

**SYMMETRIC SUMS OF SQUARES OVER  
 $k$ -SUBSET HYPERCUBES**

ANNIE RAYMOND, JAMES SAUNDERSON, MOHIT SINGH, AND REKHA R. THOMAS

ABSTRACT. We consider the problem of finding sum of squares (sos) expressions to establish the non-negativity of a symmetric polynomial over a discrete hypercube whose coordinates are indexed by the  $k$ -element subsets of  $[n]$ . For simplicity, we focus on the case  $k = 2$ , but our results extend naturally to all values of  $k \geq 2$ . We develop a variant of the Gatermann-Parrilo symmetry-reduction method tailored to our setting that allows for several simplifications and a connection to flag algebras.

We show that every symmetric polynomial that has a sos expression of a fixed degree also has a succinct sos expression whose size depends only on the degree and not on the number of variables. Our method bypasses much of the technical difficulties needed to apply the Gatermann-Parrilo method, and offers flexibility in obtaining succinct sos expressions that are combinatorially meaningful. As a byproduct of our results, we arrive at a natural representation-theoretic justification for the concept of flags as introduced by Razborov in his flag algebra calculus. Furthermore, this connection exposes a family of non-negative polynomials that cannot be certified with any fixed set of flags, answering a question of Razborov in the context of our finite setting.

## 1. INTRODUCTION

Polynomial optimization over discrete hypercubes plays a central role in many areas such as combinatorial optimization, decision problems and proof complexity. In many situations, it is natural to consider  $k$ -subset hypercubes by which we mean discrete hypercubes whose coordinates are indexed by the  $k$ -element subsets of a ground set  $[n]$ . For instance, a major focus in extremal graph theory is to optimize the edge (hyperedge) density in families of graphs (hypergraphs) with specified structure which can be cast as optimization problems over  $k$ -subset hypercubes. In this scenario, as in many others, the polynomial to be optimized is often symmetric which allows representation-theoretic techniques to dramatically cut down on computations. In this paper, we consider the general problem of optimizing a symmetric polynomial over a  $k$ -subset hypercube  $\mathcal{V}_{n,k} := \{0, 1\}^{\binom{n}{k}}$ . We focus on the case  $k = 2$ , and hence use the notation  $\mathcal{V}_n$  in place of  $\mathcal{V}_{n,2} = \{0, 1\}^{\binom{n}{2}}$  throughout. As we will explain in Section 7, our results extend naturally to all  $k \geq 2$ .

Phrased differently, our central problem is to certify the non-negativity of a symmetric polynomial  $p$  over  $\mathcal{V}_n$  which can be done by finding a sum of squares (sos) expression that equals  $p$  as a function on  $\mathcal{V}_n$ . In [GP04], Gatermann and Parrilo showed how to use representation theory to simplify the computations involved in finding a sos representation of a polynomial  $p$  that is invariant under the action

---

*Date:* August 9, 2016.

Saunderson and Thomas were partially supported by the National Science Foundation under the grants CCF-1409836 and DMS-1418728 respectively.

of a finite group. We propose a variant of their method that is adapted to the combinatorics in our setting, and hence, offers many simplifications and advantages. Among the highlights is a proof that if  $\mathbf{p}$  has a sos certificate of degree  $d$ , then it also has a succinct sos expression whose size depends on  $d$  but is independent of  $n$ . Secondly, we show that Razborov's theory of flags makes a natural entry into this problem as a byproduct of the representation theory of the symmetric group.

We now introduce some notation that will help us elaborate on our results. Fix a positive integer  $n$  and let  $\mathbb{R}[x] := \mathbb{R}[x_{ij} : 1 \leq i < j \leq n]$  be the polynomial ring in variables indexed by the pairs  $ij$  where  $1 \leq i < j \leq n$ . The set of polynomials that vanish on the discrete hypercube  $\mathcal{V}_n$  is the ideal  $\mathcal{I}_n$  in  $\mathbb{R}[x]$  generated by the  $\binom{n}{2}$  polynomials  $x_{ij}^2 - x_{ij}$  for all  $1 \leq i < j \leq n$ . The set of functions on  $\mathcal{V}_n$  may be identified with the quotient ring  $\mathbb{R}[x]/\mathcal{I}_n$ . Recall that two polynomials  $f, g \in \mathbb{R}[x]$  represent the same function in  $\mathbb{R}[\mathcal{V}_n]$  if and only if  $f - g \in \mathcal{I}_n$ , written as  $f \equiv g \pmod{\mathcal{I}_n}$ . This algebraic language is helpful in examples. The elements of  $\mathbb{R}[x]/\mathcal{I}_n$  are in bijection with the *square-free* polynomials in  $\mathbb{R}[x]$ , namely, those polynomials in which every monomial is square-free or multilinear. As a vector space, it will be convenient to identify  $\mathbb{R}[\mathcal{V}_n]$  with the set of all square-free polynomials in  $\mathbb{R}[x]$ . Let  $\mathbb{R}[x]_{\leq d}$  denote the polynomials of degree at most  $d$  in  $\mathbb{R}[x]$ . Denote by  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  the set of functions of degree at most  $d$  in  $\mathbb{R}[\mathcal{V}_n]$ , namely the quotient ring  $\mathbb{R}[x]_{\leq d}/\mathcal{I}_n$ . As a vector space,  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  can be identified with the set of all square-free polynomials of degree at most  $d$ .

The symmetric group  $\mathfrak{S}_n$  acts on  $\mathbb{R}[x]$  by linearly extending the action of  $\mathfrak{S}_n$  on monomials via  $\mathfrak{s} \cdot x_{ij} := x_{\mathfrak{s}(i)\mathfrak{s}(j)}$  for each  $\mathfrak{s} \in \mathfrak{S}_n$ . Since the ideal  $\mathcal{I}_n$  is invariant under this action, the action of  $\mathfrak{S}_n$  extends to  $\mathbb{R}[\mathcal{V}_n]$  and  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ . We say that a polynomial  $\mathbf{p}$  is *symmetric* or  *$\mathfrak{S}_n$ -invariant* if it is fixed under this action of  $\mathfrak{S}_n$ .

Every non-negative  $\mathbf{p}$  in  $\mathbb{R}[\mathcal{V}_n]$  is a sum of squares, i.e.,  $\mathbf{p} = \sum f_i^2$  for finitely many  $f_i \in \mathbb{R}[\mathcal{V}_n]$  [Par02], see also [Lau09, Theorem 2.4]. If the  $f_i$  are restricted to come from a fixed subspace  $V \subseteq \mathbb{R}[\mathcal{V}_n]$ , then we say that  $\mathbf{p}$  is *V-sos*. If  $\mathbf{p}$  is  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ -sos, we simply say that  $\mathbf{p}$  is *d-sos*.

If  $\mathbf{p} \in \mathbb{R}[\mathcal{V}_n]$  is *V-sos*, then a sos certificate for it can be found by solving a semi-definite program (SDP) (see, for instance, [BPT13, Chapter 3]). Indeed, this can be done by finding a positive semidefinite (psd) matrix  $Q$  such that  $\mathbf{p} = \text{tr}(Q \mathbf{v} \mathbf{v}^\top)$ , where  $\mathbf{v}$  is a vector containing the elements in a basis of  $V$ , and  $\text{tr}(\cdot)$  stands for trace. In particular, if  $\mathbf{p}$  is *d-sos*, we could choose the entries of  $\mathbf{v}$  to consist of all square-free monomials of degree at most  $d$ . Since the number of such monomials is

$$\sum_{i=0}^d \binom{n}{i},$$

the size of  $Q$  becomes unwieldy as  $n$  grows.

If  $\mathbf{p}$  is *d-sos* and invariant with respect to a finite group, Gatermann and Parrilo [GP04] show how to find sos expressions for  $\mathbf{p}$  by solving a (potentially) much smaller SDP than the one mentioned above. In our setting, their starting point is the isotypic decomposition of the vector space  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$  under the action of  $\mathfrak{S}_n$ . This decomposition has the form  $V = \bigoplus_{\lambda} V_{\lambda}$  where  $\lambda$  is a partition of  $n$ , and each isotypic  $V_{\lambda}$  is in turn a direct sum of  $m_{\lambda}$  copies of a single irreducible representation  $S_{\lambda}$  for  $\mathfrak{S}_n$ . Applying the results of [GP04] to our setting, one can show that there exist  $m_{\lambda} \times m_{\lambda}$  psd matrices  $\tilde{Q}_{\lambda}$  such that

$$(1) \quad \mathbf{p} = \sum_{\lambda} \text{tr} \left( \tilde{Q}_{\lambda} \text{sym}(\tilde{\mathbf{v}}_{\lambda} \tilde{\mathbf{v}}_{\lambda}^\top) \right).$$

Here, for each partition  $\lambda$ , the entries of  $\tilde{\mathbf{v}}_\lambda$  consist of elements in a basis for its  $m_\lambda$ -dimensional multiplicity space (which is discussed in Section 2 and Appendix A). The operation  $\text{sym}(\cdot)$  symmetrizes its argument under the group action. Hence, if  $\mathbf{p}$  is symmetric and  $d$ -sos, we can find a symmetry-reduced sos description of  $\mathbf{p}$  by solving a SDP of size  $\sum_\lambda m_\lambda$ .

A priori, it is unclear how much smaller the size of the Gatermann-Parrilo SDP is compared to the dimension of  $V$ . Another major challenge is that constructing this SDP requires the explicit knowledge of a basis for each of the multiplicity spaces (i.e., appropriate vectors  $\tilde{\mathbf{v}}_\lambda$ ). The algorithm outlined in [GP04] requires constructing a symmetry-adapted basis for  $V$  (see, e.g., [FS92]), which is computationally costly, having complexity that depends on both  $n$  and  $d$ .

In this paper, we focus on finding sos certificates for symmetric polynomials over discrete hypercubes  $\mathcal{V}_{n,k}$  when  $k \geq 2$ . In these cases, little is explicitly known about how the space  $\mathbb{R}[\mathcal{V}_{n,k}]$  of square-free polynomials in  $\binom{n}{k}$  variables decomposes under the action of  $\mathfrak{S}_n$  on  $k$ -tuples. This makes a direct application of the Gatermann-Parrilo method infeasible in this setting. This is in sharp contrast with the simpler case  $k = 1$  for which the isotypic decomposition of  $\mathbb{R}[\mathcal{V}_{n,1}]_{\leq d}$ , and hence all the information required for the Gatermann-Parrilo SDP, is explicitly known (see, e.g., [BGP]). This decomposition underpins a number of recent results related to non-negative  $\mathfrak{S}_n$ -invariant polynomials on  $\{0, 1\}^n$  [BGP, LPdWY16, KLM16].

**1.1. Our results.** Our first contribution is to show that if  $\mathbf{p}$  is symmetric and  $d$ -sos, then it has a symmetry-reduced sos certificate that can be obtained by solving a SDP of size independent of  $n$ . Throughout, we will informally refer to such a sos expression as being *succinct*. We establish this result in two steps. First, we show that the number of partitions needed in (1) is bounded above by  $p(0) + p(1) + p(2) + \dots + p(2d)$  (where  $p(i)$  is the number of partitions of  $i$ ), a quantity that is independent of  $n$ . Second, we show that each  $m_\lambda$  is also bounded above by a quantity that is independent of  $n$  (see Proposition 3.21 for a precise statement).

Our next contribution is to develop a variant of the Gatermann-Parrilo method tailored to our setting, that circumvents the representation-theoretic difficulties inherent in constructing bases for the multiplicity spaces. In particular, we show how to construct subspaces that contain the multiplicity spaces, have dimension depending only on  $d$ , and that are spanned by combinatorially meaningful polynomials that can be enumerated easily. One such spanning set arises as polynomial analogs of a specific collection of partially labeled graphs present in both the theory of graph homomorphisms [Lov12] and flag algebras [Raz07]. Partially labeled graphs are called *flags* in the latter. We will use this terminology for brevity and since we make several connections to the work in [Raz07].

A key result coming from the connections we make to flags is that if  $\mathbf{p}$  is symmetric and  $d$ -sos, then there is an sos expression for  $\mathbf{p}$  that uses only flag polynomials from flags on at most  $2d$  vertices. Further, we prove that even particular restricted flag sos expressions commonly found in the flag algebra literature suffice for this result. Together with [RST15], this implies the surprising result that flag methods are just as capable as standard symmetry-reduction methods for providing sos certificates for symmetric polynomials over  $\mathcal{V}_n$ . Moreover, this connection to flags offers a family of symmetric polynomials of fixed degree that cannot be certified with any fixed set of flags, answering a question in [Raz07] in the finite setting.

The theories of flag algebras and graph homomorphisms have emerged as powerful tools for establishing asymptotic (symmetric) inequalities among graph densities in extremal graph theory, by expressing them as sums of squares up to an error that vanishes asymptotically. Viewed as optimization problems, these are instances of polynomial optimization with infinitely many variables. Our results show that there is much to be gained by pausing at a finite  $n$ . We show that flags can be used to exactly certify the non-negativity of *any* symmetric polynomial over the discrete hypercube  $\mathcal{V}_n$  and its subsets. We illustrate our methods through a problem from Ramsey theory and a Turán problem for 4-cycles.

**1.2. Organization of the paper.** This paper is organized as follows. In Section 2.1, we introduce some concepts from the representation theory of  $\mathfrak{S}_n$ . In particular, we define a key ingredient for the paper, namely subspaces  $W_{\tau_\lambda}$  indexed by tableaux of shape  $\lambda$ , that are isomorphic to the multiplicity spaces referred to in Section 1.1. These subspaces play a crucial role in the Gatermann-Parrilo method tailored to our setting, which we present in Section 2.2. A proof can be found in Appendix A. We prove in Section 2.3 that if  $d$  is fixed, then the number of partitions that are needed in the resulting sos certificate for  $\mathbf{p}$  is bounded by a function of  $d$  that is independent of  $n$ . In Section 2.4, we provide a variant of the Gatermann-Parrilo method that only requires spanning sets for the subspaces  $W_{\tau_\lambda}$  as opposed to bases. In Section 3, we construct three such interrelated spanning sets, all of which have the property that their sizes depend on  $d$  but not  $n$ . In the process, we naturally encounter flags and their densities as studied by Razborov. In Section 4, we prove that the flag sos expressions that emerge from our method can be simplified to take on the more restricted form commonly found in the flag algebra literature. We also show that every symmetric  $d$ -sos expression can be retrieved from flags on at most  $2d$  vertices. Our methods naturally extend to subsets of  $\mathcal{V}_n$ . We illustrate this in Section 5 by giving sos proofs of two known results. The first is an upper bound on the edge density of a  $n$ -vertex graph that avoids 4-cycles, and the second is that the Ramsey number  $R(3, 3) = 6$ . The technical details of the Ramsey proof are in Appendix B. In Section 6, we give a family of non-negative symmetric polynomials that cannot be certified by a fixed collection of flags. We close the paper in Section 7 with a discussion of our results, and their extension to hypergraphs.

**Acknowledgments.** Several people offered valuable input to this paper. In particular, we thank Albert Atserias, Monty McGovern, Pablo Parrilo and Sasha Razborov. We especially thank Greg Blekherman for his help with several facets of the representation theory that underlies this work.

## 2. SUMS OF SQUARES CERTIFICATES FOR SYMMETRIC POLYNOMIALS

Suppose that  $\mathbf{p} \in \mathbb{R}[\mathcal{V}_n]$  is  $d$ -sos. In other words, suppose that  $\mathbf{p} = \sum f_i^2$  where  $f_i \in \mathbb{R}[\mathcal{V}_n]_{\leq d}$  for all  $i$ . If, in addition,  $\mathbf{p}$  is fixed by the action of  $\mathfrak{S}_n$  on  $\mathbb{R}[\mathcal{V}_n]$ , Gatermann and Parrilo [GP04] showed how to find simpler, symmetry-reduced, sos expressions for  $\mathbf{p}$  by exploiting the isotypic decomposition of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  under the action of  $\mathfrak{S}_n$ . In this section, we first introduce preliminary definitions and describe the structure of these symmetry-reduced sos expressions of Gatermann and Parrilo (Theorem 2.2). To search for these sos expressions via semidefinite programming requires knowledge of bases for certain subspaces  $W_{\tau_\lambda}$  of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$

(defined in Section 2.1). In Section 2.3, we show that the number of partitions  $\lambda$  of  $n$  for which these subspaces are non-zero is bounded by a quantity that depends on  $d$  but is independent of  $n$  (Corollary 2.5). In Section 2.4, we introduce the more flexible variant of Theorem 2.2 that does not require explicit knowledge of bases for the  $W_{\tau_\lambda}$ , but instead allows us to work with spanning sets for these spaces.

**2.1. Preliminaries.** Recall that the symmetric group  $\mathfrak{S}_n$  acts on  $\mathbb{R}[x]_{\leq d}$ , the ring of polynomials of degree at most  $d$  in  $\binom{n}{2}$  variables, by sending each variable  $x_{ij}$  to  $\mathfrak{s} \cdot x_{ij} = x_{\mathfrak{s}(i)\mathfrak{s}(j)}$  for each  $\mathfrak{s} \in \mathfrak{S}_n$ . This descends to an action on  $\mathbb{R}[\mathcal{V}_n]_{\leq d} = \mathbb{R}[x]_{\leq d} / \mathcal{I}_n$  because  $\mathcal{I}_n$  is invariant under the action of  $\mathfrak{S}_n$ . By extending these actions linearly, we can regard  $\mathbb{R}[x]_{\leq d}$  and  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  as  $\mathfrak{S}_n$ -modules.

We briefly summarize standard facts and terminology related to the representation theory of  $\mathfrak{S}_n$  (see, e.g., [Sag01, Chapters 1& 2]). A  $\mathfrak{S}_n$ -module  $U$  is *irreducible* if the only subspaces of  $U$  that are invariant under the action of  $\mathfrak{S}_n$  are  $\{0\}$  and  $U$ . The irreducible  $\mathfrak{S}_n$ -modules are indexed by *partitions* of  $n$ , namely sequences of positive integers  $\lambda = (\lambda_1, \dots, \lambda_k)$  such that  $\lambda_1 \geq \dots \geq \lambda_k > 0$  and  $\lambda_1 + \dots + \lambda_k = n$ . Each  $\lambda_i$  is called a *part*. We use the shorthand  $\lambda \vdash n$  to indicate that  $\lambda$  is a partition of  $n$ . We denote the irreducible  $\mathfrak{S}_n$ -module indexed by the partition  $\lambda \vdash n$  by  $S_\lambda$ . Its dimension is denoted by  $n_\lambda$ .

Any  $\mathfrak{S}_n$ -module  $V$  has an *isotypic decomposition*

$$(2) \quad V = \bigoplus_{\lambda \vdash n} V_\lambda,$$

which expresses  $V$  as a direct sum of  $\mathfrak{S}_n$ -modules  $V_\lambda$ , called the *isotypic components*. Each  $V_\lambda$  is the span of all possible isomorphic copies of the irreducible  $\mathfrak{S}_n$ -module  $S_\lambda$  in  $V$ . (We define these notions precisely in Appendix A.)

Combinatorial objects, called tableaux, and related subgroups of  $\mathfrak{S}_n$ , called row groups, play an important role in the representation theory of  $\mathfrak{S}_n$ , and appear throughout this paper. A partition  $\lambda \vdash n$  has a *shape* (or *Young diagram*) with rows of size  $\lambda_1 \geq \dots \geq \lambda_k$ . A *tableau* of shape  $\lambda$ , denoted  $\tau_\lambda$ , is a filling of the  $n$  boxes in the diagram of  $\lambda$  with the numbers  $1, 2, \dots, n$  bijectively. A tableau is *standard* if the numbering increases from left to right along each row and down each column. The number of standard tableaux of shape  $\lambda$  is  $n_\lambda$ , the dimension of the irreducible  $\mathfrak{S}_n$ -module  $S_\lambda$  [Sag01, Theorem 2.5.2].

If  $\tau_\lambda$  is a tableau of shape  $\lambda$ , let  $\text{row}_i(\tau_\lambda)$  denote the set of labels in the  $i$ th row of  $\tau_\lambda$ . The *row group*  $\mathfrak{R}_{\tau_\lambda}$  of the tableau  $\tau_\lambda$  is the subgroup of  $\mathfrak{S}_n$  that leaves each  $\text{row}_i(\tau_\lambda)$  invariant, i.e.,

$$\mathfrak{R}_{\tau_\lambda} := \{\mathfrak{s} \in \mathfrak{S}_n : \mathfrak{s} \cdot \text{row}_i(\tau_\lambda) = \text{row}_i(\tau_\lambda) \text{ for all } i = 1, 2, \dots, k\}.$$

If  $U$  is a  $\mathfrak{S}_n$ -module and  $\tau_\lambda$  is a tableau, let  $U^{\mathfrak{R}_{\tau_\lambda}}$  denote the subspace of  $U$  consisting of points fixed by the row group  $\mathfrak{R}_{\tau_\lambda}$ , i.e.,

$$U^{\mathfrak{R}_{\tau_\lambda}} := \{u \in U : \mathfrak{s} \cdot u = u \text{ for all } \mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}\}.$$

The following definition plays a central role in the paper.

**Definition 2.1.** If  $\mathbb{R}[\mathcal{V}_n]_{\leq d} = \bigoplus_{\lambda} V_\lambda$  is the isotypic decomposition of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ , and  $\tau_\lambda$  is a tableau of shape  $\lambda$ , define

$$W_{\tau_\lambda} := V_\lambda^{\mathfrak{R}_{\tau_\lambda}}$$

to be the subspace of the isotypic  $V_\lambda$  fixed by the action of the row group  $\mathfrak{R}_{\tau_\lambda}$ .

Since  $V_\lambda$  is a subspace of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ , it follows that  $W_{\tau_\lambda}$  is a subspace of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}^{\mathfrak{R}_{\tau_\lambda}}$ , the subspace of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  that is fixed by the action of the row group  $\mathfrak{R}_{\tau_\lambda}$ . This simple observation is the main property of  $W_{\tau_\lambda}$  that we use in subsequent sections of the paper.

In Lemma A.10 of Appendix A, we show that for any choice of tableau  $\tau_\lambda$  of shape  $\lambda$ , there is a vector space isomorphism between  $W_{\tau_\lambda}$  and the *multiplicity space* of the irreducible  $\mathfrak{S}_n$ -module  $S_\lambda$  in  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  (defined in Appendix A). As such,  $\dim(W_{\tau_\lambda})$  is the same for any tableau  $\tau_\lambda$  of shape  $\lambda$ . We denote this dimension by  $m_\lambda$ , and often refer to  $m_\lambda$  as the *multiplicity* of  $S_\lambda$  in  $V$ . Some of the multiplicities  $m_\lambda$  may be zero as we will see in Section 2.3.

**2.2. Symmetry-reduced sos expressions of Gatermann and Parrilo.** We are now in a position to state a version of the main result of [GP04] tailored to our setting. This result tells us the structure of symmetry-reduced sos expressions for  $\mathfrak{S}_n$ -invariant  $d$ -sos polynomials, and shows that we can search for such sos expressions by solving a SDP of size  $\sum_{\lambda \vdash n} m_\lambda$ .

**Theorem 2.2.** [GP04] *Suppose  $\mathfrak{p} \in \mathbb{R}[\mathcal{V}_n]$  is  $\mathfrak{S}_n$ -invariant and  $d$ -sos. For each partition  $\lambda \vdash n$ , fix a tableau  $\tau_\lambda$  of shape  $\lambda$  and choose a vector space basis  $\{\mathfrak{b}_1^{\tau_\lambda}, \dots, \mathfrak{b}_{m_\lambda}^{\tau_\lambda}\}$  for  $W_{\tau_\lambda}$ . Then for each partition  $\lambda$  of  $n$ , there exists a  $m_\lambda \times m_\lambda$  psd matrix  $Q_\lambda$  such that*

$$(3) \quad \mathfrak{p} = \sum_{\lambda \vdash n} \text{tr}(Q_\lambda Y^{\tau_\lambda}),$$

where  $Y_{ij}^{\tau_\lambda} := \text{sym}(\mathfrak{b}_i^{\tau_\lambda} \mathfrak{b}_j^{\tau_\lambda})$ .

We provide a full proof of Theorem 2.2 and the background needed in Appendix A. Here, we make various comments about the statement of this theorem.

Notice that the matrix  $Y^{\tau_\lambda}$  in the statement of Theorem 2.2 is filled with polynomials in  $\mathfrak{x}$ , and that it is psd for all evaluations of  $\mathfrak{x}$  on  $\mathcal{V}_n$ . Hence the right-hand side of (3) is clearly non-negative as a function on  $\mathcal{V}_n$ . The content of Theorem 2.2 is that all  $d$ -sos  $\mathfrak{S}_n$ -invariant polynomials have certificates of non-negativity of this form. Our statement of Theorem 2.2 does not require a full symmetry-adapted basis (see [FS92]) for  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$ . Instead it requires, for each partition  $\lambda$ , a basis for  $W_{\tau_\lambda}$  for a single fixed tableau of shape  $\lambda$ . That is why we have a SDP of size  $\sum_{\lambda \vdash n} m_\lambda$  rather than  $\dim(V) = \sum_{\lambda \vdash n} m_\lambda n_\lambda$ .

We also note that the original Gatermann-Parrilo method is about sos certificates for globally non-negative invariant polynomials, but it is easily adapted to the situation of non-negativity over an algebraic variety. In addition, [GP04] gives a more refined symmetry-reduced sos result in terms of invariant theory.

In the next two subsections, we will see two important improvements to the statement of Theorem 2.2 that will eventually establish one of our main results that, whenever  $\mathfrak{p}$  is symmetric and  $d$ -sos, then  $\mathfrak{p}$  has a succinct sos expression whose size is independent of  $n$ .

**2.3. Partitions needed in the Gatermann-Parrilo sos.** Recall that the vector space  $W_{\tau_\lambda}$  has dimension  $m_\lambda$  and contributes the sos expression  $\text{tr}(Q_\lambda Y^{\tau_\lambda})$  to the sos certificate of  $\mathfrak{p}$  in (3) where  $Q_\lambda$  has size  $m_\lambda \times m_\lambda$ . We now investigate which multiplicities  $m_\lambda$  are non-zero, or equivalently, which partitions are needed in the Gatermann-Parrilo symmetry-reduced sos expression. We show that the number of partitions that are needed in (3) depends on  $d$  but is independent of  $n$ .

We begin by investigating which partitions  $\lambda \vdash n$  can appear in the isotypic decomposition of  $\mathbb{R}[x]_{\leq d}$ . Our main tool is the following consequence of Young's rule, established in [RST15, Theorem 4.9]. In what follows, we write  $\lambda \geq_{\text{lex}} \mu$  for  $\lambda, \mu \vdash n$  if  $\lambda$  is lexicographically greater than or equal to  $\mu$ , and  $\lambda \supseteq \mu$  if  $\lambda \geq_{\text{lex}} \mu$  and, in addition,  $\lambda$  has at most as many parts as  $\mu$ .

**Lemma 2.3.** *If  $V$  is a finite-dimensional  $\mathfrak{S}_n$ -module and  $\tau_\mu$  is a tableau of shape  $\mu$ , then*

$$V^{\mathfrak{R}_{\tau_\mu}} \subseteq \bigoplus_{\lambda \supseteq \mu} V_\lambda.$$

Recall that a *hook partition* of  $n$  is one in which all parts except the first has size one, which we denote as  $(\lambda_1, 1^{n-\lambda_1})$  where  $\lambda_1$  is the size of the first part, and where the remaining  $n - \lambda_1$  parts constitute the *tail* of the hook. We now use Lemma 2.3 to show that only partitions lexicographically at least as large as the hook partition  $(n-2d, 1^{2d})$  can appear in the isotypic decomposition of  $V = \mathbb{R}[x]_{\leq d}$ . In Proposition 2.7, we will see that the number of such partitions is bounded by a function of  $d$  that is independent of  $n$ .

**Theorem 2.4.** *The multiplicity  $m_\lambda$  of  $S_\lambda$  in the decomposition of  $V = \mathbb{R}[x]_{\leq d}$  into irreducible  $\mathfrak{S}_n$ -modules is zero unless  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ , i.e.,*

$$V = \bigoplus_{\lambda \geq_{\text{lex}} (n-2d, 1^{2d})} V_\lambda.$$

*Proof.* Let  $x_{i_1 j_1}^{\alpha_1} x_{i_2 j_2}^{\alpha_2} \cdots x_{i_k j_k}^{\alpha_k}$  be a monomial of degree  $\sum_{i=1}^k \alpha_i \leq d$ . The set  $L = \{i_1, j_1, i_2, j_2, \dots, i_k, j_k\} \subseteq [n]$  of indices appearing in the monomial has size at most  $2d$ . Let  $\tau$  be a tableau of shape  $(n-2d, 1^{2d})$  with the elements of  $L$  put into the tail and the remaining  $n - |L|$  labels filled arbitrarily into the rest of the Young diagram of  $\tau$ . Since  $\mathfrak{R}_\tau$  fixes every element of  $L$ , clearly  $x_{i_1 j_1}^{\alpha_1} x_{i_2 j_2}^{\alpha_2} \cdots x_{i_k j_k}^{\alpha_k} \in V^{\mathfrak{R}_\tau}$ . Since  $V$  is spanned by monomials of degree at most  $d$ , we have that

$$V \subseteq \sum_{\text{shape}(\tau)=(n-2d, 1^{2d})} V^{\mathfrak{R}_\tau} \subseteq \bigoplus_{\lambda \supseteq (n-2d, 1^{2d})} V_\lambda,$$

where the second inclusion follows from Lemma 2.3. To finish the proof, we need only show that for the hook shape  $(n-2d, 1^{2d})$ , we have  $\lambda \supseteq (n-2d, 1^{2d})$  if and only if  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ . It is enough to argue that if  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ , then  $\lambda$  has at most as many parts as  $(n-2d, 1^{2d})$ , i.e.,  $\lambda$  has at most  $2d+1$  parts. This is clearly the case since  $\lambda \vdash n$  and  $\lambda_1 \geq n-2d$ .  $\square$

By the same argument (but only considering square-free monomials), the same conclusion holds for the isotypic decomposition of  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$ .

**Corollary 2.5.** *The multiplicity  $m_\lambda$  of  $S_\lambda$  in the decomposition of  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$  into irreducible  $\mathfrak{S}_n$ -modules or, equivalently, the dimension of  $W_{\tau_\lambda}$  for any tableau of shape  $\lambda$ , is zero unless  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ , i.e.,*

$$V = \bigoplus_{\lambda \geq_{\text{lex}} (n-2d, 1^{2d})} V_\lambda.$$

For the rest of the body of the paper, we will fix

$$\Lambda := \{\lambda \vdash n : \lambda \geq_{\text{lex}} (n-2d, 1^{2d})\}.$$

Then, using Theorem 2.4, we obtain the following corollary to Theorem 2.2.

**Corollary 2.6.** *Suppose  $\mathfrak{p} \in \mathbb{R}[\mathcal{V}_n]$  is  $\mathfrak{S}_n$ -invariant and  $d$ -sos. For each partition  $\lambda \vdash n$ , fix a tableau  $\tau_\lambda$  of shape  $\lambda$  and choose a vector space basis  $\{\mathbf{b}_1^{\tau_\lambda}, \dots, \mathbf{b}_{m_\lambda}^{\tau_\lambda}\}$  for  $W_{\tau_\lambda}$ . Then for each partition  $\lambda \in \Lambda$ , there exists a  $m_\lambda \times m_\lambda$  psd matrix  $Q_\lambda$  such that*

$$\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(Q_\lambda Y^{\tau_\lambda}).$$

We now give an explicit upper bound on the number of partitions that are lexicographically greater than or equal to the hook  $(n-2d, 1^{2d})$ , and hence a bound on the number of partitions  $\lambda$  that appear in the isotypic decomposition of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ .

**Proposition 2.7.** *Let  $m_\lambda$  be the multiplicity of  $S_\lambda$  in the isotypic decomposition of  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$ . Then the number of partitions  $\lambda$  of  $n$  such that  $m_\lambda$  is non-zero is bounded above by  $p(0) + p(1) + p(2) + \dots + p(2d)$ , where  $p(i)$  is the number of partitions of  $i$ . In particular, the size of  $\Lambda$  is independent of  $n$ .*

*Proof.* By Corollary 2.5, the number of partitions  $\lambda$  such that  $m_\lambda$  is non-zero is bounded above by  $|\Lambda|$ . If  $\lambda \in \Lambda$ , then it has the form  $(n-2d+i, \mu)$ , where  $\mu \vdash 2d-i$  and  $i$  is some integer satisfying  $0 \leq i \leq 2d$ . The number of such partitions  $\mu$  is bounded above by  $p(0) + p(1) + p(2) + \dots + p(2d)$ .  $\square$

By Proposition 2.7, we see that even though the number of partitions of  $n$  grows with  $n$ , the number of those that are needed in the Gatermann-Parrilo symmetry-reduced sos expression for a  $d$ -sos symmetric polynomial  $\mathfrak{p}$  is bounded by a function of  $d$  that is independent of  $n$ . This is our first improvement to the statement of Theorem 2.2 in our setting.

**2.4. Symmetry-reduction via spanning sets.** Next, we show that one does not need bases for the  $W_{\tau_\lambda}$  to find a sos certificate as in Theorem 2.2. Instead, it suffices to have a set of polynomials whose span contains  $W_{\tau_\lambda}$ . This relaxation offers a great deal of flexibility in setting up the SDP in (3).

**Theorem 2.8.** *Suppose  $\mathfrak{p} \in \mathbb{R}[\mathcal{V}_n]$  is  $\mathfrak{S}_n$ -invariant and  $d$ -sos. For each partition  $\lambda \vdash n$ , fix a tableau  $\tau_\lambda$  of shape  $\lambda$  and let  $\{\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{l_{\tau_\lambda}}^{\tau_\lambda}\}$  be a set of polynomials whose span contains  $W_{\tau_\lambda}$ . Then for each partition  $\lambda \in \Lambda$ , there exists a  $l_{\tau_\lambda} \times l_{\tau_\lambda}$  psd matrix  $R_\lambda$  such that*

$$\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(R_\lambda Z^{\tau_\lambda}),$$

where  $Z_{ij}^{\tau_\lambda} := \text{sym}(\mathfrak{p}_i^{\tau_\lambda} \mathfrak{p}_j^{\tau_\lambda})$ .

*Proof.* By Corollary 2.6, we know that under the hypotheses of the theorem, there exist  $m_\lambda \times m_\lambda$  real psd matrices  $Q_\lambda$  such that  $\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(Q_\lambda Y^{\tau_\lambda})$ . Recall that the matrix  $Y^{\tau_\lambda}$ , defined as  $Y_{ij}^{\tau_\lambda} = \text{sym}(\mathbf{b}_i^{\tau_\lambda} \mathbf{b}_j^{\tau_\lambda})$ , is constructed from a basis for  $W_{\tau_\lambda}$  after fixing one tableau  $\tau_\lambda$  of shape  $\lambda$ . Since each  $\mathbf{b}_i^{\tau_\lambda}$  is in the span of  $\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{l_{\tau_\lambda}}^{\tau_\lambda}$ , there is a real matrix  $M_{\tau_\lambda}$  of size  $l_{\tau_\lambda} \times m_\lambda$  such that

$$(\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{l_{\tau_\lambda}}^{\tau_\lambda}) M_{\tau_\lambda} = (\mathbf{b}_1^{\tau_\lambda}, \dots, \mathbf{b}_{m_\lambda}^{\tau_\lambda}).$$

Defining  $R_\lambda := M_{\tau_\lambda} Q_\lambda M_{\tau_\lambda}^\top$ , we see that

$$(\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{m_\lambda}^{\tau_\lambda}) Q_\lambda (\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{m_\lambda}^{\tau_\lambda})^\top = (\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{l_{\tau_\lambda}}^{\tau_\lambda}) R_\lambda (\mathfrak{p}_1^{\tau_\lambda}, \dots, \mathfrak{p}_{l_{\tau_\lambda}}^{\tau_\lambda})^\top.$$



Symmetrizing both sides we conclude that  $\text{tr}(Q_\lambda Y^{\tau_\lambda}) = \text{tr}(R_\lambda Z^{\tau_\lambda})$ , and hence,

$$\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(Q_\lambda Y^{\tau_\lambda}) = \sum_{\lambda \in \Lambda} \text{tr}(R_\lambda Z^{\tau_\lambda}).$$

□

Note that the sos in Theorem 2.8 retains the same structure as Corollary 2.6 in the sense that for each  $\lambda \in \Lambda$  we only need to consider  $W_{\tau_\lambda}$  for a single tableau  $\tau_\lambda$  of shape  $\lambda$ .

### 3. SPANNING SETS AND RAZBOROV'S FLAGS

Recall that our goal is to find succinct sos expressions for  $\mathfrak{S}_n$ -invariant  $d$ -sos polynomials. Theorem 2.8 implies that to achieve this, one needs to find, for a fixed tableau  $\tau_\lambda$  of shape  $\lambda \in \Lambda$ , a spanning set for  $W_{\tau_\lambda}$  with size independent of  $n$ . In Section 3.1, we show how to construct a naïve spanning set for  $W_{\tau_\lambda}$  with the property that the number of polynomials in it is independent of  $n$ . From that first spanning set, we derive two more spanning sets that have the same property, and which naturally introduce the notion of flags introduced by Razborov in [Raz07]. This connection to flags provides a nice combinatorial interpretation for these spanning sets which will be explained in Section 3.2.

**3.1. Spanning sets.** Recall that the vector space  $W_{\tau_\lambda}$  is a subspace of  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$  which is spanned by square-free monomials in  $\binom{n}{2}$  variables of degree at most  $d$ . We denote such a square-free monomial by  $x^m := \prod_{1 \leq i < j \leq n} x_{ij}^{m_{ij}}$  with  $m_{ij} \in \{0, 1\}$ .

A very naïve spanning set for  $W_{\tau_\lambda}$  can easily be generated as follows. For a tableau  $\tau_\lambda$  of shape  $\lambda = (\lambda_1, \lambda_2, \dots)$  and a monomial  $x^m$ , let

$$\text{sym}_{\tau_\lambda}(x^m) := \frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \sum_{\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}} \mathfrak{s} \cdot x^m$$

be the symmetrization of  $x^m$  under the row group  $\mathfrak{R}_{\tau_\lambda}$ . The polynomials  $\text{sym}_{\tau_\lambda}(x^m)$  as  $x^m$  varies over square-free monomials of degree at most  $d$  form a natural spanning set for  $W_{\tau_\lambda}$ . Indeed, recall that  $W_{\tau_\lambda}$  consists of  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomials. Therefore, if  $\mathfrak{p} \in W_{\tau_\lambda}$ , then  $\mathfrak{p} = \text{sym}_{\tau_\lambda}(\mathfrak{p})$  and so  $\mathfrak{p}$  is a linear combination of  $\text{sym}_{\tau_\lambda}(x^m)$  for different square-free monomials  $x^m$  of degree at most  $d$ .

We will show in Lemma 3.2 that the symmetrized monomials  $\text{sym}_{\tau_\lambda}(x^m)$  are superfluous when  $\lambda$  is not a hook. Indeed, for any  $\tau_\lambda$ , even when  $\lambda$  is not a hook,  $W_{\tau_\lambda}$  is in the span of symmetrizations of monomials under row groups of hooks.

**Definition 3.1.** Given a tableau  $\tau_\lambda$  of shape  $\lambda = (\lambda_1, \lambda_2, \dots)$ , define  $\text{hook}(\tau_\lambda)$  to be the tableau of shape  $(\lambda_1, 1^{n-\lambda_1})$  such that the first row of  $\text{hook}(\tau_\lambda)$  is the same as that of  $\tau_\lambda$ , and the remaining labels in  $\tau_\lambda$  are put in the tail of the hook in increasing order.

**Lemma 3.2.** *For the tableau  $\tau_\lambda$ , the vector space  $W_{\tau_\lambda}$  is spanned by the polynomials  $\text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , as  $x^m$  varies over square-free monomials of degree at most  $d$ .*

*Proof.* From the definition of  $W_{\tau_\lambda}$ , we have that  $W_{\tau_\lambda} = V_{\lambda}^{\mathfrak{R}_{\tau_\lambda}} \subseteq V^{\mathfrak{R}_{\tau_\lambda}}$ , the subspace of  $\mathfrak{R}_{\tau_\lambda}$ -fixed elements of  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$ . Since  $\mathfrak{R}_{\text{hook}(\tau_\lambda)}$  is a subgroup of  $\mathfrak{R}_{\tau_\lambda}$ , it follows directly that  $V^{\mathfrak{R}_{\tau_\lambda}} \subseteq V^{\mathfrak{R}_{\text{hook}(\tau_\lambda)}}$ . Hence  $W_{\tau_\lambda} \subseteq V^{\mathfrak{R}_{\text{hook}(\tau_\lambda)}}$ , which is certainly spanned by the polynomials  $\text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , as  $x^m$  varies over square-free monomials of degree at most  $d$ . □

One can think about these hook polynomials in a combinatorial manner starting with the following definition.

**Definition 3.3.** The *graph* of a square-free monomial  $x^m$ , denoted as  $G(x^m)$ , is the labeled subgraph of the complete graph  $K_n$  containing the edge  $\{i, j\}$  if  $m_{ij} = 1$ .

Therefore, for any square-free monomial  $x^m$  and any tableau  $\tau_\lambda$  of shape  $\lambda = (\lambda_1, \lambda_2, \dots)$ , we have that

$$\text{sym}_{\text{hook}(\tau_\lambda)}(x^m) = \frac{1}{|\mathfrak{R}_{\text{hook}(\tau_\lambda)}|} \sum_{\sigma \in \mathfrak{R}_{\text{hook}(\tau_\lambda)}} \prod_{\{i, j\} \in E(G(x^m))} x_{\sigma(i)\sigma(j)}.$$

Note that any vertex of  $G(x^m)$  that is in the tail of  $\text{hook}(\tau_\lambda)$  is fixed by  $\mathfrak{R}_{\text{hook}(\tau_\lambda)}$ . This partitions the vertices of  $G(x^m)$  into two sets: those fixed by  $\mathfrak{R}_{\text{hook}(\tau_\lambda)}$ , and those that are not. We now see how this phenomenon is mirrored by flags, partially labeled graphs where the labeled part is fixed. We can thus view graphs  $G(x^m)$  through the lens of flags and show that  $\text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$  can be thought of as polynomials coming from flags instead. To formalize this connection, we introduce flags and density polynomials arising from flags and then show their equivalence to the symmetric polynomials defined above.

**Definition 3.4.** Consider three fixed integers  $0 \leq t \leq f \leq n$ .

- (1) An *intersection type* of size  $t$  is a simple graph  $T$  on  $t$  vertices in which every vertex is labeled with a distinct element of  $[t]$ .
- (2) A  *$T$ -flag*  $F$  of size  $f$  is a simple graph on  $f$  vertices with  $t$  of its vertices labeled  $1, \dots, t$  such that these labeled vertices induce a copy of  $T$  in  $F$  with identical labels for the vertices.
- (3) Let  $\mathcal{F}_T^f$  be the set of all  $T$ -flags of size  $f$  up to isomorphism.

**Example 3.5.** Consider  $T = 2 \bullet \overset{1}{\curvearrowright} 3$  and  $f = 4$ . Then  $\mathcal{F}_T^f$  consists of the  $T$ -flags:

$$\begin{aligned} F_0 &= 2 \bullet \overset{1}{\curvearrowright} 3, F_1 = 2 \bullet \overset{1}{\curvearrowleft} 3, F_2 = 2 \overset{1}{\curvearrowright} 3, F_3 = 2 \overset{1}{\curvearrowleft} 3, \\ F_4 &= 2 \overset{1}{\curvearrowright} 3, F_5 = 2 \overset{1}{\curvearrowleft} 3, F_6 = 2 \overset{1}{\curvearrowright} 3, F_7 = 2 \overset{1}{\curvearrowleft} 3. \end{aligned}$$

**Definition 3.6.** For two sets  $A$  and  $B$ , let  $\text{Inj}(A, B)$  denote the set of injective maps  $h : A \rightarrow B$ . Suppose  $F$  is a  $T$ -flag of size  $f$  and  $\Theta \in \text{Inj}([t], [n])$ . We say that  $h \in \text{Inj}(V(F), [n])$  respects the labeling  $\Theta$  if  $h(v) = \Theta(i)$  for any vertex  $v \in V(F)$  labeled  $i \in [t]$ . Let  $\text{Inj}_\Theta(V(F), [n])$  denote the set of all injective maps  $h : V(F) \rightarrow [n]$  that respect the labeling  $\Theta$ .

**Example 3.7.** Suppose  $T = 2 \bullet \overset{1}{\curvearrowright} 3$  and  $F = 2 \overset{1}{\curvearrowright} 3$ . Identify  $[n]$  with  $\{a_1, a_2, a_3, \dots\}$ . Fix  $\Theta \in \text{Inj}([3], [n])$  to be the map  $\Theta(1) = a_2$ ,  $\Theta(2) = a_1$  and  $\Theta(3) = a_4$ . Let  $v$  denote the unlabeled vertex in  $F$ . Then  $h \in \text{Inj}(V(F), [n])$  such that  $h(1) = a_2$ ,  $h(2) = a_1$ ,  $h(3) = a_4$ ,  $h(v) \in [n] \setminus \{a_1, a_2, a_4\}$  is  $\Theta$ -preserving, while  $h$  such that  $h(1) = a_3$ ,  $h(2) = a_1$ ,  $h(3) = a_4$ ,  $h(v) = a_2$  is not.

**Definition 3.8.** Fix  $T, f, \Theta$ . Then for every flag  $F \in \mathcal{F}_T^f$ , we let

$$\mathfrak{g}_F^\Theta = \sum_{h \in \text{Inj}_\Theta(V(F), [n])} \prod_{\{i, j\} \in E(F)} x_{h(i)h(j)}.$$

Note that  $\mathfrak{g}_F^\Theta$  is always square-free since  $F$  is assumed to be a simple graph. So the polynomials  $\mathfrak{g}_F^\Theta$  cannot span  $\mathbb{R}[x]_{\leq d}$  but can possibly span  $\mathbb{R}[\mathcal{V}_n]_{\leq d}$ .

**Example 3.9.** Consider our running example with the labeling  $\Theta$  fixed to be  $\Theta(1) = a_i$ ,  $\Theta(2) = a_j$  and  $\Theta(3) = a_k$  where  $a_i, a_j, a_k \in [n]$  are distinct and fixed.

$F$	$\mathfrak{g}_F^\Theta$
$F_0 = \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ 2 \quad 3 \end{array}$	$(n-3) \times_{ij} \times_{ik}$
$F_1 = \begin{array}{c} \bullet \\ \swarrow \quad \uparrow \quad \searrow \\ 2 \quad 1 \quad 3 \end{array}$	$\times_{ij} \times_{ik} \sum_{l \in [n] \setminus \{i,j,k\}} \times_{il}$
$F_4 = \begin{array}{c} \bullet \\ \swarrow \quad \uparrow \quad \downarrow \quad \searrow \\ 2 \quad 1 \quad 3 \end{array}$	$\times_{ij} \times_{ik} \sum_{l \in [n] \setminus \{i,j,k\}} \times_{il} \times_{jl}$
$F_7 = \begin{array}{c} \bullet \\ \swarrow \quad \uparrow \quad \downarrow \quad \uparrow \quad \searrow \\ 2 \quad 1 \quad 3 \end{array}$	$\times_{ij} \times_{ik} \sum_{l \in [n] \setminus \{i,j,k\}} \times_{il} \times_{jl} \times_{kl}$

The polynomials  $\mathfrak{g}_F^\Theta$  can be interpreted in terms of tableaux and their row groups.

**Definition 3.10.** For  $\Theta \in \text{Inj}([t], [n])$ , let  $\text{hook}^\Theta$  be the tableau of shape  $(n-t, 1^t)$  where the labels in  $\Theta([t])$  are placed in increasing order down the tail and the remaining labels in  $[n]$  are put in increasing order in the first row.

**Example 3.11.** Let  $\Theta$  be as in Example 3.9 and assume  $a_i < a_j < a_k$ . Then

$$\text{hook}^\Theta = \begin{array}{|c|c|c|c|} \hline a_{l_1} & a_{l_2} & \cdots & a_{l_{n-3}} \\ \hline a_i & & & \\ \hline a_j & & & \\ \hline a_k & & & \\ \hline \end{array}$$

where  $a_{l_r} < a_{l_{r+1}}$  for all  $r$ , and  $\{a_{l_1}, \dots, a_{l_{n-3}}\} = [n] \setminus \{a_i, a_j, a_k\}$ .

Since  $\mathfrak{A}_{\text{hook}^\Theta}$  can be identified with the symmetric group on the labels  $[n] \setminus \Theta([t])$ , we immediately get the following connection between  $\mathfrak{g}_F^\Theta$  and  $\text{sym}_{\text{hook}^\Theta}(\mathbf{x}^m)$ .

**Lemma 3.12.** Fix any  $F \in \mathcal{F}_T^f$  and  $\Theta \in \text{Inj}([t], [n])$ . Then the polynomial  $\mathfrak{g}_F^\Theta$  is equal to  $\frac{(n-t)!}{(n-f)!} \cdot \text{sym}_{\text{hook}^\Theta}(\mathbf{x}^m)$  where  $\mathbf{x}^m = \prod_{\{i,j\} \in E(F)} \times_{h^*(i)h^*(j)}$  for any fixed  $h^* \in \text{Inj}_\Theta(V(F), [n])$ .

*Proof.* This follows from the fact that  $|\text{Inj}_\Theta(V(F), [n])| = \binom{n-t}{f-t} (f-t)!$  and that, for each  $h \in \text{Inj}_\Theta(V(F), [n])$ , there are  $(n-f)!$  permutations  $\sigma \in \mathfrak{A}_{\text{hook}^\Theta}$  such that  $\sigma(h(\mathbf{x}^m)) = h(\mathbf{x}^m)$ .  $\square$

**Example 3.13.** From the previous example, we have that

$$\mathfrak{g}_{F_0}^\Theta = (n-3) \cdot \text{sym}_{\text{hook}^\Theta}(\times_{ij} \times_{ik}), \quad \mathfrak{g}_{F_1}^\Theta = (n-3) \cdot \text{sym}_{\text{hook}^\Theta}(\times_{ij} \times_{ik} \times_{il_1}).$$

Therefore, we will be able to reinterpret our naïve spanning set for a particular  $W_{\tau_\lambda}$  in terms of flags. Before doing so, we introduce a second set of flag-based polynomials. Applying a Möbius transformation [Lov12, A.1] on the  $\mathfrak{g}_F^\Theta$ 's, we get a new set of polynomials. In Subsection 3.2, we remark on their link to flag algebras and graph homomorphism densities.

**Definition 3.14.** (1) Define

$$\mathcal{F}_{\geq T}^f := \bigcup \left\{ F : F \text{ is a } T' \text{-flag of size } f \text{ where the vertices } T' \text{ are the same as those of } T \text{ (including the labels), and } E(T') \supseteq E(T) \right\},$$

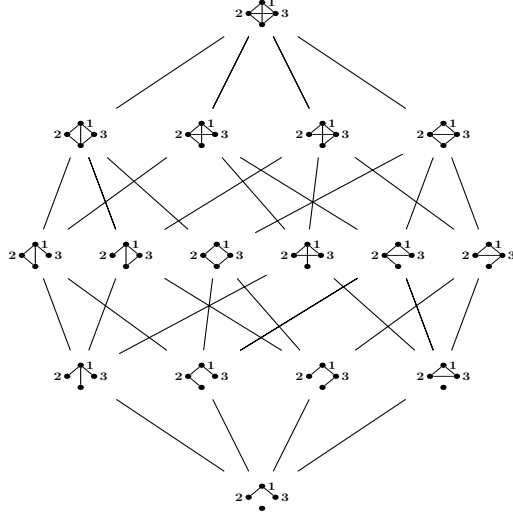
and let  $P_T^f$  be the poset on  $\mathcal{F}_{\geq T}^f$  where  $F \leq F'$  if  $E(F) \subseteq E(F')$ . This poset can be graded by the number of edges.

(2) For a flag  $F \in \mathcal{F}_{\geq T}^f$ , define the polynomial

$$d_F^\ominus := \sum_{\substack{F' \in \mathcal{P}_T^f: \\ F' \geq F}} (-1)^{|E(F')| - |E(F)|} \mathfrak{g}_{F'}^\ominus.$$

Note that unlike  $\mathcal{F}_T^f$ , the set  $\mathcal{F}_{\geq T}^f$  is not defined up to isomorphism; for example, it would contain both  $\bullet \begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \end{array} \bullet$  and  $\bullet \begin{array}{c} \bullet \\ \swarrow \\ \bullet \end{array} \bullet$  even though they are isomorphic flags.

**Example 3.15.** Consider  $T$  and  $f$  as before. They yield the following poset  $\mathcal{P}_T^f$ .



Continuing our example, for  $F_6 = 2 \begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \end{array} \bullet$ , the polynomial

$$d_{F_6}^\ominus = x_{ij}x_{ik} \sum_{l \in [n] \setminus \{i,j,k\}} (x_{jl}x_{kl} - x_{il}x_{jl}x_{kl} - x_{jk}x_{jl}x_{kl} + x_{jk}x_{il}x_{jl}x_{kl}).$$

**Definition 3.16.** Given a tableau  $\tau_\lambda$  of shape  $\lambda = (\lambda_1, \lambda_2, \dots)$ , define  $\Theta_{\tau_\lambda} \in \text{Inj}([n - \lambda_1], [n])$  such that  $\Theta_{\tau_\lambda}(i)$  is the  $i$ th smallest label not in the first row of  $\tau_\lambda$ .

**Lemma 3.17.** Let  $F \in \mathcal{F}_T^f$  and  $F' \in \mathcal{F}_T^{f'}$  where  $n \geq f' > f$  and  $E(F) = E(F')$ , i.e.,  $F'$  is obtained from  $F$  by adding isolated vertices. Then

$$\mathfrak{g}_{F'}^\ominus = \binom{n-f}{f'-f} (f'-f)! \cdot \mathfrak{g}_F^\ominus.$$

*Proof.* Since  $E(F) = E(F')$ ,

$$\begin{aligned} \mathfrak{g}_{F'}^\ominus &= \sum_{h \in \text{Inj}_\Theta(V(F'), [n])} \prod_{\{i,j\} \in E(F')} x_{h(i)h(j)} \\ &= \sum_{h \in \text{Inj}_\Theta(V(F'), [n])} \prod_{\{i,j\} \in E(F)} x_{h(i)h(j)} \\ &= \binom{n-f}{f'-f} (f'-f)! \sum_{h \in \text{Inj}_\Theta(V(F), [n])} \prod_{\{i,j\} \in E(F)} x_{h(i)h(j)}. \quad \square \end{aligned}$$

We will now show that for the vector space  $W_{\tau_\lambda}$ , the polynomials  $g_F^{\Theta_{\tau_\lambda}}$  and  $d_F^{\Theta_{\tau_\lambda}}$  form a spanning set as  $F$  varies over all graphs in  $\mathcal{F}_T^{2d}$  where  $T$  has  $n - \lambda_1$  vertices. From now on, we denote the number of vertices in  $T$  by  $|T|$ .

**Theorem 3.18.** *For the tableau  $\tau_\lambda$ , the vector space  $W_{\tau_\lambda}$  is spanned by*

- (1) the polynomials  $g_F^{\Theta_{\tau_\lambda}}$  with flags  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ , and
- (2) the polynomials  $d_F^{\Theta_{\tau_\lambda}}$  with flags  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ .

*Proof.* Recall from Lemma 3.2 that the polynomials  $\text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$  span  $W_{\tau_\lambda}$  as  $x^m$  varies over certain square-free monomials of degree at most  $d$ .

Since  $W_{\tau_\lambda}$  sits in the span of the polynomials  $\text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , it also sits in the span of  $\frac{|\mathfrak{A}_{\text{hook}(\tau_\lambda)}|}{(n-|V(G(x^m)))!} \cdot \text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , where  $x^m$  varies over square-free monomials of degree at most  $d$ . For a polynomial  $\frac{|\mathfrak{A}_{\text{hook}(\tau_\lambda)}|}{(n-|V(G(x^m)))!} \cdot \text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , consider the graph  $G(x^m)$  and note that, since  $x^m$  has degree at most  $d$ ,  $G(x^m)$  has at most  $2d$  vertices. Unlabel any vertex in  $G(x^m)$  whose label is in  $\text{row}_1(\text{hook}(\tau_\lambda))$ , and replace the labels of the remaining  $n - \lambda_1$  labeled vertices with the labels in  $[n - \lambda_1]$  through a map  $\phi$  that does so in increasing order (i.e., the smallest labeled vertex becomes vertex 1, and the largest labeled vertex becomes vertex  $n - \lambda_1$ ). Let  $F$  be that graph, and let  $\Theta = \phi^{-1}$ . Note that  $|V(F)| = |V(G(x^m))|$  and  $\Theta = \Theta_{\tau_\lambda}$ . Then, by Lemma 3.12,  $g_F^\Theta = \frac{|\mathfrak{A}_{\text{hook}(\tau_\lambda)}|}{(n-|V(G(x^m)))!} \cdot \text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ , and so  $W_{\tau_\lambda}$  sits in the span of  $g_F^\Theta$  for flags  $F \in \mathcal{F}_T^f$  where  $f \leq 2d$ ,  $|T| = n - \lambda_1$  and  $\Theta = \Theta_{\tau_\lambda}$ . See Example 3.20 for an illustration of the process just described.

By Lemma 3.17, one only needs to keep the polynomials  $g_F^\Theta$  for flags  $F \in \mathcal{F}_T^f$  where  $f = 2d$  and  $|T| = n - \lambda_1$  since, for any flag  $F^* \in \mathcal{F}_T^{f^*}$  where  $f^* < 2d$ , there exists a flag  $F^+$  with the same edge set but on  $2d$  vertices such that  $g_{F^+}^\Theta = k g_{F^*}^\Theta$  where  $k$  is a scalar.

The spans of polynomials  $g_F^\Theta$  and  $d_F^\Theta$  for  $F \in \mathcal{F}_{\leq T}^{2d}$  coincide since they are related by a Möbius transformation. Thus  $W_{\tau_\lambda}$  also sits in the span of the polynomials  $d_F^\Theta$  for flags  $F \in \mathcal{F}_T^{2d}$  with  $|T| = n - \lambda_1$  and where  $\Theta = \Theta_{\tau_\lambda}$ .  $\square$

*Remark 3.19.* Let  $\lambda, \lambda'$  be partitions of  $n$  with  $\lambda_1 = \lambda'_1$ , and let  $\tau_\lambda$  and  $\tau_{\lambda'}$  be two tableaux of shape  $\lambda$  and  $\lambda'$  such that  $\text{row}_1(\tau_\lambda) = \text{row}_1(\tau_{\lambda'})$  and thus  $\text{hook}(\tau_\lambda) = \text{hook}(\tau_{\lambda'})$ . Then, since  $\Theta_{\tau_\lambda} = \Theta_{\tau_{\lambda'}}$ , the vector spaces  $W_{\tau_\lambda}$  and  $W_{\tau_{\lambda'}}$  are both spanned by the polynomials  $g_F^{\Theta_{\tau_\lambda}}$  (respectively  $d_F^{\Theta_{\tau_\lambda}}$ ) with flags  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ . In particular, the polynomials in Theorem 3.18 that span  $W_{\text{hook}(\tau_\lambda)}$  also span  $W_{\tau_\lambda}$ .

**Example 3.20.** We give an example to illustrate the proof of the theorem above. Consider any tableau  $\tau_\lambda$  for which

$$\text{hook}(\tau_\lambda) = \begin{array}{|c|c|c|c|} \hline a_{l_1} & a_{l_2} & \cdots & a_{l_{n-3}} \\ \hline a_i & & & \\ \hline a_j & & & \\ \hline a_k & & & \\ \hline \end{array}$$

where  $a_i < a_j < a_k$  are fixed and  $\{a_{l_1}, \dots, a_{l_{n-3}}\} = [n] \setminus \{a_i, a_j, a_k\}$ . Now consider the monomial  $x^m = x_{ij} x_{ik} x_{jl} x_{kl}$ , for some fixed  $l_r$  such that  $1 \leq r \leq n - 3$ . Then

note that

$$\frac{|\mathfrak{R}_{\text{hook}(\tau_\lambda)}|}{(n - |V(G(x^m))|)!} \cdot \text{sym}_{\text{hook}(\tau_\lambda)}(x^m) = x_{ij}x_{ik} \sum_{l \in [n] \setminus \{i, j, k\}} x_{jl}x_{kl}.$$

The graph  $G(x^m)$  is  $j \begin{array}{c} \swarrow \\ \diamond \\ \searrow \end{array} i^k$ . Unlabeling the vertex labeled  $l_r$  since  $l_r$  is in the first row of  $\text{hook}(\tau_\lambda)$ , we get the graph  $j \begin{array}{c} \swarrow \\ \diamond \\ \searrow \end{array} i^k$ . Now we replace the labels  $i, j, k$  with  $1, 2, 3$  in that order to get  $F_6 = 2 \begin{array}{c} \swarrow \\ \diamond \\ \searrow \end{array} 3$ . Thus the map  $\phi$  in the proof of the lemma sends  $i \mapsto 1, j \mapsto 2, k \mapsto 3$ . Hence take  $\Theta$  to be the map that sends  $1 \mapsto i, 2 \mapsto j, 3 \mapsto k$  and  $F$  to be the flag  $F_6$ . Then the polynomial  $\mathbf{g}_{F_6}^\Theta$  is exactly  $(n-3) \cdot \text{sym}_{\text{hook}(\tau_\lambda)}(x^m)$ .

We now observe that, when  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ , we have been successful in generating spanning sets for  $W_{\tau_\lambda}$  whose sizes are independent of  $n$ .

**Proposition 3.21.** *Fix some tableau  $\tau_\lambda$  of shape  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ . Then the number of polynomials in each spanning set of  $W_{\tau_\lambda}$  described in Lemma 3.2 and Theorem 3.18 is independent of  $n$ .*

*Proof.* Suppose  $\lambda = (\lambda_1, \lambda_2, \dots)$ . Then, since  $\lambda \geq_{\text{lex}} (n-2d, 1^{2d})$ , we have that  $n - \lambda_1 \leq 2d$ . Thus, the polynomials  $\mathbf{g}_F^{\Theta, \tau_\lambda}$  (or  $\mathbf{d}_F^{\Theta, \tau_\lambda}$ ) with flags  $F \in \mathcal{F}_T^{2d}$  span  $W_{\tau_\lambda}$  as  $T$  varies over all intersection types on  $n - \lambda_1 \leq 2d$  vertices. Note that the number of intersection types on at most  $2d$  vertices does not depend on  $n$ . Moreover, note that the number of flags in  $\mathcal{F}_T^{2d}$  for any intersection type  $T$  such that  $|T| \leq 2d$  is independent of  $n$  by Definition 3.4. Thus the number of polynomials  $\mathbf{g}_F^{\Theta, \tau_\lambda}$  (or  $\mathbf{d}_F^{\Theta, \tau_\lambda}$ ) does not depend on  $n$ . By the bijection in Lemma 3.12, the result holds.  $\square$

**3.2. Relationship with flag and graph densities.** We now discuss how to interpret combinatorially the spanning sets involving  $\mathbf{g}_F^{\Theta, \tau_\lambda}$ 's or  $\mathbf{d}_F^{\Theta, \tau_\lambda}$ 's.

The *characteristic vector* of a graph  $G$ , denoted as  $\mathbf{1}_G$ , is the vector in  $\{0, 1\}^{\binom{[n]}{2}}$  whose  $ij$ -entry is one if and only if  $\{i, j\} \in E(G)$ . The polynomial  $\mathbf{g}_F^\Theta$ , when evaluated on  $\mathbf{1}_G$  for a graph  $G$  on  $n$  vertices, calculates the number ways of picking uniformly at random a set of  $t$  vertices of  $G$  to label by  $\Theta$  and  $f-t$  additional unlabeled vertices in such a way that they produce a copy of  $F$  in  $G$  (not necessarily as an induced subgraph). These numbers are at the core of the theory of graph homomorphisms. In fact, the polynomials  $\mathbf{g}_F^\Theta$  and  $\mathbf{d}_F^\Theta$ , and the relationship between them, exactly matches the graph densities  $t_{\text{inj}}$  and  $t_{\text{ind}}$  in the theory of graph limits [Lov12, Chapter 5].

On the other hand, we now observe that  $\mathbf{d}_F^\Theta(\mathbf{1}_G)$  for a graph  $G$  on  $n$  vertices, is the number of ways in which one can induce a copy of  $F$  in  $G$  by picking uniformly at random  $t$  vertices in  $G$  to label by  $\Theta$  along with  $f-t$  unlabeled vertices i.e.,  $\binom{n-t}{f-t} (f-t)!$  times the *flag density of  $F$  in  $G$* , as in the theory of flag algebras [Raz07]. This brings us to a combinatorial expression for the polynomial  $\mathbf{d}_F^\Theta$ .

**Lemma 3.22.** *The polynomial*

$$\mathbf{d}_F^\Theta = \sum_{h \in \text{Inj}_\Theta(V(F), [n])} \prod_{\{i, j\} \in E(F)} x_{h(i)h(j)} \prod_{\{i, j\} \in \binom{V(F)}{2} \setminus E(F)} (1 - x_{h(i)h(j)}).$$

*Proof.* Expanding the expression on the right-hand side we obtain

$$\sum_{h \in \text{Inj}_\Theta(V(F), [n])} \sum_{F' \supseteq F} a_{F'} \prod_{\{i, j\} \in E(F')} x_{h(i)h(j)},$$

where  $a_{F'} = (-1)^{|E(F')| - |E(F)|} \cdot |\{F'' \mid F'' \text{ is isomorphic to } F'\}|$  and  $F' \supseteq F$  is a graph on the labeled and unlabeled vertices of  $F$  and where any edge in  $F$  is also an edge in  $F'$ . In other words,  $F' \in \mathcal{F}_{\geq T}^f$  where  $T$  is the intersection type of  $F$  and  $f = |V(F)|$ , and the result holds.  $\square$

**Example 3.23.** Recall from Example 3.15 that

$$d_{F_6}^{\ominus} = x_{ij}x_{ik} \sum_{l \in [n] \setminus \{i,j,k\}} (x_{jl}x_{kl} - x_{il}x_{jl}x_{kl} - x_{jk}x_{jl}x_{kl} + x_{jk}x_{il}x_{jl}x_{kl}).$$

However, this expression can be rewritten to see that

$$d_{F_6}^{\ominus} = \sum_{l \in [n] \setminus \{i,j,k\}} x_{ij}x_{ik}x_{jl}x_{kl}(1 - x_{il})(1 - x_{jk}).$$

#### 4. FLAG SUMS OF SQUARES FOR THE HYPERCUBE

Applying Theorem 2.8 to the spanning sets for  $W_{\tau_{\lambda}}$  consisting of the  $d_F^{\ominus}$  and  $g_F^{\ominus}$  polynomials from the previous section, we immediately get the following corollaries.

**Corollary 4.1.** *Suppose  $\mathfrak{p}$  is symmetric and  $d$ -sos. For each partition  $\lambda \in \Lambda$ , fix a tableau  $\tau_{\lambda}$  of shape  $\lambda$ . Then there exists psd matrices  $R_{\lambda}$  such that*

$$\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(R_{\lambda} Z^{\tau_{\lambda}}),$$

where  $Z^{\tau_{\lambda}} := \text{sym}(\mathbf{g}_{\tau_{\lambda}} \mathbf{g}_{\tau_{\lambda}}^{\top})$  and  $\mathbf{g}_{\tau_{\lambda}}$  is the vector of polynomials  $g_F^{\ominus, \tau_{\lambda}}$  such that  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ .

**Corollary 4.2.** *Suppose  $\mathfrak{p}$  is symmetric and  $d$ -sos. For each partition  $\lambda \in \Lambda$ , fix a tableau  $\tau_{\lambda}$  of shape  $\lambda$ . Then there exists psd matrices  $R_{\lambda}$  such that*

$$\mathfrak{p} = \sum_{\lambda \in \Lambda} \text{tr}(R_{\lambda} Z^{\tau_{\lambda}}),$$

where  $Z^{\tau_{\lambda}} := \text{sym}(\mathbf{d}_{\tau_{\lambda}} \mathbf{d}_{\tau_{\lambda}}^{\top})$  and  $\mathbf{d}_{\tau_{\lambda}}$  is the vector of polynomials  $d_F^{\ominus, \tau_{\lambda}}$  such that  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ .

By Proposition 3.21, the size of each  $R_{\lambda}$ , which equals the size of the spanning set being used for  $W_{\tau_{\lambda}}$ , is independent of  $n$ . By Proposition 2.7, the number of partitions  $\lambda \in \Lambda$  is also independent of  $n$ . Therefore, the sizes of the SDPs that need to be solved in Corollaries 4.1 and 4.2 are independent of  $n$ . Thus, we have established that symmetric non-negative polynomials over the discrete hypercube  $\mathcal{V}_n$  have succinct sos certificates that come from the flag polynomials  $d_F^{\ominus}$  and their cousins  $g_F^{\ominus}$ . In particular, flag algebra techniques extend beyond the realm of extremal graph theory and can be used to establish symmetric inequalities on  $\mathcal{V}_n$ .

The sos expressions that typically appear in the flag algebra literature, e.g., [Raz10], [Raz14], [FRV13], are more restricted than those in Corollary 4.2. The main result of this section is to show that even these rather special looking sos expressions, which we will refer to as *flag sos* expressions, are sufficient to establish the non-negativity of symmetric polynomials over  $\mathcal{V}_n$ . The first step in passing to these restricted sos expressions is the following observation.

**Lemma 4.3.** *For two flags  $F$  and  $F'$  with different intersection types of equal size  $t$ , and any labeling  $\Theta \in \text{Inj}([t], [n])$ ,  $d_F^{\ominus} d_{F'}^{\ominus}$  is the zero function on  $\mathcal{V}_n$ .*

*Proof.* Suppose  $F$  is a  $T$ -flag,  $F'$  is a  $T'$ -flag and  $|E(T)| \geq |E(T')|$ . Since  $T \neq T'$ , there exist vertices  $i$  and  $j$  such that  $\{i, j\} \in E(T)$  and  $\{i, j\} \notin E(T')$ . Therefore, there is some polynomial  $\mathbf{q}$  such that

$$\mathbf{d}_F^\Theta \mathbf{d}_{F'}^\Theta = x_{ij}(1 - x_{ij})\mathbf{q} = 0.$$

The last equality follows from the fact that  $x_{ij}(1 - x_{ij}) = 0$  on  $\mathcal{V}_n$ .  $\square$

Note that the hypercube ideal  $\mathcal{I}_n$  was crucial for the proof of Lemma 4.3. Using this lemma, we now obtain a refined version of Corollary 4.2 (and similarly, Corollary 4.1, which we omit). This will allow us to translate our  $d$ -sos expressions into flag sums of squares afterwards.

**Proposition 4.4.** *Suppose  $\mathbf{p}$  is symmetric and  $d$ -sos. For each intersection type  $T$  of size  $t \leq 2d$ , fix  $\Theta_T \in \text{Inj}([t], [n])$ . Then there exists psd matrices  $R_T$  such that*

$$\mathbf{p} = \sum_{t=0}^{2d} \sum_{T:|T|=t} \text{tr} \left( R_T \text{sym}(\mathbf{d}_T \mathbf{d}_T^\top) \right),$$

where  $\mathbf{d}_T$  is the vector of polynomials  $\mathbf{d}_F^{\Theta_T}$  such that  $F \in \mathcal{F}_T^{2d}$ .

*Proof.* Suppose  $\mathbf{p}$  is symmetric and  $d$ -sos. For each partition  $\lambda \in \Lambda$  (where  $\Lambda = \{\lambda \mid \lambda \geq_{\text{lex}} (n - 2d, 1^{2d})\}$ ), fix a tableau  $\tau_\lambda$  of shape  $\lambda = (\lambda_1, \lambda_2, \dots)$ . Then, by Corollary 4.2, there exist psd matrices  $R_\lambda$  such that

$$\mathbf{p} = \sum_{\lambda \in \Lambda} \text{tr}(R_\lambda Z^{\tau_\lambda}),$$

where  $Z^{\tau_\lambda} = \text{sym}(\mathbf{d}_{\tau_\lambda} \mathbf{d}_{\tau_\lambda}^\top)$  and  $\mathbf{d}_{\tau_\lambda}$  is the vector containing all the polynomials  $\mathbf{d}_F^{\Theta_{\tau_\lambda}}$  such that  $F \in \mathcal{F}_T^{2d}$  where  $|T| = n - \lambda_1$ .

For each intersection type  $T$  with  $|T| = n - \lambda_1$ , let  $\mathbf{d}_{\tau_\lambda, T}$  be the restriction of  $\mathbf{d}_{\tau_\lambda}$  to entries corresponding to flags in  $\mathcal{F}_T^{2d}$ . Similarly, let  $R_{\lambda, T}$  be the principal submatrix of  $R_\lambda$  corresponding to rows and columns indexed by flags in  $\mathcal{F}_T^{2d}$ . Since  $R_\lambda$  is psd, we get that  $R_{\lambda, T}$  is psd for each  $T$ . From Lemma 4.3, we know that for any labeling  $\Theta$ , if  $F$  and  $F'$  have different intersection types, then  $\mathbf{d}_F^\Theta \mathbf{d}_{F'}^\Theta = 0$ . Hence,  $Z_{FF'}^{\tau_\lambda} = 0$  if  $F$  and  $F'$  have different intersection types, and so  $Z^{\tau_\lambda}$  is block diagonal, with one block for each intersection type. Thus, we have

$$\mathbf{p} = \sum_{\lambda \in \Lambda} \sum_{\substack{T: \\ |T|=n-\lambda_1}} \text{tr} \left( R_{\lambda, T} \text{sym}(\mathbf{d}_{\tau_\lambda, T} \mathbf{d}_{\tau_\lambda, T}^\top) \right).$$

Since  $\lambda_1 \geq n - 2d$  for each partition  $\lambda \in \Lambda$  and  $|T| = n - \lambda_1 \leq 2d$ , by switching the sums, we may write the above expression as follows:

$$\mathbf{p} = \sum_{\substack{T: \\ 0 \leq |T| \leq 2d}} \sum_{\substack{\lambda: \\ \lambda_1 = n - |T|}} \text{tr} \left( R_{\lambda, T} \text{sym}(\mathbf{d}_{\tau_\lambda, T} \mathbf{d}_{\tau_\lambda, T}^\top) \right).$$

Note that the vectors  $\mathbf{d}_{\tau_\lambda, T}$  and  $\mathbf{d}_{\tau_{\lambda'}}, T$  are equal if  $\text{row}_1(\tau_\lambda) = \text{row}_1(\tau_{\lambda'})$  since  $\Theta_{\tau_\lambda} = \Theta_{\tau_{\lambda'}}$ . For every  $\lambda$  in the inner sum above, we could choose the first row of each corresponding  $\tau_\lambda$  to be the same. Combining these two observations, we can reindex the vectors  $\mathbf{d}_{\tau_\lambda, T}$  by just  $\mathbf{d}_T$ . Therefore, the expression for  $\mathbf{p}$  can be further



rewritten as

$$\mathbf{p} = \sum_{\substack{T: \\ 0 \leq |T| \leq 2d}} \text{tr} \left( R_T \text{sym}(\mathbf{d}_T \mathbf{d}_T^\top) \right) = \sum_{t=0}^{2d} \sum_{T: |T|=t} \text{tr} \left( R_T \text{sym}(\mathbf{d}_T \mathbf{d}_T^\top) \right),$$

where  $R_T = \sum_{\substack{\lambda: \\ \lambda_1 = n - |T|}} R_{\tau_\lambda, T}$  which is again psd.  $\square$

The next step in getting to flag sos expressions is to switch from the operator  $\text{sym}(\cdot)$  to the language of expectations. This does not change the expressions we care about as we will see in Lemma 4.7.

**Definition 4.5.** Fix  $\Theta_0 \in \text{Inj}([t], [n])$ , and  $F, F' \in \mathcal{F}_T^f$  where  $|T| = t$ . We define

$$\mathbb{E}_\Theta[\mathbf{d}_F^{\Theta_0} \mathbf{d}_{F'}^{\Theta_0}] := \frac{1}{|\text{Inj}([t], [n])|} \sum_{\Theta \in \text{Inj}([t], [n])} \mathbf{d}_F^\Theta \mathbf{d}_{F'}^\Theta$$

**Lemma 4.6.** Fix  $\Theta_0 \in \text{Inj}([t], [n])$ , and  $F, F' \in \mathcal{F}_T^f$  where  $|T| = t$ , then

$$\text{sym}(\mathbf{d}_F^{\Theta_0} \mathbf{d}_{F'}^{\Theta_0}) = \frac{1}{|\mathfrak{S}_n|} \sum_{\mathfrak{s} \in \mathfrak{S}_n} \mathbf{d}_F^{\mathfrak{s} \cdot \Theta_0} \mathbf{d}_{F'}^{\mathfrak{s} \cdot \Theta_0},$$

where  $\mathfrak{s} \cdot \Theta_0(i) = \mathfrak{s}(\Theta_0(i))$ .

*Proof.* This follows from the definitions of the operator  $\text{sym}$  and polynomials  $\mathbf{d}_F^\Theta$ .  $\square$

**Lemma 4.7.** Fix  $\Theta_0 \in \text{Inj}([t], [n])$ , and  $F, F' \in \mathcal{F}_T^f$  where  $|T| = t$ , then

$$\text{sym}(\mathbf{d}_F^{\Theta_0} \mathbf{d}_{F'}^{\Theta_0}) = \mathbb{E}_\Theta[\mathbf{d}_F^{\Theta_0} \mathbf{d}_{F'}^{\Theta_0}]$$

*Proof.* This follows from the fact that  $|\text{Inj}([t], [n])| = \binom{n}{t} t!$  and that, for each  $\Theta \in \text{Inj}([t], [n])$ , there are  $(n-t)!$  permutations  $\mathfrak{s} \in \mathfrak{S}_n$  such that  $\mathfrak{s} \cdot \Theta_0 = \Theta$ .  $\square$

**Definition 4.8.** Let  $\mathbf{d}^{\Theta, T, f} = (\mathbf{d}_F^\Theta)_{F \in \mathcal{F}_T^f}$  be the vector of flag polynomials for a fixed intersection type  $T$ , flag size  $f$ , and labeling  $\Theta$ . A *flag sos* is a sos expression of the form

$$\sum_{T, f} \text{tr} \left( R_{T, f} \mathbb{E}_\Theta \left[ \mathbf{d}^{\Theta, T, f} \mathbf{d}^{\Theta, T, f \top} \right] \right),$$

where the sum is indexed by some chosen  $T$  and  $f$ , and the matrices  $R_{T, f}$  are psd. This expression is called  *$f_{\max}$ -flag sos* if every flag present has size at most  $f_{\max}$ .

Note that the main difference between the sos expression in Corollary 4.2 and a flag sos expression is that each summand in the latter uses only one intersection type. However, we saw in Proposition 4.4 that we can refine the sos in Corollary 4.2 to have the same property. The only step left is to bring in the language of expectations to obtain flag sos expressions.

**Theorem 4.9.** If  $\mathbf{p}$  is symmetric and  $d$ -sos, then  $\mathbf{p}$  is also a  $2d$ -flag sos.

*Proof.* By Proposition 4.4, if  $\mathbf{p}$  is symmetric and  $d$ -sos, then

$$\mathbf{p} = \sum_{t=0}^{2d} \sum_{T: |T|=t} \text{tr} \left( R_T \text{sym}(\mathbf{d}_T \mathbf{d}_T^\top) \right),$$

where  $\mathbf{d}_T$  is the vector of polynomials  $\mathbf{d}_F^{\Theta_T}$  such that  $F \in \mathcal{F}_T^{2d}$  and  $\Theta_T \in \text{Inj}([t], [n])$ . By Lemma 4.7, we then have that

$$\mathfrak{p} = \sum_{t=0}^{2d} \sum_{T:|T|=t} \text{tr} \left( R_T \mathbb{E}_{\Theta} [\mathbf{d}_T \mathbf{d}_T^T] \right).$$

□

In [RST15], we proved that any flag sos polynomial of degree  $2d$  can be written as a sos coming from  $\mathbb{R}[x]_{\leq d}$  using the techniques of [GP04]. We have now shown that the reverse is also true in the sense that any symmetric sos over  $\mathcal{V}_n$  can also be written as a flag sos. Thus flag methods are equivalent to general methods for finding sos expressions over  $\mathcal{V}_n$ .

## 5. EXAMPLES

In this section, we illustrate our methods on two examples. The first is a sos proof of an upper bound on the edge density of a  $n$ -vertex graph that does not contain any 4-cycles, and the second is a sos proof that the Ramsey number  $R(3, 3) = 6$ . These example illustrates various features of our method. The first is that we can work with additional constraints beyond those defining the discrete hypercube. The second is that in both examples,  $n$  is fixed, and hence they showcase the merits of considering flags in a finite setting. Thirdly, the polynomials  $\mathbf{g}_F^{\Theta}$  and  $\mathbf{d}_F^{\Theta}$  that appear in these proofs are inspired by the combinatorics of the problems at hand. Normally, the bases used in symmetry-reduction, such as the Gatermann-Parrilo method, does not preserve this sort of combinatorial information.

**5.1. Avoiding 4-cycles.** In [KST54], Kővari, Sós and Turán proved that the number of edges in a  $n$ -vertex graph not containing  $C_4$ , the cycle on four vertices, is bounded above by  $\frac{1}{2}n^{3/2} + O(n)$ . Observe that the bound implies that the edge density of an extremal graph in this setting tends to zero, and thus the standard application of the flag algebra method cannot recover the precise lower order term. We give a succinct sos proof for this result using our methods which follows the proof outline in [KST54].

**Theorem 5.1.** [KST54] *Let  $G$  be a  $n$  vertex graph not containing a  $C_4$ . Then the number of edges in  $G$  is at most  $\frac{1}{2}n^{3/2} + O(n)$ .*

*Proof.* Fix  $n$  and make the following definitions:

$$\begin{aligned} \mathfrak{s} &:= \sum_{1 \leq i < j \leq n} x_{ij}, \text{ and} \\ \mathcal{I} &:= \langle x_{ij}^2 - x_{ij} \mid \forall 1 \leq i < j \leq n, x_{ij}x_{ik}x_{lj}x_{lk} \mid \forall i, j, k, l \in [n] \rangle. \end{aligned}$$

Note that the variety of  $\mathcal{I}$  is the subset of  $\mathcal{V}_n$  that consists of characteristic vectors of  $n$ -vertex graphs that avoid a  $C_4$ . We will show that  $n + \frac{2}{n-1}\mathfrak{s} - \frac{2}{\binom{n}{2}}\mathfrak{s}^2$  is 2-sos modulo the ideal  $\mathcal{I}$ . This will imply that  $\mathfrak{s} \leq \frac{n + \sqrt{4n^3 - 3n^2}}{4}$  for all  $C_4$ -free graphs on  $n$  vertices giving the bound claimed in the theorem.

We first consider the spanning set given by the  $\mathbf{g}_F^{\Theta}$  polynomials. Consider flags of size three with the intersection type  $\overset{1}{\bullet} \underset{2}{\bullet}$ , and let  $\Theta_{jk} \in \text{Inj}([2], [n])$  be such that  $\Theta(1) = j$  and  $\Theta(2) = k$ . Then,

$$\begin{aligned}
\mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} &= n - 2, & \mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} &= \sum_{i \in [n] \setminus \{j, k\}} x_{ij}, \\
\mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} &= \sum_{i \in [n] \setminus \{j, k\}} x_{ik}, & \mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} &= \sum_{i \in [n] \setminus \{j, k\}} x_{ij} x_{ik}.
\end{aligned}$$

We now claim that the following sos certificate is correct and thus prove the theorem.

$$\begin{aligned}
n + \frac{2}{n-1} \mathbf{s} - \frac{2}{\binom{n}{2}} \mathbf{s}^2 \\
\equiv \mathbb{E}_{\Theta_{jk}} \left[ n \left( \frac{1}{n-2} \mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} - \mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} \right)^2 + \frac{1}{2} \left( \mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} - \mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} \right)^2 \right] \pmod{\mathcal{I}}.
\end{aligned}$$

The above expression follows from the following equalities. First, for each  $j \neq k$ , we have the following equivalence:

$$\begin{aligned}
& \left( \frac{1}{n-2} \mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} - \mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} \right)^2 \\
&= \left( 1 - \sum_{i \in [n] \setminus \{j, k\}} x_{ij} x_{ik} \right)^2 \\
&= 1 + \sum_{i \in [n] \setminus \{j, k\}} x_{ij}^2 x_{ik}^2 - 2 \sum_{i \in [n] \setminus \{j, k\}} x_{ij} x_{ik} + \sum_{i \neq l: i, l \in [n] \setminus \{j, k\}} x_{ij} x_{ik} x_{lj} x_{lk} \\
&\equiv \left( 1 - \sum_{i \in [n] \setminus \{j, k\}} x_{ij} x_{ik} \right) \pmod{\mathcal{I}}.
\end{aligned}$$

Thus,

$$\mathbb{E}_{\Theta_{jk}} \left[ n \left( \frac{1}{n-2} \mathbf{g}_{1 \bullet \bullet 2}^{\ominus_{jk}} - \mathbf{g}_{1 \bullet \blacktriangleright 2}^{\ominus_{jk}} \right)^2 \right] \equiv n - \frac{1}{n-1} \sum_{j, k \in [n], j \neq k} \sum_{i \in [n] \setminus \{j, k\}} x_{ij} x_{ik} \pmod{\mathcal{I}}.$$

Next,

$$\begin{aligned}
& \mathbb{E}_{\Theta_{jk}} \left[ \frac{1}{2} \left( g_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} - g_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} \right)^2 \right] \\
&= \frac{1}{n(n-1)} \sum_{\substack{j,k \in [n]: \\ j \neq k}} \frac{1}{2} \left( \sum_{i \in [n] \setminus \{j\}} x_{ij} - \sum_{i \in [n] \setminus \{k\}} x_{ik} \right)^2 \\
&= \frac{1}{n(n-1)} \left( n \sum_{j \in [n]} \left( \sum_{i \in [n] \setminus \{j\}} x_{ij} \right)^2 - \left( \sum_{\substack{i,j \in [n]: \\ i \neq j}} x_{ij} \right)^2 \right) \\
&= \frac{1}{n-1} \sum_{\substack{j,k \in [n]: \\ j \neq k}} \left( \sum_{i \in [n] \setminus \{j,k\}} x_{ij} x_{ik} \right) + \frac{1}{n-1} 2s - \frac{1}{n(n-1)} 4s^2.
\end{aligned}$$

In order to check the above equations, it helps to think of  $\sum_{i \in [n] \setminus \{j\}} x_{ij}$  for a fixed  $j$  as the degree of vertex  $j$ , and  $\sum_{i,j \in [n]: i \neq j} x_{ij}$  as the sum of the degrees of all vertices. Combining the two equalities above, yields the desired sos expression.

Similarly, for the  $d_F^\Theta$ 's, we again consider flags of size three with the intersection types  $\mathbf{1} \bullet \bullet_2$  and  $\mathbf{1} \bullet \bullet_2$ . We also again let  $\Theta_{jk} \in \text{Inj}([2], [n])$  be such that  $\Theta(1) = j$  and  $\Theta(2) = k$ . This gives us the following sos expression

$$\begin{aligned}
& n + \frac{2}{n-1} s - \frac{2}{\binom{n}{2}} s^2 \equiv \\
& \mathbb{E}_{\Theta_{jk}} \left[ n \left( d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} + d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} + d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} \right)^2 + n \left( d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} + d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} + d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} \right)^2 \right. \\
& \quad \left. + \frac{1}{2} \left( d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} - d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} \right)^2 + \frac{1}{2} \left( d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} - d_{\mathbf{1} \bullet \bullet_2}^{\Theta_{jk}} \right)^2 \right] \text{ mod } \mathcal{I},
\end{aligned}$$

where each summand involves a single intersection type.  $\square$

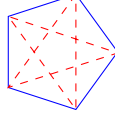
**5.2. A Ramsey example.** Next we give a sos proof that the Ramsey number  $R(3, 3) = 6$ . The Ramsey  $R(r, s)$  problem consists of finding the minimum number of vertices  $n$  such that it is impossible to color all the edges of the complete graph  $K_n$  blue and red without creating either a blue clique of size  $r$  or a red clique of size  $s$  [Rad94]. Ramsey's theorem states that this number is finite for any  $r, s \in \mathbb{Z}$ . Several values of  $R(r, s)$  are known. For example,  $R(3, 3) = 6$ : no matter how one colors the edges of  $K_6$ , one is guaranteed to have a monochromatic triangle.

We consider the following model. We let  $x_e$  be one if edge  $e$  is colored red and zero if edge  $e$  is colored blue;  $x \in \mathbb{R}^{\binom{6}{2}}$ . Consider the ideal

$$\begin{aligned}
\mathcal{I} = & \langle x_e^2 - x_e \ \forall e \in E(K_6) \rangle + \langle x_{ij} x_{ik} x_{jk} \ \forall i < j < k \in [6] \rangle \\
& + \langle (1 - x_{ij})(1 - x_{ik})(1 - x_{jk}) \ \forall i < j < k \in [6] \rangle.
\end{aligned}$$

The variety of  $\mathcal{I}$  consists of those vectors in  $\{0, 1\}^{\binom{6}{2}}$  that record a red/blue coloring of  $K_6$  without monochromatic triangles since the first part ensures that each  $x_e$  is

either 0 or 1, the second part, that no triangle is red, and the third part, that no triangle is blue. Showing that the variety of  $\mathcal{I}$  is empty will prove that  $R(3, 3) \leq 6$ . The coloring of  $K_5$  below establishes that  $R(3, 3) > 5$ .



To prove that the variety of  $\mathcal{I}$  is empty, it suffices to show that  $-1$  is equivalent to a sos mod  $\mathcal{I}$ . We now provide sos expressions that do this job in our setting.

Consider the  $d_F^\Theta$ 's from the the following flags:  $\bullet$ ,  $\bullet$ ,  $\bullet_1$  and  $\bullet_1$ . Note that for the first two,  $\Theta$  is irrelevant and unique since the intersection type is empty. For the last two, we assume that  $\Theta_i$  sends 1 to vertex  $i$ . We have

$$\begin{aligned} d_{\bullet}^\Theta &= 2 \sum_{1 \leq i < j \leq 6} x_{ij}, & d_{\bullet}^\Theta &= 2 \sum_{1 \leq i < j \leq 6} (1 - x_{ij}), \\ d_{\bullet_1}^{\Theta_i} &= \sum_{j \in [6] \setminus \{i\}} x_{ij}, & d_{\bullet_1}^{\Theta_i} &= \sum_{j \in [6] \setminus \{i\}} (1 - x_{ij}). \end{aligned}$$

Similarly, for the  $g_F^\Theta$ 's, we need the flags:  $\bullet$ ,  $\bullet_1$  and  $\bullet_1$ . Note that for the first one,  $\Theta$  is irrelevant and unique since the intersection type is empty. For the last two, we assume that  $\Theta_i$  sends 1 to vertex  $i$ . We have

$$g_{\bullet}^\Theta = 30, \quad g_{\bullet_1}^{\Theta_i} = \sum_{j \in [6] \setminus \{i\}} x_{ij}, \quad g_{\bullet_1}^{\Theta_i} = 5.$$

**Theorem 5.2.** *The following expressions*

$$\begin{aligned} &\frac{1}{8 \binom{6}{2}^2} \left( d_{\bullet}^\Theta + d_{\bullet}^\Theta \right)^2 + \mathbb{E}_{\Theta_i} \left[ \frac{1}{2} \left( d_{\bullet_1}^{\Theta_i} - d_{\bullet_1}^{\Theta_i} \right)^2 \right] \\ &\text{and } \frac{1}{2} \left( \frac{1}{30} g_{\bullet}^\Theta \right)^2 + \mathbb{E}_{\Theta_i} \left[ \left( \sqrt{2} g_{\bullet_1}^{\Theta_i} - \frac{1}{\sqrt{2}} g_{\bullet_1}^{\Theta_i} \right)^2 \right] \end{aligned}$$

are each equivalent to  $-1$  modulo  $\mathcal{I}$  and hence the variety of  $\mathcal{I}$  is empty.

*Proof.* Check that the expressions in the  $d_F^\Theta$ 's and  $g_F^\Theta$ 's are each equal to

$$\text{sym} \left( \left( \frac{1}{\sqrt{2}} \right)^2 + \left( \sqrt{2} \sum_{i=2}^6 x_{1i} - \frac{5}{\sqrt{2}} \right)^2 \right).$$

This sos is equal to another sos

$$\text{sym} \left( \left( 2 - \sum_{i=2}^6 x_{1i} \right)^2 + \left( 2 - \sum_{i=2}^6 (1 - x_{1i}) \right)^2 \right),$$

which is equivalent to  $-1$  modulo  $\mathcal{I}$  by Lemma B.1. □

## 6. AN EXAMPLE NEEDING INFINITELY MANY FLAGS

For every positive integer  $n$ , define the degree two polynomial

$$(4) \quad f_n = \frac{1}{\binom{n}{2}^2} \left( \sum_{e \in E(K_n)} x_e - \left\lfloor \frac{\binom{n}{2}}{2} \right\rfloor \right) \left( \sum_{e \in E(K_n)} x_e - \left\lfloor \frac{\binom{n}{2}}{2} \right\rfloor - 1 \right).$$

Observe that  $f_n$  satisfies  $0 \leq f_n(x) \leq 1$  for all  $x \in \mathcal{V}_n$ . (Here the upper bound is not tight.) In this section, we show that there is no finite collection of flags  $\mathcal{F}$  such that each  $f_n$  is equivalent to a sos (modulo  $\mathcal{I}_n$ ) of flag polynomials  $d_F^\ominus$  for  $F \in \mathcal{F}$ .

Despite being of degree two and non-negative on  $\mathcal{V}_n$ , Grigoriev [Gri01] showed that any sos  $h_n$ , such that  $h_n(x) = f_n(x)$  for all  $x \in \mathcal{V}_n$ , necessarily has degree  $\Omega\left(\binom{n}{2}\right)$  (see also [BGP] for an extension to the case in which  $h$  is rational). Recently, Lee, Prakesh, de Wolf, and Yuen [LPdWY16] extended this degree lower bound to the case in which  $h_n$  approximates  $f_n$  in the  $\ell_\infty$  sense. The following is a special case of [LPdWY16, Theorem 1.1].

**Theorem 6.1.** *There exists a (sufficiently small) positive constant  $c$  such that for any positive integer  $n$ , if  $h_n$  is  $d$ -sos and satisfies*

$$|h_n(x) - f_n(x)| \leq \frac{(1/50)}{\binom{n}{2}^2}$$

for all  $x \in \mathcal{V}_n$ , then  $d \geq c\binom{n}{2}$ .

Since the minimum value of  $f_n$  on  $\mathbb{R}^{\binom{n}{2}}$  is  $-(1/4)/\binom{n}{2}^2$ , it follows that  $f_n(x) + (1/4)/\binom{n}{2}^2$  is a non-negative quadratic on  $\mathbb{R}^{\binom{n}{2}}$ , and so it is a sos of degree one polynomials. As such, we certainly cannot hope to tolerate errors larger than  $(1/4)/\binom{n}{2}^2$  in Theorem 6.1.

Any symmetric sos of flag polynomials  $d_F^\ominus$  using a fixed collection of flags  $F$  necessarily has bounded degree. Hence the (growing) lower bound on degree from Theorem 6.1 implies that sums of squares of flag polynomials using a fixed collection of flags cannot be used to certify non-negativity of every  $f_n$ ,  $n \geq 1$ .

**Proposition 6.2.** *For any fixed positive integer  $f$ , there does not exist a  $f$ -flag sos  $\mathbf{p}$  such that  $|\mathbf{p}(x) - f_n(x)| \leq \frac{(1/50)}{\binom{n}{2}^2}$  for all  $x \in \mathcal{V}_n$  and all sufficiently large  $n$ .*

*Proof.* If  $\mathbf{p}$  is a  $f$ -flag sos, then  $\mathbf{p}$  is  $\binom{f}{2}$ -sos. If  $n$  is large enough, then  $c\binom{n}{2} > \binom{f}{2}$ , where  $c$  is the constant in Theorem 6.1.  $\square$

This shows that there is a sequence of symmetric polynomials of degree two, all of which are non-negative and bounded on  $\mathcal{V}_n$ , that cannot be approximated within  $O(1/n^4)$  by  $f$ -flag sums of squares for any fixed  $f$ . We make several comments.

Razborov's application of flag density polynomials typically allows much larger errors of size  $O(1/n)$ , suggesting the following question. Is there a sequence of symmetric polynomials taking values in  $[0, 1]$  that cannot be approximated within  $O(1/n)$  by sums of squares from a finite collection of flags?

In [Raz07], Razborov asked whether every asymptotic graph density inequality can be certified by a finite set of flags. As an answer to this, Hatami and Norine proved that deciding the non-negativity of a linear inequality in graph densities is in general undecidable [HN11]. Proposition 6.2 can be viewed as an answer to Razborov's question in our setting of the discrete hypercube for finite  $n$ .

At a quick glance, Proposition 6.2 may seem to contradict our earlier result that it is possible to write a flag sos expression that depends only on the degree of a given sos expression and independent of  $n$ . For the family of quadratic polynomials  $f_n$ , the degree of any sos expression is a function of  $n$  and therefore, by our results, the number of flags needed in a flag sos expression also depends on  $n$ .

### 7. SUMMARY AND DISCUSSION

Our results give the following framework for searching for a succinct sos expression for a symmetric polynomial  $\mathbf{p}$  that is  $d$ -sos.

- (1) For each partition  $\lambda \in \Lambda = \{\lambda \vdash n : \lambda \geq_{\text{lex}} (n - 2d, 1^{2d})\}$ , fix a standard tableau  $\tau_\lambda$  of shape  $\lambda$ .
- (2) Pick polynomials  $p_1^{\tau_\lambda}, \dots, p_{l_{\tau_\lambda}}^{\tau_\lambda}$  whose span contains  $W_{\tau_\lambda}$ .
- (3) Introduce a matrix of variables  $R_\lambda$  of size  $l_{\tau_\lambda} \times l_{\tau_\lambda}$ .
- (4) Formulate a SDP using the following polynomial equivalence

$$\mathbf{p} = \sum_{\lambda \in \Lambda} \text{tr}(R_\lambda Z^{\tau_\lambda})$$

where  $Z_{ij}^{\tau_\lambda} := \text{sym}(p_i^{\tau_\lambda} p_j^{\tau_\lambda})$ .

We summarize the computational savings in each of the above steps. First of, the symmetry of  $\mathbf{p}$  allows us to work with a *single* standard tableau  $\tau_\lambda$  of shape  $\lambda \in \Lambda$  as in Corollary 2.6. By Proposition 2.7, the size of  $\Lambda$ , and hence the number of partitions that are needed for sos expressions in our method, is independent of  $n$ . Our framework gives us the flexibility to work with a spanning set for  $W_{\tau_\lambda}$  as opposed to a basis. We constructed spanning sets in Section 3 whose cardinalities are independent of  $n$  as proved in Proposition 3.21. The size of the formulated SDP depends on the total sum of dimensions  $m_\lambda$  of  $W_{\tau_\lambda}$  for  $\lambda \in \Lambda$ . Combining Propositions 2.7 and 3.21, we get that  $\sum_{\lambda \in \Lambda} m_\lambda$  is also independent of  $n$ .

Our spanning sets for  $W_{\tau_\lambda}$  can have combinatorial meaning and structure relevant to the polynomial  $\mathbf{p}$  in some instances. For example, in extremal combinatorics, the density polynomials from flag algebras have simple combinatorial interpretations as compared to the basis necessary for the Gatermann-Parrilo method. We also saw in Section 3 that the concept of flags arise naturally in our theory from the action of  $\mathfrak{S}_n$  on square-free monomials.

We now make several comments on various features, connections and extensions of our method. While we presented our results for the hypercube  $\mathcal{V}_n$ , the arguments naturally generalize to subsets of  $\mathcal{V}_n$  by working with ideals that contain  $\mathcal{I}_n$  and are invariant under  $\mathfrak{S}_n$ . There are several ways to approach this general situation. If the ideal in question is  $\mathcal{I}$ , then for instance, one could look for symmetry-reduced sos expressions from the vector space  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$  as we have been doing in this paper. In this case, our methods work as is and we can further reduce the obtained sos expression by the ideal  $\mathcal{I}$  as seen in the Ramsey example in Section 5.2. Alternately, one could work with  $V = \mathbb{R}[x]_{\leq d}/\mathcal{I}$  or  $V = (\mathbb{R}[x]/\mathcal{I})_{\leq d}$ . In both these cases, our strategy would work in principle but we would need to understand the isotypic decomposition of  $V$  and the associated vector spaces  $W_{\tau_\lambda}$ .

In contrast to the flag algebra and graph homomorphism frameworks, our methods work on graphs of finite size when applied to graph density problems. This allows us to obtain new types of exact sos proofs that apply to finite, as opposed to asymptotic situations. For example, in Section 5.2, we gave a sos proof for an

upper bound on the Ramsey number  $R(3, 3)$  in which case  $n = 6$ . Our methods also naturally give upper bounds on extremal graph theoretic problems in the sparse regime. For example, we saw how to use our method to obtain a succinct sos proof of the result of Kővari, Sós and Turán [KST54] that a  $n$ -vertex graph not containing  $C_4$  has at most  $\frac{1}{2}n^{3/2} + O(n)$  edges. Since the extremal graphs are not dense in this case, applying the flag algebra framework gives an upper bound of zero on the limiting edge density but our method can give a precise estimate of the lower order terms.

Finally, as promised in the Introduction, our results extend naturally to the hypercube  $\mathcal{V}_{n,k} := \{0, 1\}^{\binom{n}{k}}$ . We would use the polynomial ring with variables  $x_{i_1 \dots i_k}$  indexed by the edges in the complete  $k$ -uniform hypergraph  $K_n^k$  with  $\mathfrak{S}_n$  acting on monomials via  $\mathfrak{s} \cdot x_{i_1 \dots i_k} := x_{\mathfrak{s}(i_1) \dots \mathfrak{s}(i_k)}$  for each  $\mathfrak{s} \in \mathfrak{S}_n$ . The ideal needed here is  $\mathcal{I}_{n,k} := \langle x_{i_1 \dots i_k}^2 - x_{i_1 \dots i_k} \forall \{i_1, \dots, i_k\} \in \binom{[n]}{k} \rangle$ . All of our results and their proofs can be translated in a straightforward manner to this setting and we give a few samples below.

- In Theorem 2.4, the multiplicity  $m_\lambda$  of  $S_\lambda$  in the decomposition of  $V = \mathbb{R}[x_{i_1 \dots i_k}]_{\leq d}$  into irreducible  $\mathfrak{S}_n$ -modules is zero unless  $\lambda \geq_{\text{lex}} (n - kd, 1^{kd})$ , i.e.,  $V = \bigoplus_{\lambda \geq_{\text{lex}} (n - kd, 1^{kd})} V_\lambda$ . Thus the number of partitions with non-zero  $m_\lambda$  can be bounded by  $p(0) + p(1) + \dots + p(kd)$  generalizing Proposition 2.7.
- In Theorem 3.18, the polynomials  $\mathfrak{g}_F^\ominus$  (respectively  $\mathfrak{d}_F^\ominus$ ) with flags  $F \in \mathcal{F}_T^{kd}$  (with the same restrictions) span  $W_{\tau_\lambda}$ . The sizes of these spanning sets will be independent of  $n$  as in Proposition 3.21.
- In Theorem 4.9, if a non-negative symmetric polynomial  $\mathfrak{p} \in \mathbb{R}[\mathcal{V}_{n,k}]$  is  $d$ -sos, then  $\mathfrak{p}$  also has a flag sos certificate coming from flags with at most  $kd$  vertices.

## REFERENCES

- [BGP] G. Blekherman, J. Gouveia, and J. Pfieffer. Sums of squares on the hypercube. *Mathematische Zeitschrift*. to appear.
- [BPT13] G. Blekherman, P.A. Parrilo, and R.R. Thomas, editors. *Semidefinite Optimization and Convex Algebraic Geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013.
- [FRV13] V. Falgas-Ravry and E.R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Comput.*, 22(1):21–54, 2013.
- [FS92] A. Fässler and E. Stiefel. *Group Theoretical Methods and Their Applications*. Birkhäuser, 1992.
- [GP04] K. Gatermann and P.A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra*, 192(1-3):95–128, 2004.
- [Gri01] D. Grigoriev. Complexity of Positivstellensatz proofs for the knapsack. *Computational Complexity*, 10(2):139–154, 2001.
- [HN11] H. Hatami and S. Norine. Undecidability of linear inequalities in graph homomorphism densities. *J. Amer. Math. Soc.*, 24(2):547–565, 2011.
- [KLM16] A. Kurpisz, S. Leppänen, and M. Mastrolilli. Sum-of-squares hierarchy lower bounds for symmetric formulations. In *Integer Programming and Combinatorial Optimization*, pages 362–374. Springer, 2016.
- [KST54] T. Kővari, V. Sós, and P. Turán. On a problem of K. Zarankiewicz. In *Colloquium Mathematicae*, volume 3, pages 50–57. Institute of Mathematics Polish Academy of Sciences, 1954.
- [Lau09] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, New York, 2009.



- [Lov12] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.
- [LPdWY16] T. Lee, A. Prakash, R. de Wolf, and H. Yuen. On the sum-of-squares degree of symmetric quadratic functions. arXiv:1601.02311, 2016.
- [Par02] P.A. Parrilo. An explicit construction of distinguished representations of polynomials nonnegative over finite sets. Technical Report IfA AUT 02-02, ETH Zürich, 2002.
- [Rad94] Stanisław P. et al. Radziszowski. Small Ramsey numbers. *Electron. J. Combin*, 1(7), 1994.
- [Raz07] A.A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.
- [Raz10] A.A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM Journal on Discrete Mathematics*, 24(3):946–963, 2010.
- [Raz14] A.A. Razborov. On Turán’s (3, 4)-problem with forbidden subgraphs. *Mathematical Notes*, 95(1-2):245–252, 2014.
- [RST15] A. Raymond, M. Singh, and R.R. Thomas. Symmetry in Turán sums of squares polynomials from flag algebras. arXiv:1507.03059, 2015.
- [Sag01] B.E. Sagan. *The Symmetric Group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.
- [Ser77] J-P. Serre. *Linear representations of finite groups*. Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, BOX 354350, SEATTLE, WA 98195, USA

*E-mail address:* raymonda@uw.edu

DEPARTMENT OF ELECTRICAL ENGINEERING, UNIVERSITY OF WASHINGTON, BOX 352500, SEATTLE, WA 98195, USA

*E-mail address:* jamesfs@uw.edu

MICROSOFT RESEARCH, 99/2934, 1 MICROSOFT WAY, REDMOND, WA 98052

*E-mail address:* mohits@microsoft.com

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, BOX 354350, SEATTLE, WA 98195, USA

*E-mail address:* rrrthomas@uw.edu

## APPENDIX A. PROOF OF THEOREM 2.2

The aim of this appendix is to prove Theorem 2.2, describing the structure of symmetry-reduced certificates of non-negativity for  $\mathfrak{S}_n$ -invariant polynomials that are  $d$ -sos. Rather than proving Theorem 2.2 directly, we establish a slightly more general result (Theorem A.8 to follow) and then specialize to give Theorem 2.2. This more general result is essentially the main result of Gatermann and Parrilo [GP04]. In Section A.1, we summarize some basic facts about real representations needed for the appendix. In Section A.2, we examine what happens when we symmetrize products and squares of polynomials under a finite group action, leading to a proof of Theorem A.8 and of Theorem 2.2.

**A.1. Preliminaries on real representations.** Let  $G$  be a finite group and let  $U$  be a finite-dimensional  $\mathbb{R}G$ -module, (i.e., a real vector space with an action of  $G$  by linear transformations). An  $\mathbb{R}G$ -module  $U$  is *irreducible* if the only real vector subspaces of  $U$  invariant under the action of  $G$  are  $\{0\}$  and  $U$ . Given any pair  $U, W$  of  $\mathbb{R}G$ -modules, let  $\text{Hom}_G(U, W)$  be the vector space of  $\mathbb{R}$ -linear maps  $\phi : U \rightarrow W$  such that  $\phi(g \cdot u) = g \cdot \phi(u)$  for all  $u \in U$ . If  $U, W$  are  $\mathbb{R}G$ -modules, then the tensor product (over  $\mathbb{R}$ ),  $U \otimes W$ , is a  $\mathbb{R}G$ -module with the action  $g \cdot (u \otimes w) = (g \cdot u) \otimes (g \cdot w)$ . If  $U$  is a  $\mathbb{R}G$ -module, then the space  $U^*$  of  $\mathbb{R}$ -valued linear functionals on  $U$  is a  $\mathbb{R}G$ -module via  $(g \cdot \ell)(u) = \ell(g^{-1} \cdot u)$ .

If  $U$  is a  $\mathbb{R}G$ -module, define the linear map  $\text{sym}_G : U \rightarrow U$  by

$$\text{sym}_G(u) = \frac{1}{|G|} \sum_{g \in G} g \cdot u.$$

Clearly, the image of  $\text{sym}_G$  is the set of elements of  $U$  fixed by the action of  $G$ , i.e.,  $U^G := \{u \in U : g \cdot u = u \text{ for all } g \in G\}$ . We omit the subscript  $G$  from  $\text{sym}_G$  if the group is clear from the context. We also note that  $\text{sym}_G \in \text{Hom}_G(U, U^G)$  where  $U^G$  carries the trivial action of  $G$ .

We now recall the the isotypic decomposition of a  $\mathbb{R}G$ -module  $V$ . Let  $(S_\lambda)_{\lambda \in \Lambda}$  be an enumeration of inequivalent irreducible  $\mathbb{R}G$ -modules and let  $n_\lambda = \dim(S_\lambda)$ . For each  $\lambda \in \Lambda$ , let  $\text{Hom}_G(S_\lambda, V)$  denote the  $m_\lambda$ -dimensional *multiplicity space* of  $S_\lambda$  in  $V$ . Let

$$V_\lambda = \text{span}\{\phi(s) : \phi \in \text{Hom}_G(S_\lambda, V), s \in S_\lambda\}$$

denote the *isotypic component* of  $V$  corresponding to  $S_\lambda$ . The decomposition  $V = \bigoplus_{\lambda \in \Lambda} V_\lambda$  is the *isotypic decomposition* of  $V$ .

There are two natural ways to construct subspaces of the isotypic components  $V_\lambda$ . If we fix a non-zero  $\phi \in \text{Hom}_G(S_\lambda, V)$  and consider  $\{\phi(s) : s \in S_\lambda\}$ , we obtain a  $G$ -invariant subspace of  $V_\lambda$  that is isomorphic to  $S_\lambda$  and has dimension  $n_\lambda$ . On the other hand, if we fix a non-zero  $s_\lambda \in S_\lambda$ , we can define

$$(5) \quad W_{s_\lambda} := \{\phi(s_\lambda) : \phi \in \text{Hom}_G(S_\lambda, V)\} \subseteq V_\lambda.$$

This is a vector subspace of  $V_\lambda$  of dimension  $m_\lambda$ , but is not  $G$ -invariant. It is however isomorphic to the multiplicity space  $\text{Hom}_G(S_\lambda, V)$  as a vector space.

**Lemma A.1.** *Fix some non-zero  $s_\lambda \in S_\lambda$ . Then the map  $\Phi_{s_\lambda} : \text{Hom}_G(S_\lambda, V) \rightarrow W_{s_\lambda}$  defined by  $\Phi_{s_\lambda}(\phi) = \phi(s_\lambda)$  is an isomorphism of vector spaces.*

*Proof.* The map  $\Phi_{s_\lambda}$  is clearly linear. It is surjective by the definition of  $W_{s_\lambda}$ . It remains to show that  $\Phi_{s_\lambda}$  is injective. Let  $\phi \in \text{Hom}_G(S_\lambda, V)$  and suppose that  $\Phi_{s_\lambda}(\phi) = \phi(s_\lambda) = 0$ . Since  $\phi \in \text{Hom}_G(S_\lambda, V)$ , the kernel of  $\phi$  is an invariant

subspace of  $S_\lambda$  that contains the non-zero element  $s_\lambda$ . Since  $S_\lambda$  is irreducible, the kernel of  $\phi$  must then be all of  $S_\lambda$ , and so  $\phi = 0$ . Hence  $\Phi_{s_\lambda}$  is injective and so an isomorphism of vector spaces.  $\square$

The following standard fact about invariant bilinear forms on  $\mathbb{R}G$ -modules, is central to our discussion.

**Lemma A.2.** *If  $G$  is a finite group and  $U$  is a finite-dimensional  $\mathbb{R}G$ -module then  $U$  has a non-degenerate,  $G$ -invariant, symmetric, bilinear form  $B : U \times U \rightarrow \mathbb{R}$  satisfying  $B(u, u) > 0$  for all non-zero  $u \in U$ . Moreover, if  $U$  is irreducible, then any  $G$ -invariant bilinear form on  $U$  is a scalar multiple of  $B$ .*

*Proof.* Let  $\langle \cdot, \cdot \rangle$  be any inner product on  $U$ , i.e.,  $\langle \cdot, \cdot \rangle$  is a symmetric, non-degenerate, positive definite, bilinear form. Then  $B : U \times U \rightarrow \mathbb{R}$ , defined by

$$B(u, v) = \frac{1}{|G|} \sum_{g \in G} \langle g \cdot u, g \cdot v \rangle,$$

is  $G$ -invariant, and is a convex combination of non-degenerate, symmetric, positive definite, bilinear forms. Hence,  $B$  also has these properties. Now suppose  $U$  is irreducible and  $B'$  is another  $G$ -invariant bilinear form on  $U$ . Since  $B(\cdot, \cdot)$  is positive definite, we can choose  $\lambda$  such that the subspace  $\{v \in U : \lambda B(u, v) = B'(u, v) \text{ for all } u \in U\}$  is non-zero. (Here, we can take  $\lambda$  to be any generalized eigenvalue for the pair of symmetric matrices representing the bilinear forms.) Since  $B$  and  $B'$  are both  $G$ -invariant, this is a  $G$ -invariant subspace of  $U$ . Since  $U$  is irreducible, we must have  $\{v \in U : \lambda B(u, v) = B'(u, v) \text{ for all } u \in U\} = U$ . Hence  $\lambda B = B'$ , as required.  $\square$

We now record basic facts about  $G$ -invariant elements of tensor products.

**Lemma A.3.** *If  $G$  is a finite group and  $U$  and  $W$  be non-isomorphic irreducible finite-dimensional  $\mathbb{R}G$ -modules, then  $(U \otimes W)^G = \{0\}$ .*

*Proof.* Suppose  $W$  and  $U$  are irreducible and not isomorphic. By Schur's lemma (see, for instance, [Ser77, Section 2.2]),  $\text{Hom}_G(W, U) = \{0\}$ . Since  $U$  is a finite-dimensional  $\mathbb{R}G$ -module, by Lemma A.2 it has a non-degenerate,  $G$ -invariant, bilinear form  $B : U \times U \rightarrow \mathbb{R}$ . Hence the map  $U \ni u \mapsto B(u, \cdot) \in U^*$  gives an isomorphism between  $U$  and  $U^*$ . Then  $(U \otimes W)^G \cong (U^* \otimes W)^G \cong \text{Hom}_G(U, W) = \{0\}$ .  $\square$

The following result tells us that if  $U$  is irreducible, then  $(U \otimes U)^G$  is one-dimensional. It also explicitly describes the map  $\text{sym} : U \otimes U \rightarrow (U \otimes U)^G$ .

**Lemma A.4.** *If  $G$  is a finite group and  $U$  is a finite-dimensional irreducible  $\mathbb{R}G$ -module then for all  $s \in U$  and all  $u, u' \in U$ ,*

$$B(s, s) \text{sym}(u \otimes u') = B(u, u') \text{sym}(s \otimes s)$$

where  $B(\cdot, \cdot)$  is the (unique up to scale) non-zero  $G$ -invariant bilinear form on  $U$ .

*Proof.* If  $s = 0$  then the statement clearly holds. Assume  $s \neq 0$  and let  $\eta \in (U \otimes U)^*$  be an arbitrary linear functional on  $U \otimes U$ . Then the map  $(u, u') \mapsto \eta(\text{sym}(u \otimes u'))$  is a  $G$ -invariant bilinear form on  $U$ . By Lemma A.2,  $\eta(\text{sym}(u \otimes u')) = \kappa(\eta)B(u, u')$  for some  $\kappa(\eta) \in \mathbb{R}$ . By putting  $u = u' = s$ , we see that  $\kappa(\eta) = \eta(\text{sym}(s \otimes s))/B(s, s)$ . Together these imply that

$$B(s, s) \eta(\text{sym}(u \otimes u')) = B(u, u') \eta(\text{sym}(s \otimes s)).$$

Since  $\eta$  was arbitrary, it follows that  $B(s, s) \operatorname{sym}(u \otimes u') = B(u, u') \operatorname{sym}(s \otimes s)$  for all  $u, u' \in U$ .  $\square$

**A.2. Symmetrizing sums of squares.** Let  $\mathbb{R}[x]$  be the polynomial ring in  $q$  indeterminates and let  $G$  be a finite group acting linearly on  $\mathbb{R}^q$ . Then  $G$  acts on  $\mathbb{R}[x]$  via  $(g \cdot \mathfrak{p})(x) = \mathfrak{p}(g \cdot x)$ . If  $\mathcal{I}$  be an ideal invariant under the action of  $G$ , then the action of  $G$  descends to the quotient  $\mathbb{R}[x]/\mathcal{I}$ . Let  $V$  be a finite-dimensional  $G$ -invariant subspace of  $\mathbb{R}[x]/\mathcal{I}$ . Let  $V \otimes V$  be the tensor product (over  $\mathbb{R}$ ) of  $V$  with itself. Given  $f_1, f_2 \in V$ , their product  $f_1 f_2 \in \mathbb{R}[x]/\mathcal{I}$  is a  $\mathbb{R}$ -bilinear map  $V \times V \rightarrow \mathbb{R}[x]/\mathcal{I}$ . Hence there is a linear map  $M_V : V \otimes V \rightarrow \mathbb{R}[x]/\mathcal{I}$  such that

$$f_1 f_2 = M_V(f_1 \otimes f_2).$$

The actions of  $G$  on  $V \otimes V$  and on  $\mathbb{R}[x]/\mathcal{I}$  are such that  $M_V$  is a  $\mathbb{R}G$ -module homomorphism.

As before, let  $(S_\lambda)_{\lambda \in \Lambda}$  be an enumeration of inequivalent irreducible  $\mathbb{R}G$ -modules. Let  $V = \bigoplus_{\lambda \in \Lambda} V_\lambda$  be the isotypic decomposition of  $V$ . We now show that if we take functions from two different isotypic components of  $V$  and symmetrize their product, the result is zero.

**Lemma A.5.** *If  $\lambda \neq \mu$ ,  $f_\lambda \in V_\lambda$ , and  $f_\mu \in V_\mu$ , then  $\operatorname{sym}(f_\lambda f_\mu) = 0$ .*

*Proof.* First, let  $\phi_\mu \in \operatorname{Hom}_G(S_\mu, V)$  and  $\phi_\lambda \in \operatorname{Hom}_G(S_\lambda, V)$ , and let  $u_\mu \in S_\mu$  and  $u_\lambda \in S_\lambda$ . Then, since  $M_V$  and  $\phi_\mu \otimes \phi_\lambda$  are both  $\mathbb{R}$ -linear and commute with the action of  $G$ , we have that

$$\operatorname{sym}(\phi_\mu(u_\mu)\phi_\lambda(u_\lambda)) = \operatorname{sym}(M_V((\phi_\mu \otimes \phi_\lambda)(u_\mu \otimes u_\lambda))) = M_V((\phi_\mu \otimes \phi_\lambda)(\operatorname{sym}(u_\mu \otimes u_\lambda))).$$

Since  $\mu \neq \lambda$ , Lemma A.3 tells us that  $\operatorname{sym}(u_\mu \otimes u_\lambda) \in (S_\mu \otimes S_\lambda)^G = \{0\}$  and so  $\operatorname{sym}(\phi_\mu(u_\mu)\phi_\lambda(u_\lambda)) = 0$ . Finally, because any  $f_\mu \in V_\mu$  is a linear combination of elements of the form  $\phi_\mu(u_\mu)$  for  $\phi_\mu \in \operatorname{Hom}_G(S_\mu, V)$  and  $u_\mu \in S_\mu$  (and similarly for any  $f_\lambda \in V_\lambda$ ), we have that  $\operatorname{sym}(f_\mu f_\lambda) = 0$  by bilinearity.  $\square$

We now investigate what happens when we symmetrize products of certain elements of  $V_\lambda$ .

**Lemma A.6.** *Let  $\phi, \psi \in \operatorname{Hom}_G(S_\lambda, V)$  and  $u, u' \in S_\lambda$ . Let  $B$  be the unique (up to scale),  $G$ -invariant bilinear form on  $S_\lambda$ . Then, for any  $s_\lambda \in S_\lambda$ ,*

$$B(s_\lambda, s_\lambda) \operatorname{sym}(\phi(u)\psi(u')) = B(u, u') \operatorname{sym}(\phi(s_\lambda)\psi(s_\lambda)).$$

*Proof.* Since  $M_V$  and  $\phi \otimes \psi$  are both  $\mathbb{R}$ -linear and commute with the action of  $G$ ,

$$\operatorname{sym}(\phi(u)\psi(u')) = \operatorname{sym}(M_V((\phi \otimes \psi)(u \otimes u'))) = M_V((\phi \otimes \psi)(\operatorname{sym}(u \otimes u'))).$$

By substituting  $u = u' = s_\lambda$ , we have

$$\operatorname{sym}(\phi(s_\lambda)\psi(s_\lambda)) = M_V((\phi \otimes \psi)(\operatorname{sym}(s_\lambda \otimes s_\lambda))).$$

Relating  $\operatorname{sym}(u \otimes u')$  and  $\operatorname{sym}(s_\lambda \otimes s_\lambda)$  via Lemma A.4, and using the fact that  $M_V$  and  $\phi \otimes \psi$  are both  $\mathbb{R}$ -linear, we obtain the stated result.  $\square$

Our next aim is to describe the structure of  $\operatorname{sym}(f^2)$  for  $f \in V_\lambda$ .

**Proposition A.7.** *Let  $\phi_1, \phi_2, \dots, \phi_{m_\lambda}$  be a basis for  $\operatorname{Hom}_G(S_\lambda, V)$ , let  $s_\lambda \in S_\lambda$  be non-zero, and let*

$$Y_{j\ell}^\lambda = \operatorname{sym}(\phi_j(s_\lambda)\phi_\ell(s_\lambda))$$

for all  $j, \ell \in [m_\lambda]$ . If  $\mathbf{f} \in V_\lambda$ , then there exists a  $m_\lambda \times m_\lambda$  psd matrix  $Q^\lambda$  such that

$$\text{sym}(\mathbf{f}^2) = \sum_{j, \ell \in [m_\lambda]} Q_{j\ell}^\lambda Y_{j\ell}^\lambda.$$

*Proof.* Let  $\mathbf{f} \in V_\lambda = \{\phi(u) : \phi \in \text{Hom}_G(S_\lambda, V), u \in S_\lambda\}$ . Then a straightforward argument shows that there are  $u_1, u_2, \dots, u_{m_\lambda} \in S_\lambda$  such that

$$\mathbf{f} = \sum_{k \in [m_\lambda]} \phi_k(u_k).$$

We expand  $\text{sym}(\mathbf{f}^2)$  in terms of  $\text{sym}(\phi_i(u_i)\phi_j(u_j))$  and apply Lemma A.6 to get

$$\text{sym}(\mathbf{f}^2) = \sum_{i, j \in [m_\lambda]} \text{sym}(\phi_i(u_i)\phi_j(u_j)) = \sum_{i, j \in [m_\lambda]} \frac{B(u_i, u_j)}{B(s_\lambda, s_\lambda)} \text{sym}(\phi_i(s_\lambda)\phi_j(s_\lambda)).$$

Since  $u \mapsto B(u, u)$  is a positive definite quadratic form and  $s_\lambda \neq 0$ , if we define  $Q_{ij}^\lambda := B(u_i, u_j)/B(s_\lambda, s_\lambda)$ , we see that  $Q^\lambda$  is psd. Since  $Y_{ij}^\lambda = \text{sym}(\phi_i(s_\lambda)\phi_j(s_\lambda))$ , we have obtained an expression for  $\text{sym}(\mathbf{f}^2)$  in the required form.  $\square$

We have seen how to symmetrize products from different isotypic components of  $V$  and how to symmetrize squares of polynomials from the same isotypic component. We can now combine these earlier arguments in a straightforward way to describe symmetry-reduced non-negativity certificates for sos that are invariant under the action of  $G$ . This is the main result of the appendix (which is essentially the main result of [GP04]).

**Theorem A.8.** *Let  $V$  be a finite-dimensional  $G$ -invariant subspace of  $\mathbb{R}[x]/\mathcal{I}$  with isotypic decomposition  $V = \bigoplus_{\lambda \in \Lambda} V_\lambda$  and corresponding multiplicities  $m_\lambda$ . For each  $\lambda \in \Lambda$ , fix a non-zero element  $s_\lambda \in S_\lambda$ . Let  $\mathbf{b}_1, \dots, \mathbf{b}_{m_\lambda}$  be a basis for the subspace  $W_{s_\lambda}$ . Define for each  $\lambda \in \Lambda$*

$$Y_{ij}^\lambda = \text{sym}(\mathbf{b}_i \mathbf{b}_j)$$

for  $i, j \in [m_\lambda]$ . Suppose  $\mathbf{p} \in \mathbb{R}[x]/\mathcal{I}$  is invariant under the action of  $G$  and is  $V$ -sos. Then there exist  $m_\lambda \times m_\lambda$  psd matrices  $Q^\lambda$  such that

$$\mathbf{p} = \sum_{\lambda \in \Lambda} \sum_{i, j \in [m_\lambda]} Q_{ij}^\lambda Y_{ij}^\lambda.$$

*Proof.* Fix  $\lambda \in \Lambda$ . By Lemma A.1, the map  $\phi \mapsto \phi(s_\lambda)$  gives an isomorphism between  $W_{s_\lambda}$  and  $\text{Hom}_G(S_\lambda, V)$ . Hence, for each  $i \in [m_\lambda]$ , there is  $\phi_i \in \text{Hom}_G(S_\lambda, V)$  such that  $\mathbf{b}_i = \phi_i(s_\lambda)$ . Moreover, the  $\phi_i$  for  $i \in [m_\lambda]$  form a basis for  $\text{Hom}_G(S_\lambda, V)$ . As such,

$$Y_{ij}^\lambda = \text{sym}(\mathbf{b}_i \mathbf{b}_j) = \text{sym}(\phi_i(s_\lambda)\phi_j(s_\lambda))$$

for  $i, j \in [m_\lambda]$ . Now suppose that  $\mathbf{p}$  is  $G$ -invariant and  $V$ -sos. Then, since  $V = \bigoplus_{\lambda \in \Lambda} V_\lambda$ , we have that  $\mathbf{f}_k = \sum_{\lambda \in \Lambda} \mathbf{f}_{k\lambda}$  where each  $\mathbf{f}_{k\lambda} \in V_\lambda$ . Hence

$$\mathbf{p} = \sum_k \sum_{\lambda, \mu \in \Lambda} \mathbf{f}_{k\lambda} \mathbf{f}_{k\mu}.$$

From Lemma A.5, we have that  $\text{sym}(\mathbf{f}_{k\lambda} \mathbf{f}_{k\mu}) = 0$  whenever  $\lambda \neq \mu$ . Using the fact that  $\mathbf{p}$  is fixed by the action of  $G$ ,

$$\mathbf{p} = \text{sym}(\mathbf{p}) = \sum_k \sum_{\lambda, \mu \in \Lambda} \text{sym}(\mathbf{f}_{k\lambda} \mathbf{f}_{k\mu}) = \sum_k \sum_{\lambda \in \Lambda} \text{sym}(\mathbf{f}_{k\lambda}^2)$$

where each  $f_{k\lambda} \in V_\lambda$ . Applying Proposition A.7 for each  $\lambda$  and each  $k$ , we see that there exist psd matrices  $Q^{\lambda,k}$  such that

$$\mathbf{p} = \sum_k \sum_{\lambda \in \Lambda} \sum_{i,j \in [m_\lambda]} Y_{ij}^\lambda Q_{ij}^{\lambda,k} = \sum_{\lambda \in \Lambda} \sum_{j, \ell \in [m_\lambda]} Y_{ij}^\lambda \left( \sum_k Q_{ij}^{\lambda,k} \right).$$

Defining  $Q^\lambda = \sum_k Q^{\lambda,k}$ , which is again psd, gives an expression for  $\mathbf{p}$  as

$$\mathbf{p} = \sum_{\lambda \in \Lambda} \sum_{i,j \in [m_\lambda]} Y_{ij}^\lambda Q_{ij}^\lambda$$

as we require.  $\square$

**A.3. Specialization to Theorem 2.2.** This specializes to give Theorem 2.2 as follows. Let  $q = \binom{n}{2}$  and let  $\mathcal{I}_n = \langle x_{ij}^2 = x_{ij} \mid 1 \leq i < j \leq n \rangle$  denote the square-free ideal in  $\mathbb{R}[x]$ . Let  $\mathbb{R}[x]/\mathcal{I}_n = \mathbb{R}[\mathcal{V}_n]$  be the corresponding quotient of the polynomial ring. Let  $G = \mathfrak{S}_n$  and suppose that  $\mathfrak{S}_n$  acts on monomials by  $\mathfrak{s} \cdot x_{ij} = x_{\mathfrak{s}(i)\mathfrak{s}(j)}$ . Let  $V = \mathbb{R}[\mathcal{V}_n]_{\leq d}$  be the  $\mathfrak{S}_n$ -invariant subspace of square-free polynomials of degree at most  $d$ .

To relate Theorem 2.2 to Theorem A.8, we need to show that the subspaces  $W_{\tau_\lambda}$  defined in Section 2.1 are isomorphic as vector spaces to  $W_s$  for some non-zero  $s \in S_\lambda$ , and hence to the multiplicity space  $\text{Hom}_{\mathfrak{S}_n}(S_\lambda, V)$  (via Lemma A.1). The argument requires the following fact, which follows from Young's rule (see, e.g., [RST15, Theorem 4.6] for a statement of Young's rule in the required form) and the fact that the diagonal Kostka numbers  $K_{\lambda\lambda}$  are one (see, e.g., [Sag01, Example 2.11.4]).

**Lemma A.9.** *If  $\tau_\lambda$  is a tableau of shape  $\lambda$ , then  $\dim(S_\lambda^{\mathfrak{R}\tau_\lambda}) = 1$ .*

**Lemma A.10.** *Fix a standard tableau  $\tau_\lambda$  of shape  $\lambda$ . If  $s_{\tau_\lambda} \in S_\lambda^{\mathfrak{R}\tau_\lambda}$  is non-zero, then*

$$V_\lambda^{\mathfrak{R}\tau_\lambda} = W_{\tau_\lambda} = W_{s_{\tau_\lambda}} = \{\phi(s_{\tau_\lambda}) : \phi \in \text{Hom}_{\mathfrak{S}_n}(S_\lambda, V)\}.$$

*Proof.* If  $w \in W_{s_{\tau_\lambda}}$  then there is  $\phi \in \text{Hom}_{\mathfrak{S}_n}(S_\lambda, V)$  such that  $\phi(s_{\tau_\lambda}) = w$ . If  $\mathfrak{s} \in \mathfrak{R}\tau_\lambda$  then

$$\mathfrak{s} \cdot w = \mathfrak{s} \cdot \phi(s_{\tau_\lambda}) = \phi(\mathfrak{s} \cdot s_{\tau_\lambda}) = \phi(s_{\tau_\lambda}) = w.$$

Hence  $W_{s_\lambda} \subseteq V_\lambda^{\mathfrak{R}\tau_\lambda} = W_{\tau_\lambda}$ . On the other hand, if  $w \in W_{\tau_\lambda}$ , then there are  $v_1, \dots, v_{m_\lambda} \in S_\lambda$  and  $\phi_1, \dots, \phi_{m_\lambda}$  such that  $w = \sum_{i \in [m_\lambda]} \phi_i(v_i) \in V_\lambda$  and  $w = \text{sym}_{\mathfrak{R}\tau_\lambda}(w)$ . Then

$$w = \text{sym}_{\mathfrak{R}\tau_\lambda}(w) = \sum_{i \in [m_\lambda]} \text{sym}_{\mathfrak{R}\tau_\lambda}(\phi_i(v_i)) = \sum_{i \in [m_\lambda]} \phi_i(\text{sym}_{\mathfrak{R}\tau_\lambda}(v_i)).$$

Since  $S_\lambda^{\mathfrak{R}\tau_\lambda}$  is one-dimensional, it is spanned by  $s_{\tau_\lambda}$ . Hence for any  $v_i \in S_\lambda$ , we have that  $\text{sym}_{\mathfrak{R}\tau_\lambda}(v_i)$  is a scalar multiple of  $s_{\tau_\lambda}$ . It then follows that  $w \in W_{s_{\tau_\lambda}}$ , giving the reverse inclusion.  $\square$

Applying Theorem A.8 in this setting, and choosing  $s_\lambda := s_{\tau_\lambda}$  (defined in Lemma A.10) for each  $\lambda$ , yields the statement of Theorem 2.2.

APPENDIX B. RAMSEY NUMBER  $R(3, 3)$ 

In this section, we prove the following lemma.

**Lemma B.1.**

$$\left(2 - \sum_{2 \leq i \leq 6} x_{1i}\right)^2 + \left(2 - \sum_{2 \leq i \leq 6} (1 - x_{1i})\right)^2 \equiv -1 \pmod{\mathcal{I}}.$$

Recall that the ideal in Section 5.2 was

$$\begin{aligned} \mathcal{I} = & \langle x_e^2 - x_e \ \forall e \in E(K_6) \rangle + \langle x_{ij}x_{ik}x_{jk} \ \forall i < j < k \in [6] \rangle \\ & + \langle (1 - x_{ij})(1 - x_{ik})(1 - x_{jk}) \ \forall i < j < k \in [6] \rangle. \end{aligned}$$

The following fact saves us from duplicating all of our arguments.

**Lemma B.2.**  $\mathfrak{p}((x_{ij})_{1 \leq i < j \leq 6}) \in \mathcal{I}$  if and only if  $\mathfrak{p}((1 - x_{ij})_{1 \leq i < j \leq 6}) \in \mathcal{I}$ .

*Proof.* This follows by exchanging the colors of the edges. This can also be observed algebraically by working with the ideal.  $\square$

We now note that claws are in the ideal, allowing us to relate certain degree two and degree one elements of the ideal.

**Lemma B.3.** For  $i \in [6]$ , and  $j < k < l \in [6] \setminus \{i\}$ ,

$$x_{ij}x_{ik}x_{il} \in \mathcal{I} \text{ and } (1 - x_{ij})(1 - x_{ik})(1 - x_{il}) \in \mathcal{I}.$$

Consequently,

$$1 - \sum_{2 \leq i \leq 6} x_{1i} + \sum_{2 \leq i < j \leq 6} x_{1i}x_{1j} \in \mathcal{I}.$$

*Proof.* Indeed, combinatorially, if  $x_{ij}x_{ik}x_{il}$  were one for some  $i, j, k, l$ , then edges  $\{i, j\}$ ,  $\{i, k\}$ , and  $\{i, l\}$  would be colored red. Then if we color any of  $\{j, k\}$ ,  $\{k, l\}$  and  $\{j, l\}$  red, we would have a red triangle spanned by that edge and vertex  $i$ , but if we color all of them blue, then we would have a blue  $j, k, l$ -triangle. We can observe the same thing symmetrically for blue edges.

Algebraically, observe that  $x_{jk}x_{jl}x_{kl} \in \mathcal{I}$  and  $(1 - x_{jk})(1 - x_{jl})(1 - x_{kl}) \in \mathcal{I}$  implies that

$$\begin{aligned} 1 & \equiv x_{jk} + x_{jl} + x_{kl} - (x_{jk}x_{jl} + x_{jk}x_{kl} + x_{jl}x_{kl}) + x_{jk}x_{jl}x_{kl} \\ & \equiv x_{jk} + x_{jl} + x_{kl} - (x_{jk}x_{jl} + x_{jk}x_{kl} + x_{jl}x_{kl}) \pmod{\mathcal{I}}. \end{aligned}$$

Hence

$$\begin{aligned} x_{ij}x_{ik}x_{il}(1) & \equiv x_{ij}x_{ik}x_{il}(x_{jk} + x_{jl} + x_{kl} - (x_{jk}x_{jl} + x_{jk}x_{kl} + x_{jl}x_{kl})) \\ & \equiv 0 \pmod{\mathcal{I}} \end{aligned}$$

since every monomial on the right hand side contains either  $x_{ij}x_{ik}x_{jk}$  or  $x_{ij}x_{il}x_{jl}$  or  $x_{ik}x_{il}x_{kl}$ , each of which is in the ideal. Thus  $x_{ij}x_{ik}x_{il} \in \mathcal{I}$ . By Lemma B.2,  $(1 - x_{ij})(1 - x_{ik})(1 - x_{il}) \in \mathcal{I}$  as well.

Since  $(1 - x_{1i})(1 - x_{1j})(1 - x_{1l}) \in \mathcal{I}$  whenever  $1 < i < j < k \leq 6$ , we have that

$$0 \equiv \prod_{2 \leq i \leq 6} (1 - x_{1i}) \pmod{\mathcal{I}}.$$

Expanding the right hand side and using the fact that  $x_{1i}x_{1j}x_{1k} \in \mathcal{I}$  whenever  $1 < i < j < k \leq 6$ , we obtain

$$0 \equiv 1 - \sum_{2 \leq i \leq 6} x_{1i} + \sum_{2 \leq i < j \leq 6} x_{1i}x_{1j} \pmod{\mathcal{I}}$$

since all of the higher order terms must vanish.  $\square$

*Proof of Lemma B.1:*

$$\begin{aligned} \left(2 - \sum_{2 \leq i \leq 6} x_{1i}\right)^2 &= 4 - 4 \sum_{2 \leq i \leq 6} x_{1i} + \sum_{2 \leq i \leq 6} x_{1i}^2 + 2 \sum_{2 \leq i < j \leq 6} x_{1i}x_{1j} \\ &\equiv 4 - 4 \sum_{2 \leq i \leq 6} x_{1i} + \sum_{2 \leq i \leq 6} x_{1i} + 2(-1 + \sum_{2 \leq i \leq 6} x_{1i}) \pmod{\mathcal{I}} \\ &= 2 - \sum_{2 \leq i \leq 6} x_{1i}. \end{aligned}$$

By Lemma B.2,  $\left(2 - \sum_{2 \leq i \leq 6} (1 - x_{1i})\right)^2 \equiv 2 - \sum_{2 \leq i \leq 6} (1 - x_{1i})$ . Therefore,

$$\begin{aligned} \left(2 - \sum_{2 \leq i \leq 6} x_{1i}\right)^2 + \left(2 - \sum_{2 \leq i \leq 6} (1 - x_{1i})\right)^2 &\equiv 2 - \sum_{2 \leq i \leq 6} x_{1i} + 2 - \sum_{2 \leq i \leq 6} (1 - x_{1i}) \\ &= 4 - 5 = -1 \pmod{\mathcal{I}}. \end{aligned}$$

$\square$

For completeness, we explicitly wrote out how we used the ideal in the above sum of squares, putting the elements of the ideal in red. These calculations can be found online: [http://www.math.washington.edu/~thomas/ramsey\\_calculations.pdf](http://www.math.washington.edu/~thomas/ramsey_calculations.pdf).