# The degree 4 sum-of-squares relaxation of the clique number of Paley graphs


Tim Kunisky
(with Xifan Yu)

Yale University


AMS Special Session—October 23, 2022

# Paley Graph

$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ = finite field on $p$ elements for $p \equiv 1 \bmod 4$.

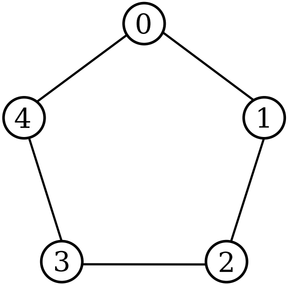
$G_p$ a graph on vertices $\mathbb{F}_p$ with $i \sim j$ iff $j - i$ is a **square** mod $p$ (for some $x \neq 0$, $j - i = x^2$).

# Paley Graph

$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ = finite field on $p$ elements for $p \equiv 1 \bmod 4$.

$G_p$ a graph on vertices $\mathbb{F}_p$ with $i \sim j$ iff $j - i$ is a **square** mod $p$ (for some $x \neq 0$, $j - i = x^2$).

**Example:** $p = 5 \rightsquigarrow$ squares are $\{1, 4 \equiv -1\}$.



# Paley Graph

$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ = finite field on $p$ elements for $p \equiv 1 \bmod 4$.

$G_p$ a graph on vertices $\mathbb{F}_p$ with $i \sim j$ iff $j - i$ is a **square** mod $p$ (for some $x \neq 0$, $j - i = x^2$).

$\implies \mathsf{deg}(i) = \frac{p-1}{2} \sim \frac{1}{2}p$ for each $i \in \mathbb{F}_p$.

# Paley Graph

$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ = finite field on $p$ elements for $p \equiv 1 \bmod 4$.

$G_p$ a graph on vertices $\mathbb{F}_p$ with $i \sim j$ iff $j - i$ is a **square** mod $p$ (for some $x \neq 0$, $j - i = x^2$).

$\implies \mathsf{deg}(i) = \frac{p-1}{2} \sim \frac{1}{2}p$ for each $i \in \mathbb{F}_p$.

**Heuristic:** $G_p$ is **pseudorandom**, behaving in many ways like $\mathsf{ER}(p, \frac{1}{2})$, i.i.d. random graph with edge probability $\frac{1}{2}$.

**Example:** For any fixed graph $H$, as $p \to \infty$,

$$\begin{aligned} \text{occurrences in } G_p &\sim \mathbb{E}\left[\text{occurrences in } \mathsf{ER}\right] \\ &\sim n^{|V(H)|}\left(\frac{1}{2}\right)^{\binom{|V(H)|}{2}} \end{aligned}$$

# Paley Graphs: The Clique Number

**Question:** What about $H$ growing slowly with $p$?

**Example:** $\omega(G)$ := **largest clique** in $G$. Easy calculations $\Longrightarrow$

$$\mathbb{E}[\omega(\mathsf{ER})] \sim 2\log_2 p.$$

# Paley Graphs: The Clique Number

**Question:** What about $H$ growing slowly with $p$?

**Example:** $\omega(G)$ := **largest clique** in $G$. Easy calculations $\Longrightarrow$

$$\mathbb{E}[\omega(\mathrm{ER})] \sim 2\log_2 p.$$

Same for $\omega(G_p)$? Not quite...

$$\omega(G_{p_i}) \geq \log p_i \log\log\log p_i \qquad \text{[Graham, Ringrose '90]}$$

$$\omega(G_p) \overset{?}{\sim} (\log p)^2 \qquad \text{(random heuristic)}$$

And, in any case, the best **upper bounds** we have are

$$\omega(G_p) \leq \sqrt{p} \qquad \text{(spectral/Hoffman/trivial bound)}$$

$$\omega(G_p) \leq \sqrt{p/2} + 1 \qquad \text{[Hanson, Petridis '21]}$$

**Big Question 1: How can we break the "square root barrier" and prove**

$$\omega(G_p) = O(p^{1/2-\varepsilon}) \ ?$$

(Formally similar to controlling the restricted isometry property for the Paley ETF.)

# Sum-of-Squares (SOS) Relaxations

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max \left\{ \sum_{i \in V} y_i \ : \ \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i,j\} \notin E \right\}$$

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max\left\{ \sum_{i \in V} y_i \;:\; \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i,j\} \notin E \right\}$$

Semidefinite programming upper bound recipe:

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max\left\{\sum_{i \in V} y_i \ : \ \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i, j\} \notin E\right\}$$

Semidefinite programming upper bound recipe:

1. Write $\boldsymbol{y}^{\otimes \leq d} = [1 \ \ \boldsymbol{y} \ \ \boldsymbol{y}^{\otimes 2} \ \cdots \ \boldsymbol{y}^{\otimes d}]$ and $\boldsymbol{X} = \boldsymbol{y}^{\otimes \leq d} \boldsymbol{y}^{\otimes \leq d^\top}$.

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max\left\{\sum_{i \in V} y_i \;:\; \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i,j\} \notin E\right\}$$

Semidefinite programming upper bound recipe:

1. Write $\boldsymbol{y}^{\otimes \leq d} = [1 \;\; \boldsymbol{y} \;\; \boldsymbol{y}^{\otimes 2} \;\; \cdots \;\; \boldsymbol{y}^{\otimes d}]$ and $\boldsymbol{X} = \boldsymbol{y}^{\otimes \leq d} \boldsymbol{y}^{\otimes \leq d^\top}$.
2. Find some tractable constraints on $\boldsymbol{X}$ for feasible $\boldsymbol{y}$:

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max \left\{ \sum_{i \in V} y_i \;:\; \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i,j\} \notin E \right\}$$

Semidefinite programming upper bound recipe:

1. Write $\boldsymbol{y}^{\otimes \leq d} = [1 \;\; \boldsymbol{y} \;\; \boldsymbol{y}^{\otimes 2} \;\; \cdots \;\; \boldsymbol{y}^{\otimes d}]$ and $\boldsymbol{X} = \boldsymbol{y}^{\otimes \leq d} \boldsymbol{y}^{\otimes \leq d^\top}$.
2. Find some tractable constraints on $\boldsymbol{X}$ for feasible $\boldsymbol{y}$:
    - $\boldsymbol{X} \succeq \boldsymbol{0}$
    - $X_{\boldsymbol{i},\boldsymbol{j}} = X(S)$ depends only on index set $S$ in $\boldsymbol{i}, \boldsymbol{j}$
    - $X(\emptyset) = 1$, $X(S) = 0$ for all $S$ not a clique in $G$

# Sum-of-Squares (SOS) Relaxations

For any graph $G = (V, E)$, have Boolean optimization formulation,

$$\omega(G) = \max \left\{ \sum_{i \in V} y_i \;:\; \boldsymbol{y} \in \{0,1\}^V, \quad y_i y_j = 0 \text{ if } \{i,j\} \notin E \right\}$$

Semidefinite programming upper bound recipe:

1. Write $\boldsymbol{y}^{\otimes \leq d} = [1 \;\; \boldsymbol{y} \;\; \boldsymbol{y}^{\otimes 2} \;\; \cdots \;\; \boldsymbol{y}^{\otimes d}]$ and $\boldsymbol{X} = \boldsymbol{y}^{\otimes \leq d} \boldsymbol{y}^{\otimes \leq d^\top}$.
2. Find some tractable constraints on $\boldsymbol{X}$ for feasible $\boldsymbol{y}$:
    - $\boldsymbol{X} \succeq \boldsymbol{0}$
    - $X_{\boldsymbol{i},\boldsymbol{j}} = X(S)$ depends only on index set $S$ in $\boldsymbol{i}, \boldsymbol{j}$
    - $X(\emptyset) = 1$, $X(S) = 0$ for all $S$ not a clique in $G$
3. Optimize $\sum_{i \in V} X(\{i\})$ over that enlarged set.

# Degree 2 $=: \mathsf{SOS}_2(G)$ (Case $d = 1$)

$$\text{maximize} \quad \sum_{i=1}^{p} X(\{i\}) \quad \text{subject to}$$

$$\boldsymbol{X} = \left[\begin{array}{c|cccc} 1 & X(\{1\}) & X(\{2\}) & \cdots & X(\{p\}) \\ \hline X(\{1\}) & X(\{1\}) & X(\{1,2\}) & \cdots & X(\{1,p\}) \\ X(\{2\}) & X(\{1,2\}) & X(\{2\}) & \cdots & X(\{2,p\}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X(\{p\}) & X(\{1,p\}) & X(\{2,p\}) & \cdots & X(\{p\}) \end{array}\right] \succeq \mathbf{0},$$

$X(\{i,j\}) = 0$ whenever $i \not\sim_G j$.

# Degree 2 $=: \mathsf{SOS}_2(G)$ (Case $d = 1$)

$$\text{maximize} \quad \sum_{i=1}^{p} X(\{i\}) \quad \text{subject to}$$

$$\boldsymbol{X} = \left[\begin{array}{c|cccc} 1 & X(\{1\}) & X(\{2\}) & \cdots & X(\{p\}) \\ \hline X(\{1\}) & X(\{1\}) & X(\{1,2\}) & \cdots & X(\{1,p\}) \\ X(\{2\}) & X(\{1,2\}) & X(\{2\}) & \cdots & X(\{2,p\}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X(\{p\}) & X(\{1,p\}) & X(\{2,p\}) & \cdots & X(\{p\}) \end{array}\right] \succeq \mathbf{0},$$

$X(\{i,j\}) = 0$ whenever $i \not\sim_G j$.

This has been studied earlier as the **Lovász function** $\vartheta(\overline{G})$.

# Degree 2 =: $\mathsf{SOS}_2(G)$ (Case $d = 1$)

maximize $\sum_{i=1}^{p} X(\{i\})$ subject to

$$\boldsymbol{X} = \left[\begin{array}{c|cccc} 1 & X(\{1\}) & X(\{2\}) & \cdots & X(\{p\}) \\ \hline X(\{1\}) & X(\{1\}) & X(\{1,2\}) & \cdots & X(\{1,p\}) \\ X(\{2\}) & X(\{1,2\}) & X(\{2\}) & \cdots & X(\{2,p\}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X(\{p\}) & X(\{1,p\}) & X(\{2,p\}) & \cdots & X(\{p\}) \end{array}\right] \succeq \mathbf{0},$$

$X(\{i,j\}) = 0$ whenever $i \not\sim_G j$.

This has been studied earlier as the **Lovász function** $\vartheta(\overline{G})$.

$d \geq 2 \rightsquigarrow \mathsf{SOS}_{2d}(G) \geq \omega(G)$, tighter bounds in time $p^{O(d)}$.

# Lower Bounds for Random Graphs

To study average-case difficulty of $\omega(\cdot)$, people wanted to understand how hard it is to compute $\omega(\mathsf{ER}(p, \frac{1}{2}))$.

# Lower Bounds for Random Graphs

To study average-case difficulty of $\omega(\cdot)$, people wanted to understand how hard it is to compute $\omega(\mathsf{ER}(p, \frac{1}{2}))$.

**Theorem:** [MW '13]...[BHKKP '19] For any fixed $d$, as $p \to \infty$,

$$\mathbb{E}\,[\mathsf{SOS}_{2d}(\mathsf{ER})] = \Omega(p^{1/2}) \gg O(\log p) = \mathbb{E}[\omega(\mathsf{ER})].$$

# Lower Bounds for Random Graphs

To study average-case difficulty of $\omega(\cdot)$, people wanted to understand how hard it is to compute $\omega(\mathsf{ER}(p, \frac{1}{2}))$.

**Theorem:** [MW '13]...[BHKKP '19] For any fixed $d$, as $p \to \infty$,

$$\mathbb{E}[\mathsf{SOS}_{2d}(\mathsf{ER})] = \Omega(p^{1/2}) \gg O(\log p) = \mathbb{E}[\omega(\mathsf{ER})].$$

SOS is a strong family of algorithms, so can view this as one **specific** demonstration of **average-case hardness** of $\omega(\cdot)$.

# Lower Bounds for Random Graphs

To study average-case difficulty of $\omega(\cdot)$, people wanted to understand how hard it is to compute $\omega(\mathsf{ER}(p, \frac{1}{2}))$.

**Theorem:** [MW '13]...[BHKKP '19] For any fixed $d$, as $p \to \infty$,

$$\mathbb{E}[\mathsf{SOS}_{2d}(\mathsf{ER})] = \Omega(p^{1/2}) \gg O(\log p) = \mathbb{E}[\omega(\mathsf{ER})].$$

SOS is a strong family of algorithms, so can view this as one **specific** demonstration of **average-case hardness** of $\omega(\cdot)$.

**Question:** How important is the distribution of $\mathsf{ER}(p, \frac{1}{2})$? What properties of a graph does this really depend on?

**Big Question 2: How can we find deterministic graphs $H_p$ with**

$$\omega(H_p) = O(\log p)$$
$$\mathsf{SOS}_{2d}(H_p) = \Omega(p^{1/2}) \text{ ?}$$

# Our Results

**Main message:** Paley graphs achieve a partial derandomization of SOS lower bounds for $\mathsf{ER}(p, \frac{1}{2})$.

# Our Results

**Main message:** Paley graphs achieve a partial derandomization of SOS lower bounds for $\mathsf{ER}(p, \frac{1}{2})$.

**Easy to show:** $\mathsf{SOS}_2(G_p) = \Omega(p^{1/2})$.

# Our Results

**Main message:** Paley graphs achieve a partial derandomization of SOS lower bounds for $\mathsf{ER}(p, \frac{1}{2})$.

**Easy to show:** $\mathsf{SOS}_2(G_p) = \Omega(p^{1/2})$.

**Main theorem:** [KY] $\mathsf{SOS}_4(G_p) = \Omega(p^{1/3})$.

# Our Results

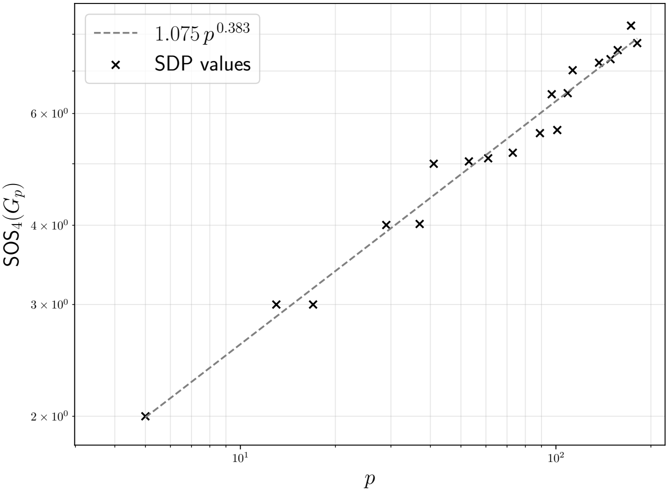
**Main message:** Paley graphs achieve a partial derandomization of SOS lower bounds for $\mathsf{ER}(p, \frac{1}{2})$.

**Easy to show:** $\mathsf{SOS}_2(G_p) = \Omega(p^{1/2})$.

**Main theorem:** [KY] $\mathsf{SOS}_4(G_p) = \Omega(p^{1/3})$.

**Remark:** Derandomizes an early result on the random graph case: [DM '15] showed $\mathbb{E}[\mathsf{SOS}_4(\mathsf{ER})] = \widetilde{\Omega}(p^{1/3})$.

# Ancillary Results I: Improve to $\Omega(p^{1/2})$?




**Exciting observation:** Appear to have $\mathsf{SOS}_4(G_p) \sim p^{0.38...}$.

# Ancillary Results II: Improve to $\Omega(p^{0.38\ldots})$?

We use a simple $\boldsymbol{X}$, first used by [FK '03], later by [MW '13], but ultimately found to be insufficient by [BHKKP '19]:

$$X(S) := f(|S|) \cdot \mathbb{1}\{S \text{ is a clique in } G\}.$$

# Ancillary Results II: Improve to $\Omega(p^{0.38...})$?

We use a simple $\boldsymbol{X}$, first used by [FK '03], later by [MW '13], but ultimately found to be insufficient by [BHKKP '19]:

$$X(S) := f(|S|) \cdot \mathbb{1}\{S \text{ is a clique in } G\}.$$

**Theorem:** [Kelner '15] For $\mathsf{ER}(p, \frac{1}{2})$ graphs, such proves only

$$\mathbb{E}[\mathsf{SOS}_{2d}(\mathsf{ER})] = \widetilde{\Omega}(p^{1/(d+1)}).$$

# Ancillary Results II: Improve to $\Omega(p^{0.38...})$?

We use a simple $\boldsymbol{X}$, first used by [FK '03], later by [MW '13], but ultimately found to be insufficient by [BHKKP '19]:

$$X(S) := f(|S|) \cdot \mathbb{1}\{S \text{ is a clique in } G\}.$$

**Theorem:** [Kelner '15] For $\mathsf{ER}(p, \frac{1}{2})$ graphs, such proves only

$$\mathbb{E}[\mathsf{SOS}_{2d}(\mathsf{ER})] = \widetilde{\Omega}(p^{1/(d+1)}).$$

**Theorem:** [KY] For Paley graphs, such proves only

$$\mathsf{SOS}_4(G_p) = \Omega(p^{1/3}),$$

i.e., our main result cannot be improved without a fancier choice of $\boldsymbol{X}$ $\rightsquigarrow$ probably significantly harder to analyze.

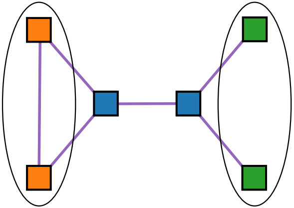
# Ancillary Results III: SOS and the "$\sqrt{p}$ Barrier"

Theoretical evidence: [BHKKP '19] proof depends on norm bounds for **graph matrices** formed from the $\{\pm 1\}$ adjacency matrix $\boldsymbol{A}$.

# Ancillary Results III: SOS and the "$\sqrt{p}$ Barrier"

Theoretical evidence: [BHKKP '19] proof depends on norm bounds for **graph matrices** formed from the $\{\pm 1\}$ adjacency matrix $A$.

**Example:** For a graph with sets of "left" and "right" vertices

$H =$ 

we get a matrix

$$M^H_{(a,b),(c,d)}(G) \approx \sum_{i,j} A_{a,b} A_{a,i} A_{b,i} A_{i,j} A_{j,c} A_{j,d}.$$

# Ancillary Results III: SOS and the "$\sqrt{p}$ Barrier"

Theoretical evidence: [BHKKP '19] proof depends on norm bounds for **graph matrices** formed from the $\{\pm 1\}$ adjacency matrix $\boldsymbol{A}$.

**Theorem:** [KY] There are some $H$ for which

$$\|\boldsymbol{M}^H(G_p)\| \gg \mathbb{E}\left[\|\boldsymbol{M}^H(\mathsf{ER})\|\right],$$

i.e., the key technical tool does not derandomize **in general** (luckily it does **for small** $H$ to get our lower bound).

# Ancillary Results III: SOS and the "$\sqrt{p}$ Barrier"

Theoretical evidence: [BHKKP '19] proof depends on norm bounds for **graph matrices** formed from the $\{\pm 1\}$ adjacency matrix $\boldsymbol{A}$.

**Theorem:** [KY] There are some $H$ for which

$$\|\boldsymbol{M}^H(G_p)\| \gg \mathbb{E}\left[\|\boldsymbol{M}^H(\mathsf{ER})\|\right],$$

i.e., the key technical tool does not derandomize **in general** (luckily it does **for small** $H$ to get our lower bound).

Basically, can build these by taking advantage of the discrepancy between

$$\boldsymbol{A}_{G_p}^2 = p\boldsymbol{I} - \mathbf{1}\mathbf{1}^\top,$$
$$\boldsymbol{A}_{\mathsf{ER}}^2 = p\boldsymbol{I} + \sqrt{p} \cdot (\text{random matrix}).$$

# Proof Idea

Also boils down to bounding $\|\boldsymbol{M}^H(G_p)\|$ for various $H$, but with different tools.

[AMP '16], [BHKKP '19]: **trace method** using $\mathbb{E}[\mathsf{Tr}((\boldsymbol{M}^H(\mathsf{ER}))^k)]$

[KY]: number-theoretic **character sum estimates**

# Proof Idea

Also boils down to bounding $\|\boldsymbol{M}^H(G_p)\|$ for various $H$, but with different tools.

[AMP '16], [BHKKP '19]: **trace method** using $\mathbb{E}[\mathsf{Tr}((\boldsymbol{M}^H(\mathsf{ER}))^k)]$

[KY]: number-theoretic **character sum estimates**

For $\chi : \mathbb{F}_p \to \mathbb{C}$ the **Legendre symbol** character,

$$(\boldsymbol{A}_{G_p})_{i,j} = \left\{ \begin{array}{ll} +1 & \text{if } i \sim j \\ -1 & \text{if } i \not\sim j \end{array} \right\} = \chi(i - j),$$

so polynomials in $\chi$ appear in entries of $\boldsymbol{M}^H$. Not many good tools for handling $\mathsf{Tr}((\boldsymbol{M}^H(G_p))^k)$ character sums, but we can use other case-by-case tricks to avoid these.

# In $\mathbb{F}_p^k$ Nobody Can Hear You Scream

However, in practice it is not always so easy to compute $G_{\text{geom}}$, even when the parameter space is a curve. We often have only meager global information about the sheaf in question, and so we try first to extract and then to exploit information about its local monodromy around each of the points at infinity of the parameter curve. One striking way in which pure lisse sheaves arising from exponential sums differ from the more traditional pure lisse sheaves arising as "cohomology along the fibres, with constant coefficients, of a proper smooth morphism" is that the local monodromy of the former can be quite wildly ramified, and can be so in quite interesting ways. This possibility can often be exploited to impose some very severe restrictions on $G_{\text{geom}}$. The underlying mechanisms of wild ramification and the restrictions it can impose are discussed in Chapter I.

One way in which the invariants and covariants of local monodromy can be detected and analyzed is through their interpretation as the difference between the compactly supported and the ordinary cohomology groups of the parameter curve with coefficients in the sheaf under discussion. This relation, together with a thorough discussion of the basic general facts about curves and their cohomology, is given in Chapter II, and systematically exploited in Chapter VII.

[Katz '88]

# Open Questions

# Open Questions

1. What is the exponent $\eta \in [\frac{1}{3}, \frac{1}{2}]$ in $\mathsf{SOS}_4(G_p) \sim p^{\eta}$?

# Open Questions

1. What is the exponent $\eta \in [\frac{1}{3}, \frac{1}{2}]$ in $\mathsf{SOS}_4(G_p) \sim p^\eta$?

2. If $\eta < \frac{1}{2}$, what other graphs can fully derandomize the Erdős-Rényi lower bound?

# Open Questions

1. What is the exponent $\eta \in [\frac{1}{3}, \frac{1}{2}]$ in $\mathsf{SOS}_4(G_p) \sim p^\eta$?

2. If $\eta < \frac{1}{2}$, what other graphs can fully derandomize the Erdős-Rényi lower bound?

3. If $\eta < \frac{1}{2}$, how can we extract formal proofs from numerical experiments with SOS?

# Open Questions

1. What is the exponent $\eta \in [\frac{1}{3}, \frac{1}{2}]$ in $\mathsf{SOS}_4(G_p) \sim p^\eta$?

2. If $\eta < \frac{1}{2}$, what other graphs can fully derandomize the Erdős-Rényi lower bound?

3. If $\eta < \frac{1}{2}$, how can we extract formal proofs from numerical experiments with SOS?

4. Higher degree sum-of-squares relaxations?

# Open Questions

1. What is the exponent $\eta \in [\frac{1}{3}, \frac{1}{2}]$ in $\mathsf{SOS}_4(G_p) \sim p^\eta$?

2. If $\eta < \frac{1}{2}$, what other graphs can fully derandomize the Erdős-Rényi lower bound?

3. If $\eta < \frac{1}{2}$, how can we extract formal proofs from numerical experiments with SOS?

4. Higher degree sum-of-squares relaxations?

5. How much of the structure of “clique space” of the Paley graph behaves like random graphs?

# Thank you!