The Lovász number of random circulant graphs

Afonso S. Bandeira, *ETH Zürich*
Jarosław Błasiok, *ETH Zürich*
Daniil Dmitriev, *ETH Zürich*
Ulysse Faure, *ETH Zürich*
Anastasia Kireeva, *ETH Zürich*
Dmitriy Kunisky, *Johns Hopkins University*

***Abstract*—This paper addresses the behavior of the Lovász number for dense random circulant graphs. The Lovász number is a well-known semidefinite programming upper bound on the independence number. Circulant graphs, an example of a Cayley graph, are highly structured vertex-transitive graphs on integers modulo n, where the connectivity of pairs of vertices depends only on the difference between their labels. While for random circulant graphs the asymptotics of fundamental quantities such as the clique and the chromatic number are well-understood, characterizing the exact behavior of the Lovász number remains open. In this work, we provide upper and lower bounds on the expected value of the Lovász number and show that it scales as the square root of the number of vertices, up to log log factor. Our proof reduces the semidefinite program formulation of the Lovász number to a linear program with random objective and constraints via diagonalization of the adjacency matrix of a circulant graph by the discrete Fourier transform (DFT). This leads to a problem about controlling the norms of vectors with sparse Fourier coefficients, which we study using results on the restricted isometry property of subsampled DFT matrices.**

***Index Terms*—Semidefinite programming, random graphs, restricted isometry property.**

I. Introduction

The Lovász number ϑ is a well-known statistic of an arbitrary simple undirected graph G. As Lovász first observed in [18], one can define a number $\vartheta(G)$ as the value of a certain semidefinite program (SDP) whose constraints depend on the adjacency matrix of G. The Lovász number provides an upper bound on the Shannon capacity of the graph and satisfies the following inequalities:

$$\omega(G) \le \vartheta(\overline{G}) \le \chi(G), \tag{1}$$

where $\omega(G)$ is the size of the largest clique in G, $\chi(G)$ is the chromatic number of G, and $\overline{G}$ is the complement of G. This observation is remarkable, since ϑ is computable in polynomial time, while ω and χ are famously NP-hard to compute.

The Lovász number has been studied for a variety of random graph models including the classical Erdős-Rényi (ER) random graph $G(n,p)$. Its expected value was first studied by Juhász [15], who showed that $\mathbf{E}\,\vartheta(G) = \Theta(\sqrt{n/p})$ for $\frac{\log^6 n}{n} \le p \le 1/2$. For $p = 1/2$, Arora and Bhaskara [1] showed that $\vartheta(G)$ concentrates around its median in an interval of polylogarithmic length. In the sparse regime $p < n^{-1/2}$, it has been further shown that $\vartheta(G)$ concentrates around its median in an interval of constant length [9]. To the best of our knowledge, determining the correct constant in the $\Theta(\sqrt{n/p})$ asymptotic remains an open question.

In this work, we focus on a class of random *circulant* graphs (RCGs), a family of vertex-transitive graphs with a circulant adjacency matrix; see Definitions 3 and 4. We emphasize that RCGs are fully determined by the connectivity of any given single vertex. Therefore, a dense RCG can be generated with $\frac{n-1}{2}$ random bits, where each bit affects the presence of n edges, in contrast to the $\frac{n(n-1)}{2}$ random bits in $G(n,1/2)$, each affecting just one edge. In this sense, RCGs may be viewed as a "partial derandomization" of ER graphs. Indeed, circulant graphs are precisely Cayley graphs on the group $\mathbb{Z}_n$, and general random Cayley graphs have long been studied for similar purposes in theoretical computer science.

It is therefore of interest to understand to what extent the above results for ER graphs also apply to RCGs. For dense RCGs, the asymptotics of the clique number and the chromatic number are well-understood: [12] showed a high-probability upper bound on the clique number $\omega(G) = O(\log n)$, and later [11] proved that the chromatic number is at most $(1 + o(1))\frac{n}{2\log_2 n}$ with high probability. These results imply bounds on the Lovász number through (1), but the resulting upper and lower bounds are far apart.

In this work, we prove much sharper upper and lower bounds on the expected Lovász number of a dense RCG.

Theorem 1. *There exists a constant $C > 0$ such that, for a dense random circulant graph G on n vertices (Definition 4),*

$$\sqrt{n} \le \mathbf{E}\,\vartheta(G) \le C\sqrt{n \log\log n}. \tag{2}$$

Proof Strategy: Our proof of the upper bound in Theorem 1 relies on the algebraic structure of circulant graphs. First, following [19], we transform the SDP formulation of $\vartheta(G)$ to a linear program (LP) using the fact that the circulant matrices are diagonalizable by a discrete Fourier transform (DFT) . Lemma 2 gives the resulting LP:

$$\vartheta(G) = \max_{(y_0,\dots,y_{n-1})\in\mathbf{R}^n} \langle y, g\rangle,$$

$$\text{subject to } \begin{cases} y_k = y_{n-k} \text{ for } k = 1,\dots,n-1, \\ \|y\|_1 = 1, y \ge \mathbf{0}, \\ \langle y, f_k\rangle = 0 \text{ for all edges } (0,k). \end{cases} \tag{3}$$

Here, f_k is the k-th row of the DFT matrix F, and $g := Fb$ for $b \in \{\pm 1\}^n$ with $b_0 = 1$ and $b_k = 1$ if $(0,k)$ is not an edge, and -1 otherwise, for $1 \le k \le n-1$. We denote $\mathbf{0} := (0,\dots,0)$ and $y \ge \mathbf{0}$ stands for entrywise positivity of y.

The last constraint in (3) requires the Fourier transform of y to have a specific sparsity pattern. Uncertainty principles for the Fourier transform (see, e.g., [3]) then suggest that all feasible vectors y must be dense [10]. A quantitative version of this "density" would be enough to bound the LP. To illustrate, suppose that y is a feasible vector with $\|y\|_1 = 1$ and its mass is spread almost uniformly among its coordinates, i.e., that $\|y\|_2 \le \frac{c}{\sqrt{n}}\|y\|_1 = \frac{c}{\sqrt{n}}$, for some constant $c > 0$. Since $\|g\|_2 = n$, Cauchy-Schwarz inequality would give $\langle y, g\rangle \le \|y\|_2\|g\|_2 \le c\sqrt{n}$, proving upper bound in Theorem 1 without the extra $\sqrt{\log\log n}$ factor.

The second part of our proof, Lemma 5, makes the aforementioned intuition rigorous, relying on the *restricted isometry property* (RIP, Definition 5). The f_k in our constraints form a so-called *subsampled DFT basis*, which is a random subset of the Fourier basis. The RIP for such bases is in fact a celebrated topic in the compressed sensing literature. RIP was first introduced and studied for subsampled DFT bases in seminal work of Candès and Tao [7], and since then, one of the central questions for compressed sensing is the number of f_k needed for RIP to hold. Lemma 8 describes a simplified version of the current best bound due to [14] which is sufficient for our purposes. Interestingly, our upper bound proof only uses the fact that feasible solutions of (3) lie on a (random) nullspace of a subsampled DFT matrix, and omits the positivity constraint $y \ge \mathbf{0}$. However, as we discuss in Section IV, we believe that this constraint is important for tighter results.

II. Preliminaries

Notation: For $n \in \mathbf{N}$, let $[n] := \{0, \dots, n-1\}$. We index vectors and matrices by $[n]$: for $x \in \mathbf{R}^n, x = (x_0, \dots, x_{n-1})$. We write $x \ge \mathbf{0}$ for entrywise positivity. For $n \in \mathbf{N}$, we denote by $G = (V, E)$ a graph with vertex set $V = [n]$ and edge set $E \subseteq (V \times V) \setminus \{(k,k) \text{ for } k \in V\}$. For a graph $G = (V, E)$ we define its complement $\overline{G} = (V, E')$, where $E' = \{(u,v) \text{ s.t. } u \ne v \text{ and } (u,v) \notin E\}$. We use the standard asymptotic notation, $O(\cdot), \Omega(\cdot)$, and $\Theta(\cdot)$ to describe the order of the growth of functions associated with the limit of the graph dimension n. For $x \in \mathbf{R}^n$, we denote $\|x\|_1 := \sum_{i=0}^{n-1}|x_i|$, $\|x\|_2 := \left(\sum_{x=0}^{n-1} x_i^2\right)^{1/2}$, and $\|x\|_\infty := \max_k |x_k|$.

Discrete Fourier Transform: Let $F \in \mathbf{C}^{n\times n}$ be the discrete Fourier transform matrix: $F_{jk} = \exp(-2\pi i jk/n)$ for $j, k \in [n]$. For $k \in [n]$, let f_k denote the k-th row of F. We associate a matrix $\widetilde{F} \in \mathbf{R}^{m\times n}$ to any RCG G consisting of subsampled rows of F.

Definition 1. *For any RCG G, let $\widetilde{F} \equiv \widetilde{F}(G) \in \mathbf{C}^{m\times n}$ (with m the number of neighbors of* 0 *in G) be defined as a submatrix of F, including row f_k if $(0,k) \in E(G)$.*

Definition 2. *The* Lovász theta number *$\vartheta(G)$ is defined as the solution to the following SDP (J is the all-ones matrix),*

$$\vartheta(G) := \max_{X\in\mathbf{R}^{n\times n}} \Big\{\langle X, J\rangle, \textit{ such that } X \succeq 0, \operatorname{Tr} X = 1, \\ X_{ij} = 0 \textit{ for all } (i,j) \in E(G)\Big\}. \tag{4}$$

Definition 3. *A graph on n vertices is called* circulant *if there is an ordering of its vertices such that its adjacency matrix is circulant. Equivalently, a circulant graph is a Cayley graph of a cyclic group $\mathbf{Z}_n$.*

This definition implies that a circulant graph is described by listing the neighbors of a single root vertex (say vertex 0), since $(i,j) \in E \iff (0, i-j) \in E$. In this text, we focus on *dense random* circulant graphs.

Definition 4. *For odd n, a* dense random circulant graph *(RCG) is a random Cayley graph of a cyclic group $\mathbf{Z}_n$. It is obtained in the following way: uniformly sample $x \in \{0,1\}^m, m = \frac{n-1}{2}$, and define the first row of the adjacency matrix as*

$$R = (0 \quad x \quad \overleftarrow{x}), \tag{5}$$

where $\overleftarrow{x}_i := x_{m-i-1}$. Circulate R to obtain the complete adjacency matrix.

For a circulant graph G we define a vector $g := Fb$, where $b \in \{\pm 1\}^n$ with $b_0 = 1$ and $b_k = 1$ if $(0,k)$ is not an edge, and -1 otherwise, for $1 \le k \le n-1$.

Definition 5 (Restricted isometry property)**.** *A matrix $A \in \mathbf{C}^{q\times n}$ is said to satisfy (k,ε)-restricted isometry property, for $k \le n$ and $\varepsilon \in (0,1)$, if for all k-sparse $x \in \mathbf{C}^n$ we have that*

$$(1-\varepsilon)\|x\|_2^2 \le \|Ax\|_2^2 \le (1+\varepsilon)\|x\|_2^2. \tag{6}$$

III. Proof of Main Theorem

Let G be a circulant graph. As noted in [19], for circulant graphs the SDP formulation of the Lovász number can be rewritten as the following linear program:

$$\vartheta(G) = \max_{x\in\mathbf{R}^n} \sum_{i\in[n]} x_i, \\ \text{subject to } \begin{cases} x_k = x_{n-k} \text{ for all } k \in [n]\setminus\{0\}, \\ x_0 = 1, Fx \ge \mathbf{0}, \\ x_k = 0 \text{ for all edges } (0,k), \end{cases} \tag{7}$$

Table I shows four equivalent linear programs, arising from strong duality (see, e.g., [4]) and switching between 'time' and 'frequency' domains. For the latter, we perform the change of variables, $y := Fx$ and $t := Fz$ respectively.

All formulations share the same structure: the optimization objective is deterministic, while the set of feasible solutions is random through the random circulant graph structure. The following proposition introduces randomness to the objective, which is a crucial part of our argument.

Lemma 2. *Let G be a dense RCG and $\widetilde{F}$ be a subsampled DFT matrix, see Definition 1. Let $g := Fb \in \mathbf{R}^n$, for $b \in \{\pm 1\}^n$ with $b_k = 1$ if $(0,k)$ is not an edge and -1 otherwise. Then,*

$$\vartheta(G) = \max_{y\in\mathbf{R}^n} \langle y, g\rangle, \\ \textit{subject to } \begin{cases} y_k = y_{n-k} \textit{ for all } k \in [n]\setminus\{0\}, \\ \|y\|_1 = 1, y \ge \mathbf{0}, \\ y \in \ker \widetilde{F}, \end{cases} \tag{8}$$

TABLE I
FOUR EQUIVALENT LPS FOR $\vartheta(G)$.

’time’ domain

Primal	**Dual**
$\max_{x\in\mathbf{R}^n} \sum_i x_i$	$\min_{z\in\mathbf{R}^n} 1+\sum_i z_i$
s.t. $x_k = x_{n-k}$	s.t. $z_k = z_{n-k}$
for all $k \in [n]\setminus\{0\}$,	for all $k \in [n]\setminus\{0\}$,
$x_0 = 1,\ Fx \geq \mathbf{0}$,	$z \geq \mathbf{0}$,
$x_k = 0$	$\langle z, f_k\rangle = -1$
for all $(0,k) \in E(G)$.	for all $(0,k) \in E(\overline{G})$.

’frequency’ domain

Primal	**Dual**
$\max_{y\in\mathbf{R}^n} ny_0$	$\min_{t\in\mathbf{R}^n} 1+nt_0$
s.t. $y_k = y_{n-k}$	s.t. $t_k = t_{n-k}$
for all $k \in [n]\setminus\{0\}$,	for all $k \in [n]\setminus\{0\}$,
$\|y\|_1 = 1,\ y \geq \mathbf{0}$,	$Ft \geq \mathbf{0}$,
$\langle y, f_k\rangle = 0$	$t_k = -1/n$
for all $(0,k) \in E(G)$.	for all $(0,k) \in E(\overline{G})$.

Proof. We use the primal formulation in the frequency domain and observe that $ny_0 = \langle y, \sum_{k\in[n]} f_k\rangle$. Since feasible vectors y are orthogonal to $\widetilde{F}$, i.e., $y \in \ker \widetilde{F}$, after subtracting $2\sum_{(0,k)\in E(G)}\langle y, f_k\rangle$ from $\langle y, \sum_{k\in[n]} f_k\rangle$ we obtain

$$\langle y, \sum_{k\in[n]} f_k\rangle = \langle y, \sum_{(0,k)\notin E(G)} f_k - \sum_{(0,k)\in E(G)} f_k\rangle = \langle y, g\rangle. \tag{9}$$

□

By the definition of graph G, $b_0 = 1$, and $b_1, b_2, \ldots, b_{\frac{n-1}{2}} \overset{\text{iid}}{\sim} \text{Unif}\{-1,1\}$. Since $\max_{jk}|F_{jk}| = 1$, we can bound $\|g\|_\infty$, leading to the following upper bound on ϑ.

Lemma 3. *Let G be a dense RCG. Then,*

$$\mathbf{P}(\vartheta(G) \leq 1 + 4\sqrt{n\log n}) \geq 1 - \frac{2}{n}. \tag{10}$$

Proof. We show that each entry of g is small with high probability. Indeed, for any $k \in [n]$,

$$\begin{aligned}\mathbf{P}(|g_k| > 1 + 4\sqrt{n\log n}) &= \mathbf{P}(|\langle f_k, b\rangle| > 1 + 4\sqrt{n\log n}) \\ \leq \mathbf{P}\left(\left|\sum_{j=1}^{(n-1)/2} X_j\right| > 2\sqrt{n\log n}\right) &\leq \frac{2}{n^2},\end{aligned} \tag{11}$$

where $X_j := \Re(F_{kj})b_j \in [-1,1]$, and the last step follows from Hoeffding’s inequality (Lemma 7). Applying union bound over $k \in [n]$, we obtain

$$\mathbf{P}(\|g\|_\infty > 1 + 4\sqrt{n\log n}) \leq \frac{2}{n}. \tag{12}$$

Thus, on a complement event, for any feasible vector y of (3), we can simply upper bound $\langle y, g\rangle \leq \|y\|_1\|g\|_\infty \leq 1 + 4\sqrt{n\log n}$, which finishes the proof. □

The upper bound in Theorem 1 would follow if we could show $\max_k g_k = O(\sqrt{n\log\log n})$ with high probability. However, this is too optimistic: since we expect that the coordinates of g behave like standard Gaussian random variables and are uncorrelated, we also expect that $\max_k g_k = \Theta(\sqrt{n\log n})$. Fortunately, as the next lemma shows, only a vanishing fraction of entries is of order at least $\sqrt{n\log\log n}$.

Lemma 4. *There exists a constant $C > 0$, such that for $\mathcal{I} := \{k \in [n] : |g_k| \geq C\sqrt{n\log\log n}\}$, it holds*

$$\mathbf{P}\left(|\mathcal{I}| \leq \frac{n}{\log^{10} n}\right) \geq 1 - \frac{1}{\log^{10} n}. \tag{13}$$

Proof. We express $|\mathcal{I}| = \sum_{k=0}^{n-1} Y_k$, where $Y_k = \mathbb{I}\{|g_k| \geq C\sqrt{n\log\log n}\}$. Using Hoeffding’s inequality we obtain, for C large enough,

$$\mathbf{E}|\mathcal{I}| = \sum_{k=0}^{n-1}\mathbf{P}(|g_k| \geq C\sqrt{n\log\log n}) \leq \frac{n}{\log^{20} n}, \tag{14}$$

where the constant on the right hand side is absorbed into logarithm, and its power is chosen for the technical reasons. Plugging this bound into Markov’s inequality we get

$$\mathbf{P}\left(|\mathcal{I}| \geq \frac{n}{\log^{10} n}\right) \leq \frac{1}{\log^{10} n}. \tag{15}$$

□

The constraint $y \in \ker \widetilde{F}$ was so far only used to change the objective function from ny_0 to $\langle y, g\rangle$. Next lemma highlights another important consequence of this constraint, namely, an upper bound on the $\|y\|_2$.

Lemma 5. *For large enough n, with probability at least $1 - \frac{1}{n}$ all $x \in \ker \widetilde{F}$ satisfy $\|x\|_2 \leq \frac{\log^2 n}{\sqrt{n}}\|x\|_1$.*

Proof. We adapt the existing results in the literature regarding the RIP of the subsampled Fourier basis.

Consider the following coupling: let $\hat{b} \in \{0,1\}^n$ with $\hat{b}_0 = 0$ and $\hat{b}_k \overset{\text{iid}}{\sim} \text{Ber}\left(\frac{\sqrt{2}-1}{\sqrt{2}}\right)$ for $k = 1, \ldots, n-1$. Let $\widetilde{b} \in \{0,1\}^n$ be defined as follows:

$$\widetilde{b}_k = \begin{cases} 0, & \text{for } k = 0, \\ 1, & \text{if } \hat{b}_k = 1 \text{ or } \hat{b}_{n-k} = 1, \text{ for } k \geq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

Note that (i) the distribution of $\widetilde{b}$ is the same as the distribution of the adjacency vector for the vertex 0 in the random circulant graph G and (ii) $\widetilde{b}_i = 0$ implies $\hat{b}_i = 0$. Let $q := \sum_k \hat{b}_k$. Define $\widehat{F} \in \mathbf{C}^{q\times n}$ to be the matrix consisting of subsampled rows of F rescaled by $1/\sqrt{q}$, where the k-th row is included if and only if $\hat{b}_k = 1$.

To show that $\widehat{F}$ satisfies the RIP, we apply Lemma 8. To ensure its requirements, we condition on the following two events. First, since we do not include row 0 in our construction, we condition on the event that among the uniformly subsampled rows, row 0 is not present; this increases the probability of a bad event by at most a constant factor. Additionally, we condition on a high probability event that $q \ge \lceil n/4 \rceil$. Lemma 8 then implies that there exist constants $c > 0$ and $0 < \varepsilon < 1/3$, such that with probability at least $1 - 1/n$, $\widehat{F}$ satisfies the RIP with parameters $k = \frac{cn}{\log^3 n}$ and ε.

On this event, by Lemma 9, it follows that

$$\|x\|_2 \le \frac{C(\varepsilon)\log^{3/2} n}{\sqrt{cn}}\|x\|_1 \le \frac{\log^2 n}{\sqrt{n}}\|x\|_1, \tag{17}$$

for all $x \in \ker \widehat{F}$ and large enough n, where we absorbed the constants in the additional $(\log n)^{1/2}$ factor in the numerator. Since $\widehat{F}$ consists of a subset of rows of $\widetilde{F}$, all $x \in \ker \widetilde{F}$ are also in $\ker \widehat{F}$, so the proof is complete. □

Remark 6 (Alternative proof technique)**.** *Lemma 5 also follows from an intermediate step in the proof of RIP of the subsampled Fourier matrix in [14]. More specifically, in our notation [14, Theorem 3.1] implies that* $\|\widehat{F}x\| \ge (1-\varepsilon)\|Fx\|_2^2 - C\varepsilon/k\|x\|_1^2$ *with high probability, and since* $x \in \ker \widehat{F}$*, it follows that* $\|x\|_2 \le \frac{\log^2 n}{\sqrt{n}}$.

Now we present the proof of our main result.

Proof of Theorem 1. We begin with the lower bound $\mathbf{E}\,\vartheta(G) \ge \sqrt{n}$. Since G is vertex-transitive, it holds that $\vartheta(G)\vartheta(\overline{G}) = n$, see [18, Theorem 8]. Therefore,

$$\log n = \mathbf{E}\log\vartheta(G)\vartheta(\overline{G}) = 2\,\mathbf{E}\log\vartheta(G) \le 2\log\mathbf{E}\,\vartheta(G), \tag{18}$$

where we used the fact that G equals in distribution to $\overline{G}$ together with Jensen's inequality and linearity of the expected value. Upon exponentiating we obtain

$$\mathbf{E}\,\vartheta(G) \ge \sqrt{n}. \tag{19}$$

To prove the upper bound, we use the LP formulation of the Lovász number as in Lemma 2. Let A denote the intersection of the events of Lemmas 3 and 5, with $\mathbf{P}(A) \ge 1 - \frac{3}{n}$ from union bound, and let B denote the event of Lemma 4. Since $\mathbf{E}[\vartheta|\overline{A} \text{ or } \overline{B}]\,\mathbf{P}(\overline{A} \text{ or } \overline{B}) = O(1)$, we condition on A and B in the following. For constant C defined in Lemma 4, we split g into two parts, g_{small} and g_{large}, where

$$(g_{\text{small}})_k = \begin{cases} g_k & \text{if } |g_k| < C\sqrt{n\log\log n}, \\ 0 & \text{otherwise,} \end{cases} \tag{20}$$

and $g_{\text{large}} = g - g_{\text{small}}$. Then, $\langle y, g\rangle = \langle y, g_{\text{small}}\rangle + \langle y, g_{\text{large}}\rangle$. We bound each term separately: first,

$$\langle y, g_{\text{small}}\rangle \le \|y\|_1\|g_{\text{small}}\|_\infty = O(\sqrt{n\log\log n}). \tag{21}$$

On the event B we have that g_{large} is $\frac{n}{\log^{10} n}$-sparse. From (12) $\|g_{\text{large}}\|_\infty = O(\sqrt{n\log n})$, which implies that $\|g_{\text{large}}\|_2 = O(n/\log^4 n)$. Using Cauchy-Schwartz inequality together with Lemma 5, we bound the second term as follows:

$$\langle y, g_{\text{large}}\rangle \le \|y\|_2\|g_{\text{large}}\|_2 \le \frac{\log^2 n}{\sqrt{n}} \cdot \frac{n}{\log^4 n} = O(\sqrt{n}), \tag{22}$$

which completes the proof. □

IV. Discussion

Based on numerical observations, we formulate the following conjecture.

Conjecture 1. *Let G be a dense random circulant graph. Then,*

$$\mathbf{E}\,\vartheta(G) = (1 + o(1))\sqrt{n}. \tag{23}$$

Existing lower bounds against RIP (see [3, 5]) do not allow us to use our proof strategy for showing Conjecture 1. Indeed, there exist $\frac{n}{\log n}$-sparse vectors in the kernel of $\widetilde{F}$, which contradicts the desired inequality $\|y\|_2 \le \frac{C}{\sqrt{n}}\|y\|_1$ for $y \in \ker \widetilde{F}$. However, it is still possible that no cn-sparse *entrywise positive* vector exists in the kernel of $\widetilde{F}$, for small enough constant $c > 0$. It is also plausible that constructing a feasible vector for the dual programs in Table I may lead to tighter upper bounds. We leave these questions for the future work.

Paley graph: A classical example of a circulant graph is *Paley* graph. For a prime $p \equiv 1 \mod 4$, it is defined as the graph on p vertices with vertices i and j connected if and only if $i - j$ is a quadratic residue modulo p, see [8, 2]. Paley graphs are believed to exhibit certain *pseudorandom* properties, and bounding its independence number is a long-standing open problem in number theory and combinatorics [13]. This quantity can be upper bounded by the Lovász number of a certain subgraph called 1-localization which is circulant [17].

Recently, several optimization based approaches were considered, see [16, 17, 20]. In [19], a numerical evidence similar to Conjecture 1 regarding subgraphs of Paley graph was observed, which if true, recovers the best known upper bound on the independence number due to [13].

V. Useful definitions and inequalities

Lemma 7 (Hoeffding's inequality)**.** *Let $X_1, \dots, X_n$ be independent random variables, such that $\mathbf{E}\,X_i = 0$ and $a \le X_i \le b$ almost surely. Then,*

$$\mathbf{P}\left(\left|\sum_{i=1}^n X_i\right| \ge t\right) \le 2\exp\left(-\frac{2t^2}{n(b-a)^2}\right) \tag{24}$$

Lemma 8 (RIP of subsampled DFT matrix, [14])**.** *Let $F \in \mathbf{C}^{n\times n}$ be a DFT matrix: $F_{jk} = \exp(-2\pi ijk/n)$ for $j, k \in [n]$. There exist $c > 0$ and $0 < \varepsilon < 1/3$, such that for all n large enough, a matrix consisting of $q \ge \lceil n/4 \rceil$ uniformly subsampled rows of F and rescaled by $1/\sqrt{q}$ has (k,ε)-RIP for $k = \frac{cn}{\log^3 n}$, with probability at least $1 - 2^{\Omega(-\log^2 n)}$.*

Lemma 9 (e.g. [6], Theorem 11)**.** *If $A \in \mathbf{C}^{m\times n}$ satisfies the RIP with parameters k and $\varepsilon < 1/3$, then there exists $C = C(\varepsilon)$, such that for any $x \in \ker A$ we have that*

$$\|x\|_2 \le \frac{C}{\sqrt{k}}\|x\|_1. \tag{25}$$

References

[1] S. Arora and A. Bhaskara, *A note on the lovász theta number of random graphs*,

[2] R. Baker, G. Ebert, J. Hemmeter, and A. Woldar, *Maximal cliques in the paley graph of square order*, Journal of statistical planning and inference **56**, 33–38 (1996).

[3] A. S. Bandeira, M. E. Lewis, and D. G. Mixon, *Discrete uncertainty principles and sparse signal processing*, Journal of Fourier Analysis and Applications **24**, 935–956 (2018).

[4] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear programming and network flows* (John Wiley & Sons, 2011).

[5] J. Blasiok, P. Lopatto, K. Luh, J. Marcinek, and S. Rao, "An improved lower bound for sparse reconstruction from subsampled hadamard matrices", 2019 ieee 60th annual symposium on foundations of computer science (focs) (IEEE, 2019), pp. 1564–1567.

[6] J. Cahill and D. G. Mixon, *Robust width: a characterization of uniformly stable and robust compressed sensing*, Excursions in Harmonic Analysis, Volume 6: In Honor of John Benedetto's 80th Birthday, 343–371 (2021).

[7] E. J. Candes and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies?*, IEEE transactions on information theory **52**, 5406–5425 (2006).

[8] S. D. Cohen, *Clique numbers of paley graphs*, Quaestiones Mathematicae **11**, 225–231 (1988).

[9] A. Coja-Oghlan, *The lovász number of random graphs*, Combinatorics, Probability and Computing **14**, 439–465 (2005).

[10] L. Demanet and P. Hand, *Scaling law for recovering the sparsest element in a subspace*, Information and Inference: A Journal of the IMA **3**, 295–309 (2014).

[11] B. Green and R. Morris, *Counting sets with small sumset and applications*, Combinatorica **36**, 129–159 (2016).

[12] B. Green*, *Counting sets with small sumset, and the clique number of random cayley graphs*, Combinatorica **25**, 307–326 (2005).

[13] B. Hanson and G. Petridis, *Refined estimates concerning sumsets contained in the roots of unity*, Proceedings of the London Mathematical Society **122**, 353–358 (2021).

[14] I. Haviv and O. Regev, "The restricted isometry property of subsampled fourier matrices", Geometric aspects of functional analysis: israel seminar (gafa) 2014–2016 (Springer, 2017), pp. 163–179.

[15] F. Juhász, *The asymptotic behaviour of lovász'theta function for random graphs*, Combinatorica **2**, 153–155 (1982).

[16] V. A. Kobzar and K. Mody, "Revisiting block-diagonal sdp relaxations for the clique number of the paley graphs", 2023 international conference on sampling theory and applications (sampta) (IEEE, 2023), pp. 1–5.

[17] D. Kunisky, *Spectral pseudorandomness and the road to improved clique number bounds for paley graphs*, Experimental Mathematics, 1–28 (2024).

[18] L. Lovász, *On the shannon capacity of a graph*, IEEE Transactions on Information theory **25**, 1–7 (1979).

[19] M. Magsino, D. G. Mixon, and H. Parshall, "Linear programming bounds for cliques in paley graphs", Wavelets and sparsity xviii, Vol. 11138 (SPIE, 2019), pp. 440–447.

[20] Y. Wang, Y. Shen, and V. A. Kobzar, *Lower bounds on block-diagonal sdp relaxations for the clique number of the paley graphs and their localizations*,