Lower Bound on the Block-Diagonal SDP Relaxation for the Clique Number of the Paley Graph

Vlad Kobzar, joint work with

Yifan Wang

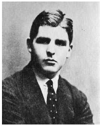
Yanling Shen

Outline

- The clique number problem and the Paley graphs
- Compressed sensing and sparse recovery motivations
- Block-diagonal (L^t), SOS and Lovasz-Schrivjer SDPs
- Our contributions
  - L^t lower bounds via FK pseudomoments
  - Localization lower bounds, and relaxation-localization trade-off
- Conclusion and future work

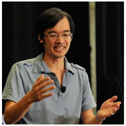
Paley graph clique number

- Classic problem in number theory and additive combinatorics
- Connected to Ramsey theory, random matrices, computational complexity and optimization, to name a few research areas
- Links to deterministic restricted isometries in compressed sensing and sparse recovery



Paley

Ramsey



Tao

Vallentin

Laurent

Gvozdenovic

Background

For any $G = (V, E)$:

- $K \subseteq V$ is a *clique* if each $i, j \in K$ are adjacent
- *Clique number*

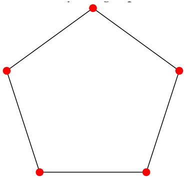
$$\omega(G) = \text{the size of a largest clique}$$

- $I \subseteq V$ is an *independence set* if each $i, j \in I$ are not adjacent
- *Independence number*

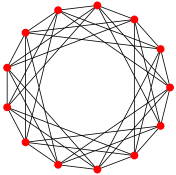
$$\alpha(G) = \text{the size of a largest independence set}$$

Finding $\omega(G)$ and $\alpha(G)$ is NP-hard for general graphs

Paley graph



$\omega(G_5) = 2$



$\omega(G_{13}) = 3$

Image credit: Wolfram

- A *Paley graph* $G_p = (V, E)$
  - $|V| = p$ where $p \equiv 1 \mod 4$ (Pythagorean prime)
  - $\{i, j\} \in E$ iff $i - j = a^2 \mod p$ for some $a \in \mathbb{Z}_p$
  - Strongly regular and self-complementary
- (We're not considering Paley graphs of prime power order p^s)

Connections to compressed sensing and sparse recovery

- SLOGAN: **compressible high-dimensional signal can be recovered from very few measurements**
- $x \in \mathbb{R}^n$ is s-sparse if it has no more than s nonzero entries
- When can you recover x exactly from few measurements y
- Sparse recovery experiment design of $A \in \mathbb{C}^{m \times n}$

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} \quad & A & \quad \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix}$$

where

$$s < m \ll n$$

Restricted isometry property (RIP)

- Guarantees that sparse recovery is robust to noise
- $A \in \mathbb{C}^{m \times n}$ satisfies RIP with distortion $0 < \delta < 1$ if for any s-sparse x

$$(1-\delta)\|x\|^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|^2$$

- Matrices with Gaussian i.i.d. entries satisfy RIP w.h.p. if

$$s \sim m/\log(n)$$

Square root bottleneck

- Deterministic constructions based on controlling "spikeness" or "localization" (coherence) of rows achieve

$$s \approx \sqrt{m}$$

- Include those based on the eigenvectors corresponding to $\lambda_1(A_{G_p})$ and $\lambda_2(A_{G_p})$ [Arash Amini and Marvasti, 2015]
- A combinatorial construction overcomes this bottleneck with

$$s = \Omega(m^{\frac{1}{2}+\epsilon})$$

for small $\epsilon > 0$ [Bourgain et al., 2011b, Bourgain et al., 2011a]
- Accordingly, random constructions are abundant but deterministic constructions are hard to find ("hay in the haystack")

Paley matrices

- Matrices constructed from rows of the DFT matrix corresponding to QR's mod p [Bandeira et al., 2013] support

$$s \sim \sqrt{p}$$

- Conditioned on a conjecture about the # of edges in any subgraph of G_p [Bandeira et al., 2016], these matrices support

$$s \sim p/\text{polylog}(p)$$

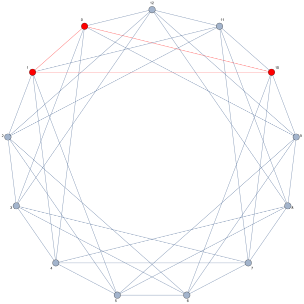
- Unconditional [Kaplan et al., 2019] for signals with a certain sparse structure

$$s = \Omega(m^{\frac{1}{2}+\frac{9}{40}})$$

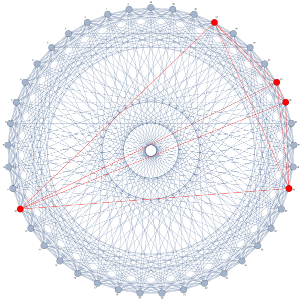
- A lower bound on $\omega(G_p)$ would lead to a lower bound on the distortion constant δ

Paley graph clique number

- Classic problem in number theory and additive combinatorics
- G_p share similarities with *Erdos-Renyi graphs* $\mathcal{G}(1/2, p)$
- Is $\omega(G_p) = O(\text{polylog } p)$, i.e., is G_p roughly a Ramsey graph?
- Note $\omega(\mathcal{G}(1/2, n)) \sim 2\log_2 n$



$\omega(G_{13}) = 3$



$\omega(G_{41}) = 5$

Existing bounds

- Upper bounds [Hanson and Petridis, 2021, Benedetto et al., 2021]
$$\omega(G_p) \leq (\sqrt{2p-1}+1)/2$$
- Improves on $\sqrt{p}$ by a constant prefactor.
- Lower bound for infinitely many primes [Graham and Ringrose, 1990]
$$\log p \cdot \log\log\log p \leq \omega(G_p)$$
- Conditioned on GRH [Montgomery, 1971],
$$\log p \cdot \log\log p \leq \omega(G_p)$$
- Numerical experiments [Bachoc et al., 2014]
$$\omega(G_p) \approx \mathrm{polylog}(p)$$

Integer program

- Easier to see in the context of the independence number $\alpha(G)$

$$\omega(G_p) = \max_{x \in \mathbb{R}^p} \sum_i x_i$$
$$\text{s.t. } x_i^2 = x_i \text{ for all } i \in V$$
$$x_i x_j = 0 \text{ for all } \{i, j\} \in E$$

- We focus on the clique problem $\omega(G)$ (i.e., take $x_i x_j = 0$ for all $\{i, j\} \notin E$)
- It makes connections to A_{G_p} more apparent

Nonconvex semidefinite matrix optimization

$$\max_{Y \in \mathbb{S}^{p+1 \times p+1}} \sum_{i \in \mathbb{Z}_p} Y_{\emptyset i}$$

$$\text{s.t. } Y_{ii}^2 = Y_{ii} \text{ for all } i \in V$$

$$Y_{ij} = 0 \text{ if } \{i, j\} \notin E$$

$$Y \succeq 0, \;\; Y_{\emptyset\emptyset} = 1$$

$$\text{rank}(Y) = 1$$

- This is equivalent to the previous program for $\omega(G_p)$
- Let $y = (1, x_1, \ldots, x_n)$ and reparametrize:

$$Y = yy^T = \begin{pmatrix} 1 & x_1 & x_2 & \ldots & x_p \\ x_1 & x_1 & x_1 x_2 & \ldots & x_1 x_p \\ x_2 & x_1 x_2 & x_2 & \ldots & x_1 x_p \\ \vdots & & & \ddots & \\ x_p & & & & x_p \end{pmatrix}$$

SOS-2 = Lovasz-Schrivjer$_0$ = L^1 convex relaxation

- Then we drop the nonconvex constraints

$$
\begin{aligned}
&\max \sum_{i \in V} y_i \\
&s.t. y \in \mathbb{R}^p, Y \in \mathbb{R}^{p \times p} \\
&Y_{ij} = 0 \text{ if } i \neq j, \{i,j\} \notin E \\
&Y_{ii} = y_i, \;\; i \in V \\
&\begin{pmatrix} 1 & y^\top \\ y & Y \end{pmatrix} \succeq 0
\end{aligned}
$$

- One can show this is equivalent to the Lovász ϑ function

SOS / Laserre-Parrilo hierarchy

- Denote the power sets of V of size $\leq t$ by $\mathcal{P}_t = \{S \subset V \mid |S| \leq t\}$.
- Now let $y = \left(\prod_{i \in S} x_i\right)_{S \in \mathcal{P}_t}$, $Y = yy^T$.
  - For example for $t = 2$, $y = (1, x_1, \ldots, x_n, x_1 x_2, \ldots x_2 x_1, \ldots x_p x_p)$.
- For $y \in \mathcal{P}_{2t}(V)$, $M_t(y)$ with $(M_t(y))_{I,J} = y_{I \cup J}$, $I, J \in \mathcal{P}_t(V)$ is called the moment matrix of y.

$$SOS_{2t}(G) = \max \sum_{i \in V} y_i$$
$$\text{s.t. } M_t(y) \succeq 0, y_0 = 1, y_{ij} = 0 \text{ if } \{i, j\} \in E.$$

Sum of squares relaxations

- An open problem proposed by Mixon and Bandeira is whether the SOS-4 relaxation of the Paley graph clique number breaks this barrier
- Xu & Kunisky
  - provided numerical evidence that $SOS_4(G_p)$ relaxation are $O(p^{\frac{1}{2}-\epsilon})$
  - proved an $\Omega(p^{\frac{1}{3}})$ lower bound
- However, $SOS_4(G_p)$ appears to be computationally intractable even for moderate $p \approx 250$.
- Gvozdenovic et al. introduced a more computationally efficient block-diagonal hierarchy of SDPs (L^t)

$$SOS_{2t}(G_p) \leq L^t(G_p)$$

Block Diagonal Hierarchy

- For $T \in \mathcal{P}_{t-1}(V)$, introduce $M(T;y)$, a principal sub-matrix of $M_t(y)$ indexed by $\bigcup_{S \subseteq T} \{S, S \cup \{i\}, i \in V\}$.

$$L^t(G) = \max \sum_{i \in V} y_i$$
$$M(T;y) \succeq 0 \ \forall \ |T| = t-1$$
$$y_0 = 1, y_{ij} = 0, \{i,j\} \in E.$$

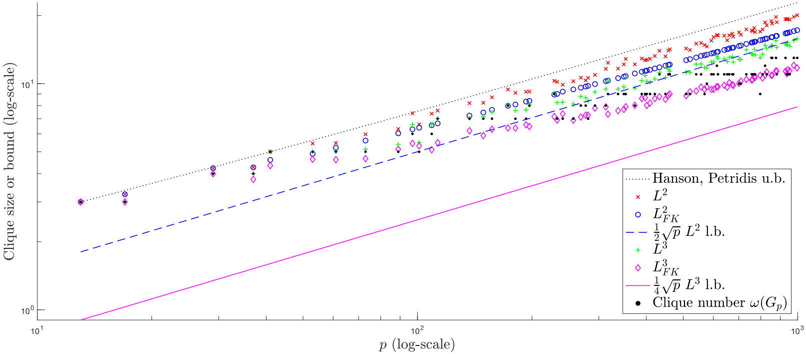
- Less computationally expensive than $SOS^{2t}(G)$.
- A relaxation of SOS because $M_t(y) \succeq 0$ is requires every submatrix to be PSD
- Block-diagonalized by zeta matrices - it is sufficient to use $p+1 \times p+1$ matrices in the constraints.

Main result: Lower bound on $L^t(G_p)$

- We proved the following lower bound

$$L^t(\overline{G}_p) \geq \frac{\sqrt{p}}{2^{t-1}} + o(\sqrt{p}).$$

- This shows L^t does not break $\sqrt{p}$ bottleneck for fixed t, but may beat it if $t(p)$ is a slowly increasing function of p.



Localization-relaxation trade-off

- Localization $G_{p,K}$ of subgraphs induced on vertices K adjacent to all vertices in G_p is another technique used to strengthen convex relaxations [Passuello, 2013, Magsino et al., 2019] and, more recently, spectral bounds on $\omega(G_p)$ [Kunisky, 2023].
- for any clique K of size a,

$$L^t(\overline{G}_{p,K}) \geq \frac{\sqrt{p}}{2^{a+t-1}} + o(\sqrt{p}). \tag{1}$$

- This shows L^t does not break $\sqrt{p}$ bottleneck for fixed t, but may beat it if $a(p)$ is a slowly increasing function of p.

Proof idea

- We construct a feasible point of L^2 using *Feige-Krauthgamer (FK) pseudomoments*, similarly to such construction in [Kunisky and Yu, 2022] for SOS_{2t}.
- The FK program $L^2_{FK}(G_p)$ corresponding to $L^2(G_p)$ is defined by replacing $A_{\{0\}}$ with:

$$A_{\{0\}} = \left(\begin{array}{c|c|c} y_{\{0\}} & y_{\{0\}} & y_{\{0,1\}}(A_{G_p})_{0,1:\text{end}} \\ \hline y_{\{0\}} & y_{\{0\}} & y_{\{0,1\}}(A_{G_p})_{0,1:\text{end}} \\ \hline \begin{array}{c} y_{\{0,1\}}\times \\ (A_{G_p})_{1:\text{end},0} \end{array} & \begin{array}{c} y_{\{0,1\}}\times \\ (A_{G_p})_{1:\text{end},0} \end{array} & y_{\{0,1\}}\text{diag}(A_{G_p})_{1:\text{end},0} + \alpha_3 M' \end{array}\right)$$

  where M' is the indicator matrix of triangles in G_p of the form $\{0, i, j\}$ for $1 \leq i, j < p$, and reducing the number of scalar optimization variables $y_{\{0,\alpha,\beta\}}$ corresponding to the orbits of triangles to the single $\alpha_3 \in \mathbb{R}$.
- Use the Schur complements to reduce the PSD constraints to a system of scalar inequalities for $y_{\{0\}}$ $y_{\{0,1\}}$ and α_3.

Future direction - symmetries and upper bounds

- We plan to upper bound L^2 and L^3, and therefore $\omega(G_p)$, by constructing feasible points of the corresponding dual programs.
- Since the edges and the edges triples (triangles) form orbits under $Aut(G_p)$, the number of optimization variables is proportional to the number of the representatives of such orbits
- Since a Paley graph is edge-transitive, the representatives of such orbits are given by $\{0, 1, \beta\}$ where both β and $\beta - 1$ are squares in $\mathbb{Z}_p$; there are approximately $(p-5)/24$ orbits.
- The upper bound problem can be reduced to a problem of studying the e.s.d of indicators of orbits as $p \to \infty$

Acknowledgements

V.A.K. acknowledges support by NSF grant DMS-1937254. We thank Afonso Bandeira for suggesting the research direction that led to this paper; we are also very grateful to him, Nebosja Gvozdenovic, Tim Kunisky and Monique Laurent for their helpful input.

Arash Amini, H. B.-S. and Marvasti, F. (2015).
From Paley graphs to deterministic sensing matrices with real-valued Gramians.
In *2015 International Conference on Sampling Theory and Applications (SampTA)*.

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

Bandeira, A. S., Fickus, M., Mixon, D. G., and Wong, P. (2013).
The road to deterministic matrices with the restricted isometry property.
J. Fourier Anal. Appl., 19:1123–1149.

Bandeira, A. S., Mixon, D. G., and Moreira, J. (2016).
A conditional construction of restricted isometries.
International Mathematics Research Notices, 2017:2:372–381.

Benedetto, D. D., Solymosi, J., and White, E. P. (2021).
On the directions determined by a Cartesian product in an affine Galois plane.
Combinatorica, 41(6):755–763.

Bourgain, J., Dilworth, S., Ford, K., Konyagin, S., and Kutzarova, D. (2011a).
Breaking the k^2 barrier for explicit RIP matrices.
In *STOC*, pages 637–644.

Bourgain, J., Dilworth, S., Ford, K., Konyagin, S., and Kutzarova, D. (2011b).
Explicit constructions of RIP matrices and related problems.
Duke Math. J., 159:145 – 185.

Graham, S. and Ringrose, C. (1990).
Lower bounds for least quadratic non-residues.
In *Analytic number theory*, pages 269–309. Springer.

Hanson, B. and Petridis, G. (2021).

Refined estimates concerning sumsets contained in the roots of unity.
In *Proceedings of the London Mathematical Society*, volume 122(3), pages 353–358.

Kaplan, A., Pohl, V., and Boche, H. (2019).
Deterministic matrices with a restricted isometry property for partially structured sparse signals.
In *13th International conference on Sampling Theory and Applications (SampTA)*.

Kunisky, D. (2023).
Spectral pseudorandomness and the road to improved clique number bounds for Paley graphs.
https://arxiv.org/abs/2211.02713.

Kunisky, D. and Yu, X. (2022).
A degree 4 sum-of-squares lower bound for the clique number of the Paley graph.
https://arxiv.org/abs/2211.02713.

Magsino, M., Mixon, D. G., and Parshall, H. (2019).

Linear programming bounds for cliques in Paley graphs.
In *Proc. SPIE 11138, Wavelets and Sparsity XVIII*.

Montgomery, H. L. (1971).
Topics in multiplicative number theory, volume 227.
Springer.

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