regress}$$

$$(y_i, t_i) \text{ feasible for (P) with cost } \leq f_{n_0}^{\mathrm{true}}(x_i) + \varepsilon_i, \tag{PC}$$

$$z_i \text{ feasible for (D) with cost } \geq f_{n_0}^{\mathrm{true}}(x_i) - \varepsilon_i, \tag{DC}$$

$$A_{n_0}, B_{n_0} \text{ satisfy } (17). \tag{Ext}$$

Constraints (PC) and (DC) are required to hold for all $i \in [D]$.

The constraints (PC) and (DC) ensure that $f_{n_0}^{\mathrm{true}} - \varepsilon_i \leq f_{n_0}(x_i; A_n, B_n, \lambda) \leq f_{n_0}^{\mathrm{true}}(x_i) + \varepsilon_i$, hence minimizing $\|\varepsilon\|$ fits the data. The constraint (Ext) ensures that the fitted f_{n_0} extends to a freely-described and compatible sequence $\mathfrak{f} = \{f_n \colon V_n \to \mathbb{R}\}$. We enforce the last constraint by finding a basis for feasible A_{n_0} and B_{n_0} using the algorithm from Section 5.1 and optimizing over the coefficients in that basis. We omit constraint (Ext) when optimizing over freely-described (but not necessarily compatible) sequences.

As (Regress) involves bilinear constraints, we tackle that program via *alternating minimization*, where we alternate between fixing $A_{n_0}, B_{n_0}, \lambda$ and $\{(t_i, y_i)\}, \{z_i\}$ while optimizing over the rest of the variables. Note that Slater's condition holds in (Regress) for both steps of alternating minimization when $u_n \in \mathrm{int}(K_n)$.

The regularization parameter λ is kept above a positive threshold $\lambda_{\min}$ to prevent numerical issues during the alternation.

5.3 Numerical results

We apply our algorithm to learn semidefinite approximations of two non-SDP-representale functions, comparing the results with and without imposing compatibility (Ext). The first function we approximate is the ℓ_π norm $\|x\|_\pi = (\sum_i |x_i|^\pi)^{1/\pi}$, which is not SDP-representable because π is irrational. We view the ℓ_π norm

(a) Learning ℓ_π.
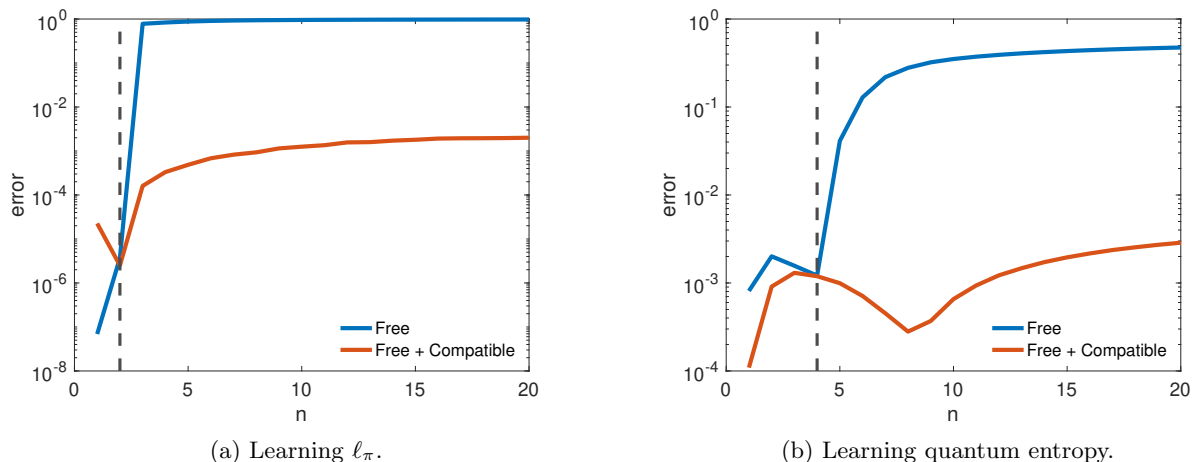
(b) Learning quantum entropy.

Figure 4: Errors for learning the ℓ_π norm and the quantum entropy variant (22). The dashed vertical lines denote the max n for which data is available.

as defined on the sequence $\mathscr{V} = \{\mathbb{R}^n\}$ with $G_n = \mathsf{B}_n$ from Example 2.2(a). It satisfies both intersection and projection compatibility. We choose description spaces $\mathscr{W} = \mathscr{U} = \{\mathrm{Sym}^2(\mathrm{Sym}^{\leq 1}\mathbb{R}^{2n+1}) = \mathbb{S}^{2n+2}\}$ as in Example 5.2(c), with the corresponding PSD cones $\{K_n = \mathbb{S}_+^{2n+2}\}$. We used 50 data points in $\mathbb{R}^n$ for $n \leq n_0 = 2$. We remark that $n_0 = 2$ exceeds the presentation degrees of $\mathscr{V}, \mathscr{W}, \mathscr{U}$, hence we can uniquely extend the learned two-dimensional set to a freely-described and compatible sequence when imposing (Ext). If we do not impose (DC) and search over all freely-described (possibly incompatible) sequences, the dimension of equivariant A_n, B_n only stabilizes from $n = 4$, the presentation degree of $\mathscr{W} \otimes \mathscr{U}$. As we only have two-dimensional data, the maps A_{n_0}, B_{n_0} we obtain from (Regress) do not uniquely extend to higher dimensions. In the experiments below, we find a basis for equivariant maps in dimension $n = 4$, and zero out the coefficients of all the basis elements on which (Regress) does not depend. This highlights another advantage of imposing compatibility–it allows us to uniquely identify a free description from lower-dimensional data.

The second sequence of functions we approximate is the nonnegative and positively-homogeneous variant of the quantum entropy given by

$$f_n(X) = \mathrm{Tr}\left[\left(X + \mathrm{Tr}(X)I\right) \log\left(X/\mathrm{Tr}(X) + I\right)\right] : \mathbb{S}_+^n \to \mathbb{R}, \tag{22}$$

defined on the sequence $\{\mathbb{S}^n\}$ with $G_n = \mathsf{O}_n$ from Example 2.2(b). Once again, the function (22) and the quantum entropy itself cannot be evaluated using semidefinite programming, though a family of semidefinite approximations for the quantum entropy is analytically derived in [68]. We aim to learn such an approximation entirely from evaluation data. To that end, we choose description spaces $\mathscr{W} = \{W_n = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 1}\mathbb{R}^n) = \mathbb{S}^{n+1}\}$ and $\mathscr{U} = \{U_n = \mathrm{Sym}^2(\mathrm{Sym}^{\leq 2}\mathbb{R}^n) = \mathbb{S}^{\binom{n+2}{2}}\}$, with corresponding PSD cones $\{K_n = \mathbb{S}_+^{\binom{n+2}{2}}\}$. Our data consists of 200 PSD matrices in $\mathbb{S}^n$ for $n \leq n_0 = 4$. Without a calculus for presentation degrees for $G_n = \mathsf{O}_n$, our theory does not guarantee the existence of an O_n-invariant extension of our learned description. Our theory does however guarantee a unique B_n-invariant extension, and in practice we observe that the extension is, in fact, O_n-invariant.

To approximate the above functions, we used (Regress) with 100 random initializations to fit the data in degree n_0. For the above two examples, not only is $f_{n_0}^{\mathrm{true}}$ positively-homogeneous and nonnegative, but also $f_{n_0}^{\mathrm{true}}(x) \neq 0$ for $x \neq 0$ in the domain. We therefore normalize the data x_i by $x_i \mapsto x_i/f_{n_0}^{\mathrm{true}}(x_i)$, so that $f_{n_0}^{\mathrm{true}}(x_i) = 1$ for all i and all points contribute equally to the objective of (Regress). We used the ℓ_2 norm for the cost function and $\lambda_{\min} = 10^{-3}$ in (Regress). To evaluate the results, we extended our learned descriptions to $n = 20$, sampled 10^3 unit-norm points (also PSD for the quantum entropy example) and computed the average normalized errors $\frac{|f_n(x) - f_n^{\mathrm{true}}(x)|}{f_n^{\mathrm{true}}(x)}$ in each n up to 20.

The errors are plotted in Figure 4, which shows that imposing compatibility conditions yields larger errors

35

in dimensions in which data is available, but ensures that the error increases gracefully when extending to higher dimensions. Since imposing compatibility conditions decreases the search space in (Regress), we expect the optimal solution of (Regress) with compatibility conditions to exhibit larger errors in dimensions in which data is available compared to the optimal freely-described (but possibly incompatible) solution. That is not the case in Figure 4(b), illustrating the nonconvexity of the fitting problem (Regress) and demonstrating another advantage of imposing compatibility–the resulting smaller parametric family allows our algorithm to better fit the data.

6 Constant-sized invariant conic programs

In the previous sections, we studied freely-described and compatible sequences of convex sets $\{C_n\}$ contained in a consistent sequence $\mathscr{V}$. As observed in Remark 2.17, these convex sets are often group-invariant because they are given by (ConicSeq) as projections of the intersection of invariant cones K_n with invariant affine spaces.

In this section, we further consider optimizing invariant linear functions over such sets. Such programs can be simplified by restricting their domain to invariant vectors [77, §3]. As we have seen in Proposition 2.7, when $\mathscr{V}$ is finitely-generated the dimensions of its spaces of invariants stabilize, so the size of the variables in such programs stabilizes as well. However, the size of the constraints may not stabilize, because the cones of invariants $\{K_n^{G_n}\}$ may grow in complexity. For example, if K_n is the cone of n-variate degree k polynomials that are nonnegative over all of $\mathbb{R}^n$ and $G_n = \mathsf{S}_n$, the best-known description of $K_n^{G_n}$ has complexity which is a (nonconstant) polynomial in n [78, 79, 80]. We therefore seek conditions for the existence of *constant-sized* descriptions for a family of convex cones $\{K_n^{G_n}\}$, and bounds on the value of n after which the size stabilizes, in the sense of the following definition.

Definition 6.1 (Constant-sized descriptions). *Let $\{U_n\}$ be a consistent sequence of $\{G_n\}$ representations and $\{K_n \subseteq U_n\}$ a sequence of convex cones. For $t \in \mathbb{N}$, we say that the sequence $\{K_n^{G_n} \subseteq U_n^{G_n}\}$ admits a constant-sized description for $n \geq t$ if there exists a single vector space U containing a cone K, linear maps $T_n : U \to U_n^{G_n}$, and subspaces $L_n \subseteq U$ such that $K_n^{G_n} = T_n(K \cap L_n)$ for all $n \geq t$.*

In other words, constant-sized descriptions express all cones in the sequence as *images* of linear slices of the same cone. If $\{K_n^{G_n}\}$ admit constant-sized descriptions for $n \geq t$ as in Definition 6.1, then optimization over convex sets given by (ConicSeq) can be rewritten as an optimization problem over an affine slice of the fixed cone K. In order to apply standard software to solve such a problem, we might express K in terms of standard cones such as the PSD and relative entropy cones, as these are among the most expressive families for which such software is available. Constant-sized descriptions for symmetric PSD and relative entropy cones have been obtained on a case-by-case basis in the literature [22, 23, 24, 25]. In this section, we explain how these results can be generalized and derived systematically from an interplay between representation stability and the structure of the cones in question.

6.1 Background: Uniform representation stability, permutation modules

Let $\{U_n\}$ be a finitely-generated consistent sequence of $\{G_n\}$-representations. We derive constant-sized descriptions for $\{K_n^{G_n} \subseteq U_n^{G_n}\}$ by finding constant-sized bases for $U_n^{G_n}$ in terms of which membership in $K_n^{G_n}$ is simply expressed. For PSD cones and their variants, such bases are obtained by viewing symmetric matrices as self-adjoint operators on a group representation, which can be block-diagonalized using the decomposition of their domains into irreducibles. For the relative entropy cones and their variants, such bases are obtained by viewing vectors as functions on a finite set on which the group acts, and considering indicators for orbits of that set. We proceed to state a few additional results from the representation stability literature pertaining to these two cases.

Proposition 2.7 shows that $\dim V_n^{G_n}$ stabilizes whenever $\{V_n\}$ is a finitely-generated consistent sequence of $\{G_n\}$-representations. In fact, the theory of [38, 39, 41] and others shows that for many standard families $\{G_n\}$ of groups, the entire decomposition of V_n into irreducibles stabilizes. This phenomenon was called *uniform representation stability* in [15], and implies that the sizes of the blocks in PSD matrices stabilize by Schur's lemma, a fact that we shall use below to obtain constant-sized descriptions for PSD cones. The following is a concrete instance of this phenomenon that we shall use below.

Theorem 6.2 ([38, Thm. 1.13],[39, Thm. 4.27]). *Let $\mathscr{V}_0 = \{\mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n be the consistent sequence from Example 2.2(a) and let $\mathscr{V} = \{V_n\}$ be a $\mathscr{V}_0$-module generated in degree d and presented in degree k. Then there exists a finite set Λ and integers $m_\lambda \in \mathbb{N}$, together with an assignment $\lambda \mapsto W_{\lambda[n]}$ of a distinct G_n-irreducible $W_{\lambda[n]}$ to each $\lambda \in \Lambda$ for $n \geq k+d$ such that $V_n \cong \bigoplus_{\lambda \in \Lambda} W_{\lambda[n]}^{m_\lambda}$ as G_n-representations.*

Proof. The irreducibles of the groups $\mathsf{S}_n, \mathsf{D}_n, \mathsf{B}_n$ are indexed as in [39, §2.1], and the consistent labelling of irreducibles for different n is given in [39, §2.2]. Under this labelling, the $\mathscr{V}_0$-module $\mathscr{V}$ is uniformly representation stable with stable range $n \geq k+d$ by [39, Thms. 4.4, 4.27], which precisely says that the set of irreducibles appearing in the decomposition of the V_n and their multiplicities become constant for $n \geq k+d$ by [15, Def. 2.6]. $\qquad\square$

Example 6.3. *Irreducibles of S_n are indexed by partitions of n. If $\lambda_1[n] = (n)$ is the trivial partition and $\lambda_2[n] = (n-1,1)$, then $\mathbb{R}^n = W_{\lambda_1[n]} \oplus W_{\lambda_2[n]}$ for all $n \geq 1$, where $W_{\lambda_1[n]} = \mathrm{span}\{\mathbb{1}_n\}$ and $W_{\lambda_2[n]} = \{x \in \mathbb{R}^n : \mathbb{1}_n^\top x = 0\}$ are distinct irreducible representations of S_n.*

Next, to study relative entropy cones and their variants, we introduce a class of particularly simple consistent sequences on which the group acts by permuting basis elements. If a group G acts on a (finite) set $\mathcal{A}$, define $\mathbb{R}^\mathcal{A} = \bigoplus_{\alpha \in \mathcal{A}} e_\alpha$ to be the ($|\mathcal{A}|$-dimensional) vector space spanned by orthonormal basis elements $\{e_\alpha\}_{\alpha \in \mathcal{A}}$, which is a G-representation with action $g \cdot e_\alpha = e_{g \cdot \alpha}$.

Definition 6.4 (Permutation modules). *Let $\mathscr{V} = \{V_n\}$ be a consistent sequence of $\{G_n\}$-representations. Let $\{\mathcal{A}_n \subseteq V_n\}$ be finite G_n-invariant sets satisfying $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ for all n. Then the* permutation $\mathscr{V}$-module *generated by the sets $\{\mathcal{A}_n\}$ is the $\mathscr{V}$-module $\{\mathbb{R}^{\mathcal{A}_n}\}_n$.*

Permutation modules can be analyzed in terms of the orbits in the sets $\mathcal{A}_n$. In particular, indicators of orbits form a basis for the space of invariants in a permutation module.

Proposition 6.5. *Let $\mathscr{V}$ be a consistent sequence of $\{G_n\}$-representations, let $\{\mathcal{A}_n \subseteq V_n\}$ be a nested sequence of finite group-invariant sets, and let $\mathscr{U} = \{\mathbb{R}^{\mathcal{A}_n}\}_n$ be the corresponding permutation $\mathscr{V}$-module.*

(a) *$\mathscr{U}$ is generated in degree d if and only if $\mathcal{A}_n = \bigcup_{g \in G_n} g\mathcal{A}_d$ for all $n \geq d$.*

(b) *The projections $\mathcal{P}_{U_n} : U_{n+1}^{G_{n+1}} \to U_n^{G_n}$ are isomorphisms for all $n \geq d$ if and only if (a) holds and the number of orbits in $\mathcal{A}_n$, which equals $\dim U_n^{G_n}$, is constant for all $n \geq d$.*

(c) *If $\mathscr{V} = \{\mathbb{R}^n\}$ and $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a), then $\mathscr{U}$ is an algebraically free $\mathscr{V}$-module.*

Proof. (a) We have $\mathbb{R}[G_n]V_d = \sum_{\substack{\alpha \in \mathcal{A}_d \\ g \in G_n}} \mathbb{R}e_{g\alpha} = \bigoplus_{\alpha \in \bigcup_{g \in G_n} g\mathcal{A}_d} \mathbb{R}e_\alpha$, which equals V_n if and only if $\bigcup_{g \in G_n} g\mathcal{A}_d = \mathcal{A}_n$.

(b) If $\hat{\mathcal{A}}_n \subseteq \mathcal{A}_n$ is a set of G_n-orbit representatives, then $U_n^{G_n} \cong \mathbb{R}^{\hat{\mathcal{A}}_n}$ has a basis consisting of $\mathbb{1}_\alpha^{(n)} = \sum_{g \in G_n / \mathrm{Stab}_{G_n}(\alpha)} e_{g\alpha}$ for $\alpha \in \hat{\mathcal{A}}_n$, where $\mathrm{Stab}_{G_n}(\alpha) = \{g \in G_n : g \cdot \alpha = \alpha\}$. Finally, $\mathcal{P}_{U_n} \mathbb{1}_\alpha^{(n+1)} = \mathbb{1}_\alpha^{(n)}$ for such α.

(c) Let $\hat{\mathcal{A}} \subseteq \mathcal{A}_d$ be a set of minimal degree G_d-orbit representatives. We argue that these are also the G_n-orbit representatives of $\mathcal{A}_n$ for all $n \geq d$. Indeed, since $\mathcal{A}_n = \bigcup_{g \in G_n} g\mathcal{A}_d = \bigcup_{g \in G_n} g\hat{\mathcal{A}}$, it suffices to show that distinct elements in $\hat{\mathcal{A}}$ lie in distinct G_n-orbits. To that end, observe that $\mathscr{V}$ satisfies the property

$$g \cdot \alpha \in V_d \text{ for } \alpha \in V_d, \ g \in G_n \implies \exists \tilde{g} \in G_d \text{ s.t. } g \cdot \alpha = \tilde{g} \cdot \alpha. \tag{23}$$

Indeed, define $\tilde{g}$ to act as g on the coordinates $\{i \in [d] : \alpha_i \neq 0\}$ and act trivially on all the others. Therefore, if $\alpha, \alpha' \in \mathcal{A}_d$ lie in the same G_n-orbit for $n > d$, then they also lie in the same G_d-orbit. This also shows that $\hat{\mathcal{A}} \cap \mathcal{A}_j$ is a set of minimal degree G_j-orbit representatives for each $j \leq d$, hence that $U_n = \bigoplus_{\alpha \in \hat{\mathcal{A}}} \mathbb{R}[G_n]e_\alpha$ for all n.

Next, we argue that $\mathscr{U}$ is an algebraically free $\mathscr{V}$-module. Observe that $\mathscr{V}$ satisfies the additional property

$$\alpha \in V_d \setminus V_{d-1} \text{ has min. degree in its } G_d\text{-orbit} \implies \mathrm{Stab}_{G_n}(\alpha) = \mathrm{Stab}_{G_d}(\alpha)H_{n,d}, \tag{24}$$

37

Indeed, if $\alpha \in V_d \setminus V_{d-1}$ has minimal degree then all its d entries are nonzero, hence any $g \in G_n$ fixing α must map the first d coordinates to themselves. Therefore, if $\alpha \in V_d \setminus V_{d-1}$ has minimal degree and $g_1, g_2, \ldots, g_M$ are coset representatives for $G_n / \mathrm{Stab}_{G_d}(\alpha) H_{n,d}$, then

$$\mathbb{R}[G_n] e_\alpha = \bigoplus_{m=1}^{M} \mathbb{R} e_{g_m \cdot \alpha} = \bigoplus_{m=1}^{M} g_m \cdot \mathbb{R} e_\alpha = \mathrm{Ind}_{\mathrm{Stab}_{G_d}(\alpha) H_{n,d}}^{G_n}(\mathbb{R} e_\alpha) = \mathrm{Ind}_{G_d H_{n,d}}^{G_n}\left(\mathrm{Ind}_{\mathrm{Stab}_{G_d}(\alpha)}^{G_d} \mathbb{R} e_\alpha \right),$$

where the first equality follows by (24), the second by the definition of a permutation representation (Definition 6.4), and the third equality follows by (Ind), and the last equality follows because $G_d / \mathrm{Stab}_{G_d}(\alpha) \cong (G_d H_{n,d})/(\mathrm{Stab}_{G_d}(\alpha) H_{n,d})$. Thus, if $\alpha \in \hat{\mathcal{A}}$ has degree d_α and we define $W_\alpha = \mathrm{Ind}_{\mathrm{Stab}_{G_{d_\alpha}}(\alpha)}^{G_{d_\alpha}} \mathbb{R} e_\alpha$, then $\mathscr{U} = \bigoplus_{\alpha \in \hat{\mathcal{A}}} \mathrm{Ind}_{G_{d_\alpha}}(W_\alpha)$ is indeed free. $\qquad \square$

Note that Proposition 6.5(c) applies to permutation $\mathscr{V}$-modules for any $\mathscr{V}$ satisfying (23) and (24). Finally, we shall need the following result, showing stabilization of invariants under stabilizing subgroups.

Proposition 6.6 ([39, Lemma 4.19]). *Let $\mathscr{V} = \{\mathbb{R}^n\}$ and $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a) and let $\mathscr{U}$ be a $\mathscr{V}$-module presented in degree k. If $\{H_{n,d}\}$ are the stabilizing subgroups of $\mathscr{V}$ and $\ell \in \mathbb{N}$, then the projections $\mathcal{P}_{U_n} \colon U_{n+1}^{H_{n+1,\ell}} \to U_n^{H_{n,\ell}}$ are isomorphisms for all $n \geq \ell + k$.*

Corollary 6.7. *Let $\mathscr{V} = \{\mathbb{R}^n\}$ and $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a) and let $\mathscr{U}$ be a $\mathscr{V}$-module presented in degree k. If $\beta \in \mathbb{R}^d \setminus \mathbb{R}^{d-1}$ has minimal degree in its G_d-orbit, then the projections $\mathcal{P}_{U_n} \colon U_{n+1}^{\mathrm{Stab}_{G_{n+1}}(\beta)} \to U_n^{\mathrm{Stab}_{G_n}(\beta)}$ are isomorphisms for all $n \geq d + k$.*

Proof. As shown in (24), we have $\mathrm{Stab}_{G_n}(\beta) = \mathrm{Stab}_{G_d}(\beta) H_{n,d}$ for all $n \geq d$, hence

$$U_n^{\mathrm{Stab}_{G_n}(\beta)} = U_n^{H_{n,d}} \cap U_n^{\mathrm{Stab}_{G_d}(\beta)}.$$

By Proposition 6.6, the projections $\mathcal{P}_{U_n} \colon U_{n+1}^{H_{n+1,d}} \to U_n^{H_{n,d}}$ are isomorphisms for all $n \geq d + k$. It thus suffices to show that if $u \in U_{n+1}^{H_{n+1,d}}$ satisfies $\mathcal{P}_{U_n} u \in U_n^{\mathrm{Stab}_{G_d}(\beta)}$, then $u \in U_{n+1}^{\mathrm{Stab}_{G_d}(\beta)}$. For any such u and $g \in \mathrm{Stab}_{G_d}(\beta)$ we have $g \cdot u \in U_{n+1}^{H_{n+1,d}}$ because $H_{n+1,d}$ and $\mathrm{Stab}_{G_d}(\beta) \subseteq G_d$ commute for our specific $\mathscr{V}$. As $\mathcal{P}_{U_n}(u - g \cdot u) = 0$ and $\mathcal{P}_{U_n}$ is injective on $U_{n+1}^{H_{n+1,d}}$, we get $u = g \cdot u$. $\qquad \square$

Proposition 6.6 is one of the main ingredients in the proof of Theorem 6.2. More generally, properties of *shifted* consistent sequences, which are sequences with group actions restricted to stabilizing subgroups, yield many of the phenomena in the representation stability literature, see [81] and references therein.

6.2 The PSD cone and its variants

We begin by giving constant-sized descriptions for certain sequences of PSD cones. We do so by deriving constant-sized bases for spaces of invariants in terms of which membership in the cones are simply expressed.

Theorem 6.8. *Let $\mathscr{V}_0 = \{\mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n$, or S_n as in Example 2.2(a), and let $\mathscr{V} = \{V_n\}$ be a $\mathscr{V}_0$-module generated in degree d and presented in degree k. Then the cones $\{\mathrm{Sym}_+^2(V_n)\}_n$ admit constant-sized descriptions for $n \geq k + d$.*

Proof. By Theorem 6.2, there exists a finite set Λ satisfying

$$V_n = \bigoplus_{\lambda \in \Lambda} \underbrace{W_{\lambda[n]}^{m_\lambda}}_{=:V_{\lambda[n]}}$$

where $W_{\lambda[n]}$ is a G_n-irreducible and $V_{\lambda[n]}$ is the corresponding isotypic component. Invariant elements of $U_n = \mathrm{Sym}^2(V_n)$ are equivariant and self-adjoint endomorphisms of V_n. If $X \in U_n^{G_n}$ is such an endomorphism and $\lambda \neq \mu \in \Lambda$ index distinct irreducibles, then $\mathcal{P}_{V_{\mu[n]}} X|_{V_{\lambda[n]}} = 0$ by Schur's Lemma [51, §1.2]. Because the irreducibles of $G_n = \mathsf{B}_n, \mathsf{D}_n$, and S_n are of real type [82] (meaning they remain irreducible when complexified), Schur's Lemma also implies that $\mathcal{P}_{W_{\lambda[n]}} X|_{W_{\lambda[n]}}$ is a multiple of the identity for each $\lambda \in \Lambda$, hence

$\mathcal{P}_{V_{\lambda[n]}} X|_{V_{\lambda[n]}} = X_\lambda \otimes I_{\dim W_{\lambda[n]}}$ for some $X_\lambda \in \mathbb{S}^{m_\lambda}$. We conclude that there exists an orthogonal matrix Q_n depending on the irreducible decomposition of V_n satisfying

$$U_n^{G_n} = \left\{ Q_n \bigoplus_{\lambda \in \Lambda} (X_\lambda \otimes I_{\dim W_{\lambda[n]}}) Q_n^* : X_\lambda \in \mathbb{S}^{m_\lambda} \right\}, \tag{25}$$

hence

$$\mathrm{Sym}_+^2(V_n)^{G_n} = \left\{ X \in U_n^{G_n} : X \succeq 0 \right\} = \left\{ Q_n \bigoplus_{\lambda \in \Lambda_n} (X_\lambda \otimes I_{\dim W_\lambda}) Q_n^* : X_\lambda \in \mathbb{S}_+^{m_\lambda} \right\}. \tag{26}$$

Thus, we obtain constant-sized descriptions by defining $U = \bigoplus_{\lambda \in \Lambda} \mathbb{S}^{m_\lambda}$ and $T_n \colon U \to U_n^{G_n}$ sending $(X_\lambda)_{\lambda \in \Lambda}$ to $Q_n \bigoplus_{\lambda \in \Lambda} (X_\lambda \otimes I_{\dim W_\lambda}) Q_n^*$, which maps $K = \bigoplus_{\lambda \in \Lambda} \mathbb{S}_+^{m_\lambda}$ onto $\mathrm{Sym}_+^2(V_n)^{G_n}$. $\qquad\square$

We now instantiate $\mathscr{V}$ to obtain more concrete corollaries.

Corollary 6.9. *If $G_n = \mathsf{S}_n, \mathsf{D}_n$ or B_n acts on $\mathbb{R}^n$ as in Example 2.2(a), then the cones $\mathrm{Sym}_+^2(\mathrm{Sym}^{\leq d}\mathbb{R}^n)^{G_n} \cong \left(\mathbb{S}_+^{\binom{n+k}{k}} \right)^{G_n}$ admit constant-sized descriptions by (26) for $n \geq 2d$.*

Proof. The sequence $\mathscr{V} = \{\mathrm{Sym}^{\leq d}\mathbb{R}^n\}$ is generated and presented in degree d by Theorem 4.11. $\qquad\square$

Remark 6.10. *The basis for $U_n^{G_n}$ obtained by mapping the standard basis for U via T_n in the proof of Theorem 6.8 is not freely-described because $\mathcal{P}_{V_n} Q_{n+1} \neq Q_n$ in general, as $\mathcal{P}_{V_n}$ does not map isotypic components $V_{\lambda[n+1]}$ to $V_{\lambda[n]}$. For example, if $\mathscr{V}$ is the sequence from Example 2.2(a), which decomposes into irreducibles as in Example 6.3, we have $\mathcal{P}_{V_n} W_{\lambda_2[n+1]} = \mathbb{R}^n$. In contrast, the basis that we shall use in Section 6.3 to give a constant-sized description of the relative entropy cone is freely-described. The choice of basis here is dictated by compatibility with the cone rather than by free descriptions.*

To obtain further corollaries, we consider equivariant images of cones.

Proposition 6.11. *Let $\mathscr{U} = \{U_n\}$ and $\mathscr{W} = \{W_n\}$ be consistent sequences of $\{G_n\}$-representations. Let $\{K_n \subseteq U_n\}$ be a sequence of convex cones such that K_n is G_n-invariant for each n. If $\{K_n^{G_n}\}$ admits constant-sized descriptions for $n \geq t$, then so does $\{\pi_n(K_n \cap L_n)^{G_n} \subseteq W_n\}$ for any sequence $\{\pi_n \in \mathcal{L}(U_n, W_n)^{G_n}\}$ and any sequence of G_n-invariant subspaces $L_n \subseteq U_n$.*

Proof. Suppose $K_n^{G_n} = T_n(K \cap L'_n)$ for $n \geq t$ where $T_n \colon U \to U_n^{G_n}$ and $L'_n \subseteq U$ are subspaces as in Definition 6.1. Because π_n is G_n-equivariant,

$$\pi_n(K_n \cap L_n)^{G_n} = \pi_n(K_n^{G_n} \cap L_n^{G_n}) = (\pi_n \circ T_n)(K \cap L'_n \cap T_n^{-1}(L_n^{G_n})).$$

Noting that $L'_n \cap T_n^{-1}(L_n^{G_n})$ is a subspace of U, we get the desired constant-sized descriptions. $\qquad\square$

We now show that various cones of symmetric sums of squares polynomials admit constant-sized descriptions, generalizing a number of results in the literature. For the following theorem, a polynomial f is a sum of squares modulo an ideal $\mathcal{I}$ if there exist finitely-many polynomials g_j such that $f - \sum_j g_j^2 \in \mathcal{I}$.

Theorem 6.12. *Let $G_n = \mathsf{B}_n, \mathsf{D}_n$ or S_n, with their standard action on $\mathbb{R}^n$. Let*

$$\mathcal{I}_n \subseteq \bigoplus_{d \geq 0} \mathrm{Sym}^d \left(\bigwedge^k \mathbb{R}^n \right) \cong \mathbb{R}[x_{i_1,\dots,i_k}]_{1 \leq i_1 < \cdots < i_k \leq n},$$

be a G_n-invariant ideal, and let $U_n = \mathrm{Sym}^{\leq 2d}(\bigwedge^k \mathbb{R}^n)/\mathcal{I}_n$ together with the sums-of-squares cone $\mathrm{SOS}_{U_n} = \{f \in U_n : f \text{ is a sum of squares mod } \mathcal{I}_n\}$. Then the sequence $\{\mathrm{SOS}_{U_n}^{G_n}\}$ admits a constant-sized description for $n \geq 2kd$.

39

Proof. If $v(x)$ is the vector whose coordinates are all the monomials in the $\binom{n}{k}$ variables $x_{i_1,\dots,i_k}$ of degree at most d, then [54, Thm. 3.39] yields

$$\text{SOS}_{U_n} = \pi_n\left(\text{Sym}^2_+\left(\text{Sym}^{\leq d}\left(\bigwedge^k \mathbb{R}^n\right)\right)\right), \quad \pi_n(M) = v(x)^\top M v(x) + \mathcal{I}_n.$$

The map $\pi_n \colon \text{Sym}^2\left(\text{Sym}^{\leq d}\left(\bigwedge^k \mathbb{R}^n\right)\right) \to U_n$ is equivariant by definition of the action of G_n and the invariance of $\mathcal{I}_n$. If $\mathscr{V}_0 = \{\mathbb{R}^n\}$ as in Example 2.2(a), then $\mathscr{V} = \text{Sym}^{\leq d}\left(\bigwedge^k \mathscr{V}_0\right)$ is a $\mathscr{V}_0$-module generated and presented in degree kd by Theorem 4.11. Thus, the result follows from Theorem 6.8 and Proposition 6.11. $\square$

If $k = 1$, $\mathcal{I}_n = (0)$, and $G_n = \mathsf{S}_n$ then we recover [22, Thms. 4.7, 4.10], and when $G_n = \mathsf{D}_n$ or B_n we recover [24, Cor. 3.23] with improved range of n for which the cones admit constant-sized descriptions for $G_n = \mathsf{D}_n$ from $n > 2d$ to $n \geq 2d$. If $k \geq 2$, if $\mathcal{I}_n = (x_I - x_I^2)_{I \subseteq \binom{[n]}{k}}$ is the ideal generated by $x_I^2 - x_I$ where I ranges over all k-subsets of $[n]$, and if $G_n = \mathsf{S}_n$, we recover [23, Thm. 2.4]. Theorem 6.12 generalizes all of these results to include any of the classical Weyl groups and any sequence of invariant ideals.

An application of Theorem 6.12 is obtaining constant-sized SDPs to certify graph homomorphism density inequalities [23, 83]. Many problems in extremal combinatorics can be recast as proving polynomial inequalities between homomorphism densities of graphs, which is the fraction of maps between the vertex sets of two graphs that define graph homomorphisms. A simple example is Mantel's theorem, which states that the maximum number of edges in a triangle-free graph is $\lfloor n^2/4 \rfloor$. Razborov proposed a method of certifying such inequalities using *flag algebras* [84], which were shown in [23, 83] to be sums-of-squares certificates of certain symmetric polynomial inequalities. Razborov's flags are interpreted there as "free" spanning sets for spaces of symmetric polynomials. Formally, they are freely-described elements in the sense of Definition 2.12. Our development above shows that both the existence of such freely-described spanning sets and the resulting constant-sized SDPs are consequences of representation stability.

6.3 The relative entropy cone and its variants

Let $\mathscr{V}_0$ be a consistent sequence of $\{G_n\}$ representations and $\mathscr{V} = \{\mathbb{R}^{\mathcal{A}_n}\}$ be a permutation $\mathscr{V}_0$-module. Let $\mathscr{U} = \mathscr{V}^{\oplus 2} \oplus \mathbb{R}$, and define the relative entropy cone

$$\text{RE}_{\mathcal{A}_n} = \{(\nu, c, t) \in \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R} : \nu, c \geq 0, \ D(\nu, c) \leq t\}, \tag{27}$$

where $D(\nu, c) = \sum_{\alpha \in \mathcal{A}_n} \nu_\alpha \log\left(\frac{\nu_\alpha}{c_\alpha}\right)$ is the relative entropy (see Section 3.4).

Proposition 6.13. *Let $\mathscr{V} = \{\mathbb{R}^{\mathcal{A}_n}\}$ be a permutation $\mathscr{V}_0$-module such that $\dim V_n^{G_n}$ is constant for all $n \geq d$. Then the cones $\{\text{RE}_{\mathcal{A}_n}^{G_n}\}$ given by (27) admit constant-sized descriptions for $n \geq d$.*

Proof. Let $M = \dim(\mathbb{R}^{\mathcal{A}_d})^{G_d}$ and fix $n \geq d$. Let $\{\alpha_j\}_{j \in [M]} \subseteq \mathcal{A}_n$ be a set of G_n-orbit representatives, and let $\mathbb{1}_{j,n} = \sum_{g \in G_n/\text{Stab}_{G_n}(\alpha_j)} e_{g\alpha_j}$ for each $j \in [M]$, so that $\{\mathbb{1}_{j,n}\}_{j \in [M]}$ is a basis for $(\mathbb{R}^{\mathcal{A}_n})^{G_n} \cong \mathbb{R}^M$. Then a basis for $U_n^{G_n}$ consists of $\{(\mathbb{1}_{j,n}, 0, 0), (0, \mathbb{1}_{j,n}, 0)\}_{j \in [M]} \cup \{(0, 0, 1)\}$.

If $(\nu, c, t) \in U_n^{G_n}$ for $n \geq d$ is expanded as $\nu = \sum_{j=1}^M \hat{\nu}_j \mathbb{1}_{j,n}$ and similarly for c, then

$$\text{RE}_{\mathcal{A}_n}^{G_n} = \left\{(\nu, c, t) \in U_n^{G_n} : \hat{\nu}, \hat{c} \geq 0, \ \sum_{j=1}^M |G_n/\text{Stab}_{G_n}(\alpha_j)| \hat{\nu}_j \log\left(\frac{\hat{\nu}_j}{\hat{c}_j}\right) \leq t\right\}.$$

Let $U = \mathbb{R}^L \oplus \mathbb{R}^L \oplus \mathbb{R}$, define

$$K = \left\{(\hat{\nu}, \hat{c}, t) \in \mathbb{R}^L \oplus \mathbb{R}^L \oplus \mathbb{R} : \hat{\nu}, \hat{c} \geq 0, \ \sum_{j=1}^M \hat{\nu}_j \log\left(\frac{\hat{\nu}_j}{\hat{c}_j}\right) \leq t\right\},$$

and define $T_n \colon U \to U_n^{G_n}$ sending $(e_j, 0, 0) \mapsto |G_n/\text{Stab}_{G_n}(\alpha_j)|^{-1}(\mathbb{1}_{j,n}, 0, 0)$, sending $(0, e_j, 0) \mapsto |G_n/\text{Stab}_{G_n}(\alpha_j)|^{-1}(0, \mathbb{1}_{j,n}, 0)$ and sending $(0, 0, 1) \mapsto (0, 0, 1)$. Then $\text{RE}_{\mathcal{A}_n}^{G_n} = T_n(K)$ for all $n \geq d$, giving the desired constant-sized descriptions. $\square$

Now suppose $\mathscr{W} = \{\mathbb{R}^{\mathcal{B}_n}\}$ is another permutation $\mathscr{V}_0$-module. Let $\widetilde{\mathscr{U}} = \mathscr{W} \otimes \mathscr{U} = \{\widetilde{U}_n = \mathcal{L}(\mathbb{R}^{\mathcal{B}_n}, U_n)\}$ and consider the cones of maps

$$\mathrm{REM}_{\mathcal{A}_n, \mathcal{B}_n} = \left\{ M \in \widetilde{U}_n : M(\mathbb{R}^{\mathcal{B}_n}_+) \subseteq \mathrm{RE}_{\mathcal{A}_n} \right\} = \left\{ M \in \widetilde{U}_n : Me_\beta \in \mathrm{RE}_{\mathcal{A}_n} \text{ for all } \beta \in \mathcal{B}_n \right\}. \qquad (28)$$

We obtain constant-sized descriptions for these cones for a specific $\mathscr{V}_0$.

Proposition 6.14. *Suppose $\mathscr{V}_0 = \{\mathbb{R}^n\}$ with $G_n = \mathsf{B}_n, \mathsf{D}_n,$ or S_n as in Example 2.2(a) and that $\{\mathbb{R}^{\mathcal{A}_n}\}, \{\mathbb{R}^{\mathcal{B}_n}\}$ are permutation $\mathscr{V}_0$-modules generated in degrees d_V, d_W, respectively. Then the sequence of cones $\{\mathrm{REM}^{G_n}_{\mathcal{A}_n, \mathcal{B}_n}\}$ in (28) admits constant-sized descriptions for $n \geq d_V + d_W$.*

Proof. Let $\hat{\mathcal{B}} \subseteq \mathcal{B}_{d_W}$ be a set of minimal-degree orbit representatives, which are also orbit representatives for $\mathcal{B}_n$ for all $n \geq d_W$ by Proposition 6.5(c). Any $M \in \widetilde{U}^{G_n}_n = \mathcal{L}(\mathbb{R}^{\mathcal{B}_n}, U_n)^{G_n}$ for $n \geq d_W$ is fully determined by the images $Me_\beta \in U^{\mathrm{Stab}_{G_n}(\beta)}_n$ of the basis elements e_β for $\beta \in \hat{\mathcal{B}}$ and conversely, for any collection $\left\{ u_\beta \in U^{\mathrm{Stab}_{G_n}(\beta)}_n \right\}_{\beta \in \hat{\mathcal{B}}}$ there is a unique $M \in \widetilde{U}^{G_n}_n$ satisfying $Me_\beta = u_\beta$. Moreover, $M \in \widetilde{K}_n$ if and only if $Me_\beta \in K^{\mathrm{Stab}_{G_n}(\beta)}_n$ for all $\beta \in \hat{\mathcal{B}}$. Thus, we have

$$\widetilde{U}^{G_n}_n = \bigoplus_{\beta \in \hat{\mathcal{B}}} U^{\mathrm{Stab}_{G_n}(\beta)}_n, \qquad \mathrm{REM}^{G_n}_{\mathcal{A}_n, \mathcal{B}_n} = \bigoplus_{\beta \in \hat{\mathcal{B}}} \mathrm{RE}^{\mathrm{Stab}_{G_n}(\beta)}_{\mathcal{A}_n}.$$

By Corollary 6.7, the projections $\mathcal{P}_{V_n} : (\mathbb{R}^{\mathcal{A}_{n+1}})^{\mathrm{Stab}_{G_{n+1}}(\beta)} \to (\mathbb{R}^{\mathcal{A}_n})^{\mathrm{Stab}_{G_n}(\beta)}$ are isomorphisms for all $n \geq d_V + d_W$. Proposition 6.13 then gives constant-sized descriptions for $\left\{ \mathrm{RE}^{\mathrm{Stab}_{G_n}(\beta)}_{\mathcal{A}_n} \right\}_n$ for each $\beta \in \hat{\mathcal{B}}$. $\square$

As an application of Proposition 6.14, we obtain constant-sized descriptions for SAGE cones of signomials. Indeed, if $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathbb{R}^n$ as in the above proposition, define the sequence $\mathscr{F} = \mathscr{V} \oplus \mathscr{W} = \{F_n\}$, viewed as spaces of functions on $\{\mathbb{R}^n\}$,

$$F_n = \left\{ f(x) = \sum_{\alpha \in \mathcal{A}_n} c_\alpha e^{\langle \alpha, x \rangle} + \sum_{\beta \in \mathcal{B}_n} t_\beta e^{\langle \beta, x \rangle} : c_\alpha, t_\beta \in \mathbb{R} \right\} \cong \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}^{\mathcal{B}_n},$$

with G_n acting by $g \cdot f = f \circ g^{-1}$. Sums of exponentials as in F_n are called signomials, and their optimization has a number of applications [85]. As usual, minimizing a signomial f over $\mathbb{R}^n$ can be recast as maximizing $\gamma \in \mathbb{R}$ such that $f - \gamma \geq 0$ on $\mathbb{R}^n$, hence optimizing signomials can be reduced to certifying their nonnegativity. This is NP-hard in general, but can be done efficiently if only a single coefficient of f is nonnegative, or if f is a sum of such signomials [86]. Formally, define the cones of (Sums of) nonnegative AM/GM Exponential functions, called AGE (resp., SAGE) functions, by

$$\mathrm{AGE}_{\mathcal{A}_n, \beta} = \left\{ f(x) = \sum_{\alpha \in \mathcal{A}_n} c_\alpha e^{\langle \alpha, x \rangle} + t e^{\langle \beta, x \rangle} : f \geq 0 \text{ on } \mathbb{R}^n \text{ and } c_\alpha \geq 0 \text{ for all } \alpha \in \mathcal{A}_n \right\},$$

$$\mathrm{SAGE}_{\mathcal{A}_n, \mathcal{B}_n} = \sum_{\beta \in \mathcal{B}_n} \mathrm{AGE}_{\mathcal{A}_n, \beta}.$$

Theorem 6.15. *Suppose $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathbb{R}^n$ where $\mathbb{R}^n$ is embedded in $\mathbb{R}^{n+1}$ by zero-padding and with the standard action of $G_n = \mathsf{S}_n, \mathsf{D}_n$ or B_n. If $\mathcal{A}_n = \bigcup_{g \in G_n} g\mathcal{A}_{d_A}$ for all $n \geq d_A$ and $\mathcal{B}_n = \bigcup_{g \in G_n} g\mathcal{B}_{d_B}$ for all $n \geq d_B$, then the invariant SAGE cones $\mathrm{SAGE}^{G_n}_{\mathcal{A}_n, \mathcal{B}_n}$ admit constant-sized descriptions for $n \geq d_A + d_B$.*

Proof. Identify $M \in \widetilde{U}_n$ with tuples $(Me_\beta)_{\beta \in \mathcal{B}_n} = (\nu^{(\beta)}, c^{(\beta)}, t_\beta)_{\beta \in \mathcal{B}_n} \in \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}$ for each $\beta \in \mathcal{B}_n$. The authors of [86] show that, in our notation,

$$\mathrm{SAGE}_{\mathcal{A}_n, \mathcal{B}_n} = \left\{ (c, t) \in \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}^{\mathcal{B}_n} : \exists M = (\nu^{(\beta)}, c^{(\beta)}, t_\beta)_{\beta \in \mathcal{B}_n} \in \mathrm{REM}_{\mathcal{A}_n, \mathcal{B}_n} \text{ s.t. } \sum_{\beta \in \mathcal{B}_n} c^{(\beta)} = c, \right.$$

$$\left. \sum_{\alpha \in \mathcal{A}_n} \nu^{(\beta)}_\alpha (\alpha - \beta) = 0 \text{ for all } \beta \in \mathcal{B}_n \right\}$$

$$= \pi_n(\mathrm{REM}_{\mathcal{A}_n, \mathcal{B}_n} \cap \mathcal{L}_n),$$

where

$$\mathcal{L}_n = \left\{ M = (\nu^{(\beta)}, c^{(\beta)}, t_\beta)_{\beta \in \mathcal{B}_n} \in \widetilde{U}_n : \sum_{\alpha \in \mathcal{A}_n} \nu_\alpha^{(\beta)}(\alpha - \beta) = 0 \text{ for all } \beta \in \mathcal{B}_n \right\},$$

$$\pi_n(M) = \left(\sum_{\beta \in \mathcal{B}_n} c^{(\beta)}, (t_\beta)_{\beta \in \mathcal{B}_n} \right) \in \mathbb{R}^{\mathcal{A}_n} \oplus \mathbb{R}^{\mathcal{B}_n}.$$

Note that π_n is equivariant, since

$$(g \cdot M)e_\beta = gMg^{-1}e_\beta = gMe_{g^{-1}\beta} = g \cdot (\nu^{(g^{-1}\beta)}, c^{(g^{-1}\beta)}, t_{g^{-1}\beta}) = (g \cdot \nu^{(g^{-1}\beta)}, g \cdot c^{(g^{-1}\beta)}, t_{g^{-1}\beta}),$$

hence

$$\pi_n(g \cdot M) = \left(\sum_{\beta \in \mathcal{B}_n} g \cdot c^{(g^{-1}\beta)}, (t_{g^{-1}\beta})_{\beta \in \mathcal{B}_n} \right) = \left(g \cdot \sum_{\beta \in \mathcal{B}_n} c^{(\beta)}, g \cdot (t_\beta)_{\beta \in \mathcal{B}_n} \right) = g \cdot \pi_n(M).$$

Similarly, if $\sum_{\alpha \in \mathcal{A}_n} \nu_\alpha^{(\beta)}(\alpha - \beta) = 0$ for all $\beta \in \mathcal{B}_n$ then

$$\sum_{\alpha \in \mathcal{A}_n} (g \cdot \nu^{(g^{-1}\beta)})_\alpha (\alpha - \beta) = g \sum_{\alpha \in \mathcal{A}_n} \nu_{g^{-1}\alpha}^{(g^{-1}\beta)}(g^{-1}\alpha - g^{-1}\beta) = g \sum_{\alpha \in \mathcal{A}_n} \nu_\alpha^{(g^{-1}\beta)}(\alpha - g^{-1}\beta) = 0,$$

hence $\mathcal{L}_n$ is G_n-invariant. Thus, the result follows from Proposition 6.14 and Proposition 6.11. $\qquad\square$

Theorem 6.15 generalizes [25, Thm. 5.3] beyond S_n to the other classical Weyl groups D_n and B_n. It would be interesting to further generalize to signomials defined on more general consistent sequences than $\{\mathbb{R}^n\}$, which would require generalizing the description of the AGE cone from [86].

7 Limits

In this section, we consider infinite-dimensional limits of consistent sequences and of their convex subsets. In particular, we show that free descriptions certifying compatibility conditions as in Theorem 2.23 naturally extend to descriptions of limiting convex sets. To do so, we begin by reviewing a limiting approach to representation stability.

7.1 Background: A limiting approach to representation stability

We define limits of consistent sequences, and interpret our definitions from Section 2.2 in terms of these limits.

Definition 7.1. *For a consistent sequence $\mathscr{V} = \{V_n\}$ of $\{G_n\}$-representations, define its* limiting representation *as the vector space $V_\infty = \bigcup_n V_n$, viewed as a representation of $G_\infty = \bigcup_n G_n$.*

There is an approach to representation stability studying limiting representations of limiting groups as above, instead of representations of categories as in Appendix A. For example, the authors of [41] analyze representations of five standard infinite groups, including $\mathsf{O}_\infty, \mathsf{S}_\infty$ defined above, that occur as quotients or subrepresentations of tensor powers of $\mathbb{R}^\infty$ and its dual.

Example 7.2. *We give the limits of some sequences encountered above.*

(a) *If $\mathscr{V} = \{\mathbb{R}^n\}$ as in Example 2.2(a), then*

$$V_\infty = \mathbb{R}^\infty = \{x \in \mathbb{R}^{\mathbb{N}} : \exists N \in \mathbb{N} \ s.t. \ x_n = 0 \ for \ all \ n \geq N\}.$$

This is viewed as a representation of $\mathsf{O}_\infty = \bigcup_{n \in \mathbb{N}} \mathsf{O}_n$ or the similarly-defined $\mathsf{B}_\infty, \mathsf{D}_\infty, \mathsf{S}_\infty$, viewed as $\mathbb{N} \times \mathbb{N}$ matrices differing from the identity in finitely-many entries.

(b) In Example 2.2(b), we have $V_\infty = \mathbb{S}^\infty$, the set of $\mathbb{N} \times \mathbb{N}$ symmetric matrices with finitely-many nonzero entries, with O_∞ or one of its subgroups in (a) acting by conjugation.

(c) For the consistent sequence used in Section 3.2, we have

$$V_\infty = \left\{ f \in L^2([0,1]) : \exists n \in \mathbb{N} \text{ s.t. } f \text{ is constant on } [(i-1)/m2^n, i/m2^n) \text{ for each } i \in [m2^n] \right\}.$$

Here $G_\infty = \bigcup_n \mathsf{S}_{m2^n}$ acts on $[0,1]$ by permuting intervals of the above form, and acts on $f \in L^2([0,1])$ by $g \cdot f = f \circ g^{-1}$. The limit in Example 2.2(c) is the special case $m = 1$ above, with the action of $G_\infty = \bigcup_n \mathsf{C}_{2^n}$ consisting of cyclic shifts of $[0,1]$ by $i/2^n$ for some $i \in \{0, \ldots, 2^n - 1\}$ and $n \in \mathbb{N}$.

(d) For the graphon sequence in Section 3.6, we have

$$V_\infty = \Big\{ W \in L^2([0,1]^2) : W(x,y) = W(y,x), \ \exists n \in \mathbb{N} \text{ s.t. } W \text{ is constant on}$$

$$[(k-1)/2^n, k/2^n) \times [(\ell-1)/2^n, \ell/2^n) \text{ for each } k, \ell \in [2^n] \Big\},$$

and $G_\infty = \bigcup_n \mathsf{S}_{2^n}$ as in (c). Here G_∞ acts on V_∞ by $\sigma \cdot W = W \circ (\sigma^{-1}, \sigma^{-1})$.

Given two consistent sequences $\{V_n\}, \{U_n\}$ of $\{G_n\}$-representations, a sequence of equivariant linear maps $\{A_n \in \mathcal{L}(V_n, U_n)^{G_n}\}$ *extends to the limit,* meaning there exists $A_\infty \in \mathcal{L}(V_\infty, U_\infty)^{G_\infty}$ satisfying $A_\infty|_{V_n} = A_n$ for all n, if and only if $\{A_n\}$ is a morphism of sequences. Similarly, a sequence of invariants $\{v_n \in V_n^{G_n}\}$ defines a sequence of invariant linear functionals $\ell_n(x) = \langle v_n, x \rangle \colon V_n \to \mathbb{R}$, and these linear functionals extend to a G_∞-invariant linear functional on V_∞ if and only if $\{v_n\}$ is a freely-described element. Every invariant functional on V_∞ arises in this way, so freely-described elements are in one-to-one correspondence with invariant linear functionals on the limit of a consistent sequence.

We also consider continuous limits of consistent sequences. Because we require the embeddings in a consistent sequence $\mathcal{V} = \{V_n\}$ to be isometries (Definition 2.1(b)), the inner products on the V_n extend to the limit V_∞. We can then take the completion with respect to the inner product topology to get a Hilbert space $\overline{V_\infty}$ on which G_∞ acts unitarily. It is then natural to consider sequences of linear maps which extend to *continuous* maps between these limits.

Definition 7.3 (Continuous limits). *Let $\mathcal{V} = \{V_n\}$ be a consistent sequence of $\{G_n\}$-representations and let $\overline{V_\infty}$ be the Hilbert space completion with respect to the inner product topology. We call $\overline{V_\infty}$ with its G_∞-action the* continuous limit *of the sequence V.*

A sequence of maps $\{A_n \colon V_n \to U_n\}$ extends continuously to the limit *if there exists a bounded linear operator $\overline{A_\infty} \colon \overline{V_\infty} \to \overline{U_\infty}$ such that $\overline{A_\infty}|_{V_n} = A_n$ for all n. A freely-described element $\{u_n \in U_n^{G_n}\}$* extends continuously to the limit *if there exists $u_\infty \in \overline{U_\infty}$ satisfying $\mathcal{P}_{U_n} u_\infty = u_n$ for all n.*

Because $\overline{U_\infty}$ is a Hilbert space, such u_∞ exists if and only if the linear functionals corresponding to $\{u_n\}$ extend continuously to the limit. Note that a morphism of sequences $\{A_n \colon V_n \to U_n\}$ extends to the continuous limit if and only if the sequence of operator norms $\{\|A_n\|_{\mathrm{op}}\}$ with respect to the norms on V_n and U_n is bounded. Similarly, a freely-described element $\{u_n\}$ extends continuously to the limit if and only if the sequence of norms $\{\|u_n\|\}_n$ is bounded.

Example 7.4. (a) *In the setting of Example 7.2(a), we have $\overline{\mathbb{R}^\infty} = \ell_2(\mathbb{R})$. The only sequences of maps extending continuously to the limit are $\{A_n = \alpha I_n\}$ for $\alpha \in \mathbb{R}$.*

(b) *In Example 7.2(b), we have $\overline{\mathbb{S}^\infty} = \left\{ X \in \mathbb{S}^\mathbb{N} : \sum_{i,j=1}^\infty X_{i,j}^2 < \infty \right\}$. The only sequences of maps $\mathbb{R}^n \to \mathbb{S}^n$ extending continuously to the limit are $\{A_n = \alpha \mathrm{diag}\}$ for $\alpha \in \mathbb{R}$, and similarly for $\mathbb{S}^n \to \mathbb{R}^n$. Note that the sequence $\{A_n(X) = \mathrm{diag}(X \mathbb{1}_n)\}$ of maps $\mathbb{S}^n \to \mathbb{R}^n$ extends to the limit, but not continuously.*

(c) *In Example 7.2(d), we have $\overline{V_\infty} = \{ W \in L^2([0,1]^2) : W(x,y) = W(y,x) \text{ for a.e. } (x,y) \in [0,1]^2 \}$, which is known as the space of L^2-graphons [87, Def. 2.3]. The sequences of maps $\{A_n \in \mathcal{L}(V_n)^{G_n}\}$ that*

extend continuously to this limit are linear combinations of the three maps $A_n^{(1)}(X) = X$, $A_n^{(2)}(X) = (\mathbb{1}^\top X \mathbb{1}) \mathbb{1} \mathbb{1}^\top$, *and* $A_n^{(3)}(X) = \frac{1}{n}\left(X \mathbb{1} \mathbb{1}^\top + \mathbb{1} \mathbb{1}^\top X\right)$, *which can be written in terms of graphons as*

$$A_n^{(1)}(W)(x,y) = W(x,y), \quad A_n^{(2)}(W)(x,y) = \int_{[0,1]^2} W(s,t)\, \mathrm{d}s \mathrm{d}t,$$

$$A_n^{(3)}(W)(x,y) = \int_{[0,1]} [W(x,s) + W(s,y)]\, \mathrm{d}s.$$

There are extensions of Definition 7.3 that we do not pursue in this paper. For example, one can complete V_∞ with respect to different G_∞-invariant metrics. For example, the map $A_\infty(X) = \mathrm{diag}(X\mathbb{1}_\infty)$ from $\mathbb{S}^\infty$ to itself and the trace $\mathbb{S}^\infty \to \mathbb{R}$ do no extend continuously with respect to the ℓ_2-norm above, but are continuous with respect to the ℓ_1-norm.

7.2 Limiting conic descriptions

The goal of this section is give conditions under which free descriptions extend to descriptions of continuous limits of convex sets. Certificates of compatibility as in Theorem 2.23 play a major role once again. If $\mathscr{C} = \{C_n \subseteq V_n\}$ is an intersection-compatible sequence of convex subsets of a consistent sequence $\{V_n\}$, define $C_\infty = \bigcup_n C_n$, which is a convex subset of V_∞.

Theorem 7.5 (Descriptions of limits). *Suppose $\mathscr{C} = \{C_n \subseteq V_n\}$ is given by (ConicSeq). If all the hypotheses of Theorem 2.23(a) are satisfied (so that $\mathscr{C}$ is certifiably intersection and projection compatible) and if $\{A_n\}, \{B_n\}, \{u_n\}$ extend continuously to the limit, then*

$$\left\{ x \in \overline{V_\infty} : \exists y \in \overline{W_\infty} \text{ s.t. } \overline{A_\infty} x + \overline{B_\infty} y + u_\infty \in \overline{K_\infty} \right\}, \tag{29}$$

contains C_∞ and is dense in its closure. If $B_\infty = 0$, then (29) equals $\overline{C_\infty}$.

Proof. To prove that C_∞ is contained in (29), observe that $u_\infty - u_n = \lim_{N\to\infty}(u_N - u_n) \in \overline{K_\infty}$ for all n. Therefore, if $x \in C_n$ and $y \in W_n$ satisfy $A_n x + B_n y + u_n \in K_n$, then $\overline{A_\infty} x + \overline{B_\infty} y + u_\infty = A_n x + B_n y + u_n + (u_\infty - u_n) \in \overline{K_\infty}$, proving that x is in (29).

To prove that (29) is contained in $\overline{C_\infty}$, suppose $x \in V_\infty$ and $y \in \overline{W_\infty}$ satisfy $\overline{A_\infty} x + \overline{B_\infty} y + u_\infty \in \overline{K_\infty}$. Because $\{A_n^*\}, \{B_n^*\}$ are morphisms and $\{K_n\}$ is projection-compatible, applying $\mathcal{P}_{U_n}$ we obtain $A_n(\mathcal{P}_{V_n} x) + B_n(\mathcal{P}_{W_n} y) + u_n \in K_n$, hence $\mathcal{P}_{V_n} x \in C_n$ for all n and $x = \lim_n \mathcal{P}_{V_n} x \in \overline{C_\infty}$.

If $B_\infty = 0$ then (29) is the preimage under the continuous map $x \mapsto \overline{A_\infty} x + u_\infty$ of the closed cone $\overline{K_\infty}$, hence (29) is closed and must equal $\overline{C_\infty}$. $\qquad\square$

If (29) is dense in $\overline{C_\infty}$, then optimizing a continuous function over (29) and over $\overline{C_\infty}$ are equivalent.

Example 7.6. *The following are simple examples of limiting sets $\overline{C_\infty}$.*

(a) *For the sequences $\mathscr{V}, \mathscr{W}\mathscr{U}$ used to describe the permutahedron in Section 3.2, we have $\overline{V_\infty} = L^2([0,1])$ since functions constant on intervals $[(i-1)/m2^n, i/m2^n]$ are dense in $L^2[0,1]$, and $\overline{W_\infty} = \overline{V_\infty}^{\oplus q}$, $\overline{U_\infty} = \overline{W_\infty} \oplus \overline{V_\infty}^{\oplus 2} \oplus \mathbb{R}^q$. The maps $\{A_n\}, \{B_n\}$ given there extend continuously to the limit since $\|A_n\| = 1$ and $\|B_n\| \leq \sqrt{2 + q + \|\lambda\|_2^2}$ for all n. Also, $u_n = u_{n+1}$ for all n so $\{u_n\}$ extends continuously to the limit. Thus, Theorem 7.5 implies that the following is a dense subset of $\overline{\mathrm{Perm}(\lambda)_\infty}$:*

$$\left\{ \sum_{i=1}^q \lambda_i f_i \in L^2([0,1]) : f_1, \ldots, f_q \geq 0, \ \sum_{i=1}^q f_i = 1, \ \int_{[0,1]} f_i = \frac{m_i}{m} \text{ for each } i \in [q] \right\}.$$

Similarly, Theorem 7.5 applies to the description of the Schur-Horn orbitopes in Section 3.2. The Hilbert spaces one obtains in this case appear in the construction of the hyperfinite II_1 factor [88, §1.6] in the theory of operator algebras.

(b) Theorem 7.5 yields the following parametric family of closed convex subsets of the space of L^2 graphons (see Section 3.6 and Example 7.2(b)):

$$\overline{C_\infty} = \left\{ W \in L^2([0,1]^2) \ symmetric : \right.$$

$$\left. 0 \preceq K(x,y) = \alpha_1 \int_{[0,1]^2} W(t,s)\, \mathrm{d}t\mathrm{d}s + \alpha_2 \int_{[0,1]} [W(x,t) + W(t,y)]\, \mathrm{d}t + \alpha_3 W(x,y) + \alpha_4 \right\}.$$

Other interesting examples are obtained by taking closures with respect to topologies other than the one induced by inner products. For example, the simplices and spectraplices of Example 1.1(a)-(b) admit natural closures in weak topologies (since their elements are viewed as probability measures), and Schur-Horn orbitopes in Section 3.2 admit a natural closure in a certain strong operator topology [88, §1.6] (since we want to associate spectral measures to their elements).

We also mention that the above limiting perspective can be applied to the sequences of invariant cones from Section 6. In particular, the authors of [89] study the difference between the cone of symmetric nonnegative polynomials and the cone of symmetric SOS polynomials under two isomorphisms between the relevant spaces of invariants (neither of which corresponds to our projections). Similarly, the authors of [90] study limits of certain symmetric cones and show how the geometry of these limits underlies various undecidability results. It would be interesting to relate these works to out framework and extend them to other consistent sequences.

8   Conclusions and future work

We developed a systematic framework to study convex sets that can be instantiated in different dimensions using representation stability, as well as a computational method to parametrize such sets and fit them to data. We did so by formally defining free descriptions of convex sets and compatibility conditions relating sets in different dimensions. We then gave conditions on free descriptions to certify this compatibility, and characterized descriptions in a fixed dimension that extend to free descriptions satisfying these conditions. We also used representation stability to systematically derive constant-sized descriptions for sequences of symmetric PSD and relative entropy cones. Finally, we showed that free descriptions certifying compatibility often extend to descriptions of continuous limits of sequences of convex sets. Our work can be viewed as identifying and exploiting a new point of contact between representation stability and convex geometry through conic descriptions of convex sets.

Our work suggests questions and directions for future research in several areas.

(Computational algebra) Is there an algorithm to compute the generation and presentation degrees of a given consistent sequence?

(Lie groups) Can we extend our calculus for presentation degrees in Theorem 4.11 to Lie groups such as $G_n = \mathsf{O}_n$?

(Constructing descriptions) Given a sequence of convex sets instantiable in any relevant dimension, can we systematically construct freely-described, and possibly compatible, approximations for it? When are approximations derived from sums-of-squares hierarchies such as [27] free and certify compatibility?

(Complexity) Is there a systematic framework to study the smallest possible size of a free description for a given sequence of sets, extending the slack operator-based approach for fixed convex sets [91]?

(Free separation) Under what conditions can a point outside a compatible sequence of convex sets be separated by a freely-described sequence, generalizing the Effros-Winkler theorem [92]?

(Statistical inference) How much data do we need to learn a given sequence of sets or functions, and in what dimensions should this data lie?

Acknowledgements

References

[1] Ani Kristo, Kapil Vaidya, Ugur Çetintemel, Sanchit Misra, and Tim Kraska. The case for a learned sorting algorithm. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pages 1001–1016, 2020.

[2] Zhengdao Chen, Lei Chen, Soledad Villar, and Joan Bruna. Can graph neural networks count substructures? In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 10383–10395. Curran Associates, Inc., 2020.

[3] Weichi Yao, Afonso S. Bandeira, and Soledad Villar. Experimental performance of graph neural networks on random instances of max-cut. In Dimitri Van De Ville, Manos Papadakis, and Yue M. Lu, editors, *Wavelets and Sparsity XVIII*, volume 11138, page 111380S. International Society for Optics and Photonics, SPIE, 2019.

[4] Gregory Ongie, Ajil Jalal, Christopher A Metzler, Richard G Baraniuk, Alexandros G Dimakis, and Rebecca Willett. Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory*, 1(1):39–56, 2020.

[5] Vishal Monga, Yuelong Li, and Yonina C. Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2):18–44, 2021.

[6] Maria-Florina Balcan, Travis Dick, Tuomas Sandholm, and Ellen Vitercik. Learning to branch. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 344–353. PMLR, 10–15 Jul 2018.

[7] Tom-Lukas Kriel. An introduction to matrix convex sets and free spectrahedra. *Complex Analysis and Operator Theory*, 13(7):3251–3335, 2019.

[8] J. William Helton. "positive" noncommutative polynomials are sums of squares. *Annals of Mathematics*, 156(2):675–694, 2002.

[9] J. William Helton and Scott McCullough. Every convex free basic semi-algebraic set has an lmi representation. *Annals of Mathematics*, 176(2):979–1013, 2012.

[10] Eric Evert, J William Helton, Igor Klep, and Scott McCullough. Extreme points of matrix convex sets, free spectrahedra, and dilation theory. *The Journal of Geometric Analysis*, 28:1373–1408, 2018.

[11] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization*. Society for Industrial and Applied Mathematics, 2001.

[12] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[13] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[14] J. William Helton, Igor Klep, and Scott McCullough. *Chapter 8: Free Convex Algebraic Geometry*, pages 341–405. Society for Industrial and Applied Mathematics, 2012.

[15] Thomas Church and Benson Farb. Representation theory and homological stability. *Advances in Mathematics*, 245:250–314, 2013.

[16] Venkat Chandrasekaran, Pablo A. Parrilo, and Alan S. Willsky. Convex graph invariants. *SIAM Review*, 54(3):513–541, 2012.

[17] Hamza Fawzi and James Saunderson. Optimal self-concordant barriers for quantum relative entropies. *arXiv preprint arXiv:2205.04581*, 2022.

[18] Hamza Fawzi, Omar Fawzi, Aram Harrow, and Monique Laurent. Geometry and optimization in quantum information. *Oberwolfach Reports*, 18(4):2665–2720, 2022.

[19] Haonan Zhang. From wigner-yanase-dyson conjecture to carlen-frank-lieb conjecture. *Advances in Mathematics*, 365:107053, 2020.

[20] D.L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995.

[21] Venkat Chandrasekaran, Benjamin Recht, Pablo A Parrilo, and Alan S Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.

[22] Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013.

[23] Annie Raymond, James Saunderson, Mohit Singh, and Rekha R Thomas. Symmetric sums of squares over k-subset hypercubes. *Mathematical Programming*, 167(2):315–354, 2018.

[24] Sebastian Debus and Cordian Riener. Reflection groups and cones of sums of squares. *arXiv preprint arXiv:2011.09997*, 2020.

[25] Philippe Moustrou, Helen Naumann, Cordian Riener, Thorsten Theobald, and Hugues Verdure. Symmetry reduction in AM/GM-based optimization. *SIAM Journal on Optimization*, 32(2):765–785, 2022.

[26] Hamza Fawzi, Joao Gouveia, Pablo A. Parrilo, James Saunderson, and Rekha R. Thomas. Lifting for simplicity: Concise descriptions of convex sets. *SIAM Review*, 64(4):866–918, 2022.

[27] Hamza Fawzi, James Saunderson, and Pablo A. Parrilo. Equivariant semidefinite lifts and sum-of-squares hierarchies. *SIAM Journal on Optimization*, 25(4):2212–2243, 2015.

[28] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 223–228, 1988.

[29] Hamza Fawzi, James Saunderson, and Pablo A. Parrilo. Equivariant semidefinite lifts of regular polygons. *Mathematics of Operations Research*, 42(2):472–494, 2017.

[30] J. Saunderson, P. A. Parrilo, and A. S. Willsky. Semidefinite descriptions of the convex hull of rotation matrices. *SIAM Journal on Optimization*, 25(3):1314–1343, 2015.

[31] James A Mingo and Roland Speicher. *Free probability and random matrices*, volume 35. Springer, 2017.

[32] S. Pironio, M. Navascués, and A. Acín. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM Journal on Optimization*, 20(5):2157–2180, 2010.

[33] Sabine Burgdorf, Igor Klep, Janez Povh, et al. *Optimization of polynomials in non-commuting variables*, volume 2. Springer, 2016.

[34] Sander Gribling, David de Laat, and Monique Laurent. Bounds on entanglement dimensions and quantum graph parameters via noncommutative polynomial optimization. *Mathematical programming*, 170:5–42, 2018.

[35] Sander Gribling, David De Laat, and Monique Laurent. Lower bounds on matrix factorization ranks via noncommutative polynomial optimization. *Foundations of Computational Mathematics*, 19(5):1013–1070, 2019.

[36] Andreas Bluhm and Ion Nechita. Joint measurability of quantum effects and the matrix diamond. *Journal of Mathematical Physics*, 59(11):112202, 2018.

[37] Andreas Bluhm, Anna Jenčová, and Ion Nechita. Incompatibility in general probabilistic theories, generalized spectrahedra, and tensor norms. *Communications in Mathematical Physics*, 393(3):1125–1198, 2022.

[38] Thomas Church, Jordan S. Ellenberg, and Benson Farb. FI-modules and stability for representations of symmetric groups. *Duke Mathematical Journal*, 164(9):1833 – 1910, 2015.

[39] Jennifer C.H. Wilson. FI$_\mathcal{W}$-modules and stability criteria for representations of classical Weyl groups. *Journal of Algebra*, 420:269–332, 2014.

[40] Nir Gadish. Categories of FI type: A unified approach to generalizing representation stability and character polynomials. *Journal of Algebra*, 480:450–486, 2017.

[41] Steven V Sam and Andrew Snowden. Stability patterns in representation theory. *Forum of Mathematics, Sigma*, 3:e11, 2015.

[42] Steven Sam and Andrew Snowden. GL-equivariant modules over polynomial rings in infinitely many variables. *Transactions of the American Mathematical Society*, 368(2):1097–1158, 2016.

[43] Steven Sam and Andrew Snowden. Gröbner methods for representations of combinatorial categories. *Journal of the American Mathematical Society*, 30(1):159–203, 2017.

[44] Benson Farb. Representation stability. *arXiv preprint arXiv:1404.4065*, 2014.

[45] Jenny Wilson. An introduction to FI–modules and their generalizations. *Michigan Representation Stability Week*, 2018.

[46] Steven V Sam. Structures in representation stability. *Notices of the American Mathematical Society*, 67(1), 2020.

[47] Dinh Van Le and Tim Römer. Theorems of Carath\'{e}odory, Minkowski-Weyl, and Gordan up to symmetry. *arXiv preprint arXiv:2110.10657*, 2021.

[48] Jan Draisma. *Noetherianity up to Symmetry*, pages 33–61. Springer International Publishing, Cham, 2014.

[49] Christopher H Chiu, Alessandro Danelon, Jan Draisma, Rob H Eggermont, and Azhar Farooq. Sym-noetherianity for powers of gl-varieties. *arXiv preprint arXiv:2212.05790*, 2022.

[50] Yulia Alexandr, Joe Kileel, and Bernd Sturmfels. Moment varieties for mixtures of products. *arXiv preprint arXiv:2301.09068*, 2023.

[51] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer, 2013.

[52] Jean-Pierre Serre et al. *Linear representations of finite groups*, volume 42. Springer, 1977.

[53] R Tyrrell Rockafellar. *Convex analysis*, volume 18. Princeton university press, 1970.

[54] Pablo A. Parrilo. *Chapter 3: Polynomial Optimization, Sums of Squares, and Applications*, pages 47–157. Society for Industrial and Applied Mathematics, 2012.

[55] Silvia Gandy, Benjamin Recht, and Isao Yamada. Tensor completion and low-n-rank tensor recovery via convex optimization. *Inverse Problems*, 27(2):025010, jan 2011.

[56] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[57] Fajwel Fogel, Rodolphe Jenatton, Francis Bach, and Alexandre D' Aspremont. Convex relaxations for permutation problems. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013.

[58] Shmuel Onn. Geometry, complexity, and combinatorics of permutation polytopes. *Journal of Combinatorial Theory, Series A*, 64(1):31–49, 1993.

[59] Günter M Ziegler. *Lectures on polytopes*, volume 152. Springer, 1995.

[60] Raman Sanyal, Frank Sottile, and Bernd Sturmfels. Orbitopes. *Mathematika*, 57(2):275–314, 2011.

[61] Yichuan Ding and Henry Wolkowicz. A low-dimensional semidefinite relaxation for the quadratic assignment problem. *Mathematics of Operations Research*, 34(4):1008–1022, 2009.

[62] Michel X Goemans. Smallest compact formulation for the permutahedron. *Mathematical Programming*, 153(1):5–11, 2015.

[63] Cong Han Lim and Stephen Wright. Beyond the birkhoff polytope: Convex relaxations for vector permutation problems. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[64] Utkan Onur Candogan and Venkat Chandrasekaran. Finding planted subgraphs with few eigenvalues using the Schur–Horn relaxation. *SIAM Journal on Optimization*, 28(1):735–759, 2018.

[65] Raman Sanyal and James Saunderson. Spectral polyhedra. *arXiv preprint arXiv:2001.04361*, 2020.

[66] Alberto Seeger. Convex analysis of spectrally defined matrix functions. *SIAM Journal on Optimization*, 7(3):679–696, 1997.

[67] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 1999.

[68] Hamza Fawzi, James Saunderson, and Pablo A Parrilo. Semidefinite approximations of the matrix logarithm. *Foundations of Computational Mathematics*, 19(2):259–296, 2019.

[69] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[70] Alexander F Sidorenko. Inequalities for functionals generated by bipartite graphs. *Diskretnaya Matematika*, 3(3):50–65, 1991.

[71] Joonkyung Lee and Bjarne Schülke. Convex graphon parameters and graph norms. *Israel Journal of Mathematics*, 242(2):549–563, 2021.

[72] Martin Doležal, Jan Grebík, Jan Hladký, Israel Rocha, and Václav Rozhoň. Cut distance identifying graphon parameters over weak* limits. *Journal of Combinatorial Theory, Series A*, 189:105615, 2022.

[73] Marc Finzi, Max Welling, and Andrew Gordon Gordon Wilson. A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 3318–3328. PMLR, 18–24 Jul 2021.

[74] Craig Gotsman and Sivan Toledo. On the computation of null spaces of sparse rectangular matrices. *SIAM Journal on Matrix Analysis and Applications*, 30(2):445–463, 2008.

[75] Pawel Kowal. Null space of a sparse matrix. https://www.mathworks.com/matlabcentral/fileexchange/11120-null-space-of-a-sparse-matrix, 2006. [Retrieved July 12, 2022].

[76] Christopher C Paige and Michael A Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software (TOMS)*, 8(1):43–71, 1982.

[77] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1):95–128, 2004.

[78] Vlad Timofte. On the positivity of symmetric polynomial functions.: Part I: General results. *Journal of Mathematical Analysis and Applications*, 284(1):174–190, 2003.

[79] Cordian Riener. Symmetric semi-algebraic sets and non-negativity of symmetric polynomials. *Journal of Pure and Applied Algebra*, 220(8):2809–2815, 2016.

[80] José Acevedo and Mauricio Velasco. Test sets for nonnegativity of polynomials invariant under a finite reflection group. *Journal of Pure and Applied Algebra*, 220(8):2936–2947, 2016.

[81] Wee Liang Gan and Liping Li. An inductive machinery for representations of categories with shift functors. *Transactions of the American Mathematical Society*, 371(12):8513–8534, 2019.

[82] K.S. Wang and L.C. Grove. Realizability of representations of finite groups. *Journal of Pure and Applied Algebra*, 54(2):299–310, 1988.

[83] Annie Raymond, Mohit Singh, and Rekha R Thomas. Symmetry in Turán sums of squares polynomials from flag algebras. *Algebraic Combinatorics*, 1(2):249–274, 2018.

[84] Alexander A. Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[85] Riley Murray, Venkat Chandrasekaran, and Adam Wierman. Signomial and polynomial optimization via relative entropy and partial dualization. *Mathematical Programming Computation*, 13(2):257–295, 2021.

[86] Venkat Chandrasekaran and Parikshit Shah. Relative entropy relaxations for signomial optimization. *SIAM Journal on Optimization*, 26(2):1147–1173, 2016.

[87] Christian Borgs, Jennifer Chayes, Henry Cohn, and Yufei Zhao. An l^p theory of sparse graph convergence i: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, 2019.

[88] Claire Anantharaman and Sorin Popa. An introduction to II_1 factors. `https://www.math.ucla.edu/~popa/Books/IIun.pdf`, 2017.

[89] Grigoriy Blekherman and Cordian Riener. Symmetric non-negative forms and sums of squares. *Discrete & Computational Geometry*, 65(3):764–799, 2021.

[90] Jose Acevedo, Grigoriy Blekherman, Sebastian Debus, and Cordian Riener. The wonderful geometry of the vandermonde map. *arXiv preprint arXiv:2303.09512*, 2023.

[91] Joao Gouveia, Pablo A Parrilo, and Rekha R Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013.

[92] Edward G. Effros and Soren Winkler. Matrix convexity: Operator analogues of the bipolar and Hahn–Banach theorems. *Journal of Functional Analysis*, 144(1):117–152, 1997.

A   Representations of categories

As Remark 4.13 shows, the set of embeddings from low to high dimensions in a consistent sequence $\mathscr{V}$ of $\{G_n\}$-representations, determined by $\{G_n\}$ and the stabilizing subgroups $\{H_{n,d}\}$ from Definition 4.2, play a central role in our framework. These sets of embeddings are conveniently encoded in a category, whose representations are precisely the $\mathscr{V}$-modules we used in Section 4. Morphisms between such representations in the categorical sense coincide with morphisms of sequences. This categorical approach to representation stability was introduced in [38] for the case $G_n = \mathsf{S}_n$ and the $H_{n,d}$ from Example 4.3(a), and has since been extended to other groups in [39, 40, 43].

Definition A.1. *A (real)* representation *of a category $\mathcal{C}$, also called a $\mathcal{C}$-module, is a functor $\mathcal{C} \to \mathrm{Vect}_{\mathbb{R}}$ from $\mathcal{C}$ to the category of real vector spaces.*

In other words, a $\mathcal{C}$-module is an assignment of a vector space V_n to each object $n \in \mathcal{C}$ and a linear map $\varphi_{n,N} \colon V_n \to V_N$ to each morphism in $\mathrm{Hom}_{\mathcal{C}}(n, N)$ such that compositions are respected. Each V_n is then a representation of the group $G_n = \mathrm{End}_{\mathcal{C}}(n)^{\times}$ of the automorphisms of n in $\mathcal{C}$. Every consistent sequence is a representation of a suitable category.

Definition A.2. *Given a consistent sequence $\mathscr{V} = \{(V_n, \varphi_n)\}$ of $\{G_n\}$-representations, define a category $\mathcal{C}_{\mathscr{V}}$ whose set of objects is $\mathbb{N}$ and whose morphisms are $\mathrm{Hom}_{\mathcal{C}_V}(n, N) = \{g\varphi_{N-1}\cdots\varphi_n : g \in G_N\}$ for $n \leq N$ and zero otherwise. Note that $\mathrm{Hom}_{\mathcal{C}_V}(n, N) = G_N/H_{N,n}$ where $H_{N,n} \subseteq G_N$ is the subgroup of elements acting trivially on V_n.*

This definition clearly extends to consistent sequences indexed by posets (Remark 2.3). If $\mathscr{U} = \{(U_n, \psi_n)\}$ is a $\mathscr{V}$-module (Definition 4.4), then $\mathscr{U}$ is a $\mathcal{C}_{\mathscr{V}}$-module, since sending $n \in \mathbb{N}$ to U_n and $g\varphi_{N-1}\cdots\varphi_n$ to the map $g\psi_{N-1}\cdots\psi_n$ for each $g \in G_N$ is a well-defined functor $\mathcal{C}_{\mathscr{V}} \to \mathrm{Vect}_{\mathbb{R}}$. Conversely, if $\mathscr{U}$ is a $\mathcal{C}_{\mathscr{V}}$-module then it is a $\mathscr{V}$-module since $H_{n,d}$ acts trivially on the image of $\psi_{N-1}\cdots\psi_d$ by definition of a functor. Furthermore, if $\mathscr{W}, \mathscr{U}$ are $\mathscr{C}_{\mathscr{V}}$-modules, then a morphism of functors $\mathscr{W} \to \mathscr{U}$ (also called a natural transformation) coincides with a morphism of sequences in Definition 2.8. Applying the constructions in Remark 2.4 to $\mathcal{C}$-modules yields other $\mathcal{C}$-modules.

Example A.3. *Here are some examples of the categories resulting from Definition A.2.*

(a) *The category corresponding to Examples 2.2(a)-(b) with $G_n = \mathsf{S}_n$ is (the skeleton of) $\mathcal{C} = \mathrm{FI}$, the category whose objects are finite sets and whose morphisms are injections.*

(b) *The category corresponding to Examples 2.2(a)-(b) with $G_n = \mathsf{B}_n$ (resp., D_n) is $\mathcal{C} = \mathrm{FI}_{BC}$ (resp., $\mathcal{C} = \mathrm{FI}|_D$) defined in [39], whose objects are the sets $[\pm n] := \{\pm 1, \ldots, \pm n\}$ for $n \in \mathbb{N}$ and whose morphisms are injections $f\colon [\pm n] \hookrightarrow [\pm N]$ satisfying $f(-i) = -f(i)$ (and reverse evenly-many signs if $G_n = \mathsf{D}_n$).*

(c) *The category corresponding to the graphon sequence is the opposite category $\mathcal{C} = \mathcal{P}_2^{\mathrm{op}}$ of the category $\mathcal{P}_2$ with objects $[2^n]$ and morphisms which are 2^{N-n}-to-one surjections $[2^N] \to [2^n]$, or equivalently, partitions of $[2^N]$ into 2^n equal parts.*

Following [39], we say $\mathcal{C} = \mathrm{FI}|_{\mathcal{W}}$ if $\mathcal{C} = \mathrm{FI}, \mathrm{FI}|_{BC}$ or $\mathrm{FI}|_D$.

(Algebraically) free $\mathcal{C}$-modules are defined exactly as in Definition 4.6, see [38, Def. 2.2.2] and [40, Def. 1.8,3.1] for example. The theory of [40] gives the following result for $\mathcal{C} = \mathrm{FI}|_{\mathcal{W}}$, which extends to categories of FI-type introduced in [40].

Theorem A.4 ([40, Thm. B(1)]). *Tensor products of free $\mathrm{FI}|_{\mathcal{W}}$-modules are free.*

The following result illustrates two further properties of $\mathrm{FI}|_{\mathcal{W}}$-modules.

Theorem A.5 (Noetherianity and tensor products). *Let $\mathcal{C} = \mathrm{FI}|_{\mathcal{W}}$.*

(Noetherianity) *Any submodule of a finitely-generated $\mathcal{C}$-module is finitely-generated.*

(Tensor products) *If $\mathscr{V}_1$ and $\mathscr{V}_2$ are $\mathcal{C}$-modules generated in degrees d_1 and d_2, respectively, then $\mathscr{V}_1 \otimes \mathscr{V}_2$ is generated in degree $d_1 + d_2$.*

Proof. Noetherianity is shown for FI in [38, Thm. 1.13] and for $\mathrm{FI}|_{BC}, \mathrm{FI}|_D$ in [39, Thm. 4.21]. The generation degree bound is shown in [38, Prop. 2.3.6] for FI and in [39, Prop. 5.2] for $\mathrm{FI}|_{BC}, \mathrm{FI}|_D$. □

Noetherianity helps explain the ubiquity of representation stability, while the generation degree bound for tensor products allows us to bound the generation degrees of complicated sequences from degrees of simple ones. The two properties in Theorem A.5 hold over more general categories than $\mathrm{FI}|_{\mathcal{W}}$, including for categories of FI-type and certain quasi-Gröbner categories introduced in [43]. We remark that these two types of categories only include representations of finite groups, and do not include the graphon category in Example A.3(c), whose properties would be interesting to study in future work.

Definition A.6 (Property (TFG)). *We say that a category $\mathcal{C}$ satisfies property (TFG) if Tensor products of free $\mathcal{C}$-modules are Free and satisfy the Generation degree bound in Theorem A.5.*

An example of a category not satisfying (TFG) is given in [43, Rmk. 7.4.3]. We can use property (TFG) to obtain a calculus for presentation degrees from which Theorem 4.11 may be deduced.

Proposition A.7. *Suppose $\mathcal{C}$ is a category satisfying (TFG). If $\mathscr{V}, \mathscr{U}$ are $\mathcal{C}$-modules which are generated in degrees d_V, d_U and presented in degrees k_V, k_U, respectively, then $\mathscr{V} \otimes \mathscr{U}$ is presented in degree $\max\{d_V + k_U, d_U + k_V\}$.*

Proof. Suppose $\mathscr{F}_V, \mathscr{F}_U$ are free $\mathcal{C}$-modules generated in degrees d_V, d_U, respectively, and $\mathscr{F}_V \to \mathscr{V}$ and $\mathscr{F}_U \to \mathscr{U}$ are surjective morphisms whose kernels $\mathscr{K}_V$ and $\mathscr{K}_U$ are generated in degrees r_V, r_U, respectively. Then $\mathscr{F}_V \otimes \mathscr{F}_U$ is a free $\mathcal{C}$-module generated in degree $d_V + d_U$ by (TFG), and the morphism $\mathscr{F}_V \otimes \mathscr{F}_U \to \mathscr{V} \otimes \mathscr{U}$ is surjective with kernel $\mathscr{K}_V \otimes \mathscr{F}_U + \mathscr{F}_V \otimes \mathscr{K}_U$. Since $\mathscr{K}_V \otimes \mathscr{F}_U$ is generated in degree $r_V + d_U$ and similarly for $\mathscr{F}_U \otimes \mathscr{K}_V$, their sum is generated in degree $\max\{r_V + d_U, d_V + r_U\}$. $\qquad\square$

Our next goal is to understand presentation degrees for images, and in particular, for Schur functors. We begin with a number of elementary lemmas.

Lemma A.8. *Let $\mathscr{V}$ and $\mathscr{U}$ be $\mathcal{C}$-modules. If $\mathscr{V}, \mathscr{U}$ are generated in degrees d_V, d_U and presented in degrees k_V, k_U, respectively, then $\mathscr{V} \oplus \mathscr{U}$ is generated in degree $\max\{d_V, d_U\}$ and presented in degree $\max\{k_V, k_U\}$.*

Proof. The claim about the generation degree is immediate from its definition. Suppose that $\mathscr{F}_V \to \mathscr{V}$ and $\mathscr{F}_U \to \mathscr{U}$ are surjective morphisms with $\mathscr{F}_V, \mathscr{F}_U$ being free $\mathcal{C}$-modules generated in degrees d_V, d_U with kernels $\mathscr{K}_V, \mathscr{K}_U$ generated in degrees k_V, k_U, respectively. Then $\mathscr{F}_V \oplus \mathscr{F}_U$ is free (by definition) and generated in degree $\max\{d_V, d_U\}$, and surjects onto $\mathscr{V} \oplus \mathscr{U}$ with kernel $\mathscr{K}_V \oplus \mathscr{K}_U$ which is generated in degree $\max\{k_V, k_U\}$. $\qquad\square$

Lemma A.9. *Let $\mathscr{V} = \{V_n\}$ and $\mathscr{U} = \{U_n\}$ be two $\mathcal{C}$-modules, let $\mathscr{A} = \{A_n\}\colon \mathscr{V} \to \mathscr{U}$ be a surjective morphism, and let $\mathscr{W} = \{W_n \subseteq U_n\}$ be a $\mathcal{C}$-submodule of $\mathscr{U}$. If $\ker \mathscr{A}$ is generated in degree d and $\mathscr{W}$ is generated in degree d_W, then $\mathscr{A}^{-1}(\mathscr{W}) = \{A_n^{-1}(W_n)\}$ is a $\mathcal{C}$-module generated in degree $\max\{d, d_W\}$.*

Proof. Define the consistent sequence $Z_n = \mathbb{R}[G_n]\left(A_{d_W}^{\dagger} W_{d_W}\right) \subseteq V_n$ if $n \geq d_W$ and $Z_n = 0$ otherwise, where $A_{d_W}^{\dagger}$ is the pseudoinverse of A_{d_W}. Note that $\{Z_n\}$ is generated in degree d_W. Moreover, $A_n^{-1}(W_n) = \ker A_n + Z_n$. Indeed, we have $A_n A_{d_W}^{\dagger} = A_d A_d^{\dagger} = \mathrm{id}_{U_d}$ because $\{A_n\}$ is a surjective morphism, hence $A_n(\ker A_n + Z_n) = A_n(Z_n) = \mathbb{R}[G_n]W_{d_W} = W_n$. Conversely, if $A_n x \in W_n = \mathbb{R}[G_n]W_{d_W}$ then we can write $A_n x = \sum_i g_i w_i$ for $g_i \in G_n$ and $w_i \in W_{d_W}$. Then $\hat{x} = \sum_i g_i A_{d_W}^{\dagger} w_i \in Z_n$ and $A_n(x - \hat{x}) = 0$, so $x \in \ker A_n + Z_n$. Since $\ker \mathscr{A}$ is generated in degree d and $\{Z_n\}$ is generated in degree d_W, their sum is generated in degree $\max\{d, d_W\}$. $\qquad\square$

Lemma A.10. *Suppose $\mathscr{V} = \{V_n\}$ and $\{U_n\}$ are two $\mathcal{C}$-modules, and $\mathscr{A} = \{A_n\}\colon \mathscr{V} \to \mathscr{U}$ is a morphism. If $\mathscr{V}$ is generated in degree d, then $\mathrm{im}\,\mathscr{A}$ is generated in degree d. If, moreover, $\mathscr{A}^* = \{A_n^*\}$ is a morphism, then $\ker \mathscr{A}$ is also generated in degree d.*

Proof. The first claim follows from $A_n(V_n) = A_n(\mathbb{R}[G_n]V_d) = \mathbb{R}[G_n]A_n(V_d) = \mathbb{R}[G_n]A_d(V_d)$, where we used the equivariance of A_n and the fact that $A_n|_{V_d} = A_d$. For the second claim, note that if $\mathscr{A}^*$ is a morphism, then $\{\mathrm{im}A_n^* = (\ker A_n)^{\perp}\}$ is a $\mathcal{C}$-submodule of $\mathscr{V}$. Therefore, $\{\mathcal{P}_{\ker A_n}\}\colon \mathscr{V} \to \mathscr{V}$ is a morphism, and its image is precicely $\ker \mathscr{A}$. $\qquad\square$

Proposition A.11. *Suppose $\mathscr{V} = \{V_n\}, \mathscr{U} = \{U_n\}$ are two $\mathcal{C}$-modules and both $\mathscr{A} = \{A_n\colon V_n \to U_n\}$ and $\{A_n^*\colon U_n \to V_n\}$ are morphisms. If $\mathscr{V}$ is generated in degree d and presented in degree k, then $\mathrm{im}\,\mathscr{A} = \{A_n(V_n)\}$ is generated in degree d and presented in degree k.*

Proof. Let $\mathscr{F} = \{F_n\}$ be a free $\mathcal{C}$-module generated in degree d and let $\mathscr{B} = \{B_n\}\colon \mathscr{F} \to \mathscr{V}$ be a surjective morphism whose kernel $\mathscr{K} = \{K_n\}$ is generated in degree k. The composition $\mathscr{F} \xrightarrow{\mathscr{B}} \mathscr{V} \xrightarrow{\mathscr{A}} \mathrm{im}\,\mathscr{A}$ is a surjective morphism from the free $\mathcal{C}$-module $\mathscr{F}$ whose kernel is $\mathscr{B}^{-1}(\ker \mathscr{A})$ and is generated in degree $\max\{d, k\} = k$ by Lemmas A.9 and A.10. $\qquad\square$

Corollary A.12. *Suppose $\mathscr{C}$ satisfies property (TFG). If $\mathscr{V}$ is a $\mathcal{C}$-module generated in degree d and presented in degree k, and λ is a partition, then $\mathbb{S}^{\lambda}\mathscr{V}$ is generated in degree $d|\lambda|$ and presented in degree $k + d(|\lambda| - 1)$.*

Schur functors generalize symmetric and alternating algebras, see [51, §6.1]. Their generation degree for $\mathcal{C} = \mathrm{FI}$ was bounded using a similar approach in [38, Prop. 3.4.3].

Proof. By Proposition A.7, the $\mathcal{C}$-module $\mathscr{V}^{\otimes|\lambda|}$ is generated in degree $d|\lambda|$ and presented in degree $d(|\lambda| - 1) + k$. Let $\mathsf{S}_{|\lambda|}$ act on each $V_n^{\otimes|\lambda|}$ by permuting its factors $\sigma \cdot (v_1 \otimes \cdots \otimes v_{|\lambda|}) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(|\lambda|)}$, which is an orthogonal action commuting with the emebddings $V_n^{\otimes|\lambda|} \subseteq V_{n+1}^{\otimes|\lambda|}$. In this way, any element $c_\lambda \in \mathbb{R}[\mathsf{S}_{|\lambda|}]$ defines a morphism $c_\lambda \colon \mathscr{V}^{\otimes|\lambda|} \to \mathscr{V}^{\otimes|\lambda|}$ such that $c_\lambda^* \in \mathbb{R}[\mathsf{S}_{|\lambda|}]$ is also a morphism. If $c_\lambda \in \mathbb{R}[\mathsf{S}_{|\lambda|}]$ is the Young symmetrizer corresponding to partition λ, then $\operatorname{im} c_\lambda = \mathbb{S}^\lambda \mathscr{V}$. Hence the result follows from Proposition A.11. $\qquad\square$

We conjecture that, moreover, Schur functors of free modules remain free. This is true if $\mathcal{C} = \mathrm{FI}|_{\mathcal{W}}$ and $\mathscr{V} = \{\mathbb{R}^n\}$ by Proposition 6.5(c). We conclude this appendix by summarizing our calculus for generation and presentation degrees. Instantiating the following theorem with $\mathcal{C} = \mathrm{FI}|_{\mathcal{W}}$ yields Theorem 4.11.

Theorem A.13 (Calculus for generation and presentation degrees)**.** *Let $\mathscr{V} = \{V_n\}, \mathscr{U} = \{U_n\}$ be $\mathcal{C}$-modules generated in degrees d_V, d_U and presented in degrees k_V, k_U, respectively.*

(Sums) *$\mathscr{V} \oplus \mathscr{U}$ is generated in degree $\max\{d_V, d_U\}$ and presented in degree $\max\{k_V, k_U\}$.*

(Images and kernels) *If $\mathscr{A} \colon \mathscr{V} \to \mathscr{U}$ and $\mathscr{A}^*$ are morphisms, then $\operatorname{im} \mathscr{A}$ and $\ker \mathscr{A}$ are generated in degree d_V and presented in degree k_V.*

Suppose $\mathcal{C}$ satisfies (TFG). Then

(Tensors) *$\mathscr{V} \otimes \mathscr{U}$ is generated in degree $d_V + d_U$ and presented in degree $\max\{k_V + d_U, k_U + d_V\}$.*

(Schur functors) *$\mathbb{S}^\lambda \mathscr{V}$ is generated in degree $d_V|\lambda|$ and presented in degree $d_V(|\lambda| - 1) + k_V$ for any partition λ.*