Asymptotics of $r(4, t)$

Talk by Dingding Dong

September 29, 2023

(This paper is by Mattheus and Verstraete.)

§1 Introduction

Definition 1.1. $r(4, t)$ is the smallest number r such that if we color the edges of K_r with red and blue, then there is either a red K_4 or a blue K_t.

The main theorem we'll prove today is the following lower bound:

Theorem 1.2

There exists $c > 0$ such that

$$r(4, t) \geq \frac{ct^3}{\log^4 t}.$$

Remark 1.3. There's also an upper bound (Shearer 1995) — there exists c' such that

$$r(4, t) \leq \frac{c' t^3}{\log^2 t}.$$

Combining the upper bound and the lower bound we'll see today, we know the asymptotic exponent of $r(4, t)$.

Remark 1.4. One can get an exponent of t^3 (forgetting about log factors) just from the Erdős–Szekerés bound of

$$r(s, t) \leq \binom{s + t - 2}{t - 1}$$

(here $4 - 1 = 3$).

Remark 1.5. The previous best lower bound was roughly $t^{5/2}$.

§1.1 Proof Sketch

We'll take q to be a prime, or prime power. Our overall goal is to find some graph G with $|V(G)| = \Theta(q^3 \log^2 q)$, such that:

- G does not contain K_4.

- $\alpha(G) < \Theta(q \log^2 q)$.

Then G does not contain a 4-clique, and it doesn't contain an independent set of size $q \log^2 q$, providing a counterexample graph for this value of r.

We'll gradually construct this G using a few steps.

(1) The first step is to use finite fields to construct a graph H with the following properties:
   - Its number of vertices is $n = q^4 - q^3 + q^2$, and it is d-regular with $d = (q+1)(q^2-1)$. (In the end, the small terms won't matter.)
   - There exists q^3+1 maximal q^2-cliques that pairwise intersect at exactly 1 vertex. (We will use this property to build our eventual graph G.) (These maximal cliques should partition the edge-set of the graph.)
   - If there exists some $K_4 \subseteq H$, then at least three of the vertices in this K_4 belong to the same maximal clique (in the previous point).

(2) The next step is to deduce from these properties that for all vertex subsets $X \subseteq V(H)$ of size $|X| = 2^{24}q^2$ (let $m = 2^{24}q^2$ be this quantity), X intersects 'many' cliques only in a small number of edges. This is kind of because of the abundance of maximal cliques.

(3) The third step is to construct a random subgraph $H^* \subseteq H$ by doing the following: for every maximal clique C, we take a random partition of its vertex set $V(C)$ into two random parts $A \cup B$. We then include the resulting bipartite graph $K(A_C, B_C)$ into H^*. (So we randomly partition our clique into two parts, and put all edges between them into H^*.)

   After this step, there is no K_4 in H^* — for every K_4 in H, we have three vertices in the same clique, creating a triangle. But after we do this, within every clique there will be no triangle; so no matter how we partition in this step, there won't be any K_4 anymore in H^*.

   We then want to find a good G out of this H^*. We use the properties in (2) plus Azuma's inequality to show that with positive probability, for all vertex sets X of size $|X| = m$, we have

$$e(H^*[X]) \geq 2^{24}q^3.$$

   Why do we expect this? The edge density of our big graph H is roughly $\frac{1}{q}$, and now X has size around q^2, so we expect there to be q^3 edges in H^* induced on X. So this is essentially a concentration inequality. And now because X intersects many cliques in a small number of edges, we can use martingale ideas to deduce that the parts every vertex in the partition belong to don't matter that much, when we gradually reveal these parts. We're able to do that by using (2).

   We then use this to deduce that for all $|Y| \geq m$, we have

$$e(H^*[Y]) \geq \frac{|Y|^2}{256q}$$

   (using a purely averaging argument).

(4) We then use some graph container lemma along with (3) to bound the number of independent sets of size $t := 2^{30} q \log^2 q$ in H^*. (The bound on $e(H^*[Y])$ leads to a good upper bound on the number of large independent sets.)

(5) Finally, we sample a random subset $W \subseteq V(H^*)$, and show that with positive probability, there is no independent set of size t in the induced subgraph $H^*[W]$, and that $|W| = \Theta(q^3 \log^2 q)$.

The main work occurs in (1) and (2); the rest of the argument involves fairly standard things.

§2 Defining the graph H (Step 1)

Before defining H, let's recall the following definition:

Definition 2.1. The *projective geometry* $\mathbb{PG}(2,q)$ is:

- a collection of 'points' — which are actually lines through the origin in $\mathbb{F}_{q^2}^3$ (the 3-dimensional space over $\mathbb{F}_{q^2}$), or in other words
$$\{\langle x\rangle \mid x \in \mathbb{F}_{q^2}^3 \setminus \{0\}\}.$$
This has size q^4+q^2+1.
- A collection of 'lines,' which are actually planes through the origin in $\mathbb{F}_{q^2}^3$ (each pair of lines forms a plane), in other words
$$\{x^\perp \mid x \in \text{points}\}$$
(each plane is orthogonal to a unique line). This also has size q^4+q^2+1.
- There is an incidence structure between points and lines (a point is contained in a line if it actually belongs to the line — i.e., the line lies (fully) in the plane). In other words, $x \sim y^\perp$ if $\langle x, y\rangle = 0$.

Definition 2.2. The *Hermitian unital* $\mathcal{H}$ is defined as
$$\{\langle (a,b,c)\rangle \in \text{points} \mid a^{q+1}+b^{q+1}+c^{q+1}=0\}.$$

(So $\mathcal{H}$ is a collection of points, where each $\langle (a,b,c)\rangle$ is a point (i.e., line through the origin in $\mathbb{F}_{q^2}^3$).)

Here it's important that we're working over $\mathbb{F}_{q^2}$ — this doesn't exist over a prime field.

Remark 2.3. If you think of $\mathbb{F}_{q^2}$ as $\mathbb{C}$, you can consider in the complex plane the set of points (z,w) with $|z|^2+|w|^2=1$, which is topologically a 3-dimensional sphere; this is a finite-field analog.

Fact 2.4 — We have the following observations about $\mathcal{H}$:

- We have $|\mathcal{H}| = q^3+1$.
- Every line in $\mathbb{PG}(2,q^2)$ intersects $\mathcal{H}$ at either 1 or $q+1$ points.

Proof sketch. The proof is a combination of direct computation and the fact that there are exactly $q+1$ elements in $\mathbb{F}_{q^2}$ which satisfy $x^{q+1}=1$. We also know $x \mapsto x^{q+1}$ is a norm function in $\mathbb{F}_{q^2}$. Then we use some symmetry of $\mathcal{H}$ in $\mathbb{PGU}(3,q^2)$ — because $\mathcal{H}$ is symmetric under this object, to check the second condition it suffices to check whether the x used to define our line (as $x^\perp$) lies in the Hermitian unital or not, which allows us to do direct computation by choosing two different points. $\square$

Remark 2.5. A *norm map* is $\mathcal{N}(a) = a\overline{a}$. In complex numbers $\overline{a}$ is the conjugate; but here the Galois conjugate is a^q (in $\mathbb{F}_{q^2}$). So the equation for the Hermitian unital is
$$\mathcal{H} = \{a\overline{a} + b\overline{b} + c\overline{c} = 1\}.$$
This is the analog of a sphere. This is also why we call it Hermitian; and it lets you derive the symmetry under $\mathbb{PGU}$. But this is why it's no longer some completely magical algebra (it's still somewhat magical that it translates to finite fields, but at least you are more willing to believe it).

Definition 2.6. We define *secants* to be those lines that intersect $\mathcal{H}$ at $q+1$ points.

(The name may have to do with the sphere understanding.)

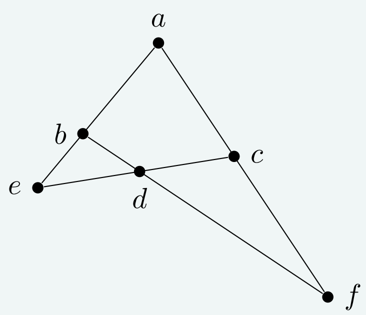
We will come back to this definition later.

Definition 2.7. Define Q to be the graph $\mathbb{PG}(2,q^2)[\mathcal{H}]$.

(We can understand the incidence structure as a graph, and we restrict it to $\mathcal{H}$.)

Proposition 2.8

Q does not contain the following structure (with $a,b,c,d,e,f \in \mathcal{H}$).



(In this configuration, we assume that e.g. a, b, and c are not on the same line.)

Remark 2.9. This might also have a complex analog.

Proof. Recall that $\mathcal{H}$ is the collection of points $\langle x\rangle$ such that $x\cdot\overline{x}=0$ — i.e.,

$$\mathcal{H}=\{\langle x\rangle \mid x\cdot\overline{x}=0\}$$

(as we said earlier). Now assume for contradiction that the above structure exists.

Then because of the collinear structure, we can find scalars $\alpha,\beta,\gamma,\delta\in\mathbb{F}_{q^2}$ such that

$$\begin{cases} d=a+\alpha b\\ e=a+\beta c\\ d=f+\gamma c\\ e=f+\delta b.\end{cases}$$

(Because we're in projective space, we don't care about scaling, so we don't need more scalars for the first terms.) So we can deduce that

$$\begin{cases} a+\alpha b=f+\gamma c\\ a+\beta c=f+\delta b.\end{cases}$$

This gives

$$(\alpha+\delta)b=(\beta+\gamma)c.$$

But because b and c are distinct points in the projective plane, they cannot be scalar multiples of each other; this means we can conclude $\alpha+\delta=\beta+\gamma=0$.

Now up to replacing all the vectors by their scalar multiples, without loss of generality we can let $\alpha=\beta=1$ and $\delta=\gamma=-1$.

Then consider the following two matrices: let

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \text{ and } B = \begin{bmatrix} \overline{a_1} & \overline{b_1} & \overline{c_1} \\ \overline{a_2} & \overline{b_2} & \overline{c_2} \\ \overline{a_3} & \overline{b_3} & \overline{c_3} \end{bmatrix}$$

(where (a_1, a_2, a_3) are the coordinates of the vector a, and so on). Because of linear independence, both A and B are nonsingular (because a, b, and c are not collinear). But

$$AB = \begin{bmatrix} 0 & a\overline{b} & a\overline{c} \\ b\overline{a} & 0 & b\overline{c} \\ c\overline{a} & c\overline{b} & 0 \end{bmatrix},$$

which ends up giving

$$\det(AB) = (a\overline{b})(b\overline{c})(c\overline{a}) + (a\overline{c})(b\overline{a})(c\overline{b}).$$

(We consider all the 'diagonals' and these are the only ones that don't have 0's.)

(In the matrix, $a\overline{b}$ is actually a dot product $a \cdot \overline{b}$.)

But on the other hand we know $d = a + b$ and $d\overline{d} = 0$; that tells us

$$a \cdot \overline{b} + b \cdot \overline{a} = 0.$$

Similarly $e = a = c$ and $e\overline{e} = 0$, which tells us that

$$a \cdot \overline{c} + c \cdot \overline{a} = 0.$$

Finally, $f = c + d = a + b + c$, and we know $f\overline{f} = 0$; combining these, we can conclude that

$$b \cdot \overline{c} + c \cdot \overline{b} = 0.$$

Then the terms in the second factor in $\det(AB)$ are all negations of each other, so we get

$$\det(AB) = 0,$$

which is a contradiction. □

We now construct H in the following way:

Definition 2.10. Let $H = H_q$ be the graph whose vertex set is the set of all the secants in this projective plane — i.e.,

$$V(H) = \{\text{secants}\} = \{x^\perp \mid \left|x^\perp \cap \mathcal{H}\right| = q + 1\},$$

and whose edges are all the pairs of secants whose intersection belongs to $\mathcal{H}$ — i.e.,

$$E(H_q) = \{(x^\perp, y^\perp) \mid x^\perp \cap y^\perp \in \mathcal{H}\}.$$

We need to check that H has the desired properties.

First we check that $|V(H)|$ is correct; let n be the number of secants. Every point in $\mathbb{PG}(2, q^2)$ is contained in $q^2 + 1$ lines. So by a double-counting argument (counting all pairs of a vertex in the Hermitian unital and lines containing that point), we get

$$(q^3 + 1)(q^2 + 1) = n(q + 1) + (q^4 + q^2 + 1 - n),$$

which gives that $n = q^4 - q^3 + q^2$.

(Here $|V(H_q)| = |\{(x, y^\perp) \mid x \in \mathcal{H}, x \in y^\perp\}|$; then q^3+1 is the size of $\mathcal{H}$, and every point in $\mathcal{H}$ is counted in q^2+1 lines; on the right-hand side, we choose the line first and say, either it's a secant or it isn't a secant (every line has either 1 or $q+1$ intersections).)

Next we check that it's d-regular. We know every point in the projective geometry is contained in

$$\frac{n(q+1)}{q^3+1} = q^2$$

secants. (The top is the number of points contained in every secant, and the bottom is the size of the Hermitian unital.)

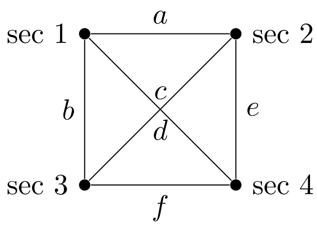
Finally, for every secant $y^\perp$ and every point $x \in y^\perp$, there are q^2-1 other secants containing x. This tells us that $\deg(y^\perp)$ in our graph H is $(q+1)(q^2-1)$. (For every secant and every point on it, we count the number of other secants that go through that point; that's the total degree of this $y^\perp$ in H.)

Finally, what's the collection of cliques? We define

$$\mathcal{C} = \{K_x := \{\text{secants through } x\} \mid x \in \mathcal{H}\}$$

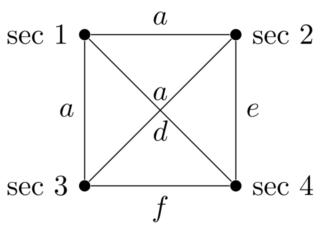
(we name our cliques K_x, where each K_x consists of all secants through x, and we take these cliques for all $x \in \mathcal{H}$). Because every point is contained in q^2 secants, we know that all these cliques have size q^2. And now because $\mathcal{H}$ has size q^3+1, this is the number of maximal such cliques. Finally, for two cliques (each represented by a point), there's only one secant that goes through both points; so these cliques only intersect in one point.

Finally, we want to check why the fourth property holds. Suppose that there is some $K_4 \subseteq H$, and consider our four secants.



We have all six edges between these four secants, where each edge represents a point. But these points cannot be distinct, or else we'd have the forbidden structure.

So either we must have a triangle all consisting of the same label, or all the edges consist of the same label; this gives the desired property.



§3 The clique structure of H_q (Step 2)

The rest of the proof doesn't use anything other than the properties defined in step (1) — starting from here, there's no finite fields happening.

Definition 3.1. For all $X \subseteq V(H)$ with $|X| = k$, consider collection of intersections of the vertex set X with different cliques in our collection, i.e., $\{X \cap C \mid C \in \mathcal{C}\}$. We first exclude intersections of size 1. Then we can partition the remaining intersections into 'small' intersections, 'medium' intersections, and 'large' intersections:

- $S = \{|X \cap C| \mid 2 \leq |X \cap C| \leq \sqrt{2k}/\log n\}$
- $M = \{|X \cap C| \mid \sqrt{2k}/\log n \leq |X \cap C| \leq \sqrt{2k}\}$.
- $L = \{|X \cap C| \mid |X \cap C| \geq \sqrt{2k}\}$.

So we have a bunch of cliques, and we intersect them with our set X. We take all the cliques X intersects at more than one points (if X intersects a clique at just one point we ignore it) and classify them into S, M, and L.

Definition 3.2. For every $U \subseteq \{S, M, L\}$, we define

$$v(U) = \sum_{T \in U} |T| \text{ and } e(U) = \sum_{T \in U} \binom{|T|}{2}.$$

Note that sometimes we count one point twice when computing $v(U)$, since we could have a point in two cliques.

Theorem 3.3

If $k = 2^{24}q^2 = m$, then either

$$e(S) \geq \frac{k^2}{64q} \text{ or } e(M) \geq \frac{qk^{3/2}}{16\log^2 n}.$$

So either X intersects small cliques with a lot of edges, or it intersects medium cliques with a large number of edges.

Proof sketch. First note that $v(S) + v(M) + v(C)$ is big — we have

$$v(S) + v(M) + v(C) \geq (q + 1)k - q^3 - 1,$$

since you can't expect X to intersect many cliques at just one vertex (so it'll intersect most cliques in at least 2).

Then we upper-bound vertices involved in large cliques — we prove

$$v(L) \leq 2k$$

by a double-counting argument.

Then there are two cases. If $v(S) \geq v(M)$, then we prove the first bound, again by some double-counting argument. If $v(M) \geq v(S)$, then we prove the second.

Here's the proof in the first case: we know that

$$e(S) = \sum_{T \in S} \binom{|T|}{2} \geq |S| \binom{v(S)/|S|}{2} \geq \frac{v(S)^2}{4|S|}$$

(by convexity). But we know that $|S| \leq q^3 + 1$; so we ca lower-bound this by

$$\frac{v(S)^2}{q^3} \geq \frac{k^3}{64q}$$

and deduce some lower bound on the number of edges involved in small cliques. (Basically, what happens is that the number of vertices in large cliques cannot be large, and then we can deduce the lower bounds on edge numbers by these averaging arguments.) $\square$

§4 The random partition (Step 3)

We now take our random partition and define the martingale we consider. We lower bound the number of edges by using intersections with small cliques or with big cliques. If

$$e(S) \geq \frac{m^2}{64q},$$

then we let $\{v_1, \dots, v_\ell\}$ be all vertices contained in some clique in S. Now we let

$$Z = \mathbb{E}[\sum_{T \in S} e[H^*(A_C, B_C)]],$$

and for each i, let

$$Z_i = \mathbb{E}[Z \mid \text{knowledge of which parts } v_1,\, \dots,\, v_{i-1} \text{ belong to}].$$

Then all the Z_i's form a martingale, and we know that

$$|Z_i - Z_{i-1}| \leq |T| - 1,$$

where $T \in S$ is the clique that v_i belongs to. Let this quantity $|T| - 1$ be c_i.

Finally, we want to bound $\sum c_i^2$; we have

$$\sum c_i^2 \leq \sum_{T \in S} (|T| - 1)^2 \, |T|$$

(for every clique T, there are at most $|T|$ vertices in that clique, and then we sum $(|T| - 1)^2$), and we can apply concentration to give an upper bound on this in terms of $e(S)$.

Remark 4.1. Why do we need Azuma — doesn't this follow from Chernoff, because every clique is independent? In the M case this might be true, but in the S case you have bigger range; but you *do* have a sum of independent variables.

Remark 4.2. You have $k \asymp q^2$, so $e(S) \asymp q^3$ and $e(M) \asymp q^4/\log^2 n$. This is kind of critical. Why are the powers different? There are two cases.

In $e(M)$, you just have so many edges that you're winning — the density is $1/q$, so $e(S)$ is a true density set, and you're winning the log so that you can do a union bound. In $e(M)$ you win a q in the density so you can lose these log factors. This is important.

Here we assume $|S| \leq \sqrt{2k}/\log n$; this lets you run the union bound in the $e(S)$ case. In the $e(M)$ you win a factor of q in the density, and that's why you don't need the log. If you don't save this log, then you don't get good control — otherwise the union bound doesn't go through.

§5 The container theorem and finish (Step 4 and 5)

Here's the container theorem we'll apply:

Theorem 5.1

Let G be a graph such that:

(1) For all vertex subsets X with $|X| \geq R$, we have $2e(G[X]) \geq \alpha |X|^2$.

(2) $e^{-\alpha r} n \leq R$, and $t \geq r$.

Then the number of independent sets of size t in G is

$$i_t(G) \leq \binom{n}{r}\binom{R+r}{t-r}.$$

In our graph H^*, we have $n \approx q^4$, and we know that for every $|Y| \geq m = 2^{24}q^2$, we have

$$e(H^*[Y]) \geq \frac{|Y|^2}{256q}.$$

Then we take $R = 2^{24}q^2$, $r = 2^{10}q\log q$, and $\alpha = \frac{1}{2^8 q}$, and $t = 2^{30}q\log^2 q$. From here we can conclude that

$$i_t(H^*) \leq \binom{n}{r}\binom{R+r}{t-r} \leq \left(\frac{q}{\log^2 q}\right)^t.$$

Finally, we take a $\frac{\log^2 q}{q}$-vertex random subset $W \subseteq V(H^*)$. We can check using expectation that with positive probability, we have

$$|W| \geq \frac{q^3 \log^2 q}{2} \text{ and } \alpha(W) < t = 2^{30} q \log^2 q.$$

Remark 5.2. For background, there was an earlier paper that had many ideas except the finite field construction — if you don't have too many independent sets, by random sampling you can destroy them. So if you have some nice graph with nice properties, that implies bounds on Ramsey numbers. (3), (4), and (5) are identical to that paper except that they assume H is a good expander; here you just have a lower bound on the number of edges.

Then how do you get this nice graph? This is the point of the first part. There the authors had some idea that it'd have something to do with finite geometries; his coauthor was a grad student or postdoc who was an expert in finite geometries.

The way you guess this graph is magic is that it has no large negative eigenvalues. You can compute its spectrum; its least eigenvalue is $-q$. And that means the spectral bound gives you all sets of size larger than q^2 have at least this density. (This is something about expander mixing.) All the random sampling says you preserve this bound.

The splitting of cliques into halves is not in the Mubayi–Verstraete paper; but the same trick comes in in another paper.