Small maximal independent sets

Jeroen Schillewaert
(joint with Michael Tait and Jacques Verstraëte)

Department of Mathematics
University of Auckland
New Zealand

Table of Contents

Ramsey's theorem (for 2 colors)

Theorem (Ramsey)

There exists a least positive integer $R(r, s)$ for which every blue-red edge coloring of the complete graph on $R(r, s)$ vertices contains a blue clique on r vertices or a red clique on s vertices.

- $R(3, 3)$: least integer N for which each blue-red edge coloring on K_N contains either a red or a blue triangle.
- $R(3, 3) \leq 6$: friends and strangers.
- $R(3, 3) > 5$: Pentagon with red edges, then color "inside" edges blue.

The probabilistic method (Erdős)

- Color each edge of K_N independently with $\mathbb{P}(R) = \mathbb{P}(B) = \frac{1}{2}$.
- For $|S| = r$ vertices define $X(S) = 1$ if monochromatic, else 0.
- Number of monochromatic subgraphs is $X = \sum_{|S|=r} X(S)$.
- Linearity of expectation: $\mathbb{E}(X) = \binom{n}{r} 2^{1-\binom{r}{2}}$.
- If $\mathbb{E}(X) < 1$ then a non-monochromatic example exists, so $R(r, r) \geq 2^{r/2}$.
- Can one explicitly (pol. time algorithm in nr. of vertices) construct for some fixed $\epsilon > 0$ a 2-edge coloring of the complete graph on $N > (1 + \epsilon)^n$ vertices with no monochromatic clique of size n?

Sum free sets

- A subset S of an Abelian group is called sum-free if there are no elements a, b and c in S such that $a + b = c$.
- In $\mathbb{Z}_{3k+2}$, the set $\{k+1, k+2, \cdots, 2k+1\}$ is sum free.

Theorem (Erdős)

Every finite set B of positive integers has a sum-free subset of size more than $\frac{1}{3}|B|$.

Remark: The largest c for which every set B of positive integers has a sum-free subset of size at least $c|B|$ satisfies $\frac{1}{3} < c < \frac{12}{29}$.

Proof of the sum free set theorem

- Pick an integer $p = 3k + 2$ such that $p > 2\max(B)$.
- $I = \{k+1, \cdots, 2k+1\}$ is sum-free in $\mathbb{Z}_p$, and $|I| > \frac{|B|}{3}$.
- Choose $x \neq 0$ uniformly at random in $\mathbb{Z}_p$.
- The map $\sigma_x : b \mapsto xb$ is an injection from B into $\mathbb{Z}_p$.
- Denote $A_x = \{b \in B : \sigma_x(b) \in I\}$. (note A_x is sumfree)
- $\mathbb{P}(\sigma_x(b) \in I) = \frac{|I|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}$.
- $\mathbb{E}(|A_x|) = \sum_{b \in B} \mathbb{P}(\sigma_x(b) \in I) > \frac{|B|}{3}$.
- Hence there exists an $A^\star \subset B$ of size larger than $\frac{|B|}{3}$ which is sum free since $A_x = xA^\star$ is.

Main Result

- A d-regular graph is called δ-sparse if the number of paths of length two joining any pair of vertices is at most $d^{1-\delta}$.
- *independent set I*: no two vertices in I form an edge of the graph.

Main Result

Let $\delta, \varepsilon \in \mathbb{R}^+$ and let G be a v-vertex d-regular δ-sparse graph. If d is large enough relative to δ and ε, then G contains a maximal independent set of size at most

$$\frac{(1+\varepsilon)v \log d}{d}.$$

Table of Contents

The classical generalized quadrangles

- non-singular quadric of Witt index 2 in $\mathrm{PG}(3, q)$ ($O^+(4, q)$), $\mathrm{PG}(4, q)$ ($O(5, q)$) and $\mathrm{PG}(5, q)$ ($O^-(6, q)$).
- non-singular Hermitian variety in $\mathrm{PG}(3, q^2)$ ($U(4, q^2)$) or $\mathrm{PG}(4, q^2)$ ($U(5, q^2)$).
- Symplectic quadrangle $W(q)$, of order q ($\mathrm{Sp}(4, q)$).
- Not all GQs are classical (e.g. Tits, Kantor, Payne).

Small maximal partial ovoids in GQs

$\mathcal{Q}$	Previous range for $\gamma(\mathcal{Q})$	Theorem	Ref.
$Q^-(5,q)$	$[2q, q^2/2]$	$[2q, 3q\log q]$	[DBKMS,EH,MS]
$Q(4,q)$, q odd	$[1.419q, q^2]$	$[1.419q, 2q\log q]$	[CDWFS,DBKMS]
$H(4,q^2)$	$[q^2, q^5]$	$[q^2, 5q^2\log q]$	[MS]
$DH(4,q^2)$	$[q^3, q^5]$	$[q^3, 5q^3\log q]$	/
$H(3,q^2)$, q odd	$[q^2, 2q^2\log q]$	$[q^2, 3q^2\log q]$	[AEL,M]

- $\gamma(\mathcal{Q})$: Minimal size of maximal partial ovoid.
- *ovoid* : set of points, no two of which are collinear.
- Main theorem: any GQ of order (s,t) has a maximal partial ovoid of size roughly $s\log(st)$.

Small maximal partial ovoids in polar spaces

Q	Known prior	Range from MT	Ref.
$Q(2n, q)$, q odd	$[q, q^n]$	$[q, (2n-2)q \log q]$	[BKMS]
$Q(2n, q)$, q even	$= q + 1$		[BKMS]
$Q^+(2n+1, q)$	$[2q, q^n]$, $n \geq 3$	$[2q, (2n-1)q \log q]$	[BKMS]
$Q^-(2n+1, q)$	$[2q, \frac{1}{2}q^{n+1}]$, $n \geq 3$	$[2q, (2n-1)q \log q]$	[BKMS]
$W(2n+1, q)$	$= q + 1$		[BKMS]
$H(2n, q^2)$	$[q^2, q^{2n+1}]$, $n \geq 3$	$[q^2, (4n-3)q^2 \log q]$	[JDBKL]
$H(2n+1, q^2)$	$[q^2, q^{2n+1}]$, $n \geq 2$	$[q^2, (4n-1)q^2 \log q]$	[JDBKL]

Other examples

- Small maximal partial spreads in polar spaces.
- Maximal partial spreads in projective space $\mathrm{PG}(n, q), n \geq 3$.
- For the latter: vertices=lines, edges=intersecting lines.
- δ-sparse system with $v = q^{2n-2}$, $d = q^n$, so maximal partial spread of size $(n-2)q^{n-2} \log q$.

Problem: How to prove lower bounds?

Theorem (Weil)

Let ξ be a character of $\mathbb{F}_q$ of order s. Let $f(x)$ be a polynomial of degree d over $\mathbb{F}_q$ such that $f(x) \neq c(h(x))^s$, where $c \in \mathbb{F}_q$. Then

$$|\sum_{a \in \mathbb{F}_q} \xi(f(a))| \leq (d-1)\sqrt{q}.$$

- Gács and Szőnyi: In a Miquelian $3-(q^2+1, q+1, 1)$ design, q odd the minimal number of circles through a given point needed to block all circles is always at least or order $\frac{1}{2}\log q$ using Weil's theorem.
- This case involves estimates of quadratic character sums, becomes very/too complicated for other examples.
- Moreover many problems do not have an algebraic description.

Table of Contents

A technical condition for GQs

A GQ of order (s, t) is called *locally sparse* if for any set of three points, the number of points collinear with all three points is at most $s + 1$.

- Any GQ of order (s, s^2) is locally sparse (Bose-Shrikhande, Cameron)
- In particular, $Q^-(5, q)$ is locally sparse.
- $H(4, q^2)$ is **not** locally sparse.

A weaker theorem for GQs

Theorem

For any $\alpha > 4$*, there exists* $s_0(\alpha)$ *such that if* $s \geq s_0(\alpha)$ *and* $t \geq s(\log s)^{2\alpha}$*, then any locally sparse generalized quadrangle of order* (s, t) *has a maximal partial ovoid of size at most* $s(\log s)^{\alpha}$*.*

First round

- Fix a point $x \in \mathcal{P}$ and for each line l through x independently flip a coin with heads probability $p_s = \frac{s \log t - \alpha s \log \log s}{t}$, where $\alpha > 4$.
- On each line l where the coin turned up heads, select uniformly a point of $l \setminus \{x\}$ and denote the set of selected points by S.
- $U = \mathcal{P} \setminus (S \cup \{x\})^{\bowtie}$ (uncovered points not collinear with x).

Second round

Let $x^\star \in x^\perp \setminus S^{\bowtie}$. On each line $l \in \mathcal{L}$ through $x^\star$ with $l \cap U \neq \emptyset$, uniformly and randomly select a point of $l \cap U$. Moreover select a point x^+ on the line M through $x^\star$ and x different from x, and call this set of selected points T. Then clearly $S \cup T \cup \{x^+\}$ is a partial ovoid. So we will need to show that $S \cup T \cup \{x^+\}$ is maximal, and small.

A form of the Chernoff bound

A sum of independent random variables is concentrated according to the so-called Chernoff Bound. We shall use the Chernoff Bound in the following form. We write $X \sim \text{Bin}(n, p)$ to denote a binomial random variable with probability p over n trials.

Proposition

Let $X \sim Bin(n, p)$. Then for $\delta \in [0, 1]$,

$$\mathbb{P}(|X - pn| \geq \delta pn) \leq 2e^{-\delta^2 pn/2}.$$

Proof for GQs i

First we show $|S| \lesssim s \log t$ using the Chernoff Bound. There are $t+1$ lines through x, and we independently selected each line with probability ps and then one point on each selected line. So $|S| \sim \text{Bin}(t+1, ps)$ and $\mathbb{E}(|S|) = ps(t+1) \sim s \log t$. By Chernoff, for any $\delta > 0$,

$$\mathbb{P}(|S| \geq (1+\delta)s \log t) \leq 2\exp(-\tfrac{1}{2}\delta^2 s \log t) \to 0.$$

Therefore a.a.s. $|S| \lesssim s \log t$.

Three key properties

We can show that in selecting S, Properties I – III described below occur simultaneously a.a.s. as $s \to \infty$:

I. *For all lines* $\ell \in \mathcal{L}$ *disjoint from* x, $|\ell \cap U| < \lceil \log s \rceil$.
II. *For all* $u \in x^{\perp} \backslash S$, $|u^{\perp} \cap U| \lesssim s(\log s)^{\alpha}$
III. *For* $v, w \notin S \cup \{x\}; v \not\sim w$, $|\{v, w\}^{\perp} \cap U| \gtrsim (\log s)^{\alpha}$.

Proof for GQs ii

Assuming that a.a.s., S satisfies Properties I – III, we fix an instance of such a partial ovoid S with $|S| \lesssim s \log t$ and let T be as before. By Property II, $|T| \leq X_{x^\star} \lesssim s(\log s)^\alpha$. Therefore

$$|S \cup T| \leq |S| + X_{x^\star} + 1 \lesssim s \log t + s(\log s)^\alpha \lesssim s(\log s)^\alpha$$

Proof for GQs iii

For $v \in (x^\perp \backslash S^{\bowtie}) \cup U$ not collinear with $x^\star$, a.a.s., $X_{vx^\star} \geq \frac{1}{2}(\log s)^\alpha$ by Property III. By Property I, the probability that v is not collinear with any point in T is at most

$$\left(\frac{\log s - 1}{\log s}\right)^{X_{vx^\star}} \leq \left(1 - \frac{1}{\log s}\right)^{\frac{1}{2}(\log s)^\alpha} \leq e^{-\frac{1}{2}(\log s)^3} < \frac{1}{s^5}$$

since $\alpha > 4$. Hence the expected number of points in $(x^\perp \backslash S^{\bowtie}) \cup U$ not collinear with any point in T is at most

$$s^{-5}|\mathcal{P}| \lesssim \frac{1}{s}.$$

It follows that a.a.s. T covers all points not yet covered by S except those on the line $xx^\star$. So $S \cup T \cup \{x^+\}$ is a maximal partial ovoid.

Definition of Random variables I

For $u \in x_\circ^\perp$, let $U(u)$ denote the set of points in $\mathcal{P} \backslash x^\perp$ which are not covered by $S \backslash \{u\}$, and define the random variable:

$$X_u = |u^\perp \cap U(u)|.$$

In the case $u \in x^\perp \backslash S$, note that $U(u) = U$, so then $X_u = |u^\perp \cap U|$.

Definition of Random variables II

For $v, w \in \mathcal{P}\backslash\{x\}$ non-collinear, let $U(v, w)$ denote the set of points in $\mathcal{P}\backslash x^{\perp}$ which are not covered by $S\backslash\{v, w\}$, and define the random variable:

$$X_{vw} = |\{v, w\}^{\perp}_{\circ} \cap U(v, w)|.$$

In the case $v, w \notin S \cup \{x\}$, $U(v, w) = U$ and so $X_{vw} = |\{v, w\}^{\perp}_{\circ} \cap U|$.

Expected values

Lemma

Let $u \in x_\circ^\perp$, *and let* $v, w \in \mathcal{P} \backslash \{x\}$ *be a pair of non-collinear points. Then*

$$\mathbb{E}(X_u) \sim s(\log s)^\alpha \quad \textit{and} \quad \mathbb{E}(X_{vw}) \sim (\log s)^\alpha.$$

In addition, if $j \in \mathbb{N}$ *and* $jtp^2 \to 0$ *as* $s \to \infty$, *then* $\mathbb{E}(X_u)^j \sim s^j (\log s)^{\alpha j}$.

Proof of property I-i

Fix a line $\ell \in \mathcal{L}$ disjoint from x, and let Y_ℓ be the number of sequences of $a = \lceil \log s \rceil$ distinct points in $U \cap (\ell \backslash x^\perp)$. Let $R \subset \ell \backslash x^\perp$ be a set of a distinct points. Then

$$\left| \bigcup_{y \in R} \{x, y\}_\circ^\perp \right| = at + 1$$

and hence

$$\mathbb{E}(Y_\ell) = s(s-1)(s-2)\dots(s-a+1) \cdot (1-p)^{at+1}.$$

Since $atp^2 \to 0$ and $a^2/s \to 0$, we obtain

$$\mathbb{E}(Y_\ell) \sim \frac{s^a (\log s)^{a\alpha}}{t^a}.$$

Proof of property I-ii

Let $A_s = \bigcup_{\substack{\ell \in \mathcal{L} \\ x \notin \ell}} [Y_\ell \geq 1]$. Since $|\mathcal{L}| = (t+1)(st+1) \sim st^2$ is the total number of lines,

$$\mathbb{P}(A_s) \leq \sum_{\substack{\ell \in \mathcal{L} \\ x \notin \ell}} \mathbb{P}(Y_\ell \geq 1) \lesssim st^2 \cdot \mathbb{E}(Y_\ell) \sim \frac{s^{a+1}(\log s)^{a\alpha}}{t^{a-2}}.$$

Since $t \geq s(\log s)^{2\alpha}$ and $a = \lceil \log s \rceil$, $\mathbb{P}(A_s) \to 0$ as $s \to \infty$, as required for Property I.

Practical implementation

The randomized algorithm could be implemented, and we believe it is effective in finding maximal partial ovoids even in (s, t)-quadrangles where s is not too large. In addition, it can be deduced from the proof that the probability that the algorithm does not return a maximal partial ovoid of size at most $s(\log s)^{\alpha}, \alpha > 4$, is at most $s^{-\log s}$ if s is large enough.

Table of Contents

Set systems

- $X = X(S)$ is a set of *atoms*.
- Set system $\mathcal{S}$: family of subsets of X referred to as *blocks*.
- $\mathcal{S}$ is an (n, d, r)-*system* if $|X| = n$, every atom is contained in d blocks, every block contains r atoms.
- A *maximal independent set* in a set system $\mathcal{S}$ is a set I of atoms containing no block but such that the addition of any atom to I results in a set containing some block of $\mathcal{S}$.
- General problem: find the smallest possible size $\gamma_0(S)$ of a maximal independent set in $\mathcal{S}$.

Segre's Problem I.

- What is the smallest possible size for a complete arc in a projective plane?
- $\mathcal{S}$: family of triples of collinear points in the plane; the atoms are the points of the projective plane.
- Kim-Vu: There are positive constants c and M such that the following holds. In every projective plane of order $q \geq M$, there is a complete arc of size at most $q^{1/2} \log^c q (c = 300)$.
- If the plane has order q, then $\mathcal{S}$ is an (n, d, r)-system with $n = q^2 + q + 1$, $r = 3$ and $d = (q+1)\binom{q}{2}$.

Sparseness conditions

Let $\delta > 0$, $n \geq r \geq 2$ and $d \geq 1$. An (n, d, r)-system $\mathcal{S}$ is δ**-sparse** if

- δ**-1** for $x, y \in X(\mathcal{S})$, the number of pairs of blocks $e, f \in \mathcal{S}$ such that $x \in e$, $y \in f$, $(e \setminus \{x\}) = (f \setminus \{y\})$ is at most $d^{1-\delta}$.
- δ**-2** for $a \in [2, r-1]$ and any $A \subseteq X(\mathcal{S})$ with $|A| = a$, the number of blocks in $\mathcal{S}$ containing A is at most $d^{(r-a)/(r-1)-\delta}$.

Main result Note: Work in progress

Theorem

For all $\delta > 0$ and $r \geq 2$, there exist a constants $c_1(r,\delta), c_2(r,\delta) > 0$ such that for any δ-sparse (n,d,r)-system $\mathcal{S}$, there exists a maximal independent set $I \subseteq X(\mathcal{S})$ such that

$$c_1(r,\delta)n\left(\frac{\log d}{d}\right)^{1/(r-1)} \leq |I| \leq c_2(r,\delta)n\frac{\log d}{d^{1/(r-1)}}.$$

- Bohman-Bennett; randomized greedy algorithm.
- Our approach is iterative greedy using the Lovàsz local lemma
- In fact, we prove a result on (ϵ,δ)-sparse systems.

Segre's problem II.

- $\gamma_\circ(\mathcal{S})$ is roughly at most $\sqrt{3q}\log q$ if q is large enough.
- Best lower bound is roughly $2\sqrt{q}$, by Lunelli and Sce.
- Computational evidence by Fisher that the average size of a complete arc in $PG(2, q)$ is close to $\sqrt{3q\log q}$.
- Main open problem: find lower bounds; in particular does every complete arc have size at least $\sqrt{q}\omega(q)$ for some unbounded function $\omega(q)$.