FINDING PLANTED SUBGRAPHS WITH FEW EIGENVALUES USING THE SCHUR–HORN RELAXATION[*]

UTKAN ONUR CANDOGAN[†] AND VENKAT CHANDRASEKARAN[‡]

Abstract. Extracting structured subgraphs inside large graphs—often known as the planted subgraph problem—is a fundamental question that arises in a range of application domains. This problem is NP-hard in general and, as a result, significant efforts have been directed towards the development of tractable procedures that succeed on specific families of problem instances. We propose a new computationally efficient convex relaxation for solving the planted subgraph problem; our approach is based on tractable semidefinite descriptions of majorization inequalities on the spectrum of a symmetric matrix. This procedure is effective at finding planted subgraphs that consist of few distinct eigenvalues, and it generalizes previous convex relaxation techniques for finding planted cliques. Our analysis relies prominently on the notion of *spectrally comonotone* matrices, which are pairs of symmetric matrices that can be transformed to diagonal matrices with sorted diagonal entries upon conjugation by the same orthogonal matrix.

Key words. convex optimization, distance-regular graphs, induced subgraph isomorphism, majorization, orbitopes, semidefinite programming, strongly regular graphs

AMS subject classifications. 90C25, 90C22, 90C90, 90C35

DOI. 10.1137/16M1075144

1. Introduction. In application domains ranging from computational biology to social data analysis, graphs are frequently used to model relationships among large numbers of interacting entities. A commonly encountered question across many of these application domains is that of identifying structured subgraphs inside larger graphs. For example, identifying specific motifs or substructures inside gene regulatory networks is useful in revealing higher-order biological function [3, 15, 30]. Similarly, extracting completely connected subgraphs in social networks is useful for determining communities of people that are mutually linked to each other [29, 31, 33]. In this paper, we propose a new algorithm based on convex optimization for finding structured subgraphs inside large graphs, and we give conditions under which our approach succeeds in performing this task.

Formally, suppose Γ and $\mathfrak{G}$ are graphs[1] on k nodes and n nodes (here $n > k$), respectively, with the following property: there exists a subset of vertices $V \subset \{1, \ldots, n\}$ with $|V| = k$ such that the induced subgraph of $\mathfrak{G}$ corresponding to the vertex set V is isomorphic to Γ. The *planted subgraph* problem is to identify the vertex subset V given the graphs $\mathfrak{G}$ and Γ; see Figure 1 for an example. The decision version of the planted subgraph problem is known as the induced subgraph isomorphism problem in the theoretical computer science literature, and it has been shown to be NP-hard [26]. Nevertheless, as this problem arises in a wide range of application domains as

[†]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 (utkan@caltech.edu).

[‡]Departments of Computing and Mathematical Sciences and of Electrical Engineering, California Institute of Technology, Pasadena, CA 91106 (venkatc@caltech.edu).

[1]Throughout this paper we consider undirected, unweighted, loopless graphs.
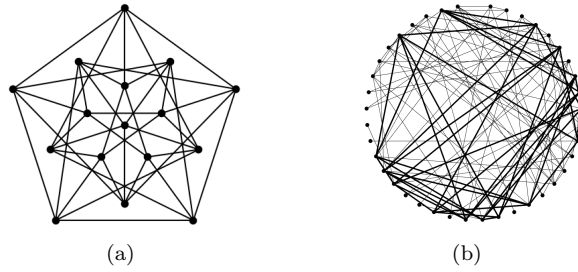
(a)   (b)

FIG. 1. *The Clebsch graph (16 nodes) on the left. An example on the right of a 40-node graph containing the Clebsch graph as an induced subgraph; the thick edges correspond to a 16-node induced subgraph that is isomorphic to the Clebsch graph.*

described above, significant efforts have been directed towards the development of computationally tractable procedures that succeed on certain families of problem instances. Much of the focus of this attention has been on the special case of the *planted clique* problem in which the subgraph Γ is fully connected. Alon, Krivelevich, and Sudakov [1] and Feige and Krauthgamer [20] developed a spectral algorithm for the planted clique problem, and subsequently Ames and Vavasis [2] described an approach based on semidefinite programming with similar performance guarantees to the earlier work based on spectral algorithms. Conceptually, these methods are based on a basic observation about the spectrum of a clique, namely, that the adjacency matrix of a clique on k nodes has two distinct eigenvalues, one with multiplicity equal to one and the other with multiplicity equal to $k - 1$. We describe a new semidefinite programming technique that generalizes the method of Ames and Vavasis [2] to planted subgraphs Γ that are not fully connected, with the spectral properties of Γ playing a prominent role in our algorithm and our analysis.

1.1. Our contributions. Let $A_\Gamma \in \mathbb{S}^k$ and $A_{\mathfrak{G}} \in \mathbb{S}^n$ represent the adjacency matrices of Γ and of $\mathfrak{G}$ with $\mathbb{S}^q$ denoting the space of $q \times q$ real symmetric matrices. Given any matrix $M \in \mathbb{S}^k$, we let $[M]_{k \to n} \in \mathbb{S}^n$ for $n > k$ denote an $n \times n$ symmetric matrix with the leading principal minor of order k equal to M and all the other entries equal to zero. The following combinatorial optimization problem is a natural first approach to phrase the planted subgraph problem in a variational manner:

$$
\begin{aligned}
\hat{A}_{co} = \underset{A \in \mathbb{S}^n}{\arg\max} \ & \operatorname{trace}(A \cdot A_{\mathfrak{G}}) \\
\text{s.t. } & A_{i,j} = 0 \text{ if } (A_{\mathfrak{G}})_{i,j} = 0 \text{ and } i \neq j, \\
& A \in \{\Pi[A_\Gamma]_{k \to n}\Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}.
\end{aligned}
\tag{1}
$$

Assuming that there is no other subgraph of $\mathfrak{G}$ that is isomorphic to Γ, one can check that the optimal solution $\hat{A}_{co}$ of this problem identifies the vertices $V \subset \{1, \dots, n\}$ whose induced subgraph in $\mathfrak{G}$ is isomorphic to Γ, i.e., the unique optimal solution $\hat{A}_{co}$ is equal to zero everywhere except for the principal minor corresponding to the indices in V and $(\hat{A}_{co})_{V,V} = \tilde{\Pi}A_\Gamma\tilde{\Pi}'$ for some $k \times k$ permutation matrix $\tilde{\Pi}$. However, solving (1) is intractable in general. Replacing the combinatorial constraint $A \in \{\Pi[A_\Gamma]_{k \to n}\Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}$ with the convex constraint $A \in \operatorname{conv}\{\Pi[A_\Gamma]_{k \to n}\Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}$ does not lead to a tractable problem as checking membership in the polytope $\operatorname{conv}\{\Pi[A_\Gamma]_{k \to n}\Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}$ is intractable for general planted graphs Γ (unless P = NP).

We describe next a convex outer approximation of the set $\{\Pi[A_\Gamma]_{k\to n}\Pi' \mid \Pi \text{ is an }$ $n \times n$ permutation matrix$\}$ that leads to a tractable convex program. For any matrix $M \in \mathbb{S}^n$, the *Schur–Horn orbitope* $\mathcal{SH}(M) \subset \mathbb{S}^n$ is defined as [36]

$$(2) \qquad \mathcal{SH}(M) = \text{conv}\{UMU' \mid U \text{ is an } n \times n \text{ orthogonal matrix}\}.$$

The term "orbitope" was coined by Sanyal, Sottile, and Sturmfels in their work on convex hulls of orbits generated by the actions of groups, and the Schur–Horn orbitope was so named by these authors due to its connection to the Schur–Horn theorem in linear algebra [36]. In combinatorial optimization, approximations based on replacing permutations matrices by orthogonal matrices have also been employed to obtain bounds on the quadratic assignment problem [21]. The set $\mathcal{SH}(M)$ depends only on the eigenvalues of M, and it is clearly an outer approximation of the set $\{\Pi M \Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}$. Crucially for our purposes, the Schur–Horn orbitope $\mathcal{SH}(M)$ for any $M \in \mathbb{S}^n$ has a tractable semidefinite description via majorization inequalities on the spectrum of a symmetric matrix [6, 36]; see section 4.1. Hence, we propose the following tractable semidefinite programming relaxation for the planted subgraph problem:

$$\hat{A}_{sh} = \arg\max_{A \in \mathbb{S}^n} \text{trace}(A \cdot A_{\mathfrak{G}})$$

$$(P) \qquad\qquad \text{s.t. } A_{i,j} = 0 \text{ if } (A_{\mathfrak{G}})_{i,j} = 0 \text{ and } i \neq j,$$
$$A \in \mathcal{SH}([A_\Gamma - \gamma I_k]_{k\to n}).$$

Here $I_k \in \mathbb{S}^k$ is the $k \times k$ identity matrix. We refer to this convex program as the *Schur–Horn relaxation*, and this problem can be solved to a desired precision in polynomial time. This relaxation only requires knowledge of the eigenvalues of the planted graph Γ. The parameter $\gamma \in \mathbb{R}$ is to be specified by the user, and we discuss suitable choices for γ in the following. Note that changing A_Γ to $A_\Gamma - \gamma I_k$ in the constraints of (1) essentially leaves that problem unchanged (the nonzero principal minor of the optimal solution simply changes from $\hat{A}_{co}$ to $\hat{A}_{co} - \gamma I_k$). However, the additional degree of freedom provided by the parameter γ plays a more significant role in the Schur–Horn relaxation as it allows for shifts of the spectrum of A_Γ to more favorable values, which is essential for the solution of various planted subgraph problems; see section 2.1 for further details, as well as the experiments in section 4 for numerical illustrations. We say that the Schur–Horn relaxation succeeds in recovering the planted subgraph Γ if the optimal solution $\hat{A}_{sh} \in \mathbb{S}^n$ satisfies the following conditions: the optimal solution $\hat{A}_{sh}$ is unique, the submatrix $(\hat{A}_{sh})_{V,V} = \tilde{\Pi} A_\Gamma \tilde{\Pi}' - \gamma I_k$ for some $k \times k$ permutation matrix $\tilde{\Pi}$, and the remaining entries of $\hat{A}_{sh}$ are equal to zero.

In section 2 we study the geometric properties of the Schur–Horn orbitope as these pertain to the optimality conditions of the Schur–Horn relaxation. Our analysis relies prominently on the notion of spectrally comonotone matrices, which refers to a pair of symmetric matrices that can be transformed to diagonal matrices with sorted diagonal entries upon conjugation by the same orthogonal matrix. Spectral comonotonicity is a more restrictive condition than simultaneous diagonalizability, and it enables a precise characterization of the normal cones at extreme points of the Schur–Horn orbitope (Proposition 6). This discussion leads directly to the central observation of our paper that the Schur–Horn relaxation is useful for finding planted graphs Γ that consist of *few distinct eigenvalues*. Cliques form the simplest examples of such graphs as their spectrum consists of two distinct eigenvalues. There are numerous other graph families whose spectrum consists of few distinct eigenvalues, and the study of such

graphs is a significant topic in graph theory [8, 17, 18, 32, 39, 40, 41]. For example, strongly regular graphs are (an infinite family of) regular graphs with three distinct eigenvalues; the Clebsch graph of Figure 1 is a strongly regular graph on 16 nodes with eigenvalues in the set $\{5, 1, -3\}$ and degree equal to five. For a more extensive list of graphs with few eigenvalues, see section 2.2.

We state and prove the main theoretical result of this paper in section 3.4—see Theorem 19. If the planted subgraph Γ and its complement are both *symmetric*— Γ and its complement are both vertex- and edge-transitive—and if Γ is connected, then this theorem takes on a simpler form (Corollary 20) . Specifically, the success of the Schur–Horn relaxation (P) relies on the existence of a suitable eigenspace $\mathcal{E} \subset \mathbb{R}^k$ of A_Γ. Concretely, let $\mathcal{P}_\mathcal{E} \in \mathbb{S}^k$ denote the projection onto $\mathcal{E}$, and let $\mu(\mathcal{E}) = \max_{i,j,\ i \neq j} \frac{|(\mathcal{P}_\mathcal{E})_{i,j}|}{\sqrt{|(\mathcal{P}_\mathcal{E})_{i,i}||(\mathcal{P}_\mathcal{E})_{j,j}|}}$ denote the coherence of $\mathcal{E}$. Assuming that the edges in $\mathfrak{G}$ outside the induced subgraph Γ are placed independently and uniformly at random with probability $p \in [0, \frac{1}{\mu(\mathcal{E})k})$ (i.e., the Erdős–Rényi random graph model), we show in Corollary 20 that the Schur–Horn relaxation (P) with parameter[2] $\gamma = \lambda_\mathcal{E}$ (the eigenvalue associated with $\mathcal{E}$) succeeds with high probability provided

$$n \lesssim \min_{\substack{\lambda \text{ eigenvalue of } A_\Gamma \\ \lambda \neq \lambda_\mathcal{E}}} \min \left\{ |\lambda - \lambda_\mathcal{E}|^2 \, \frac{\dim(\mathcal{E})^2 \big(1 - kp\mu(\mathcal{E})\big)}{k^2 \, p}, (|\lambda - \lambda_\mathcal{E}| - 2|\lambda_\mathcal{E}|)^2 \right\} + k.$$

The coherence parameter $\mu(\mathcal{E})$ lies in $(0, 1]$, and it appears prominently in results on sparse signal recovery via convex optimization [16]. In analogy to that literature, a small value of $\mu(\mathcal{E})$ is useful in our context (informally) to ensure that the planted graph Γ looks sufficiently "different" from the remainder of $\mathfrak{G}$ (see section 3 for details). Thus, the Schur–Horn relaxation succeeds if the planted graph Γ consists of few distinct eigenvalues that are well-separated, and in which one of the eigenspaces has a small coherence parameter associated to it. For more general nonsymmetric graphs, our main result (Theorem 19) is stated in terms of a parameter associated with an eigenspace $\mathcal{E}$ of A_Γ called the *combinatorial width*, which roughly measures the average conditioning over all minors of $\mathcal{P}_\mathcal{E}$ of a certain size.

Specialization to the planted clique problem. The sum of the adjacency matrix of a clique and the identity matrix has rank equal to one, and consequently the planted clique problem may be phrased as one of identifying a rank-one submatrix inside a larger matrix (up to shifts of the diagonal by the identity matrix). In her thesis [19], Fazel proposed the nuclear norm as a tractable convex surrogate for identifying low-rank matrices in convex sets, and subsequent efforts provided theoretical support for the effectiveness of this relaxation in a range of rank minimization problems [12, 34]. Building on these ideas, Ames and Vavasis [2] proposed a nuclear norm minimization approach for the planted clique problem. The Schur–Horn relaxation (P) specializes to the relaxation in [2] when Γ is the clique. Specifically, letting $A_{\mathrm{clique}} \in \mathbb{S}^k$ denote the adjacency matrix of a k-clique, one can check that

$$(3) \qquad \mathcal{SH}([A_{\mathrm{clique}} + I_k]_{k \to n}) = \{P \in \mathbb{S}^n \mid \mathrm{trace}(P) = k, \ P \succeq 0\}.$$

As the nuclear norm of a positive semidefinite matrix is equal to its trace, the Schur–Horn orbitope $\mathcal{SH}([A_\Gamma + I_k]_{k \to n})$ is simply a face of the nuclear norm ball in $\mathbb{S}^n$ scaled by a factor k. Thus, the Schur–Horn relaxation (P) with $\gamma = -1$ is effectively a nuclear

---

[2]In our experiments in section 4, we set γ equal to the eigenvalue of A_Γ with the largest multiplicity. See section 3 for further discussion.

norm relaxation when the planted subgraph of interest is the clique.[3] Further, our main result (Theorem 19) can be specialized to the case of a planted clique to obtain the main result in [2]; see Corollary 21.

1.2. Paper outline. In section 2 we discuss the geometric properties of the Schur–Horn orbitope and their connection to the optimality conditions of the Schur–Horn relaxation, along with an extensive list of families of graphs with few eigenvalues. Section 3 contains our main theoretical results, while in section 4 we demonstrate the utility of the Schur–Horn relaxation in practice via numerical experiments. We conclude in section 5 with a discussion of further research directions.

Notation. The *normal cone* at a point $x \in \mathcal{C}$ for a closed, convex set $\mathcal{C} \subset \mathbb{R}^n$ is denoted by $\mathcal{N}_{\mathcal{C}}(x)$ and it is the collection of linear functionals that attain their maximal value over $\mathcal{C}$ at x [35]. The projection operator onto a subspace $\mathcal{E} \subset \mathbb{R}^n$ is denoted by $\mathcal{P}_{\mathcal{E}}$. The restriction of a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ to an invariant subspace $\mathcal{E}$ of A is denoted by $A|_{\mathcal{E}} : \mathcal{E} \to \mathcal{E}$. The orthogonal complement of a subspace $\mathcal{E}$ is denoted by $\mathcal{E}^{\perp}$. The notation $\dim(\mathcal{E})$ denotes the dimension of a subspace $\mathcal{E}$. The *eigengap* of a symmetric matrix $M \in \mathbb{S}^n$ associated with an invariant subspace $\mathcal{E} \subset \mathbb{R}^n$ of M is defined as

$$\mathrm{eigengap}(M, \mathcal{E}) = \min \big\{ |\lambda_{\mathcal{E}} - \lambda_{\mathcal{E}^{\perp}}| \mid \lambda_{\mathcal{E}} \text{ an eigenvalue of } M|_{\mathcal{E}},$$
$$\lambda_{\mathcal{E}^{\perp}} \text{ an eigenvalue of } M|_{\mathcal{E}^{\perp}} \big\}.$$

The smallest and largest eigenvalues of a symmetric matrix A are represented by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$, respectively. The norms $\|\cdot\|, \|\cdot\|_2$, and $\|\cdot\|_F$ denote the vector ℓ_2 norm, the matrix operator/spectral norm, and the matrix Frobenius norm, respectively. The vector $1_{\ell} \in \mathbb{R}^{\ell}$ denotes the all-ones vector of length ℓ. We denote the identity matrix of size k by I_k. The matrix $I_{\Omega} \in \mathbb{R}^{|\Omega| \times k}$ denotes the matrix whose rows are the rows of I_k indexed by $\Omega \subset \{1, \ldots, k\}$, so that the rows of $I_{\Omega}A$ are the rows of A indexed by Ω for any $A \in \mathbb{R}^{k \times q}$. The matrix $A_{\Omega,\Omega} \in \mathbb{R}^{|\Omega| \times |\Omega|}$ denotes the principal minor of A indexed by the set Ω. The group of $n \times n$ orthogonal matrices is denoted by $\mathcal{O}_n \subset \mathbb{R}^{n \times n}$. The set $\mathrm{relint}(\mathcal{C})$ specifies the relative interior of any convex set $\mathcal{C}$. The column space of a matrix A is denoted by $\mathrm{col}(A)$. The quantity $\mathbb{E}[\cdot]$ denotes the usual expected value, where the distribution is clear from context.

2. Geometric properties of the Schur–Horn orbitope. In this section, we analyze the optimality conditions of the Schur–Horn relaxation from a geometric perspective. In particular, the notion of a pair of *spectrally comonotone matrices* plays a central role in our development, and we elaborate on this point in the next subsection. Based on this discussion, we observe that the Schur–Horn relaxation is especially useful for finding planted graphs consisting of few distinct eigenvalues, and we give examples of graphs with this property in section 2.2. The main theoretical results formalizing the utility of the Schur–Horn relaxation are presented in section 3.

2.1. Optimality conditions of the Schur–Horn relaxation. We state the optimality conditions of the Schur–Horn relaxation in terms of the normal cones at extreme points of the Schur–Horn orbitope.

LEMMA 1. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_{\mathfrak{G}}$ of order k is equal to A_{Γ}. Suppose there exists a matrix $M \in \mathbb{S}^n$ with the following properties:*

---

[3]The nuclear norm relaxation in [2] is formulated in a slightly different fashion compared to the Schur–Horn relaxation (P) for the case of the planted clique; specifically, one can show that our relaxation succeeds whenever the nuclear norm relaxation in [2] succeeds.

1. $M_{i,j} = (A_{\mathfrak{G}})_{i,j}$ if $(A_{\mathfrak{G}})_{i,j} = 1$ or if $i = j$;
2. $M \in \text{relint}\left(\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})\right).$

Then the Schur–Horn relaxation succeeds at identifying the planted subgraph Γ *inside the larger graph* $\mathfrak{G}$*, i.e., the unique optimal solution of the convex program (P) is* $\hat{A}_{sh} = \left(\begin{smallmatrix} A_\Gamma - \gamma I_k & 0 \\ 0 & 0 \end{smallmatrix}\right).$

Proof. From standard results in convex analysis [35], we have that $\left(\begin{smallmatrix} A_\Gamma - \gamma I_k & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is the unique optimal solution of (P) if $A_{\mathfrak{G}}$ can be decomposed as $A_{\mathfrak{G}} \in K +$ relint $\left(\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})\right)$ for some matrix $K \in \mathbb{S}^n$ that satisfies

$$K_{i,j} = 0 \text{ if either } (A_{\mathfrak{G}})_{i,j} = 1 \text{ or } i = j.$$

Letting $K = A_{\mathfrak{G}} - M$ we have the desired result.   □

The assumption on the node labeling is made purely for the sake of notational convenience in our analysis (to avoid clutter in having to keep track of additional permutations), and our algorithmic methodology does not rely on such a labeling. Based on this characterization of the optimality conditions, the success of the Schur–Horn relaxation relies on the existence of a suitable dual variable $M \in \mathbb{S}^n$ that satisfies two conditions. The first of these conditions relates to the structure of the noise edges in $\mathfrak{G}$, while the second condition relates to the structure of the planted graph Γ via the normal cone $\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})$. From the viewpoint of Lemma 1, favorable problem instances for the Schur–Horn relaxation are, informally speaking, those in which there are not too many noise edges in $\mathfrak{G}$ (implying a less restrictive first requirement on M) and in which the normal cone $\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})$ is large (entailing a more flexible second condition for M). The interplay between these two conditions forms the basis of our analysis and results presented in section 3. In the remainder of the present section, we investigate spectral properties of planted graphs Γ that result in a large normal cone $\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})$.

The normal cones at the extreme points of the Schur–Horn orbitope are conveniently described based on the following notion (see Proposition 6 in the following).

DEFINITION 2. *A pair of symmetric matrices* $A, B \in \mathbb{S}^n$ *is* spectrally comonotone *if there exists an orthogonal matrix* $U \in \mathbb{R}^{n \times n}$ *such that* $U'AU$ *and* $U'BU$ *are both diagonal matrices with the diagonal entries sorted in nonincreasing order.*

The stipulation that two matrices be spectrally comonotone is a stronger condition than the requirement that the matrices be simultaneously diagonalizable, due to the additional restriction on the ordering of the diagonal entries upon conjugation by an orthogonal matrix.

Example 3. Consider the matrices

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The matrices A and B are spectrally comonotone, while A and C are only simultaneously diagonalizable and are not spectrally comonotone.

As Proposition 1 states the optimality conditions of the Schur–Horn relaxation in terms of the *relative interiors* of normal cones at extreme points of the Schur–Horn orbitope, we need the following "strict" analog of spectral comonotonicity.

DEFINITION 4. *A matrix* $A \in \mathbb{S}^n$ *is* strictly spectrally comonotone *with a matrix* $B \in \mathbb{S}^n$*, if for every* $P \in \mathbb{S}^n$ *that is simultaneously diagonalizable with* B*, there exists* $\epsilon > 0$ *such that* $A + \epsilon P$ *and* B *are spectrally comonotone.*

Strict spectral comonotonicity is more restrictive than spectral comonotonicity. Further, the definition of strict spectral comonotonocity is not a symmetric one, unlike that of spectral comonotonicity, i.e., even if $A \in \mathbb{S}^n$ is strictly spectrally comonotone with $B \in \mathbb{S}^n$, it may be that B is not strictly spectrally comonotone with A.

Example 5. Consider the matrices

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix A is strictly spectrally comonotone with the matrix B, but B is not strictly spectrally comonotone with A.

The following result provides a characterization of normal cones at extreme points of the Schur–Horn orbitope in terms of spectrally comonotone matrices.

PROPOSITION 6. *For any matrix $M \in \mathbb{S}^n$ and the associated Schur–Horn orbitope $\mathcal{SH}(M)$, the normal cone $\mathcal{N}_{\mathcal{SH}(M)}(W)$ and its relative interior at an extreme point W of $\mathcal{SH}(M)$ are given by*

$$\mathcal{N}_{\mathcal{SH}(M)}(W) = \{Q \in \mathbb{S}^n \mid Q \text{ and } W \text{ are spectrally comonotone}\},$$
$$\text{relint}\left(\mathcal{N}_{\mathcal{SH}(M)}(W)\right) = \{Q \in \mathbb{S}^n \mid Q \text{ is strictly spectrally comonotone with } W\}.$$

Note. For any matrix $M \in \mathbb{S}^n$, the extreme points of $\mathcal{SH}(M)$ are the elements of the set $\{UMU' \mid U \in \mathcal{O}_n\}$, as each of the matrices UMU' for $U \in \mathcal{O}_n$ has the same Frobenius norm.

Proof. Let $W = M$ without loss of generality. We have that

$$\mathcal{N}_{\mathcal{SH}(M)}(M) = \left\{ Y \in \mathbb{S}^n \mid \sup_{Z \in \mathcal{SH}(M)} \text{trace}(YZ) \leq \text{trace}(YM) \right\}$$

$$= \left\{ Y \in \mathbb{S}^n \mid \sup_{Z = UMU' \text{ for } U \in \mathcal{O}_n} \text{trace}(YZ) \leq \text{trace}(YM) \right\}$$

$$= \left\{ Y \in \mathbb{S}^n \mid \sup_{U \in \mathcal{O}_n} \text{trace}(U'YUM) = \text{trace}(YM) \right\}.$$

The last line follows from the inequality $\text{trace}(YM) \leq \sup_{U \in \mathcal{O}_n} \text{trace}(U'YUM)$. Considering the case of equality in the Von Neumann trace inequality [43], we have that $\sup_{U \in \mathcal{O}_n} \text{trace}(U'YUM) = \text{trace}(YM)$ if and only if Y and M are spectrally comonotone. The claim about the relative interior of the normal cone follows immediately from the definition of strict spectral comonotonicity. □

If a matrix $M \in \mathbb{S}^n$ has few distinct eigenvalues, the normal cone at an extreme point UMU' (for U orthogonal) of $\mathcal{SH}(M)$ is larger as there are many more matrices that are spectrally comonotone with UMU'. Based on Proposition 6, this observation suggests that planted graphs Γ with few distinct eigenvalues have large normal cones $\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})$, and such graphs are especially amenable to recovery in planted subgraph problems via the Schur–Horn relaxation. We make this insight more precise with our analysis in section 3.4. Proposition 6 also points to the utility of employing the parameter γ in the Schur–Horn relaxation (P). Specifically, multiplicities in the spectrum of the matrix $[A_\Gamma - \gamma I_k]_{k \to n} \in \mathbb{S}^n$ may be increased via suitable choices of γ, which in turn makes the normal cone $\mathcal{N}_{\mathcal{SH}([A_\Gamma - \gamma I_k]_{k \to n})}([A_\Gamma - \gamma I_k]_{k \to n})$

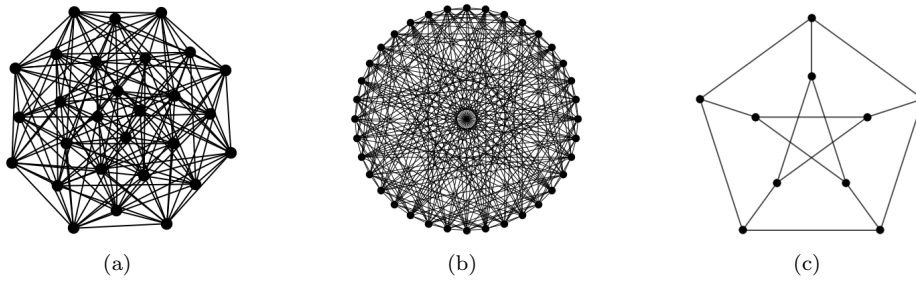FIG. 2. *From left to right:* 8-*triangular graph,* 9-*triangular graph, and Petersen graph.*

larger. In particular, setting γ equal to an eigenvalue of A_Γ increases the multiplicity of zero as an eigenvalue of $[A_\Gamma - \gamma I_k]_{k\to n}$. As detailed in section 3, the success of the Schur–Horn relaxation relies on the existence of an eigenspace $\mathcal{E} \subset \mathbb{R}^k$ of A_Γ with a small coherence parameter, and the appropriate choice of γ is the eigenvalue $\lambda_\mathcal{E}$ associated with $\mathcal{E}$. In our experiments in section 4, we set γ equal to the eigenvalue of A_Γ with largest multiplicity, so that the multiplicity of zero as an eigenvalue of $[A_\Gamma - \gamma I_k]_{k\to n}$ is as large as possible.

To conclude, we record an observation on spectral comonotonocity that is useful in section 3. The claim is straightforward and therefore we omit the proof.

LEMMA 7. *A pair of symmetric matrices* $A, B \in \mathbb{S}^n$ *is spectrally comonotone if and only if* A *and* B *are simultaneously diagonalizable and*

$$(4) \qquad \lambda_{\min}(A|_{\mathcal{E}_i}) \geq \lambda_{\max}(A|_{\mathcal{E}_{i+1}}) \quad \forall i \in \{1, \ldots, t-1\},$$

where $\mathcal{E}_i$ *for* $i \in \{1, \ldots, t\}$ *are eigenspaces of* B *ordered such that the corresponding eigenvalues of* B *are decreasing. Further,* A *is strictly spectrally comonotone with* B *if and only if* A *and* B *are simultaneously diagonalizable and each of the inequalities* (4) *holds strictly.*

Note that if A and B are simultaneously diagonalizable, then any eigenspace $\mathcal{E}$ of B is an invariant subspace of A. As a result, the restriction of A to the eigenspaces of B in (4) is consistent with the notation described in section 1.2.

2.2. Graphs with few eigenvalues. Building on the preceding section, we give examples of families of graphs consisting of few distinct eigenvalues. Such graphs have received much attention due to their connections to topics in combinatorics and design theory such as pseudorandomness [27] and association schemes [5, 22].

Triangular graphs. The triangular graph $\mathcal{T}_m$ of order m is the line graph of the complete graph on m nodes. The graph $\mathcal{T}_m$ has $\binom{m}{2}$ nodes and it has the three distinct eigenvalues $2(m-2)$ (with multiplicity 1), $m-4$ (with multiplicity $m-1$), and -2 (with multiplicity $\frac{m(m-3)}{2}$). Figure 2 gives two examples.

Kneser graphs. A Kneser graph $\mathcal{K}(m, \ell)$ is a graph on $\binom{m}{\ell}$ nodes, each corresponding to an ℓ-element subset of m elements, and it consists of edges between those pairs of vertices for which the corresponding subsets are disjoint. The graph $\mathcal{K}(m, 1)$ is the complete graph on m nodes and the graph $\mathcal{K}(5, 2)$ is the Petersen graph (Figure 2). The Kneser graph $\mathcal{K}(m, \ell)$ has $\ell + 1$ distinct eigenvalues in general.

Paley graphs. Let q be a prime power such that $q \equiv 1 (\mathrm{mod}\ 4)$. The Paley graph on q nodes is an undirected graph formed by connecting pairs of nodes $i, j \subset \{0, \ldots, q-1\}$
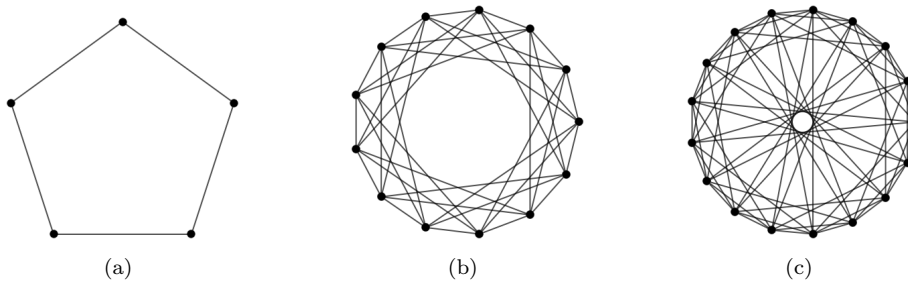
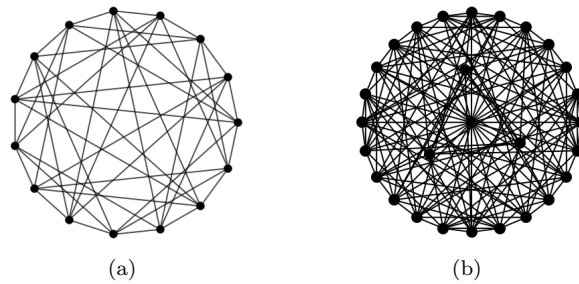FIG. 3. *From left to right:* 5-*Paley graph,* 13-*Paley graph, and* 17-*Paley graph.*

FIG. 4. *Generalized quadrangle-*(2, 2) *graph (left) and generalized quadrangle-*(2, 4) *graph (right).*

if the difference $i - j$ is a square in the finite field $\mathrm{GF}(q)$. Note that $i - j$ is a square if and only if $j - i$ is a square as -1 is a square in $\mathrm{GF}(q)$. Paley graphs have eigenvalues $\frac{1}{2}(q-1)$ (with multiplicity 1), $\frac{1}{2}(-1+\sqrt{q})$ (with multiplicity $\frac{1}{2}(q-1)$), and $\frac{1}{2}(-1-\sqrt{q})$ (with multiplicity $\frac{1}{2}(q-1)$). Paley graphs are also examples of pseudorandom graphs as they exhibit properties similar to random graphs (in the limit of large q) [27]. Figure 3 shows the three smallest Paley graphs.

Strongly regular graphs. These are regular graphs with the property that every pair of adjacent vertices has the same number d_a of common neighbors and every pair of nonadjacent vertices has the same number d_{na} of common neighbors for some integers d_a, d_{na} [7]. Strongly regular graphs that are connected have three distinct eigenvalues; conversely, connected and regular graphs with three distinct eigenvalues are necessarily strongly regular. The triangular graphs, Kneser graphs with parameter $\ell = 2$, and the Paley graphs mentioned above are examples of strongly regular graphs. The Clebsch graph shown in Figure 1(a) in the introduction is also a strongly regular graph with degree 5 and eigenvalues 5 (with multiplicity 1), -3 (with multiplicity 5), and 1 (with multiplicity 10). The generalized quadrangle graphs shown in Figure 4 are additional examples of strongly regular graphs. Strongly regular graphs form a significant topic in graph theory due to their many regularity properties [10, 11, 37].

Other examples. Unlike regular graphs with three distinct eigenvalues, graphs with four (or more) eigenvalues do not appear to have a simple combinatorial characterization [40]. Nonetheless, there are many constructions of such graphs in the literature [25, 39, 40], most notably those derived from distance-regular graphs [9] and from association schemes. The two graphs shown in Figure 5 are obtained from the Hamming scheme.
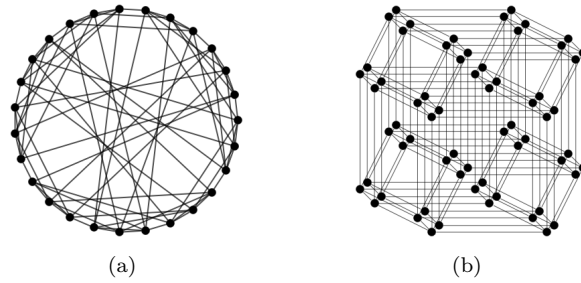
(a)   (b)

FIG. 5. *Hamming-$(3,3)$ graph with 4 distinct eigenvalues (left) and 6-hypercube graph with 7 distinct eigenvalues (right).*

3. Recovering subgraphs planted in Erdős–Rényi random graphs. In this section we discuss our theoretical results on the performance of the Schur–Horn relaxation in recovering subgraphs planted inside Erdős–Rényi random graphs. Formally, suppose without loss of generality as in the previous section that the nodes of $\mathfrak{G}$ and of Γ are labeled so that the leading principal minor of $A_{\mathfrak{G}}$ of order k is equal to A_Γ. The Erdős–Rényi model for the planted subgraph problem specifies a distribution on the edges in the remainder of the graph $\mathfrak{G}$ via a probability parameter $p \in [0, 1]$; for each $i, j \in \{1, \dots, n\}$ with $i < j$ and $k < j$, the graph $\mathfrak{G}$ contains an edge between nodes i and j with probability p (independent of the other edges):

$$(A_{\mathfrak{G}})_{i,j} = (A_{\mathfrak{G}})_{j,i} = \begin{cases} 1 \text{ with probability } p, \\ 0 \text{ with probability } 1 - p. \end{cases}$$

We begin with a sufficient condition for the optimality condition described in Lemma 1, which suggests a natural approach for constructing suitable dual variables for certifying optimality. These sufficient conditions point to the importance of the existence of an eigenspace of A_Γ with certain properties to the success of the Schur–Horn relaxation; these properties are discussed in section 3.2. In section 3.4 we state and prove the main theorem (Theorem 19) of this paper, with section 3.5 giving specializations of this result (e.g., to the planted clique problem).

3.1. A simpler sufficient condition for optimality. The following proposition provides a simpler set of conditions than those in Lemma 1 on dual variables that certify the success of the Schur–Horn relaxation. This result continues to be deterministic in nature, and the probabilistic aspects of our analysis—due to the Erdős–Rényi model—appear in the following.

PROPOSITION 8. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_{\mathfrak{G}}$ of order k is equal to A_Γ. Suppose there exists an eigenspace $\mathcal{E} \subset \mathbb{R}^k$ of A_Γ with eigenvalue $\lambda_{\mathcal{E}}$, and suppose there exists a matrix $M = \begin{pmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{pmatrix} \in \mathbb{S}^n$ with submatrices $M_{11} \in \mathbb{S}^k, M_{12} \in \mathbb{R}^{k \times (n-k)}, M_{22} \in \mathbb{S}^{n-k}$ such that the the following conditions are satisfied:*
  (i) *$M_{i,j} = (A_{\mathfrak{G}})_{i,j}$, if $(A_{\mathfrak{G}})_{i,j} = 1$ or if $i = j$.*
  (ii) *The submatrix $M_{11} \in \mathbb{S}^k$ is strictly spectrally comonotone with A_Γ.*
  (iii) *$\lambda_{\max}(M_{11}|_{\mathcal{E}}) \geq \lambda_{\mathcal{E}}$ and $\lambda_{\min}(M_{11}|_{\mathcal{E}}) \leq \lambda_{\mathcal{E}}$.*
  (iv) *Each column of the submatrix $M_{12} \in \mathbb{R}^{k \times (n-k)}$ lies in the subspace $\mathcal{E}$.*
  (v) *$\mathrm{eigengap}(M_{11}, \mathcal{E}) > \|M_{12}\|_2 + \|M_{22}\|_2 + |\lambda_{\mathcal{E}}|$.*

Then the Schur–Horn relaxation (P) *with parameter* $\gamma = \lambda_{\mathcal{E}}$ *succeeds at identifying the planted subgraph* Γ *inside the larger graph* $\mathfrak{G}$.

Proof. We establish this result by showing that the given matrix $M \in \mathbb{S}^n$ satisfies the requirements of Lemma 1. The first condition of Lemma 1 is identical to that of this proposition, and therefore it is satisfied. We prove next that the remaining conditions of this proposition ensure that the second requirement of Lemma 1 is also satisfied, i.e., $M \in \text{relint}(\mathcal{N}_{\mathcal{SH}([A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n})}([A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}))$. Based on Proposition 6, this entails showing that M is strictly spectrally comonotone with $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$. Our strategy is to employ Lemma 7.

Let $\mathcal{E}_i \subset \mathbb{R}^k$, $i = 1, \ldots, t$, be the eigenspaces of A_Γ ordered such that the corresponding eigenvalues $\lambda_{\mathcal{E}_i}$ are strictly decreasing, and suppose $\mathcal{E}_j = \mathcal{E}$, $\lambda_{\mathcal{E}_j} = \lambda_{\mathcal{E}}$ for some $j \in \{1, \ldots, t\}$. As 0 is an eigenvalue of $A_\Gamma - \lambda_{\mathcal{E}} I_k$, one can check that the eigenspaces of $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$ are $\tilde{\mathcal{E}}_i = \mathcal{E}_i \times \{0\} \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$, $i = 1, \ldots, t$, $i \neq j$ (with corresponding eigenvalues $\lambda_{\mathcal{E}_i} - \lambda_{\mathcal{E}}$), and $\tilde{\mathcal{E}}_j = \mathcal{E} \times \mathbb{R}^{n-k} \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ (with eigenvalue 0). We now need to show that M and $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$ are simultaneously diagonalizable, and that $\lambda_{\min}(M|_{\tilde{\mathcal{E}}_i}) > \lambda_{\max}(M|_{\tilde{\mathcal{E}}_{i+1}})$ for $i \in \{1, \ldots, t-1\}$.

First, as $\mathcal{E}$ is an eigenspace of $A_\Gamma - \lambda_{\mathcal{E}} I_k$ with eigenvalue 0 and as every column of M_{12} belongs to $\mathcal{E}$, one can check that $(A_\Gamma - \lambda_{\mathcal{E}} I_k) \cdot M_{12} = 0 \in \mathbb{R}^{k \times (n-k)}$. Further, from Lemma 7 we note that M_{11} and $A_\Gamma - \lambda_{\mathcal{E}} I_k$ are simultaneously diagonalizable because M_{11} is strictly spectrally comonotone with A_Γ (and hence with $A_\Gamma - \lambda_{\mathcal{E}} I_k$). From these two observations one can check that M and $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$ commute with each other, and therefore are simultaneously diagonalizable.

As M and $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$ are simultaneously diagonalizable, we have that the eigenspaces $\tilde{\mathcal{E}}_i$, $i = 1, \ldots, t$, of $[A_\Gamma - \lambda_{\mathcal{E}} I_k]_{k \to n}$ are invariant subspaces of M. Similarly, as M_{11} is strictly spectrally comonotone with A_Γ, the eigenspaces $\mathcal{E}_i$ are invariant subspaces of M_{11}. Based on the structure of these eigenspaces as described above, one can check that the eigenvalues of $M|_{\tilde{\mathcal{E}}_i}$ are equal to those of $M_{11}|_{\mathcal{E}_i}$ for each $i = 1, \ldots, t$, $i \neq j$. Hence, $\lambda_{\min}(M|_{\tilde{\mathcal{E}}_i}) > \lambda_{\max}(M|_{\tilde{\mathcal{E}}_{i+1}})$ for $i > j$ and for $i < j - 1$.

All that remains to be verified is that $\lambda_{\min}(M|_{\tilde{\mathcal{E}}_j}) > \lambda_{\max}(M|_{\tilde{\mathcal{E}}_{j+1}})$ and that $\lambda_{\min}(M|_{\tilde{\mathcal{E}}_{j-1}}) > \lambda_{\max}(M|_{\tilde{\mathcal{E}}_j})$. As each column of M_{12} belongs to $\mathcal{E}$ and as $\tilde{\mathcal{E}}_j = \mathcal{E} \times \mathbb{R}^{n-k} \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$, we have for $x \in \mathcal{E}, y \in \mathbb{R}^{n-k}$ that

$$(5) \quad M|_{\tilde{\mathcal{E}}_j} \begin{pmatrix} x \\ y \end{pmatrix} = \left[\begin{pmatrix} M_{11}|_{\mathcal{E}} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & M_{12} \\ M_{12}' & M_{22} \end{pmatrix} \right] \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} M_{11}|_{\mathcal{E}} x + M_{12} y \\ M_{12}' x + M_{22} y \end{pmatrix} \in \tilde{\mathcal{E}}_j.$$

Consequently, recalling that $\mathcal{E}_j = \mathcal{E}$ we have

$$\lambda_{\max}(M|_{\tilde{\mathcal{E}}_j}) \leq \max\{\lambda_{\max}(M_{11}|_{\mathcal{E}}), 0\} + \|M_{12}\|_2 + \|M_{22}\|_2$$
$$< \max\{\lambda_{\max}(M_{11}|_{\mathcal{E}}), 0\} - |\lambda_{\mathcal{E}}| + \text{eigengap}(M_{11}, \mathcal{E})$$
$$\leq \max\{\lambda_{\max}(M_{11}|_{\mathcal{E}}), 0\} - |\lambda_{\mathcal{E}}| + \lambda_{\min}(M_{11}|_{\mathcal{E}_{j-1}}) - \lambda_{\max}(M_{11}|_{\mathcal{E}})$$
$$= \max\{0, -\lambda_{\max}(M_{11}|_{\mathcal{E}})\} - |\lambda_{\mathcal{E}}| + \lambda_{\min}(M_{11}|_{\mathcal{E}_{j-1}})$$
$$\leq \max\{0, -\lambda_{\mathcal{E}}\} - |\lambda_{\mathcal{E}}| + \lambda_{\min}(M_{11}|_{\mathcal{E}_{j-1}})$$
$$\leq \lambda_{\min}(M_{11}|_{\mathcal{E}_{j-1}})$$
$$= \lambda_{\min}(M|_{\tilde{\mathcal{E}}_{j-1}}).$$

The first inequality follows from (5), the second inequality from condition (v), the third inequality from the definition of eigengap (see section 1.2) as $\mathcal{E}_j = \mathcal{E}$, the fourth

inequality from condition (iii), and the second equality from the fact that the eigenvalues of $M|_{\tilde{\mathcal{E}}_i}$ are equal to those of $M_{11}|_{\mathcal{E}_i}$ for each $i = 1, \ldots, t$, $i \neq j$. Similarly, one can check that $\lambda_{\min}(M|_{\tilde{\mathcal{E}}_j}) > \lambda_{\max}(M|_{\tilde{\mathcal{E}}_{j+1}})$. This concludes the proof. $\qquad\square$

This result provides a concrete approach for constructing dual variables to certify the optimality of the Schur–Horn relaxation (P) at the desired solution. In the remainder of this section, we give conditions on the eigenstructure of the planted graph Γ, the probability p of the Erdős–Rényi model, and the size n of the larger graph $\mathfrak{G}$ under which the Schur–Horn relaxation (P) succeeds with high probability.

3.2. Invariants of graph eigenspaces. In this section, we investigate properties of eigenspaces of graphs which ensure that the conditions of Proposition 8 can be satisfied. For notational clarity in the discussion in this section, we let $\Omega_j \subset \{1, \ldots, k\}$ for $j = 1, \ldots, n-k$ denote the locations of the entries equal to one in the submatrix $(A_{\mathfrak{G}})_{i,j+k}$, $i = 1, \ldots, k, j = 1, \ldots, n-k$, i.e., $(A_{\mathfrak{G}})_{i,j+k} = 1 \Leftrightarrow i \in \Omega_j$.

A requirement of Proposition 8 is the existence of a suitable eigenspace $\mathcal{E} \subset \mathbb{R}^k$ of A_Γ such that one can obtain a matrix $M_{12} \in \mathbb{R}^{k \times (n-k)}$ (a submatrix of a larger dual certificate) that satisfies three conditions: (i) every column of M_{12} lies in $\mathcal{E}$; (ii) for each $i = 1, \ldots, k$ and $j = 1, \ldots, n-k$ we have that $(M_{12})_{i,j} = 1$ if $(A_{\mathfrak{G}})_{i,j+k} = 1$; and (iii) the operator norm $\|M_{12}\|_2$ is as small as possible.

We begin by analyzing the first two conditions and the restrictions they impose on $\mathcal{E}$. Consider the jth column of M_{12} for a fixed $j \in \{1, \ldots, n-k\}$ as an illustration. Then conditions (i) and (ii) are simultaneously satisfied if the coordinate subspace of vectors in $\mathbb{R}^k$ with support on the indices in Ω_j has a *transverse intersection* with $\mathcal{E}^\perp$. More generally, a natural sufficient condition for the first two requirements on M_{12} to be satisfied (for every column) is for $\mathcal{E}^\perp$ to have a transverse intersection with the coordinate subspaces specified by each of the subsets Ω_j for $j = 1, \ldots, k$. This observation leads to the following invariant that characterizes the transversality of a subspace with all coordinate subspaces of a certain dimension.

DEFINITION 9 (see [28]). *The* Kruskal rank *of a subspace* $\mathcal{S} \subseteq \mathbb{R}^k$, *denoted* kruskal$(\mathcal{S})$, *is the largest* $m \in \mathbb{Z}$ *such that for any* $\Omega \subseteq \{1, \ldots, k\}$ *with* $|\Omega| = m$ *we have*

$$\mathcal{S}^\perp \cap \{v \in \mathbb{R}^k \mid v_i = 0 \text{ if } i \notin \Omega\} = \{0\}.$$

In other words, the Kruskal rank of a subspace $\mathcal{S} \subset \mathbb{R}^k$ is one less than the size of the support of the sparsest nonzero vector in $\mathbb{R}^k$ that is orthogonal to $\mathcal{S}$. The Kruskal rank of a matrix—the largest m such that all subsets of m columns of the matrix are linearly independent—was first introduced in [28] in the context of tensor decompositions. This version in terms of matrices is equivalent to our definition in terms of subspaces. One can check that all principal minors of $\mathcal{P}_{\mathcal{S}}$ of size up to kruskal$(\mathcal{S})$ are nonsingular.

Recall that the entries $(A_{\mathfrak{G}})_{i,j+k}$ for $i = 1, \ldots, k$ and $j = 1, \ldots, n-k$ correspond to edges (or lack thereof) between nodes in $\mathfrak{G}$ outside the induced subgraph corresponding to Γ and those of Γ. Therefore, if we employ the Schur–Horn relaxation with parameter $\gamma = \lambda_{\mathcal{E}}$ (the eigenvalue associated to $\mathcal{E}$), then the Kruskal rank of $\mathcal{E}$ provides a bound on the number of noise edges that can be tolerated between these two sets of nodes. As such kruskal$(\mathcal{E})$ plays a central role in our main result (see Theorem 19) in providing an upper bound on the probability of a noise edge in $\mathfrak{G}$ under the Erdős–Rényi model.

Returning to the three conditions on M_{12} stated at the beginning of this section, if an eigenspace $\mathcal{E}$ of A_Γ has large Kruskal rank and if the size of each Ω_j, $j = 1, \ldots, n-k$,

is smaller than kruskal($\mathcal{E}$), then there is an affine space (of dimension potentially larger than zero) of matrices in $\mathbb{R}^{k \times (n-k)}$ that satisfy the first two requirements on M_{12}. The third condition on M_{12} requires that we find the element of this affine space with the smallest spectral norm:

$$\hat{M}_{12}^{\text{spectral}} = \underset{X \in \mathbb{R}^{k \times (n-k)}}{\arg\min} \quad \|X\|_2$$
$$\text{s.t. } X_{i,j} = 1 \text{ if } i \in \Omega_j \text{ for } j = 1, \ldots, n-k,$$
$$\text{col}(X) \subseteq \mathcal{E}.$$

As long as $|\Omega_j| \leq \text{kruskal}(\mathcal{E})$ for each $j = 1, \ldots, n-k$, this problem is feasible. However, analytically characterizing the optimal value and solution of this problem is challenging, especially in the context of problem instances that arise from the Erdős–Rényi model, as the subsets Ω_j, $j = 1, \ldots, n-k$, are random. As a result, a common approach is to replace the objective in the above problem with the Frobenius norm:

(6)
$$\hat{M}_{12}^{\text{Frobenius}} = \underset{X \in \mathbb{R}^{k \times (n-k)}}{\arg\min} \quad \|X\|_F$$
$$\text{s.t. } X_{i,j} = 1 \text{ if } i \in \Omega_j \text{ for } j = 1, \ldots, n-k,$$
$$\text{col}(X) \subseteq \mathcal{E}.$$

One of the virtues of this latter formulation in comparison to the earlier one is that the spectral norm of the optimal solution $\|\hat{M}_{12}^{\text{Frobenius}}\|_2$ is more tractable to bound, primarily since the optimization problem (6) decomposes into $n-k$ separable problems, one for each column of the decision variable X. In particular, for any subspace $\mathcal{S} \subseteq \mathbb{R}^k$ and any $\Omega \subset \{1, \ldots, k\}$ with $|\Omega| \leq \text{kruskal}(\mathcal{S})$, consider the following minimum Euclidean-norm completion,

(7)
$$q_\Omega(\mathcal{S}) \triangleq \underset{q \in \mathbb{R}^k}{\arg\min} \quad \|q\| \quad \text{s.t. } q \in \mathcal{S} \text{ and } q_i = 1 \text{ for } i \in \Omega$$
$$= \mathcal{P}_\mathcal{S} I_\Omega{}'((\mathcal{P}_\mathcal{S})_{\Omega,\Omega})^{-1} 1_{|\Omega|}.$$

With this notation, the jth column of $\hat{M}_{12}^{\text{Frobenius}}$ is given by $q_{\Omega_j}(\mathcal{E})$. Further, under the Erdős–Rényi model, the entries $(A_\mathfrak{G})_{i,j+k}$, $i = 1, \ldots, k, j = 1, \ldots, n-k$, are independent and identically distributed Bernoulli random variables. In such a family of problem instances, the columns of $\hat{M}_{12}^{\text{Frobenius}}$, i.e., $q_{\Omega_j}(\mathcal{E}) \in \mathbb{R}^k$, $j = 1, \ldots, k$, are independent and identically distributed random vectors. These observations in conjunction with the following tail bound on the spectral norm of a random matrix suggest a natural invariant of $\mathcal{E}$ that leads to bounds on $\|\hat{M}_{12}^{\text{Frobenius}}\|_2$:

LEMMA 10 (see [42]). *Let A be a $d \times N$ matrix $(d < N)$ with columns A_i and let $\Sigma = \mathbb{E}[A_i A_i{}^T]$ denote the correlation matrix of the A_i's. Further, suppose there exists $m \in \mathbb{R}$ such that $\|A_i\| \leq \sqrt{m}$ almost surely for all i. Then $\forall x \geq (N\|\Sigma\|_2)^{1/2}$ we have that*

(8)
$$\mathbb{P}(\|A\|_2 \geq x) \leq 2d \exp\left(-\frac{3(x^2 - N\|\Sigma\|_2)^2}{4m(x^2 + 2N\|\Sigma\|_2)}\right).$$

Proof. The proof follows that of [42, Theorem 5.41] with minor modifications. We apply the noncommutative Bernstein inequality to $\frac{1}{N}x^2 - \|\Sigma\|_2$ rather than to $\max(\delta, \delta^2)$ on [42, p. 27], and we don't make the isotropy assumption. □

To apply Lemma 10 to obtain a bound on $\|\hat{M}_{12}^{\mathrm{Frobenius}}\|_2$, we describe next the second key invariant of $\mathcal{E}$, which is essentially the correlation matrix in Lemma 10.

DEFINITION 11. *Let $\mathcal{S} \subseteq \mathbb{R}^k$ be a subspace. Then the* combinatorial width *of $\mathcal{S}$ for each $\ell = 1, \dots, \mathrm{kruskal}(\mathcal{S})$ and $p \in [0, 1)$ is defined as*

$$\omega(\mathcal{S}, \ell, p) \triangleq \left\| \mathbb{E}[q_{\Omega}(\mathcal{S}) \, q_{\Omega}(\mathcal{S})' \,|\, |\Omega| \leq \ell] \right\|_2$$

with the expectation taken over Ω, where each element of $\{1, \dots, k\}$ is contained in Ω independently with probability p.

The conditioning in the definition ensures that $q_{\Omega}(\mathcal{S})$ is well-defined as $|\Omega| \leq \mathrm{kruskal}(\mathcal{S})$. We utilize this terminology as a parallel to analogous notions such as "mean width" that are prominent in the convex geometry literature. The explicit appearance of ℓ in this definition allows for a more fine-grained analysis in our main result Theorem 19; see section 3.4. Based on the following result, the Kruskal rank and the combinatorial width play a central role in Theorem 19 as the success of the Schur–Horn relaxation (P) relies on the existence of an eigenspace $\mathcal{E}$ of A_{Γ} that has large Kruskal rank and small combinatorial width.

PROPOSITION 12. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_{\mathfrak{G}}$ of order k is equal to A_{Γ}, and the remaining edges in $\mathfrak{G}$ are drawn according to the Erdős–Rényi model with probability $p \in [0, \frac{\mathrm{kruskal}(\mathcal{E})}{k})$. Fix any $\ell \in \mathbb{Z}$ satisfying $kp < \ell \leq \mathrm{kruskal}(\mathcal{E})$, and denote $\zeta := \min_{\substack{\Omega \subset \{1,\dots,k\} \\ |\Omega| \leq \ell}} \lambda_{\min}((\mathcal{P}_{\mathcal{E}})_{\Omega,\Omega})$. For any $\delta \geq \sqrt{(n-k)\,\omega(\mathcal{E},\ell,p)}$, there exists a matrix $M_{12} \in \mathbb{R}^{k \times (n-k)}$ satisfying the following properties:*
  1. *Each column of M_{12} lies in $\mathcal{E}$,*
  2. *$(M_{12})_{i,j} = (A_{\mathfrak{G}})_{i,\,j+k}$ if $(A_{\mathfrak{G}})_{i,\,j+k} = 1$,*
  3. *$\|M_{12}\|_2 < \delta$*
with probability at least $(1 - 2k \exp\left(-\frac{3\zeta(\delta^2 - (n-k)\,\omega(\mathcal{E},\ell,p))^2}{4\,\ell(\delta^2 + 2(n-k)\omega(\mathcal{E},\ell,p))}\right))(1 - \exp\left(-\frac{(\ell-kp)^2}{\ell+kp}\right))^{n-k}$.

Proof. We bound the probability that $\hat{M}_{12}^{\mathrm{Frobenius}}$ obtained as the optimal solution of (6) satisfies the requirements of this proposition.

We begin by bounding the cardinality of each Ω_j for $j = 1, \dots, n - k$. Under the Erdős–Rényi model, each $|\Omega_j|$ follows a binomial distribution. Consequently, using the Chernoff bound we have for each $j = 1, \dots, n - k$ that

$$\mathbb{P}(|\Omega_j| \geq \ell + 1) \leq \mathbb{P}(|\Omega_j| \geq \ell) = \mathbb{P}\left(|\Omega_j| \geq \left(1 + \frac{\ell - kp}{kp}\right) kp\right) \leq \exp\left(-\frac{(\ell - kp)^2}{\ell + kp}\right).$$

The first inequality is not essential and it is simply used to avoid notational clutter. Based on the independence of the Ω_j's,

$$(9) \qquad \mathbb{P}(|\Omega_j| \leq \ell, \ j = 1, \dots, n - k) \geq \left(1 - \exp\left(-\frac{(\ell - kp)^2}{\ell + kp}\right)\right)^{n-k}.$$

This inequality provides a bound on the probability that the optimization problem (6) is feasible.

In our next step we bound $\|\hat{M}_{12}^{\mathrm{Frobenius}}\|_2$ via Lemma 10. As $\ell \leq \mathrm{kruskal}(\mathcal{E})$ one can check that $\zeta > 0$. Further, from (7) we have that $\|q_{\Omega_j}\|^2 \leq \frac{|\Omega_j|}{\zeta}$. Thus, by

applying Lemma 10, we deduce that

$$(10) \quad \mathbb{P}(\|M_{12}\|_2 < \delta \mid |\Omega_j| \leq \ell \; \forall j) \geq 1 - 2k \exp\left(-\frac{3\zeta(\delta^2 - (n-k)\omega(\mathcal{E}, \ell, p))^2}{4\,\ell(\delta^2 + 2(n-k)\omega(\mathcal{E}, \ell, p))}\right).$$

The final result follows by combining (9) and (10). □

3.3. Properties of Kruskal rank and combinatorial width. Beyond the utility of the Kruskal rank and combinatorial width in characterizing the performance of the Schur–Horn relaxation, these graph parameters are also of intrinsic interest and we discuss next their relationship to structural properties of Γ.

3.3.1. Invariance under complements for regular graphs. Both the Kruskal rank and the combinatorial width are preserved under graph complements for connected regular graphs. Suppose Γ is a connected regular graph on k vertices, and let $A_\Gamma \in \mathbb{S}^k$ be an adjacency matrix representing Γ for some labeling of the nodes. Then the eigenspaces of A_Γ are the same as those of the adjacency matrix A_{Γ^c} of the complement Γ^c based on the following relation,

$$(11) \qquad\qquad A_{\Gamma^c} = 1_k 1_k' - I_k - A_\Gamma.$$

As Γ is connected and regular, the vector 1_k is an eigenvector of A_Γ. Thus, the Kruskal ranks and the combinatorial widths associated with the eigenspaces of A_Γ are the same as those associated with the eigenspaces of A_{Γ^c}.

3.3.2. Combinatorial width for symmetric graphs. For graphs Γ that are symmetric—vertex- and edge-transitive—and also have symmetric complements Γ^c, the combinatorial width of any eigenspace $\mathcal{E}$ of A_Γ can be characterized in terms of the minimum singular values of minors of $\mathcal{P}_\mathcal{E}$. In particular, we establish our result by demonstrating that the correlation matrix $\mathbb{E}[q_\Omega(\mathcal{S})\, q_\Omega(\mathcal{S})' \mid |\Omega| \leq \ell]$ in the definition of the combinatorial width has the property that all its nonzero eigenvalues are equal to each other, which leads to bounds on the combinatorial width via bounds on the trace of the correlation matrix.

PROPOSITION 13. *Let $A_\Gamma \in \mathbb{S}^k$ be an adjacency matrix of a (connected) symmetric graph Γ with a symmetric complement Γ^c, and let $\mathcal{E} \subset \mathbb{R}^k$ be an eigenspace of A_Γ. Fix any $\ell \in \mathbb{Z}$ and $p \in [0,1)$ such that $kp \leq \ell \leq \mathrm{kruskal}(\mathcal{E})$, and let $\zeta := \min_{\substack{\Omega \subset \{1,\dots,k\} \\ |\Omega| \leq \ell}} \lambda_{\min}((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})$. Then,*

$$\omega(\mathcal{E}, \ell, p) \;\leq\; \frac{2kp}{\zeta \dim(\mathcal{E})}.$$

Proof. Denote the correlation matrix in the definition of the combinatorial width as follows:

$$(12) \qquad \Sigma = \mathbb{E}\left[q_\Omega(\mathcal{E}) q_\Omega(\mathcal{E})' \mid |\Omega| \leq \ell\right] = \sum_{i=0}^{\ell} c_{p,\ell}\, p^i (1-p)^{k-i} \sum_{|\Omega|=i} q_\Omega(\mathcal{E}) q_\Omega(\mathcal{E})',$$

where the term $c_{p,\ell} = (\sum_{i=0}^{\ell} \binom{k}{i} p^i (1-p)^{k-i})^{-1}$ is the normalization constant.

The main element of the proof is to show that the rank of $\Sigma \in \mathbb{S}^k$ is equal to $\dim(\mathcal{E})$ (it is easily seen that $\mathrm{col}(\Sigma) \subseteq \mathcal{E}$) and that all the nonzero eigenvalues of Σ are

equal to each other. After this step is completed, one can bound the combinatorial width using the following relation,

$$\omega(\mathcal{E}, \ell, p) = \frac{\text{trace}(\Sigma)}{\dim(\mathcal{E})}. \tag{13}$$

In particular, we have $q_\Omega(\mathcal{E}) = \mathcal{P}_\mathcal{E} I_\Omega{}'((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})^{-1} 1_{|\Omega|}$ with $|\Omega| \le \text{kruskal}(\mathcal{E})$. One can check that $\|q_\Omega(\mathcal{E})\|^2 \le \frac{|\Omega|}{\zeta}$, and then obtain that

$$\text{trace}(\Sigma) = \sum_{i=0}^{\ell} c_{p,\ell}\, p^i (1-p)^{k-i} \sum_{|\Omega|=i} \|q_\Omega\|^2 \le 2 \sum_{i=0}^{\ell} \binom{k}{i} p^i (1-p)^{k-i} \frac{i}{\zeta} \le \frac{2\,k\,p}{\zeta}. \tag{14}$$

The first inequality follows from the implication that $kp \le \ell \Rightarrow c_{p,\ell} \le 2$. The second inequality is obtained by bounding the sum from above with the expectation of a binomial random variable with parameters k and p. Combining (13) and (14) we have the desired result.

To complete the proof, we need to show that $\text{rank}(\Sigma) = \dim(\mathcal{E})$ and that all the nonzero eigenvalues of Σ are equal to each other. For each $i \le \ell$ denote $S^{(i)} := \sum_{|\Omega|=i} I_\Omega{}'((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})^{-1} 1_{|\Omega|} 1_{|\Omega|}{}'((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})^{-1} I_\Omega$ so that $\sum_{|\Omega|=i} q_\Omega(\mathcal{E}) q_\Omega(\mathcal{E})' = \mathcal{P}_\mathcal{E} S^{(i)} \mathcal{P}_\mathcal{E}$. Let $\Pi \in \mathbb{R}^{k \times k}$ be a permutation matrix such that $\Pi A_\Gamma \Pi' = A_\Gamma$, i.e., Π corresponds to an element of the automorphism group of Γ. It is easily seen that $\Pi \mathcal{P}_\mathcal{E} \Pi' = \mathcal{P}_\mathcal{E}$. Consequently, if a vertex subset $\Omega \subset \{1, \ldots, k\}$ is mapped to $\hat{\Omega}$ under the automorphism represented by Π, then we have that $(\mathcal{P}_\mathcal{E})_{\Omega,\Omega} = (\mathcal{P}_\mathcal{E})_{\hat{\Omega},\hat{\Omega}}$. In turn, one can check that $\Pi I_\Omega{}'((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})^{-1} 1_{|\Omega|} 1_{|\Omega|}{}'((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})^{-1} I_\Omega \Pi' = I_{\hat{\Omega}}{}'((\mathcal{P}_\mathcal{E})_{\hat{\Omega},\hat{\Omega}})^{-1} 1_{|\hat{\Omega}|} 1_{|\hat{\Omega}|}{}'((\mathcal{P}_\mathcal{E})_{\hat{\Omega},\hat{\Omega}})^{-1} I_{\hat{\Omega}}$. Based on these observations and the fact that $|\Omega| = i \Leftrightarrow |\hat{\Omega}| = i$, we note that a summand of $S^{(i)}$ gets mapped to another summand of $S^{(i)}$ under conjugation by Π. Moreover, automorphisms are injective functions, and hence distinct summands of $S^{(i)}$ must be mapped to distinct summands of $S^{(i)}$. Thus, we conclude that $\Pi S^{(i)} \Pi' = S^{(i)}$ for each $i \le \ell$ and for any permutation matrix $\Pi \in \mathbb{R}^{k \times k}$ representing an automorphism of Γ.

As Γ is vertex- and edge-transitive, and as Γ^c is also edge-transitive, each $S^{(i)}$ is of the following form:

$$(S^{(i)})_{p,q} = \begin{cases} \alpha_1 & \text{if } (A_\Gamma)_{p,q} = 1 \text{ and } p \ne q, \\ \alpha_2 & \text{if } (A_\Gamma)_{p,q} = 0 \text{ and } p \ne q \\ \alpha_3 & \text{if } p = q \end{cases} \implies S^{(i)} = \alpha_1 A_\Gamma + \alpha_2 A_{\Gamma^c} + \alpha_3 I_k, \tag{15}$$

for some $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. Since Γ is vertex-transitive it is also a regular graph and, consequently, the discussion from section 3.3.1 implies that the eigenspaces of A_Γ and A_{Γ^c} are the same. As Γ is assumed to be connected, we have from (11), (15), and from the equality $\sum_{|\Omega|=i} q_\Omega(\mathcal{E}) q_\Omega(\mathcal{E})' = \mathcal{P}_\mathcal{E} S^{(i)} \mathcal{P}_\mathcal{E}$ that

$$\sum_{|\Omega|=i} q_\Omega(\mathcal{E}) q_\Omega(\mathcal{E})' = \begin{cases} [\alpha_1 \lambda_\mathcal{E} + \alpha_2(k - \lambda_\mathcal{E} - 1) + \alpha_3] \frac{1_k 1_k{}^T}{k} & \text{if } \mathcal{E} = \text{span}\{1_k\}, \\ [\alpha_1 \lambda_\mathcal{E} - \alpha_2(\lambda_\mathcal{E} + 1) + \alpha_3] \mathcal{P}_\mathcal{E} & \text{otherwise.} \end{cases}$$

Since this holds for each index $i \le \ell$, we deduce that every summand of Σ in (12) is a scalar multiple of $\mathcal{P}_\mathcal{E}$ and, consequently, so is Σ. Furthermore, by construction, Σ cannot be the all-zeros matrix. Therefore, $\text{rank}(\Sigma) = \dim(\mathcal{E})$ and all nonzero eigenvalues of Σ are equal to each other. The result follows immediately. □

3.3.3. Simplifications based on coherence. The Kruskal rank of a subspace is intractable to compute in general; as a result, a number of subspace parameters have been considered in the literature to obtain tractable bounds on the Kruskal rank. The most prominent of these is the coherence parameter of a subspace. In our context, the additional analytical simplification provided by the coherence of a subspace along with Proposition 13 lead to simple performance guarantees on the Schur–Horn relaxation for symmetric planted graphs.

DEFINITION 14. *Let $\mathcal{S} \subseteq \mathbb{R}^k$ be a subspace. The* coherence *of $\mathcal{S}$, denoted $\mu(\mathcal{S})$, is defined as*

$$\mu(\mathcal{S}) := \max_{\substack{1 \leq i,j \leq k \\ i \neq j}} \frac{|(\mathcal{P}_\mathcal{S})_{i,j}|}{((\mathcal{P}_\mathcal{S})_{i,i})^{1/2}((\mathcal{P}_\mathcal{S})_{j,j})^{1/2}}.$$

The coherence parameter of a subspace can be computed efficiently, and it can be used to bound the Kruskal rank from below.

PROPOSITION 15 (see [16]). *For any subspace $\mathcal{S} \in \mathbb{R}^k$, $\mathrm{kruskal}(\mathcal{S}) \geq \frac{1}{\mu(\mathcal{S})}$.*

Further, for symmetric planted graphs Γ, the following result provides a bound on the minimum eigenvalue of minors of $\mathcal{P}_\mathcal{E}$ for eigenspaces $\mathcal{E}$ of A_Γ. Recall that this result is directly relevant in the context of Proposition 13.

PROPOSITION 16. *Suppose Γ is a vertex-transitive graph with adjacency matrix $A_\Gamma \in \mathbb{S}^k$, and let $\mathcal{E}$ denote an eigenspace of A_Γ. For any $\ell \in \mathbb{Z}$ with $\ell < \frac{1}{\mu(\mathcal{E})} + 1$, we have that $\min_{\substack{\Omega \subset \{1,\ldots,k\} \\ |\Omega| \leq \ell}} \lambda_{\min}((\mathcal{P}_\mathcal{E})_{\Omega,\Omega}) \geq \frac{\dim(\mathcal{E})}{k}(1 - (\ell - 1)\mu(\mathcal{E}))$.*

Proof. One can check that $\Pi \mathcal{P}_\mathcal{E} \Pi' = \mathcal{P}_\mathcal{E}$ for permutation matrices $\Pi \in \mathbb{R}^{k \times k}$ that correspond to automorphisms of Γ. Therefore, by vertex transitivity, the diagonal entries of $\mathcal{P}_\mathcal{E}$ are all equal to each other. As $\mathrm{trace}(\mathcal{P}_\mathcal{E}) = \dim(\mathcal{E})$, we conclude that $(\mathcal{P}_\mathcal{E})_{i,i} = \frac{\dim(\mathcal{E})}{k}$ for each $i = 1, \ldots, k$. Every row of $(\mathcal{P}_\mathcal{E})_{\Omega,\Omega}$ has at most $\ell - 1$ off-diagonal entries, and each of these entries is bounded above by $\frac{\dim(\mathcal{E})}{k}\mu(\mathcal{E})$. We obtain the desired result by applying the Gershgorin circle theorem. □

3.4. Main result. Building on the preceding discussion, we state and prove our main result Theorem 19. The proof of this result relies on an intermediate step regarding the M_{22} submatrix of the dual variable $M = \begin{pmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{pmatrix}$ from Proposition 8. From that result, we are required to obtain an $M_{22} \in \mathbb{S}^{n-k}$ such that (i) for each $i, j = 1, \ldots, n - k$ we have $(M_{22})_{i,j} = 1$ if $(A_\mathfrak{G})_{i+k,j+k} = 1$ or if $i = j$, and (ii) the operator norm $\|M_{22}\|_2$ is as small as possible.

We present the following result from [4], which we utilize subsequently in Lemma 18 to establish a bound on $\|M_{22}\|_2$.

LEMMA 17 (see [4]). *Let $X \in \mathbb{S}^d$ be a symmetric matrix whose entries $X_{i,j}$ are independent and centered random variables. For each $\epsilon \in (0, 1/2]$, there exists a constant $\tilde{c}_\epsilon$ such that for all $x \geq 0$*

$$\mathbb{P}(\|X\|_2 \geq (1 + \epsilon)2\tilde{\sigma} + x) \leq d \exp\left(-\frac{x^2}{\tilde{c}_\epsilon \tilde{\sigma}_*^2}\right),$$

where $\tilde{\sigma} := \max_i \sqrt{\sum_j \mathbb{E}[X_{i,j}^2]}$ and each $|X_{i,j}| \leq \tilde{\sigma}_*$ almost surely.

LEMMA 18. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_\mathfrak{G}$ of order k is equal to A_Γ, and the remaining edges in $\mathfrak{G}$ are drawn according to the Erdős–Rényi model with*

probability $p \in [0,1)$. For constants $c_1 = \sqrt{\frac{9p}{1-p}}$ and c_2 depending only on p and for $\alpha \geq c_1\sqrt{n-k}$, there exists $M_{22} \in \mathbb{S}^{n-k}$ satisfying

    1. *$(M_{22})_{i,j} = 1$ if $(A_\mathfrak{G})_{i+k,j+k} = 1$ or $i = j$,*
    2. *$\|M_{22}\|_2 < \alpha$*

with probability at least $1 - (n-k)\exp(-c_2(\alpha - c_1\sqrt{n-k})^2)$.

    Proof. Our proof is inspired by the approach in [2]. Consider the following matrix $M_{22} \in \mathbb{S}^{n-k}$:

$$(16) \qquad (M_{22})_{i,j} = \begin{cases} 1 \text{ if } (A_\mathfrak{G})_{i+k,j+k} = 1, i \neq j, \\ \frac{-p}{1-p} \text{ if } (A_\mathfrak{G})_{i+k,j+k} = 0, i \neq j, \\ 0 \text{ if } i = j. \end{cases}$$

As the submatrix $(A_\mathfrak{G})_{i+k,j+k}$, $i,j = 1,\ldots,n-k$, consists of independent and centered entries (in the off-diagonal locations) and zeros on the diagonal, one can check that M_{22} is a random matrix that satisfies the requirements of Lemma 17. Further, the first part of the present lemma is satisfied. The second claim follows from an application of Lemma 17 with $\epsilon = 1/2$. $\qquad\square$

    Combining Propositions 8 and 12 and Lemma 18, we now state and prove the main result of this paper.

    THEOREM 19. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_\mathfrak{G}$ of order k is equal to A_Γ, and the remaining edges in $\mathfrak{G}$ are drawn according to the Erdős–Rényi model. Suppose $\mathcal{E} \subset \mathbb{R}^k$ is an eigenspace of A_Γ with associated eigenvalue $\lambda_\mathcal{E}$, and we employ the Schur–Horn relaxation (P) with parameter $\gamma = \lambda_\mathcal{E}$. Further suppose that*

    1. *$p \in [0, \frac{\mathrm{kruskal}(\mathcal{E})}{k})$, and*
    2. *for some $\ell \in \mathbb{Z}$ satisfying $kp < \ell \leq \mathrm{kruskal}(\mathcal{E})$,*

$$n < \min\left(\frac{\mathrm{eigengap}(A_\Gamma, \mathcal{E})^2}{4\omega(\mathcal{E},\ell,p)}, \frac{(\mathrm{eigengap}(A_\Gamma, \mathcal{E}) - 2|\lambda_\mathcal{E}|)^2}{4c_1^2}\right) + k.$$

Then the Schur–Horn relaxation succeeds at identifying the planted subgraph Γ inside $\mathfrak{G}$ with probability at least $1 - p_1 - p_2$, where

$$p_1 = 1 - \left[\left(1 - \exp\left(-\frac{(\ell - kp)^2}{\ell + kp}\right)\right)^{n-k}\right.$$
$$\left.\times\left(1 - 2k\exp\left(-\frac{3\zeta(\frac{1}{4}\mathrm{eigengap}(A_\Gamma,\mathcal{E})^2 - (n-k)\omega(\mathcal{E},\ell,p))^2}{4\ell(\frac{1}{4}\mathrm{eigengap}(A_\Gamma,\mathcal{E})^2 + 2(n-k)\omega(\mathcal{E},\ell,p))}\right)\right)\right]$$

and

$$p_2 = (n-k) \times \exp(-c_2(\tfrac{1}{2}\mathrm{eigengap}(A_\Gamma, \mathcal{E}) - |\lambda_\mathcal{E}| - c_1\sqrt{n-k})^2).$$

Here, $\zeta = \min_{\substack{\Omega \subset \{1,\ldots,k\} \\ |\Omega| \leq \ell}} \lambda_{\min}((\mathcal{P}_\mathcal{E})_{\Omega,\Omega})$. Further, the constants c_1 and c_2 depend only on p.

    Proof. As discussed previously, since $\ell \leq \mathrm{kruskal}(\mathcal{E})$ we have that $\zeta > 0$. We establish the result by constructing a dual certificate M satisfying the conditions of Proposition 8.

    We start by setting $M_{11} = A_\Gamma$. This ensures that conditions (ii) and (iii) of Proposition 8 are satisfied. Next, we choose M_{12} as discussed in Proposition 12, with the parameter $\delta = \frac{1}{2}\mathrm{eigengap}(\Gamma, \mathcal{E})$, which satisfies $\delta \geq \sqrt{(n-k)\omega(\mathcal{E},\ell,p)}$ due to the upper bound on n. Such an M_{12} exists with probability at least $1 - p_1$, and it satisfies

condition (iv) of Proposition 8 as well as the bound $\|M_{12}\|_2 < \frac{1}{2}\text{eigengap}(A_\Gamma, \mathcal{E})$. Finally, we set M_{22} as discussed in Lemma 18, with $\alpha = \frac{1}{2}\text{eigengap}(A_\Gamma, \mathcal{E}) - |\lambda_\mathcal{E}|$, which satisfies $\alpha \geq c_1\sqrt{n-k}$ due to the upper bound on n. Such an M_{22} exists with probability at least $1 - p_2$ and satisfies the bound $\|M_{22}\|_2 < \frac{1}{2}\text{eigengap}(A_\Gamma, \mathcal{E}) - |\lambda_\mathcal{E}|$.

Based on this construction, the matrix $M = \begin{pmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{pmatrix}$ satisfies conditions (i) and (v) of Proposition 8. Thus, if M_{12} and M_{22} with the stated properties exist, then all the conditions of Proposition 8 are satisfied. By the union bound, the desired M_{12} and M_{22} exist concurrently with probability at least $1 - p_1 - p_2$. $\square$

Remark 1. The parameter ℓ arises in multiple aspects of this result. We discuss specific choices of ℓ in the corollaries in the next section.

Theorem 19 provides a nonasymptotic bound on the performance of the Schur–Horn relaxation (P). In words, this relaxation succeeds with high probability in identifying a subgraph Γ planted inside a larger graph $\mathfrak{G}$ (under the Erdős–Rényi model) provided A_Γ has an eigenspace $\mathcal{E}$ satisfying four conditions: (i) the eigenspace $\mathcal{E}$ has large Kruskal rank; (ii) the eigenspace $\mathcal{E}$ has small combinatorial width; (iii) A_Γ has a large eigengap with respect to $\mathcal{E}$; and (iv) the projection matrix $\mathcal{P}_\mathcal{E}$ has the property that all sufficiently large principal minors are well-conditioned. In practice, larger dimensional eigenspaces of A_Γ may be expected to satisfy these conditions more easily and, therefore, we set γ equal to the eigenvalue of A_Γ of largest multiplicity in our experimental demonstrations in section 4.

3.5. Specializations of Theorem 19. We appeal to the discussion in section 3.3 on the properties of the Kruskal rank and the combinatorial width to obtain specializations of Theorem 19 to certain graph families. We begin by considering the case of symmetric planted graphs with symmetric complements.

COROLLARY 20. *Consider a planted subgraph problem instance in which the nodes of $\mathfrak{G}$ and Γ are labeled so that the leading principal minor of $A_\mathfrak{G}$ of order k is equal to A_Γ, and the remaining edges in $\mathfrak{G}$ are drawn according to the Erdős–Rényi model. Suppose $\mathcal{E} \subset \mathbb{R}^k$ is an eigenspace of A_Γ with associated eigenvalue $\lambda_\mathcal{E}$, and we employ the Schur–Horn relaxation (P) with parameter $\gamma = \lambda_\mathcal{E}$. Further suppose that the following three conditions hold:*

1. Γ is a connected symmetric graph with a symmetric complement,

2. $p \in [0, \frac{1}{\mu(\mathcal{E})k})$,

3. $n < \min\left(\frac{\text{eigengap}(A_\Gamma, \mathcal{E})^2 \dim(\mathcal{E})^2 \left(1 - kp\mu(\mathcal{E})\right)}{16k^2p}, \frac{(\text{eigengap}(A_\Gamma, \mathcal{E}) - 2|\lambda_\mathcal{E}|)^2}{4c_1^2}\right) + k$.

Then the Schur–Horn relaxation succeeds in identifying the planted subgraph Γ inside the larger graph $\mathfrak{G}$ with probability at least $1 - p_1 - p_2$, where p_1 and p_2 are as stated in Theorem 19 (one can substitute $\frac{\dim(\mathcal{E})}{2k}(1 - kp\mu(\mathcal{E}))$ for the ζ term and $\frac{4k^2p}{\dim(\mathcal{E})^2(1-kp\mu(\mathcal{E}))}$ for the $\omega(\mathcal{E}, \ell, p)$ term appearing in p_1).

Proof. This result follows by a combination of Theorem 19, and Propositions 13, 15, 16. Set $\ell = \lceil \frac{1}{2}(kp + \frac{1}{\mu(\mathcal{E})}) \rceil$. This choice satisfies $kp < \ell \leq \text{kruskal}(\mathcal{E})$ based on Proposition 15. One can also check that the inequality $\ell < \frac{1}{\mu(\mathcal{E})} + 1$ holds. The vertex transitivity of Γ implies that one can appeal to Proposition 16 to conclude that

$$(17) \quad \min_{\substack{\Omega \subset \{1,\ldots,k\} \\ |\Omega| \leq \ell}} \lambda_{\min}\left((\mathcal{P}_\mathcal{E})_{\Omega, \Omega}\right) \geq \frac{\dim(\mathcal{E})}{k}\left(1 - (\ell - 1)\mu(\mathcal{E})\right) > \frac{\dim(\mathcal{E})}{2k}(1 - kp\mu(\mathcal{E})).$$

Based on the condition on p, this lower bound is strictly positive. As Γ is symmetric

and has a symmetric complement (and is connected), we conclude from Proposition 13 that $\omega(\mathcal{E}, \ell, p) \le \frac{4k^2 p}{\dim(\mathcal{E})^2(1-kp\mu(\mathcal{E}))}$.

Finally, one can check that conditions (2) and (3) of the corollary imply that both of the requirements of Theorem 19 are met and, hence, the Schur–Horn relaxation succeeds in identifying the planted subgraph Γ with probability at least $1 - p_1 - p_2$, where p_1 and p_2 are as stated in Theorem 19—one can substitute $\frac{4k^2 p}{\dim(\mathcal{E})^2(1-kp\mu(\mathcal{E}))}$ as an upper bound for $\omega(\mathcal{E}, \ell, p)$ and $\frac{\dim(\mathcal{E})}{2k}(1 - kp\mu(\mathcal{E}))$ as a lower bound for ζ, which yields a lower bound on $1 - p_1 - p_2$ from Theorem 19. ☐

As the coherence parameter of an eigenspace is more tractable to compute than the Kruskal rank, this result provides an efficiently verifiable set of conditions that guarantee the success of the Schur–Horn relaxation (P) for symmetric planted graphs Γ. This result specialized to the case of the planted clique problem yields the result of Ames and Vavasis [2].

COROLLARY 21. *Fix $p \in [0, 1)$ and consider a family of planted clique problem instances $\{\Gamma_k, \mathfrak{G}_k\}_{k=1}^{\infty}$ generated according to the Erdős–Rényi model, where Γ_k is the k-clique and $\mathfrak{G}_k$ is a graph on n_k nodes. There exists a constant $\beta > 0$ only depending on p such that if $n_k \le \beta k^2$, the Schur–Horn relaxation with $\gamma = -1$ succeeds in identifying Γ_k inside $\mathfrak{G}_k$ with probability approaching one exponentially fast in k.*

Proof. The k-clique is a connected symmetric graph with a complement that is also symmetric; hence the first condition of Corollary 20 is satisfied. Each $A_{\Gamma_k} \in \mathbb{S}^k$ has a $(k-1)$-dimensional eigenspace $\mathcal{E}$ such that $\mu(\mathcal{E}) = \frac{1}{k-1}$, $\dim(\mathcal{E}) = k - 1$, and eigengap$(A_\Gamma, \mathcal{E}) = k$.

Based on the choice $\ell = \lceil \frac{1}{2}(kp + \frac{1}{\mu(\mathcal{E})}) \rceil$ as in Corollary (20), one can check that $\frac{\dim(\mathcal{E})}{2k}(1 - kp\mu(\mathcal{E})) = \Theta(1)$, that $\omega(\mathcal{E}, \ell, p) \le \frac{4k^2 p}{\dim(\mathcal{E})^2 (1-kp\mu(\mathcal{E}))} = \Theta(1)$, and that $\ell - kp = \Theta(k)$.

Set $n_k = \frac{k^2}{32} \min(\frac{1-kp\mu(\mathcal{E})}{2p}, \frac{1}{c_1{}^2}) + k$. One can check that the third condition of Corollary 20 is satisfied with this choice. Moreover, this value of n_k (or any smaller value) yields $\frac{1}{4}$eigengap$(A_\Gamma, \mathcal{E})^2 - (n_k - k)\omega(\mathcal{E}, \ell, p) = \Theta(k^2)$ and $\frac{1}{2}$eigengap$(A_\Gamma, \mathcal{E}) - 1 - c_1\sqrt{n_k - k} = \Theta(k)$.

By Corollary 20, we conclude that the Schur–Horn relaxation (P) with parameter $\gamma = -1$ identifies a hidden k-clique with probability $1 - p_1 - p_2$, where

$$p_1 = 1 - \left(1 - \exp(-c_4 k)\right)^{n-k}\left(1 - 2k\exp(-c_3 k)\right) \longrightarrow 0, \text{ as } n, \ k \to \infty, \text{ and}$$
$$p_2 = (n - k)\exp\left(-c_5 k^2\right) \longrightarrow 0, \text{ as } n, \ k \to \infty,$$

for some constants $c_3 > 0$, $c_4 > 0$, and $c_5 > 0$. ☐

Thus, Theorem 19 can be specialized to obtain the result presented in [2].

Remark 2. Theorem 19 can also be specialized to obtain recovery results for other families of planted subgraph problems. As an example, for some fixed integer $t > 1$ and probability p, consider a family of planted subgraph problem instances $\{\Gamma_k, \mathfrak{G}_k\}_{k=t,2t,\ldots}$ in which Γ_k is the complete t-partite graph on k vertices whose independent sets each contain k/t vertices—this graph is the complement of a disjoint union of t cliques with size k/t—and $\mathfrak{G}_k$ is a graph on n_k vertices. By suitably specializing Theorem 19, one can show that the Schur–Horn relaxation succeeds in identifying the planted complete t-partite graph with probability approaching 1 exponentially fast as $k \to \infty$ provided that $p \in [0, \frac{1}{t})$ and $n_k \lesssim k^2$.

4. Numerical experiments.

4.1. Semidefinite descriptions of the Schur–Horn orbitope. We begin with a discussion of semidefinite representations of the Schur–Horn orbitope $\mathcal{SH}(M)$ for $M \in \mathbb{S}^n$. Specifically, suppose $s_\ell : \mathbb{S}^n \to \mathbb{R}$ denotes the sum of the ℓ-largest eigenvalues of a symmetric matrix for $\ell = 1, \ldots, n$. Then the Schur–Horn orbitope $\mathcal{SH}(M)$ can be described via majorization inequalities on the spectrum [36]:
(18)
$$\mathcal{SH}(M) = \{N \in \mathbb{S}^n \mid s_\ell(N) \leq s_\ell(M) \text{ for } 1 \leq \ell \leq n - 1, \text{ and trace}(N) = \text{trace}(M)\}.$$

As the sublevel sets of the convex functions s_ℓ have tractable semidefinite descriptions [6], one can obtain a lifted polynomial-sized semidefinite representation of $\mathcal{SH}(M)$ for arbitrary $M \in \mathbb{S}^n$. However, specifications of $\mathcal{SH}(M)$ via semidefinite representations of the sublevels sets of s_ℓ involve a total of $O(n)$ additional matrix variables in $\mathbb{S}^n$ and $O(n)$ semidefinite constraints (one for each of the majorization inequalities in (18)); in particular, these do not take advantage of any structure in the spectrum of M, such as multiplicities in the eigenvalues.

We discuss next an alternative semidefinite representation of $\mathcal{SH}(M)$ that is based on a modification of the description of $\mathcal{SH}(M)$ presented in [14], and it exploits the multiplicities in the eigenvalues of M so that both the number of additional matrix variables and semidefinite constraints scale with the number of distinct eigenvalues of M rather than the ambient size n of M. Suppose M has q distinct eigenvalues $\lambda_1, \ldots, \lambda_q$ with multiplicities $m_1, \ldots, m_q$. Then one can check that [14]
(19)
$$\mathcal{SH}(M) = \left\{ N \in \mathbb{S}^n \ \middle| \ \exists Y_i \in \mathbb{S}^n, Y_i \succeq 0, \ i = 1, \ldots, q, \text{ such that} \right.$$
$$\left. N = \sum_{i=1}^q \lambda_i Y_i, \ \sum_{i=1}^q Y_i = I_n, \ \text{trace}(Y_i) = m_i \text{ for } i = 1, \ldots, q \right\}.$$

In this latter description of the Schur–Horn orbitope, both the number of additional matrix variables in $\mathbb{S}^n$ and the number of semidefinite constraints are on the order of the number of distinct eigenvalues of M, which can be far smaller than n for the adjacency matrices of graphs considered in this paper. In the numerical experiments presented next, we employ the description (19) of the Schur–Horn orbitope.

4.2. Experimental results. We investigate the performance of the Schur–Horn relaxation (P) in planted subgraph problems with the four planted subgraphs Γ listed in Table 1. For each of these graphs, we set γ equal to the eigenvalue corresponding to the largest eigenspace of the corresponding graph. We vary n (the size of the larger graph $\mathfrak{G}$ inside which Γ is planted) and p (the probability of a noise edge in $\mathfrak{G}$), and we obtain 10 random instances of planted subgraph problems for each value of n and p. In Figure 6, we plot the empirical probability of success of the Schur–Horn relaxation for these random trials; the white cells represent a probability of success of one and the black cells represent a probability of success of zero. Our results were obtained using the CVX parser [23, 24] and the SDPT3 solver [38]. In each of the four cases, the Schur–Horn relaxation (P) succeeds in solving the underlying planted subgraph problem for suitably small n and p.

5. Discussion.
In this paper, we introduce a new convex relaxation approach for the planted subgraph problem, and we describe families of problem instances for which our method succeeds. Our method generalizes previous convex optimization

Planted subgraphs (and associated parameters) for which we demonstrate the utility of the Schur–Horn relaxation. See Figure 6 for the associated phase transitions.

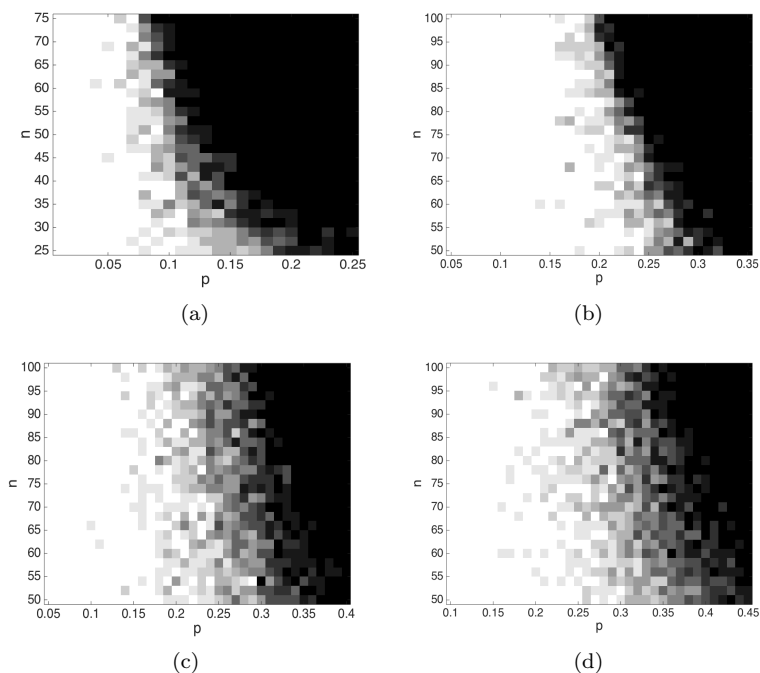| Planted graph Γ [with # vertices]	Eigenvalues [with multiplicity]	Kruskal rank of the largest eigenspace
Clebsch [$k = 16$] (Figure 1(a))	$5[\times 1], -3[\times 5], 1[\times 10]$	5
Generalized quadrangular-$(2, 4)$ [$k = 27$] (Figure 4(b))	$10[\times 1], -5[\times 6], 1[\times 20]$	10
8-triangular [$k = 28$] (Figure 2(a))	$12[\times 1], 4[\times 7], -2[\times 20]$	6
9-triangular [$k = 36$] (Figure 2(b))	$14[\times 1], 5[\times 8], -2[\times 27]$	7

FIG. 6. *Phase transition plots based on the experiment described in section* 4.2 *for the* (a) *Clebsch graph,* (b) *Generalized quadrangle-*$(2, 4)$ *graph,* (c) 8-*triangular graph, and* (d) 9-*triangular graph.*

techniques for identifying planted cliques based on nuclear norm minimization [2], and it is useful for identifying planted subgraphs consisting of few distinct eigenvalues. There are several further directions that arise from our investigation, and we mention a few of these here.

Spectrally comonotone matrices with sparsity constraints. One of the ingredients in the proof of our main result Theorem 19 is to find a matrix $M_{11} \in \mathbb{S}^k$ that is spectrally comonotone with $A_\Gamma \in \mathbb{S}^k$, and that further satisfies the condition that $(M_{11})_{i,j} = 1$ whenever $(A_\Gamma)_{i,j} = 1$. In the proof of Theorem 19 we simply choose $M_{11} = A_\Gamma$. This choice does not exploit the fact that the entries of M_{11} corresponding to those where $(A_\Gamma)_{i,j} = 0$ are not constrained (and in particular can be nonzero).

With a different choice of M_{11}, one could replace eigengap$(A_\Gamma, \mathcal{E})$ in Theorem 19 by eigengap$(M_{11}, \mathcal{E})$ (recall that $\mathcal{E}$ is an eigenspace of A_Γ). Consequently, our main result could be improved via principled constructions of matrices $M_{11} \in \mathbb{S}^k$ that satisfy the conditions of Theorem 19 and for which eigengap$(M_{11}, \mathcal{E}) >$ eigengap$(A_\Gamma, \mathcal{E})$.

Sparse graphs with eigenspaces with large Kruskal rank. One of the central questions concerning the planted subgraph problem is the possibility of identifying "sparse" planted subgraphs inside "dense" noise via computationally tractable approaches. Concretely, suppose Γ is a regular graph with degree d. Under the Erdős–Rényi model for the noise, the average degree of any k-node subgraph of the larger graph $\mathfrak{G}$ is about $(k-1)p$. From Theorem 19, we have that the Schur–Horn relaxation succeeds (with high probability) in identifying Γ if $p \in [0, \frac{\text{kruskal}(\mathcal{E})}{k})$, where $\mathcal{E} \subset \mathbb{R}^k$ is one of the eigenspaces of Γ. In other words, (for suitably large k) if $d < \text{kruskal}(\mathcal{E})$ then the Schur–Horn relaxation succeeds in identifying Γ inside $\mathfrak{G}$ despite the fact that Γ is sparser than a typical k-node subgraph in $\mathfrak{G}$. Of the graphs we have investigated in this paper, the Clebsch graph from Figure 1(a), is an example in which both the degree and the Kruskal rank of the largest subspace are equal to 5. For some of the other small graphs discussed in this paper, the degree is larger than the Kruskal ranks of the eigenspaces. For larger graphs, the computation of the Kruskal rank of the large eigenspaces quickly becomes computationally intractable. Therefore, it is of interest to identify graph families in which (by construction) the degree is smaller than the Kruskal rank of one of the eigenspaces.

Convex geometry and graph theory. In developing convex relaxations for the planted subgraph problem (based on the formulation (1)) as well as other inverse problems involving unlabeled graphs, the key challenge is one of obtaining tractable convex outer approximations of the set $\mathcal{A}(B) = \{\Pi B \Pi' \mid \Pi \text{ is an } n \times n \text{ permutation matrix}\}$ for some given adjacency matrix $B \in \mathbb{S}^n$. In particular, a convex approximation $\mathcal{C}$ that contains $\mathcal{A}(B)$ is useful if the normal cone $\mathcal{N}_\mathcal{C}(B)$ is large; as an example, the Schur–Horn relaxation $\mathcal{SH}(B)$ has this property for adjacency matrices B with few distinct eigenvalues. More generally, what is an appropriate convex relaxation for other structured graph families such as low-treewidth graphs (arising in inference in statistical graphical models), or graphs with a specified degree distribution (arising in social network analysis)? Recent work [13] provides a catalog of *convex graph invariants* that are useful for obtaining computationally tractable convex relaxations of $\mathcal{A}(B)$. A deeper investigation of the interaction between convex-geometric aspects of these invariants (such as the normal cones of the associated convex relaxations) and the structural properties of the graph specified by the adjacency matrix B has the potential to yield new convex relaxations for general inverse problems on graphs.

Finding subgraphs versus induced subgraphs. A planted clique is both a subgraph and an induced subgraph of an underlying graph $\mathfrak{G}$. However, for other planted graphs, this is not necessarily the case as there could be additional noise edges between the vertices of the planted graph Γ. In this paper we focus on identifying planted subgraphs that are induced subgraphs, and we investigate the theoretical properties as well as the empirical utility of the Schur–Horn relaxation (P) for this task. For identifying planted subgraphs that may not be induced subgraphs, the experimental results in Figure 7 suggest that Schur–Horn relaxation (P) continues to be useful. Specifically, Figure 7(a) gives the performance of the Schur–Horn relaxation in a setting in which the generalized quadrangle-$(2, 4)$ graph is a subgraph of a larger graph on n nodes and noise edges (both outside the planted graph and between disconnected

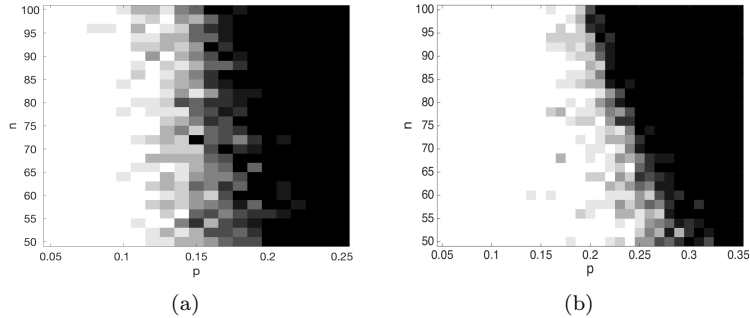(a)   (b)

Fig. 7. *Phase transition plot for the problem of identifying the generalized quadrangle-$(2, 4)$ graph: on the left is a plot in which this graph is a subgraph and on the right is a plot in which this graph is an induced subgraph (as in section 4.2).*

vertices of the planted graph) occur with probability p. In comparison with the result of section 4.2 (when the generalized quadrangle-$(2, 4)$ graph is an induced subgraph), we notice that the Schur–Horn relaxation performs slightly worse. A reason for this is that identifying structured subgraphs is inherently harder than identifying induced subgraphs due to the additional noise edges. Analyzing the performance of the Schur–Horn relaxation (P)—both theoretically and empirically—in a more extensive manner for identifying subgraphs that may not be induced remains an open question.

REFERENCES

[1] N. ALON, M. KRIVELEVICH, AND B. SUDAKOV, *Finding a large hidden clique in a random graph*, Random Structures Algorithms, 13 (1998), pp. 457–466.

[2] B. P. W. AMES AND S. A. VAVASIS, *Nuclear norm minimization for the planted clique and biclique problems*, Math. Program., 129 (2011), pp. 69–89.

[3] P. J. ARTYMIUK, A. R. POIRRETTE, H. M. GRINDLEY, D. W. RICE, AND P. WILLETT, *A graph-theoretic approach to the identification of three-dimensional patterns of amino acid side-chains in protein structures*, J. Mol. Biol., 243 (1994), pp. 327–344.

[4] A. S. BANDEIRA AND R. VAN HANDEL, *Sharp nonasymptotic bounds on the norm of random matrices with independent entries*, Ann. Probab., 44 (2016), pp. 2479–2506.

[5] E. BANNAI AND T. ITO, *Algebraic Combinatorics*, Benjamin/Cummings, Menlo Park, CA, 1984.

[6] A. BEN-TAL AND A. NEMIROVSKI, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Ser. Optim. 2, SIAM, Philadelphia, 2001.

[7] R. C. BOSE, *Strongly regular graphs, partial geometries and partially balanced designs*, Pacific J. Math., 13 (1963), pp. 389–419.

[8] W. BRIDGES AND R. MENA, *Multiplicative cones—a family of three eigenvalue graphs*, Aequationes Math., 22 (1981), pp. 208–214.

[9] A. E. BROUWER AND W. H. HAEMERS, *Distance-Regular Graphs*, Springer, Berlin, 2012.

[10] A. E. BROUWER AND J. H. VAN LINT, *Strongly regular graphs and partial geometries*, in Enumeration and Design, Academic Press, London, 1984, pp. 85–122.

[11] P. J. CAMERON, *Strongly regular graphs*, in Selected Topics in Graph Theory, Academic Press, London, 1978, pp. 337–360.

[12] E. J. CANDÈS AND B. RECHT, *Exact matrix completion via convex optimization*, Found. Comput. Math., 9 (2009), pp. 717–772.

[13] V. CHANDRASEKARAN, P. A. PARRILO, AND A. S. WILLSKY, *Convex graph invariants*, SIAM Rev., 54 (2012), pp. 513–541.

[14] Y. DING AND H. WOLKOWICZ, *A low-dimensional semidefinite relaxation for the quadratic assignment problem*, Math. Oper. Res., 34 (2009), pp. 1008–1022.

[15] R. DOBRIN, Q. K. BEG, A.-L. BARABÁSI, AND Z. N. OLTVAI, *Aggregation of topological motifs in the escherichia coli transcriptional regulatory network*, BMC Bioinform., 5 (2004), 10.

[16] D. L. DONOHO AND M. ELAD, *Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ1 minimization*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 2197–2202.

[17] M. Doob, *Graphs with a small number of distinct eigenvalues*, Ann. New York Acad. Sci. 175 (1970), pp. 104–110.

[18] M. Doob, *On characterizing certain graphs with four eigenvalues by their spectra*, Linear Algebra Appl., 3 (1970), pp. 461–482.

[19] M. Fazel, *Matrix Rank Minimization with Applications*, Ph.D. thesis, Stanford University, Stanford, CA, 2002.

[20] U. Feige and R. Krauthgamer, *Finding and certifying a large hidden clique in a semirandom graph*, Random Structures Algorithms, 16 (2000), pp. 195–208.

[21] G. Finke, R. E. Burkard, and F. Rendl, *Quadratic assignment problems*, in Surveys in Combinatorial Optimization, North-Holland Math Stud. 132, North-Holland, Amsterdam, 2011, pp. 61–82.

[22] C. Godsil, *Algebraic Combinatorics*, Chapman & Hall Math. Ser. 6, CRC Press, New York, 1993.

[23] M. C. Grant and S. P. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, http://cvxr.com/cvx/ (2017).

[24] M. C. Grant and S. P. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, Springer, London, 2008, pp. 95–110.

[25] W. H. Haemers and V. D. Tonchev, *Spreads in strongly regular graphs*, Des. Codes Cryptogr., 8 (1996), pp. 145–157.

[26] R. M. Karp, *Reducibility Among Combinatorial Problems*, University of California, Berkeley, CA, 1972.

[27] M. Krivelevich and B. Sudakov, *Pseudo-random graphs*, in More Sets, Graphs and Numbers, Springer, Berlin, 2006, pp. 199–262.

[28] J. B. Kruskal, *Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra Appl., 18 (1977), pp. 95–138.

[29] J. Leskovec, K. J. Lang, and M. Mahoney, *Empirical comparison of algorithms for network community detection*, in Proceedings of the 19th International Conference on World Wide Web, ACM, New York, 2010, pp. 631–640.

[30] O. Mason and M. Verwoerd, *Graph theory and networks in biology*, IET Syst. Biol., 1 (2007), pp. 89–119.

[31] N. Mishra, R. Schreiber, I. Stanton, and R. E. Tarjan, *Clustering social networks*, in Algorithms and Models for the Web-Graph, Springer, Berlin, 2007, pp. 56–67.

[32] M. Muzychuk and M. Klin, *On graphs with three eigenvalues*, Discrete Math., 189 (1998), pp. 191–207.

[33] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, *Defining and identifying communities in networks*, Proc. Natl. Acad. Sci. USA, 101 (2004), pp. 2658–2663.

[34] B. Recht, M. Fazel, and P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev., 52 (2010), pp. 471–501.

[35] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 2015.

[36] R. Sanyal, F. Sottile, and B. Sturmfels, *Orbitopes*, Mathematika, 57 (2011), pp. 275–314.

[37] J. J. Seidel, *Strongly regular graphs*, in Recent Progress in Combinatorics, Academic Press, London, 1969, pp. 185–198.

[38] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3 – a matlab software package for semidefinite programming, version 1.3*, Optim. Methods Softw., 11 (1999), pp. 545–581.

[39] E. R. van Dam, *Regular graphs with four eigenvalues*, Linear Algebra Appl., 226 (1995), pp. 139–162.

[40] E. R. van Dam, *Graphs with few eigenvalues. An Interplay between Combinatorics and Algebra*, Ph.D. thesis, Tilburg University, Tilburg, The Netherlands, 1996.

[41] E. R. van Dam and E. Spence, *Small regular graphs with four eigenvalues*, Discrete Math., 189 (1998), pp. 233–257.

[42] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, in Compressed Sensing, Cambridge University Press, Cambridge, 2012, pp. 210–268.

[43] J. von Neumann, *Some matrix inequalities and metrization of matrix space*, Tomsk Univ. Rev, 1 (1937), pp. 286–300.