

# ALGEBRAIC COMBINATORICS


Annie Raymond, Mohit Singh & Rekha R. Thomas

**Symmetry in Turán sums of squares polynomials from flag algebras**

Volume 1, issue 2 (2018), p. 249-274.

[http://alco.centre-mersenne.org/item/ALCO\\_2018\\_\\_1\\_2\\_249\\_0](http://alco.centre-mersenne.org/item/ALCO_2018__1_2_249_0)

© The journal and the authors, 2018.  
*Some rights reserved.*

 This article is licensed under the  
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.  
<http://creativecommons.org/licenses/by/4.0/>

Access to articles published by the journal *Algebraic Combinatorics* on  
the website <http://alco.centre-mersenne.org/> implies agreement with the  
Terms of Use (<http://alco.centre-mersenne.org/legal/>).



*Algebraic Combinatorics* is member of the  
Centre Mersenne for Open Scientific Publishing  
[www.centre-mersenne.org](http://www.centre-mersenne.org)



# Symmetry in Turán sums of squares polynomials from flag algebras

Annie Raymond, Mohit Singh & Rekha R. Thomas

**ABSTRACT** Turán problems in extremal combinatorics ask to find asymptotic bounds on the edge densities of graphs and hypergraphs that avoid specified subgraphs. The theory of flag algebras proposed by Razborov provides powerful methods based on semidefinite programming to find sums of squares that establish edge density inequalities in Turán problems. Working with polynomial analogs of the flag algebra entities, we prove that such sums of squares created by flag algebras can be retrieved from a restricted version of the symmetry-adapted semidefinite program proposed by Gatermann and Parrilo. This involves using the representation theory of the symmetric group for finding succinct sums of squares expressions for invariant polynomials. The connection reveals several combinatorial and structural properties of flag algebra sums of squares, and offers new tools for Turán and other related problems.

## 1. INTRODUCTION

The *Turán problem* from extremal combinatorics asks the following question: given a graph  $A$ , what is the maximum number of edges in a graph on  $n$  vertices not containing  $A$  as a subgraph? Turán [26] answered this question for  $A = K_s$ , the complete graph on  $s$  vertices, generalizing a classical result of Mantel [14] for triangle-free graphs, and establishing the field of extremal graph theory. In general, for any graph  $A$ , Erdős and Stone [5] identified the maximum possible density of edges in any  $A$ -free graph asymptotically. The *hypergraph Turán problem* asks the same question for hypergraphs, but the current understanding of this problem is far from satisfactory. In particular, even asymptotically, tight bounds on the maximum number of edges in a  $n$ -vertex 3-uniform hypergraph<sup>(1)</sup> not containing a complete graph of size four is not known. A variety of general techniques have been developed to prove bounds for this long-standing hypergraph Turán problem; see for example [3, 8, 13, 16, 25, 11] for a survey.

Recently, semidefinite programming methods arising from the powerful theory of *flag algebras* introduced by Razborov [18] have led to significant progress on this problem. Indeed, many of the previous bounds can be proven via this technique and several new results giving the tightest known bounds have been obtained [19, 20, 21, 6]. For instance, Razborov in [19] proved that the (maximum) asymptotic edge density of a 3-uniform hypergraph without a 4-clique is 0.561666 (Turán [27] conjectured it to be  $\frac{5}{9}$ ). Moreover, he also showed in the same paper that, if one forbids an additional

---

*Manuscript received 24th August 2017, accepted 9th October 2017.*

**KEYWORDS.** sums of squares, semidefinite programming, flag algebras, extremal combinatorics, symmetry.

<sup>(1)</sup>A  $r$ -uniform hypergraph has (hyper)edges of size  $r$ .

subgraph, then the asymptotic edge density is indeed  $\frac{5}{9}$ . Razborov's method relies on establishing inequalities involving densities of suitably chosen subgraphs in any  $n$ -vertex graph/hypergraph. This is done by lower bounding density expressions with a scalar sum of squares (sos) coming from flags. A suitable sos is found by formulating a semidefinite program (SDP) whose size depends on the flags that are used. The key to the success of this method is that the size of the SDP is thus independent of the number of vertices, which is particularly helpful for asymptotic results. However, deciding which flags are needed to construct the flag sos expressions is an art.

Our work is motivated by the basic question as to whether there is a fundamental connection between Razborov's scalar flag sos methods and the more standard sos theory for polynomials. Expressing a polynomial with real coefficients as a sos of polynomials in order to certify its nonnegativity is a well-established technique in real algebraic geometry going back at least to the 19th century. In recent years, these ideas have acquired new life following the realization that sos polynomials can be found via the modern tool of semidefinite programming which has led to remarkable progress in optimization and algorithm design. For an introduction to these methods, see one of [2, Chapters 1 & 2] or [12, 15].

In this paper, we show that indeed there is a deep connection between the sos methods coming from flags and those for polynomials in real algebraic geometry. We show that symmetry-reduction in polynomial optimization is precisely the right framework through which this relationship can be established. This brings in tools from the representation theory of the symmetric group, highlighting the many combinatorial features of flag sos expressions.

Symmetry-reduction in polynomial optimization, or more generally semidefinite programming, is a powerful technique and has been useful in many settings [24, 1, 4, 10]. When a nonnegative polynomial is invariant under the action of a finite group, the representation theory of the group can be used to simplify the SDP used to obtain the sos certificate for its nonnegativity. In [10], Gatermann and Parrilo show that in this invariant setting the original SDP breaks into several smaller (but coupled) SDPs, each indexed by an *irreducible representation* of the group, leading to tremendous computational savings. We appeal to this framework to establish our results.

Our main technical result shows that the flag algebra method for establishing graph density inequalities embeds naturally in a restricted version of the Gatermann-Parrilo symmetry-adapted SDP. Our results rely on the rich combinatorics hidden in the sos expressions coming from flag algebras that we expose using the representation theory of the symmetric group. We give a precise description of the symmetry-reduced SDP in terms of the irreducible representations of the symmetric group and show that only certain irreducibles are needed. Consequently, we prove that the size of this SDP is independent of the number of vertices in the associated graphs, as in Razborov's methods. This offers a systematic way of establishing graph density inequalities, and more generally, Cauchy-Schwarz proofs coming from flags, through standard sos methods where no sophisticated choices are necessary.

1.1. OUR RESULTS IN DETAIL AND THE ORGANIZATION OF THE PAPER. For notational simplicity, we restrict the exposition in this paper to Turán-type problems in the setting of graphs. Our results extend naturally to the broader realm of hypergraphs, digraphs, tournaments, etc, which carry many open problems. In the conclusion of this paper, we will elaborate on the modifications needed for these extensions.

Restricting to the setting of graphs, assume that there is only one graph  $A$  that must be avoided in a Turán problem. The more general setting of avoiding all graphs in a family is treated similarly. Also, we will work in the setting of avoiding  $A$  as an

induced subgraph. Note that this setting is general, and that the non-induced setting can be modeled through it by forbidding every induced graph containing  $A$ . The main technical challenge in Turán problems is to show an upper bound on the edge density of any  $A$ -free graph.

The first step in linking flag algebra methods for Turán problems to the symmetry-reduction techniques of [10] is to view graph density expressions as polynomials modulo an ideal, and flag sos expressions as polynomial sos expressions. This is done in Section 2.

We begin with a few basic definitions. For a fixed positive integer  $n$ , the polynomials we work with lie in  $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_{ij} : 1 \leq i < j \leq n]$ , the polynomial ring over  $\mathbb{R}$  in  $\binom{n}{2}$  variables indexed by the edges in the complete graph  $K_n$ . If  $A$  denotes the graph that must be avoided in the Turán problem we are interested in, then the ideal we need,  $\mathcal{I}_n^A$ , is precisely the set of polynomials in  $\mathbb{R}[\mathbf{x}]$  that vanish on the characteristic vectors of all graphs on  $n$  vertices that do not contain  $A$  as an induced subgraph. A polynomial is nonnegative on this finite set of characteristic vectors if and only if it is equivalent to a sos polynomial modulo  $\mathcal{I}_n^A$ .

Two polynomials  $f$  and  $g$  are equivalent modulo  $\mathcal{I}_n^A$  if and only if  $f - g \in \mathcal{I}_n^A$ , written as  $f \equiv g \pmod{\mathcal{I}_n^A}$ . This equivalence relation differentiates between functions on the zeros of  $\mathcal{I}_n^A$ , i.e.,  $f \equiv g \pmod{\mathcal{I}_n^A}$  if and only if  $f(\mathbf{v}) = g(\mathbf{v})$  for every zero  $\mathbf{v}$  of  $\mathcal{I}_n^A$ . Therefore, if  $g = \sum h_j^2$  is a sos, then  $f$  is nonnegative on the characteristic vectors of all  $A$ -free graphs on  $n$  vertices. Further, we say that  $f$  is  $d$ -sos mod  $\mathcal{I}_n^A$  if each  $h_j$  has degree at most  $d$ . Every  $d$ -sos polynomial has the form  $\mathbf{y}^T Q \mathbf{y}$  where  $Q$  is a  $l \times l$  positive semidefinite (psd) matrix and  $\mathbf{y}$  is a  $l$ -dimensional vector whose coordinates are polynomials of degree at most  $d$ . This allows us to use semidefinite programming to search for  $d$ -sos expressions for a given polynomial modulo  $\mathcal{I}_n^A$ .

In Section 2, we describe the polynomial analogs of the ingredients in a sos proof coming from flag algebras. This allows us to translate flag sos expressions to polynomial sos expressions for density polynomials as shown in Propositions 2.4 and 2.5. We illustrate our polynomial translation on Mantel's theorem. The complete proof, including the flag algebra approach for Turán problems, can be found in the Appendix.

Given the polynomial formulation, an approach to showing an upper bound on the graph density polynomial is to use the standard sos method in polynomial optimization. This raises several natural questions. Firstly, whether flag sos expressions can be retrieved via this approach? Secondly, whether the size of the SDP formulated in this approach will be independent of  $n$  as in the flag algebra framework? In this paper, we answer both questions affirmatively.

A priori, searching for a  $d$ -sos proof leads to a SDP formulation whose size grows with  $n$ . The first step in establishing our result is to notice that the graph density polynomial whose nonnegativity we need to establish is invariant under an action induced by the symmetric group  $\mathfrak{S}_n$  on  $n$  letters acting on the vertices of  $K_n$ . Therefore, one can use the symmetry-reduction techniques in [10] to simplify the computational cost of searching for its sos certificate. This in turn relies on the representation theory of  $\mathfrak{S}_n$ ; we explain the basics of this theory in Section 3.1. We then describe the symmetry-reduction strategy of [10] in Section 3.2 which breaks the SDP that searches for a sos expression into smaller SDPs that are indexed by the irreducible representations of  $\mathfrak{S}_n$  or, equivalently, the partitions of  $n$ . This section is largely expository but it is crucial for understanding our main results in Section 4. For efficiency, we tailor all discussion of [10] to  $\mathfrak{S}_n$  which in turn creates a new set of combinatorial tools for problems to which flag sos methods apply.

In Section 4, we come to our main results. Suppose we fix a maximum degree  $d$  for the sos polynomials we are searching for, and let  $\mathbb{R}[\mathbf{x}]_{\leq d}$  denote the vector space of

all polynomials in  $\mathbb{R}[\mathbf{x}]$  of degree up to  $d$ . The group  $\mathfrak{S}_n$  breaks  $\mathbb{R}[\mathbf{x}]_{\leq d}$  into a direct sum of subspaces indexed by the partitions of  $n$ , called the isotypic decomposition of  $\mathbb{R}[\mathbf{x}]_{\leq d}$ . We first establish that the atomic pieces of the sos polynomials that come out of flag algebras are invariant with respect to the *row group* of a *tableau* defined from the flags that are chosen by the flag algebra method (Theorem 4.3). Next, using the theory of restricted representations, we decide which subset of subspaces in the isotypic decomposition of  $\mathbb{R}[\mathbf{x}]_{\leq d}$  contain the polynomials in a flag algebra sos in their span. Finally, we show that the sos expressions from flag algebras can be retrieved from the Gatermann-Parrilo SDP restricted to those partitions that survive in the previous step. The key point to note here is that the number of partitions indexing these necessary subspaces is not a function of  $n$ , but rather of  $d$ . So if we fix  $d$  and let  $n$  go to infinity, the number of partitions, and the sizes of the corresponding SDPs, will stay fixed. This answers the two questions raised above and links flag algebra methods to symmetry-reduction in semidefinite programming. Under the usual action of  $\mathfrak{S}_n$  on the polynomial ring  $\mathbb{R}[x_1, \dots, x_n]$ , [22] has also shown that one can get sos expressions for symmetric polynomials whose size is independent of  $n$ .

In Section 5, we illustrate our main results on the Mantel example. We will see that for any  $n$ , just two specific partitions are enough to obtain the sos expression from flag algebras. Since there are many simple proofs of Mantel's result, the goal of using this example is simply to illustrate the chain of results that make up this paper. It is both simple and rich enough for this purpose.

We conclude in Section 6 where we explain how to apply our techniques in other settings such as directed graphs, hypergraphs and tournaments, and also discuss different future directions for this work.

## 2. SUMS OF SQUARES FROM FLAG ALGEBRAS

In this section we present the polynomial analogs of Razborov's flag algebra method for proving inequalities on graph densities. As mentioned in the Introduction, we restrict our attention to graphs. For generalizations, see the conclusion of the paper. The key new feature distinguishing this work from the literature ([6, 18, 21]) is that we think of all densities as polynomials that can be evaluated on characteristic vectors of graphs. If the reader is unfamiliar with flag algebras, we highly recommend reading Section 2.2 of [6] first to see a concrete application of the flag method to Mantel's theorem. The polynomial version may appear more difficult to parse at first but they are simply functions on the (finite) set of characteristic vectors of the graphs allowed by the problem that evaluate to density expressions as in [6] and [19].

The polynomial translation illustrates how certificates based on flag algebras can be interpreted as polynomial sos proofs. The polynomials appearing in Razborov's sos proofs and their specific symmetries will be key in obtaining our main results efficiently via the Gatermann-Parrilo framework introduced in Section 3. To illustrate the method on an example, we use the polynomial version of flag algebras to prove Mantel's theorem at the end of this section.

Consider the general problem of certifying an inequality involving graph densities over all graphs with a certain property. We first fix  $n$ , the number of vertices in our graphs. Let  $\mathcal{G}$  be the set of all undirected graphs up to isomorphism with the desired property, and let  $\mathcal{G}_n$  be the graphs in  $\mathcal{G}$  that have  $n$  vertices. Also, let  $V(G)$  and  $E(G)$  denote the sets of vertices and edges of  $G$ . We represent a graph  $G$  by its characteristic vector  $\mathbb{1}_G \in \{0, 1\}^{\binom{n}{2}}$  whose  $ij$ -th coordinate is 1 if and only if  $\{i, j\} \in E(G)$ . Throughout, we work with the polynomial ring  $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_{ij} : 1 \leq i < j \leq n]$  described earlier.

Fix an integer  $m$  such that  $m < n$ , and let  $H \in \mathcal{G}_m$ . Furthermore, let  $\text{Inj}(S, [n])$  denote the set of injective maps  $\alpha : S \rightarrow [n]$  for any set  $S$ . For a fixed  $\alpha \in \text{Inj}(V(H), [n])$ , define the polynomial

$$(1) \quad \mathbf{q}_H^\alpha := \prod_{\{i,j\} \in E(H)} x_{\alpha(i)\alpha(j)} \prod_{\{i,j\} \in \binom{V(H)}{2} \setminus E(H)} (1 - x_{\alpha(i)\alpha(j)}) \in \mathbb{R}[\mathbf{x}].$$

For a graph  $G \in \mathcal{G}_n$ ,  $\mathbf{q}_H^\alpha(\mathbb{1}_G) = 0$  if and only if at least one of the following is true: (i) for some  $\{i, j\} \in E(H)$ ,  $\{\alpha(i), \alpha(j)\}$  is not an edge of  $G$  or (ii) for some  $\{i, j\} \in \binom{V(H)}{2} \setminus E(H)$ ,  $\{\alpha(i), \alpha(j)\}$  is an edge of  $G$ . If  $\mathbf{q}_H^\alpha(\mathbb{1}_G) = 1$ , we say that the vertices  $\alpha(V(H))$  label-*induce*  $H$  in  $G$ . Being label-induced is stronger than being induced, since the set of vertices  $\alpha(V(H))$  might induce (in the usual sense of the word)  $H$  in  $G$  even if  $\mathbf{q}_H^\alpha(\mathbb{1}_G) = 0$ . If  $\alpha(V(H))$  does not induce  $H$  in  $G$ , then  $\mathbf{q}_H^\alpha(\mathbb{1}_G) = 0$ .

EXAMPLE 2.1. Let  $H = \overset{0}{\bullet} \overset{3}{\bullet} \overset{2}{\bullet}$ . Suppose that  $\alpha(i) = i$  for every  $i \in \{0, 1, 2, 3\}$ . Then  $\mathbf{q}_H^\alpha(\mathbb{1}_H) = 1$ , i.e.,  $\alpha(V(H))$  label-*induces*  $H$  in  $H$ . However, if  $\alpha(0) = 0$ ,  $\alpha(1) = 1$ ,  $\alpha(2) = 3$  and  $\alpha(3) = 2$ , then  $\mathbf{q}_H^\alpha(\mathbb{1}_H) = 0$  (since, for example,  $\alpha(1)$  and  $\alpha(2)$  do not form an edge) even though  $\alpha(V(H))$  does induce  $H$  in  $H$ .

We can use the polynomials  $\mathbf{q}_H^\alpha$  for different injections  $\alpha$  to calculate the density of  $H$  in  $G$ . Define

$$(2) \quad \rho_H := \frac{1}{\binom{n}{|V(H)|}} \cdot \sum_{\substack{S \subseteq [n]: \\ |S|=|V(H)|}} \frac{1}{a_H} \sum_{\alpha \in \text{Inj}(V(H), S)} \mathbf{q}_H^\alpha$$

$$(3) \quad = \frac{1}{a_H \binom{n}{|V(H)|}} \cdot \sum_{\alpha \in \text{Inj}(V(H), [n])} \mathbf{q}_H^\alpha$$

where

$$a_H = \sum_{\alpha \in \text{Inj}(V(H), V(H))} \mathbf{q}_H^\alpha(\mathbb{1}_H).$$

The quantity  $a_H$  is the number of label-induced copies of  $H$  in itself. Thus, for any subset  $S \subseteq [n]$  of size  $|V(H)|$ , if  $S$  induces  $H$  in  $G$ , then

$$\sum_{\alpha \in \text{Inj}(V(H), S)} \mathbf{q}_H^\alpha(\mathbb{1}_G) = a_H.$$

Otherwise,  $\sum_{\alpha \in \text{Inj}(V(H), S)} \mathbf{q}_H^\alpha(\mathbb{1}_G) = 0$ . Therefore,  $\frac{1}{a_H} \sum_{\alpha \in \text{Inj}(V(H), S)} \mathbf{q}_H^\alpha(\mathbb{1}_G) \in \{0, 1\}$  and it is one if and only if  $S$  induces  $H$  in  $G$ .

EXAMPLE 2.2. Again, let  $H = \overset{0}{\bullet} \overset{3}{\bullet} \overset{2}{\bullet}$ . Out of the  $4! = 24$  maps in  $\text{Inj}(V(H), V(H))$ , eight of them label-*induce*  $H$  in itself, namely the maps  $\alpha_j$  and  $\beta_j$  for  $j \in \{0, 1, 2, 3\}$  such that  $\alpha_j(i) \equiv i + j \pmod 4$  and  $\beta_j(i) \equiv 4 - i + j \pmod 4$  for every  $i \in \{0, 1, 2, 3\}$ . Thus  $a_H = 8$ .

Hence, evaluated on  $\mathbb{1}_G$ ,  $\rho_H$  yields the density of  $H$  in  $G$ , i.e., the probability that a collection of  $m$  vertices in  $G$  chosen uniformly at random will induce a copy of  $H$  in  $G$  up to isomorphism. Note that  $\rho_H(\mathbb{1}_G) = p(H, G)$  in [18].

A *type* of size  $k$  ( $\leq n$ ) is a  $k$ -vertex graph  $\sigma$  in which every vertex is labeled with a distinct element of  $[k]$ . For an integer  $l \geq k$ , a  $\sigma$ -*flag*  $F$  of size  $l$  is a graph in  $\mathcal{G}_l$  which has  $k$  vertices labeled  $1, \dots, k$  and the bijective map that sends vertex labeled  $i$  in  $\sigma$  to the vertex labeled  $i$  in  $F$  label-*induces* a copy of  $\sigma$  in  $F$  with identical labels for the vertices. Let  $\mathcal{F}_l^\sigma$  be the set of all  $\sigma$ -flags of size  $l$  up to isomorphism.

EXAMPLE 2.3. Let  $\sigma = 2 \curvearrowright^1 3$ . Then,  $2 \curvearrowright^1 3$  is in  $\mathcal{F}_4^\sigma$ . However,  $1 \curvearrowright^2 3$  is not since  $\sigma$  is mislabeled;  $2 \curvearrowright^1 3$  also is not since vertices 1, 2, 3 do not induce  $\sigma$ .

Fix a type  $\sigma$  of size  $k$  and  $l \geq k$ . For a flag  $F \in \mathcal{F}_l^\sigma$  and an injective map  $\alpha : V(F) \rightarrow [n]$ , we have the polynomial  $\mathbf{q}_F^\alpha$  as in (1). Now suppose  $\theta$  is a fixed labeling of  $k$  vertices in  $G$  using all the labels in  $[k]$ , i.e.,  $\theta \in \text{Inj}([k], [n])$  is an injective map from  $[k]$  to  $[n] = V(G)$ . We say  $\alpha : V(F) \rightarrow [n]$  respects the labeling  $\theta$  if  $\alpha(v) = \theta(i)$  for any vertex  $v \in V(F)$  labeled  $i \in [k]$ . Let  $\text{Inj}_\theta(V(F), [n])$  denote the set of injective maps  $\alpha : V(F) \rightarrow [n]$  that respect the labeling  $\theta$ . Then we define the following density polynomial

$$(4) \quad \mathbf{p}_F^\theta := \frac{1}{\binom{n-k}{l-k}} \sum_{\substack{S \subseteq [n]: \\ |S|=|V(F)|, \\ S \supseteq \text{im}(\theta)}} \frac{1}{a_F^\sigma} \sum_{\alpha \in \text{Inj}_\theta(V(F), S)} \mathbf{q}_F^\alpha$$

$$(5) \quad = \frac{1}{a_F^\sigma \binom{n-k}{l-k}} \sum_{\alpha \in \text{Inj}_\theta(V(F), [n])} \mathbf{q}_F^\alpha,$$

where

$$a_F^\sigma = \sum_{\substack{\alpha \in \text{Inj}(V(F), V(F)): \\ \alpha(v)=v \ \forall v \in V(\sigma)}} \mathbf{q}_F^\alpha(\mathbb{1}_F).$$

The quantity  $a_F^\sigma$  is the number of label-induced copies of  $F$  in  $F$  such that every vertex in  $\sigma$  is sent to itself. Note that  $\mathbf{p}_F^\theta(\mathbb{1}_G) = p(F_1; F_2)$  in [18] where  $F_1 = (F, \theta) \in \mathcal{F}_l^\sigma$  and  $F_2 = (G, \theta) \in \mathcal{F}_n^\sigma$  where  $|\sigma| = k$ . Indeed,  $\mathbf{p}_F^\theta(\mathbb{1}_G)$  is the probability that the  $k$  vertices of  $G$  labeled by  $\theta$  along with the remaining  $l - k$  unlabeled vertices picked uniformly at random induce a copy of  $F$  in  $G$ .

Razborov’s flag algebra methods can be used to certify the nonnegativity of graph density functions, i.e, functions involving  $\mathbf{p}_H(\mathbb{1}_G)$  (or Razborov’s  $p(H, G)$ ) for different graphs  $H$ . For example, to retrieve Mantel’s theorem, we want to show that  $r = \frac{1}{2} - p(e, G)$  is nonnegative over all triangle-free graphs on  $n$  vertices where  $n \rightarrow \infty$ . Here,  $e$  is the 2-vertex graph consisting of one edge. This is done by expressing the given graph density function  $r$  as a sos of linear combinations of flag densities, thus establishing the nonnegativity of  $r$ . This in turn involves finding a psd matrix  $M$  such that  $r$  is the average over  $\theta$ ’s of  $(f_1(G), \dots, f_s(G))M(f_1(G), \dots, f_s(G))^\top$  where  $f_i(G)$  is a linear combination of  $\sigma$ -flag densities  $\mathbf{p}_F^\theta(\mathbb{1}_G)$  for different flags  $F$ . Razborov refers to such certificates as *Cauchy-Schwarz* proofs (for the nonnegativity of  $r$ ).

The main result of this section is that one can interpret Razborov’s Cauchy-Schwarz certificates as sum of squares of polynomials modulo an ideal, using the polynomial analogs of densities that we constructed in this section. First we give a high level version of this result.

PROPOSITION 2.4. *Let  $r$  be a function of graph densities that is nonnegative over a family of graphs as the the number of vertices goes to infinity, and let  $\mathbf{r}$  be the polynomial analog of  $r$ . Suppose we are given a Cauchy-Schwarz proof of the nonnegativity of  $r$  where the flags used have type  $\sigma$  and size at most  $l$ . Then  $\mathbf{r}$  can be written as a sum of squares of polynomials modulo the vanishing ideal of the family of graphs under consideration.*

*Proof.* Suppose  $r$  has a certificate of nonnegativity in a Cauchy-Schwarz proof using flags of size  $l$  and type  $\sigma$ . Then this certificate is the average of an expression of the form

$$(f_1(G), \dots, f_s(G))M(f_1(G), \dots, f_s(G))^\top$$

over all  $\theta$ , where each  $f_i(G)$  is a linear combination of  $\sigma$ -flags densities and  $M$  is psd. In Razborov's language this means that  $f_i = \sum_j b_j^i p(B_j, B)$  where  $B_j = (F_j, \theta) \in \mathcal{F}_l^\sigma$  and  $B = (G, \theta) \in \mathcal{F}_n^\sigma$ .

We now replace these expressions by their polynomial analogs. Let  $\mathcal{I}$  be the vanishing ideal of the characteristic vectors all graphs in the family being considered. Since  $f_i(G) = \sum_j b_j^i \mathbf{p}_{F_j}^\theta(\mathbf{1}_G)$ , for a graph  $G$  in our family,  $f_i(G)$  is the evaluation of the polynomial  $\mathbf{f}_i := \sum_j b_j \mathbf{p}_{F_j}^\theta$  on the characteristic vector of  $G$  which is a zero of  $\mathcal{I}$ . The averaging in the sos expression for  $r$  amounts to averaging  $(\mathbf{f}_1, \dots, \mathbf{f}_s)M(\mathbf{f}_1, \dots, \mathbf{f}_s)$  over  $\theta$ . This averaging in turn, is equivalent to taking the expectation over all maps  $\theta$ . Thus overall, we get that

$$r \equiv \mathbb{E}_\theta [(\mathbf{f}_1, \dots, \mathbf{f}_s)M(\mathbf{f}_1, \dots, \mathbf{f}_s)^\top] \pmod{\mathcal{I}}$$

which says that when evaluated on the zeros of  $\mathcal{I}$ , which are precisely the characteristic vectors of graphs in our family, the function  $r$  and the expression on the right hand side are equal. This is exactly what the scalar sos in the Cauchy-Schwarz proof was saying.  $\square$

We now write a more precise version of Proposition 2.4 that will be helpful in later sections.

PROPOSITION 2.5. *Assume the same hypotheses as in Proposition 2.4, and consider the vector of flag density polynomials  $\mathbf{p}^{\theta, \sigma, l} = (\mathbf{p}_F^\theta)_{F \in \mathcal{F}_l^\sigma}$  for all flags with type  $\sigma$ , size  $l$ , a fixed numbering  $\theta$ . Then there exists a psd matrix  $Q \in \mathbb{R}^{|\mathcal{F}_l^\sigma| \times |\mathcal{F}_l^\sigma|}$  and a sos certificate for the nonnegativity of  $r$  of the following form:*

$$(6) \quad r \equiv \mathbb{E}_\theta \left[ \mathbf{p}^{\theta, \sigma, l \top} Q \mathbf{p}^{\theta, \sigma, l} \right]$$

modulo the vanishing ideal  $\mathcal{I}$  of the family of graphs under consideration.

*Proof.* As in the proof of Proposition 2.4, we arrive at the expression

$$r \equiv \mathbb{E}_\theta [(\mathbf{f}_1, \dots, \mathbf{f}_s)M(\mathbf{f}_1, \dots, \mathbf{f}_s)^\top] \pmod{\mathcal{I}},$$

where  $\mathbf{f}_i = \sum_j b_j \mathbf{p}_{F_j}^\theta$  and  $M$  is a psd matrix.

Writing  $M = (M^{\frac{1}{2}})^\top M^{\frac{1}{2}}$  where  $M^{\frac{1}{2}} = (m_{ji})$  is a  $p \times s$  matrix, we have

$$(\mathbf{f}_1, \dots, \mathbf{f}_s)M(\mathbf{f}_1, \dots, \mathbf{f}_s)^\top = \sum_{j=1}^p \left( \sum_{i=1}^s m_{ji} \mathbf{f}_i \right)^2.$$

Since each  $\mathbf{f}_i$  is a linear combination of  $\mathbf{p}_{F_t}^\theta$ 's, there exists  $q_{ti} \in \mathbb{R}$  such that

$$(\mathbf{f}_1, \dots, \mathbf{f}_s)M(\mathbf{f}_1, \dots, \mathbf{f}_s)^\top = \sum_{j=1}^p \left( \sum_{F_t \in \mathcal{F}_l^\sigma} q_{ti} \mathbf{p}_{F_t}^\theta \right)^2 = \mathbf{p}^{\theta, \sigma, l \top} Q \mathbf{p}^{\theta, \sigma, l}$$

where  $\mathbf{p}^{\theta, \sigma, l} = (\mathbf{p}_F^\theta)_{F \in \mathcal{F}_l^\sigma}$  and  $Q$  is a psd matrix of size  $p \times |\mathcal{F}_l^\sigma|$ .  $\square$

Note that the polynomials that are squared in the sos (6), namely the components of the vector  $Q^{\frac{1}{2}} \mathbf{p}^{\theta, \sigma, l}$ , are linear combinations of the flag density polynomials  $\{\mathbf{p}_F^\theta : F \in \mathcal{F}_l^\sigma\}$ . This immediately yields some quantitative bounds on the size of  $Q$  and the degree of the sos.

COROLLARY 2.6. *The degree of the above sos-proof equals the maximum degree of the polynomials of the form  $\mathbf{p}_F^\theta$  which is at most  $\binom{l}{2}$  (for graphs) where  $l$  is the number of vertices in the  $\sigma$ -flag. Moreover, the size of  $Q$  in Proposition 2.5 depends only on the size of  $|\mathcal{F}_l^\sigma|$  and not on  $n$  (when  $k$  and  $l$  are fixed).*



We remark that certain flag Cauchy-Schwarz proofs require choosing several types  $\sigma$  and sizes of flags  $l$ , and then taking a conic combination of sos for each  $\sigma$  and  $l$ . This however does not change the explanations above since the argument stands for each sos.

We illustrate the above Propositions by providing a polynomial version of the Cauchy-Schwarz proof of Mantel’s theorem presented in [6].

EXAMPLE 2.7. Consider the problem of finding the maximum edge density of a graph which does not contain any triangle (which we denote by  $K_3$ ). Mantel’s famed result states that the maximum edge density of a triangle-free graph goes to  $\frac{1}{2}$  as the number of vertices goes to infinity. Let  $\mathcal{G}$  be the family of triangle-free graphs and  $\mathcal{G}_n$  be the set of triangle-free graphs on  $n$  vertices for a fixed  $n$ . We need to choose  $\sigma$  and  $l$  to obtain a flag Cauchy-Schwarz proof of the non-negativity of  $\frac{1}{2} - p(e, G)$  for  $G \in \mathcal{G}_n$  as  $n \rightarrow \infty$ .

The characteristic vectors of  $K_3$ -free graphs in  $\mathcal{G}_n$  are precisely the zeros of the ideal

$$\mathcal{I}_n^{K_3} = \langle x_{ij}^2 - x_{ij}, 1 \leq i < j \leq n \rangle + \langle x_{ij}x_{ik}x_{il} : 1 \leq i < j < k \leq n \rangle.$$

The polynomial  $p_n := \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij}$  evaluated on  $\mathbf{1}_G$  is equal to the edge density,  $\frac{|E(G)|}{\binom{|V(G)|}{2}}$ , of the graph  $G \in \mathcal{G}_n$ . In order to show that the edge density of any  $G \in \mathcal{G}_n$  is at most  $\beta$  (here we would like  $\beta$  to be  $\frac{1}{2} + O(\frac{1}{n})$ ), it suffices to find polynomials  $r_j$  such that

$$\beta - p_n \equiv \sum_j r_j^2 \pmod{\mathcal{I}_n^{K_3}}.$$

Translating the flag sos proof in [6], we obtain the following polynomial sos certificate

$$(7) \quad \frac{1}{2} - p_n + \text{err} \equiv \mathbb{E}_\theta \left[ \left( \begin{matrix} p_{F_0}^\theta & p_{F_1}^\theta \\ -\frac{1}{2} & \frac{1}{2} \end{matrix} \right) \left( \begin{matrix} p_{F_0}^\theta \\ p_{F_1}^\theta \end{matrix} \right) \right] + \frac{1}{3} p_{H_1} \pmod{\mathcal{I}_n^{K_3}},$$

where

$$F_0 = \mathbf{1} \bullet \bullet, \quad F_1 = \mathbf{1} \bullet \bullet, \quad H_1 = \bullet \bullet \bullet,$$

and  $\text{err}(\mathbf{1}_G)$  has value  $O(\frac{1}{n})$  for every  $G \in \mathcal{G}_n$ . The first expression on the right hand side is a sos by construction. The second expression is also a sum of squares because  $(q_H^\alpha)^2 \equiv q_H^\alpha \pmod{\mathcal{I}_n^{K_3}}$  since  $x_{ij}$  and  $(1 - x_{ij})$  are equivalent to their squares mod  $\mathcal{I}_n^{K_3}$ . Therefore,

$$(8) \quad p_H \equiv \frac{1}{a_H \binom{n}{|V(H)|}} \sum_{\alpha \in \text{Inj}(V(H), [n])} (q_H^\alpha)^2 \pmod{\mathcal{I}_n^{K_3}},$$

for every  $H$ . In particular,  $p_{H_1}$  is a sos mod  $\mathcal{I}_n^{K_3}$ . We have expressed the edge density expression on the left hand side as a polynomial sos modulo the ideal  $\mathcal{I}_n^{K_3}$ . For a verification of this equivalence as well as a full translation of [6] to polynomials, see the Appendix.

### 3. SUM OF SQUARES REPRESENTATIONS OF INVARIANT POLYNOMIALS

In [10], Gatermann and Parrilo use methods from representation theory to organize the computation of sos expressions for polynomials that are invariant with respect to a finite group. These symmetry-reduction techniques allow the SDP that provides the potential sos to be broken into several smaller SDPs that are coupled together, often leading to tremendous computational savings. Since the graph density polynomials

that arise in Turán problems are invariant under an induced action of the symmetric group  $\mathfrak{S}_n$  on  $n$  letters, we can apply the methods in [10] to provide an alternate and systematic method for establishing graph density inequalities.

Our main aim in this section is to describe the strategy and mechanics in [10] specialized to our setting of  $\mathfrak{S}_n$  acting on  $\mathbb{R}[\mathbf{x}]$ . For a full proof of the Gattermann-Parrilo method specialized to our setting, see Appendix A of [17]. To keep the paper self-contained, we assume very little background.

**3.1. REPRESENTATION THEORY OF THE SYMMETRIC GROUP.** There are many excellent expositions of the representation theory of  $\mathfrak{S}_n$ , and our brief account below is based on [23, Chapter 1]. We will reference general theorems from [23] even if we only state their specialized versions for  $\mathfrak{S}_n$ .

Recall that  $\mathbb{R}[\mathbf{x}]$  denotes the polynomial ring in the variables  $x_{ij}, 1 \leq i < j \leq n$ . Since we will be searching for  $d$ -sos polynomials for a fixed degree  $d$ , we will focus on  $\mathbb{R}[\mathbf{x}]_{\leq d}$ , the set of all polynomials in  $\mathbb{R}[\mathbf{x}]$  of degree at most  $d$ . A natural basis for this vector space is the set of all monomials of degree at most  $d$ , and hence the dimension of  $V := \mathbb{R}[\mathbf{x}]_{\leq d}$  is  $D := \binom{e+d}{d}$  where  $e := \binom{n}{2}$ . The symmetric group  $\mathfrak{S}_n$  acts on monomials in  $\mathbb{R}[\mathbf{x}]$  via  $\mathfrak{s}x_{ij} := x_{\mathfrak{s}(i)\mathfrak{s}(j)}$  for each  $\mathfrak{s} \in \mathfrak{S}_n$ . Extending this action linearly to the vector space  $V$  makes  $V$  a  $\mathfrak{S}_n$ -module. This means that the multiplication  $\mathfrak{s}f$  for  $\mathfrak{s} \in \mathfrak{S}_n$  and  $f \in V$  satisfies the following properties:

$$(i) \mathfrak{s}f \in V, \quad (ii) (\mathfrak{st})f = \mathfrak{s}(\mathfrak{t}f), \quad (iii) \mathfrak{e}f = f, \quad (iv) \mathfrak{s}(\alpha f + \beta g) = \alpha \mathfrak{s}f + \beta \mathfrak{s}g$$

for all  $\mathfrak{s}, \mathfrak{t} \in \mathfrak{S}_n, f, g \in V, \alpha, \beta \in \mathbb{R}$ , and where  $\mathfrak{e}$  is the identity permutation in  $\mathfrak{S}_n$ . The  $\mathfrak{S}_n$ -module  $V$  is called the *permutation representation* of  $\mathfrak{S}_n$  associated to the monomials of degree at most  $d$  for reasons we will see below.

The  $\mathfrak{S}_n$ -module  $V$  gives rise to a homomorphism  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$ , where  $\text{GL}(V)$  is the set of invertible linear transformations from  $V$  to itself, by defining  $\vartheta(\mathfrak{s})$  to be the linear transformation of  $V$  corresponding to multiplication by  $\mathfrak{s}$ . The matrices realizing  $\vartheta(\mathfrak{s})$  for all  $\mathfrak{s} \in \mathfrak{S}_n$ , with respect to a fixed basis of  $V$ , form a set of *representing matrices* of  $\vartheta$ . For example, the representing matrices of  $\vartheta$ , with respect to the monomial basis of  $V$ , are the permutation matrices of size  $D \times D$ . This follows since for each  $\mathfrak{s} \in \mathfrak{S}_n$ ,  $\vartheta(\mathfrak{s})$  sends a monomial to a monomial. Let  $P_{\mathfrak{s}} \in \mathbb{R}^{D \times D}$  denote the permutation matrix representing  $\vartheta(\mathfrak{s})$ . Note that the matrices  $P_{\mathfrak{s}}$  are orthonormal.

Conversely, a homomorphism  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$  makes  $V$  a  $\mathfrak{S}_n$ -module via the multiplication  $\mathfrak{s}f := \vartheta(\mathfrak{s})f$  for each  $\mathfrak{s} \in \mathfrak{S}_n$  and  $f \in V$ . By this discussion, a representation of  $\mathfrak{S}_n$  refers to the  $\mathfrak{S}_n$ -module  $V$ , or the homomorphism  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$ , or even a set of representing matrices of  $\vartheta$  with respect to a fixed basis of  $V$ . The *trivial representation* of  $\mathfrak{S}_n$  is the homomorphism  $\vartheta(\mathfrak{s}) = 1$  for all  $\mathfrak{s} \in \mathfrak{S}_n$ . Equivalently, a  $\mathfrak{S}_n$ -module  $V$  is a trivial representation of  $\mathfrak{S}_n$  if  $V$  is one-dimensional and  $\mathfrak{s}f = f$  for all  $\mathfrak{s} \in \mathfrak{S}$  and  $f \in V$ .

A subspace  $W$  of  $V$  is a  $\mathfrak{S}_n$ -submodule if it is invariant under the action of  $\mathfrak{S}_n$ , i.e.,  $\mathfrak{s}f \in W$  for all  $f \in W$  and  $\mathfrak{s} \in \mathfrak{S}_n$ . A  $\mathfrak{S}_n$ -module is *irreducible* if it does not contain any nontrivial submodules, and the associated homomorphism  $\vartheta$  is also said to be irreducible. The irreducible  $\mathfrak{S}_n$ -modules are indexed by the partitions  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$  of  $n$  denoted as  $\lambda \vdash n$ . There is a canonical irreducible  $\mathfrak{S}_n$ -module indexed by the partition  $\lambda$  called the *Specht module*  $S^\lambda$  whose dimension is  $n_\lambda$ , the number of *standard tableaux* of shape  $\lambda$ . All irreducible representations of  $\mathfrak{S}_n$  are isomorphic to one of these Specht modules. See [23, Chapter 2] for a detailed account of the combinatorics underlying the representation theory of  $\mathfrak{S}_n$ . In Section 4, we give more details of the specific items we will need.

One of the fundamental results in the representation theory of finite groups specialized to our setting says the following.

THEOREM 3.1 (Maschke’s theorem). [23, Theorem 1.5.3] *The  $\mathfrak{S}_n$ -module  $V = \mathbb{R}[\mathbf{x}]_{\leq d}$  breaks into a direct sum of irreducible submodules.*

Let  $V_\lambda := \bigoplus_{i=1}^{m_\lambda} V_\lambda^i$  denote the sub-sum of all the isomorphic copies of the irreducible  $S^\lambda$  in a full irreducible decomposition of  $V$  from Maschke’s theorem. While this decomposition is not unique, the *multiplicity*  $m_\lambda$  of  $S^\lambda$  in the decomposition is. The subspace  $V_\lambda$  is called an isotypic of  $V$ , and the decomposition

$$(9) \quad V = \bigoplus_{\lambda \vdash n} V_\lambda$$

called the *isotypic decomposition* of  $V$ , is unique. A useful fact to note is that a  $\mathfrak{S}_n$ -invariant polynomial  $f$  (i.e.,  $\mathfrak{s}f = f$  for each  $\mathfrak{s} \in \mathfrak{S}_n$ ) must lie in the isotypic corresponding to the trivial representation of  $\mathfrak{S}_n$ .

The irreducible decomposition of  $V$  guaranteed by Maschke’s theorem creates several block-diagonal structures that are the key to the methods in [10]. Let

$$(10) \quad V = \bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} V_\lambda^i$$

be the full decomposition of  $V$  into irreducibles where  $V_\lambda^i \cong S^\lambda$  for  $i = 1, \dots, m_\lambda$ .

The first instance of block structure arises at the level of representing matrices of the homomorphism  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$ . Suppose  $\mathcal{B}$  is a basis of  $V$  obtained by concatenating bases of the different irreducible submodules in (10). Then the representing matrices of  $\vartheta(\mathfrak{s})$  with respect to  $\mathcal{B}$  are block-diagonal with a block corresponding to each irreducible  $V_\lambda^i$  in (10) of size  $n_\lambda \times n_\lambda$ . The basis  $\mathcal{B}$  is said to be *symmetry-adapted* if the following stronger property holds: for a fixed  $\lambda$ , the  $n_\lambda \times n_\lambda$  sized blocks corresponding to the  $m_\lambda$  irreducibles  $V_\lambda^i$  are *exactly the same*, i.e., the representing matrix of  $\vartheta(\mathfrak{s})$  with respect to  $\mathcal{B}$  has the form:

$$(11) \quad B_\mathfrak{s} = \left[ \begin{array}{ccc|ccc|ccc} B_{\lambda_1} & \cdots & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & B_{\lambda_1} & \cdots & 0 & 0 & 0 & \cdots & \cdots & \cdots \\ \hline 0 & 0 & \ddots & 0 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & B_{\lambda_2} & \cdots & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & B_{\lambda_2} & \cdots & \cdots & \cdots & \cdots \\ \hline 0 & 0 & 0 & 0 & 0 & \ddots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{array} \right].$$

It is usual to notate this as

$$(12) \quad B_\mathfrak{s} = \bigoplus_{\lambda \vdash n} \bigoplus_{i=1}^{m_\lambda} B_\lambda, \quad \text{where } B_\lambda \in \mathbb{R}^{n_\lambda \times n_\lambda}.$$

A symmetry-adapted basis of  $V$  always exists and an algorithm to find it is given in [7, Chapter 5.2]. Recall that the permutation matrices  $P_\mathfrak{s}$  were representing matrices for  $\vartheta$ , but they are not block-diagonal. If  $M$  is the change of basis matrix from the monomial basis of  $V$  to  $\mathcal{B}$ , then the new representing matrices are  $B_\mathfrak{s} := MP_\mathfrak{s}M^{-1}$ . If  $M$  is orthogonal, then  $B_\mathfrak{s} = MP_\mathfrak{s}M^\top$  is also orthogonal.

Next we consider the set of matrices that commute with every  $B_\mathfrak{s}$ . This is the *commutant algebra*

$$(13) \quad \text{Com } B := \{Q \in \mathbb{R}^{D \times D} : QB_\mathfrak{s} = B_\mathfrak{s}Q \ \forall \ \mathfrak{s} \in \mathfrak{S}_n\}.$$

Matrices in  $\text{Com } B$  have a very special structure as a consequence of the block-diagonal nature of  $B_\mathfrak{s}$  and Schur’s Lemma [23, Theorem 1.6.5].

THEOREM 3.2. [23, Thm 1.7.8 (2)] *If  $Q$  lies in the commutant algebra  $\text{Com } B$ , then  $Q$  is block-diagonal with a block  $Q_\lambda$  for each partition  $\lambda \vdash n$ . Further,  $Q_\lambda$  is a block matrix with  $m_\lambda$  row and column blocks each of size  $n_\lambda \times n_\lambda$ . The matrices in each block of  $Q_\lambda$  are multiples of the identity matrix  $I_{n_\lambda}$ .*

For example, if  $m_\lambda = 3$  and  $n_\lambda = 2$ , then  $Q_\lambda$  is the  $3 \times 3$  block matrix shown below on the left. Now notice that by permuting rows and columns, we can transform  $Q_\lambda$  to a block-diagonal matrix with  $n_\lambda$  equal blocks each of size  $m_\lambda \times m_\lambda$ . This is the block-diagonalization of  $Q$  needed in [10].

$$\left[ \begin{array}{c|c|c} c_1 & 0 & c_2 & 0 & c_3 & 0 \\ \hline 0 & c_1 & 0 & c_2 & 0 & c_3 \\ \hline c_4 & 0 & c_5 & 0 & c_6 & 0 \\ \hline 0 & c_4 & 0 & c_5 & 0 & c_6 \\ \hline c_7 & 0 & c_8 & 0 & c_9 & 0 \\ \hline 0 & c_7 & 0 & c_8 & 0 & c_9 \end{array} \right] \longrightarrow \left[ \begin{array}{c|c|c} c_1 & c_2 & c_3 & 0 & 0 & 0 \\ \hline c_4 & c_5 & c_6 & 0 & 0 & 0 \\ \hline c_7 & c_8 & c_9 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & c_1 & c_2 & c_3 \\ \hline 0 & 0 & 0 & c_4 & c_5 & c_6 \\ \hline 0 & 0 & 0 & c_7 & c_8 & c_9 \end{array} \right].$$

The above discussion can be phrased more generally as follows.

**THEOREM 3.3 (The Fundamental Theorem).** [7, pp 40] *Consider the decomposition of  $V = \mathbb{R}[\mathbf{x}]_{\leq d}$  as in (10) under the representation  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$  with representing matrices  $B_\mathfrak{s}$  for each  $\mathfrak{s} \in \mathfrak{S}_n$  computed with respect to a symmetry-adapted basis of  $V$ . Suppose  $Q \in \mathbb{R}^{D \times D}$  is such that  $QB_\mathfrak{s} = B_\mathfrak{s}Q$  for all  $\mathfrak{s} \in \mathfrak{S}_n$ . Then there is a reordering of the symmetry-adapted basis with respect to which  $Q$  is block-diagonal of the form:*

$$(14) \quad Q = \oplus_{\lambda \vdash n} \oplus_{i=1}^{n_\lambda} Q_\lambda.$$

Note that the structure of  $Q$  in (14) is doubly block-diagonal similar to (11). The matrix  $Q_\lambda$  has size  $m_\lambda \times m_\lambda$ . Also, since the algorithm creates a bijection between the matrices  $Q_\lambda$  in (14) and the  $n_\lambda$  standard tableaux  $\tau_\lambda^1, \dots, \tau_\lambda^{n_\lambda}$ , we may rewrite (14) as

$$(15) \quad Q = \oplus_{\lambda \vdash n} \oplus_{\tau_\lambda} Q_\lambda.$$

For the particular ordering of the symmetry-adapted basis required to block-diagonalize  $Q$ , we refer the reader to [7, pp 40]. As a consequence of the reordering we get a different direct sum decomposition of each isotypic  $V_\lambda$  in (10), indexed by the standard tableaux of shape  $\lambda$ :

$$(16) \quad V_\lambda = \oplus_{i=1}^{n_\lambda} W_{\tau_\lambda^i} \quad \text{leading to } V = \oplus_{\lambda \vdash n} \oplus_{i=1}^{n_\lambda} W_{\tau_\lambda^i}.$$

Note that this vector space decomposition is not a  $\mathfrak{S}_n$ -module decomposition. For more information on the vector spaces  $W_{\tau_\lambda}$ , see Section 2 and Appendix A of [17].

**3.2. SUM OF SQUARES FOR AN INVARIANT POLYNOMIAL.** The main message of [10] is that the computation of sos certificates of degree at most  $d$  for an invariant polynomial can be helped greatly by the irreducible decomposition (10) and the block structures discussed above. We present their strategy in our setting.

A representation  $\vartheta$  of  $\mathfrak{S}_n$  on  $\mathbb{R}[\mathbf{x}]_{\leq d}$  induces a representation of  $\mathfrak{S}_n$  on  $\mathcal{S}^D$ , the vector space of  $D \times D$  real symmetric matrices as follows:

$$(17) \quad \mathfrak{s}X := \vartheta(\mathfrak{s})^\top X \vartheta(\mathfrak{s}) \quad \forall \mathfrak{s} \in \mathfrak{S}_n, X \in \mathcal{S}^D.$$

Here we identify  $\vartheta(\mathfrak{s})$  with a matrix representation of it, and assume that these matrices are all orthonormal. This action preserves the cone of psd matrices. The set of all invariant matrices under the action (17) are those of the form  $X = \vartheta(\mathfrak{s})^\top X \vartheta(\mathfrak{s})$  for all  $\mathfrak{s} \in \mathfrak{S}_n$ . Note that these are precisely the matrices that commute with every  $\vartheta(\mathfrak{s})$  and hence they can be block-diagonalized as in (15). An easy way to construct an invariant matrix from any matrix  $X \in \mathcal{S}^D$  is to pass to its *symmetrization*  $\bar{X} := \sum_{\mathfrak{s} \in \mathfrak{S}_n} \vartheta(\mathfrak{s})^\top X \vartheta(\mathfrak{s})$ . If  $X$  is psd then so is  $\bar{X}$ .

Suppose we are given a  $\mathfrak{S}_n$ -invariant polynomial  $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$  for which we wish to find a  $d$ -sos representation modulo an ideal  $\mathcal{I} \subset \mathbb{R}[\mathbf{x}]$  that is also  $\mathfrak{S}_n$ -invariant. Let

$[\mathbf{x}]_{\leq d}$  denote the vector of monomials in  $\mathbb{R}[\mathbf{x}]_{\leq d}$ . The polynomial  $f(\mathbf{x})$  is  $d$ -sos modulo  $\mathcal{I}$  if and only if there exists a psd matrix  $Q$  such that

$$f(\mathbf{x}) \equiv [\mathbf{x}]_{\leq d}^\top Q [\mathbf{x}]_{\leq d} \pmod{\mathcal{I}} \quad :\Leftrightarrow \quad f(\mathbf{x}) - \mathbf{h} = [\mathbf{x}]_{\leq d}^\top Q [\mathbf{x}]_{\leq d} \text{ for some } \mathbf{h} \in \mathcal{I}.$$

Since  $f$  and  $\mathcal{I}$  are both  $\mathfrak{S}_n$ -invariant, we may symmetrize both sides of the above equation to get another sos expression for  $f \pmod{\mathcal{I}}$ . Therefore, we can assume that the sos on the right is  $\mathfrak{S}_n$ -invariant. In other words, we may assume that the psd matrix  $Q$  in  $[\mathbf{x}]_{\leq d}^\top Q [\mathbf{x}]_{\leq d}$  is symmetrized since symmetrizing the sos yields

$$\begin{aligned} \sum_{\mathfrak{s} \in \mathfrak{S}_n} (\vartheta(\mathfrak{s})[\mathbf{x}]_{\leq d})^\top Q (\vartheta(\mathfrak{s})[\mathbf{x}]_{\leq d}) &= \sum_{\mathfrak{s} \in \mathfrak{S}_n} ([\mathbf{x}]_{\leq d})^\top (\vartheta(\mathfrak{s})^\top Q \vartheta(\mathfrak{s})) [\mathbf{x}]_{\leq d} \\ &= ([\mathbf{x}]_{\leq d})^\top \left( \sum_{\mathfrak{s} \in \mathfrak{S}_n} \vartheta(\mathfrak{s})^\top Q \vartheta(\mathfrak{s}) \right) [\mathbf{x}]_{\leq d}. \end{aligned}$$

Now suppose  $M$  denotes the change of basis matrix from the monomial basis of  $\mathbb{R}[\mathbf{x}]_{\leq d}$  to a symmetry-adapted basis with respect to (10). Then  $f(\mathbf{x})$  is sos mod  $\mathcal{I}$  if and only if there is an invariant psd matrix  $Q$  such that

$$\begin{aligned} f(\mathbf{x}) &\equiv [\mathbf{x}]_{\leq d}^\top Q [\mathbf{x}]_{\leq d} \pmod{\mathcal{I}} \\ &\equiv [\mathbf{x}]_{\leq d}^\top M M^\top Q M M^\top [\mathbf{x}]_{\leq d} \pmod{\mathcal{I}} \\ &\equiv (M^\top [\mathbf{x}]_{\leq d})^\top (M^\top Q M) (M^\top [\mathbf{x}]_{\leq d}) \pmod{\mathcal{I}} \\ &\equiv \mathbf{y}^\top \tilde{Q} \mathbf{y} \pmod{\mathcal{I}} \end{aligned}$$

where  $\mathbf{y} := M^\top [\mathbf{x}]_{\leq d}$  is the vector of elements in the symmetry-adapted basis of  $V$ . The matrix  $\tilde{Q} = M^\top Q M$  has the block structure in (15) since  $Q$  is invariant and  $M^\top Q M$  is precisely the transformation that block diagonalizes  $Q$ . The components of  $\mathbf{y}$  will be referred to as (*symmetry-adapted*) *basis polynomials*. The block structure of  $\tilde{Q}$  endows a block structure on  $\mathbf{y}$ , with big blocks indexed by the partitions  $\lambda$  of  $n$ , and each block  $\mathbf{y}_\lambda$  broken further into blocks indexed by the standard tableaux of shape  $\lambda$ .

For simplicity, we rename  $\tilde{Q}$  by  $Q$  and rewrite the sos expression for  $f$  using (15). Thus

$$(18) \quad f(\mathbf{x}) \equiv \mathbf{y}^\top Q \mathbf{y} \equiv \sum_{\lambda} \sum_{\tau_\lambda} \mathbf{y}_{\tau_\lambda}^\top Q_{\lambda} \mathbf{y}_{\tau_\lambda}$$

$$(19) \quad \equiv \sum_{\lambda} \sum_{\tau_\lambda} \langle Q_{\lambda}, \mathbf{y}_{\tau_\lambda} \mathbf{y}_{\tau_\lambda}^\top \rangle \equiv \sum_{\lambda} n_{\lambda} \langle Q_{\lambda}, Y_{\lambda} \rangle \pmod{\mathcal{I}}$$

where  $Q_{\lambda}$  is an unknown psd matrix of size  $m_{\lambda} \times m_{\lambda}$ , and

$$Y_{\lambda} := \frac{1}{|\mathfrak{S}_n|} \sum_{\mathfrak{s} \in \mathfrak{S}_n} \mathfrak{s}(\mathbf{y}_{\tau'_\lambda} \mathbf{y}_{\tau'_\lambda}^\top)$$

is a matrix of the same size where  $\tau'_\lambda$  is any tableau of shape  $\lambda$  (symmetrization here is possible since  $f$  is  $\mathfrak{S}_n$ -invariant). The choice of tableaux  $\tau'_\lambda$  does not affect  $Y_{\lambda}$ . For further explanations, see Appendix A of [17]. Therefore, in order to check if  $f$  is a sos of the above type, we need to search for a psd matrix  $Q_{\lambda}$  of size  $m_{\lambda}$  for each  $\lambda$  such that the linear equations that come from equating  $f$  to the sos expression (19) hold. For details on how to set up this SDP, we refer the reader to [2, Chapter 7.2.1].

It might be possible to certify the nonnegativity of  $f(\mathbf{x})$  by using only a subset  $\Lambda$  of the partitions of  $n$ . Since we will rely on such subsets in the next two sections, we make a formal definition to say precisely what we mean.

DEFINITION 3.4. We say that a sos expression for  $f$  modulo the ideal  $\mathcal{I}$  can be obtained through the Gatermann-Parrilo SDP restricted to partitions in  $\Lambda$  if there exists psd matrices  $Q_\lambda$  such that

$$(20) \quad f(\mathbf{x}) \equiv \sum_{\lambda \in \Lambda} \langle Q_\lambda, Y_\lambda \rangle \pmod{\mathcal{I}}.$$

#### 4. MAIN RESULTS

In this section, we establish the connection between the sos certificate (6) obtained from the flag algebra method to those that can be obtained from the Gatermann-Parrilo symmetry-adapted SDP described in the previous section. In particular, we show that this sos certificate can be obtained from a SDP restricted to a fixed number of known partitions as in Definition 3.4 above.

Note that the sos certificate (6) is a nested sos (sums of sums of squares really) which makes it cumbersome to work with. Therefore, we work with the inner sums of squares in (6) whenever possible. In Theorem 4.3 and Corollary 4.10, we use the innermost sos by fixing some  $\theta, \sigma$  and  $l$ , namely

$$(21) \quad \mathbf{P}^{\theta, \sigma, l \top} Q_{\sigma, l} \mathbf{P}^{\theta, \sigma, l} =: \sum_j r_j^2,$$

to prove that each  $r_j$  is invariant under a particular subgroup of  $\mathfrak{S}_n$ , and as such, lies in the direct sum of finitely many  $V_\mu$  which we explicitly describe.

In Theorem 4.13, we need the sos (6) obtained from fixing only  $\sigma$  and  $l$  and which, with a slight abuse of notation, we denote by

$$(22) \quad \sum_\theta \sum_j r_{\theta, j}^2 := \mathbb{E}_\theta \mathbf{P}^{\theta, \sigma, l \top} Q_{\sigma, l} \mathbf{P}^{\theta, \sigma, l}.$$

We prove that such a sos can be retrieved through the Gatermann-Parrilo method restricted to partitions said to be lexicographically greater or equal to  $(n - k, 1^k)$  where  $k$  is the size of  $\sigma$  (see definition below).

Corollary 4.14 involves the conic combination of (6) for problems where several types  $\sigma$  and flags of different size are necessary. It shows that this whole flag sos can also be retrieved through the Gatermann-Parrilo method restricted to partitions lexicographically greater or equal to  $(n - k^*, 1^{k^*})$  where  $k^*$  is the maximum size among all types present. A key feature is that the number of such partitions is independent of  $n$ .

Throughout this section, we assume that  $V = \mathbb{R}[\mathbf{x}]_{\leq d}$  is such that  $d$  is at least as big as the maximum number of edges in the flags considered in the conic combination of sos expressions of the form (6).

4.1. INVARIANCE OF DENSITY POLYNOMIALS. A *partition* of  $n$  is a way of writing  $n$  as a sum of positive integers; each summand is called a *part*. We denote a partition  $\lambda$  by a vector containing the parts  $\lambda_i$  in non-decreasing order, i.e.,  $(\lambda_1, \dots, \lambda_t)$  such that  $\lambda_1 \geq \dots \geq \lambda_t > 0$  and  $\lambda_1 + \dots + \lambda_t = n$ . A key partition for us is the *hook partition* for which all the parts but one are 1; we denote the hook partition with  $k + 1$  parts as  $(n - k, 1^k)$ . There is a lexicographic order on the partitions of  $n$ ; for  $\lambda = (\lambda_1, \dots, \lambda_{t_1})$  and  $\mu = (\mu_1, \dots, \mu_{t_2})$ , we write  $\lambda \geq_{\text{lex}} \mu$  if the vector  $(\lambda_1, \dots, \lambda_{t_1})$  is lexicographically greater than or equal to the vector  $(\mu_1, \dots, \mu_{t_2})$ . A partition  $\lambda$  has a shape (Young diagram) with rows of size  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_t$ . A *tableau* of shape  $\lambda$ , denoted as  $\tau_\lambda$ , is a filling of the boxes in the diagram of  $\lambda$  by the numbers  $1, \dots, n$ .

The tableau is *standard* if the numbering increases along each row and column. The *row group* of a tableau  $\tau_\lambda$  is the subgroup of  $\mathfrak{S}_n$  defined as

$$\mathfrak{R}_{\tau_\lambda} := \{\mathfrak{s} \in \mathfrak{S}_n : \mathfrak{s} \text{ fixes the set of numbers in each row of } \tau_\lambda\}.$$

Note that this group is isomorphic to  $\mathfrak{S}_\lambda := \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_t}$ .

We will show that the polynomials  $\mathfrak{p}_F^\theta$  from (4) are invariant under a particular row group for all flags  $F \in \mathcal{F}_l^\sigma$ . Recall that these polynomials were defined from the choice of a  $\sigma$ -flag  $F$  of size  $l$ , and an injective map  $\theta \in \text{Inj}([k], [n])$ . In particular, no partitions, tableaux or row groups were involved. We first present an example.

EXAMPLE 4.1. Consider  $n = 5$  and the hook partition  $\lambda = (2, 1, 1, 1)$ . The row group of the tableau

$$\tau_\lambda = \begin{array}{|c|c|} \hline 4 & 5 \\ \hline 1 & \\ \hline 2 & \\ \hline 3 & \\ \hline \end{array} \text{ is } \mathfrak{R}_{\tau_\lambda} = \{1, (4, 5)\}.$$

Suppose we choose the type  $\sigma = 2 \bullet \nearrow_3^1$ , set  $l = 4$ , and consider the  $\sigma$ -flag  $F = 2 \bullet \nearrow_3^1$ . Since  $n = 5$ , assume  $\theta : [3] \rightarrow [5]$  is such that  $\theta(1) = 1, \theta(2) = 2$  and  $\theta(3) = 3$ . The set  $\text{Inj}_\theta(V(F), [5])$  contains two maps, both preserving  $\theta$  and hence, sending 1, 2, 3 to themselves. Suppose  $\alpha_1 \in \text{Inj}_\theta(V(F), [5])$  sends the unlabeled vertex in  $F$  to 4 and  $\alpha_2 \in \text{Inj}_\theta(V(F), [5])$  sends it to 5.

Using the formula in (1), one can see that

$$\begin{aligned} \mathfrak{q}_F^{\alpha_1} &= x_{12}x_{13}x_{14}(1 - x_{23})(1 - x_{24})(1 - x_{34}), \text{ and} \\ \mathfrak{q}_F^{\alpha_2} &= x_{12}x_{13}x_{15}(1 - x_{23})(1 - x_{25})(1 - x_{35}). \end{aligned}$$

Similarly, from (4) we get that

$$\mathfrak{p}_F^\theta = \frac{1}{|\text{Inj}_\theta(V(F), [n])|} \sum_{\alpha \in \text{Inj}_\theta(V(F), [n])} \mathfrak{q}_F^\alpha = \frac{1}{2}(\mathfrak{q}_F^{\alpha_1} + \mathfrak{q}_F^{\alpha_2}).$$

Note that  $\mathfrak{p}_F^\theta$  is invariant under  $\mathfrak{R}_{\tau_\lambda}$ . Indeed, the action 1  $\in \mathfrak{R}_{\tau_\lambda}$  applied to  $\mathfrak{p}_F^\theta$  doesn't change anything, and the action  $(4, 5) \in \mathfrak{R}_{\tau_\lambda}$  sends  $\mathfrak{q}_F^{\alpha_1}$  to  $\mathfrak{q}_F^{\alpha_2}$  and vice-versa, thus leaving  $\mathfrak{p}_F^\theta$  unchanged.

We now prove that this observation is not accidental; a polynomial  $\mathfrak{p}_F^\theta$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant for some tableau  $\tau_\lambda$  of shape  $\lambda$  where  $\lambda$  is a specific hook partition.

PROPOSITION 4.2. *Let  $F$  be a  $\sigma$ -flag where  $|\sigma| = k$ , and  $\theta : [k] \rightarrow [n]$  be an injective map. Consider the hook partition  $\lambda = (n - k, 1^k)$  and a tableau  $\tau_\lambda$  where we fill the first row by numbers from  $[n] \setminus \{\theta(i) : i \in [k]\}$  in any order, and the  $k$  remaining rows of size one with numbers from  $\{\theta(i) : i \in [k]\}$  in any order. Then the polynomial  $\mathfrak{p}_F^\theta$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant.*

*Proof.* Consider any  $\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda} \cong \mathfrak{S}_{n-k}$  and  $\alpha : V(F) \rightarrow [n]$  in  $\text{Inj}_\theta(V(F), [n])$ . Then  $(\mathfrak{s} \circ \alpha) : V(F) \rightarrow [n]$  is also an injective map preserving  $\theta$ . Indeed, for any  $\alpha \in \text{Inj}_\theta(V(F), [n])$ , the composition gives a map from  $\mathfrak{R}_{\tau_\lambda} \rightarrow \text{Inj}_\theta(V(F), [n])$  which is  $\frac{|\mathfrak{R}_{\tau_\lambda}|}{|\text{Inj}_\theta(V(F), [n])|}$ -to-one surjective. This is because there are  $n - l$  elements of  $[n]$  outside the range of  $\alpha$  and  $\alpha = \mathfrak{s} \circ \alpha$  for any  $\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}$  that fixes the range of  $\alpha$ . Since there are

$$(n - l)! = \frac{(n - k)!}{(n - k)(n - k - 1) \cdots (n - l + 1)} = \frac{|\mathfrak{R}_{\tau_\lambda}|}{|\text{Inj}_\theta(V(F), [n])|}$$

such permutations  $\mathfrak{s}$ , we get that the  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomial

$$\frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \sum_{\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}} \mathfrak{s} \cdot \mathfrak{q}_F^\alpha = \frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \cdot \frac{|\mathfrak{R}_{\tau_\lambda}|}{|\text{Inj}_\theta(V(F), [n])|} \sum_{\alpha \in \text{Inj}_\theta(V(F), [n])} \mathfrak{q}_F^\alpha = \mathfrak{p}_F^\theta$$

by (4), which implies that  $\mathfrak{p}_F^\theta$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant. □

**THEOREM 4.3.** *Each polynomial  $r_j$  in the sos (21) is invariant under the row group  $\mathfrak{R}_{\tau_\lambda}$  corresponding to the tableau  $\tau_\lambda$  and hook partition  $\lambda$  as in Proposition 4.2.*

*Proof.* Since each  $r_j$  is a linear combination of  $\mathfrak{p}_F^\theta$  with  $F \in \mathcal{F}_l^\sigma$ , and since  $\tau_\lambda$  only depends on  $\theta$ , all  $\mathfrak{p}_F^\theta$  for any  $F \in \mathcal{F}_l^\sigma$  are invariant under the same  $\mathfrak{R}_{\tau_\lambda}$ . Therefore,  $r_j$  is also  $\mathfrak{R}_{\tau_\lambda}$ -invariant. □

**REMARK 4.4.** Note that the hook  $\lambda$  in Proposition 4.2 and Theorem 4.3 only depends on  $k$ , the size of the type  $\sigma$  of the flags  $F$  and not on their size  $l$ .

**4.2.  $\mathfrak{R}_{\tau_\lambda}$ -INVARIANT POLYNOMIALS AND THE ISOTYPIC DECOMPOSITION.** Our next goal is to show that a  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomial  $f$  lies in the span of certain specific isotypics in the isotypic decomposition (9) of  $V = \mathbb{R}[\mathbf{x}]_{\leq d}$ . We will use this in the next subsection to prove that Razborov’s sos can be obtained by restricting the Gatermann-Parrilo SDP to the subblocks indexed by these isotypics/partitions.

Recall that the induced  $\mathfrak{S}_n$ -action we have decomposes  $V$  into a direct sum as in (10). Therefore, our polynomial  $f \in V$  decomposes as

$$(23) \quad f = \sum_{\mu \vdash n} \sum_{i=1}^{m_\mu} f_{\mu,i}$$

where  $f_{\mu,i} \in V_\mu^i$ . Since  $f$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant, we may assume without loss of generality that each  $f_{\mu,i}$  in (23) is also  $\mathfrak{R}_{\tau_\lambda}$ -invariant. Indeed,

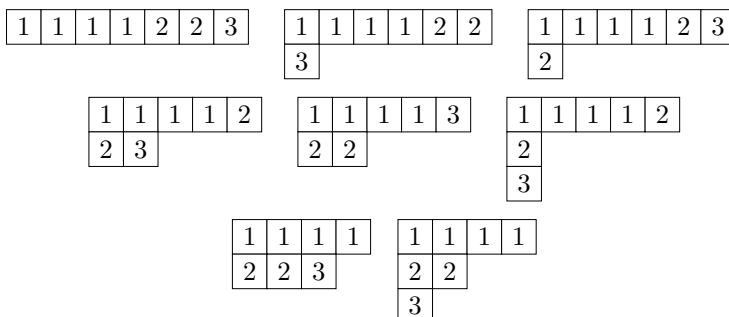
$$f = \frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \sum_{\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}} \mathfrak{s}f = \sum_{\mu \vdash n} \sum_{i=1}^{m_\mu} \left( \frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \sum_{\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}} \mathfrak{s}f_{\mu,i} \right),$$

and since  $V_\mu^i$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant (because it is  $\mathfrak{S}_n$ -invariant),  $\frac{1}{|\mathfrak{R}_{\tau_\lambda}|} \sum_{\mathfrak{s} \in \mathfrak{R}_{\tau_\lambda}} \mathfrak{s}f_{\mu,i}$  lies in  $V_\mu^i$ . In the case when  $f$  is some  $r_j$  in the sos (21), we are interested in knowing when a  $f_{\mu,i}$  is non-zero, or equivalently, in determining which parts of  $V$  contain  $f$ . For this, we rely on the theory of *restricted representations* [9, Chapter 4].

The restricted representation of the subgroup  $\mathfrak{R}_{\tau_\lambda}$  on  $V$  is the representation of  $\mathfrak{R}_{\tau_\lambda}$  on  $V$  obtained by restricting the  $\mathfrak{S}_n$  representation  $\vartheta : \mathfrak{S}_n \rightarrow \text{GL}(V)$  to the elements of  $\mathfrak{R}_{\tau_\lambda}$ . This has the effect of refining the direct sum decomposition in (10) since an irreducible  $V_\mu^i$ , while being  $\mathfrak{R}_{\tau_\lambda}$ -invariant, may not be irreducible with respect to  $\mathfrak{R}_{\tau_\lambda}$  and hence will decompose into irreducible representations of  $\mathfrak{R}_{\tau_\lambda}$ . The  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomials in a  $V_\mu^i$  are contained in the copies of the trivial representation of  $\mathfrak{R}_{\tau_\lambda}$  in  $V_\mu^i$ . Therefore, the component  $f_{\mu,i}$  in (23) is nonzero only if  $V_\mu^i$  contains at least one copy of the trivial representation of  $\mathfrak{R}_{\tau_\lambda}$  in the restricted representation of  $\mathfrak{R}_{\tau_\lambda}$  on  $V_\mu^i$ . To characterize such partitions  $\mu$ , we need the following definition. Let  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_t)$  be a partition and  $N_\lambda$  be the sequence containing  $\lambda_p$  copies of the number  $p$  for  $p = 1, \dots, t$ . A semistandard tableau of shape  $\mu$  and type  $\lambda$  is a tableau of  $\mu$  with numbers coming from  $N_\lambda$  such that the numbers are non-decreasing along rows and increasing along columns.

**EXAMPLE 4.5.** If  $\lambda = (4, 2, 1)$ , then  $N_\lambda = (1, 1, 1, 1, 2, 2, 3)$  and the semistandard tableaux of type  $\lambda$  are





Among these, there is one of shape (7), two of shape (6, 1), two of shape (5, 2), one of shape (5, 1, 1), one of shape (4, 3), and one of shape (4, 2, 1).  $\square$

The following theorem presented in many books (including in Corollary 4.39 and the discussion following it in [9]) is exactly what we need.

**THEOREM 4.6 (Young’s Rule).** *Let  $\lambda = (\lambda_1, \dots, \lambda_t)$  and  $\mu$  be partitions of  $n$ . The multiplicity of the trivial representation of  $\mathfrak{R}_{\tau_\lambda}$  in the restriction of some irreducible representation  $V_\mu^i$  to  $\mathfrak{R}_{\tau_\lambda}$  is equal to the number of semistandard tableaux of shape  $\mu$  and type  $\lambda$ .*

Note that in [9], this result is given in terms of *induced representations*, but by *Frobenius reciprocity* [9, Corollary 3.20], the above is an equivalent version.

The number of semistandard tableaux of shape  $\mu$  and type  $\lambda$  is called the *Kostka number*  $K_{\mu\lambda}$ . We need to know when  $K_{\mu\lambda}$  is non-zero, and the following standard fact gives a necessary characterization.

**LEMMA 4.7.** *If  $\lambda$  is lexicographically greater than  $\mu$  or if  $\mu$  has more parts than  $\lambda$ , then  $K_{\mu\lambda} = 0$ .*

*Proof.* Suppose  $\lambda = (\lambda_1, \dots, \lambda_{t_1})$  and  $\mu = (\mu_1, \dots, \mu_{t_2})$ . Note that since the numbers are increasing in the columns of any semistandard tableau of shape  $\mu$  and type  $\lambda$ , the minimum number in row  $i$  is  $i$ .

Suppose  $\lambda_i = \mu_i$  for all  $1 \leq i < i^*$  and  $\lambda_{i^*} > \mu_{i^*}$ , i.e.,  $\lambda$  is lexicographically greater than  $\mu$ . Then any semistandard tableau of shape  $\mu$  and type  $\lambda$  will have to have  $\lambda_1$  1’s in row 1,  $\dots$ , and  $\lambda_{i^*-1}$  ( $i^* - 1$ )’s in row  $i^* - 1$  in order to have increasing numbers along the columns. Then one would attempt to put  $\lambda_{i^*}$   $i^*$ ’s in row  $i^*$ , but that is not possible since  $\mu_{i^*} < \lambda_{i^*}$ , and so at least one  $i^*$  will have to be in row  $i^* + 1$ , meaning that we don’t have a semistandard tableau since the columns are not strictly increasing.

Now suppose that  $\mu$  has more parts than  $\lambda$ , i.e.,  $t_2 > t_1$ . Look at the first column of a semistandard tableau: there needs to be at least  $t_2$  distinct numbers in it, but  $N_\lambda$  contains only  $t_1$  distinct numbers and so there is no semistandard tableau.  $\square$

**DEFINITION 4.8.** *We define the ordering  $\succeq$  on partitions as follows:  $\mu \succeq \lambda$  if  $\mu$  is lexicographically greater than or equal to  $\lambda$  and has at most as many parts as  $\lambda$ .*

Using Theorem 4.6 and Lemma 4.7, we can determine which irreducibles in (10) contribute components to a given  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomial.

**THEOREM 4.9.** *A  $\mathfrak{R}_{\tau_\lambda}$ -invariant polynomial  $f$  in  $V = \mathbb{R}[\mathbf{x}]_{\leq \mathbf{d}}$  lies in  $\bigoplus_{\mu \succeq \lambda} V_\mu$ .*

*Proof.* Since  $f$  is  $\mathfrak{R}_{\tau_\lambda}$ -invariant, by Theorem 4.6, the multiplicity of the trivial representation of  $\mathfrak{R}_{\tau_\lambda}$  in  $V_\mu^i$  restricted to  $\mathfrak{R}_{\tau_\lambda}$  is the number of semistandard tableaux of shape  $\mu$  and type  $\lambda$ . By Lemma 4.7, this multiplicity is zero if  $\mu$  is lexicographically smaller than  $\lambda$  or if it has more parts than  $\lambda$ . Therefore,  $f \in \bigoplus_{\mu \succeq \lambda} V_\mu$ .  $\square$

Since  $r_j$  in  $\sum_j r_j^2$  from (21) is  $\mathfrak{R}_{\tau_\lambda}$ -invariant for a tableau  $\tau_\lambda$  of shape  $(n - k, 1^k)$  by Theorem 4.3, and since for the hook  $\lambda = (n - k, 1^k)$ ,  $\mu \triangleright \lambda$  if and only if  $\mu \geq_{\text{lex}} \lambda$ , we get the following offshot.

**COROLLARY 4.10.** *Any polynomial  $r_k \in V$  that contributes to the sos (21) lies in  $\bigoplus_{\mu \geq_{\text{lex}} \lambda} V_\mu$  where  $\lambda = (n - k, 1^k)$ .*

**4.3. FROM RAZBOROV TO GATERMANN-PARRILO.** We are now ready for the final step. We will show that a sos arising from flag algebras as in Proposition 2.5 can be retrieved from the Gatermann-Parrilo SDP by restricting to certain blocks in the sense of Definition 3.4. We first define formally the concept of symmetrization, which we already saw a few times.

**DEFINITION 4.11.** *The symmetrization of a polynomial  $f \in \mathbb{R}[\mathbf{x}]$  with respect to  $\mathfrak{S}_n$  is denoted by  $\text{sym}(f)$  and defined to be*

$$\text{sym}(f) = \frac{1}{n!} \sum_{s \in \mathfrak{S}_n} sf.$$

**PROPOSITION 4.12.** *If some  $\mathfrak{S}_n$ -invariant sos  $\sum_j f_j^2$  is such that  $f_j \in V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$  for all  $j$ , then this sos can be obtained through the Gatermann-Parrilo SDP restricted to  $\lambda_1, \dots, \lambda_s$ .*

*Proof.* Suppose  $\sum_j f_j^2$  is a sos such that  $f_j = f_{j,1} + \dots + f_{j,s}$  where  $f_{j,i} \in V_{\lambda_i}$  for  $i = 1, \dots, s$  for all  $j$ . Then

$$(24) \quad f_j^2 = (f_{j,1})^2 + \dots + (f_{j,s})^2 + 2 \sum_{i_1 < i_2} f_{j,i_1} f_{j,i_2}.$$

Since the sos  $\sum_j f_j^2$  is  $\mathfrak{S}_n$ -invariant, symmetrizing it with respect to  $\mathfrak{S}_n$  leaves it unchanged. Therefore,

$$\begin{aligned} \sum_j f_j^2 &= \text{sym} \left( \sum_j f_j^2 \right) = \sum_j \text{sym} (f_j^2) = \sum_j \text{sym} ((f_{j,1} + \dots + f_{j,s})^2) \\ &= \sum_j \left( \sum_i \text{sym} (f_{j,i}^2) + 2 \sum_{i_1 < i_2} \text{sym} (f_{j,i_1} f_{j,i_2}) \right). \end{aligned}$$

Note that  $f_{j,i_1} f_{j,i_2} \in V_{\lambda_{i_1}} \otimes V_{\lambda_{i_2}}$ . By Exercise 4.51 in [9], the trivial representation of  $\mathfrak{S}_n$  is present in  $V_{\lambda_{i_1}} \otimes V_{\lambda_{i_2}}$  once if  $i_1 = i_2$  and zero times otherwise. For  $i_1 \neq i_2$ ,  $\text{sym}(f_{j,i_1} f_{j,i_2})$  is  $\mathfrak{S}_n$ -invariant as well by symmetrization and therefore, must be the zero vector.

Thus,  $f_j^2 = f_{j,1}^2 + \dots + f_{j,s}^2$ , so each square in  $\sum_j f_j^2$  can be obtained from restricting the Gatermann-Parrilo SDP to partitions  $\lambda_1, \dots, \lambda_s$ . □

**THEOREM 4.13.** *The sos (22) can be obtained by restricting the Gatermann-Parrilo SDP to the partitions  $\mu$  where  $\mu \geq_{\text{lex}} (n - k, 1^k)$ .*

*Proof.* First note that the sos  $\sum_\theta \sum_j r_{\theta,j}^2$  is invariant under  $\mathfrak{S}_n$  since we are taking the expectation over all maps  $\theta$ . Moreover,  $r_{\theta,j}$  is invariant under  $\mathfrak{R}_{\tau_\lambda}$  for some tableau  $\tau_\lambda$  of shape  $(n - k, 1^k)$  for all  $\theta$  and  $j$ . Indeed, even though different  $\theta$ 's will require different tableaux, all will be of shape  $(n - k, 1^k)$ . Thus all  $r_{\theta,k}$  lie in  $\bigoplus_{\mu >_{\text{lex}} (n - k, 1^k)} V_\mu$  by (4.10). Therefore, by (4.12), the sos (22) can be obtained from the Gatermann-Parrilo SDP restricted to the partitions  $\mu \geq_{\text{lex}} \lambda$ . □

Finally, we tackle the general Razborov sos for proofs involving several  $\sigma$  and  $l$ .

COROLLARY 4.14. Consider a conic combination of sos (6) for different choices of  $\sigma$  and  $l$ . Let  $k^*$  be the maximum size of all types  $\sigma$  present. Then this sos can be obtained through the Gatermann-Parrilo program restricted to partitions lexicographically greater or equal to the hook partition  $(n - k^*, 1^{k^*})$ .

*Proof.* In Theorem 4.13, we proved that the sos (21) can be obtained by restricting the Gatermann-Parrilo SDP to partitions  $\mu$  for  $\mu \geq_{\text{lex}} (n - |\sigma|, 1^{|\sigma|})$ . Thus every sos in (6) for each  $\theta, \sigma, l$  can be obtained from partitions  $\mu$  for  $\mu \geq_{\text{lex}} (n - k^*, 1^{k^*})$  where  $k^*$  is the maximum size of all types considered. Since the final sos is a conic combination of these smaller sos, it can also be obtained by restricting the Gatermann-Parrilo SDP to those same partitions.  $\square$

We have established the relationship between the Gatermann-Parrilo framework and the flag algebra sos for Turán problems. We conclude with a few remarks that will be helpful in implementing the Gatermann-Parrilo framework.

REMARK 4.15. Note that flag algebra Turán sos uses fixed  $k$  and  $l$  (independent of  $n$ ) and thus the number of partitions indexing blocks in the restricted SDP is also independent of  $n$ . Indeed, the number of partitions lexicographically greater than or equal to  $(n - k, 1^k)$  is at most twice the number of partitions of  $k$ . Moreover, each small subblock in a block corresponding to a  $\lambda$  has size  $m_\lambda \times m_\lambda$ , which depends on  $d$  but not  $n$ . Finally, even though the number of subblocks per block (i.e., the number of standard tableaux) increases as  $n$  increases, by (19), we only need one subblock. Thus, the size of the restricted SDP does not depend on  $n$ .

REMARK 4.16. In Theorem 4.3, we showed that a  $r_j$  in the sos (21) is invariant under  $\mathfrak{R}_{\tau_\lambda}$  for some tableau  $\tau_\lambda$  where  $\lambda$  is a hook partition. One can obtain further savings in the size of the corresponding SDP by using symmetries of the  $\sigma$ -flags which allows us to replace the hook partition with partitions that are lexicographically larger than it. For example, in Example 4.1,  $\mathfrak{p}_F^\theta$  is also invariant under the row group of

$$\tau_\lambda = \begin{array}{|c|c|} \hline 4 & 5 \\ \hline 2 & 3 \\ \hline 1 & \\ \hline \end{array}, \text{ i.e., } \mathfrak{R}_{\tau_\lambda} = \{1, (2, 3), (4, 5), (2, 3)(4, 5)\}.$$

We also note that we have stated our results based on the isotypic decomposition of  $V = \mathbb{R}[\mathbf{x}]_{\leq d}$  and not of the quotient vector space  $\mathbb{R}[\mathbf{x}]_{\leq d} / \mathcal{I}_n^A$ . This is again for simplicity and will not change the flavor of the results. The ideal is used implicitly in some places such as in the derivation of the sos (6).

### 5. THE SYMMETRY-ADAPTED SDP FOR MANTEL'S THEOREM

Recall the proof of Mantel's theorem from Section 2 using flag algebra calculus (see also the Appendix). The sos was divided into two parts: one involving  $F_0$  and  $F_1$ , and one involving  $H_1$ . We show how to retrieve the former; the latter is similar in flavor. Since the flags  $F_0$  and  $F_1$  needed in the sos expression (36) have a type of size  $k = 1$ , Theorem 4.13 implies that we can retrieve this sos expression by restricting the Gatermann-Parrilo SDP to partitions  $(n)$  and  $(n - 1, 1)$ . Moreover,  $\mathfrak{p}_F^\theta$  has degree at most one for  $F \in \{F_0, F_1\}$  and therefore we only need to consider  $V = \mathbb{R}[\mathbf{x}]_{\leq 1}$ . For illustration, we verify this.

Observe that  $n_\lambda$ , the number of standard tableaux of shape  $\lambda$ , is 1 for  $\lambda = (n)$  and  $n - 1$  for  $\lambda = (n - 1, 1)$ . Restricting the expression in (19) to partitions  $(n)$  and  $(n - 1, 1)$  implies that there exist psd matrices  $Q_{(n)}$  and  $Q_{(n-1,1)}$  such that

$$(25) \quad \mathfrak{f} = 1 \cdot \langle Q_{(n)}, Y_{(n)} \rangle + (n - 1) \cdot \langle Q_{(n-1,1)}, Y_{(n-1,1)} \rangle$$

where

$$f = E_\theta \left[ \left( \mathfrak{p}_{F_0}^\theta(G) \ \mathfrak{p}_{F_1}^\theta(G) \right) \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \mathfrak{p}_{F_0}^\theta(G) \\ \mathfrak{p}_{F_1}^\theta(G) \end{pmatrix} \right]$$

is the sos expression obtained using flag algebras in (36). Recall that the expression (25) relies on the decomposition  $V_\lambda = \bigoplus_{i=1}^{n_\lambda} W_{\tau_\lambda^i}$  as in Equation (16). The matrices  $Y_\lambda$  come from the polynomials in the symmetry-adapted basis and can be computed using the algorithm in [7, Chapter 5.2, pp. 113–114]. Also recall that  $Y_\lambda$  depends only on one standard tableau of shape  $\lambda$ .

The algorithm yields the basis polynomials

$$p_{0,0} := 1, \quad p_{0,1} := \frac{1}{\sqrt{\binom{n}{2}}} \cdot \sum_{1 \leq i < j \leq n} x_{ij}$$

of degrees zero and one for  $V_{(n)} = W_{\tau_{(n)}}$  where  $\tau_{(n)} = \boxed{1 \ 2 \ \dots \ n}$ . Therefore,

$$Y_{(n)} = (p_{0,0} \ p_{0,1}) \begin{pmatrix} p_{0,0} \\ p_{0,1} \end{pmatrix}.$$

Similarly, the algorithm yields the basis polynomial

$$p_{1,n} := \frac{1}{\sqrt{\binom{n-1}{2} \cdot 1^2 + (n-1) \cdot \left(\frac{n-2}{2}\right)^2}} \left( \sum_{1 \leq i < j \leq n-1} x_{ij} - \frac{n-2}{2} \cdot \sum_{1 \leq i \leq n-1} x_{in} \right)$$

of degree one for  $W_{\tau_{(n-1,1)}}$  where  $\tau_{(n-1,1)} = \boxed{1 \ 2 \ 3 \ \dots \ n-1} \ \boxed{n}$ . Thus, we

have that

$$Y_{(n-1,1)} = \frac{1}{|\mathfrak{S}_n|} \sum_{\mathfrak{s} \in \mathfrak{S}_n} \mathfrak{s}(p_{1,n})(p_{1,n})^T = \frac{1}{n} \sum_{i=1}^n (p_{1,i})^2.$$

Altogether, we obtain a block SDP consisting of one block of size  $2 \times 2$  for  $Q_{(n)}$  and another block of size  $1 \times 1$  for  $Q_{(n-1,1)}$ . For instance choosing psd matrices

$$Q_{(n)} = \begin{pmatrix} \frac{(n-1)^2}{2} & -\frac{\sqrt{2(n-1)^3}}{\sqrt{n}} \\ -\frac{\sqrt{2(n-1)^3}}{\sqrt{n}} & \frac{4(n-1)}{n} \end{pmatrix}$$

and

$$Q_{(n-1,1)} = \left( \frac{2(n-1)(n-2)}{n} \right)$$

yields the expression (36) as desired.

## 6. CONCLUSION

The main result of this paper is that standard symmetry-reduction methods in polynomial sos theory can retrieve Razborov’s sos proofs using flag densities that arise in the context of Turán problems. For the sake of notational simplicity, we presented our results in the context of graphs. However, the same techniques can be used for Turán and non-Turán problems over hypergraphs, digraphs, tournaments, etc. Below, we give a few examples of the changes that need to be made in different cases.

- (1) *Hypergraphs*: For any optimization problem over  $a$ -uniform hypergraphs on  $n$  vertices, we can use the ideal

$$\langle x_{i_1 \dots i_a}^2 - x_{i_1 \dots i_a} \ \forall 1 \leq i_1 < \dots < i_a \leq n \rangle.$$

This ideal is  $\mathfrak{S}_n$ -invariant, which implies that our theory can still be applied. Thus any flag Cauchy-Schwarz proof in this setting can be retrieved by using symmetry-reduction and restricting to partitions lexicographically greater or equal to  $(n - k^*, k^*)$  where  $k^*$  is the size of the biggest  $\sigma$  involved and by letting the degree be as big as the biggest flag present.

For example, we could use this for the Turán hypergraph problem of maximizing the hyperedge density in an  $a$ -uniform hypergraph on  $n$  vertices where  $a$ -uniform hypercliques of size  $b$  are forbidden by adding

$$\prod_{\{i_1, \dots, i_a\} \in E(Q)} x_{i_1 \dots i_a} \quad \forall b\text{-clique } Q \text{ in } K_n^a$$

to our ideal. In [19], Razborov gives a Cauchy-Schwarz proof that the maximum edge density in a 3-uniform hypergraph without 4-cliques is at most 0.561666. This result can be translated in the polynomial language using the above-mentioned framework.

- (2) *Digraphs*: For any optimization problem over digraphs on  $n$  vertices, we can use the ideal

$$\langle x_{ij}^2 - x_{ij} \quad \forall 1 \leq i, j \leq n \text{ such that } i \neq j \rangle.$$

Again, the ideal is  $\mathfrak{S}_n$ -invariant, and so our techniques can again be used here.

For example, we could use this for the Caccetta-Häggkvist conjecture which states that every simple digraph of order  $n$  with minimum outdegree of at least  $r$  has a cycle of length at most  $\lceil \frac{n}{r} \rceil$  by adding

$$\prod_{(i,j) \in E(C)} x_{ij} \quad \forall \text{ cycles } C \text{ such that } |C| \leq \lceil \frac{n}{r} \rceil$$

$$\prod_{j \in S} (1 - x_{ij}) \quad \forall i, \forall S \subseteq [n] \setminus \{i\} \text{ such that } |S| = n - r + 1$$

and check for feasibility. Indeed, the first set of constraints ensures that we get a directed graph on  $n$  vertices, the second set of constraints forbids cycles of length less or equal to  $\lceil \frac{n}{r} \rceil$ , and the third set of constraints sets the outdegree of every vertex to be at least  $r$ . Of course, here the degree of these constraints being quite high makes it unlikely that this technique would yield any interesting results unless  $r$  is very big.

- (3) *Tournaments*: For an optimization problem over all tournaments of size  $n$ , we can use the ideal

$$\begin{array}{ll} \langle x_{ij}^2 - x_{ij} & \forall 1 \leq i, j \leq n \text{ such that } i \neq j \\ x_{ij}x_{ji} & \forall 1 \leq i < j \leq n \\ (1 - x_{ij})(1 - x_{ji}) & \forall 1 \leq i < j \leq n \end{array}$$

since exactly one of the arcs  $(i, j)$  and  $(j, i)$  are present for any  $1 \leq i < j \leq n$ , as well as any additional constraints forbidding certain structures in the problem. Again, we have an ideal that is  $\mathfrak{S}_n$ -invariant, and our techniques can again be used here.

The general theory of flag algebras [18] is couched in a broad setting using tools from algebra, topology, and probability. It is a formal calculus that came about from an attempt to systematize and distill the many ad hoc methods in extremal combinatorics. In this paper, we have shown that sos proofs that arise from flag densities

are equivalent to polynomial sos proofs using symmetry-reduction. This connection allows for systematic ways to search for flag sos proofs which was part of our original motivation in undertaking this project. There are likely much further, and deeper, connections between Razborov’s theory in its full generality, and the fundamental structures of real algebraic geometry. We hope to delve deeper in this direction.

In [17], the authors, along with James Saunderson, established the converse to the result in this paper. By a  $k$ -subset hypercube we mean a hypercube whose coordinates are indexed by the  $k$ -element subsets of  $[n]$ . Thus the usual hypercube  $\{0, 1\}^n$  is the 1-subset hypercube. The 2-subset hypercube  $\{0, 1\}^{\binom{[n]}{2}}$  arises in the context of optimization over (the edges of) a graph. The main result of [17] is that flag methods can be used to provide sos certificates for the nonnegativity of symmetric polynomials over  $k$ -subset hypercubes. This extends their use beyond the realm of extremal combinatorics into general polynomial optimization. The two papers together establish that flag methods are equivalent to standard symmetry-reduction methods in polynomial sos theory over the class of  $k$ -subset hypercubes.

### 7. APPENDIX

We present here the full translation of [6] to polynomials, culminating with a verification of the equivalence (7). We first discuss the (non-trivial) inequality whose nonnegativity establishes the theorem. A few other quantities are also needed for the proof in [6] which we define here.

Fix  $m < n$ , and consider  $H \in \mathcal{G}_m$ . Then we obtain the following equivalences:

$$(26) \quad \rho_n \equiv \sum_{H \in \mathcal{G}_m} \rho_H \rho_m(\mathbb{1}_H) \pmod{\mathcal{I}_n^{K_3}},$$

$$(27) \quad 1 \equiv \sum_{H \in \mathcal{G}_m} \rho_H \pmod{\mathcal{I}_n^{K_3}}.$$

Both equivalences follow from the fact that two polynomials are equivalent mod  $\mathcal{I}_n^{K_3}$  if and only if they have the same value on  $\mathbb{1}_G$  for each  $G \in \mathcal{G}_n$ . The first equivalence is since the edge density of  $G$  is the sum of the edge densities of  $H \in \mathcal{G}_m$  weighted by the density of  $H$  in  $G$ . The second is since  $\rho_H(\mathbb{1}_G)$  is the probability of  $H$  in  $G$ .

Note that  $(\mathbf{q}_H^\alpha)^2 \equiv \mathbf{q}_H^\alpha \pmod{\mathcal{I}_n^{K_3}}$  since  $x_{ij}$  and  $(1 - x_{ij})$  are equivalent to their squares mod  $\mathcal{I}_n^{K_3}$ , which implies that

$$(28) \quad \rho_H \equiv \frac{1}{a_H \binom{n}{|V(H)|}} \sum_{\alpha \in \text{Inj}(V(H), [n])} (\mathbf{q}_H^\alpha)^2 \pmod{\mathcal{I}_n^{K_3}},$$

and thus  $\rho_H$  is a sos mod  $\mathcal{I}_n^{K_3}$ , a fact that will be useful later. By multiplying 1 by  $\max_{H \in \mathcal{G}_m} \rho_m(\mathbb{1}_H)$  and then using equations (26) and (27), we get that

$$(29) \quad \max_{H \in \mathcal{G}_m} \rho_m(\mathbb{1}_H) - \rho_n \equiv \sum_{H \in \mathcal{G}_m} \underbrace{\left( \max_{H' \in \mathcal{G}_m} \rho_m(\mathbb{1}_{H'}) - \rho_m(\mathbb{1}_H) \right)}_{\geq 0} \rho_H \pmod{\mathcal{I}_n^{K_3}}.$$

By (28), the right-hand side of the above expression is a sos polynomial mod  $\mathcal{I}_n^{K_3}$ , and hence,  $\max_{H \in \mathcal{G}_m} \rho_m(\mathbb{1}_H)$  is an upper bound on the edge density of any  $G \in \mathcal{G}_n$ . Typically, this bound does not give a sufficiently tight answer.

Now suppose we also had a sos polynomial mod  $\mathcal{J}_n^{K_3}$  of the following type:

$$(30) \quad \sum_{H \in \mathcal{G}_m} c_H \mathbf{p}_H \equiv \sum_i s_i^2 \quad \text{mod } \mathcal{J}_n^{K_3}$$

where  $c_H \in \mathbb{R}$  for each  $H \in \mathcal{G}_m$ . Then adding inequality (30) to (26) and by a similar calculation to that in (29), we get

$$\begin{aligned} & \max_{H \in \mathcal{G}_m} (\mathbf{p}_m(\mathbf{1}_H) + c_H) - \mathbf{p}_n \equiv \\ & \sum_{H \in \mathcal{G}_m} \left( \underbrace{\left( \max_{H' \in \mathcal{G}_m} \mathbf{p}_m(\mathbf{1}_{H'}) + c_{H'} \right) - (\mathbf{p}_m(\mathbf{1}_H) + c_H)}_{\geq 0} \right) \mathbf{p}_H + \sum_k s_k^2 \quad \text{mod } \mathcal{J}_n^{K_3}. \end{aligned}$$

Again, since the right-hand side of the equivalence above is sos, we obtain the improved bound of

$$(31) \quad \max_{H \in \mathcal{G}_m} (\mathbf{p}_m(\mathbf{1}_H) + c_H)$$

on the edge density of any  $G \in \mathcal{G}_n$ . So the goal becomes to find equivalences as in (30) which then feeds into (31) to yield the desired bound of  $\frac{1}{2} + O(\frac{1}{n})$ . We now explain how this is done using flag algebras.

We must first define the density of two flags. Fix a type  $\sigma$  of size  $k$  and  $l \geq k$ . For a single flag  $F \in \mathcal{F}_l^\sigma$  and a fixed labeling  $\theta$  of  $k$  vertices in  $G$  using all the labels in  $[k]$ , we have the density polynomial  $\mathbf{p}_F^\theta$ . Recall that  $\mathbf{p}_F^\theta(\mathbf{1}_G)$  is the probability that the  $k$  vertices of  $G$  labeled by  $\theta$  along with  $l - k$  unlabeled vertices picked uniformly at random induce a copy of  $F$  in  $G$ . Now continue to select a disjoint set of  $l' - k$  unlabeled vertices in  $G$  uniformly at random. Then given a second  $\sigma$ -flag  $F' \in \mathcal{F}_{l'}^\sigma$ ,  $\mathbf{p}_{F, F'}^\theta(\mathbf{1}_G)$  is the probability that the two sets of unlabeled vertices each taken separately with the graph induced by  $\theta$  induce copies of  $F$  and  $F'$  respectively, where

$$\begin{aligned} \mathbf{p}_{F, F'}^\theta &= \frac{1}{\binom{n-k}{l-k} \binom{n-l}{l'-k}} \sum_{\substack{S, S' \subseteq [n]: \\ |S|=l, |S'|=l' \\ S \cap S' = \text{im}(\theta)}} \left( \frac{1}{a_F^\sigma} \sum_{\alpha \in \text{Inj}_\theta(V(F), S)} \mathbf{q}_F^\alpha \right) \left( \frac{1}{a_{F'}^\sigma} \sum_{\alpha \in \text{Inj}_\theta(V(F'), S')} \mathbf{q}_{F'}^\alpha \right) \\ &= \frac{1}{a_F^\sigma a_{F'}^\sigma \binom{n-k}{l-k} \binom{n-l}{l'-k}} \sum_{\substack{\alpha_1 \in \text{Inj}_\theta(V(F), [n]) \\ \alpha_2 \in \text{Inj}_\theta(V(F'), [n] \setminus \alpha_1(V(F) \setminus \sigma))}} \mathbf{q}_F^{\alpha_1} \mathbf{q}_{F'}^{\alpha_2}. \end{aligned}$$

Note that  $\mathbf{p}_{F, F'}^\theta(\mathbf{1}_G) = p(F_1, F_2; F_3)$  in [18] where  $F_1 = (F, \theta) \in \mathcal{F}_l^\sigma$ ,  $F_2 = (F', \theta) \in \mathcal{F}_{l'}^\sigma$  and  $F_3 = (G, \theta) \in \mathcal{F}_n^\sigma$  where  $|\sigma| = k$ .

Next, let  $\mathbf{p}_F := \mathbb{E}_\theta \mathbf{p}_F^\theta \quad \text{mod } \mathcal{J}_n^{K_3}$  where  $\mathbf{p}_F$  starts by choosing a map  $\theta : [k] \rightarrow [n]$  uniformly at random and then computes  $\mathbf{p}_F^\theta$  for it. Similarly, let  $\mathbf{p}_{F, F'} := \mathbb{E}_\theta \mathbf{p}_{F, F'}^\theta \quad \text{mod } \mathcal{J}_n^{K_3}$  where  $\mathbf{p}_{F, F'}$  starts by choosing a map  $\theta : [k] \rightarrow [n]$  uniformly at random and then computes  $\mathbf{p}_{F, F'}^\theta$  for it.

If  $k, l$  and  $l'$  are fixed and  $n$  is large, then picking two random extensions of  $\theta$  of sizes  $l - k$  and  $l' - k$  is essentially the same as picking two disjoint extensions. Indeed, the probability that the two random extensions share  $d$  vertices is roughly  $\frac{d}{n}$  where  $d \leq \max\{l, l'\} - k$ . Therefore, the probability of the two random extensions overlapping is  $O(\frac{1}{n})$  where we hide factors depending on  $k$  and  $\max\{l, l'\}$ . This implies

that for all partial labelings  $\theta : [k] \rightarrow [n]$  there is a polynomial  $\text{err}_{F,F'}^\theta(\mathbf{x})$  such that

$$(32) \quad \mathfrak{p}_F \mathfrak{p}_{F'}^\theta \equiv \mathfrak{p}_{F,F'}^\theta + \text{err}_{F,F'}^\theta \pmod{\mathcal{I}_n^{K_3}}, \text{ and}$$

$$(33) \quad \text{err}_{F,F'}^\theta \equiv O\left(\frac{1}{n}\right) \pmod{\mathcal{I}_n^{K_3}}.$$

Taking expectation with respect to the uniform distribution over all possible partial labelings  $\theta$ , and letting  $\text{err}_{F,F'}(\mathbf{x}) = \mathbb{E}_\theta \text{err}_{F,F'}^\theta(\mathbf{x})$ , we obtain that

$$(34) \quad \mathbb{E}_\theta \mathfrak{p}_F \mathfrak{p}_{F'}^\theta \equiv \mathfrak{p}_{F,F'} + \text{err}_{F,F'} \pmod{\mathcal{I}_n^{K_3}}$$

where again  $\text{err}_{F,F'} \equiv O\left(\frac{1}{n}\right) \pmod{\mathcal{I}_n^{K_3}}$ .

Going back to  $m$ -vertex graphs  $H \in \mathcal{G}_m$ , if  $m \geq l + l' - k$ , then by the same reasoning as in (26), we also have

$$(35) \quad \mathfrak{p}_{F,F'} \equiv \sum_{H \in \mathcal{G}_m} \mathfrak{p}_H \mathfrak{p}_{F,F'}(\mathbb{1}_H) \pmod{\mathcal{I}_n^{K_3}}.$$

We can now use these observations to produce an expression of the form (30) as follows. For any psd matrix  $Q$  of size  $|\mathcal{F}_l^\sigma| \times |\mathcal{F}_l^\sigma|$ , the polynomial  $\sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_F \mathfrak{p}_{F'}^\theta$ , and hence,  $\mathbb{E}_\theta \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_F \mathfrak{p}_{F'}^\theta$ , are sos mod  $\mathcal{I}_n^{K_3}$ . Now note that

$$\begin{aligned} \mathbb{E}_\theta \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_F \mathfrak{p}_{F'}^\theta &\equiv \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} (\mathfrak{p}_{F,F'} + \text{err}_{F,F'}) \pmod{\mathcal{I}_n^{K_3}} \\ &\equiv \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_{F,F'} + \text{err} \pmod{\mathcal{I}_n^{K_3}} \\ &\equiv \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \left( \sum_{H \in \mathcal{G}_m} \mathfrak{p}_H \mathfrak{p}_{F,F'}(\mathbb{1}_H) \right) + \text{err} \pmod{\mathcal{I}_n^{K_3}} \\ &\equiv \sum_{H \in \mathcal{G}_m} \mathfrak{p}_H \underbrace{\left( \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_{F,F'}(\mathbb{1}_H) \right)}_{c_H} + \text{err} \pmod{\mathcal{I}_n^{K_3}} \end{aligned}$$

where  $\text{err}(\mathbb{1}_G)$  has value  $O\left(\frac{1}{n}\right)$  for every  $G \in \mathcal{G}_n$ , i.e., on the zeros of  $\mathcal{I}_n^{K_3}$ . Therefore, just as we derived the bound on edge density following (30), we get that modulo  $\mathcal{I}_n^{K_3}$ ,

$$\begin{aligned} \max_{H \in \mathcal{G}_m} (\mathfrak{p}_m(\mathbb{1}_H) + c_H) - \mathfrak{p}_n + \text{err} &\equiv \\ \sum_{H \in \mathcal{G}_m} \left( \max_{H' \in \mathcal{G}_m} (\mathfrak{p}_m(\mathbb{1}_{H'}) + c_{H'}) - (\mathfrak{p}_m(\mathbb{1}_H) + c_H) \right) \mathfrak{p}_H &+ \mathbb{E}_\theta \sum_{F,F' \in \mathcal{F}_l^\sigma} Q_{F,F'} \mathfrak{p}_F \mathfrak{p}_{F'}^\theta. \end{aligned}$$

Since the right-hand side is a sos expression for any graph  $G \in \mathcal{G}_n$ , we have  $\mathfrak{p}_n(\mathbb{1}_G) \leq \mathfrak{p}_m(\mathbb{1}_H) + c_H + O\left(\frac{1}{n}\right)$  which is an upper bound on the edge density of  $G$ .

In summary, to apply the flag Cauchy-Schwarz approach to the Mantel problem, one chooses parameters  $k, l$  and  $m$  appropriately. Then the choice of the psd matrix  $Q$  that obtains the best possible bound on the triangle-free density as in (31) reduces to solving a semidefinite program whose size depends on  $k, l$  and  $m$  and is independent of  $n$ . Note that in certain proofs, one must choose several  $k, l, m$  and take a linear combination of (30) to obtain the desired upper bound; this however doesn't change the explanations above in a significant way.



We now illustrate the above procedure on the Mantel example. Set  $k = 1$  and take  $\sigma$  to be a single vertex labeled “1”. Taking  $l = 2$ , we have the set of flags  $\mathcal{F}_l^\sigma = \{F_0, F_1\}$  where

$$F_0 = \begin{array}{c} \bullet \\ | \\ \bullet \end{array}, \quad F_1 = \begin{array}{c} \bullet \\ / \\ \bullet \end{array}.$$

Now we choose  $m = 3$  so that  $\mathcal{H}$  is the set of all triangle-free graphs on three vertices. Up to isomorphism, there is a unique graph on three vertices with  $i$  edges for  $i = 0, 1, 2, 3$ . Call these  $H_0, H_1, H_2$ , and  $H_3$ , i.e.,

$$H_0 = \bullet \bullet \bullet, \quad H_1 = \begin{array}{c} \bullet \\ / \\ \bullet \end{array} \bullet, \quad H_2 = \begin{array}{c} \bullet \\ / \\ \bullet \end{array} \begin{array}{c} \bullet \\ \backslash \\ \bullet \end{array}, \quad H_3 = \begin{array}{c} \bullet \\ / \\ \bullet \end{array} \begin{array}{c} \bullet \\ \backslash \\ \bullet \end{array}.$$

Therefore,  $\mathcal{G}_3 = \{H_0, H_1, H_2\}$  since  $H_3$  is not triangle-free.

The density polynomials  $\mathfrak{p}_{H_i}$  for each  $0 \leq i \leq 3$  can be defined according to equation (2). For example, we have  $\mathfrak{p}_{H_0} = \frac{1}{\binom{n}{3}} \sum_{i < j < k} (1 - x_{ij})(1 - x_{ik})(1 - x_{jk})$ .

Using the fact that  $\mathfrak{p}_{H_3} \equiv 0 \pmod{\mathcal{I}_n^{K_3}}$ , we rewrite equations (26) and (27) as:

$$\begin{aligned} \mathfrak{p}_n &\equiv \frac{1}{3}\mathfrak{p}_{H_1} + \frac{2}{3}\mathfrak{p}_{H_2} \pmod{\mathcal{I}_n^{K_3}}, \\ 1 &\equiv \mathfrak{p}_{H_0} + \mathfrak{p}_{H_1} + \mathfrak{p}_{H_2} \pmod{\mathcal{I}_n^{K_3}}. \end{aligned}$$

We now calculate  $\mathfrak{p}_{F_i, F_j}(\mathbb{1}_{H_k})$  for every  $i, j \in \{0, 1\}$  and  $k \in \{0, 1, 2, 3\}$  and use (35) to obtain the following relations

$$\begin{aligned} \mathfrak{p}_{F_0, F_0} &= \mathfrak{p}_{H_0} + \frac{1}{3}\mathfrak{p}_{H_1}, \\ \mathfrak{p}_{F_0, F_1} = \mathfrak{p}_{F_1, F_0} &= \frac{1}{3}\mathfrak{p}_{H_1} + \frac{1}{3}\mathfrak{p}_{H_2}, \\ \mathfrak{p}_{F_1, F_1} &= \frac{1}{3}\mathfrak{p}_{H_2} + \mathfrak{p}_{H_3} \equiv \frac{1}{3}\mathfrak{p}_{H_2} \pmod{\mathcal{I}_n^{K_3}}. \end{aligned}$$

Choosing  $Q = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ , we obtain the following polynomial identity

$$(36) \quad \frac{1}{2}\mathfrak{p}_{H_0} - \frac{1}{6}\mathfrak{p}_{H_1} - \frac{1}{6}\mathfrak{p}_{H_2} + \text{err} \equiv \mathbb{E}_\theta \left[ \begin{pmatrix} \mathfrak{p}_{F_0}^\theta & \mathfrak{p}_{F_1}^\theta \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \mathfrak{p}_{F_0}^\theta \\ \mathfrak{p}_{F_1}^\theta \end{pmatrix} \right]$$

as in the sos expression (30). Thus we get the following proof:

$$\begin{aligned} \frac{1}{2} - \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij} + \text{err} &\equiv \frac{1}{2}\mathfrak{p}_{H_0} - \frac{1}{6}\mathfrak{p}_{H_1} - \frac{1}{6}\mathfrak{p}_{H_2} + \frac{1}{3}\mathfrak{p}_{H_1} + \text{err} \pmod{\mathcal{I}_n^{K_3}} \\ &\equiv \mathbb{E}_\theta \left[ \begin{pmatrix} \mathfrak{p}_{F_0}^\theta & \mathfrak{p}_{F_1}^\theta \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \mathfrak{p}_{F_0}^\theta \\ \mathfrak{p}_{F_1}^\theta \end{pmatrix} \right] + \frac{1}{3}\mathfrak{p}_{H_1} \pmod{\mathcal{I}_n^{K_3}}. \end{aligned}$$

*Acknowledgements.* We thank Greg Blekherman for several important inputs to this paper relating to representation theory. They came at crucial junctures and helped us along greatly. We also thank Andrew Berget, Monty McGovern, Pablo Parrilo, James Pfeiffer, Paul Smith and Vasu Tewari for helpful conversations and suggestions. We also thank the referees of this paper for their valuable comments which have improved the content and exposition.

## REFERENCES

- [1] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin, *Invariant semidefinite programs*, in Handbook on Semidefinite, Conic and Polynomial Optimization, Internat. Ser. Oper. Res. Management Sci., vol. 166, Springer, New York, 2012, pp. 219–269.
- [2] G. Blekherman, P. A. Parrilo, and R. R. Thomas (eds.), *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Series on Optimization, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013.
- [3] F. Chung and L. Lu, *An upper bound for the Turán number  $t_3(n, 4)$* , J. Combin. Theory Ser. A **87** (1999), no. 2, 381–389.
- [4] E. de Klerk, F. M. de Oliveira Filho, and D.V. Pasechnik, *Relaxations of combinatorial problems via association schemes*, in Handbook on Semidefinite, Conic and Polynomial Optimization, Internat. Ser. Oper. Res. Management Sci., vol. 166, Springer, New York, 2012, pp. 171–199.
- [5] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc **52** (1946), no. 3, 1087–1091.
- [6] 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** (2013), no. 1, 21–54.
- [7] A. Fässler and E. Stiefel, *Group theoretical methods and their applications*, Birkhäuser, 1992.
- [8] P. Frankl and Z. Füredi, *An exact result for 3-graphs*, Discrete Math. **50** (1984), no. 2-3, 323–328.
- [9] W. Fulton and J. Harris, *Representation theory: A first course*, Graduate Texts in Mathematics, no. 129, Springer, 1991.
- [10] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128.
- [11] P. Keevash, *Hypergraph Turán problems*, Surveys in combinatorics **392** (2011), 83–140.
- [12] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging Applications of Algebraic Geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.
- [13] L. Lu and Y. Zhao, *An exact result for hypergraphs and upper bounds for the Turán density of  $K_{r+1}^T$* , SIAM J. Discrete Math. **23** (2009), no. 3, 1324–1334.
- [14] W. Mantel, *Problem 28*, Wiskundige Opgaven **10** (1907), no. 320, 60–61.
- [15] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., Ser. B **96** (2003), no. 2, 293–320.
- [16] O. Pikhurko, *An exact Turán result for the generalized triangle*, Combinatorica **28** (2008), no. 2, 187–208.
- [17] A. Raymond, J. Saunderson, M. Singh, and R.R. Thomas, *Symmetric sums of squares over  $k$ -subset hypercubes*, Mathematical Programming **167** (2018), no. 2, 315–354.
- [18] A. A. Razborov, *Flag algebras*, J. Symbolic Logic **72** (2007), no. 4, 1239–1282.
- [19] ———, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM Journal on Discrete Mathematics **24** (2010), no. 3, 946–963.
- [20] ———, *Flag algebras: an interim report*, in The Mathematics of Paul Erdős II, Springer, 2013, pp. 207–232.
- [21] ———, *On Turán’s  $(3, 4)$ -problem with forbidden subgraphs*, Mathematical Notes **95** (2014), no. 1-2, 245–252.
- [22] C. Riemer, T. Theobald, L. Jansson Andrén, and J. B. Lasserre, *Exploiting symmetries in sdp-relaxations for polynomial optimization*, Mathematics of Operations Research **38** (2013), no. 1, 122–141.
- [23] B. E. Sagan, *The Symmetric Group*, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001.
- [24] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inform. Theory **25** (1979), 425–429.
- [25] A. F. Sidorenko, *Asymptotic solution for a new class of forbidden  $r$ -graphs*, Combinatorica **9** (1989), no. 2, 207–215.
- [26] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.
- [27] P. Turán, *Research problem*, Közl. MTA Mat. Kutató Int. **6** (1961), 417–423.

ANNIE RAYMOND, Department of Mathematics and Statistics, University of Massachusetts Amherst,  
Amherst, MA 01003  
*E-mail* : `raymond@math.umass.edu`

MOHIT SINGH, H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute  
of Technology, Atlanta, GA 30332  
*E-mail* : `mohit.singh@isye.gatech.edu`

REKHA R. THOMAS, Department of Mathematics, University of Washington, Box 354350, Seattle,  
WA 98195  
*E-mail* : `rrthomas@uw.edu`