Asymptotics of $r(4, t)$

Talk by Dingding Dong

Notes by Sanjana Das

September 29, 2023

This is based on a paper by Mattheus and Verstraete.

§1 Introduction

Definition 1.1. The *Ramsey number* $r(4, t)$ is the smallest integer r such that if we color the edges of K_r with red and blue in any way, 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 we have

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

In 1995, Shearer proved the following *upper bound* on $r(4, t)$ — there exists $c' > 0$ such that

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

Combining these two bounds, we now know that the 'asymptotic exponent' of $r(4, t)$ is 3. (Before, the previous best lower bound was roughly $t^{5/2}$.)

Remark 1.3. If we just want an upper bound with an exponent of t^3, then the Erdős–Szekerés bound

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

suffices (this gives an upper bound of t^3 without the log factors).

§1.1 Proof Outline

We'll start by outlining the proof. Let q be a prime. Our goal is to find a graph G with $|V(G)| = \Theta(q^3 \log^2 q)$ such that G does not contain K_4 and $\alpha(G) < \Theta(q \log^2 q)$. Then G doesn't contain a clique of size 4 or an independent set of size $t \asymp q \log^2 q$, which gives the desired bound on $r(4, t)$.

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

(1) First, we'll use finite fields to construct a graph H with the following properties.

- It has $n = q^4 - q^3 + q^2$ vertices, and is d-regular with $d = (q+1)(q^2-1)$. (Only the leading terms of these expressions are important for the analysis.)
- We can find a collection $\mathcal{C}$ of $q^3 + 1$ maximal q^2-cliques in H, such that each pair intersects at exactly one vertex. (We'll use this property to eventually build G. In fact, these maximal cliques will partition the edge-set of H, though we may not need this.)
- For each K_4 contained in H, one of the maximal q^2-cliques in $\mathcal{C}$ contains at least three of its four vertices.

(2) Let $k = 2^{24}q^2$. In the next step, we use these properties to deduce that for each vertex subset $X \subseteq V(H)$ of size $|X| = k$, roughly speaking, X intersects many cliques only in a small number of edges. (This sort of happens because of the abundance of maximal cliques.)

(3) We then construct a random subgraph $H^* \subseteq H$ in the following way: for each of our maximal cliques $C \in \mathcal{C}$, we randomly partition its vertex set $V(C)$ into two sets A and B, and then include the complete bipartite graph $K(A, B)$ in H^*. (In other words, we randomly partition our clique into two parts, and place all edges between these two parts in H^*.)

This ensures that there is no K_4 in H^* — for every K_4 in H, there is a clique $C \in \mathcal{C}$ containing three of its vertices (which form a triangle inside this clique). But when we build H^*, no matter how we partition C, we'll destroy all its triangles; this means H^* can't contain any K_4. (We'll eventually build G out of H^*.)

We then use the property in (2) together with Azuma's inequality to show that with positive probability, for all vertex sets X of size k, we have

$$e(H^*[X]) \geq 2^{40}q^3.$$

As some intuition for why we might expect this, the edge density of H is on the order of $1/q$ (so the edge density of H^* is as well), and X has size on the order of q^3, so in *expectation* $H^*[X]$ should have about q^3 edges. This means it suffices to prove a concentration inequality — that $H^*[X]$ is usually not too far below its expectation. And we can do this using martingales, because of (2) — we can imagine gradually revealing what H^* is by revealing which of A and B each vertex is in for every clique. Because X intersects many cliques in a small number of edges, the information we learn at each step doesn't affect the expectation of $H^*[X]$ too much, allowing us to use martingale concentration inequalities.

Then using an averaging argument, we can deduce that for all vertex sets Y of size $|Y| \geq k$, we have

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

(4) We then use a graph container lemma together with (3) (specifically, the bound on $e(H^*[Y])$) to upper-bound the number of independent sets of size $t = 2^{30}q\log^2 q$ in H^*.

(5) Finally, we sample a random subset $W \subseteq V(H^*)$ and show that with positive probability, we have $|W| = \Theta(q^3 \log^2 q)$ and there is no independent set of size t in $H^*[W]$. We then take G to be $H^*[W]$, and we're done.

§2 Step (1) — Defining H

§2.1 Some definitions

Before we define H, we need a few definitions.

Definition 2.1. The *projective geometry* $\mathbb{PG}(2,q^2)$ consists of a collection of points $\mathcal{P}$ and lines $\mathcal{L}$, together with an incidence structure between points and lines, defined in the following way:

- The 'points' are lines through the origin in $\mathbb{F}_{q^2}^3$ (the 3-dimensional vector space over $\mathbb{F}_{q^2}$) — in other words,
$$|\mathcal{P}| = \{\langle x\rangle \mid x \in \mathbb{F}_{q^2}^3 \setminus \{0\}\}$$
(where $\langle x\rangle$ denotes the line in $\mathbb{F}_{q^2}^3$ spanned by x).
- The 'lines' are planes through the origin in $\mathbb{F}_{q^2}^3$ — in other words,
$$\mathcal{L} = \{x^\perp \mid x \in \mathcal{P}\}$$
(since each plane in $\mathbb{F}_{q^2}^3$ is orthogonal to a unique line).
- We say a point is contained in a line if the corresponding line is contained in the corresponding plane in $\mathbb{F}_{q^2}^3$ — in other words, for $x, y \in \mathcal{P}$, we say $x \sim y^\perp$ if $\langle x, y\rangle = 0$.

Note that $|\mathcal{P}| = |\mathcal{L}| = q^4 + q^2 + 1$.

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

(Here each (a,b,c) is an element of $\mathbb{F}_{q^2}^3$, and $\langle (a,b,c)\rangle$ is a point — i.e., a line through the origin in $\mathbb{F}_{q^2}^3$.) It's important that we're working over $\mathbb{F}_{q^2}$ and not $\mathbb{F}_q$ — such an object doesn't exist for a field of prime order.

We have the following observations about $\mathcal{H}$.

Fact 2.3 — We have $|\mathcal{H}| = q^3 + 1$, and 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 in $\mathbb{F}_{q^2}$ there are exactly $q+1$ solutions to $x^{q+1} = 1$, and the fact that the map $x \mapsto x^{q+1}$ is a norm function in $\mathbb{F}_{q^2}$. We then use the symmetry of $\mathcal{H}$ under $\mathbb{PGU}(3,q^2)$ — because $\mathcal{H}$ is symmetric under this object, to check the second statement it suffices to check whether the point $x \in \mathcal{P}$ used to define our line (as $x^\perp$) lies in $\mathcal{H}$ or not; this allows us to do direct computation by choosing two different points. □

Remark 2.4. In general, a *norm map* is a map of the form $\mathcal{N}(a) = a\overline{a}$. If we were working over $\mathbb{C}$, then $\overline{a}$ would represent the complex conjugate of a; over $\mathbb{F}_{q^2}$ it represents the *Galois conjugate* a^q. This means we can think of the equation for $\mathcal{H}$ as
$$a\overline{a} + b\overline{b} + c\overline{c} = 1,$$
so $\mathcal{H}$ is a finite-field analog of a sphere over $\mathbb{C}$. (This is why we call it *Hermitian*, and it lets you derive the symmetry of $\mathcal{H}$ under $\mathbb{PGU}$.)

The connection may make this seem less like completely magical algebra; it's still magical that it translates to finite fields, but you may at least be more willing to believe it.

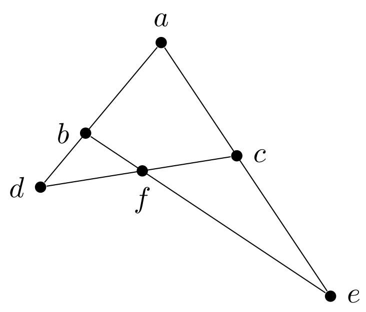
Definition 2.5. We define a *secant* to be a line that intersects $\mathcal{H}$ at $q+1$ points.

(We will come back to this definition later.)

Proposition 2.6

There do not exist six distinct points $a, b, c, d, e, f \in \mathcal{H}$ such that each of the triples abd, ace, bef, and cdf is collinear, and no other triples are.

In other words, $\mathcal{H}$ doesn't contain the following structure.



Remark 2.7. This statement might also have an analog over $\mathbb{C}$.

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

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

(as stated in Remark 2.4). Now assume for contradiction that $\mathcal{H}$ does contain the above structure. Then the collinearities mean that we can find scalars $\alpha, \beta, \gamma, \delta \in \mathbb{F}_{q^2}$ such that

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

(The reason we don't need scalars in front of the first terms — e.g., $d = \alpha_1 a + \alpha_2 b$ — is that we're in projective space, so we don't care about scaling and can just scale so that $\alpha_1 = 1$.) This gives

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

which we can rearrange to

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

But b and c are distinct points in the projective plane, so they cannot be scalar multiples of each other; this means we can conclude $\alpha + \delta = \beta + \gamma = 0$. Then by scaling all our points appropriately, we can assume without loss of generality that $\alpha = \beta = 1$ and $\delta = \gamma = -1$.

Now consider the two matrices

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

(where a has coordinates (a_1, a_2, a_3), and so on). Because a, b, and c are not collinear, both A and B are nonsingular (as their columns are linearly independent). But we can compute

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

(using the fact that $a \in \mathcal{H}$, so $a \cdot \overline{a} = 0$, and the same is true for b and c), which means

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

(we can consider all the 'diagonals' of the matrix, and these are the only ones that don't contain a 0).

But we know $d = a + b$ and $d \cdot \overline{d} = 0$, which implies that

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

Similarly $e = a + c$ and $e \cdot \overline{e} = 0$, which gives that

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

Finally, we have $f = c + d = a + b + c$ and $f \cdot \overline{f} = 0$, and by combining these equations we eventually get

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

This means the factors in the second term of $\det(AB)$ are all negations of the factors in the first term, and therefore we have

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

which contradicts the fact that A and B are nonsingular. □

§2.2 The construction of H

We now construct our graph H in the following way.

Definition 2.8. Let H be the graph whose vertices are the secants in $\mathbb{PG}(2, q)$, with an edge between every pair of secants whose intersection is in $\mathcal{H}$ — in other words,

$$V(H) = \{x^\perp \mid |x^\perp \cap \mathcal{H}| = q + 1\},$$
$$E(H) = \{(x^\perp, y^\perp) \mid x^\perp \cap y^\perp \in \mathcal{H}\}.$$

We need to check that H satisfies all the properties described in (1).

First, we'll check that H has the correct number of vertices — let $n = |V(H)|$ be the number of secants. Every point in $\mathbb{PG}(2, q)$ is contained in exactly $q^2 + 1$ lines, so by a double-counting argument — counting pairs consisting of a point in $\mathcal{H}$ and a line containing our point, i.e., $\#\{(x, y^\perp) \mid x \in \mathcal{H},\ x \in y^\perp\}$ — we get

$$(q^3 + 1)(q^2 + 1) = n(q + 1) + (q^4 + q^2 + 1 - n) \cdot 1.$$

(The left-hand side counts by first choosing x, which can be done in $|\mathcal{H}| = q^3 + 1$ ways, and then choosing $y^\perp$, which can be done in $q^2 + 1$ ways (as every point is contained in exactly $q^2 + 1$ lines). The right-hand side counts by first choosing $y^\perp$; if it's a secant then there's $q + 1$ choices for x, and if it's not a secant then there's 1 choice.) This rearranges to

$$n = q^4 - q^3 + q^2,$$

as desired.

Next, we'll check that our graph is d-regular with $d = (q+1)(q^2-1)$. To see this, we first know that every point in $\mathcal{H}$ is contained in exactly

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

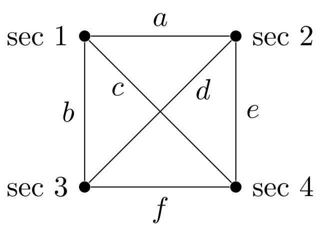
secants. (The numerator is the total number of point-secant pairs — we have n secants and each contains $q+1$ points in $\mathcal{H}$ — and the denominator is the total number of points in $\mathcal{H}$.) So for each secant $y^\perp$ and each of the $q+1$ points $x \in y^\perp \cap \mathcal{H}$, there are exactly q^2-1 other secants through x; these secants are precisely the neighbors of $y^\perp$ in our graph, so we get that $\deg(y^\perp) = (q+1)(q^2-1)$.

Next, what is the collection $\mathcal{C}$ of cliques? For each $x \in \mathcal{H}$, we define K_x as the set of secants through x (which form a clique, as every pair of such secants intersects at x); and we let

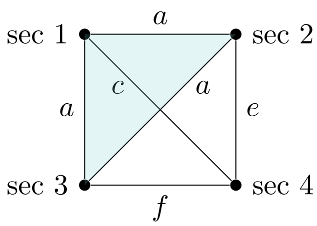
$$\mathcal{C} = \{K_x \mid x \in \mathcal{H}\}$$

be the collection of these cliques over all $x \in \mathcal{H}$. Because every point $x \in \mathcal{H}$ is contained in exactly q^2 secants, we know all these cliques have size q^2; and because $|\mathcal{H}| = q^3+1$, we have q^3+1 cliques in $\mathcal{C}$. Finally, for any two of these cliques K_x and K_y, there's only one secant through both x and y, so the two cliques only intersect at one vertex.

Finally, we'll check that the fourth property holds. Consider some K_4 contained in H, and let the six points corresponding to its six edges be $a, b, c, d, e, f \in \mathcal{H}$. (Recall that an edge in H between two secants corresponds to a point in $\mathcal{H}$ that those two secants intersect at.)



If these six labels $a, \ldots, f$ are all distinct, then this corresponds to the structure forbidden by Proposition 2.6. So either we must have a triangle all consisting of the same label, or *all* the edges must have the same label; this implies the last property (as if we have a triangle with label a, then its three vertices are in K_a).



§3 Step (2) — the clique structure of H

The rest of the proof doesn't use any properties of H other than the ones stated in Step (1) — from here, there's nothing more involving finite fields.

Given a vertex set $X \subseteq V(H)$, in order to prove (2) for X, we'll consider the intersections of X with each of our cliques in $\mathcal{C}$; we'll ignore intersections of size 1, and partition the remaining intersections into 'small' intersections, 'medium' intersections, and 'large' intersections in the following way.

Definition 3.1. Given a set $X \subseteq V(H)$ of size $|X| = k$, we partition the intersections $X \cap C$ (over all cliques $C \in \mathcal{C}$) of size at least 1 into three sets — *small*, *medium*, and *large* intersections — in the following way:

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

In other words, we have a bunch of cliques $C \in \mathcal{C}$, and we intersect all of them with X. We ignore the intersections of size 1, and classify the remaining intersections into $\mathcal{S}$, $\mathcal{M}$, and $\mathcal{L}$ depending on their size.

Definition 3.2. For each $\mathcal{U} \in \{\mathcal{S}, \mathcal{M}, \mathcal{L}\}$, we define $v(\mathcal{U}) = \sum_{T \in \mathcal{U}} |T|$ and $e(\mathcal{U}) = \sum_{T \in \mathcal{U}} \binom{|T|}{2}$.

Intuitively, $v(\mathcal{S})$ sort of counts the number of vertices in X that are in cliques with 'small' intersections — though we may count a vertex multiple times, if it's in multiple such cliques — and $e(\mathcal{S})$ counts the number of *edges* within X in such cliques.

Theorem 3.3

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

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

In other words, either lots of edges in X are in cliques with small intersections with X, or lots of edges in X are in cliques with medium intersections (for different meanings of 'lots').

Proof sketch. First, we can show that $v(\mathcal{S}) + v(\mathcal{M}) + v(\mathcal{L})$ is large by showing that X shouldn't intersect many cliques at just one vertex (so it'll intersect most cliques in at least two vertices) — more precisely, we should have

$$v(\mathcal{S}) + v(\mathcal{M}) + v(\mathcal{L}) \geq (q+1)k - q^3 - 1.$$

Then we can upper-bound the number of vertices involved in cliques with large intersections — we can show

$$v(\mathcal{L}) \leq 2k$$

by using a double-counting argument.

Now we consider two cases — when $v(\mathcal{S}) \geq v(\mathcal{M})$ we'll prove the first bound, and when $v(\mathcal{M}) \geq v(\mathcal{S})$ we'll prove the second. We'll only explain the proof in the first case; the proof is again a double-counting argument. First, by convexity, we have

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

But $|\mathcal{S}| \leq q^3 + 1$ (since there are $q^3 + 1$ *total* cliques), so we end up getting

$$e(\mathcal{S}) \geq \frac{v(\mathcal{S})^2}{4(q^3+1)} \geq \frac{k^3}{64q}.$$

□

§3.1 Step (3) — the random partitions

We'll now perform Step (3) — where we randomly partition each clique C into A and B in order to obtain H^*, and use martingale concentration inequalities to show that with positive probability, every m-vertex set X has lots of edges in H^*.

Fix some set X with $|X| = m$. The idea is that we lower-bound the number of edges in $H^*[X]$ by considering the intersections of X with either small cliques or medium cliques, depending on which one has lots of edges (as in the previous step); here we'll focus on the case $e(\mathcal{S}) \geq k^2/64q$.

In this case, we let $\{v_1, \ldots, v_\ell\} \subseteq X$ be all the vertices in X contained in some clique in $\mathcal{S}$ (where we count vertices multiple times if they're contained in multiple such cliques, so $\ell = v(\mathcal{S})$). Now we define a martingale as follows. We define the random variable

$$Z = \sum_{T \in \mathcal{S}} e[H(A_T, B_T)],$$

where A_T and B_T are the sets that T gets partitioned into when we partition its clique into two sets A and B — so Z represents the number of edges of X in *small* intersections that end up in H^*. Now for each i, let

$$Z_i = \mathbb{E}[Z \mid \text{knowledge of which parts } v_1, \ldots, v_i \text{ belong to}].$$

In other words, we imagine going through the vertices $v_1, \ldots, v_\ell$ in order, where at v_i, if $T \in \mathcal{S}$ is the clique that v_i is contained in, then we reveal which side v_i belongs to in the random partition of T; and we define Z_i as the expected number of edges given the information that's been revealed so far. (This is why we count a vertex multiple times if it's contained in multiple cliques; in one step we only consider one clique.)

Then the Z_i's form a martingale, and if $T \in \mathcal{S}$ is the clique that v_i belongs to, then we have

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

(since moving v_i to the opposite side of the partition of T can change the number of edges in T that we keep by at most $|T| - 1$, and doesn't change anything else). Let c_i be this quantity $|T| - 1$.

Finally, in order to apply concentration inequalities we need to bound $\sum c_i^2$. To do so, we have

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

(since for every clique T, there are at most $|T|$ vertices in the clique, and each contributes $(|T| - 1)^2$), and we have an upper bound on $|T|$ from the definition of $\mathcal{S}$. Then we can apply concentration bounds such as Azuma's inequality to show that Z is very unlikely to be much smaller than $\frac{1}{4}e(\mathcal{S}) \geq 2^{40}q^3$ (its expectation is $\mathbb{E}[Z] = \frac{1}{2}e(\mathcal{S})$), as desired.

> **Remark 3.4.** It's critical that the bounds we have on $e(\mathcal{S})$ and $e(\mathcal{M})$ in the two cases are different — we have $k \asymp q^2$, so our bounds are $e(\mathcal{S}) \gtrsim q^3$ and $e(\mathcal{M}) \gtrsim q^4/\log^2 n$.
>
> The bound on $e(\mathcal{S})$ corresponds to the true density of our graph (the graph has edge density $1/q$, and X has $k \asymp q^2$ vertices), and the factor of $\log n$ in our definition of $\mathcal{S}$ (where we have $|T| \leq \sqrt{2k}/\log n$) is what allows us to do a union bound (over all X).
>
> Meanwhile, in the bound for $e(\mathcal{M})$, we have a lot more edges — we essentially win a factor of q in the density. This extra factor of q is why we don't need the log factor in $|T|$ anymore (here we just have $|T| \leq \sqrt{2k}$) — if we didn't have this, then we wouldn't get good enough control without the log factor and the union bound wouldn't go through.

§3.2 Steps (4) and (5) — the container theorem and finish

Finally, here's the graph container theorem we'll use.

Theorem 3.5

Let G be a graph with the following two properties (for some R, r, α, and t).

- For all vertex subsets X of size $|X| \geq R$, we have $2e(G[X]) \geq \alpha |X|^2$.
- We have $e^{-\alpha r} n \leq R$ and $t \geq r$.

Then the number of independent sets of size t in G is at most $\binom{n}{r}\binom{R+r}{t-r}$.

In our situation (where we're applying this theorem to H^*), we have $n \approx q^4$, and we know that for every $|Y| \geq k = 2^{24}q^2$ we have

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

Now we take $R = 2^{24}q^2$, $r = 2^{10}q \log q$, $\alpha = 2^{-8}q^{-1}$, and $t = 2^{30}q \log^2 q$. Then the conditions in the theorem apply, and we conclude that

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

(where $i_t(H^*)$ denotes the number of independent sets of size t).

Finally, we take a random subset $W \subseteq V(H^*)$ where each vertex is included with probability $q^{-1}\log^2 q$; we can check (using expected value) that with positive probability, we have

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

and we're done (we take our Ramsey construction to be $H^*[W]$ for this choice of W).