

A Degree 4 Sum-Of-Squares Lower Bound for the Clique Number of the Paley Graph

Dmitriy Kunisky ✉🏠

Department of Computer Science, Yale University, New Haven, CT, USA

Xifan Yu ✉

Department of Computer Science, Yale University, New Haven, CT, USA

Abstract

We prove that the degree 4 sum-of-squares (SOS) relaxation of the clique number of the Paley graph on a prime number p of vertices has value at least $\Omega(p^{1/3})$. This is in contrast to the widely believed conjecture that the actual clique number of the Paley graph is $O(\text{polylog}(p))$. Our result may be viewed as a derandomization of that of Deshpande and Montanari (2015), who showed the same lower bound (up to $\text{polylog}(p)$ terms) with high probability for the Erdős-Rényi random graph on p vertices, whose clique number is with high probability $O(\log(p))$. We also show that our lower bound is optimal for the Feige-Krauthgamer construction of pseudomoments, derandomizing an argument of Kelner. Finally, we present numerical experiments indicating that the value of the degree 4 SOS relaxation of the Paley graph may scale as $O(p^{1/2-\varepsilon})$ for some $\varepsilon > 0$, and give a matrix norm calculation indicating that the pseudocalibration construction for SOS lower bounds for random graphs will not immediately transfer to the Paley graph. Taken together, our results suggest that degree 4 SOS may break the “ \sqrt{p} barrier” for upper bounds on the clique number of Paley graphs, but prove that it can at best improve the exponent from $1/2$ to $1/3$.

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization; Theory of computation → Semidefinite programming; Mathematics of computing → Combinatorial optimization

Keywords and phrases convex optimization, sum of squares, Paley graph, derandomization

Digital Object Identifier 10.4230/LIPIcs.CCC.2023.30

Related Version *Full Version:* <https://arxiv.org/abs/2211.02713>

Funding This work was partially supported by ONR Award N00014-20-1-2335 and a Simons Investigator Award from the Simons Foundation to Daniel Spielman.

Acknowledgements We thank Afonso Bandeira, Chris Jones, and Daniel Spielman for helpful discussions, and the anonymous reviewers for their careful reading of the paper.

1 Introduction

1.1 Maximum and Planted Clique Problems in Random Graphs

For a graph G , we denote by $\omega(G)$ the number of vertices in the largest *clique* or complete subgraph in G . Computing $\omega(G)$ is a classical NP-hard problem in combinatorial optimization, which is moreover hard to approximate within any polynomial factor $n^{1-\varepsilon}$ for $\varepsilon > 0$ [31, 24]. Aside from this worst-case hardness, an average-case setting of computing $\omega(G)$ was proposed by Karp [32]. In this setting, the input graph is an Erdős-Rényi (ER) random graph G on n vertices, where each edge is present independently with probability $\frac{1}{2}$. We denote this by distribution by $G \sim \mathcal{G}(n, \frac{1}{2})$. It is known that (see, e.g., [5, Section 11.1]), with high probability,

$$\omega(G) \in [(2 - o(1)) \log_2 n, (2 + o(1)) \log_2 n]. \quad (1)$$



© Dmitriy Kunisky and Xifan Yu;
licensed under Creative Commons License CC-BY 4.0
38th Computational Complexity Conference (CCC 2023).
Editor: Amnon Ta-Shma; Article No. 30; pp. 30:1–30:25



Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



In [32], Karp showed that a simple greedy algorithm with high probability finds a clique of size roughly $\log_2 n$, and asked whether a polynomial-time algorithm can with high probability find a clique of size $(1 + \varepsilon) \log n$ for any constant $\varepsilon > 0$. The problem remains open, but, perhaps surprisingly, evidence has accumulated that such an algorithm does not exist [28, 17].

A natural related problem is that of algorithmically *bounding* the size of the largest clique in G , outputting a number that is always an upper bound on $\omega(G)$. For example, under $G \sim \mathcal{G}(n, \frac{1}{2})$, a simple algorithm based on the maximum degree can produce a $O(\sqrt{n \log n})$ bound [34]. Spectral algorithms operating on the eigenvalues of the adjacency matrix of G can improve this to $O(\sqrt{n})$ (for instance, using Haemers’ generalization to irregular graphs of Hoffman’s classical spectral bound on the clique number [22]).

The question of algorithmically bounding the clique number is also related to the problem of *hypothesis testing* between $G \sim \mathcal{G}(n, \frac{1}{2})$ and G drawn from another distribution where a typical G contains a *planted* clique of size much larger than $2 \log_2 n$, since if we have an algorithm that always produces a valid bound on $\omega(G)$ and this bound is typically small for $G \sim \mathcal{G}(n, \frac{1}{2})$, then we can use its output to detect the planting of a sufficiently large clique. The above then shows that we may detect the presence of a clique of size $C\sqrt{n}$ for sufficiently large C ; [2] moreover showed that an efficient spectral algorithm can even *recover* the vertex set of a planted clique of this size.¹

A long line of work considered whether using convex relaxations of $\omega(G)$ that produce bounds that are in general stronger than spectral bounds can break this “ \sqrt{n} barrier” for $G \sim \mathcal{G}(n, \frac{1}{2})$, with a particular focus on semidefinite programming (SDP) relaxations. [30] showed that Lovász’s ϑ function [37] also has value $\Omega(\sqrt{n})$; [14] later considered further aspects of using the ϑ function for detecting and recovering planted cliques. [15] showed the same $\Omega(\sqrt{n})$ lower bound for any constant level of the Lovász-Schrijver hierarchy of SDPs, of which the ϑ function is merely the first and weakest. The stronger sum-of-squares (SOS) hierarchy of relaxations proved harder to analyze. The pioneering but flawed analysis of [40] was fixed by [39], albeit at the cost of falling short of an $\Omega(\sqrt{n})$ lower bound. Many subsequent works, first on the degree 4 SOS relaxation [9, 45, 27] and culminating in the development of the *pseudocalibration* technique for larger degrees [4], ultimately established an $\Omega(n^{1/2-o(1)})$ lower bound for any constant degree of the SOS hierarchy.²

All of these results apply, as we have mentioned, to the average case of computing the clique number over $G \sim \mathcal{G}(n, \frac{1}{2})$. Some recent literature has revisited other average-case SOS lower bounds and identified *deterministic* instances over which the same quality of lower bound holds (see in particular the work of [11, 26], derandomizing the result of [19] on refuting 3-XORSAT instances).³ In this paper, we initiate the study of the same question for the clique problem, by derandomizing the SOS lower bound of [9] for the degree 4 SOS relaxation of $\omega(G)$ with $G \sim \mathcal{G}(n, \frac{1}{2})$. The deterministic graphs that achieve this derandomization are the *Paley graphs*, whose clique number is a question of independent interest in number theory. We first review some background on the Paley graphs, and then describe our results.

¹ Observe that, while a brute force search can both detect and recover a planted clique of any size $(2 + \varepsilon) \log_2 n$, this brute force search does not run in polynomial time.

² The SOS hierarchy consists of a sequence of SDPs producing smaller and smaller upper bounds on $\omega(G)$, indexed by an even number called the *degree*. See Section 2.2.1 for a precise definition.

³ Here we are interested in quantitative lower bounds showing large integrality gaps, rather than arbitrarily small integrality gaps – deterministic explicit examples giving the latter for high degrees of SOS have been shown before for several problems in works such as [19, 35].

1.2 Paley Graphs, Pseudorandomness, and Derandomization

The Paley graphs are an infinite family of graphs that exhibit certain *pseudorandom* properties, behaving in some regards similarly to a typical $G \sim \mathcal{G}(n, \frac{1}{2})$. They are defined on vertex sets identified with finite fields \mathbb{F}_q of order $q \equiv 1 \pmod{4}$, where edges connect pairs of elements of \mathbb{F}_q whose differences are quadratic residues. We denote the Paley graph on \mathbb{F}_q by G_q ; the reader may see Section 2.2.2 for a more precise definition.

Many quantities that may be computed from Paley graphs are the same as those of typical graphs drawn from $\mathcal{G}(q, \frac{1}{2})$. In the simplest instance, Paley graphs are regular of degree $\frac{q-1}{2}$, roughly the average degree of the corresponding random graph. [7] showed that the same holds for the number of occurrences of any subgraph of constant size, for the first eigenvalue being asymptotically $\frac{q}{2}$, and the second eigenvalue being $o(q^{\frac{1}{2}+\epsilon})$ for any $\epsilon > 0$.

How far can we take this analogy? It is natural to ask for subgraph counts of graphs of size growing slowly with q , and the clique number is just such a question: under $G \sim \mathcal{G}(q, \frac{1}{2})$ we have $\mathbb{E}[\omega(G)] \sim 2 \log_2 q$, and we might expect the same for $\omega(G_q)$.

However, the clique number of Paley graphs is not well understood. Let us review what is currently known. Hoffman's spectral bound [25, 22] implies the upper bound

$$\omega(G_q) \leq \sqrt{q}. \quad (2)$$

In fact, this is easy to derive by elementary combinatorial means (see, e.g., [47]) and for this reason is sometimes called the *trivial* upper bound on $\omega(G_q)$. This is tight for $q = p^{2k}$ an even power of a prime, as $\mathbb{F}_{\sqrt{q}}$ may be realized as a subfield of \mathbb{F}_q all of whose elements are quadratic residues in this case [6].

However, for odd prime powers, and even the simplest case $q = p$ a prime, the clique number is believed to be much lower. The upper bound on the diagonal Ramsey number established by [12] implies that

$$\omega(G_p) \geq \left(\frac{1}{2} + o(1) \right) \log_2 p. \quad (3)$$

By a number-theoretic analysis of the least quadratic non-residue modulo p , [18] improved this, showing that for infinitely many primes p ,

$$\omega(G_p) \geq \log p \log \log \log p. \quad (4)$$

Moreover, conditional on the Generalized Riemann Hypothesis, the $\log \log \log p$ term may be improved to $\log \log p$ [41, Theorem 13.5].⁴

On the other hand, the best known upper bound [23, 10] improves only by a constant factor on the spectral bound (2),

$$\omega(G_p) \leq \frac{\sqrt{2p-1} + 1}{2} \sim \frac{\sqrt{p}}{\sqrt{2}}. \quad (5)$$

In contrast to this state-of-the-art bound, $\omega(G_p)$ is widely believed to actually scale at most polylogarithmically with p based on computations of $\omega(G_p)$ for small p . We express this in the following conjecture; see [46, 3, 47, 33] as well as our Figure 2.

► **Conjecture 1.** *For some $C, K > 0$ and all $p \equiv 1 \pmod{4}$ prime, $\omega(G_p) \leq C(\log p)^K$.*

Numerical evidence suggests that we might in fact expect to be able to take $K = 2$, as discussed by [3, 33] and illustrated in our Figure 2.

⁴ It is still possible to reconcile these results with the proposal that G_p behaves like a random graph, so long as we adopt a more sophisticated random model than $\mathcal{G}(p, \frac{1}{2})$ [42].

Moreover, these graphs are believed to be good constructions for lower bounds on the diagonal Ramsey numbers $R(k, k)$. For example, the Paley graph of order 17 is the unique largest graph that contains neither a clique of size 4 nor an independent set of size 4, which shows that $R(4, 4) = 18$ [13]. The current best known bound $R(6, 6) \geq 102$ is established by the Paley graph of order 101, which contains neither a clique of size 6 nor an independent set of size 6 [44].

Because of this application among others, it is a long-standing open problem in additive combinatorics and number theory to improve the upper bound for clique numbers of Paley graphs of prime orders, and in particular to break the “ \sqrt{p} barrier” and prove an upper bound scaling as $p^{1/2-\varepsilon}$ for some $\varepsilon > 0$.⁵ Some recent work has begun to explore whether convex relaxations of the clique number can lead to such improvements. For instance, [21, 33] explored using a hierarchy of SDPs producing bounds between that of the Lovász-Schrijver hierarchy and the SOS hierarchy for this purpose, and [38] empirically found that a modification of the Lovász ϑ function SDP can recover and sometimes slightly improve on the best-known upper bound (5).

1.3 Our Contributions

Our main result contributes to both of the lines of work outlined above. On the one hand, it shows (conditional on Conjecture 1) that the Paley graph gives a derandomization of the SOS lower bound of [9] for ER random graphs. On the other hand, it shows that a powerful convex optimization approach to upper-bounding the clique number cannot be too effective when applied to G_p .

► **Theorem 2.** *There is a constant $c > 0$ such that the value of the degree 4 SOS relaxation of the clique number $\text{SOS}_4(G)$, as defined in Section 2.2.1, evaluated with G_p the Paley graph on p vertices for p any prime number with $p \equiv 1 \pmod{4}$, as defined in Section 2.2.2, satisfies*

$$\text{SOS}_4(G_p) \geq cp^{1/3}. \quad (6)$$

The main ingredients in our proof are new norm bounds for certain *graph matrices* (as appear in the analysis of SOS relaxations for random graphs; see, e.g., [1]) formed from Paley graphs and certain character sum estimates for the Legendre symbol.

To elaborate on this result, we provide three further pieces of more detailed analysis. Note that Theorem 2 does not exclude the possibility that $\text{SOS}_4(G_p) = o(\sqrt{p})$. In Section 4.1, however, we show that at least the lower bound construction we use to prove Theorem 2, involving the simple class of *Feige-Krauthgamer pseudomoments* (see Definition 6), cannot improve on the $p^{1/3}$ scaling of our lower bound.

On the other hand, in Section 4.2, we present some numerical evidence that $\text{SOS}_4(G_p) \sim p^\eta$ for a constant $\eta \in (0, \frac{1}{2})$, with value $\eta \approx 0.4$. As we discuss in Section 4.2, these results are similar to earlier numerical studies of [21], who consider a weaker class of SDPs than the SOS hierarchy, and results of [33], who consider the same weaker SDPs and extract a prediction of the power scaling of their values with p from numerical results. We thus have reason to believe that our lower bound cannot be improved all the way to a scaling of $p^{1/2}$. Unfortunately, we have not found a way to convert these numerical results into a proof of an improved bound on the clique number, but we leave this as a tantalizing open problem for future work.

⁵ For instance, this is mentioned as “probably a very hard problem” in the problem list [8].

Finally, to accompany these empirical results, we provide some modest theoretical evidence that the SOS hierarchy may break the \sqrt{p} barrier for upper bounds on $\omega(G_p)$. The tight analysis showing that $\mathbb{E}[\text{SOS}_{2d}(G)] = \Omega(n^{1/2-o(1)})$ for $G \sim \mathcal{G}(n, \frac{1}{2})$ and any constant d uses a construction satisfying a property called *pseudocalibration* [4], whose analysis hinges on norm bounds for the aforementioned graph matrices built from the adjacency matrix of G [1]. In Section 4.3, we show that some of these norm bounds *fail* for the Paley graph. Thus, the analysis of the pseudocalibration construction for random graphs cannot be directly adapted to the case of Paley graphs.⁶

2 Preliminaries and Proof Overview

2.1 Notations

Throughout the paper, p will denote a prime number, and q a prime power $q = p^k$. The finite field of order q (unique up to isomorphism) is denoted by \mathbb{F}_q , and its group of units by \mathbb{F}_q^\times . A nonzero element y of \mathbb{F}_q is called a *quadratic residue* of \mathbb{F}_q if $y = x^2$ for some $x \in \mathbb{F}_q$, and a *quadratic nonresidue* otherwise. We write $(\mathbb{F}_q^\times)^2$ for the set of quadratic residues. We will also freely identify \mathbb{F}_p with $\mathbb{Z}/p\mathbb{Z}$, with representatives $\{0, 1, \dots, p-1\}$.

We write $[n] := \{1, 2, \dots, n\}$. For a finite set X , we write 2^X for the power set, and $\binom{X}{k}$ and $\binom{X}{\leq k}$ to denote the sets of subsets of X with exactly k elements and at most k elements respectively. We also use $X_{(k)}$ to denote the set of tuples of elements of X of length k with all entries distinct.

When the discussion involves variables $\{x_i\}_{i \in \mathcal{I}}$ indexed by \mathcal{I} , for a subset $S \subset \mathcal{I}$, we will use x^S to denote the monomial $\prod_{i \in S} x_i$.

We use $\mathbf{1} \in \mathbb{R}^n$ to denote the all-ones vector. We use $I \in \mathbb{R}^{n \times n}$ to denote the identity matrix, $J \in \mathbb{R}^{n \times n}$ to denote the all-ones matrix, and $\mathbf{0} \in \mathbb{R}^{n \times n}$ to denote the all-zeros matrix. The dimensions of these objects will be clear from context. For a real symmetric or Hermitian matrix A , we use $\text{spec}(A)$ to denote its spectrum, which we write in double braces $\{\{\dots\}\}$ to indicate that the spectrum is a multiset. For matrices $A, B \in \mathbb{C}^{n \times n}$ and $C \in \mathbb{C}^{m \times m}$, we use $A \circ B \in \mathbb{C}^{n \times n}$ to denote the Hadamard product (entrywise product) of A and B , and $A \otimes C \in \mathbb{C}^{nm \times nm}$ to denote the Kronecker product (tensor product) of A and C .

For a graph $G = (V, E)$, we use $V(G)$ to denote its vertex set and $E(G)$ to denote its edge set. We use \overline{G} to denote the complement of G . For vertices $u, v \in V(G)$, we use $u \sim_G v$ to indicate that u and v are adjacent in G and $u \not\sim_G v$ to indicate that they are not adjacent. We will use A_G to denote the $\{0, 1\}$ adjacency matrix of G , and S_G to denote the Seidel or $\{\pm 1\}$ adjacency matrix. We drop the subscript G when the graph is clear from context. Conventionally, the Seidel adjacency matrix is -1 on pairs of adjacent vertices, $+1$ on pairs of nonadjacent vertices, and 0 on the diagonal. In this paper, we abuse this term to mean the matrix that is 1 on pairs of adjacent vertices, -1 on pairs of nonadjacent vertices, and 0 on the diagonal, as this is more conveniently written in terms of the Legendre symbol in the context of Paley graphs (see Section 3.4). It is easy to see that the A_G and S_G are related by $S_G = 2A_G - J + I$. We write $\mathcal{K}(G)$ for the set of subsets of $V(G)$ that form cliques in G .

We will use the standard asymptotic notations $O(\cdot)$, $\Omega(\cdot)$, $\Theta(\cdot)$, and $o(\cdot)$. We will use $\tilde{O}(\cdot)$ and $\tilde{\Omega}(\cdot)$ to additionally suppress polylogarithmic factors.

⁶ We note that the initial premise of pseudocalibration, which involves comparing a pair of “null” and “alternative” random graph distributions, is not sensible to apply to the deterministic Paley graph. But, ultimately, the pseudocalibration argument yields a function mapping a graph to a matrix that one hopes will be feasible for a high-degree SOS program, and one may simply substitute the Paley graph into this function and consider the result.

2.2 Problem Setup

Let us now specify in full detail the SOS relaxations SOS_{2d} of the clique number, and the Paley graphs G_p .

2.2.1 Sum-Of-Squares Relaxations of the Clique Number

Let G be a graph of order n . The clique number $\omega(G)$ of G is equal to the value of the following polynomial optimization program:

$$\omega(G) = \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i \in V(G)} x_i \\ \text{subject to} \quad x_i^2 = x_i \text{ for all } i \in V(G), \\ \quad \quad \quad x_i x_j = 0 \text{ for all } i, j \in V(G) \text{ with } i \neq j \text{ and } i \not\sim_G j \end{array} \right\}. \quad (7)$$

It is easy to see that the feasible solutions of the program above are in one-to-one correspondence with the indicator vectors of the cliques in G . Before we introduce the SOS relaxations of the clique number, let us first define the *pseudoexpectation* operators over which the SOS relaxations optimize.

► **Definition 3** (Pseudoexpectation). *We say $\tilde{\mathbb{E}} : \mathbb{R}[x_1, \dots, x_n]_{\leq 2d} \rightarrow \mathbb{R}$ is a degree $2d$ pseudoexpectation with respect to polynomial constraints $\{f_i(x) = 0\}_{i=1}^a$, $\{g_j(x) \geq 0\}_{j=1}^b$ if the following properties hold:*

- $\tilde{\mathbb{E}}$ is linear.
- $\tilde{\mathbb{E}}[1] = 1$.
- $\tilde{\mathbb{E}}[f_i(x)p(x)] = 0$, for all $p(x) \in \mathbb{R}[x_1, \dots, x_n]$ such that $\deg(f_i p) \leq 2d$, for all $1 \leq i \leq a$.
- $\tilde{\mathbb{E}}[p(x)^2] \geq 0$, for all $p(x) \in \mathbb{R}[x_1, \dots, x_n]_{\leq d}$.
- $\tilde{\mathbb{E}}[g_j(x)p(x)^2] \geq 0$, for all $p(x) \in \mathbb{R}[x_1, \dots, x_n]$ such that $\deg(g_j p^2) \leq 2d$, for all $1 \leq j \leq b$.

In the case of the maximum clique program (7), the polynomial constraints are generated by the Boolean constraints $x_i^2 - x_i = 0$ for $i \in V(G)$ and the clique constraints $x_i x_j = 0$ for $i, j \in V(G)$ with $i \neq j$ and $i \not\sim_G j$. For convenience, let us identify the vertex set $V(G)$ with $[n]$ where $n = |V(G)|$. Then, the degree $2d$ SOS relaxation of the polynomial optimization program (7) written in terms of pseudoexpectations is

$$\text{SOS}_{2d}(G) = \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i=1}^n \tilde{\mathbb{E}}[x_i] \\ \text{subject to} \quad \tilde{\mathbb{E}} : \mathbb{R}[x_1, \dots, x_n]_{\leq 2d} \rightarrow \mathbb{R} \text{ linear,} \\ \quad \quad \quad \tilde{\mathbb{E}}[1] = 1, \\ \quad \quad \quad \tilde{\mathbb{E}}[(x_i^2 - x_i)p(x)] = 0 \text{ for all } i \in [n], \deg(p) \leq 2d - 2, \\ \quad \quad \quad \tilde{\mathbb{E}}[x_i x_j p(x)] = 0 \text{ for all } i \not\sim_G j, \deg(p) \leq 2d - 2, \\ \quad \quad \quad \tilde{\mathbb{E}}[p(x)^2] \geq 0 \text{ for all } \deg(p) \leq d. \end{array} \right\}. \quad (8)$$

To see that this is indeed a relaxation of the clique program (7), observe that for any probability measure $\mu : 2^{[n]} \rightarrow \mathbb{R}^{\geq 0}$ supported on the cliques of the graph G , the corresponding expectation operator \mathbb{E}_μ is a pseudoexpectation of any degree.

For every monomial x^S for $S \in \binom{[n]}{\leq 2d}$, $\tilde{\mathbb{E}}[x^S]$ is called the *pseudomoment* of S of the corresponding pseudoexpectation $\tilde{\mathbb{E}}$. By linearity, every pseudoexpectation of degree $2d$ is uniquely determined by its pseudomoments of degree at most $2d$, i.e., by the set $\{\tilde{\mathbb{E}}[x^S] : S \subseteq [n], |S| \leq 2d\}$. We may therefore encode the pseudoexpectation in the *pseudomoment matrix* $M \in \mathbb{R}_{\text{sym}}^{\binom{[n]}{\leq 2d} \times \binom{[n]}{\leq 2d}}$ with entries

$$M_{S,T} = \tilde{\mathbb{E}}[x^S x^T]. \quad (9)$$

This is especially convenient since the positivity of $\tilde{\mathbb{E}}$ on squared polynomials is equivalent to positive semidefiniteness of M . We can then rewrite the above program (8) in the form of an SDP:

$$\text{SOS}_{2d}(G) = \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i=1}^n M_{\emptyset, \{i\}} \\ \text{subject to} \quad M \in \mathbb{R}^{\binom{[n]}{\leq d} \times \binom{[n]}{\leq d}} \\ M_{\emptyset, \emptyset} = 1, \\ M_{S,T} \text{ depends only on } S \cup T, \\ M_{S,T} = 0 \text{ whenever } S \cup T \notin \mathcal{K}(G), \\ M \succeq 0. \end{array} \right\}. \quad (10)$$

We will not verify in detail the equivalence of (10) and (8); the reader may consult [36] for an overview of this pseudomoment matrix framework, or the papers [9, 45, 27, 4] on SOS relaxations of $\omega(G)$ for further details.

► **Remark 4 (Pseudomoment matrix compression).** We note that the row and column of M indexed by any $S \notin \mathcal{K}(G)$ is forced by the constraints to be identically zero. These entries do not affect the positivity of M and do not play a role in the objective function, so we may just as well take M to be indexed by *cliques* of size at most d rather than arbitrary subsets of vertices.

In the special case $2d = 2$, the SDP in (10) takes the form

$$\text{SOS}_2(G) = \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i=1}^n y_i \\ \text{subject to} \quad y \in \mathbb{R}^n, Y \in \mathbb{R}_{\text{sym}}^{n \times n}, \\ Y_{i,i} = y_i \text{ for all } i \in [n], \\ Y_{i,j} = 0 \text{ for all } i, j \in [n] \text{ with } i \neq j \text{ and } i \not\sim_G j, \\ M = \begin{bmatrix} 1 & y^\top \\ y & Y \end{bmatrix} \succeq 0. \end{array} \right\}. \quad (11)$$

One can show (see [20, 16]) that the program above is equivalent to the *Lovász ϑ function* of the complement graph \overline{G} , a well-known upper bound on $\omega(G)$ due to [37]:

$$\text{SOS}_2(G) = \vartheta(\overline{G}). \quad (12)$$

This SDP enjoys many special properties, some of which we will mention below; the reader may consult the above references for further information.

On the other hand, once the degree increases to $2d = 4$, the resulting SDP is not as well understood. This SDP, which we study in the remainder of the paper, takes the form

$$\text{SOS}_4(G) = \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i=1}^n M_{\emptyset, \{i\}}^{0,1} \\ \text{subject to} \quad M^{r,c} \in \mathbb{R}^{\binom{[n]}{r} \times \binom{[n]}{c}} \text{ for } r, c \in \{0, 1, 2\}, \\ M_{S,T}^{r,c} \text{ depends only on } S \cup T, \\ M_{S,T}^{r,c} = 0 \text{ whenever } S \cup T \notin \mathcal{K}(G), \\ M = \begin{bmatrix} 1 & M^{0,1} & M^{0,2} \\ M^{1,0} & M^{1,1} & M^{1,2} \\ M^{2,0} & M^{2,1} & M^{2,2} \end{bmatrix} \succeq 0 \end{array} \right\}. \quad (13)$$

2.2.2 Paley Graphs

We now give the definition and some useful basic properties of the Paley graphs.

► **Definition 5** (Paley graph). *Let $q = p^k$ be a prime power such that $q \equiv 1 \pmod{4}$. The Paley graph G_q of order q then has vertex set $V(G_q) := \mathbb{F}_q$ and edge set*

$$E(G_q) := \left\{ \{a, b\} \in \binom{\mathbb{F}_q}{2} : a - b \in (\mathbb{F}_q^\times)^2 \right\}. \tag{14}$$

The condition $q \equiv 1 \pmod{4}$ ensures that -1 is a square in \mathbb{F}_q . As a result, $a - b \in (\mathbb{F}_q^\times)^2$ if and only if $b - a \in (\mathbb{F}_q^\times)^2$, so the edge set is well-defined.

We will study the SOS relaxations of the clique number of Paley graphs, $\text{SOS}_{2d}(G_q)$. Recall that the degree 2 SOS relaxation of the clique number of the Paley graph G_q is equal to the Lovász theta function of its complement, $\text{SOS}_2(G_q) = \vartheta(\overline{G_q})$. Since G_q is self-complementary (under the automorphism $x \mapsto gx$ for g a multiplicative generator of \mathbb{F}_q^\times), $\vartheta(\overline{G_q}) = \vartheta(G_q)$. Since G_q is vertex-transitive, by Lovász’s result in [37],

$$\vartheta(\overline{G_q})\vartheta(G_q) = |V(G_q)| = q, \tag{15}$$

whereby combining our observations shows that

$$\text{SOS}_2(G_q) = \sqrt{q}. \tag{16}$$

This is the same as the upper bound of the clique number given by Hoffman’s spectral bound. Thus, degree 2 SOS does not improve on the spectral bound, and degree 4 SOS, which we begin to analyze with Theorem 2, is the first more interesting degree.

2.3 Proof Overview

To prove Theorem 2, we will construct a feasible pseudomoment matrix M for the program (13) that has objective value $\Omega(p^{1/3})$. We will consider the following type of pseudomoments, which we call *Feige-Krauthgamer (FK) pseudomoments*, first studied by Feige and Krauthgamer [15] to prove lower bounds on Lovász-Schrijver relaxations for the maximum independent set of random graphs (sometimes these are called *MPW pseudomoments* after their use by the later paper [39]).

► **Definition 6** (Feige-Krauthgamer pseudomoments). *Consider the degree $2d$ SOS relaxation of the clique number of a graph G . We say the pseudomoments of a degree $2d$ pseudoexpectation $\tilde{\mathbb{E}}$ are Feige-Krauthgamer (FK) pseudomoments if there exists a sequence $1 = \alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{2d} \in \mathbb{R}$ such that*

$$\tilde{\mathbb{E}}[x^S] = \begin{cases} \alpha_{|S|} & \text{if } S \in \mathcal{K}(G) \text{ (i.e., if } S \text{ is a clique in } G) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

We note that FK pseudomoments automatically satisfy all conditions on a pseudoexpectation other than positivity.

The line of work beginning with [40] sought to use FK pseudomoments to prove lower bounds on SOS relaxations of $\omega(G)$ for random graphs G .⁷ While eventually in [27, 4] it was found that FK pseudomoments could *not* prove optimal $\Omega(\sqrt{n})$ lower bounds, earlier works still proved polynomial $\Omega(n^\eta)$ lower bounds with $\eta < \frac{1}{2}$ using FK pseudomoments,

⁷ Some works, wanting to study an SOS relaxation that included the “exact” constraint $\sum_{i=1}^n x_i = k$ for some k , adjusted the FK pseudomoments to satisfy the consequences of this constraint (see, e.g., [27, 43]). We do not take this route here.

which are simpler to define and to work with than the alternatives developed later. In particular, our analysis will closely follow that of [9], who used FK pseudomoments to prove that $\text{SOS}_4(G) = \tilde{\Omega}(n^{\frac{1}{3}})$ with high probability for $G \sim \mathcal{G}(n, \frac{1}{2})$. [27] later showed that, up to polylogarithmic factors, this is optimal over any choice of FK pseudomoments for the degree 4 relaxation.

► **Remark 7 (Partial symmetry).** By vertex transitivity and edge transitivity of Paley graphs, there always exists an optimal degree 4 pseudoexpectation giving all $\tilde{\mathbb{E}}[x_i]$ the same value and all $\tilde{\mathbb{E}}[x_i x_j]$ with $i \sim j$ in G_p the same value, regardless of whether $\tilde{\mathbb{E}}$ is given by FK pseudomoments or not. This strong symmetry of course fails to hold for ER random graphs.

Recall that in the degree 4 SOS program (13), we write the pseudomoment matrix M in the block form

$$M = \begin{bmatrix} 1 & M^{0,1} & M^{0,2} \\ M_{1,0} & M^{1,1} & M^{1,2} \\ M_{2,0} & M^{2,1} & M^{2,2} \end{bmatrix}. \quad (18)$$

We will follow the strategy of [9] to successively check the Schur complement conditions for positive semidefiniteness of M . Namely, we will rely on the following fact.

► **Proposition 8.** *Let*

$$M = \begin{bmatrix} A & B^\top \\ B & C \end{bmatrix} \in \mathbb{R}^{(a+b) \times (a+b)} \quad (19)$$

be a real symmetric matrix written in block form, with $A \in \mathbb{R}^{a \times a}$ and $C \in \mathbb{R}^{b \times b}$. If $A \succ 0$ and $C - BA^{-1}B^\top \succeq 0$, then $M \succeq 0$. We call the matrix $C - BA^{-1}B^\top$ the Schur complement of the block A in M .

3 Proof of Theorem 2

We restate Theorem 2 in more detailed terms of the FK pseudomoments that we will construct.

► **Theorem 9.** *There exists a constant $c > 0$ so that, setting $\alpha_1 := cp^{-2/3}$, $\alpha_2 := 4\alpha_1^2$, $\alpha_3 := 8\alpha_1^3$, and $\alpha_4 := 512\alpha_1^4$, the FK pseudomoments defined by these parameters give a feasible solution to the degree 4 SOS relaxation (13) of the clique number of the Paley graphs G_p for all sufficiently large p .*

Theorem 2 follows, since the above gives, for all sufficiently large p ,

$$\text{SOS}_4(G_p) \geq p \cdot cp^{-2/3} = cp^{1/3}. \quad (20)$$

To remind the reader of the notations we set in the previous section, the pseudomoment matrix in the degree 4 SOS relaxation (13) is denoted

$$M = \begin{bmatrix} 1 & M^{0,1} & M^{0,2} \\ M^{1,0} & M^{1,1} & M^{1,2} \\ M^{2,0} & M^{2,1} & M^{2,2} \end{bmatrix}, \quad (21)$$

and we take this to be given by the FK pseudomoments proposed in Theorem 9. Recall that $M^{r,c} \in \mathbb{R}^{\binom{p}{r} \times \binom{p}{c}}$ for all $r, c \in \{0, 1, 2\}$. We will use

$$N = \begin{bmatrix} N^{1,1} & N^{1,2} \\ N^{2,1} & N^{2,2} \end{bmatrix} \quad (22)$$

to denote the Schur complement of the top left 1×1 block in M .

3.1 Filling Zero Rows and Columns

As mentioned before, we will fill in the zero rows and columns of N in order to make use of graph matrices. In this section, we define the matrix

$$H = \begin{bmatrix} H^{1,1} & H^{1,2} \\ H^{2,1} & H^{2,2} \end{bmatrix} \quad (23)$$

that will achieve this filling.

► **Definition 10.** We write $\mathbb{1}_k : \binom{\mathbb{F}_p}{k} \rightarrow \{0, 1\}$ for the function with $\mathbb{1}_k(S) = 1$ if S is a clique in G_p and $\mathbb{1}_k(S) = 0$ otherwise.

We now expand the $N^{\bullet,\bullet}$ matrices in terms of this indicator function.

► **Proposition 11.** Under the FK pseudomoments proposed in Theorem 9, the matrix N can be written as

$$N = \begin{bmatrix} N^{1,1} & N^{1,2} \\ N^{2,1} & N^{2,2} \end{bmatrix}, \quad (24)$$

where $N^{1,1} \in \mathbb{R}^{\mathbb{F}_p \times \mathbb{F}_p}$, $N^{1,2} \in \mathbb{R}^{\mathbb{F}_p \times \binom{\mathbb{F}_p}{2}}$, $N^{2,1} = N^{1,2^\top} \in \mathbb{R}^{\binom{\mathbb{F}_p}{2} \times \mathbb{F}_p}$, $N^{2,2} \in \mathbb{R}^{\binom{\mathbb{F}_p}{2} \times \binom{\mathbb{F}_p}{2}}$ have entries

$$N_{a,b}^{1,1} = \begin{cases} \alpha_1 - \alpha_1^2 & \text{if } a = b, \\ \alpha_2 \mathbb{1}_2(\{a, b\}) - \alpha_1^2 & \text{if } a \neq b, \end{cases} \quad (25)$$

$$N_{a,\{b,c\}}^{1,2} = \begin{cases} (\alpha_2 - \alpha_1 \alpha_2) \mathbb{1}_2(\{b, c\}) & \text{if } a \in \{b, c\}, \\ \alpha_3 \mathbb{1}_3(\{a, b, c\}) - \alpha_1 \alpha_2 \mathbb{1}_2(\{b, c\}) & \text{if } a \notin \{b, c\}, \end{cases} \quad (26)$$

$$N_{\{a,b\},\{c,d\}}^{2,2} = \begin{cases} (\alpha_2 - \alpha_2^2) \mathbb{1}_2(\{a, b\}) & \text{if } \{a, b\} = \{c, d\}, \\ \alpha_3 \mathbb{1}_3(\{a, b\} \cup \{c, d\}) - \alpha_2^2 \mathbb{1}_2(\{a, b\}) \mathbb{1}_2(\{c, d\}) & \text{if } |\{a, b\} \cap \{c, d\}| = 1, \\ \alpha_4 \mathbb{1}_4(\{a, b, c, d\}) - \alpha_2^2 \mathbb{1}_2(\{a, b\}) \mathbb{1}_2(\{c, d\}) & \text{if } \{a, b\} \cap \{c, d\} = \emptyset. \end{cases} \quad (27)$$

Per Remark 4, rows and columns indexed by pairs are identically zero in any of these matrices for all pairs that are not edges in G_p .

Next, we define matrices $H^{\bullet,\bullet}$ based on the $N^{\bullet,\bullet}$ by replacing the clique indicator functions with “bipartite” versions of those indicator functions, that only depend on the presence of edges between two subsets of vertices.

► **Definition 12.** We write $\mathbb{1}_{\ell,r} : \binom{\mathbb{F}_p}{\ell} \times \binom{\mathbb{F}_p}{r} \rightarrow \{0, 1\}$ for the function with

$$\mathbb{1}_{\ell,r}(L, R) = \begin{cases} 1 & \text{if } v \sim_{G_p} w \text{ for all } v \in L \setminus R, w \in R \setminus L, \\ 0 & \text{otherwise.} \end{cases} \quad (28)$$

In other words, $\mathbb{1}_{\ell,r}(L, R) = 1$ if and only if all pairs of vertices in $\binom{L \cup R}{2}$ that don't belong simultaneously to L or R are connected in G_p .

Now we are ready to state what matrix H is: it is given by blocks $H^{1,1} \in \mathbb{R}^{\mathbb{F}_p \times \mathbb{F}_p}$, $H^{1,2} \in \mathbb{R}^{\mathbb{F}_p \times \binom{\mathbb{F}_p}{2}}$, $H^{2,1} = H^{1,2^\top}$, and $H^{2,2} \in \mathbb{R}^{\binom{\mathbb{F}_p}{2} \times \binom{\mathbb{F}_p}{2}}$ having entries

$$H_{a,b}^{1,1} = \begin{cases} \alpha_1 - \alpha_1^2 & \text{if } a = b \\ \alpha_2 \mathbb{1}_{1,1}(\{a\}, \{b\}) - \alpha_1^2 & \text{if } a \neq b \end{cases} \tag{29}$$

$$H_{a,\{b,c\}}^{1,2} = \begin{cases} \alpha_2 - \alpha_1 \alpha_2 & \text{if } a \in \{b, c\} \\ \alpha_3 \mathbb{1}_{1,2}(\{a\}, \{b, c\}) - \alpha_1 \alpha_2 & \text{if } a \notin \{b, c\} \end{cases} \tag{30}$$

$$H_{\{a,b\},\{c,d\}}^{2,2} = \begin{cases} \alpha_2 - \alpha_2^2 & \text{if } \{a, b\} = \{c, d\} \\ \alpha_3 \mathbb{1}_{2,2}(\{a, b\}, \{c, d\}) - \alpha_2^2 & \text{if } |\{a, b\} \cap \{c, d\}| = 1 \\ \alpha_4 \mathbb{1}_{2,2}(\{a, b\}, \{c, d\}) - \alpha_2^2 & \text{if } \{a, b\} \cap \{c, d\} = \emptyset \end{cases} \tag{31}$$

It is easy to see that proving positive semidefiniteness for H also proves N is positive semidefinite, due to the following observation.

► **Proposition 13.** *Up to permutation of rows and columns, N is the direct sum of the principal submatrix of H indexed by singletons and the edges of G_p with a zero matrix.*

The proof is simply that, for $|L|, |R| \leq 2$, we have $\mathbb{1}_{|L \cup R|}(L \cup R) = \mathbb{1}_{|L|, |R|}(L, R)$ so long as L is an edge if $|L| = 2$ and R is an edge if $|R| = 2$.

3.2 Second Schur Complement Bounds

Next, the goal is to prove under the same setting of Theorem 9 that $H \succeq 0$. The argument for this analysis is included in the full version of this paper and is similar to that of [9].

We will use $Q_0 = \frac{1}{p}J \in \mathbb{R}^{\mathbb{F}_p \times \mathbb{F}_p}$ to denote the orthogonal projection matrix to the constant vector, and $Q_1 = I - Q_0$ to denote the projection matrix to the orthogonal complement.

► **Proposition 14.** *Under the FK pseudomoments specified by $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ in Theorem 9, for any constant $\varepsilon > 0$, the matrix $H^{1,1}$ satisfies*

$$H^{1,1} \succeq \left(\alpha_1 + \frac{p-1}{2} \alpha_2 - p \alpha_1^2 \right) Q_0 + (1 - \varepsilon) \alpha_1 Q_1 \succ 0 \tag{32}$$

for all sufficiently large primes p .

So, if moreover we can show $H^{2,2} - H^{2,1}(H^{1,1})^{-1}H^{1,2} \succeq 0$, we can conclude the positive semidefiniteness of H . Our last simplification before proceeding to the main technical analysis is to remove the $(H^{1,1})^{-1}$ term above. Fix some constant $\varepsilon > 0$ for all future discussions, say $\varepsilon := \frac{1}{2}$. Then,

$$H^{1,1} \succeq \left(\alpha_1 + \frac{p-1}{2} \alpha_2 - p \alpha_1^2 \right) Q_0 + (1 - \varepsilon) \alpha_1 Q_1 \succ 0 \tag{33}$$

for all sufficiently large primes p , so

$$(H^{1,1})^{-1} \preceq \left(\alpha_1 + \frac{p-1}{2} \alpha_2 - p \alpha_1^2 \right)^{-1} Q_0 + ((1 - \varepsilon) \alpha_1)^{-1} Q_1, \tag{34}$$

and substituting this into the term appearing in the inequality we need to show,

$$H^{2,1}(H^{1,1})^{-1}H^{1,2} \preceq H^{2,1} \left[\left(\alpha_1 + \frac{p-1}{2} \alpha_2 - p \alpha_1^2 \right)^{-1} Q_0 + ((1 - \varepsilon) \alpha_1)^{-1} Q_1 \right] H^{1,2}. \tag{35}$$

30:12 Degree 4 SOS Lower Bound for Clique Number of Paley Graph

Note that the column sum (row sum) of $H^{2,1}$ is the same across each column (nonzero row indexed by edges of Paley graphs) due to the partial symmetry of Paley graphs. As a result, $\mathbf{1}$ is an eigenvector of $H^{2,1}H^{1,2}$, and $H^{2,1}Q_0H^{1,2} = P_0H^{2,1}H^{1,2}P_0$, where we use $P_0 = \frac{2}{p(p-1)}J \in \mathbb{R}^{\binom{p}{2} \times \binom{p}{2}}$ to denote the orthogonal projection matrix to the constant vector. Moreover, since $\mathbf{1}$ is an eigenvector of $H^{2,1}H^{1,2}$, $(I - P_0)H^{2,1}H^{1,2}P_0 = 0$. We therefore have

$$\begin{aligned} & H^{2,1} \left[\left(\alpha_1 + \frac{p-1}{2}\alpha_2 - p\alpha_1^2 \right)^{-1} Q_0 + ((1-\varepsilon)\alpha_1)^{-1} Q_1 \right] H^{1,2} \\ &= H^{2,1} \left[\left(\left(\alpha_1 + \frac{p-1}{2}\alpha_2 - p\alpha_1^2 \right)^{-1} - ((1-\varepsilon)\alpha_1)^{-1} \right) Q_0 + ((1-\varepsilon)\alpha_1)^{-1} I \right] H^{1,2} \\ &= \left(\alpha_1 + \frac{p-1}{2}\alpha_2 - p\alpha_1^2 \right)^{-1} P_0H^{2,1}H^{1,2}P_0 + ((1-\varepsilon)\alpha_1)^{-1} (I - P_0)H^{2,1}H^{1,2}(I - P_0). \end{aligned} \quad (36)$$

Thus, to show $H^{2,2} \succeq H^{2,1}(H^{1,1})^{-1}H^{1,2}$ holds for all sufficiently large primes p , it is sufficient to prove the following proposition:

► **Proposition 15.** *Under the FK pseudomoments specified by $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ in Theorem 9, for any constant $\varepsilon > 0$,*

$$\begin{aligned} H^{2,2} \succeq & \left(\alpha_1 + \frac{p-1}{2}\alpha_2 - p\alpha_1^2 \right)^{-1} P_0H^{2,1}H^{1,2}P_0 \\ & + ((1-\varepsilon)\alpha_1)^{-1} (I - P_0)H^{2,1}H^{1,2}(I - P_0) \end{aligned} \quad (37)$$

holds for all sufficiently large primes p .

3.3 Ribbons and Graph Matrices

To organize the remaining calculation, now let us review the construction of *graph matrices* that has played a role in many SOS lower bound analyses in previous literature. We will use the following definitions as appeared in the work of [29].

► **Definition 16 (Ribbon).** *A ribbon on a ground set V is a tuple $R = (V(R), E(R), A_R, B_R)$, where $(V(R), E(R))$ is a graph, and $A_R, B_R \subseteq V(R) \subseteq V$.*

► **Definition 17 (Matrix for a Ribbon).** *Let $G \in \mathbb{R}^{V \times V}$ be a real symmetric matrix whose off-diagonal entries are ± 1 and whose diagonal entries are zero. For $R = (V(R), E(R), A_R, B_R)$ a on V , the corresponding matrix $M_G(R) \in \mathbb{R}^{\binom{V}{|A_R|} \times \binom{V}{|B_R|}}$ has rows and columns indexed by the subsets of V of sizes $|A_R|$ and $|B_R|$, respectively. The entries of $M_G(R)$ is given by*

$$M_G(R)_{I,J} = \begin{cases} \prod_{\{i,j\} \in E(R)} G_{i,j} & \text{if } I = A_R \text{ and } J = B_R \\ 0 & \text{otherwise} \end{cases}. \quad (38)$$

In other words, there is only one nonzero entry of $M_G(R)$, and it is located at the row and the column corresponding to A_R and B_R .

► **Definition 18 (Isomorphisms Between Ribbons).** *Two ribbons R, S are isomorphic, or have the same shape, if there is a bijection $f : V(R) \rightarrow V(S)$ which is a graph isomorphism between $(V(R), E(R))$ and $(V(S), E(S))$ and also a bijection from A_R to A_S and from B_R to B_S .*

If we ignore the labels on the vertices of a ribbon, what remains is the shape of the ribbon.

► **Definition 19** (Shape). A shape is an equivalence class of ribbons of the same shape. Each shape has associated with it a representative $\beta = (V(\beta), E(\beta), A_\beta, B_\beta)$.

► **Definition 20** (Embedding of a Shape). Given a shape β on V and an injective function $f : V(\beta) \rightarrow V$, we let $f(\beta)$ be the ribbon by labeling the vertices $V(\beta)$ in the natural way.

► **Definition 21** (Graph Matrix). Let $G \in \mathbb{R}^{V \times V}$ be a real symmetric matrix whose off-diagonal entries are ± 1 and whose diagonal entries are zero. For a shape β on V , the graph matrix $M_G(\beta) \in \mathbb{R}^{\binom{V}{|A_\beta|} \times \binom{V}{|B_\beta|}}$ is defined as the sum of all ribbon matrices over ribbons with shape β :

$$M_G(\beta) = \sum_{R \text{ ribbon of shape } \beta} M_G(R). \tag{39}$$

► **Definition 22** (Automorphism of a Shape). For a shape β , $\text{Aut}(\beta)$ is the group of bijection from $V(\beta)$ to itself such that A_β and B_β are fixed as sets and the map is a graph automorphism of $(V(\beta), E(\beta))$.

It is easy to see that if we sum over ribbon matrices of all ribbons obtained from injective labelings of β , we obtain the graph matrix $M_G(\beta)$ multiplied by $|\text{Aut}(\beta)|$. Thus,

$$M_G(\beta) = \sum_{R \text{ ribbon of shape } \beta} M_G(R) = \frac{1}{|\text{Aut}(\beta)|} \sum_{f: V(\beta) \rightarrow V \text{ injective}} M_G(f(\beta)). \tag{40}$$

3.4 Graph Matrix Decomposition

► **Definition 23** (Legendre Symbol). Let \mathbb{F}_p be the finite field of order p . The Legendre symbol is defined as

$$\chi(a) = \chi_p(a) := \begin{cases} 0 & \text{if } a \equiv 0 \pmod{p}, \\ 1 & \text{if } a \text{ is a quadratic residue in } \mathbb{F}_p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue in } \mathbb{F}_p. \end{cases} \tag{41}$$

When the underlying finite field \mathbb{F}_p is fixed and clear from context, we will omit the subscript.

► **Remark 24**. Recall that all the primes p in our discussion are congruent to 1 modulo 4. This ensures that $\chi(-1) = 1$, and thus $\chi(a) = \chi(-a)$ for any $a \in \mathbb{F}_p$.

► **Proposition 25**. We have $\mathbb{1}_{\ell,r}(L, R) = \frac{1}{2^{|L \setminus R| \times |R \setminus L|}} \prod_{(a,b) \in (L \setminus R) \times (R \setminus L)} (1 + \chi(a - b))$ for all $\ell, r \geq 0$, $L \in \binom{\mathbb{F}_p}{\ell}$, and $R \in \binom{\mathbb{F}_p}{r}$.

Proof. The result follows from observing that, for $a, b \in \mathbb{F}_p$ distinct, $\frac{1}{2}(1 + \chi(a - b))$ is the indicator of the edge $\{a, b\}$ existing in the Paley graph. ◀

In the following few equations, let us write S for the Seidel adjacency matrix of G_p , so that $S_{a,b} := \chi(a - b)$. By substituting the indicator functions $\mathbb{1}_{\ell,r}$ in the definition of H using Proposition 25 and expanding the products, we have

$$H_{\{a,b\},\{c,d\}}^{2,2} = \begin{cases} \alpha_2 - \alpha_2^2 & \text{if } \{a,b\} = \{c,d\}, \\ \left(\frac{\alpha_3}{2} - \alpha_2^2\right) + \frac{\alpha_3}{2} S_{b,d} & \text{if } a = c \text{ and } b \neq d, \\ \left(\frac{\alpha_4}{16} - \alpha_2^2\right) + \frac{\alpha_4}{16} (S_{a,c} + S_{a,d} + S_{b,c} + S_{b,d} \\ \quad + S_{a,c}S_{a,d} + S_{b,c}S_{b,d} + S_{a,c}S_{b,c} \\ \quad + S_{a,d}S_{b,d} + S_{a,c}S_{b,d} + S_{a,d}S_{b,c} \\ \quad + S_{a,c}S_{a,d}S_{b,c} + S_{b,d}S_{a,d}S_{b,c} \\ \quad + S_{a,c}S_{a,d}S_{b,d} + S_{a,c}S_{b,c}S_{b,d} \\ \quad + S_{a,c}S_{a,d}S_{b,c}S_{b,d}) & \text{if } \{a,b\} \cap \{c,d\} = \emptyset. \end{cases} \quad (42)$$

and

$$\begin{aligned} & (H^{2,1}H^{1,2})_{\{a,b\},\{c,d\}} \\ &= \sum_{i \in \mathbb{F}_p} H_{\{a,b\},i}^{2,1} H_{i,\{c,d\}}^{1,2} \\ &= \begin{cases} 2(\alpha_2 - \alpha_1\alpha_2)^2 + (p-2)((\alpha_1\alpha_2)^2 + \frac{\alpha_3^2}{4} - \frac{\alpha_1\alpha_2\alpha_3}{2}) \\ \quad + \left(\frac{\alpha_3^2}{4} - \frac{\alpha_1\alpha_2\alpha_3}{2}\right) \sum_{i \in \mathbb{F}_p \setminus \{a,b\}} (S_{a,i} + S_{b,i} + S_{a,i}S_{b,i}) & \text{if } \{a,b\} = \{c,d\}, \\ \\ (\alpha_2 - \alpha_1\alpha_2)^2 - 2(\alpha_2 - \alpha_1\alpha_2)\alpha_1\alpha_2 + (p-3)(\alpha_1\alpha_2)^2 \\ \quad + \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{2} - (p-3)\frac{\alpha_1\alpha_2\alpha_3}{2} + (p-3)\frac{\alpha_3^2}{8} \\ \quad + \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{4} (S_{a,b}S_{b,d} + S_{a,d}S_{b,d} + S_{a,b} + S_{a,d} + 2S_{b,d}) \\ \quad + \left(\frac{\alpha_3^2}{8} - \frac{\alpha_1\alpha_2\alpha_3}{2}\right) \sum_{i \notin \{a,b,d\}} S_{a,i} \\ \quad + \left(\frac{\alpha_3^2}{8} - \frac{\alpha_1\alpha_2\alpha_3}{4}\right) \sum_{i \notin \{a,b,d\}} (S_{b,i} + S_{d,i} \\ \quad \quad + S_{a,i}S_{b,i} + S_{a,i}S_{d,i}) \\ \quad + \frac{\alpha_3^2}{8} \sum_{i \in \mathbb{F}_p \setminus \{a,b,d\}} (S_{b,i}S_{d,i} + S_{a,i}S_{b,i}S_{d,i}) & \text{if } a = c \text{ and } b \neq d, \\ \\ (\alpha_2 - \alpha_1\alpha_2)\alpha_3 - 4(\alpha_2 - \alpha_1\alpha_2)\alpha_1\alpha_2 \\ \quad + (p-4)(\alpha_1\alpha_2)^2 - (p-4)\frac{\alpha_1\alpha_2\alpha_3}{2} + (p-4)\frac{\alpha_3^2}{16} \\ \quad + \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{2} (S_{a,c} + S_{a,d} + S_{b,c} + S_{b,d}) \\ \quad + \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{4} (S_{a,c}S_{a,d} + S_{b,c}S_{b,d} + S_{a,c}S_{b,c} + S_{a,d}S_{b,d}) \\ \quad + \left(\frac{\alpha_3^2}{16} - \frac{\alpha_1\alpha_2\alpha_3}{4}\right) \sum_{i \notin \{a,b,c,d\}} (S_{a,i} + S_{b,i} + S_{c,i} + S_{d,i} \\ \quad \quad + S_{a,i}S_{b,i} + S_{c,i}S_{d,i}) \\ \quad + \frac{\alpha_3^2}{16} \sum_{i \notin \{a,b,c,d\}} (S_{a,i}S_{c,i} + S_{a,i}S_{d,i} + S_{b,i}S_{c,i} + S_{b,i}S_{d,i} \\ \quad \quad + S_{a,i}S_{b,i}S_{c,i} + S_{a,i}S_{b,i}S_{d,i} \\ \quad \quad + S_{a,i}S_{c,i}S_{d,i} + S_{b,i}S_{c,i}S_{d,i} \\ \quad \quad + S_{a,i}S_{b,i}S_{c,i}S_{d,i}) & \text{if } \{a,b\} \cap \{c,d\} = \emptyset. \end{cases} \quad (43) \end{aligned}$$

We now express this as a sum of graph matrices. We present all the matrices required for this decomposition in Table 1. Using the notations for graph matrices defined above and in the table, we can write the matrix $H^{2,2}$ and the matrix $H^{2,1}H^{1,2}$ as a weighted sum of these matrices, as follows:

$$\begin{aligned}
H^{2,2} &= (\alpha_2 - \alpha_2^2)I + \left(\frac{\alpha_3}{2} - \alpha_2^2\right)T^{3,0,1} + \frac{\alpha_3}{2}T^{3,1,1} + \left(\frac{\alpha_4}{16} - \alpha_2^2\right)T^{4,0,1} \\
&+ \frac{\alpha_4}{16}(T^{4,1,1} + T^{4,2,1} + T^{4,2,2} + T^{4,2,3} + T^{4,3,1} + T^{4,4,1}), \tag{44} \\
H^{2,1}H^{1,2} &= \left[2(\alpha_2 - \alpha_1\alpha_2)^2 + (p-2)\left((\alpha_1\alpha_2)^2 + \frac{\alpha_3^2}{4} - \frac{\alpha_1\alpha_2\alpha_3}{2}\right)\right]I \\
&+ \left(\frac{\alpha_3^2}{4} - \frac{\alpha_1\alpha_2\alpha_3}{2}\right)(U^{3,1,1} + U^{3,2,1}) \\
&+ \left[(\alpha_2 - \alpha_1\alpha_2)\left(\alpha_2 - 3\alpha_1\alpha_2 + \frac{\alpha_3}{2}\right) + (p-3)\left((\alpha_1\alpha_2)^2 - \frac{\alpha_1\alpha_2\alpha_3}{2} + \frac{\alpha_3^2}{8}\right)\right]T^{3,0,1} \\
&+ \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{4}(T^{3,2,1} + T^{3,2,2} + 2T^{3,1,1} + T^{3,1,2} + T^{3,1,3}) \\
&+ \left(\frac{\alpha_3^2}{8} - \frac{\alpha_1\alpha_2\alpha_3}{2}\right)U^{4,1,1} \\
&+ \left(\frac{\alpha_3^2}{8} - \frac{\alpha_1\alpha_2\alpha_3}{4}\right)(U^{4,1,2} + U^{4,1,3} + U^{4,2,1} + U^{4,2,2}) + \frac{\alpha_3^2}{8}(U^{4,2,3} + U^{4,3,1}) \\
&+ \left[(\alpha_2 - \alpha_1\alpha_2)(\alpha_3 - 4\alpha_1\alpha_2) + (p-4)\left(\alpha_1\alpha_2 - \frac{\alpha_3}{4}\right)^2\right]T^{4,0,1} \\
&+ \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{2}T^{4,1,1} + \frac{(\alpha_2 - \alpha_1\alpha_2)\alpha_3}{4}(T^{4,2,1} + T^{4,2,2}) \\
&+ \left(\frac{\alpha_3^2}{16} - \frac{\alpha_1\alpha_2\alpha_3}{4}\right)(U^{5,1,1} + U^{5,1,2} + U^{5,2,1} + U^{5,2,2}) \\
&+ \frac{\alpha_3^2}{16}(U^{5,2,3} + U^{5,3,1} + U^{5,3,2} + U^{5,4,1}). \tag{45}
\end{aligned}$$

3.5 Graph Matrix Norm Bounds

Now we analyze the norms of the graph matrices defined above in order to prove Proposition 15.

► **Remark 26.** Previous work of [1] established the typical norm of graph matrices when the underlying matrix G is the Seidel adjacency matrix of an ER random graph, where the quantities that characterize the norm bounds are the sizes of the minimum vertex separators of the shapes. In this work, using different techniques, we prove graph matrix norm bounds when the underlying matrix is the Seidel adjacency matrix of the Paley graph G_p .

Recall that we defined $P_0 = \frac{1}{p(p-1)}J \in \mathbb{R}^{\binom{\mathbb{F}_p}{2} \times \binom{\mathbb{F}_p}{2}}$ to denote the orthogonal projection matrix to the constant vector. Following the strategies in [9], we define the following subspaces of $\mathbb{R}^{\binom{\mathbb{F}_p}{2}}$:

$$\mathbb{V}_0 = \left\{ v \in \mathbb{R}^{\binom{\mathbb{F}_p}{2}} : v_{i,j} = v_{i',j'}, \quad \forall \{i,j\}, \{i',j'\} \in \binom{\mathbb{F}_p}{2} \right\} \tag{46}$$

$$\mathbb{V}_1 = \left\{ v \in \mathbb{R}^{\binom{\mathbb{F}_p}{2}} : \exists u \in \mathbb{R}^{\mathbb{F}_p}, \text{ s.t. } \langle \mathbf{1}, u \rangle = 0 \text{ and } v_{\{i,j\}} = u_i + u_j, \quad \forall \{i,j\} \in \binom{\mathbb{F}_p}{2} \right\} \tag{47}$$

$$\mathbb{V}_2 = (\mathbb{V}_0 \oplus \mathbb{V}_1)^\perp. \tag{48}$$

In words, \mathbb{V}_0 is the span of constant vectors, $\mathbb{V}_0 \oplus \mathbb{V}_1$ is the span of vectors v whose entries $v_{\{i,j\}}$ can be decomposed to a sum of $u_i + u_j$ for some $u \in \mathbb{R}^{\mathbb{F}_p}$, and \mathbb{V}_2 is the orthogonal

30:16 Degree 4 SOS Lower Bound for Clique Number of Paley Graph

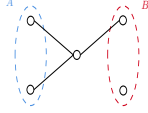
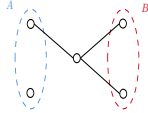
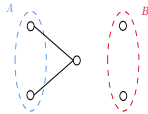
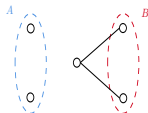
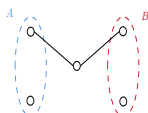
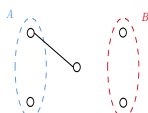
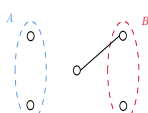
■ **Table 1** We present the graph matrices that we consider in Section 3.5 for the proof of Theorem 2, all defined on the Seidel adjacency matrix S of G_p . For each matrix, we give its name, the associated shape (see Definition 19), and the formula for the entries of the matrix. Some matrices are only non-zero on index sets satisfying certain equalities; in this case, for the sake of brevity, we indicate this “pattern” in the first column, and do not include the requisite indicator function in the third column. We also give the norm bound we prove in Section 3.5 and the norm bound for the same graph matrix evaluated on an ER random graph that follows from [1]. In these bounds we give only the order of growth; our bounds should be viewed as having an implicit $O(\cdot)$, and the bounds of [1] as having an implicit $\tilde{O}(\cdot)$.

Matrix	Shape	Entry Formula	G_p	$\mathcal{G}(p, \frac{1}{2})$
$T_{\{a,b\},\{a,c\}}^{3,2,1}$		$S_{a,b}S_{b,c}$	$p^{1/2}$	$p^{1/2}$
$T_{\{a,b\},\{a,c\}}^{3,2,2}$		$S_{a,c}S_{b,c}$	$p^{1/2}$	$p^{1/2}$
$T_{\{a,b\},\{a,c\}}^{3,1,1}$		$S_{b,c}$	$p^{1/2}$	$p^{1/2}$
$T_{\{a,b\},\{a,c\}}^{3,1,2}$		$S_{a,b}$	p	p
$T_{\{a,b\},\{a,c\}}^{3,1,3}$		$S_{a,c}$	p	p
$T_{\{a,b\},\{a,c\}}^{3,0,1}$		1	p	p
$T_{\{a,b\},\{c,d\}}^{4,4,1}$		$S_{a,c}S_{a,d}S_{b,c}S_{b,d}$	$p^{5/4}$	p

Matrix	Shape	Entry Formula	G_p	$\mathcal{G}(p, \frac{1}{2})$
$T_{\{a,b\},\{c,d\}}^{4,3,1}$		$S_{a,c}S_{a,d}S_{b,c} + S_{a,c}S_{a,d}S_{b,d} + S_{a,c}S_{b,c}S_{b,d} + S_{a,d}S_{b,c}S_{b,d}$	p	p
$T_{\{a,b\},\{c,d\}}^{4,2,1}$		$S_{a,c}S_{a,d} + S_{b,c}S_{b,d}$	$p^{3/2}$	$p^{3/2}$
$T_{\{a,b\},\{c,d\}}^{4,2,2}$		$S_{a,c}S_{b,c} + S_{a,d}S_{b,d}$	$p^{3/2}$	$p^{3/2}$
$T_{\{a,b\},\{c,d\}}^{4,2,3}$		$S_{a,c}S_{b,d} + S_{a,d}S_{b,c}$	p	p
$T_{\{a,b\},\{c,d\}}^{4,1,1}$		$S_{a,c} + S_{a,d} + S_{b,c} + S_{b,d}$	$p^{3/2}$	$p^{3/2}$
$T_{\{a,b\},\{c,d\}}^{4,0,1}$		1	p^2	p^2
$U_{\{a,b\},\{a,b\}}^{3,2,1}$		$\sum_{i \notin \{a,b\}} S_{a,i}S_{b,i}$	1	$p^{1/2}$
$U_{\{a,b\},\{a,b\}}^{3,1,1}$		$\sum_{i \notin \{a,b\}} S_{a,i} + S_{b,i}$	1	$p^{1/2}$

30:18 Degree 4 SOS Lower Bound for Clique Number of Paley Graph

Matrix	Shape	Entry Formula	G_p	$\mathcal{G}(p, \frac{1}{2})$
$U_{\{a,b\},\{a,c\}}^{4,3,1}$		$\sum_{i \notin \{a,b,c\}} S_{a,i} S_{b,i} S_{c,i}$	$p^{3/2}$	p
$U_{\{a,b\},\{a,c\}}^{4,2,1}$		$\sum_{i \notin \{a,b,c\}} S_{a,i} S_{b,i}$	p	$p^{3/2}$
$U_{\{a,b\},\{a,c\}}^{4,2,2}$		$\sum_{i \notin \{a,b,c\}} S_{a,i} S_{c,i}$	p	$p^{3/2}$
$U_{\{a,b\},\{a,c\}}^{4,2,3}$		$\sum_{i \notin \{a,b,c\}} S_{b,i} S_{c,i}$	p	p
$U_{\{a,b\},\{a,c\}}^{4,1,1}$		$\sum_{i \notin \{a,b,c\}} S_{a,i}$	p	$p^{3/2}$
$U_{\{a,b\},\{a,c\}}^{4,1,2}$		$\sum_{i \notin \{a,b,c\}} S_{b,i}$	p	$p^{3/2}$
$U_{\{a,b\},\{a,c\}}^{4,1,3}$		$\sum_{i \notin \{a,b,c\}} S_{c,i}$	p	$p^{3/2}$
$U_{\{a,b\},\{c,d\}}^{5,4,1}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} S_{b,i} S_{c,i} S_{d,i}$	p^2	p^2

Matrix	Shape	Entry Formula	G_p	$\mathcal{G}(p, \frac{1}{2})$
$U_{\{a,b\},\{c,d\}}^{5,3,1}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} S_{b,i} S_{c,i} + S_{a,i} S_{b,i} S_{d,i}$	p^2	p^2
$U_{\{a,b\},\{c,d\}}^{5,3,2}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} S_{c,i} S_{d,i} + S_{b,i} S_{c,i} S_{d,i}$	p^2	p^2
$U_{\{a,b\},\{c,d\}}^{5,2,1}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} S_{b,i}$	p^2	$p^{5/2}$
$U_{\{a,b\},\{c,d\}}^{5,2,2}$		$\sum_{i \notin \{a,b,c,d\}} S_{c,i} S_{d,i}$	p^2	$p^{5/2}$
$U_{\{a,b\},\{c,d\}}^{5,2,3}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} S_{c,i} + S_{a,i} S_{d,i} + S_{b,i} S_{c,i} + S_{b,i} S_{d,i}$	p^2	p^2
$U_{\{a,b\},\{c,d\}}^{5,1,1}$		$\sum_{i \notin \{a,b,c,d\}} S_{a,i} + S_{b,i}$	p^2	$p^{5/2}$
$U_{\{a,b\},\{c,d\}}^{5,1,2}$		$\sum_{i \notin \{a,b,c,d\}} S_{c,i} + S_{d,i}$	p^2	$p^{5/2}$

30:20 Degree 4 SOS Lower Bound for Clique Number of Paley Graph

complement of $\mathbb{V}_0 \oplus \mathbb{V}_1$. Furthermore, let P_1 and P_2 be the orthogonal projection matrices to the subspaces \mathbb{V}_1 , and \mathbb{V}_2 respectively. Note that this is consistent with the previously defined P_0 , which is the orthogonal projection matrix to the span of constant vectors \mathbb{V}_0 .

In the analysis of ER graphs, these subspaces appear because they are the decomposition of $\mathbb{R}^{\binom{p}{2}}$ into irreducible subrepresentations under the action of S_p , with respect to which the expectation of an FK pseudomoment matrix is invariant. This invariance does not hold for our deterministic FK pseudomoment matrix, but we will see that the same decomposition is still useful.

We will use the following norm bounds for the graph matrices defined earlier. The proofs may be found in the full version of the paper.

- **Proposition 27.** $\|T^{3,2,i}\| = O(\sqrt{p})$ for $i \in \{1, 2\}$.
- **Proposition 28.** $\|T^{3,1,1}\| = O(\sqrt{p})$.
- **Proposition 29.** $\|T^{3,1,i}\| = O(p)$ for $i \in \{2, 3\}$.
- **Proposition 30.** $T^{3,0,1} = 2(p-2)P_0 + (p-4)P_1 - 2P_2$.
- **Proposition 31.** $\|T^{4,3,1}\| = O(p)$.
- **Proposition 32.** $\|T^{4,i,j}\| = O(p^{3/2})$ for $(i, j) \in \{(2, 1), (2, 2), (1, 1)\}$. Moreover, all of $\|T^{4,2,1}P_2\|$, $\|P_2T^{4,2,2}\|$, $\|P_2T^{4,1,1}\|$, and $\|T^{4,1,1}P_2\|$ are $O(\sqrt{p})$.
- **Proposition 33.** $\|T^{4,2,3}\| = O(p)$.
- **Proposition 34.** $T^{4,0,1} = \frac{(p-2)(p-3)}{2}P_0 - (p-3)P_1 + P_2$.
- **Proposition 35.** $\|U^{3,i,1}\| = O(1)$ for $i \in \{1, 2\}$.
- **Proposition 36.** $\|U^{4,3,1}\| = O(p^{3/2})$.
- **Proposition 37.** $\|U^{4,i,j}\| = O(p)$ for $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$.
- **Proposition 38.** $\|U^{5,4,1}\| = O(p^2)$.
- **Proposition 39.** $\|U^{5,3,i}\| = O(p^2)$ for $i \in \{1, 2\}$.
- **Proposition 40.** $\|U^{5,i,j}\| = O(p^2)$ for $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$, where $j \neq 3$ if $i = 1$.
- **Theorem 41.** $\|T^{4,4,1}\| = O(p^{5/4})$.

Of these statements, Theorem 41 is by far the subtlest – unlike the other terms, where fairly straightforward arguments work, for $T^{4,4,1}$ it turns out that a naive bound is insufficient, and we must more carefully account for character sum cancellations. The bounds we prove are generally incomparable to those for random graphs following from [1]: for some graph matrices we expect a comparable norm bound but cannot prove one due to technical obstacles, while for other graph matrices the Paley graph exhibits stronger cancellations than a random graph and we can show a stronger norm bound. We compare the respective bounds in Table 1. Moreover, as we show in Section 4.3, there is an example of a graph matrix for which the norm when evaluated on the Paley graph is actually asymptotically larger than the norm when evaluated on a random graph; however, this example does not figure in our analysis.

3.6 Final Steps

Finally, putting all the graph matrix norm bounds together, we prove Proposition 15, which will conclude the proof of Theorem 9, as we have discussed earlier.

Proof of Proposition 15. The statements in this proof will hold for all sufficiently large primes p .

To show $H^{2,2} \succeq (\alpha_1 + \frac{p-1}{2}\alpha_2 - p\alpha_1^2)^{-1}P_0H^{2,1}H^{1,2}P_0 + ((1-\varepsilon)\alpha_1)^{-1}(I-P_0)H^{2,1}H^{1,2}(I-P_0)$, we have to show that $M_1 \succeq M_2$, where M_1 is the sum of all multiples of the identity, $T^{3,0,1}$, and $T^{4,0,1}$ (possibly conjugated by P_0 or $I-P_0$) appearing in the expressions (44) and (45), and M_2 is the sum of the remaining graph matrices of shapes having at least one edge. Note that $\mathbb{V}_0, \mathbb{V}_1, \mathbb{V}_2$ are eigenspaces of M_1 , with eigenvalues scaling as $(1-o(1))6p^2\alpha_1^4$, $(1-o(1))4p\alpha_1^3$, and $(1-o(1))4\alpha_1^2$, respectively.

It is then sufficient to show

$$\begin{bmatrix} 3p^2\alpha_1^4 & 0 & 0 \\ 0 & 2p\alpha_1^3 & 0 \\ 0 & 0 & 2\alpha_1^2 \end{bmatrix} \succeq \begin{bmatrix} \|P_0M_2P_0\| & \|P_0M_2P_1\| & \|P_0M_2P_2\| \\ \|P_1M_2P_0\| & \|P_1M_2P_1\| & \|P_1M_2P_2\| \\ \|P_2M_2P_0\| & \|P_2M_2P_1\| & \|P_2M_2P_2\| \end{bmatrix}. \quad (49)$$

Using the graph matrix norm bounds above, we have for any $i \in \{0, 1, 2\}$ and $j \in \{0, 1, 2\}$ with $(i, j) \neq (2, 2)$ that

$$\|P_iM_2P_j\| = O(p^{3/2}\alpha_1^4), \quad (50)$$

and for the remaining case

$$\|P_2M_2P_2\| = O(p^2\alpha_1^5), \quad (51)$$

so we only need to prove that the following matrix is positive semidefinite:

$$\begin{bmatrix} 3p^2\alpha_1^4 - O(p^{3/2}\alpha_1^4) & -O(p^{3/2}\alpha_1^4) & -O(p^{3/2}\alpha_1^4) \\ -O(p^{3/2}\alpha_1^4) & 2p\alpha_1^3 - O(p^{3/2}\alpha_1^4) & -O(p^{3/2}\alpha_1^4) \\ -O(p^{3/2}\alpha_1^4) & -O(p^{3/2}\alpha_1^4) & 2\alpha_1^2 - O(p^2\alpha_1^5) \end{bmatrix}, \quad (52)$$

which is verified by taking the Schur complement and using diagonal dominance when $\alpha_1 = c \cdot p^{-2/3}$ for a sufficiently small constant c . ◀

With Proposition 15 proved, we have finished proving Theorem 9.

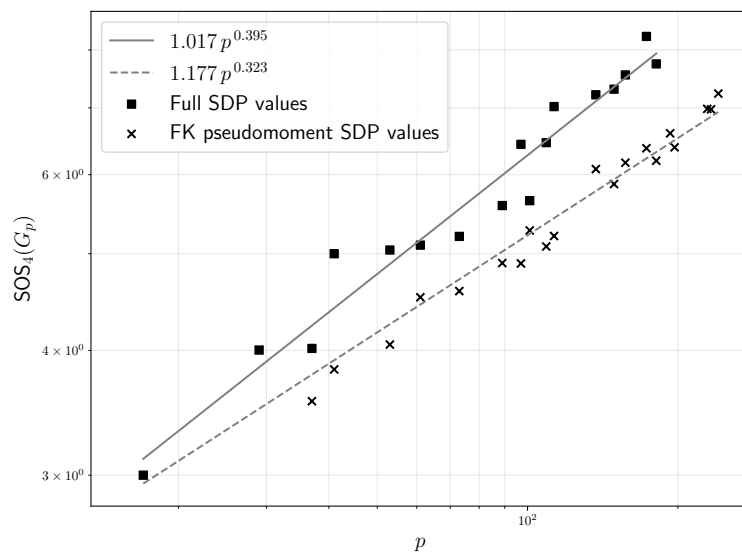
4 Ancillary Results

4.1 Optimality Over Feige-Krauthgamer Pseudomoments

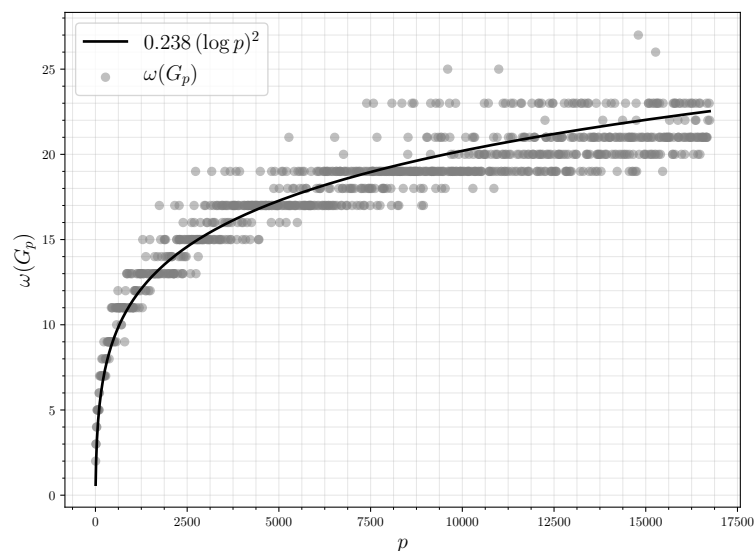
In this section, we show that our lower bound is optimal over those achievable by FK pseudomoments. To be precise, let us define a new SDP corresponding to this restricted type of pseudomoment, a variant of (13):

$$\text{FK}_4(G) := \left\{ \begin{array}{l} \text{maximize} \quad \sum_{i=1}^n M_{\emptyset,i}^{0,1} \\ \text{subject to} \quad M^{r,c} \in \mathbb{R}^{\binom{[n]}{r} \times \binom{[n]}{c}} \text{ for } r, c \in \{0, 1, 2\}, \\ \quad M_{S,T}^{r,c} \text{ depends only on } S \cup T, \\ \quad M_{S,T}^{r,c} = 0 \text{ whenever } S \cup T \notin \mathcal{K}(G), \\ \quad M_{S,T}^{r,c} \text{ depends only on } |S \cup T| \text{ when } S \cup T \in \mathcal{K}(G), \\ \quad M = \begin{bmatrix} 1 & M^{0,1} & M^{0,2} \\ M^{1,0} & M^{1,1} & M^{1,2} \\ M^{2,0} & M^{2,1} & M^{2,2} \end{bmatrix} \succeq 0 \end{array} \right\}. \quad (53)$$

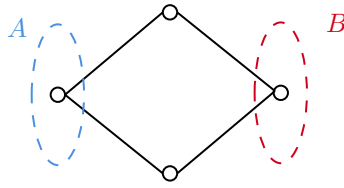
30:22 Degree 4 SOS Lower Bound for Clique Number of Paley Graph



■ **Figure 1** For primes $5 \leq p \leq 250$, we present the value of $\text{SOS}_4(G_p)$ and the value of $\text{FK}_4(G_p)$ (where the semidefinite program is restricted to optimize over only FK pseudomoments). We fit power models ap^b to the data and plot the results as well.



■ **Figure 2** For primes $5 \leq p \leq 16741$, we present computations of the true clique number $\omega(G_p)$ (taken from [46] and its online supplementary materials). We fit a model $a(\log p)^2$ to the data and plot the results as well.



■ **Figure 3** We illustrate the graph matrix used as an example in Section 4.3.

Since the conditions of this SDP are more restrictive than those of $\text{SOS}_4(G)$, we have $\text{SOS}_4(G) \geq \text{FK}_4(G)$. Our strategy has been to show that $\text{FK}_4(G)$ is large; the following shows a limitation to this approach. The proof is given in the full version of the paper.

► **Theorem 42.** *Over primes $p \equiv 1 \pmod{4}$, $\text{FK}_4(G_p) = \Theta(p^{1/3})$.*

4.2 Numerical Experiments

Given our results in Theorems 2 and 42, it is natural to ask whether a better lower bound technique than working with FK pseudomoments might prove an optimal lower bound of the form $\text{SOS}_4(G_p) = \Omega(p^{1/2})$. In Figure 1, we present some surprising numerical results suggesting that this is *not* the case. Namely, in addition to the true values of $\omega(G_p)$, we plot the values of $\text{SOS}_4(G_p)$ (the “full SDP”) and of $\text{FK}_4(G_p)$ (the “FK pseudomoment SDP”) on a log-log plot, and fit lines to these results.⁸

These results for $\text{FK}_4(G_p)$ confirm the statement of Theorem 42, with an estimated scaling of $\text{FK}_4(G_p) \sim p^{0.323}$, close to our result showing that $\text{FK}_4(G_p) \sim p^{1/3}$. For $\text{SOS}_4(G_p)$, the results still indicate a scaling below $p^{1/2}$, estimated at $\text{SOS}_4(G_p) \sim p^{0.395}$. Based on these results, it seems reasonable to conjecture that $\text{SOS}_4(G_p) = O(p^{1/2-\epsilon})$ for some $\epsilon > 0$. This prediction is compatible with that of [33], who, based experiments solving a weaker SDP than degree 4 SOS as proposed by [21], experimentally found that $\text{SOS}_4(G_p) \lesssim p^{0.456}$.

4.3 General Graph Matrix Norm Bounds Do Not Derandomize

In this section, we give a simple example of a graph matrix for which the norm bound of [1] for ER graphs fails to hold for Paley graphs. Since the bound of [1] is a crucial ingredient in the proof of the $\Omega(p^{1/2})$ SOS lower bound of [4], we take this as some evidence that a sufficiently high degree of SOS can prove a bound of the form $\omega(G_p) \leq O(p^{1/2-\epsilon})$. In particular, this gives theoretical evidence for the numerical observations above.

Let $S \in \mathbb{R}^{n \times n}$ be the Seidel adjacency matrix of a graph. We consider the graph matrix $M = M(S)$ formed from S and the shape in Figure 3, with entries $M_{xy} = \mathbb{1}\{x \neq y\} \sum_{\substack{a,b \in [n] \\ a \neq b}} S_{a,x} S_{a,y} S_{b,x} S_{b,y}$, where we do not need to include the constraints $a, b \notin \{x, y\}$ since these are automatically enacted by having $S_{a,a} = 0$ for all a .

For any such S and $x \neq y$, we have $M_{xy} = (S^2)_{x,y}^2 - (p-2)$. When S is the Seidel adjacency matrix of the Paley graph, we have $S^2 = pI - \mathbf{1}\mathbf{1}^\top$. Thus in this case we have $M(S) = (p-3)I - (p-3)\mathbf{1}\mathbf{1}^\top$, and $\|M\| = (p-1)(p-3) \sim p^2$. On the other hand, when S is the Seidel adjacency matrix of a random ER graph, then the results of [1] show that, since the shape of M has minimum vertex separator of size 1, with high probability, $\|M\| \leq \tilde{O}(p^{3/2})$. Thus, the Paley graph adjacency matrix fails to satisfy this basic graph matrix bound.

⁸ These SDPs are solved using the Mosek solver through the CVXPY interface for Python.

References

- 1 Kwangjun Ahn, Dhruv Medarametla, and Aaron Potechin. Graph matrices: Norm bounds and applications. *arXiv preprint*, 2016. [arXiv:1604.03423](#).
- 2 Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466, 1998.
- 3 Christine Bachoc, Máté Matolcsi, and Imre Z Ruzsa. Squares and difference sets in finite fields. *Integers*, 13:A77, 2013.
- 4 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.
- 5 Béla Bollobás. *Random graphs*. Cambridge University Press, second edition, 2001.
- 6 I Broere, D Döman, and JN Ridley. The clique numbers and chromatic numbers of certain Paley graphs. *Quaestiones Mathematicae*, 11(1):91–93, 1988.
- 7 Fan R. K. Chung, Ronald L. Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.
- 8 Ernie Croot and Vsevolod F Lev. Open problems in additive combinatorics. *Additive Combinatorics*, 43:207–233, 2007.
- 9 Yash Deshpande and Andrea Montanari. Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems. In *28th Annual Conference on Learning Theory (COLT 2015)*, pages 523–562, 2015.
- 10 Daniel Di Benedetto, József Solymosi, and Ethan P White. On the directions determined by a Cartesian product in an affine Galois plane. *Combinatorica*, 41(6):755–763, 2021.
- 11 Irit Dinur, Yuval Filmus, Prahladh Harsha, and Madhur Tulsiani. Explicit SoS lower bounds from high-dimensional expanders. *arXiv preprint*, 2020. [arXiv:2009.05218](#).
- 12 Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.
- 13 RJ Evans, JR Pulham, and J Sheehan. On the number of complete subgraphs contained in certain graphs. *Journal of Combinatorial Theory, Series B*, 30(3):364–371, 1981.
- 14 Uriel Feige and Robert Krauthgamer. Finding and certifying a large hidden clique in a semirandom graph. *Random Structures & Algorithms*, 16(2):195–208, 2000.
- 15 Uriel Feige and Robert Krauthgamer. The probable value of the Lovász-Schrijver relaxations for maximum independent set. *SIAM Journal on Computing*, 32(2):345–370, 2003.
- 16 Laura Galli and Adam N Letchford. On the Lovász theta function and some variants. *Discrete Optimization*, 25:159–174, 2017.
- 17 David Gamarnik and Ilias Zadik. The landscape of the planted clique problem: Dense subgraphs and the overlap gap property. *arXiv preprint*, 2019. [arXiv:1904.07174](#).
- 18 Sidney West Graham and CJ Ringrose. Lower bounds for least quadratic non-residues. In *Analytic number theory*, pages 269–309. Springer, 1990.
- 19 Dima Grigoriev. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1-2):613–622, 2001.
- 20 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.
- 21 Nebojša Gvozdenović, Monique Laurent, and Frank Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. *Operations Research Letters*, 37(1):27–31, 2009.
- 22 Willem H Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its Applications*, 226:593–616, 1995.
- 23 Brandon Hanson and Giorgis Petridis. Refined estimates concerning sumsets contained in the roots of unity. *Proceedings of the London Mathematical Society*, 122(3):353–358, 2021.
- 24 Johan Hastad. Clique is hard to approximate within $n^{1-\epsilon}$. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 627–636. IEEE, 1996.

- 25 Alan J Hoffman. On eigenvalues and colorings of graphs. In Bernard Harris, editor, *Graph Theory and its Applications*. Academic Press, 1970.
- 26 Max Hopkins and Ting-Chun Lin. Explicit lower bounds against $\omega(n)$ -rounds of sum-of-squares. *arXiv preprint*, 2022. [arXiv:2204.11469](#).
- 27 Samuel B Hopkins, Pravesh K Kothari, and Aaron Potechin. SOS and planted clique: Tight analysis of MPW moments at all degrees and an optimal lower bound at degree four. *arXiv preprint*, 2015. [arXiv:1507.05230](#).
- 28 Mark Jerrum. Large cliques elude the Metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992.
- 29 Chris Jones, Aaron Potechin, Goutham Rajendran, Madhur Tulsiani, and Jeff Xu. Sum-of-squares lower bounds for sparse independent set. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 406–416. IEEE, 2022.
- 30 Ferenc Juhász. The asymptotic behaviour of Lovász’ theta function for random graphs. *Combinatorica*, 2(2):153–155, 1982.
- 31 Richard M Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer, 1972.
- 32 Richard M. Karp. The probabilistic analysis of some combinatorial search algorithms. In *Algorithms and complexity: New Directions and Recent Results*, 1976.
- 33 Vladimir A Kobzar and Krishnan Mody. Revisiting block-diagonal sdp relaxations for the clique number of the paley graphs. *arXiv preprint*, 2023. [arXiv:2304.08615](#).
- 34 Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995.
- 35 Monique Laurent. Lower bound for the number of iterations in semidefinite hierarchies for the cut polytope. *Mathematics of Operations Research*, 28(4):871–883, 2003.
- 36 Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, pages 157–270. Springer, 2009.
- 37 László Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information theory*, 25(1):1–7, 1979.
- 38 Mark Magsino, Dustin G Mixon, and Hans Parshall. Linear programming bounds for cliques in Paley graphs. In *Wavelets and Sparsity XVIII*, volume 11138, page 111381H. International Society for Optics and Photonics, 2019.
- 39 Raghu Meka, Aaron Potechin, and Avi Wigderson. Sum-of-squares lower bounds for planted clique. In *47th Annual ACM Symposium on Theory of Computing (STOC 2015)*, pages 87–96. ACM, 2015.
- 40 Raghu Meka and Avi Wigderson. Association schemes, non-commutative polynomial concentration, and sum-of-squares lower bounds for planted clique. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 20, page 10, 2013.
- 41 Hugh L Montgomery. *Topics in multiplicative number theory*, volume 227. Springer, 1971.
- 42 Rudi Mrazović. A random model for the Paley graph. *The Quarterly Journal of Mathematics*, 68(1):193–206, 2017.
- 43 Shuo Pang. SOS lower bound for exact planted clique. In *36th Computational Complexity Conference (CCC 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.
- 44 Stanislaw Radziszowski. Small Ramsey numbers. *The Electronic Journal of Combinatorics*, 1000:DS1–Aug, 2011.
- 45 Prasad Raghavendra and Tselil Schramm. Tight lower bounds for planted clique in the degree-4 SOS program. *arXiv preprint*, 2015. [arXiv:1507.05136](#).
- 46 James B Shearer. Lower bounds for small diagonal Ramsey numbers. *Journal of Combinatorial Theory, Series A*, 42(2):302–304, 1986.
- 47 Chi Hoi Yip. On the clique number of Paley graphs of prime power order. *Finite Fields and Their Applications*, 77:101930, 2022.