Lower Bound on the Block-Diagonal SDP Relaxation of the Clique Number for Seidel Spectrum Bounded Graphs

Yifan Wang*   Yanling Shen*   Vladimir A. Kobzar†

Abstract

This work establishes a lower bound on the optimal value of the block-diagonal semidefinite program (SDP) relaxation for the clique number of a family of graphs parametrized by their order n provided that their Seidel spectrum is bounded as $O(\sqrt{n})$ as n gets large (Seidel bounded graphs or SBGs). These graphs include symmetric conference graphs, such as the Paley graphs, among others. The size of the maximal clique (clique number) of a graph is a classic NP-complete problem; the Paley graph is a deterministic graph where two vertices are connected if their difference is a quadratic residue in a finite field with the number of elements given by certain primes and prime powers. Improving on the current $O(\sqrt{p})$ upper bound for the clique number of Paley graphs of prime order p is a long-standing open problem in number theory and combinatorics. Kunisky and Yu (CCC 2023) provided numerical evidence that the upper bounds given by the sum-of-squares relaxations of degree 4 (SOS-4) are growing at an order smaller than square root of p and proved that the values of these relaxations are lower bounded as $\Omega(p^{1/3})$. Gvozdenovic, Laurent and Vallentin introduced a block-diagonal hierarchy (L^t) of SDPs that are weaker than the SOS-$2t$ SDPs. Therefore, the values of these block-diagonal SDPs of degree 2 (L^2) bound from above the values of the corresponding SOS-4 relaxations, and the $\Omega(p^{1/3})$ lower bound also applies to the L^2 relaxations of the Paley graph clique number. Building on the above-mentioned work, using Feige-Krauthgamer pseudomoments, we show that these L^t relaxations of the Paley graphs and other SBGs are bounded from below by $2^{1-t}\sqrt{n}$, at the leading order as the number of vertices n gets large. Since the L^t hierarchy is stronger than the Lovász-Schrijver hierarchy, our lower bound also applies to the latter hierarchy. Lastly, we study the subgraphs (localizations) of vertex-transitive graphs, including the Paley graphs, induced on a set of vertices extending a clique of a given size a to a maximal clique. We prove that interchanging localization degree a and relaxation degree t are equivalent for the purpose of our lower bound, which is consistent with the localization-relaxation trade-off conjectured by Kunisky (Exp. Math. 2024). More broadly, we make progress towards generalizing methods available to analyze SDPs to the block-diagonal hierarchy. This hierarchy appears to be well suited for relaxations of combinatorial optimization/graph-based problems, yet remains relatively unexplored.

1 Introduction

This work is motivated by the problem of estimating the clique number of the Paley graphs, which is a classic open problem in number theory with connections to additive combinatorics, random matrix theory, Ramsey theory, complexity theory and compressed sensing/sparse recovery. Semidefinite programming (SDP) and other convex relaxations are popular tools to bound from above the clique

*Department of Applied Physics and Applied Mathematics, Columbia University. These authors contributed equally to this work.

†Department of Mathematics, The Ohio State University.

number of various graphs of interest, including the Paley graphs. Our work focuses on the block-diagonal SDP hierarchy of relaxations developed by Gvozdenovic, Laurent and Vallentin, which are at least as powerful as the classic Lovász-Schrijver relaxations and include the Lovász theta function (which we will denote by ϑ) as the first and weakest level of the hierarchy. We establish lower bounds on these relaxations of the clique number for the Paley graphs of prime order, as well as certain subgraphs thereof, referred to as localizations.

Our lower bound also applies to any other family of graphs parametrized by their order n with Seidel spectrum bounded as $O(\sqrt{n})$, which we denote as *Seidel bounded graphs (SBGs)*. These graphs include all symmetric conference graphs.

A subset of vertices K in a graph G forms a *clique* if every pair in K is adjacent, and the *clique number* $\omega(G)$ is the size of a largest clique in G. Conversely, a subset of vertices in G is an *independent set* if no two vertices in it are connected, and the *independence number* $\alpha(G)$ is the size of the largest independent set in G (a single vertex and the empty set $\varnothing$ are deemed to be cliques and independence sets of sizes 1 and 0 respectively.) Computing each of these numbers (the *clique and independent set problems*, respectively) for a general graph is a classic NP-complete problem [Kar72]. Moreover, it is hard to approximate these numbers within any polynomial factor $n^{1-\epsilon}$ for any $\epsilon > 0$ where n is the number of vertices [Hås99].

The clique and independence numbers are among the main quantities of interest in Ramsey theory. For samples G of *Erdős-Renyi (ER) graphs*, denoted by $\mathcal{G}(\frac{1}{2}, n)$, which are random graphs on n vertices with edge probability $\frac{1}{2}$, the classic result of Erdős provides that $\mathsf{max}(\omega(G), \alpha(G)) \leq \mathsf{log}_2 n$ with positive probability [Erd47]. However, finding deterministic graphs with small independence and clique numbers remains a significant open problem, sometimes referred to as one of finding "hay in the haystack".

The Paley graphs are connected to the construction of deterministic matrices satisfying the restricted isometry property (RIP), an important problem in compressed sensing and sparse recovery [Tao08] (see also [Mix15]). References [BFMW13, BMM16] constructed a family of deterministic matrices (equitriangluar frames) from the rows of the discrete Fourier transform matrix indexed by the quadratic residues modulo a prime number p (the *Paley matrices*), which provably achieve RIP when sparsity is on the order of $\sqrt{p}$ but are *conjectured* to achieve higher sparsity, breaking this square root bottleneck (see also [AAM15] designing deterministic RIP matrices using the adjacency matrix of a Paley graph). In this conditional construction, a lower bound on $\omega(G_p)$ leads to a lower bound on the distortion in the sparse recovery (Theorem 2.3 in [BMM16]). Using a similar analysis, reference [KPB19] overcame the square root bottleneck unconditionally for signals with a certain sparse structure. Finally, reference [BFMM16] constructed a matrix using the Legendre symbol (which is closely connected to the Paley graphs) to reduce the number of random bits in a random RIP matrix.

We focus on the *Paley graphs* of prime order p, denoted by G_p, which are undirected graphs with vertices identified with the elements of the finite field $\mathbb{F}_p$ where $p \equiv 1 \pmod 4$. Any two vertices i and j in G_p are connected if and only if $i - j$ is a quadratic residue in $\mathbb{F}_p$. The Paley graphs are considered to be *pseudorandom*, sharing certain similarities with the ER graphs [CGW89, KS06].

Since the Paley graphs are self-complementary, their clique and independence numbers are equal. We focus on the former problem because it makes the connection between relaxations and localizations apparent more readily in our results. Presently it is not known whether or not the clique number $\omega(G_p)$ is $O(p^{\frac{1}{2}-\epsilon})$ for some $\epsilon > 0$; this problem is referred to as the *square root barrier*. Spectral methods provide that $\omega(G_p) \leq \sqrt{p}$, and the state-of-the-art upper bound by Hanson and Petridis

$$HP(G_p) := \frac{\sqrt{2p-1}+1}{2} \approx \sqrt{\frac{p}{2}}$$

improves on the spectral bound by a constant prefactor [HP21, BSW21]. On the other hand, reference [GR90] showed that for infinitely many primes p,

$$\omega(G_p) \geq \log p \log \log \log p.$$

Accordingly, there is significant gap between the existing upper and lower bounds on the clique number of the Paley graphs. Based on numerical experiments, this number is conjectured to be $O(\text{polylog } p)$ (see discussion of [Exo23, She23] in [BMR14]).

The maximal clique problem can be formulated as an integer $\{0, 1\}$ program. A long line of work established that classic convex relaxations, including the Lovász-Schrijver and sum-of-squares (SOS), also known as Lasserre-Parrillo SDPs, of any fixed degree, do *not* break the square root barrier in the context of ER graphs. The Lovász-Schrijver hierarchy, however, breaks the square root barrier in that context if the degree of hierarchy $t = t(n)$ is a slowly growing function of the number of vertices n (see also [FK03]).

An open problem proposed by Mixon and Bandeira is whether the SOS-4 relaxation of the Paley graph clique number breaks this barrier [Ban16]. This problem is premised on numerical experiments using the block-diagonal relaxations (which are stronger than the Lovász-Schrijver relaxations as noted previously, but weaker than the SOS relaxations) suggesting that these block-diagonal relaxations may break the square root barrier [GLV09] (see also [Gvo08, KM23]). More recent work provided numerical evidence that the SOS-4 relaxations of the Paley graph clique number may break the square root barrier, and proved that it can at best improve the exponent from 1/2 to 1/3 [KY23].

Perhaps surprisingly, our main result (Theorem 3.4) is that the block-diagonal relaxations of any constant degree do *not* break the square root barrier for the Paley graph clique number, in the sense that they all have the exponent of 1/2 (they may, however, improve the constant prefactor in front of $\sqrt{p}$). Since, as noted previously, the block diagonal relaxations are as strong as the Lovász-Schrijver relaxations, our result also applied to the latter relaxations. Accordingly, our lower bound parallels in the Paley graph setting the lower bound in [FK03] on the Lovász-Schrijver relaxations of the clique number of the ER random graphs.

Another line of work considered whether convex relaxations [Pas13, MMP19] and, more recently, spectral bounds [Kun24] for the clique number of *localizations* of the Paley graphs (which are defined below) may break the square root barrier. Our lower bounds (Corollary 3.10) cover block-diagonal relaxations of the clique number of localizations of the Paley graphs. Since the first and the weakest degree of the block-diagonal hierarchy is the Lovász theta function, which is equivalent to the sum of squares relaxation of degree 2 (SOS-2), our analysis shows that this relaxation of the clique number of a localization of the Paley graph of any constant degree does not break the square root barrier either (resolving the corresponding open question in Table 1 in [Kun24]). Our results, however, do not rule out the possibility that SOS-2, or a block-diagonal relaxation of higher constant degree, may break the square root bottleneck if the degree of localization $a = a(p)$ is a slowly growing function of p.

Our results are premised on a simplification of the positive semidefinite (PSD) constraints in the block diagonal program (Lemmas 3.1 and 3.2), which may be of interest independent from the Paley graph clique problem.

2 Background and notation

Notation In this work, p denotes a prime number $p \equiv 1 \bmod 4$, $\mathbb{F}_p$ denotes a finite field of order p, and $\mathbb{F}_p^\times$ denotes its group of units. An element y of $\mathbb{F}_p^\times$ is a *quadratic residue* if $y = x^2$ for some

$x \in \mathbb{F}_p^\times$ and is a *nonresidue* otherwise. We will denote the set of quadratic residues by $(\mathbb{F}_p^\times)^2$. We will represent $\mathbb{F}_p$ by the elements $\{0, 1, \dots, p-1\}$ of $\mathbb{Z}_p$.

For any graph $G = (V, E)$, the $\{0, 1\}$ adjacency matrix $\boldsymbol{A}_G$ has the following relationship to the $\{\pm 1\}$ adjacency matrix $\boldsymbol{U}_G$ (also called the *Seidel matrix*) :

$$\boldsymbol{A}_G = \frac{1}{2}(\boldsymbol{U}_G + \boldsymbol{J} - \boldsymbol{I})$$

where $\boldsymbol{J} = \mathbb{1}\mathbb{1}^\top$ is a matrix containing entries of all 1's. For any $S \subseteq V$, $\mathbb{1}_S$ denotes the indicator function of S being a clique in G.

For a diagonalizable matrix M, $\lambda_i(\boldsymbol{M})$ shall denote its i-th largest eigenvalue, and the spectrum of $\boldsymbol{M}$, or $\mathrm{spec}(\boldsymbol{M})$, shall denote the set of eigenvalues of $\boldsymbol{M}$.

On the other hand, $\mathbb{1}$ denotes the vector containing 1 in each entry. We denote the *power set* of V by $\mathcal{P}(V)$, and the subsets of V with at most t and exactly t elements by $\mathcal{P}_t(V) := \{I \in \mathcal{P}(V) \mid |I| \leq t\}$ and $\mathcal{P}_{=t}(V) := \{I \in \mathcal{P}(V) \mid |I| = t\}$ respectively. We will suppress the dependence on V in $\mathcal{P}_t(V)$ for brevity. Note that $\mathcal{P}_t$ contains the empty set $\varnothing$. For example, $\mathcal{P}_1 = \{\varnothing, \{i\},\ i \in V\}$. We denote the canonical basis vectors in $\mathbb{R}^{\mathcal{P}_1}$ by $e_\varnothing$ and e_i for $i \in V$. We also denote the set of all cliques of G by $\mathcal{K}(G)$ or $\mathcal{K}$ when the context is clear. Similarly, we let $\mathcal{K}_a(G)$ or $\mathcal{K}_a$ denote the set of cliques of size a in G.

Convex relaxations of the clique number Convex relaxations, including Lovász-Schrijver and sum-of-squares (SOS) SDP hierarchies defined below, are used to bound optimal values of integer programs. In the context of the clique problem, linear programming and SDP relaxations have been applied to many graphs of interest [Del72, DL98, Sch79], [Lov79, Sch05, BV08, dLV15].

For a graph G sampled from the ER graph distribution $\mathcal{G}(\frac{1}{2}, n)$, reference [FK03] showed that the value of the Lovász-Schrijver relaxation of degree t for $\omega(G)$ is approximately $\sqrt{n/2^t}$ by proving matching lower and upper bounds. To establish the lower bound, this reference introduced the so-called *Feige-Krauthgamer (FK) pseudomoments* (see Definition 3.3). Reference [BHK+19] established an $\Omega(n^{1/2-o(1)})$ lower bound on the SOS relaxations for $\omega(G)$ at any constant degree of the hierarchy, using the so-called *pseudocalibration* technique (see also [MPW15] applying the FK pseudomoments in a related setting). Recently, reference [Fau24] proved an $O(\sqrt{n \log n})$ bound on the expected value of the Lovász ϑ function on random circulant graphs.

In contrast, in the context of the Paley graphs of prime order or other SBGs, it has not been previously determined whether or not any given convex relaxation hierarchy leads to nontrivial upper bounds on $\omega(G)$, i.e., an upper bound that either breaks the square root barrier or reduces the prefactor in front of square root of the number of vertices to any constant strictly less than 1. Recent work [KY23] made progress in that direction: it provided numerical evidence that $SOS_4(G_p)$ breaks that barrier and proved that the value of these relaxations is lower bounded as $\Omega(p^{1/3})$ using the FK pseudomoments method.

Sum-of-squares relaxations The sum-of-squares (SOS) hierarchy, also known as the Lasserre-Parrilo hierarchy [Las01, Par03], of SDP relaxations has been used to analyze the Paley graph clique number. To define these relaxations, we first define the *moment matrix* $\boldsymbol{M}_t(y) \in \mathbb{R}^{\mathcal{P}_t \times \mathcal{P}_t}$: given a vector $y \in \mathbb{R}^{\mathcal{P}_{2t}}$, for $I, J \in \mathcal{P}_t$, $\boldsymbol{M}_t(y)_{IJ} := y_{I \cup J}$.[1] The *sum-of-squares* relaxation of $\omega(G)$, denoted

[1]This definition implies that $\boldsymbol{M}_t(y)_{IJ}$ depends only on $I \cup J$. Based on that fact, the moment matrices and the SOS_{2t} relaxation hierarchy can be defined without reference to y (see, e.g., [KY23]). However, we use the definition in the text accompanying this footnote to make the relationship between the SOS_{2t} and the block diagonal hierarchy discussed below more apparent.

by $SOS_{2t}(G)$, is defined as

$$SOS_{2t}(G) \;:=\; \max_{y \in \mathbb{R}^{\mathcal{P}_{2t}}} \quad \sum_{i \in V} y_{i,\varnothing} \qquad \text{s.t. } y \in Q_t(G)$$

where the constraint set (convex cone in $\mathbb{R}^{\mathcal{P}_{2t}}$) is

$$Q_t(G) = \{y \in \mathbb{R}^{\mathcal{P}_{2t}} \mid y_\varnothing = 1,\; y_{S,T} = 0 \;\; \forall S \cup T \notin \mathcal{K},\; \boldsymbol{M}_t(y) \succeq 0\}.$$

Lovász-Schrijver relaxations The *Lovász-Schrijver* relaxation of the clique number [LS90, GLV09], which we denote by $LS_t(G)$ for $t \geq 0$, is defined as:

$$LS_t(G) \;:=\; \max_{x \in \mathbb{R}^{\mathcal{P}_1}} \quad \sum_{i \in V} x_i \qquad \text{s.t. } x_\varnothing = 1,\; x \in N_+^t$$

where

$$N_+^0 := \{\; Ye_\varnothing \in \mathbb{R}^{\mathcal{P}_1} \mid Y \in M_{+,V}, Y_{ij} = 0 \text{ if } (i,j) \notin E \;\} \quad \text{and} \quad M_{+,V} := \{Y \in \mathbb{R}^{\mathcal{P}_1 \times \mathcal{P}_1} | Y \succeq 0, Y_{ii} = Y_{\varnothing i}\}$$

and the constraints of higher degree $t \geq 1$ (also convex cones in $\mathbb{R}^{\mathcal{P}_1}$) are defined iteratively by

$$N_+^t := \{Ye_\varnothing | Y \in M_{+,V}, Ye_i \in N_+^{t-1}, Y(e_\varnothing - e_i) \in N_+^{t-1}, \forall i \in V\}.$$

Note that $LS_0(G) = SOS_2(G)$, and both of these relaxations are equivalent to Lovász $\vartheta(\overline{G})$ [GL17]. For $t > 1$, the SOS_{2t} relaxations are stronger than the corresponding LS_{t-1} relaxations. This implies that in the context of the corresponding clique number relaxations $SOS_{2t}(G) \leq LS_{t-1}(G)$. However, the SOS hierarchy is more computationally expensive, and appears difficult to analyze, as can been seen in the $SOS_4(G_p)$ case considered in [KY23].

Block-diagonal hierarchy References [GLV09] and [Gvo08] introduced the block-diagonal hierarchy of SDPs, denoted by L^t, which are less computationally expensive than the SOS SDPs. These references also initiated the investigation of the L^t relaxations of independent sets of graphs, including the Paley graphs.

As noted previously, analyzing the clique number instead of the independence number makes the connection between relaxations and localizations more apparent (specifically, under the FK pseudomoment assumption, principal submatrices of the adjacency matrices of localizations emerge in Appendix D in the optimization constraints). Therefore, we define the L^t optimization problem with respect to the complement graph $\overline{G}$, which constitutes a relaxation of the clique number.

Let us first define a submatrix $\boldsymbol{M}(T; y)$ of the moment matrix $\boldsymbol{M}_t(y)$, defined above: for a subset $T \subseteq V$ of cardinality $t-1$ and vector $y \in \mathbb{R}^{\mathcal{P}_{t+1}}$ where the power set $\mathcal{P}_{t+1}$ is defined previously, let $\boldsymbol{M}(T; y) \in \mathbb{R}^{(n+1)\mathcal{P}_{t-1} \times (n+1)\mathcal{P}_{t-1}}$ denote the principal submatrix of the moment matrix $\boldsymbol{M}_t(y)$, whose rows and columns are indexed by

$$A(T) := \bigcup_{S \subseteq T} A_S, \text{ where } A_S := \{S\} \cup \{S \cup \{i\} \mid i \in V\}.$$

Here, we treat $A(T)$ as a multiset, and therefore we keep possible repeated occurrences of the same elements, e.g., S and $S \cup \{i\}$ if $i \in S$ in the indexing. The L^t relaxation of the clique number problem is given by

$$L^t(\overline{G}) := \begin{cases} \max \sum_{i \in V} y_{\{i\}} & \text{(1a)} \\ \text{s.t. } y \in \mathbb{R}^{\mathcal{P}t+1},\ y_{\varnothing} = 1 & \text{(1b)} \\ \quad y_{\{i,j\}} = 0 \ \forall\, i \not\sim j \text{ in } G & \text{(1c)} \\ \quad \boldsymbol{M}(T; y) \succeq 0 \text{ for all } T \in \mathcal{P}_{=t-1} & \text{(1d)} \end{cases}$$

Relationship between block-diagonal, SOS and Lovasz-Schrivjer hierarchies For $t = 1$, the $SOS_2(G)$ and $LS_0(G)$ constraint matrix $\boldsymbol{M}_1(y)$ matches the L^1 constraint matrix $\boldsymbol{M}(\varnothing, y)$, and therefore all three relaxations coincide: $L^1(\overline{G}) = SOS_2(G) = LS_1(G)$. When $t \geq 2$, for each $T \subset V$ of cardinality $t-1$, $\boldsymbol{M}(\varnothing, y)$ is a principal submatrix of $\boldsymbol{M}_t(y)$. If a moment matrix $\boldsymbol{M}_t(y)$ is PSD, then all of its principal submatrices, such as $\boldsymbol{M}(T; y)$, are also PSD. Therefore, L^t is a relaxation of SOS_{2t}, and the optimal $L^t(\overline{G})$ value bounds from above the optimal $SOS_{2t}(G)$ value. In the context of the Paley graph clique number $\omega(G_p)$, this relationship implies that the above-mentioned $\Omega(p^{1/3})$ lower bound on $SOS_4(G_p)$ in [KY23] also applies to $L^2(\overline{G}_p)$.

On the other hand, the relationship between $L^{t'}$ and SOS_{2t} when $t' > t$ does not appear to be well understood. In particular, as discussed, in the section entitled “Numerical experiments” below and shown in Figure 3, the numerical experiments for small p indicate that $L^3(G_p)$ provides a tighter upper bound on $\omega(G_p)$ than $SOS_4(G_p)$ when p between 137 and 197.

We note that the block-diagonal hierarchy is stronger than the Lovász-Schrijver hierarchy (see also Remark 3.5).

Remark 2.1. *As noted in [GLV09], the feasible region of $L^t(G)$ is contained in the feasible region of Lovász-Schrijver relaxation N_+^{t-1} for all $t \geq 1$. Therefore, any lower bound on $L^t(G)$ also lower bounds on $LS_{t-1}(G)$.*

Recent experiments in [KM23] numerically showed that $L^2(\overline{G}_p)$ grows as approximately $O(p^{0.456})$ for $p < 1000$. Since the exponent in this scaling estimate is fairly close to $1/2$ and the range of values of p is relatively small, these experiments does not make it clear, even empirically, whether L^2 breaks the $\sqrt{p}$ barrier.

3 Main results

Our main lower bound is based on two preliminary results: first, we simplify each PSD constraint (3d) of the $L^t(\overline{G})$ program for graph G of n vertices. This result generalizes the previous simplifications of L^2 and L^3 programs, which made it possible to use matrices of size smaller than $(n+1) \times (n+1)$ in the optimization constraints (see dimensions of the constraints in Section 4 in [GLV09], see also Section IV in [KM23] for additional details about the L^2 constraint simplifications). Second, we set forth further simplified constraints which define a feasible set that satisfies (3d); a point in that set leads to our main lower bound. These simplifications are not specific to the Paley graph clique number relaxation or the FK pseudomoments. As such, our simplifications could be of interest in other contexts where the block-diagonal hierarchy may be used.

Simplification of the block-diagonal constraints The moment matrices $\boldsymbol{M}(T; y)$ can be block-diagonalized by a congruence transformation using the zeta matrices of the $\mathcal{P}_{t-1}$ lattice.

For a subset S of T, let $\boldsymbol{A}_S(y) \in \mathbb{R}^{(n+1)\times(n+1)}$ denote the principal submatrix of $\boldsymbol{M}(T;y)$ indexed by the set $\boldsymbol{A}_S$ with entries given by:

$$\boldsymbol{A}_S(y)_{\varnothing,\varnothing} = y_S, \quad \boldsymbol{A}_S(y)_{\varnothing,i} = y_{S\cup\{i\}}, \quad \boldsymbol{A}_S(y)_{i,j} = y_{S\cup\{i,j\}} \quad (i,j \in V, \text{ where } |V| = n).$$

By Lemma 2.2 in [Gvo08], $\boldsymbol{M}(T;y)$ is PSD if and only if for all subsets S of T the matrix

$$\boldsymbol{A}(S,T)(y) := \sum_{S' : S\subseteq S'\subseteq T} (-1)^{|S'\setminus S|} \boldsymbol{A}_{S'}(y) \tag{2}$$

are PSD. The L^t relaxation of the clique number problem (1) is equivalently given by

$$L^t(\overline{G}) := \begin{cases} \max \sum_{i\in V} y_{\{i\}} & (3a)\\ \text{s.t. } y \in \mathbb{R}^{\mathcal{P}t+1},\ y_\varnothing = 1 & (3b)\\ \qquad y_{\{i,j\}} = 0\ \forall\ i \not\sim j \text{ in } G & (3c)\\ \qquad \boldsymbol{A}(S,T)(y) \succeq 0 \text{ for all } S \subseteq T \text{ and } T \in \mathcal{P}_{=t-1} & (3d) \end{cases}$$

We can remove the rows and columns of each $\boldsymbol{A}(S,T)(y)$ corresponding to the vertices in T, as well as to each vertex $i \in V(G) \setminus T$ such that $S \cup \{i\} \notin \mathcal{K}(G)$. Let $\tilde{\boldsymbol{A}}(S,T)(y)$ (resp., $\tilde{\boldsymbol{M}}(T;y)$) denote the submatrix of $\boldsymbol{A}(S,T)(y)$ (resp., $\boldsymbol{M}(T;y)$) after removing the rows and columns corresponding to all such vertices. Similarly, we let $\tilde{\boldsymbol{A}}_{S'}(y)$ denote the submatrix of $\boldsymbol{A}_{S'}(y)$ in (2) with the rows and columns corresponding to the above-mentioned vertices removed:

$$\tilde{\boldsymbol{A}}_{S'}(y)_{\varnothing,\varnothing} = y_S,\ \tilde{\boldsymbol{A}}_{S'}(y)_{\varnothing,i} = y_{S'\cup\{i\}},\ \tilde{\boldsymbol{A}}_{S'}(y)_{i,j} = y_{S'\cup\{i,j\}}$$

where the rows and columns of $\tilde{\boldsymbol{A}}_{S'}$ are indexed by vertices $i \in V \setminus T$, such that $S' \cup \{i\} \in \mathcal{K}(G)$.

Lemma 3.1. *For any undirected graph G, the constraint (3d) is equivalent to*

$$\tilde{\boldsymbol{A}}(S,T)(y) \succeq 0 \textit{ for all } S \subseteq T \textit{ where } T \in \mathcal{P}_{=t-1} \textit{ and } S \in \mathcal{K}(G) \tag{4}$$

where

$$\tilde{\boldsymbol{A}}(S,T)(y) := \sum_{\substack{S' : S\subseteq S'\subseteq T\\ S'\in\mathcal{K}(G)}} (-1)^{|S'\setminus S|} \tilde{\boldsymbol{A}}_{S'}(y). \tag{5}$$

Additional simplifications for lower bound purposes When $t \geq 3$, the following lemma shows that further simplified constraints define a feasible region of the original block-diagonal program.

Lemma 3.2. *For any undirected graph G, the constraint (4) is satisfied if*

$$\hat{\boldsymbol{A}}(S,T)(y) \succeq 0 \textit{ for all } S \subseteq T \textit{ such that } S \in \mathcal{K}(G) \textit{ and for all } T \in \mathcal{P}_{=t-1}, \tag{6}$$

where

$$\hat{\boldsymbol{A}}(S,T)(y) := \begin{cases} \tilde{\boldsymbol{A}}_S(y) - \sum_{\substack{k\in T\setminus S\\ S\cup\{k\}\in\mathcal{K}(G)}} \tilde{\boldsymbol{A}}_{S\cup\{k\}}(y) & \textit{if } S \textit{ is a proper subset } S \subsetneq T,\\ \tilde{\boldsymbol{A}}_T(y) & \textit{if } S = T. \end{cases} \tag{7}$$

FK pseudomoments According to Lemma 3.1 in [GLV09], the constraint (3c) implies that $y_S = 0$ for any subset $S \subseteq V$ with $|S| \leq t+1$ containing non-edge in G, and therefore have $(\boldsymbol{A}_{S'})_{i,j} = 0$ if $S' \cup \{i,j\} \notin \mathcal{K}(G)$. As a result, we may adopt the following *Feige-Krauthgamer (FK) pseudomoment assumptions* to reduce the number of optimization variables in SDP hierarchies such as SOS_{2t} and L^t with moment-type matrices. Following the convention in [KY23], we introduce the notation $L^t_{FK}(\overline{G})$ to indicate the program $L^t(\overline{G})$ with this additional constraint.

Definition 3.3. *Given a graph $G = (V, E)$, $y \in \mathbb{R}^{\mathcal{P}_{t+1}}$ satisfies the* FK pseudomoment assumption *if there exists a sequence $1 = \alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_t, \alpha_{t+1} \in \mathbb{R}$ such that:*

$$y_K = \begin{cases} \alpha_{|K|}, & \text{if } K \in \mathcal{P}_{t+1} \text{ is a clique in } G \\ 0, & \text{otherwise.} \end{cases}$$

For $m = |S|$, and any $i, j \in \{\varnothing, V(G) \setminus T\}$ and $k \in T \setminus S$ we have

$$|S \cup \{k,i\}| = m+1 \quad \text{and} \quad |S \cup \{k,i,j\}| = \begin{cases} m+3 & \text{if } i \neq j \\ m+2 & \text{if } i = j. \end{cases}$$

Hence, according to the FK pseudomoment assumption (Definition 3.3), we denote by $\hat{\boldsymbol{A}}(S,T)(\alpha)$, $\tilde{\boldsymbol{A}}_S(\alpha)$, $\tilde{\boldsymbol{A}}_{S\cup\{k\}}(\alpha)$ the matrices $\hat{\boldsymbol{A}}(S,T)(y)$ and $\tilde{\boldsymbol{A}}_S(y)$, $\tilde{\boldsymbol{A}}_{S\cup\{k\}}(y)$ with $y_S = \alpha_m$, $y_{S\cup\{k\}} = \alpha_{m+1}$, $y_{S\cup\{k,i\}} = \alpha_{m+2}$ if $S \cup \{k,i\} \in \mathcal{K}(G)$ and 0 otherwise, and similarly $y_{S\cup\{k,i,j\}} = \alpha_{m+3}$ if $S \cup \{k,i,j\} \in \mathcal{K}(G)$ and 0 otherwise. Since $S \cup \{k\} \subseteq T$, the set of indices i, j does not include any elements of $S \cup \{k\}$. Accordingly,

$$\tilde{\boldsymbol{A}}_S(\alpha) = \left(\begin{array}{c|c} \alpha_m & \alpha_{m+1}\mathbb{1}^\top \\ \hline \alpha_{m+1}\mathbb{1} & \alpha_{m+1}\boldsymbol{I} + \alpha_{m+2}\boldsymbol{M}_S \end{array}\right), \text{ and } \quad \tilde{\boldsymbol{A}}_{S\cup\{k\}}(\alpha) = \left(\begin{array}{c|c} \alpha_{m+1} & \alpha_{m+2}v_{S\cup\{k\}}^\top \\ \hline \alpha_{m+2}v_{S\cup\{k\}} & \begin{array}{c}\alpha_{m+2}\,\mathsf{diag}(v_{S\cup\{k\}}) \\ +\alpha_{m+3}\boldsymbol{M}_{S\cup\{k\}}\end{array} \end{array}\right)$$

where we denote the vector of all ones by $\mathbb{1}$, $(v_{S\cup\{k\}})_i := \mathbb{1}_{S\cup\{k,i\}}$ for and the indicator function of a clique by $\mathbb{1}_S$, $(\boldsymbol{M}_S)_{i,j} := \mathbb{1}_{S\cup\{i,j\}}$ if $i \neq j$ and 0 if $i = j$ for any $i, j \in V(G) \setminus T$, and $\boldsymbol{M}_{S\cup\{k\}}$ is defined similarly. The definition $\tilde{\boldsymbol{A}}_S(\alpha)$ and $\boldsymbol{M}_S$ can be extended to $\tilde{\boldsymbol{A}}_T(\alpha)$ and $\boldsymbol{M}_T$ accordingly. Notice that $\tilde{\boldsymbol{A}}_S(\alpha)$ and $\tilde{\boldsymbol{A}}_{S\cup\{k\}}(\alpha)$ have the same dimension of $|V(G) \setminus T| + 1 = n - t + 2$ (where $n = |V(G)|$). In particular, the length (denoted n_S) of the $\{0,1\}$-vector $v_{S\cup\{k\}}$ is given by

$$n_S = |V(G_S)| - t + 1. \tag{8}$$

In other words, $\boldsymbol{M}_S$ and $\boldsymbol{M}_{S\cup\{k\}}$ are the principal submatrices constructed by removing the rows and columns corresponding to $\varnothing$ from $\tilde{\boldsymbol{A}}_S(\alpha)$ and $\tilde{\boldsymbol{A}}_S(\alpha)$, respectively. Note also that $\boldsymbol{M}_S$ and $\boldsymbol{M}_{S\cup\{k\}}$ are principal submatrices of the $\{0,1\}$ adjacency matrices $\boldsymbol{A}_G$ of G and $\boldsymbol{A}_{G_{S\cup\{k\}}}$ of degree 1 localization of G respectively (the latter matrix is padded with rows and columns of all zeros in the positions of each i such that $S \cup \{k,i\} \notin \mathcal{K}(G)$). This observation facilitates our subsequent analysis of the FK pseudomoments using Schur complements, and also makes transparent the connection between the constraints in the block diagonal program and adjacency matrices of the graph and its localizations.

Lower bounds on the block-diagonal relaxations of the Paley graph clique problem Analysis of the FK pseudomoments now leads to our main result: the block-diagonal relaxation of any constant degree for the Paley graph clique number does not break the $\sqrt{p}$ barrier, resolving the corresponding open question raised by the numerical results in [Gvo08, GLV09, KM23]. As noted previously, our lower bound also applies to the L^t relaxation of the clique number of any other SBGs, which include all symmetric conference graphs.

Theorem 3.4. *Let G_n be a family of arbitrary undirected graphs parametrized by their size n with the spectrum of its Seidel matrix supported on $[-\rho\sqrt{n}, \rho\sqrt{n}]$ for some $\rho > 0$ independent of n. For any level of hierarchy $t \geq 1$, the value $L^t(\overline{G}_n)$ of the block-diagonal relaxation for the clique number satisfies*

$$L^t(\overline{G}_n) \geq \frac{\sqrt{n}}{\mathsf{max}(\rho, 1)2^{t-1}} + O\left(\frac{1}{2^t}\right). \tag{9}$$

as n gets large.

Remark 3.5. *By Remark 2.1, the lower bound (9) applies respectively to the value of $LS_{t-1}(G)$ of the Lovász-Shrijver relaxation.*

To apply the preceding theorem to the Paley graph G_p we note that the adjacency matrix $\boldsymbol{A}_{G_p}$ and the Seidel matrix $\boldsymbol{U}_{G_p}$ have the following spectra, e.g., Prop. 2.5 in [KY23].

Proposition 3.6. *For any prime $p = 1 \mod 4$, the spectra of the $\{0, 1\}$ adjacency and the Seidel matrices $\boldsymbol{A}_{G_p}$ and and $\boldsymbol{U}_{G_p}$ of the Paley graph G_p are*

$$spec(\boldsymbol{A}_{G_p}) = \{\frac{p-1}{2}, \underbrace{\frac{\sqrt{p}-1}{2}, \ldots, \frac{\sqrt{p}-1}{2}}_{\frac{p-1}{2} \text{ times}}, \underbrace{\frac{-1-\sqrt{p}}{2}, \ldots, \frac{-1-\sqrt{p}}{2}}_{\frac{p-1}{2} \text{ times}}\} \tag{10}$$

$$spec(\boldsymbol{U}_{G_p}) = \{0, \underbrace{\sqrt{p}, \ldots, \sqrt{p}}_{\frac{p-1}{2} \text{ times}}, \underbrace{-\sqrt{p}, \ldots, -\sqrt{p}}_{\frac{p-1}{2} \text{ times}}\}. \tag{11}$$

and $\mathbb{1}$ is the eigenvector corresponding to the $\frac{p-1}{2}$ and 0 eigenvalues of $\boldsymbol{A}_{G_p}$ and $\boldsymbol{U}_{G_p}$, respectively.

With the above proposition, our theorem immediately implies a lower-bound on the $L^t(G_p)$.

Corollary 3.7. *For any level of hierarchy $t \geq 1$, the optimal value $L^t(\overline{G}_p)$ of the block-diagonal relaxation for the clique number of Paley graph G_p with $p \equiv 1 \pmod 4$ satisfies*

$$L^t(\overline{G}_p) \geq \frac{\sqrt{p}}{2^{t-1}} + O\left(\frac{1}{2^t}\right) \tag{12}$$

as p gets large.

For $t = 1$, the classical result $\vartheta(G_p) = \sqrt{p}$ demonstrates the optimality of our lower bound.

Localization lower bound and relaxation-localization tradeoff Localization is another technique used to strengthen convex relaxations [Pas13, MMP19] and, more recently, spectral bounds on the clique number of the Paley graphs [Kun24]. We shall use the following definition compatible with the clique number problem.

Definition 3.8. *Given a set of vertices $X \subset V(G)$, the* localization G_X *of degree $|X|$ is a subgraph of G induced on the vertices adjacent to all vertices of X (excluding the vertices in X).*

Our lower bounds generalize to the block-diagonal relaxations of the clique number of a Paley graph localization; in this section, we show that such relaxations also do not break the $\sqrt{p}$ barrier for any fixed degree of hierarchy and any fixed degree of localization, resolving the corresponding open question in Table 1 of [Kun24].

The clique numbers of a graph and its localizations of a given degree have following basic relationship: for any $a \leq \omega(G)$,

$$\omega(G) = a + \max_{K \in \mathcal{K}_a(G)} \omega(G_K).$$

If a function $f(G_K)$ bounds $\omega(G_K)$ from above, this leads to the following upper bound:

$$\omega(G) \leq a + \max_{K \in \mathcal{K}_a(G)} f(G_K).$$

In the context of the Paley graph clique number, when $a = 1$, it is sufficient to consider $K = \{0\}$ for purposes of this maximization since G_p is vertex transitive. Reference [Pas13] observed that $G_{p,\{0\}}$ is a circulant graph, and therefore $\vartheta(G_{p,\{0\}})$ reduces to a linear program. This approach facilitated the computation of the value of this relaxation for up to $p < 20000$ and led to the observation that this value is around $\sqrt{(p-1)/2}$, which matches the leading order term of the subsequently established $HP(G_p)$ upper bound. Similarly, reference [MMP19] computed $\vartheta(G_{p,\{0\}})$, strengthened by Schrijver's entrywise non-negativity condition [Sch79], numerically for $p < 3000$ and observed that the optimal values of this problem usually coincided, and sometimes improve upon, the $HP(G_p)$ upper bound.

Reference [Kun24] proposed a program towards improving the upper bound on $\omega(G_p)$ by combining localization with spectral methods. One of the main conjectures (Conjecture 1.9 in that reference) is that minimum eigenvalue of the adjacency matrix of each Paley graph localization converges to the left edge of support of the appropriately rescaled Kesten-MacKay measure (KM) as $p \to \infty$. Assuming this conjecture holds, reference [Kun24] proved that

$$\omega(G_p) \leq \frac{\sqrt{2^a - 1}}{2^{a-1}} \sqrt{p} + o(\sqrt{p}). \tag{13}$$

This reference also conjectured that higher degree of localization can be 'traded' for weaker convex relaxations while obtaining comparable bounds on the clique number (*relaxation-localization trade-off*). Based on numerical evidence, this reference also conjectured that $\vartheta(G_{p,\{0,1\}})$ of Paley graph degree 2 localization leads to $(1/\sqrt{2} - \epsilon)\sqrt{p}$ upper bound, improving on the upper bound of Hanson and Petridis.

Our next result reveals a basic lower bound on the block-diagonal relaxation of a localization of an arbitrary graph exploiting the fact that we can construct a set of moment matrices feasible for the $L^t(\overline{G}_K)$ program as principal submatrices of the moment matrices that are feasible (optimal) for the $L^t(\overline{G})$ program.

Proposition 3.9. *Let G be a vertex transitive graph. For any level of hierarchy $t \geq 1$ and any clique K in G of size a, we have*

$$L^t(\overline{G}_K) \geq \frac{|V(G_K)|}{|V(G)|} L^t(\overline{G}) \tag{14}$$

Since the Paley graph G_p is vertex transitive and the size of its localizations can be estimated, combining (14) and our estimate of $L^t(G_p)$ in (12) gives an estimate on $L^t(\overline{G}_{p,K})$.

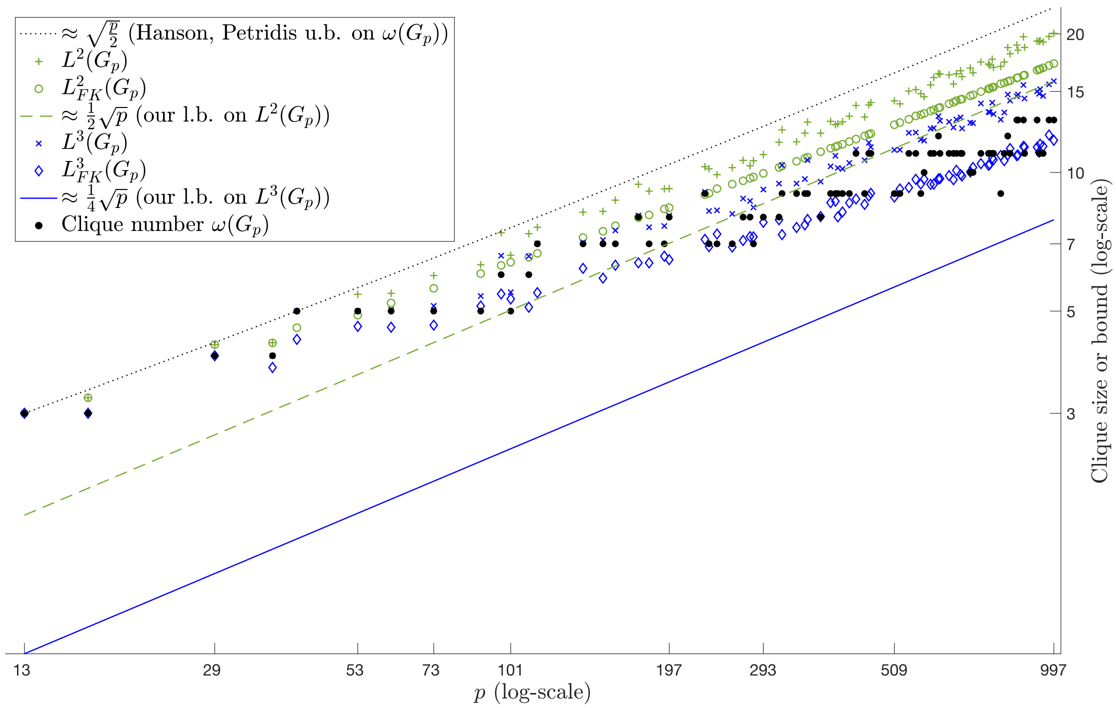

Figure 1: The $L^2(G_p)$ and $L^3(G_p)$ values for $p \le 809$ computed in [GLV09, Gvo08], $L^2(G_p)$ values for $809 < p < 1000$ computed in [KM23], $L^3(G_p)$ values for $809 < p < 1000$ computed in this work, along with the $L^2_{FK}(G_p)$ and $L^3_{FK}(G_p)$ values, are plotted relative to the corresponding lower bounds in Theorem 3.4. The values of $\omega(G_p)$ were obtained from [She23] and the Hanson, Petridis upper bound on $\omega(G_p)$ is $(\sqrt{2p-1}+1)/2$ established in [HP21].

Corollary 3.10. *For any level of hierarchy $t \ge 1$, for any clique K of size a in Paley graphs G_p, the value $L^t(\overline{G}_{p,K})$ of the block-diagonal relaxation for the clique number of the localization $G_{p,K}$ satisfies*

$$L^t(\overline{G}_{p,K}) \ge \frac{\sqrt{p}}{2^{a+t-1}} + O\left(\frac{a}{2^t}\right). \tag{15}$$

as p gets large.

For $t = 1, a = 0$, the classical result $\vartheta(G_p) = \sqrt{p}$ demonstrates the optimality of our lower bound. In particular, since $L^1 = SOS_2 = \vartheta$, Corollary 3.10 shows that SOS_2 does not break the $\sqrt{p}$ barrier for any constant degree of localization of G_p, resolving the corresponding open question in Table 1 in reference [Kun24]. Also, since the lower bound is a function of $a + t$, it is consistent with the relaxation-localization trade-off conjectured in that reference.

Finally, our results indicate that any fixed level of the Lovász-Shrijver relaxation for the Paley graph localization clique number also does not break the $\sqrt{p}$ barrier (see also Remark 3.5).

Numerical experiments Let $L^t_{FK}(\overline{G})$ denote the variant of (3) where the optimization variable y is restricted according to the FK pseudomoment assumption (see Definition 3.3). We replicated the $L^3(G_p)$ computations reported in [GLV09, Gvo08] using Matlab/CVX for primes $p \le 809$ and extended them for all $p < 1000$.[2]

[2]The corresponding code is available on the last author's website.

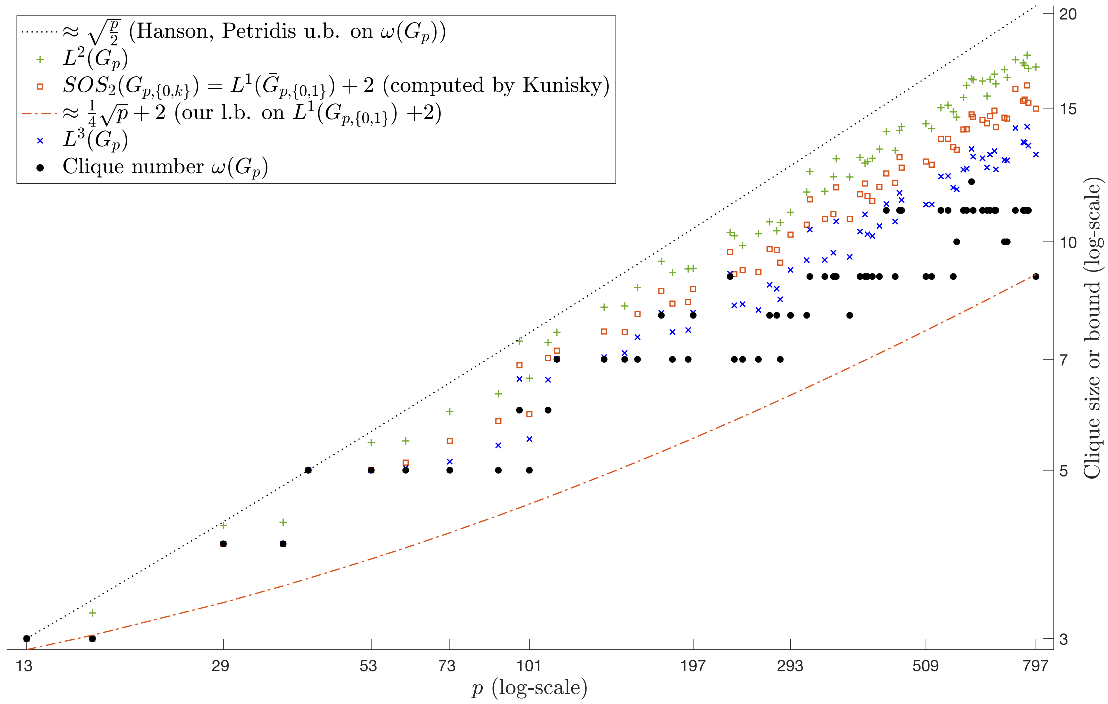

Figure 2: The $L^2(G_p)$ and $L^3(G_p)$ values for $p \leq 797$ computed in [GLV09, Gvo08], along with the $L^1(\overline{G}_{p,\{0,1\}}) = SOS_2(G_{p,\{0,k\}})$ values determined in [Kun24], relative to the lower bound for $L^1(\overline{G}_{p,\{0,1\}})$ in Corollary 3.10. The values of $\omega(G_p)$ were obtained from [She23] and the Hanson, Petridis upper bound on $\omega(G_p)$ is $(\sqrt{2p-1}+1)/2$ established in [HP21].

Figure 1 plots the $L^2(G_p)$ and $L^3(G_p)$ values for $p \leq 809$ determined in [GLV09, Gvo08], the $L^2(G_p)$ values for $809 < p < 1000$ reported in [KM23], $L^3(G_p)$ values for $809 < p < 1000$ computed in this work, along with the $L^2_{FK}(G_p)$ and $L^3_{FK}(G_p)$ values, relative to the corresponding lower bounds in Theorem 3.4. Figure 2 plots the $L^2(G_p)$ and $L^3(G_p)$ values for $p \leq 797$ determined in [GLV09, Gvo08], along with the $L^1(\overline{G}_{p,\{0,1\}}) = SOS_2(G_{p,\{0,k\}})$ values determined in [Kun24], relative to the lower bound corresponding to $L^1(\overline{G}_{p,\{0,1\}})$ in Corollary 3.10. Our figures also include the values of $\omega(G_p)$ obtained from [She23] and the Hanson, Petridis upper bound on $\omega(G_p)$ established in [HP21].

Finally, Figure 3 depicts the $L^3(G_p)$ values computed in [GLV09, Gvo08] and $L^3_{FK}(G_p)$ values computed in this work for $p \leq 241$, relative to the $SOS_2(G_p)$ values for $p \leq 197$ and $SOS_{2FK}(G_p)$ values $p \leq 241$ determined in [KY23]. Perhaps surprisingly, $L^3(G_p)$ provides a tighter upper bound on $\omega(G_p)$ than $SOS_4(G_p)$ when p between 137 and 197.

4 Conclusion and future directions

We have established that the block-diagonal L^t relaxations, and therefore the Lovász-Schrijver relaxations, of fixed level of hierarchy do not break the $\sqrt{p}$ barrier with respect to the clique number of the Paley graphs and their localizations. Our results, however, leave open the possibility that L^1/SOS-2, or a block-diagonal relaxation of some higher constant degree, may break the $\sqrt{p}$ bottleneck if the degree of localization a is a slowly growing function of p. It might also be possible to

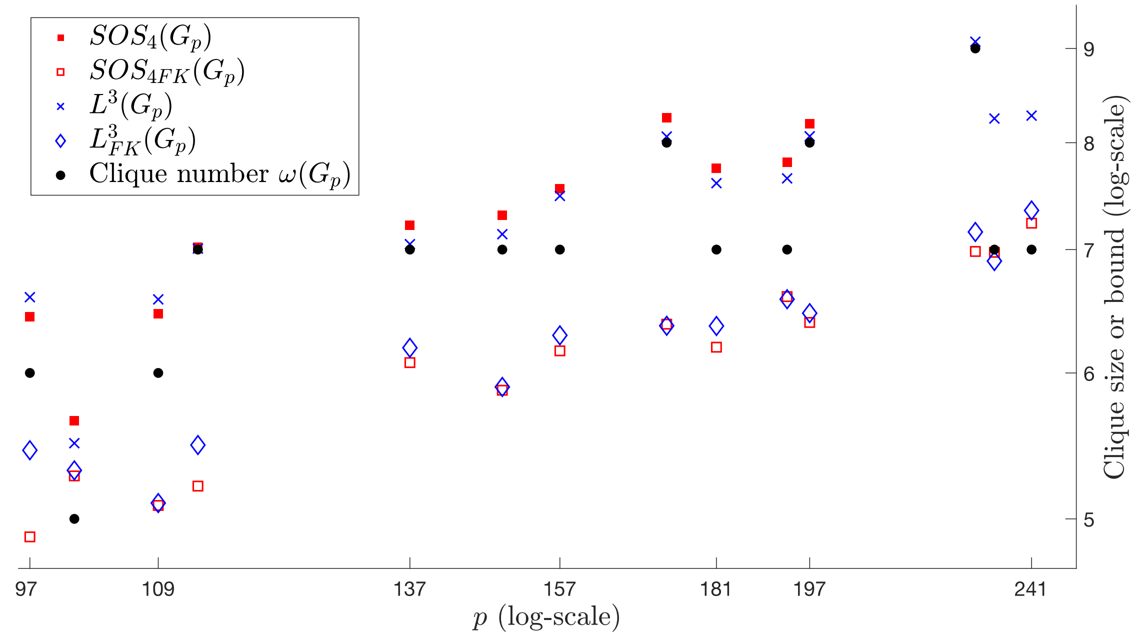

Figure 3: The $L^3(G_p)$ values computed in [GLV09, Gvo08] and $L^3_{FK}(G_p)$ values computed in this work for $p \le 241$, relative to the $SOS_4(G_p)$ values for $p \le 197$ and $SOS_{4FK}(G_p)$ values $p \le 241$ determined in [KY23]. The values of $\omega(G_p)$ were again obtained from [She23].

improve the constant prefactor in our lower bounds, especially if the bounds on the Seidel spectrum of localizations can be improved beyond the estimate given by the eigenvalue interlacing theorem. We leave these interesting questions to further work.

Our results also leave open the possibility that the block-diagonal relaxations may improve the constant prefactor of the Hanson-Petridis upper bound. In particular, since the adjacency matrices $\boldsymbol{A}_{G_p}$ and $\boldsymbol{A}_{G_{p\{0\}}}$, as well as the matrices of the indicators of orbits of triangles of the form $\{0, \alpha, \beta\}$ in G_p are circulant, Fourier-based methods may facilitate the analysis of $L^2(G_p)$ upper bounds; a possible next step in this direction would be to consider the corresponding dual programs.

Finally, the numerical experiments raise the question of whether $L^{t'}$ may provide a stronger relaxation than SOS_{2t} when $t' > t$ is sufficiently large. In this regard, understanding the relationship between $L^3(G_p)$ and $SOS_4(G_p)$ when p is large appears to be a natural direction to explore initially.

Our broad contribution is a methodological advance: we make progress towards extending methods available to analyze convex relaxations of combinatorial optimization problems to the block-diagonal relaxations, an SDP hierarchy that remains relatively unexplored. Accordingly, we hope that block-diagonal relaxations will lead to new results in the context of other combinatorial problems/problems over graphs.

References

[AAM15] Hamed Bagh-Sheikhi Arash Amini and Farokh Marvasti. From Paley graphs to deterministic sensing matrices with real-valued Gramians. In *2015 International Conference on Sampling Theory and Applications (SampTA)*, 2015.

[Ban16] Afonso S. Bandeira. Ten lectures and forty-two open problems in the mathematics of data science, Oct 2016. URL: https://people.math.ethz.ch/~abandeira/TenLecturesFortyTwoProblems.pdf.

[BFMM16] Afonso S. Bandeira, Matthew Fickus, Dustin G. Mixon, and Joel Moreira. Derandomizing restricted isometries via the Legendre symbol. *Constr. Approx.*, 43:409–424, 2016.

[BFMW13] Afonso S. Bandeira, Matthew Fickus, Dustin G. Mixon, and Percy Wong. The road to deterministic matrices with the restricted isometry property. *J. Fourier Anal. Appl.*, 19:1123–1149, 2013.

[BHK+19] Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh K Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735, 2019.

[BMM16] Afonso S. Bandeira, Dustin G. Mixon, and Joel Moreira. A conditional construction of restricted isometries. *International Mathematics Research Notices*, 2017:2:372–381, 2016.

[BMR14] C. Bachoc, M. Matolcsi, and I. Z. Ruzsa. Squares and difference sets in finite fields. *Integers: Electronic Journal of Combinatorial Number. Theory, Vol 13.*, 2014.

[BSW21] Daniel Di Benedetto, József Solymosi, and Ethan P. White. On the directions determined by a cartesian product in an affine galois plane. *Combinatorica*, 41:755– 763, 2021. doi:10.1007/s00493-020-4516-z.

[BV08] Christine Bachoc and Frank Vallentin. New upper bounds for kissing numbers from semidefinite programming. *Journal of the American Mathematical Society*, 21(3):909–924, 2008.

[CGW89] Fan R. K. Chung, Ronald L. Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica, 9(4):345–362*, 1989.

[Del72] Philippe Delsarte. Bounds for unrestricted codes, by linear programming. In *Philips Res. Rep, 27*, page 272–289, 1972.

[DL98] P. Delsarte and V.I. Levenshtein. Association schemes and coding theory. *IEEE Transactions on Information Theory*, 44(6):2477–2504, 1998. doi:10.1109/18.720545.

[dLV15] David de Laat and Frank Vallentin. A semidefinite programming hierarchy for packing problems in discrete geometry. *Mathematical Programming*, 151(2):529–553, 2015.

[Erd47] P. Erdös. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292 – 294, 1947.

[Exo23] Geoffrey Exoo. Independence numbers for Paley graphs, accessed January 11, 2023. URL: http://isu.indstate.edu/ge/PALEY/index.html.

[Fau24] Ulysse Faure. On the lovasz theta number for random circulant graphs. Master's thesis, ETH Zurich, 2024. URL: https://www.research-collection.ethz.ch/handle/20.500.11850/694627.

[FK03] Uriel Feige and Robert Krauthgamer. The probable value of the lovász-schrijver relaxations for maximum independent set. *SIAM J. Comput.*, 32(2):345–370, feb 2003. doi:10.1137/S009753970240118X.

[GL05] N. Gvozdenovic and M. Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. *Proceedings of 11th International IPCO Conference, Lecture Notes in Computer Science, 3509:136–151*, 2005.

[GL17] Laura Galli and Adam N. Letchford. On the Lovász theta function and some variants. *Discrete Optimization*, 25:159–174, 2017. doi:https://doi.org/10.1016/j.disopt.2017.04.001.

[GLV09] N. Gvozdenovic, M. Laurent, and F. Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. *Operations Research Letters 37:27-31*, 2009.

[GR90] S. W. Graham and C. J. Ringrose. *Lower Bounds for Least Quadratic Non-Residues*, pages 269–309. Birkhäuser Boston, Boston, MA, 1990. doi:10.1007/978-1-4612-3464-7_18.

[Gvo08] N. Gvozdenovic. *Approximating the stability number and the chromatic number of a graph via semidefinite programming.* PhD thesis, University of Amsterdam, 2008.

[Hås99] Johan Håstad. Clique is hard to approximate within n1-epsilon. *Acta Mathematica*, 182(1):105 – 142, 1999. doi:10.1007/BF02392825.

[HP21] Brandon Hanson and Giorgis Petridis. Refined estimates concerning sumsets contained in the roots of unity. In *Proceedings of the London Mathematical Society*, volume 122(3), pages 353–358, 2021.

[Kar72] Richard M. Karp. *Reducibility among Combinatorial Problems*, pages 85–103. Springer US, Boston, MA, 1972. doi:10.1007/978-1-4684-2001-2_9.

[KM23] Vladimir A. Kobzar and Krishnan Mody. Revisiting block-diagonal sdp relaxations for the clique number of the paley graphs. In *2023 International Conference on Sampling Theory and Applications (SampTA)*, pages 1–5, 2023. doi:10.1109/SampTA59647.2023.10301406.

[KPB19] Alihan Kaplan, Volker Pohl, and Holger Boche. Deterministic matrices with a restricted isometry property for partially structured sparse signals. In *13th International conference on Sampling Theory and Applications (SampTA)*, 2019.

[KS06] Michael Krivelevich and Benny Sudakov. *Pseudo-random Graphs*, pages 199–262. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. doi:10.1007/978-3-540-32439-3_10.

[Kun24] Dmitriy Kunisky. Spectral pseudorandomness and the road to improved clique number bounds for Paley graphs. *Experimental Mathematics*, pages 1–28, 2024. doi:10.1080/10586458.2024.2400182.

[KY23] Dmitriy Kunisky and Xifan Yu. A Degree 4 Sum-Of-Squares Lower Bound for the Clique Number of the Paley Graph. In Amnon Ta-Shma, editor, *38th Computational Complexity Conference (CCC 2023)*, volume 264 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:25, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CCC.2023.30, doi:10.4230/LIPIcs.CCC.2023.30.

[Las01] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. arXiv:https://doi.org/10.1137/S1052623400366802, doi:10.1137/S1052623400366802.

[Lau02] Monique Laurent. A comparison of the sherali-adams, lovász-schrijver and lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28, 08 2002. doi:10.1287/moor.28.3.470.16391.

[Lau07] Monique Laurent. Strengthened semidefinite programming bounds for codes. *Math. Program.*, 109(2–3):239–261, March 2007.

[Lov79] L. Lovasz. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985.

[LS90] Lovász László and A. Schrijver. Cones of matrices and set-functions and 0–1 optimization. *Siam Journal on Optimization - SIAMJO*, 1, 01 1990. doi:10.1137/0801013.

[LS91] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0–1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991. arXiv:https://doi.org/10.1137/0801013, doi:10.1137/0801013.

[Mix15] Dustin G. Mixon. Explicit matrices with the restricted isometry property: Breaking the square-root bottleneck. In Holger Boche, Robert Calderbank, Gitta Kutyniok, and Jan Vybíral, editors, *Compressed Sensing and its Applications: MATHEON Workshop 2013*, pages 389–417, Cham, 2015. Springer International Publishing.

[MMP19] Mark Magsino, Dustin G Mixon, and Hans Parshall. Linear programming bounds for cliques in paley graphs. In *Wavelets and Sparsity XVIII*, volume 11138, 2019.

[MPW15] Raghu Meka, Aaron Potechin, and Avi Wigderson. Sum-of-squares lower bounds for planted clique. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15, page 87–96, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2746539.2746600.

[Par03] Pablo Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming, Series B*, 96, 03 2003. doi:10.1007/s10107-003-0387-5.

[Pas13] N. Passuello. Semidefinite programming in combinatorial optimization with applications to coding theory and geometry, 2013. Ph.D. Thesis, Université Sciences et Technologies – Bordeaux I.

[Sch79] Alexander Schrijver. A comparison of the Delsarte and Lovasz bounds. *IEEE Trans. Inform. Theory*, 25:4:425–429, 1979.

[Sch05] Alexander Schrijver. New code upper bounds from the terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory*, 51(8):2859–2866, 2005.

[She23] J. B. Shearer. Independence numbers for Paley graphs, accessed January 11, 2023. URL: https://web.archive.org/web/20000815214259/http://www.research.ibm.com/people/s/shearer/indpal.html.

[Tao08] T. Tao. *Structure and Randomness: pages from year one of a mathematical blog*. Amer. Math. Soc., 2008. URL: http://terrytao.wordpress.com/2007/07/02/open-question-deterministic-uup-matrices/.

A Proof of Lemma 3.1

Consider any vertex $x \in S$. Then for all S' satisfying

$$S \subseteq S' \subseteq T, \tag{16}$$

and for all $i \in V$, we have

$$\boldsymbol{A}_{S'}(y)_{i,x} = y_{S' \cup \{i,x\}} = y_{S' \cup \{i\}} = \boldsymbol{A}_{S'}(y)_{\varnothing,i}.$$

This implies that each row and column of $\boldsymbol{A}_{S'}$ corresponding to $S' \cup \{i, x\}$ is the same as the one corresponding to $S' \cup \{i\}$. Therefore, the rows and columns of $A(S,T)(y)$ corresponding to x are the same as the rows corresponding to $\varnothing$. The duplicated rows and columns $A(S,T)(y)_{x,:}$ and $A(S,T)(y)_{:,x}$ can be removed for purposes of the PSD constraint according to Lemma C.1 and results in [GL05, Gvo08].

Now consider any $x \in T$, such that $x \notin S$. We also consider each pair of S' satisfying (16) containing x, and $\tilde{S} = S' \setminus x$. Given any such pair, since $x \notin S$, $\tilde{S}$ also satisfies (16). Also $(-1)^{|S' \setminus S|}$ and $(-1)^{|\tilde{S} \setminus S|}$ have the opposite signs. For all $i \in V$ and $i = \varnothing$, we have

$$\boldsymbol{A}_{\tilde{S}}(y)_{i,x} = y_{\tilde{S} \cup \{i,x\}} = y_{S' \cup \{i,x\}} = \boldsymbol{A}_{S'}(y)_{i,x}.$$

Since this result holds for every pair of $\boldsymbol{A}_{\tilde{S}}$ and $\boldsymbol{A}_{S'}$, the resulting pairwise cancellations in (2) lead $A(S,T)(y)_{x,:}$ and $A(S,T)(y)_{:,x}$ to be equal to zero. Therefore these rows and columns can be removed for purposes of the PSD constraint. This result, and the result in the previous paragraph, lead to the desired simplification of $A(S,T)(y)$.

As noted previously, according to Lemma 3.1 in [GLV09], constraints $y_{\{i,j\}} = 0, \forall \{i,j\} \notin E$ imply that $y_S = 0$ for any subset $S \subseteq V$ with $|S| \leq t+1$ containing nonedge in G. Therefore, $(\boldsymbol{A}_{S'})_{i,j} = 0$ for all $j \in V$ if $S' \cup \{i\} \notin \mathcal{K}(G)$. This implies that if $S \cup \{i\} \notin \mathcal{K}(G)$, then the i-th row and column in $\boldsymbol{A}(S,T)(y)$ contain zero entries only, and can be removed for purposes of the corresponding PSD constraints. Also $(\boldsymbol{A}_{S'})_{i,j} = 0$ if $S' \cup \{i,j\} \notin \mathcal{K}(G)$. Accordingly if $S' \notin \mathcal{K}(G)$, then the corresponding $\boldsymbol{A}_{S'}$ will be identically zero. Therefore, in (5) we can sum only over those subsets S' of vertices that are cliques. Similarly, for purposes of the constraint (3d) it is sufficient to consider only $\boldsymbol{A}(S,T)(y)$ where $S \in \mathcal{K}(G)$.

B Proof of Lemma 3.2

Zeta matrix We use tools introduced in [Lau02, Lau07] in the proof. In particular, given a finite set T such that $|T| = t-1$, we use the zeta matrix $\boldsymbol{Z}$ defined in the following way:

- $\boldsymbol{Z} \in \mathbb{R}^{(n+1)2^{t-1} \times (n+1)2^{t-1}}$ is the $\mathcal{P}(T) \times \mathcal{P}(T)$ block matrix with $\boldsymbol{Z}(S,S') := \boldsymbol{I}_{(n+1)}$ if $S \subseteq S'$ and $\boldsymbol{Z}(S,S') := 0$ otherwise, for $S, S' \subseteq T$, and

- its inverse $\boldsymbol{Z}^{-1}$ [LS91] is the $\mathcal{P}(T) \times \mathcal{P}(T)$ block matrix with $\boldsymbol{Z}^{-1}(S,S') = (-1)^{|S' \setminus S|} \boldsymbol{I}_{(n+1)}$ if $S \subseteq S'$ and $\boldsymbol{Z}^{-1}(S,S') = 0$ otherwise, for $S, S' \subseteq T$.

Note that $\boldsymbol{Z}^{\top}(S,S') = \boldsymbol{Z}(S',S) = \boldsymbol{I}_{(n+1)}$ if $S' \subseteq S$ and $\boldsymbol{Z}^{\top}(S,S') = 0$ otherwise. Since $\boldsymbol{Z}$ is invertible, $\boldsymbol{Z}^{\top}$ is also invertible, every vector w corresponds uniquely to a vector $x = \boldsymbol{Z}^{\top} w$. Define the block diagonal matrix

$$\hat{\boldsymbol{D}}(S,S) := \hat{\boldsymbol{A}}(S,T), \text{ for each } S \subseteq T, \tag{17}$$

where $\hat{\boldsymbol{A}}(S,T)$ defined as in (7), so that constraint (6) holds if and only if $\hat{\boldsymbol{D}} \succeq 0$. Notice $\boldsymbol{Z}^\top$ is a congruence mapping which preserves PSD matrices:

$$w^\top(\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top)w = (\boldsymbol{Z}^\top w)^\top \hat{\boldsymbol{D}}(\boldsymbol{Z}^\top w) = x^\top \hat{\boldsymbol{D}} x,$$

hence constraint (6) holds if and only if $\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top \succeq 0$. Moreover, constraint (4) is equivalent to $\tilde{\boldsymbol{M}}(T;y) \succeq 0$ (see discussion before Lemma 3.1), hence it suffices to prove $\tilde{\boldsymbol{M}}(T;y) \succeq \boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top$ for the purpose of Lemma 3.2.

Proving $\tilde{\boldsymbol{M}}(T;y) \succeq \boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top$ For $S_1, S_2 \subseteq T$, we consider the following cases:

(i) If $S_1 \cup S_2 = T$, then

$$\begin{aligned}
[\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) &= \sum_{S\subseteq T} \boldsymbol{Z}(S_1,S)\hat{\boldsymbol{D}}(S,S)\boldsymbol{Z}^\top(S,S_2) = \sum_{S:S_1\cup S_2\subseteq S\subseteq T} \hat{\boldsymbol{D}}(S,S) \\
&= \hat{\boldsymbol{D}}(T,T) = \tilde{\boldsymbol{A}}_T = \tilde{\boldsymbol{M}}(T,T)
\end{aligned}$$

which is PSD by assumption.

(ii) If $T = S_1 \cup S_2 \cup \{k\}$ for some $k \in T \setminus (S_1 \cup S_2)$, then

$$\begin{aligned}
[\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) &= \sum_{S\subseteq T} \boldsymbol{Z}(S_1,S)\hat{\boldsymbol{D}}(S,S)\boldsymbol{Z}^\top(S,S_2) = \sum_{S:S_1\cup S_2\subseteq S\subseteq T} \hat{\boldsymbol{D}}(S,S) \\
&= \tilde{\boldsymbol{A}}_{S_1\cup S_2} - \tilde{\boldsymbol{A}}_T = \tilde{\boldsymbol{M}}(S_1,S_2) - \tilde{\boldsymbol{M}}(T,T)
\end{aligned}$$

since $\tilde{\boldsymbol{M}}(T,T) \succeq 0$ by assumption, this implies

$$\tilde{\boldsymbol{M}}(S_1,S_2) = [\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) + \tilde{\boldsymbol{M}}(T,T) \succeq 0$$

if $[\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) \succeq 0$ for $S_1 \cup S_2 = T \setminus \{k\}$.

(iii) If $T \supsetneq S_1 \cup S_2 \cup \{k\}$ for some $k \in T \setminus (S_1 \cup S_2)$, then

$$\begin{aligned}
[\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) &= \sum_{S\subseteq T} \boldsymbol{Z}(S_1,S)\hat{\boldsymbol{D}}(S,S)\boldsymbol{Z}^\top(S,S_2) = \sum_{S:S_1\cup S_2\subseteq S\subseteq T} \hat{\boldsymbol{D}}(S,S) \\
&= \sum_{S:S_1\cup S_2\subseteq S\subseteq T} \left[\tilde{\boldsymbol{A}}_S - \sum_{k\in T\setminus S} \tilde{\boldsymbol{A}}_{S\cup\{k\}}\right] \\
&= \underbrace{\sum_{S:S_1\cup S_2\subseteq S\subseteq T} \tilde{\boldsymbol{A}}_S}_{\text{(I)}} - \underbrace{\sum_{S:S_1\cup S_2\subseteq S\subseteq T} \sum_{k\in T\setminus S} \tilde{\boldsymbol{A}}_{S\cup\{k\}}}_{\text{(II)}}
\end{aligned}$$

- Part(I): $\sum_{S:S_1\cup S_2\subseteq S\subseteq T} \boldsymbol{A}_S$
  Fix $X := S_1 \cup S_2$, then $S_1 \cup S_2 \subseteq S$ means $S = X \cup R$ for some $R \subseteq T \setminus X$, hence

  $$\sum_{S:S_1\cup S_2\subseteq S\subseteq T} \tilde{\boldsymbol{A}}_S = \sum_{R\subseteq T\setminus X} \tilde{\boldsymbol{A}}_{X\cup R} = \tilde{\boldsymbol{A}}_X + \sum_{R\subseteq T\setminus X, R\neq\varnothing} \tilde{\boldsymbol{A}}_{X\cup R}$$

  and notice $\tilde{\boldsymbol{A}}_X = \tilde{\boldsymbol{A}}_{S_1\cup S_2}$ by definition.

- Part(II): $\sum_{S:X\subseteq S\subseteq T}\sum_{k\in T\setminus S}\tilde{\boldsymbol{A}}_{S\cup\{k\}}$
  Let $U = S\cup\{k\}$ for some $k\in T\setminus S$, then for a fixed U containing X, the number of valid (S,k) pairs with $S\cup\{k\}=U$ and $X\subseteq S\subseteq T$ is $|U\setminus X|$. Therefore,

$$\sum_{S:X\subseteq S\subseteq T}\sum_{k\in T\setminus S}\tilde{\boldsymbol{A}}_{S\cup\{k\}} = \sum_{U:X\subseteq U\subseteq T}\Big[|U\setminus X|\tilde{\boldsymbol{A}}_U\Big]$$

Substituting the above into (I) and (II) gives

$$\begin{aligned}[\boldsymbol{Z}\hat{\boldsymbol{D}}\boldsymbol{Z}^\top](S_1,S_2) &= \sum_{U:X\subseteq U\subseteq T}\tilde{\boldsymbol{A}}_U - \sum_{U:X\subseteq U\subseteq T}|U\setminus X|\tilde{\boldsymbol{A}}_U = \sum_{U:X\subseteq U\subseteq T}\Big[1-|U\setminus X|\Big]\tilde{\boldsymbol{A}}_U\\ &= \tilde{\boldsymbol{A}}_X + \sum_{U:X\subsetneq U\subseteq T}\Big[1-|U\setminus X|\Big]\tilde{\boldsymbol{A}}_U\\ &= \tilde{\boldsymbol{A}}_{S_1\cup S_2} - \sum_{U:S_1\cup S_2\subsetneq U\subseteq T}\Big[|U\setminus(S_1\cup S_2)|-1\Big]\tilde{\boldsymbol{A}}_U\end{aligned}$$

Define $\tilde{\boldsymbol{M}}_{\text{sub}}(S_1,S_2) := \sum_{U:S_1\cup S_2\subsetneq U\subseteq T}\Big[|U\setminus(S_1\cup S_2)|-1\Big]\tilde{\boldsymbol{A}}_U$, and $S_1\cup S_2\subsetneq U$ implies $|U\setminus(S_1\cup S_2)|\geq 1$, thus $|U\setminus(S_1\cup S_2)|-1\geq 0$, implies the coefficient in front of each $\tilde{\boldsymbol{A}}_U$ is non-negative. By assumption, each block $\boldsymbol{A}_U\succeq 0$, and therefore each (S_1,S_2)-block of $\tilde{\boldsymbol{M}}_{\text{sub}}$ being a non-negative linear combination of PSD matrices is also PSD, i.e., $\boldsymbol{M}_{\text{sub}}(S_1,S_2)\succeq 0$ for all S_1,S_2.

Furthermore, given a vector $x\in\mathbb{R}^{(n+1)2^{(t-1)}}$, we can view it as being partitioned into 2^{t-1} blocks $x_S\in\mathbb{R}^{(n+1)}$ (one for each subset $S\subseteq T$), i.e., $x=(x_S)_{S\subseteq T}$, then

$$\begin{aligned}x^\top\tilde{\boldsymbol{M}}_{\text{sub}}x &= \sum_{S_1,S_2\subseteq T}(x_{S_1})^\top\Big[\tilde{\boldsymbol{M}}_{\text{sub}}(S_1,S_2)\Big](x_{S_2})\\ &= \sum_{U:S_1\cup S_2\subsetneq U}\Big[|U\setminus(S_1\cup S_2)|-1\Big]\sum_{S_1\subseteq U,S_2\subseteq U}(x_{S_1})^\top\tilde{\boldsymbol{A}}_U(x_{S_2})\geq 0.\end{aligned}$$

This concludes the proof of Lemma 3.2.

C Lemma C.1

Lemma C.1. *Given any $n\times n$ symmetric matrix $\boldsymbol{M}$ and some $\mathcal{J}\subsetneq[n]$, so that $\boldsymbol{M}_{j\bullet}=\boldsymbol{M}_{j'\bullet}=\boldsymbol{M}_{\bullet j}^\top=\boldsymbol{M}_{\bullet j'}^\top$ for any $j,j'\in\mathcal{J}$. Then $\boldsymbol{M}\succeq 0$ if and only if $\tilde{M}\succeq 0$, where $\tilde{\boldsymbol{M}}$ is the $m\times m$ ($m=n+1-|\mathcal{J}|$) principal submatrix of $\boldsymbol{M}$ with $\boldsymbol{M}_{j\bullet}$ and $\boldsymbol{M}_{\bullet j}$ removed for all but one $j\in\mathcal{J}$.*

Proof. The "only if" direction is trivial as $\tilde{\boldsymbol{M}}$ is a principal submatrix of $\boldsymbol{M}$. Suppose $\tilde{\boldsymbol{M}}\succeq 0$, then it's enough to consider the case of $|\mathcal{J}|=2$ as it can be easily extended to general $\mathcal{J}$ by an induction statement on its cardinality. Without loss of generality, it's enough to consider $\mathcal{J}=\{1,2\}$ as the spectrum of $\boldsymbol{M}$ is invariant under change of basis by permutation matrix.

Write

$$\boldsymbol{M}=\begin{pmatrix}c & c & v^\top\\ c & c & v^\top\\ v & v & A\end{pmatrix}\quad\text{and}\quad\tilde{\boldsymbol{M}}=\begin{pmatrix}c & v^\top\\ v & A\end{pmatrix}\succeq 0.$$

For any vector $x^\top=(x_1,x_2,\ldots,x_n)=(x_1,x_2,y^\top)\in\mathbb{R}^n$, we see that $x^\top\boldsymbol{M}x=c(x_1+x_2)^2+2(x_1+$

$x_2)\langle v, y\rangle + y^\top \boldsymbol{A} y$. Because $\tilde{\boldsymbol{M}} \succeq 0$, then for any $(z, y^\top) \in \mathbb{R}^n$, $(z, y^\top)^\top \tilde{\boldsymbol{M}}(z, y^\top) = cz^2 + 2z\langle v, y\rangle + y^\top A y \geq 0$, so we can simply take $z = x_1 + x_2$, so that $x^\top \boldsymbol{M} x = (z, y^\top)^\top \tilde{\boldsymbol{M}}(z, y^\top) \geq 0$, therefore, $\boldsymbol{M} \succeq 0$. □

D Proof of Theorem 3.4

Our overall approach is to look for a feasible point for (6) of the following form:

$$\alpha_i = c_i n^{-\frac{i}{2}} > 0 \tag{18}$$

and find the constant $c_i > 0$ for each $i > 0$ (with $c_0 = \alpha_0 = 1$ given by Definition 3.3). Then, we obtain our lower bound by lower bounding the objective function at that feasible point.

D.1 Schur complements

Letting $m := |S|$, this sum has $s := |T| - m$ terms, corresponding to the number of vertices in $T \setminus S$, i.e. $s \in [1, \ldots, |T|]$ and by construction $|T| = t - 1$. In this proof we omit the dependence of $\hat{\boldsymbol{A}}(S, T)(\alpha)$ and $\tilde{\boldsymbol{A}}_{S\cup\{k\}}(\alpha)$ on α. For $\hat{\boldsymbol{A}}(S, T)$ defined in (7), we have

$$\hat{\boldsymbol{A}}(S, T) = \tilde{\boldsymbol{A}}_S - \sum_{\substack{k \in T\setminus S \\ S\cup\{k\}\in\mathcal{K}(G)}} \tilde{\boldsymbol{A}}_{S\cup\{k\}} = \sum_{\substack{k \in T\setminus S \\ S\cup\{k\}\in\mathcal{K}(G)}} \left(\frac{1}{s}\tilde{\boldsymbol{A}}_S - \tilde{\boldsymbol{A}}_{S\cup\{k\}}\right) = \sum_{\substack{k \in T\setminus S \\ S\cup\{k\}\in\mathcal{K}(G)}} \boldsymbol{A}_k(S, T) \tag{19}$$

where

$$\boldsymbol{A}_k(S, T) = \begin{pmatrix} \overline{\alpha}_m - \alpha_{m+1} & \overline{\alpha}_{m+1}\mathbb{1}^\top - \alpha_{m+2} v_{S\cup\{k\}}^\top \\ \overline{\alpha}_{m+1}\mathbb{1} - \alpha_{m+2} v_{S\cup\{k\}} & \overline{\alpha}_{m+1}\boldsymbol{I} + \overline{\alpha}_{m+2}\boldsymbol{M}_S - \alpha_{m+2}\,\mathsf{diag}(v_{S\cup\{k\}}) - \alpha_{k+3}\boldsymbol{M}_{S\cup\{k\}} \end{pmatrix}, \tag{20}$$

with $v_{S\cup\{k\}}$, $\boldsymbol{M}_S$ and $\boldsymbol{M}_{S\cup\{k\}}$ defined in Section 3, and

$$\overline{\alpha}_m := \alpha_m / s.$$

According to equation (18) of each α_i, for all p sufficiently large, $\overline{\alpha}_m - \alpha_{m+1} \geq \alpha_m/(t-1) - \alpha_{m+1} > 0$ uniformly in s. Therefore, for each $k \in T \setminus S$, it is sufficient to consider the positive semidefiniteness of the Schur complement of $\boldsymbol{A}_k(S, T)$:

$$\boldsymbol{D}_S := \mathsf{diag}(w) + \overline{\alpha}_{m+2}\boldsymbol{M}_S - \alpha_{m+3}\boldsymbol{M}_{S\cup\{k\}} - \frac{1}{\overline{\alpha}_m - \alpha_{m+1}} w w^\top \tag{21}$$

where

$$w := \overline{\alpha}_{m+1}\mathbb{1} - \alpha_{m+2}\boldsymbol{v}_{S\cup\{k\}}. \tag{22}$$

Similarly we can decompose each $\boldsymbol{M}_S$ and $\boldsymbol{M}_{S\cup\{k\}}$ into submatrices $\boldsymbol{U}_S$ and $\boldsymbol{U}_{S\cup\{k\}}$ of the Seidel matrices $\boldsymbol{U}_{G_S}$ and $\boldsymbol{U}_{G_{S\cup\{k\}}}$ of G_S and $G_{S\cup\{k\}}$ respectively (or more precisely, in the latter case $\boldsymbol{U}_{S\cup\{k\}}$ is a submatrix of $\boldsymbol{U}_{G_{S\cup\{k\}}}$ padded with rows and columns of all zeros in the positions

$i \in V(G_S)$ corresponding to $S \cup \{k, i\} \notin \mathcal{K}(G_S)$):

$$\boldsymbol{M}_S = \frac{1}{2}(\boldsymbol{U}_S + \mathbb{1}\mathbb{1}^\top - \boldsymbol{I}) \text{ and } \boldsymbol{M}_{S\cup\{k\}} = \frac{1}{2}\left(\boldsymbol{U}_{S\cup\{k\}} + v_{S\cup\{k\}}v_{S\cup\{k\}}^\top - \mathsf{diag}(v_{S\cup\{k\}})\right) \tag{23}$$

Moreover, the definition for $\boldsymbol{U}_S$ can be extended to the case where $S = T$, so that $\boldsymbol{M}_T = \frac{1}{2}(\boldsymbol{U}_T + \mathbb{1}\mathbb{1}^\top - \boldsymbol{I})$.

D.1.1 Positive semidefiniteness of the Schur complement of $\hat{\boldsymbol{A}}(T, T)$

For $S = T$, the Schur complement of $\hat{\boldsymbol{A}}(T, T)$ is

$$\boldsymbol{D}_T = \alpha_t \boldsymbol{I} + \alpha_{t+1}\boldsymbol{M}_T - \frac{\alpha_t^2}{\alpha_{t-1}}\mathbb{1}\mathbb{1}^\top \tag{24}$$

In all these cases, it is more convenient to decompose $\boldsymbol{M}_T$ into a principal submatrix $\boldsymbol{U}_T$ of the Seidel matrix $\boldsymbol{U}_{G_T}$ according to equation (23). We can control the eigenvalues of $\boldsymbol{U}_{G_T}$, and therefore $\boldsymbol{U}_T$, by the eigenvalue interlacing theorem. Then, equation (24) becomes

$$\begin{aligned}\boldsymbol{D}_T &= \alpha_t \boldsymbol{I} + \alpha_{t+1}\boldsymbol{M}_T - \frac{\alpha_t^2}{\alpha_{t-1}}\mathbb{1}\mathbb{1}^\top \\ &= \left(\alpha_t - \frac{\alpha_{t+1}}{2}\right)\boldsymbol{I} + \frac{\alpha_{t+1}}{2}\boldsymbol{U}_T + \left(\frac{\alpha_{t+1}}{2} - \frac{\alpha_t^2}{\alpha_{t-1}}\right)\mathbb{1}\mathbb{1}^\top.\end{aligned}$$

Since the only non-zero eigenvalue of $\mathbb{1}\mathbb{1}^\top$ is the norm of $\mathbb{1}$ which is positive and asymptotically $O(p)$, we require the constant before $\mathbb{1}\mathbb{1}^\top$ be non-negative:

$$\frac{\alpha_{t+1}}{2} - \frac{\alpha_t^2}{\alpha_{t-1}} \geq 0, \tag{25}$$

so that this term does not decrease $\lambda_{\mathsf{min}}(\boldsymbol{D}_T)$. Based on that to guarantee $\boldsymbol{D}_T \succeq 0$, it suffices to require

$$\begin{aligned}\lambda_{\mathsf{min}}(\boldsymbol{D}_T) &\geq \left(\alpha_t - \frac{\alpha_{t+1}}{2}\right) + \frac{\alpha_{t+1}}{2}\lambda_{\mathsf{min}}(\boldsymbol{U}_T) + \lambda_{\mathsf{min}}\left(\left(\frac{\alpha_{t+1}}{2} - \frac{\alpha_t^2}{\alpha_{t-1}}\right)\mathbb{1}\mathbb{1}^\top\right) \\ &\geq \alpha_t - \frac{\alpha_{t+1}\rho}{2}\sqrt{n} + E_1(n) \geq 0, \end{aligned} \tag{26}$$

where the leading order term of (26) is $O(n^{-\frac{t}{2}})$ and the lower term error is $E_1(n) := -\frac{\alpha_{t+1}}{2} = O(n^{-\frac{t+1}{2}})$. The inequalities (25) and (26) for G hold for sufficiently large p if

$$\frac{c_{t+1}}{2} - \frac{c_t^2}{c_{t-1}} \geq 0 \tag{27}$$

$$c_t - \frac{c_{t+1}\rho}{2} + \hat{E}_1(n) \geq 0. \tag{28}$$

where $\hat{E}_1(n) := -\frac{c_{t+1}}{2\sqrt{n}}$.

D.1.2 Positive semidefiniteness of the Schur complement of $\hat{\boldsymbol{A}}(S,T)$ for $S \subsetneq T$

When $t > 1$, for each proper subset S of T ($0 \le |S| = m \le t-2$), we also establish the positive semidefiniteness of the Schur complement of $\boldsymbol{A}_k(S,T)$ for each $k \in T \setminus S$. This leads to the positive semidefiniteness of the corresponding $\hat{\boldsymbol{A}}(S,T)$. We simplify the notation by writing $v_{S\cup\{k\}}$ as v, denoting $v_c := \mathbb{1} - v$. Hence we can write w defined in (22) as

$$w = \overline{\alpha}_{m+1}\mathbb{1} - \alpha_{m+2}\boldsymbol{v} = (\overline{\alpha}_{m+1} - \alpha_{m+2})\boldsymbol{v} + \overline{\alpha}_{m+1}\boldsymbol{v}_c.$$

Using equation (23), the Schur complement $\boldsymbol{D}_S$ of $\boldsymbol{A}_k(S,T)$ given by (21) can be decomposed as

$$\boldsymbol{D}_S = \left(\overline{\alpha}_{m+1} - \frac{\overline{\alpha}_{m+2}}{2}\right)\boldsymbol{I} + \left(-\alpha_{m+2} + \frac{\alpha_{m+3}}{2}\right)\mathsf{diag}(\boldsymbol{v}) + \frac{\overline{\alpha}_{m+2}}{2}\boldsymbol{U}_S - \frac{\alpha_{m+3}}{2}\boldsymbol{U}_{S\cup\{k\}} + \boldsymbol{C}$$

where

$$\boldsymbol{C} := \tau\boldsymbol{v}\boldsymbol{v}^\top + \iota\boldsymbol{v}_c\boldsymbol{v}_c^\top + \eta(\boldsymbol{v}\boldsymbol{v}_c^\top + \boldsymbol{v}_c\boldsymbol{v}^\top), \text{ such that } \begin{cases} \iota := \frac{\overline{\alpha}_{m+2}}{2} - \frac{\overline{\alpha}_{m+1}^2}{\overline{\alpha}_m - \alpha_{m+1}}, \\ \tau := \frac{\overline{\alpha}_{m+2}}{2} - \frac{\alpha_{m+3}}{2} - \frac{(\overline{\alpha}_{m+1} - \alpha_{m+2})^2}{\overline{\alpha}_m - \alpha_{m+1}} = \iota + E_2, \\ \eta := \frac{\overline{\alpha}_{m+2}}{2} - \frac{\overline{\alpha}_{m+1}(\overline{\alpha}_{m+1} - \alpha_{m+2})}{\overline{\alpha}_m - \alpha_{m+1}} = \iota + E_3. \end{cases} \tag{29}$$

and

$$E_2 := -\frac{\alpha_{m+3}}{2} + \frac{\alpha_{m+2}(2\overline{\alpha}_{m+1} - \alpha_{m+2})}{\overline{\alpha}_m - \alpha_{m+1}} = O(n^{-\frac{m+3}{2}}) \tag{30}$$

$$E_3 := \frac{\overline{\alpha}_{m+1}\alpha_{m+2}}{\overline{\alpha}_m - \alpha_{m+1}} = O(n^{-\frac{m+3}{2}}) \tag{31}$$

where the asymptotic estimates are given by the order assumption (18). Since $\boldsymbol{D}_S \succeq 0$ if and only if $\lambda_{\min}(\boldsymbol{D}_S) \ge 0$, we need to estimate $\lambda_{\min}(\boldsymbol{C})$, which is given by Lemma E.1. It suffices to consider $-\zeta \le \lambda_{\min}(\boldsymbol{C})$ and minimize $\zeta > 0$. As the leading order terms of ι, η and τ are equal, we set this leading order terms to zero, which is achieved by:

$$\frac{\overline{\alpha}_{m+2}}{2} = \frac{\overline{\alpha}_{m+1}^2}{\overline{\alpha}_m} \text{ which holds if and only if } \frac{c_{m+2}}{2c_{m+1}} = \frac{c_{m+1}}{c_m}. \tag{32}$$

Therefore,

$$\iota = \frac{\overline{\alpha}_{m+1}^2}{\overline{\alpha}_m} - \frac{\overline{\alpha}_{m+1}^2}{\overline{\alpha}_m - \alpha_{m+1}} = \frac{-\overline{\alpha}_{m+1}^2\alpha_{m+1}}{\overline{\alpha}_m(\overline{\alpha}_m - \alpha_{m+1})} = \frac{-\overline{\alpha}_{m+2}\alpha_{m+1}}{2(\overline{\alpha}_m - \alpha_{m+1})}, \quad \text{by (32)} \tag{33}$$

Moreover, this ζ can be further minimized by setting the leading order term of $\tau - \iota = E_2 = 0$. By equation (30), it is enough to require:

$$\frac{\alpha_{m+3}}{2} = \frac{2\overline{\alpha}_{m+1}\alpha_{m+2}}{\overline{\alpha}_m} \text{ which holds if and only if } \frac{c_{m+3}}{4c_{m+2}} = \frac{c_{m+1}}{c_m}. \tag{34}$$

This further enforces equality between $\tau = \iota$ because

$$\tau - \iota = -\frac{2\overline{\alpha}_{m+1}\alpha_{m+2}}{\overline{\alpha}_m} + \frac{2\overline{\alpha}_{m+1}\alpha_{m+2} - \alpha_{m+2}^2}{\overline{\alpha}_m - \alpha_{m+1}} = \alpha_{m+2}\left(\frac{2\overline{\alpha}_{m+1}\alpha_{m+1} - \overline{\alpha}_m\alpha_{m+2}}{\overline{\alpha}_m(\overline{\alpha}_m - \alpha_{m+1})}\right) \tag{35}$$

where the numerator of the fraction on the right-hand side above is 0 by equation (32). Moreover, $\eta = -\iota$ because:

$$\eta + \iota = 2\iota + E_3 = \frac{-2\overline{\alpha}_{m+2}\alpha_{m+1}}{2(\overline{\alpha}_m - \alpha_{m+1})} + \frac{\overline{\alpha}_{m+1}\alpha_{m+2}}{\overline{\alpha}_m - \alpha_{m+1}} = 0. \text{ by (31) and (33)} \tag{36}$$

Let $\|\boldsymbol{v}\|_2^2 = n_v, \|\boldsymbol{v}_c\|_2^2 = n_c$, so that $n_S = n_v + n_c$, where n_S is given by (8). By our previous computation, we have

$$\iota = \tau = -\eta < 0,$$

Therefore, by Lemma E.1, the nonzero eigenvalues of the matrix $\boldsymbol{C}$ are

$$\begin{aligned}
\lambda_1 &= \frac{1}{2}\left(\iota n_S + \sqrt{\iota^2(n_v - n_c)^2 + 4\iota^2 n_v n_c}\right) = \frac{1}{2}\left(\iota n_S + |\iota|(n_v + n_c)\right) = 0 \\
\lambda_2 &= \frac{1}{2}\left(\iota n_S - |\iota|(n_v + n_c)\right) = \iota n_S = -\frac{n_S\overline{\alpha}_{m+1}\alpha_{m+2}}{2(\overline{\alpha}_m - \alpha_{m+1})}, \text{ by (33)} \\
&= -\frac{n_S\overline{\alpha}_{m+1}\alpha_{m+2}}{2\overline{\alpha}_m} + E_4' = -\frac{n_S\alpha_{m+3}}{8} + E_4', \text{ by (34)}
\end{aligned}$$

where $E_4' = -\frac{n_S\alpha_{m+2}^2}{4(\overline{\alpha}_m - \alpha_{m+1})} = O(n^{-\frac{m+2}{2}})$. Therefore, assuming that $n_S = \Theta(n)$, it is sufficient to prove that for sufficiently large n :

$$\lambda_{\min}(\boldsymbol{D}_S) \geq -\frac{n_S\alpha_{m+3}}{8} + \overline{\alpha}_{m+1} - \frac{\overline{\alpha}_{m+2}\rho}{2}\sqrt{n} - E_4(n) \geq 0, \tag{37}$$

and we write $E_4(n)$ as the collection of lower order terms in $\lambda_{\min}(\boldsymbol{D}_S)$

$$E_4(n) = \alpha_{m+2}\left(1 + \frac{1}{2s}\right) + \frac{\alpha_{m+3}\rho}{2}\sqrt{n} + \frac{n_S\alpha_{m+2}^2}{4(\overline{\alpha}_m - \alpha_{m+1})} = O(n_S \cdot n^{-\frac{m+4}{2}} + n_S^{-(m+2)/2}) \geq 0 \tag{38}$$

Note that if $n_S \ll n$, the term $-\frac{n_S\alpha_{m+3}}{8}$ will be collected into $E_4(n)$ as a lower order term, which leads to inequalities as for $S = T$. We write $\hat{E}_4(n) = n^{\frac{m+1}{2}}$ and $E_4(n) = O(n_S \cdot n^{-3/2} + n^{-1/2})$. To ensure the feasibility condition (6) under the α_i of form give by (18), we need to make sure that there exists some c_m such that for all $0 \leq m \leq t-2$,

$$\frac{c_{m+3}}{4c_{m+2}} = \frac{c_{m+2}}{2c_{m+1}} = \frac{c_{m+1}}{c_m}, \text{ by (32) and (34)} \tag{39}$$

$$\frac{c_{m+1}}{s} - \frac{n_S c_{m+3}}{8n} - \frac{c_{m+2}\rho}{2s} - \hat{E}_4(n) \geq 0, \text{ by (37)} \tag{40}$$

holds for sufficiently large n. Since $n_S < n$, it suffices to verify (40) for $n_S = n$.

D.2 Lower bound for $t = 1$

When $t = 1$, since $|T| = t - 1 = 0$, $T = \varnothing$, and it has no proper subsets. Therefore, it is sufficient to ensure that the inequalities (27) and (28) are satisfied. Theses inequalities can be collected into:

$$\frac{2\sqrt{n}}{\rho\sqrt{n} + 1} \geq \frac{c_2}{c_1} \geq \frac{2c_1}{c_0} = 2c_1. \tag{41}$$

We obtain

$$L^1(\overline{G}) \geq \frac{\sqrt{n}}{\rho} - O(1), \tag{42}$$

which proves Theorem 3.4 for $t = 1$.

The $L^1(G_p)$ case We consider the case $L^1(G_p)$, which is equivalent to the well known the Lovasz ϑ/SOS-2. In this case, $\boldsymbol{M}_T = \boldsymbol{A}_{G_p}$, which has an eigenvector $\mathbb{1} \in \mathbb{R}^p$ corresponding to the $\frac{p-1}{2}$ eigenvalue, and the smallest eigenvalue equal to $\frac{-1-\sqrt{p}}{2}$. Therefore, $\boldsymbol{D}_T$ is PSD if

$$\alpha_1 + \alpha_2 \frac{p-1}{2} - \frac{\alpha_1^2}{\alpha_0} p \geq 0, \quad \text{and} \quad \alpha_1 + \alpha_2 \frac{-1-\sqrt{p}}{2} \geq 0. \tag{43}$$

The second inequality is equivalent to $\alpha_2 \leq \frac{2}{1+\sqrt{p}}\alpha_1$, then substituting the second inequality into the first inequality gives

$$0 \geq p\alpha_1^2 - \alpha_1 - \frac{p-1}{2}\alpha_2 \geq p\alpha_1^2 - \alpha_1 - \frac{p-1}{2} \cdot \frac{2}{1+\sqrt{p}}\alpha_1$$

which yields $0 \leq \alpha_1 \leq \frac{1}{\sqrt{p}}$ and $0 \leq \alpha_2 \leq \frac{2}{p+\sqrt{p}}$. And α_1 is optimized taking $\alpha_1 = \frac{1}{\sqrt{p}}$, which gives the lower bound $L^1(\overline{G}_p) \geq p \cdot \frac{1}{\sqrt{p}} = \sqrt{p}$. In particular, since $L^1 = SOS_2 = \vartheta$, this recovers the classical result $\vartheta(G) = \sqrt{n}$ for vertex-transitive self-complementary graph G of n vertices [Lov79].

Remark D.1. *Since G_p is a strongly regular graph and every two adjacent vertices have $\frac{1}{4}(p-5)$ common neighbors, $\mathbb{1}$ is the eigenvector corresponding to the eigenvalues $\frac{p-5}{4}$ and -1 of $\boldsymbol{A}_{G_{p,\{0\}}}$ and $\boldsymbol{U}_{G_{p,\{0\}}}$, respectively.*

The $L^1(G_{p,\{k\}})$ case We now consider the consider the case $L^1(G_{p,\{k\}})$ where $G_{p,\{k\}}$ is the degree one localization of the Paley graph. Since the Paley graph is vertex transitive, it is sufficient to consider $L^1(G_{p,\{0\}})$. In this case, $\boldsymbol{M}_T = \boldsymbol{A}_{G_{p,\{0\}}}$, which also an eigenvector $\mathbb{1} \in \mathbb{R}^{\frac{p-1}{2}}$ corresponding to the $\frac{p-5}{4}$ eigenvalue, and the smallest eigenvalue is bounded by $\frac{-1-\sqrt{p}}{2}$ from below by the eigenvalue interlacing theorem. Therefore, $\boldsymbol{D}_T$ is PSD if

$$\alpha_1 + \alpha_2 \frac{p-5}{4} - \frac{\alpha_1^2}{\alpha_0}\frac{p-1}{2} \geq 0, \quad \text{and} \quad \alpha_1 + \alpha_2 \frac{-1-\sqrt{p}}{2} \geq 0.$$

The second inequality is equivalent to $\alpha_2 \leq \frac{2}{1+\sqrt{p}}\alpha_1$, then substituting the second inequality into the first inequality gives

$$0 \geq \frac{p-1}{2}\alpha_1^2 - \alpha_1 - \frac{p-5}{4}\alpha_2 \geq \frac{p-1}{2}\alpha_1^2 - \alpha_1 - \frac{p-5}{4} \cdot \frac{2}{1+\sqrt{p}}\alpha_1$$

which yields $0 \leq \alpha_1 \leq \frac{(p-5)\sqrt{p}+p+3}{(p-1)^2}$ and $0 \leq \alpha_2 \leq \frac{2p^2-12p+16\sqrt{p}-6}{(p-1)^3}$. And α_1 is optimized taking $\alpha_1 = \frac{(p-5)\sqrt{p}+p+3}{(p-1)^2}$, which gives a lower bound $L^1(\overline{G}_{p,\{k\}}) \geq \frac{p-1}{2} \cdot \frac{(p-5)\sqrt{p}+p+3}{(p-1)^2} = \frac{\sqrt{p}}{2} + O\left(\frac{1}{2}\right)$.

D.3 Lower bound for $t \geq 2$

For any fixed $t \geq 2$, to satisfy constraints (6) for sufficiently large p, we need to find a feasible set $\{c_i\}_{i=1}^{t+1}$ of positive numbers so that the inequalities (27), (28), (39), and (40) hold, in the case of the last two inequalities, for each $0 \leq m \leq t-2$. These inequalities can be collected into the following system:

$$\frac{2}{\rho} > \frac{2\sqrt{n}}{\rho\sqrt{n}+1} \geq \frac{c_{t+1}}{c_t} = \frac{2c_t}{c_{t-1}} = \frac{4c_{t-1}}{c_{t-2}} = \dots = \frac{2^{t-1}c_2}{c_1} = \frac{2^t c_1}{c_0} = 2^t c_1 \tag{44}$$

$$\frac{c_{m+3}}{c_{m+2}} + \hat{E}_4(n) \leq \frac{8}{t-m-1}\left(\frac{c_{m+1}}{c_{m+2}} - \frac{\rho}{2}\right), \text{ for } 0 \leq m \leq t-2 \tag{45}$$

where (44) is obtained by reorganizing and combining (39), (27), and (28), whereas (45) is obtained by shifting terms in (40) for all m and substituting $s = t-m-1$.

Now define for $0 \leq i \leq t$, $r_i = \frac{8}{t-i+1}$ and $q_i = \frac{c_{i+1}}{c_i}$. In particular, $q_0 = c_1$ as $c_0 = 1$ by definition. We can reorganize the above inequalities as:

$$\frac{2}{\rho} > \frac{2\sqrt{n}}{\rho\sqrt{n}+1} \geq q_t = 2q_{t-1} = 2^2 q_{t-2} = \cdots = 2^{t-i} q_i = \cdots = 2^{t-1} q_1 = 2^t q_0, \tag{46}$$

$$q_i + \hat{E}_4(n) \leq r_i\left(\frac{1}{q_{i-1}} - \frac{\rho}{2}\right), \text{ for } 2 \leq i \leq t. \tag{47}$$

For sufficiently large n, $\hat{E}_4(n) = O(n^{-1/2}) < q_i$, and we have $q_i + \hat{E}_4(n) \leq 2q_i$. Since $t \geq 2$, we have $2^{t-1} \geq 2$ and $\rho^2(2^{t-i+1}-1) \geq (t-i+1)2^{i-t}$ for any $0 \leq i \leq t, \rho \geq 1$. Equation (46) implies $2q_i \leq \frac{2^{i+2-t}}{\rho}$ for any $i \geq 0$ and $\frac{1}{q_{i-1}} \geq \rho \cdot 2^{t-i}$ for any $i \geq 1$. Hence

$$q_i + \hat{E}_4(n) \leq 2q_i \leq \frac{2^{i+2-t}}{\rho} \leq \frac{8\rho}{t-i+1}\left(2^{t-i} - \frac{1}{2}\right) \leq r_i\left(\frac{1}{q_{i-1}} - \frac{\rho}{2}\right), \text{ for } 2 \leq i \leq t$$

for sufficiently large n. Hence, equation (47) is redundant. Therefore, we can take $q_0 = c_1 = \frac{\sqrt{n}}{2^{t-1}(\rho\sqrt{n}+1)}$ by (46), we establish the following lower bound for sufficiently large n:

$$L^t(\overline{G}) \geq \frac{c_1|V(\overline{G})|}{\sqrt{n}} \geq \frac{n-t+1}{2^{t-1}(\rho\sqrt{n}+1)} = \frac{\sqrt{n}}{2^{t-1}\rho} + O\left(\frac{1}{2^t}\right)$$

This completes the proof of Theorem 3.4.

E Lemma E.1

Lemma E.1. *Let $\boldsymbol{C}$ be a matrix of the form $\tau vv^\top + \iota v_c v_c^\top + \eta(vv_c^\top + v_c v^\top)$, where each of v and v_c is a vector in $\mathbb{R}^n$ containing 0 and 1's in each entry such that $v + v_c = \mathbb{1}$. Let $\|v\|_2^2 = n_v$, $\|v_c\|_2^2 = n_c$, so that $n = n_v + n_c$. Then the eigenvalues of $\boldsymbol{C}$ are either 0 or given by*

$$\lambda_{1,2} = \frac{\tau n_v + \iota n_c \pm \sqrt{(\tau n_v - \iota n_c)^2 + 4\eta^2 n_v n_c}}{2}.$$

Proof. We can represent $\boldsymbol{C}$ in block form as:

$$\boldsymbol{C} = \begin{pmatrix} \tau \boldsymbol{J}_{n_v,n_v} & \eta \boldsymbol{J}_{n_v,n_c} \\ \eta \boldsymbol{J}_{n_c,n_v} & \iota \boldsymbol{J}_{n_c,n_c} \end{pmatrix}$$

where $\boldsymbol{J}_{m,n}$ is an $m \times n$ all-ones matrix. The range of $\boldsymbol{C}$ is in the span of v and v_c, and therefore eigenvectors corresponding to non-zero eigenvalues have the prescribed form x, i.e., a linear combination of v and v_c:

$$x = \begin{bmatrix} \alpha \mathbb{1}_{n_v} \\ \beta \mathbb{1}_{n_c} \end{bmatrix}$$

where α and β are scalars. Substituting x into the equation $\det(\boldsymbol{C} - \lambda \boldsymbol{I})x = 0$ leads to the following equivalent characteristic equation:

$$\begin{vmatrix} \tau n_v - \lambda & \eta n_c \\ \eta n_v & \iota n_c - \lambda \end{vmatrix} = 0.$$

Therefore, the eigenvalues are given by:

$$\lambda_{1,2} = \frac{\tau n_v + \iota n_c \pm \sqrt{(\tau n_v - \iota n_c)^2 + 4\eta^2 n_v n_c}}{2}.$$

□

F Proof of Proposition 3.9

By vertex transitivity of G, there exists an optimal solution y^* of $L^t(\overline{G})$ which is invariant under permutations of vertices, i.e., $y_i^* = y_j^*$ for all vertices i, j in G (see, for example, Section 2.4 of [Gvo08]). Therefore, the optimal $L^t(\overline{G})$ value is

$$\frac{|V(G_K)|}{|V(G)|} L^t(\overline{G}) = |V(G_K)| y_{\{0\}}^* = \sum_{i \in V(G_K)} y_i^* \leq L^t(\overline{G}_K)$$

where the last inequality follows from the following lemma.

Lemma F.1. *For any graph G, any clique $K \in \mathcal{K}(G)$ and any $t \geq 1$, let $y^* \in \mathcal{P}_{t+1}(G_K)$ be an optimal point for $L^t(\overline{G}_K)$ and $y \in \mathcal{P}_{t+1}(G)$ be a feasible point for $L^t(\overline{G})$. In this setting, we have the following lower bound on the optimal value of $L^t(\overline{G}_K)$:*

$$\sum_{i \in V(G_K)} y_i \leq \sum_{i \in V(G_K)} y_i^* \tag{48}$$

Proof. We construct a feasible point y^K for $L^t(\overline{G}_K)$ and the corresponding moment matrices $M_K(T; y^K)$ for all $T \in V(\overline{G}_K)$ such that $|T| = t - 1$ using a feasible point y for $L^t(\overline{G})$ and the corresponding moment matrices $M(T; y)$. Specifically, since $V(G_K) \subseteq V(G)$, every S in $\mathcal{P}_{t+1}(V(G_K))$ also belongs to $\mathcal{P}_{t+1}(V(G))$, and every T in $\mathcal{P}_{=t-1}(V(G_K))$ also belongs to $\mathcal{P}_{=t-1}(V(G))$. Therefore, we can define y^K by

$$y_S^K := y_S$$

for each S in $\mathcal{P}_{t+1}(V(G_K))$. Then, for each T in $\mathcal{P}_{=t-1}(V(G_K))$, the corresponding moment matrix $M_K(T; y^K)$ is a principal submatrix of $M(T; y)$ (with the rows and columns corresponding to vertices in $V(G)$ that are not in $V(G_K)$ removed). Since $M(T; y) \succeq 0$, $M_K(T; y^K)$ is also PSD. By construction, the $L^t(\overline{G}_K)$ objective evaluated at the $L^t(\overline{G}_K)$ feasible y^K point is

$$\sum_{i \in V(G_K)} y_i^K \leq \sum_{i \in V(G_K)} y_i$$

which lower bounds the optimal value of $L^t(\overline{G}_K)$ □

G Proof of Corollary 3.10

The Proposition G.1 below gives an estimate on the quantity $\frac{|V(G_{p,K})|}{|V(G_p)|} \geq \left(\frac{1}{2^a} - \frac{a-1}{\sqrt{p}} + \frac{a}{2p}\right)$. Combining this with the Corollary 3.7 and Proposition 3.9, we see that

$$\begin{aligned} L^t(\overline{G}_{p,K}) &\geq \left(\frac{1}{2^a} - \frac{a-1}{\sqrt{p}} + \frac{a}{2p}\right)\left(\frac{\sqrt{p}}{2^{t-1}} + O\left(\frac{1}{2^t}\right)\right) \\ &\geq \frac{\sqrt{p}}{2^{t+a-1}} + O\left(\frac{1}{2^{a+t}} - \frac{a}{2^{t-1}}\right) = \frac{\sqrt{p}}{2^{a+t-1}} + O\left(\frac{a}{2^t}\right). \end{aligned}$$

Proposition G.1. *For any clique K of G_p of size a,*

$$\left| |V(G_{p,K})| - \frac{p}{2^a} \right| \leq (a-1)\sqrt{p} + \frac{a}{2} \tag{49}$$

This proposition is a restatement of Proposition 4.3 in [Kun24], adjusted for the fact that our localization $G_{p,K}$ is induced on vertices of G_p adjacent to a clique K rather than an equivalent localization induced on vertices not adjacent to an independent set.

the fact that, for any $x \in V(G_{p,K})$ with $K \in \mathcal{K}_a$, we have $\{x\} \cup K \in \mathcal{K}_{a+1}$, and $\mathsf{deg}_x(G_{p,K}) = |V(G_{p,K \cup \{x\}})|$ satisfies the estimate (49) with $|K \cup \{x\}| = a + 1$.