2.3 Forbidding $K_{t,s}$

We have seen now many dense $K_{t,s}$-free constructions, but none so far where both parameters t and s are at least 4. While the Turán numbers of $K_{2,s}$ and $K_{3,s}$ are well-understood, the densest $K_{4,4}$-graph we know is essentially the Brown graph. This graph does not even have a $K_{3,3}$, consequently has only $\Theta(n^{5/3})$ edges—very far from the KST-upper bound of $n^{7/4}$ for $K_{4,4}$.

So what does become so hard when t and s are both at least 4? Phrasing it mysteriously, besides us being not creative enough, the problem is that $4 = 2 + 2$.

For $t \leq s$, it is natural to choose a $K_{t,s}$-free construction to live in some (affine or projective) t-dimensional space over a finite field $\mathbb{F}_q$, at least this what we have done so far. The vertex set is typically the space itself and the neighborhood of each vertex is chosen to be some hypersurface, so that the number of edges does work out well. One of course has to prove that these $(t-1)$-dimensional surfaces contain the appropriate number, $\approx q^{t-1} \approx n^{1-\frac{1}{t}}$, of points; this is sometimes easier, sometimes harder, but could always be done. Then one must also show that the intersection of any t of these $\approx q^t$ neighborhood hypersurfaces contains at most $s-1$ points. The exercises leading up to this section tried to demonstrate that this is the more problematic issue, and in fact the first and most critical point is whether these t-wise intersections contain at most *constantly many* points. Then it is usually only a bonus that by what one is able to bound this constant, let it be $t-1$ or something else. In Exercise ?? and Exercise 2.11 we investigated promising constructions, which broke down in a strong way: we found complete bipartite graphs whose order tended to infinity with the order of the graph. Part of the problem was degeneracies: a line or some low degree curve was part of several neighborhood surfaces.

Neighborhoods are $t-1$ dimensional hypersurfaces, common neighborhoods are the intersections of these. Two "average"hypersurfaces intersect in a $(t-2)$-dimensional surface, and taking one more average hypersurface to the intersection always reduces the dimension by one. Hence when one takes t average hypersurfaces in t-space one expects that their intersection is 0-dimensional, that is the union of constantly many points. Our experience shows that for $t \geq 4$ it is hard to select q^t hypersurfaces in the t-dimensional space such that *any* t of them has a 0-dimensional intersection.

In the first section we pursue the idea of choosing the neighborhood hypersurfaces randomly. We will succeed in formalizing the intuitive idea that the intersection of t "avergae" hypersurfaces has a 0-dimensional, and thus constant size, intersection.

In the two subsequent sections we introduce two related explicit constructions that give better estimates on the intersection size.

2.3.1 Random algebraic constructions

We start by revisiting the simplest random construction and take a closer look at why it fails.

Since the constant factor of the number of edges will not play any role in this section, we prefer to work in the bipartite random graph model $G(n,n,p)$, that allows a technically cleaner treatment. Let L and R be two disjoint sets of vertices of size n each. For each pair $u \in L$ and $v \in R$ of vertices we put an edge with probability $p = n^{-\frac{1}{t}}$, with the choices being mutually independent.

We choose this probability so that the number of edges in expectation n^2p, will be of the order $n^{2-\frac{1}{t}}$—the order we expect to be the right one for the Turán number of the complete bipartite graph $K_{t,s}$ for any $t \leq s$.

We already convinced ourselves that at such high probability the random graph will be filled with copies of $K_{t,t}$ and even the alteration methods works only for much smaller edge probability.

Let us fix a set $U \subseteq L$ of size t and study the distribution of the size of the common neighborhood $N(U) = \cap_{u\in U} N(u)$ of U. For each vertex $v \in R$ let $I(v) = I_U(v)$ denote the indicator random variable of the event that $v \in N(U)$. Then $\mathbb{P}[I(v) = 1] = p^t = \frac{1}{n}$. The random variables $I(v)$ are mutually independent for $v \in V$, since they depend on pairwise disjoint sets of pairs of vertices.

This means that the random variable $|N(U)| = \sum_{v\in R} I(v)$ has binomial distribution with parameters n and $\frac{1}{n}$. Its expectation is $n \cdot \frac{1}{n} = 1$, so on the average U does *not* even form the larger side of a $K_{t,2}$, not to mention forming one of the sides of a $K_{t,t}$. So, at least on average, everything looks rosy! How come then, that still, with overwhelming probability, $G(n,n,p)$ contains not only $K_{t,t}$, but even $K_{t,g(n)}$ for some function g of n tending to infinity?

Having a binomial distribution with constant expectation, the distribution of $|N(U)|$ can be estimated by the Poisson distribution of the same expectation. In particular, for every constant $s \in \mathbb{N}$,

$$\mathbb{P}[|N(U)| = s] = \binom{n}{s}\frac{1}{n^s}\left(1 - \frac{1}{n}\right)^{n-s} \to \frac{1}{e}\cdot\frac{1}{s!},$$

This means that while on the average each of the $\binom{n}{t}$ t-subsets of L will have a single common neighbor in R, at the same time each has a constant (independent of n) chance of having a common neighborhood of size s, for any constant s (say, also for $s = t^{t^{t^t}}$).

For example when $t = 4$, the probabilty that a fixed four-element set $U \subseteq L$, has 100 common neighbor tends to a constant. This constant is tiny, something of the order 10^{-157}, but independent of n. Consequently the expected number of 4-subsets of L hosting a $K_{4,100}$ in $G(n,n,p)$ is of the same order n^4 as the number of four elements sets. Calculating the variance and applying Chebyshev's inequality to the random variable X counting such 4-sets, one can show that the value of X is likely to be very close to its expectation. Hence $\Theta(n^4)$ copies of $K_{4,100}$ do appear with overwhelming probability.

In a way the problem is the long smooth tail of the Poisson distribution: any constant value appears with constant probability. The cause of the long smooth tail is the full mutual independence of the n^2 random bits we use to create our random graph. This definitely has positive effects: it is very useful for example to have the number of edges what we want it to be with high probability, and it is certainly very pleasent that we can understand and analyse the properties of the random graph relatively easily. On the other hand we have also seen the adverse effects of this excessive independence: the very subgraph of constant size we would like to avoid, say $K_{4,100}$, also appears in it almost surely. And not only that, but its copies are populating all parts of the random graph so densely, that even plastic surgery can not get rid of them all without making the graph essentially empty.

This is what we circumvent with the introduction of algebra into a probabilistic model. Instead of using n^2 independent random bits to create our random bipartite graph on two parts of size n each, we define a random algebraic model where we use only $\sim \log n$ bits to create a random graph on the same number of vertices. This limited amount of independence will be enough to maintain the expected number of edges and also to provide effective enough bounds on the size of common neighborhoods.

The key algebraic theorem we use to cut away a significant chunk of the Poisson-tail that was killing us above, provides a quantitative correlation between two different measures of "how large" an algebraic variety over $\bar{\mathbb{F}}_q$ is. One is the algebro-geometric notion of dimension, the other is the number theoretic concept of the number of its elements with coordinates in $\mathbb{F}_q$. To illustrate, consider t linear functions $f_1, \ldots, f_t \in \mathbb{F}_q[y_1, \ldots, y_m]$. The set of their common zeroes is the intersection of t hyperplanes. This set is an affine subspace of a certain dimension d and hence its size is q^d. The *Lang-Weil bound* is a far-reaching generalization of this counting fact for the situation when the f_i are polynomials of higher degree. It states that under certain assumptions the number of common zeroes of a set of polynomials $f_1, \ldots, f_t \in \mathbb{F}_q[y_1, \ldots, y_m]$ in $\mathbb{F}_q^m$ is of the order $q^{\dim V}$, where V denotes the set of common zeroes of the f_i over the algebraic closure $\bar{\mathbb{F}}_q$. Since the dimension is an integer, the possible number of common zeroes of the polynomials is quite limited, for example it is never $\log q$ or $q^{\frac{9}{10}}$ (for q large enough).

For our purposes the following spezialized consequence of the Lang-Weil bound, with no assumptions on the polynomials, will be convenient.

Theorem 2.9 *[?] For every $t, d \in \mathbb{N}$, there exists a constant $C = C(t, d)$ such that for arbitrary polynomials $f_1, \ldots, f_t \in \mathbb{F}_q[y_1, \ldots y_t]$ of degree at most d, the number of their common zeroes $y \in \mathbb{F}_q^t$ is either less than C or more than $q - C\sqrt{q}$.*

The proof of the Lang-Weil bound and in turn of the above theorem is beyond the scope of our lecture.

This theorem implies that in any graph where the neighborhoods are hypersurfaces defined by polynomials, there are no common neighborhoods of size between C and $q - C\sqrt{q}$. Therefore this part of the tail will be gone!

Construction. The vertex set will be two copies of $\mathbb{F}_q^t$, which we denote by L and R. For a polynomial $F(x, y) \in \mathbb{F}_q[x_1, \ldots, x_t, y_1, \ldots, y_t]$ we define a graph $G = G(f)$ on the vertex set $L \cup R$. Vertices $u \in L$ and $v \in R$ form an edge if $F(u, v) = 0$. Note that neighborhood of a vertex $u \in L$ is the hypersurface $\{y \in \mathbb{F}_q^t : F(u, y) = 0\}$.

We set $d = d(t) = t^2 - t + 2$. For the randomized construction, we choose the polynomial F uniformly at random from the set $\mathcal{P}_d \subset \mathbb{F}_q[x_1, \ldots, x_t, y_1, \ldots, y_t]$ of polynomials that have degree at most d both in the variables $x_1, \ldots, x_t$ and the variables $y_1, \ldots, y_t$. The probability space is the uniform product space $\mathbb{F}_q^{C_d}$, where C_d is the number of monomials satisfying the degree condition, where the coordinates correspond to the coefficients of the monomials.

As a first sanity check, let us see that for every pair $u \in L$ and $v \in R$ the probability that they form an edge in $G(F)$ is the right one, i.e. $n^{-\frac{1}{s}} = \frac{1}{q}$. This follows immediately by writing $F = P + Q$ as the sum of its non-constant terms P and its constant term Q and noting that for any polynomial $g \in \mathcal{P}_d$ with constant term 0, the probability of $F(u, v) = 0$ conditioned on $P = g$ is exactly $\frac{1}{q}$. This is because $F(u, v) = 0$ happens if and only if Q is equal to the constant $-g(u, v) \in F_q$.

This already implies that the number of edges has the right expectation. Writing $e(G) = \sum_{u \in L, v \in R} \nVdash_{uv \in E(G)}$ as the sum of indicator random variables and using the linearity of expectations we have that

$$\mathbb{E}\left[e(G)\right] = \sum_{u \in L, v \in R} \mathbb{E}\left[\nVdash_{uv \in E(G)}\right] = \sum_{u \in L, v \in R} \mathbb{P}\left[F(u, v) = 0\right] = n^2 \cdot \frac{1}{q} = n^{2 - \frac{1}{s}}.$$

We will show that the expected number of t-sets in L which host a copy of a $K_{t,C}$ in $G(f)$ for the constant $C = C(t, d)$ from Theorem 2.9 will be much smaller than q^{t-1} and hence deleting all incident edges to these t-sets produces a $K_{t,C}$-free graph with the right number of edges.

Let us fix a vertex set $U \subseteq L$ of size $|U| = t$. For every $u \in U$, consider the polynomials $F(u, y) \in \mathbb{F}_q[y_1, \ldots y_s]$. The number of their common zeroes $v \in \mathbb{F}_q^s$ is equal to $|N(U)|$. By Theorem ?? $\mathbb{P}\left[|N(U)| \geq C\right] = \mathbb{P}\left[|N(U)| \geq q - C\sqrt{q}\right]$. Here is where the tail-cutting happnes: to bound the number of $K_{t,C}$ in G it is then sufficient to bound the number of $K_{t,q-C\sqrt{q}}$. And we will be able to prove this to be very small, since q is not a constant.

To upper bound the upper bound of random variable eventually our only tool is Markov's Inequality, the effectiveness of which one can always try to boost using a higher moment, in particular if the variable in question is the sum of random variables with at least a limited amount of independence. This is our plan as well to bound $\mathbb{P}\left[|N(U)| \geq s\right]$. Recall that for a vector $v \in R$, $I(v)$ denotes the indicator random variable of the event that $v \in N(U)$, that is, $F(u, v) = 0$ for every $u \in U$. Then $|N(U)| = \sum_{v \in V} I(v)$. We will first establish the d-wise independence of the random variables $I(v)$ in order to bound $\mathbb{E}\left[|N(U)|^d\right]$ and use it in

$$\mathbb{P}\left[|N(U)| \geq s\right] = \mathbb{P}\left[|N(U)|^d \geq s^d\right] \leq \frac{\mathbb{E}\left[|N(U)|^d\right]}{s^d}.$$

The following lemma establishes the d-wise indpendence of random variables $I(v)$.

Lemma 2.9.1 *Let $t, r, d \in \mathbb{N}$ be such that $t, r \leq \min\{d, \sqrt{q}\}$. For every $U \subseteq L$ of size $|U| = t$ and $V \subseteq R$ of size $|V| = r$ we have that*

$$\mathbb{P}\left[F(u, v) = 0 \text{ for every } u \in U, v \in V\right] = \frac{1}{q^{tr}},$$

where F is polynomial chosen uniformly random from $\mathcal{P}_d$.

Proof. We call a set of vectors *simple* if the first coordinates are pairwise distinct. First we prove the statement in the case when both U and V are simple.

We write $F = P + Q$, where Q is the part containing the monomials of F the form $x_1^i y_1^j$ for $0 \le i \le t-1$ and $0 \le j \le r-1$ and P is the rest. First we generate P and then Q. Let $g \in \mathcal{P}$ be an arbitrary polynomial without the above monomials. We claim that the probability of $F(u,v) = 0$ happening for every $u \in U$ and $v \in V$, conditioned on $P = g$, is $\frac{1}{q^{tr}}$. Indeed, $F(u,v) = 0$ happens if and only if for the tr uniformly random coefficients $\beta_{i,j} \in \mathbb{F}_q$ we have that

$$\sum_{j=0}^{r-1} \sum_{i=0}^{t-1} \beta_{i,j} u_1^i v_1^j = -g(u,v),$$

which gives a linear equation system with tr equations in tr variables. Its matrix $(u_1^i v_1^j)_{u,v,i,j}$ is the tensor product of the two Vandermonde matrices $(u_1^i)_{u,i}$ and $(v_1^j)_{v,j}$. The determinant is conseuqently the product of $(\det(u_1)_i)^t$ and $(\det(v_1)_j)^r$. Both of these determinants are non-zero, since the u_1 are all distinct and the v_1 are all distinct. Therefore there is a unique solution $(\beta_{i,j}) \in \mathbb{F}_q^{tr}$ satisfying the equation system. Each vector in $\mathbb{F}_q^{tr}$ has the same probability to appear as the coefficients of Q, so the conditional probability is indeed $\frac{1}{q^{tr}}$. Since this holds for every g, we have that the unconditioned probability of $F(u,v) = 0$ is also $\frac{1}{q^{tr}}$.

Let now $U \subseteq L$ and $V \subseteq R$ be arbitrary. We will find invertible linear transformations T and S of $\mathbb{F}_q^t$, such that $T(U)$ and $S(V)$ are both simple. Let $\tilde{F}(x,y) := F(T(x), S(y))$. Note that $F \to \tilde{F}$ is one-to-one map from $\mathcal{P}_d$ to $\mathcal{P}_d$, hence the distribution of $\tilde{F}$ is also uniform on $\mathcal{P}_d$. Since $T(U)$ and $S(V)$ are simple and have sizes t and r, respectively, the probability that $\tilde{F}(u,v) = F(T(u), S(v)) = 0$ for every $u \in U$ and $v \in V$ is $\frac{1}{q^{tr}}$. This is then equal to the probability of $F(u,v) = 0$ for every $u \in U$ and $v \in V$, because the distribution of $\tilde{F}$ and F are the same.

In order to create the invertible linear function T, we define its first coordinate $T_1 : \mathbb{F}_q^t \to \mathbb{F}_q$ and extend it so the function is invertible. A linear function is determined uniquely by its values on any basis, so the number of potential T_1 is q^t. Among these, for any $u \neq u'$ there are exactly q^{t-1} for which $T_1(u) = T_1(u')$. One can see this for example by extending $u - u'$ to a basis, so the values of T_1 are freely chosen on all but the first element of this basis. Consequently there are at least $q^t - \binom{t}{2} q^{t-1} > 0$ linear functions T_1, such that $T_1(u) \neq T_1(u')$ for any $u, u' \in U$. Here we used that $t < \sqrt{q}$.

The construction of S is analogous, but we use that $r < \sqrt{q}$. $\square$

Now using the linearity of expectation,

$$\mathbb{E}\left[|N(U)|^d\right] = \mathbb{E}\left[\left(\sum_{v\in V} I(v)\right)^d\right] = \sum_{v_1\in V}\cdots\sum_{v_d\in V}\mathbb{E}\left[I(v_1)\cdots I(v_d)\right]$$
$$= \sum_{v_1\in V}\cdots\sum_{v_d\in V}\mathbb{P}\left[\{v_1,\ldots,v_d\}\subseteq N(U)\right]$$

By Lemma 2.9.1 the probability in the summand is $\frac{1}{q^{tr}}$ where r is the cardinality of the set $\{v_1,\ldots,v_d\}$. We classify the sum according to r and obtain that

$$\mathbb{E}\left[|N(U)|^d\right] = \sum_{r=1}^{d}\binom{n}{r}M_r\frac{1}{q^{tr}} < \sum_{r=1}^{d} n^r M_r\frac{1}{q^{tr}} = \sum_{r=1}^{d} M_r =: m_d$$

where M_r denotes the number of surjective functions from a d-element set to an r-element set.

We now proved that similar to the expectation, the dth moment of $|N(U)|$ is also bounded by a constant.

This is good because then we have that

$$\sum_{U\subseteq L,|U|=t}\mathbb{P}\left[|N(U)|\geq s\right]\leq\frac{\mathbb{E}\left[|N(U)|^d\right]}{s^d}\leq\frac{m_d}{s^d}.$$

So for the expected number X of t-sets that form a side of a $K_{t,C(t,C(t,d(t)))}$ we have

$$\mathbb{E}\left[X\right] = 2\binom{n}{t}\frac{m_d}{(q/2)^d}\leq c_d' n^t q^d = c_d' q^{t^2} q^{-t^2+t-2} = c' q^{t-2}$$

for large q.

Let us now choose a ploynomial $f\in\mathcal{P}_d$ for which the value of the random variable $e(G)-q^t X$ is at least its expectation $q^{2t-1}-c' q^{2t-2}$. From the graph $G(f)$ delete all edges incidedent to one vertex in every bad t-set. In the resulting graph H there are obviously no $K_{t,C(t,C(t,d(t)))}$, yet the number of edges is at least $q^{2t-1}-c' q^{2t-2} = \Omega\left(n^{2-\frac{1}{t}}\right)$.

2.3.2 The norm-graphs

The next exercise demonstrates some more what can occur when we step into the fourth dimension. It is a prelude for what is coming in the following section.

Exercise 2.15 *Let the vertex set of a graph G be $\mathbb{F}_p^4$. Let (a,b,c,d) be adjacent to (a',b',c',d') if and only if $(a+a')(b+b')(c+c')(d+d')=1$. Prove that G contains a $K_{n^{1/4},n^{1/4}}$.*