Chapter 6

The asymmetric Ramsey-problem

6.1 Constructive $R(K_3, K_k)$

We continue our study of explicit constructions for the Ramsey problem with the asymmetric case. We want lower bound $R(3, k)$ constructively. The general bound of Erdős and Szekeres gives

$$R(3,k) < \binom{k+1}{2} = O(k^2).$$

This was improved by Ajtai, Komlós and Szemerédi (1981) to $k^2/\log k$. Their argument is considered the first application of the "nibble-method", which then went on to become one of the more succesful technique of probabilistic combinatorics.

From the other side, an ugly, but routine application of the Local Lemma gives a lower bound of $\left(\frac{k}{\log k}\right)^2$. In 1995 Kim, in remarkable tour de force of probabilistic combinatorics, managed to remove one log-factor and established the correct order of magnitude $\frac{k^2}{\log k}$. He received the Fulkerson prize for his paper. Curiously, Kim is using the very same nibble method that was invented while proving the matching upper bound.

The first explicit construction of triangle-free graph where the vertex set is of superlinear size in terms of the independence number is due to Erdős who showed

$$R(3,k) = \Omega\left(k^{2\log 2/3(\log 3 - \log 2)}\right) = \Omega\left(k^{1.13}\right)$$

constructively. Later this has been improved by Chung, Cleve and Dagum to

$$\Omega\left(k^{\log 6/\log 4}\right) = \Omega\left(k^{1.29}\right),$$

and even further improved by Alon to $\Omega\left(k^{4/3}\right)$. Later Alon polished his approach, giving a construction with $\Omega(k^{3/2})$ vertices. His proof bounds the independence number through the second eigenvalue of the adjacency matrix of the graph. The second part of this section is devoted to this construction, but first we look at a more recent, simpler one, found by Codenotti, Pudlak and Resta [?].

6.1.1 A simple one

We start by describing a weaker construction that demonstrates the basic trick of [?]; this will later be generalized in Section 6.2. The final twist to this idea will be added in a subsequent exercise. Let B be either the Benson-graph or the Wenger-graph of girth 8 described in Section 3.4. From B we construct a new graph G defined on the edges of B. Let $V(G) := E(B)$ and two vertices x_1y_1, x_2y_2 of G with $x_1 \neq x_2$, $y_1 \neq y_2$ are adjacent if x_1y_2 or x_2y_1 is an edge of B.

First, we check that there is no triangle in G. Suppose to the contrary that there are vertices $x_1y_1, x_2y_2, x_3y_3 \in V(G)$ which form a K_3. Then the six vertices $x_1, x_2, x_3, y_1, y_2, y_3$ of B span three other edges, six altogether. Consider the subgraph of B induced by these six vertices. Were two of the neighborhoods of x_1, x_2 and x_3 intersecting $Y = \{y_1y_2, y_3\}$ in more than one vertex, we would immediately arrive at a contradiction to the C_4-freeness of B. Hence the neighborhood of any x_i in Y must contain exactly two vertices, and any pair of them should intersect in exactly one vertex. That is, the six edges must form a six-cycle in B, yet another contradiction. So G is indeed triangle-free.

Let $k = |V(B)|$. Then $|V(G)| = |E(B)| = \Theta\left(k^{4/3}\right)$. We set to prove that every independent set of G is of size $O(k)$ giving an explicit lower bound of order $k^{4/3}$ to $R(3, k)$. Let $I \subseteq V(G)$ be an independent set of G. The members of I are edges in B and we claim they form a star-forest in B. Indeed, should there be a path $e_1e_2e_3$ of three edges in I, the two terminal edges e_1 and e_3 would be adjacent in G, contradicting the independence of I. Hence $|I| \leq k - 1$ and we are done.

Exercise 6.1 *Improve further the above construction by using as B, instead of the Benson graph, the point/line incidence graph of the projective plane (introduced in Subsection 2.1.3 as Construction 1.) and modifying slightly the definition of an edge in G. Let $V(G) := E(B)$ and let P be one of the partite sets of B, say the points of the projective plane. Fix an arbitrary ordering $\prec$ on P. Two vertices x_1y_1, x_2y_2 of G are adjacent if $x_1 \prec x_2$, $y_1 \neq y_2$ and $x_1y_2 \in E(B)$. (So roughly "half" of the edges are kept compared to the previous construction.) Show that this graph is an explicit lower bound of order $k^{3/2}$ for $R(3, k)$.*

6.1.2 Alon's Construction

As in all our Ramsey-type constructions, we try to imitate randomness. We aim to bound the independence number of the constructed quasi-random graph via its second eigenvalue. Cayley graphs are very symmetric and their eigenvalues are strongly connected to the charaters of the underlying group. Let us be given a group H where the operation $+$ is written additively. We define a graph G on the vertex set $H = V(G)$. The adjacency relation will be given by a subset $S \subseteq H$ the in the following fashion: $g, h \in H$ are adjacent if $g - h \in S$. To make sure that the adjacency relation is symmetric we also require that $S = -S$. In order to avoid loops we assume $0 \notin S$. Note that the neighborhood of each vertex $g \in V(G)$ is the set $g + S$.

Cayley graphs are often good quasi-random graph provided the subset S is chosen "random enough" inside the underlying group. For example, if one chooses S to be a proper subgroup of H, then $G(H, S)$ is very not quasi-random, it is not even connected:

it contains $|H|/|S|$ disjoint cliques, each being a coset of the subgroup $S < H$. In other words, a subgroup of H is not a random-like subset of H.

Passing to Cayley graphs from general graphs represents a simplification of our task, in some sense we reduce the "dimension of our problem": instead of trying to construct all the edges of a quasi-random graph, we are content with describing a quasi-random neighborhood of one vertex and then make sure that all vertices look the same locally. Definition and analysis gets simpler, but of course, we reduce the playing field significantly.

Let us start by looking at what makes a Cayley graph triangle-free.

Claim 4 *A Cayley graph $G(S, H)$ is triangle-free if and only if the equation $s_1 + s_2 + s_3 = 0$ has no solution s_1, s_2, s_3 in S.*

Proof. Suppose that $u, v, w \in V(G)$ form a triangle. Then $v = u + s_1$ for some $s_1 \in S$, $w = v + s_2$ for some $s_2 \in S$, $u = w + s_3$ for some $s_3 \in S$. Hence $u = w + s_3 = v + s_2 + s_3 = u + s_1 + s_2 + s_3$ implying $s_1 + s_2 + s_3 = 0$. For the other direction, let us assume that $s_1 + s_2 + s_3 = 0$ for some $s_1, s_2, s_3 \in S$. Then for any $u \in V(G)$, the vertices $u, u + s_1$ and $u + s_1 + s_2$ form a triangle. It is clear that there is an edge between any pair of these three vertices, and then this also implies that they do represent three different vertices, since G is loop-free. $\square$

In a vector space over the two-element field $\mathbb{F}_2$ the linear equation $s_1 + s_2 + s_3 = 0$ is equivalent to the vectors s_1, s_2 and s_3 being linearly dependent. Hence triangle-freeness of a Cayley graph $G(\mathbb{F}_2^n, S)$ is equivalent to the set S being three-wise linearly independent (cf. the three-wise independence of sample spaces as defined in Subsection 5.1.2). Consequently, the idea of using the parity check matrices of BCH-codes arises quite naturally.

The first try

First we obtain a weaker bound on $R(3, k)$ by introducing one of the basic ideas of Alon's construction. Let G be the Cayley graph on the group $\mathbb{Z}_2^{2k}$ with a "neighborhood set" $S \subseteq \mathbb{Z}_2^{2k}$, that consists of the column vectors of the parity check matrix of the binary BCH-code of designed distance 5. To recall, this construction uses the 2^k-element field $\mathbb{F}_{2^k}$. On the one hand $\mathbb{F}_{2^k}$ is a k-dimensional vector space over $\mathbb{F}_2$, that is, its elements can be written as $0-1$-vectors of length k and the addition in $\mathbb{F}_{2^k}$ is just the usual vector addition in $\mathbb{Z}_2^k$. On the other hand, the multiplication of $\mathbb{F}_{2^k}$ completely "messes up" the additive structure. We will in fact use the vector notation to denote the elements of the field but will switch between interpreting the 0-1-vector as an element of $\mathbb{F}_{2^k}$ or $\mathbb{Z}_2^k$. For two vectors $u, v \in \mathbb{Z}_2^k$ let $[u, v]$ denote their concatenation in $\mathbb{Z}_2^{2k}$.

Let $S = S^{(2)}_{simple} = \{[z, z^3] \in Z_2^{2k} : z \in \mathbb{F}_{2^k} \setminus \{0\}\}$. Here, of course, z^3 denotes the binary vector of length k corresponding to the field element z^3.

The number of vertices in this Cayley graph is $n = 2^{2k}$, while the degree d of regularity is equal to $|S| = 2^k - 1 \approx \sqrt{n}$.

This graph certainly does not contain any triangle, as that would mean that there are three different $z_1, z_2, z_3 \in \mathbb{F}_{2^k} \setminus \{0\}$ such that

$$\begin{aligned} z_1 + z_2 + z_3 &= 0 \text{ and} \\ z_1^3 + z_2^3 + z_3^3 &= 0. \end{aligned}$$

This is of course impossible since substituting $z_1 + z_2$ for z_3 in the second equation (we used that $x = -x$ over characteristic 2) we get $0 = z_1^3 + z_2^3 + (z_1 + z_2)^3 = z_1^3 + z_2^3 + z_1^3 + 3z_1^2 z_2 + 3z_1 z_2^2 + z_2^3 = z_1 z_2 (z_1 + z_2) = z_1 z_2 z_3$. Hence one of the z_i must be 0 which is not allowed. Thus there is no triangle in G.

Exercise 6.2 *Argue why using squares instead of cubes in the definition of S would not be a good idea. That is, what is the problem with the neighborhood set $\{[z, z^2] : z \in \mathbb{F}_{2^k}\}$?*

How about the largest independent set? We intend to use Corollary A.36 and bound the largest non-principal Fourier coefficients of the characteristic vector of the neighborhood set S.

By Corollary A.36 we know that $\alpha(G(S)) \leq n \frac{\Phi(S)}{|S|}$. What is now $\Phi(S)$? For that, what is $\widehat{\mathbb{1}}_S = \sum_{s \in S} \chi(s)$? Recall that every character of Z_2^{2k} is determined by an element of Z_2^{2k} the following way: $\chi_\beta(\alpha) = (-1)^{\langle \alpha, \beta \rangle}$. Then

$$n \widehat{\mathbb{1}}_S(\chi_\beta) = \sum_{s \in S} \chi_\beta(s) = \sum_{s \in S} (-1)^{\langle \beta, s \rangle}.$$

Imagine that the elements of S written as colums of the $2k \times (2^k - 1)$ -dimensional matrix M. We are interested in the vector $\beta^T M$, whose components are the bits $\langle \beta, s \rangle$. In fact we are rather interested in the Hamming weight $w = w(\beta)$ of $\beta^T M$: for each 1-coordinate of $\beta^T M$ we have a -1 term in the corresponding Fourier coefficient, while for each 0-coordinate we have a 1. That is the corresponding Fourier coefficient is $d - 2w$. Obviously, we want this to be as small as possible, that is, we want roughly the same number of ones and zeros in the vector $\beta^T M$, for every $\beta \in \mathbb{Z}_2^{2k}, \beta \neq 0$. This is certainly easy to check for a β of the form $[\beta_1, 0]$ where $\beta_1 \in \mathbb{Z}_2^k, \beta_1 \neq$: there are exactly one more ones than zeros in $\beta^T M$. Similar is true for β of the form $[0, \beta_2]$ where $\beta_2 \in \mathbb{Z}_2^k, \beta_1 \neq$ provided $z \to z^3$ is a bijection, that is, if $3 \not| 2^k - 1$ or k is odd.

Now comes the big cannon, a theorem of Carlitz and Uchiyama (which is itself a nontrivial consequence of the theorem of Weil about the Rieman hyothesis over finite fields) tells us that the Hamming weight of $\beta^T M$ is roughly half of the length for *any* $\beta \in \mathbb{Z}_2^{2k}, \beta \neq 0$. More precisely, $|w - 2^{k-1}| \leq 2^{k/2}$. This means that for every $\beta \neq 0$, $|n \widehat{\mathbb{1}}_S(\chi_\beta)| \leq 2^{k/2}$ and thus $\Phi(S) \leq 2^{k/2} \approx \sqrt{d}$.

Combining this with our bound on the independence number we have $\alpha(G(S)) \leq n/\sqrt{d} = n^{3/4}$. Hence we have shown $ExpR(3, k) \gtrapprox k^{4/3}$.

Remark. The matrix M is the parity check matrix of a binary BCH code (with designed distance 5), and as such, the linear combinations of its rows form the dual of the code.

That is, the vector $\beta^T A$, which is just linear combination of the rows with coefficients β_i, is an element of this dual code. The theorem of Carlitz and Uchiyama then tells us that the weight of any code word in the dual of a BCH code is roughly half of the length.

Increasing the degree

Are we at the end of the road for the BCH-idea? One certainly senses the presence of some leftover in the triangle-freeness proof: the full power of the BCH-matrix is not exploited. We only use that any *three* columns of M are linearly independent, while it is also true that any *four* of them are.

Observe that our bound on the independence number depends on $\Phi(S)/|S|$. If we assume that S is perfectly quasi-random in the sense that $\Phi(S) \approx \sqrt{|S|}$, then the upper bound on $\alpha(G)$ depends solely on how big the degree $d = |S|$ is. In the previous subsection we had $d = \Theta(n^{1/2})$, so there could be some room for improvement.

How can we increase the degree? Considering larger BCH-matrices and taking not only z and z^3, but z^5, z^7, etc..., will increase the length of our vectors, together with the parameter of their linear independence property, but the degree of the graph goes down (see Exercise 6.3). The Carlitz-Uchiyama bound remains valid, however, thus the Hamming weight of $\beta^T M$ remains roughly half of the length, just the error term worsens by a constant factor.

Theorem 6.1 (Carlitz-Uchiyama) *Let $M = M_h$ be the $hk \times (2^k - 1)$ matrix whose columns are the vectors of the form $[z, z^3, z^5, \ldots, z^{2h-1}]$, with $z \in \mathbb{F}_{2^k}^*$. Then for every $\beta \in \mathbb{Z}_2^k, \beta \neq 0$, the Haming weight $w = w(\beta)$ of $\beta^T M$ satisfies*

$$\left|w - 2^{k-1}\right| \leq (h-1)2^{k/2}.$$

Remark. The matrix M_h is the parity-check matrix of the binary BCH-code of designed distance $2h+1$. As it was proved in Subsection 5.1.2, any $2h$ columns of M_h are linearly independent.

Exercise 6.3 *Describe an explicit construction showing $ExpR(C_5, K_k) = \Omega(k^{6/5})$, using the parity check matrix of the BCH-code of designed distance 7. Explain why the constructed graph is also K_3-free, but not C_4-free or C_6-free.*

In the previous exercise we worked over a larger group $\mathbb{Z}_2^{3k}$ and our neighbor set S_{simple} consisted of all vectors of the form $[z, z^3, z^5]$, $z \in \mathbb{F}_{2^k}^*$. This resulted in degree only of order $n^{1/3}$, so-so for C_5-freeness, but too little for a K_3-free construction. To increase the degree, let us partition S_{simple} into two, for the moment arbitrary, subsets W_0 and W_1 and let us define our new neighbor set as $S = \{w_0 + w_1 : w_0 \in W_0, w_1 \in W_1\}$. Note that the degree d is $|S| = |W_0||W_1|$; if $w_0 + w_1$ was equal to $w_0' + w_1'$, then, since any four elements of S_{simple} are linearly independent, we had $w_0 = w_0'$ and $w_1 = w_1'$. To maximize the degree, we will select W_0 and W_1 with almost equal size and then $d \approx n^{2/3}$.

Alon's graph is the Cayley graph G on $\mathbb{Z}_2^{3k}$ with the neighborhood set S, where the partition $W_0 \cup W_1$ of S_{simple} will be chosen later appropriately.

Why is there no triangle in G? Were there a triangle then by Claim 4 there would be three different elements $w_0 + w_1, w_0' + w_1', w_0'' + w_1''$ of S which sum up to 0. On the other hand we know that any *six* members of S_{simple} are linearly independent so $w_0 + w_0' + w_0'' + w_1 + w_1' + w_1''$ can only be 0 if every vector occurs an even number of times. Since the sets W_0 and W_1 are disjoint and the sum contains an odd number of elements from each, this is clearly impossible.

For estimating the independence number we again use the upper bound involving $\Phi(S)$.

$$\sum_{s\in S}\chi(s) = \sum_{s_0\in W_0}\sum_{s_1\in W_1}\chi(s_0+s_1) = \sum_{s_0\in W_0}\sum_{s_1\in W_1}\chi(s_0)\chi(s_1) = \left(\sum_{s_0\in W_0}\chi(s_0)\right)\left(\sum_{s_1\in W_1}\chi(s_1)\right).$$

For a character $\chi = \chi_\beta$, the sum $\sum_{s_0\in W_0}\chi_\beta(s_0) = \sum_{s_0\in W_0}(-1)^{\langle\beta,s_0\rangle}$ depends on the W_0-entries of the vector $\beta^T M$. By the same argument as in the previous subsection, $\sum_{s_0\in W_0}\chi_\beta(s_0)$ is equal to $|W_0| - 2x$ where x is the Hamming weight of the vector $\beta^T M$ restricted to the coordinates at W_0. Similarly, $\sum_{s_1\in W_1}\chi_\beta(s_1)$ is equal to $|W_1| - 2y$ where y is the Hamming weight of the vector $\beta^T M$ restricted to the coordinates at W_1. In conclusion, our main concern is to minimize the product $(|W_0| - 2x)(|W_1| - 2y)$. Of course the Carlitz-Uchiyama bound does tell us that the sum $x+y$, the Hamming weight of the vector $\beta^T A$, is roughly half of 2^k, but in principle it could happen that x and y are very non-equal resulting in $(|W_0| - 2x)(|W_1| - 2y)$ being very large.

In order to avoid this we must now specify the selection of W_0 and W_1 such that $x \approx y$. A random partition would certainly fit the bill, but considering the business we are in, we need to be, well, more explicit. We need to find a 0-1 vector of length $2^k - 1$, that is more or less independent of the previous row vectors of M. The message of Subsection 5.1.2 was that the linear independence of the columns corresponds to probabilistic independence of the rows and in particular the rows of the matrix of the BCH-code proovide a good approximation of independence. Hence the natural choice for the required quasi-random partition should be given by the first new row of the next BCH-matrix M_4. That involves the function z^7, more precisely its first digit in its expression as a bit vector of length k. For $i = 0$ and 1, let

$$W_i := \{[z, z^3, z^5] \in S_{simple} : \text{first digit of } z^7 \text{ is } i\}.$$

The easiest way to ensure that W_0 and W_1 are roughly half of 2^k is if we assume that $z \to z^7$ is a bijection of $\mathbb{F}_{2^k}$, which certainly happens if $7 \not|(2^k - 1)$ or $3 \not|k$.

After all the heuristic, how does one actually prove that this partition is a good choice? The set W_1 was chosen such that its characteristic vector of W_1 is a row of the one larger BCH-matrix (whose columns are the $4k$-vectors $[z, z^3, z^5, z^7]$). We know the Carlitz-Uchiyama bound holds for this matrix as well, maybe with a worse constant factor in the error term, but who cares: any linear combination of the rows of M_4

have Hamming-weight roughly 2^{k-1}. In particular, $[\beta, (1, 0 \ldots, 0)]M_4 = \mathbb{1}_{W_1} + \beta^T M_3$ is a linear combination of the rows of M_4. Compared to $\beta^T M_3$, the W_0-entries do not change, but all the W_1-entries flip, hence the Hamming-weight of this vector is $x + |W_1| - y = x + 2^{k-1} - 1 - y$. This is roughly 2^{k-1} by the Carlitz-Uchiyama bound, so $x \approx y$. Or more precisely, $|x - y| \leq 3 \cdot 2^{k/2}$.

We already know that $x + y \approx 2^{k-1} \pm 2 \cdot 2^{k/2}$, so we have that $x \approx y \approx 2^{k-2} \pm 4 \cdot 2^{k/2}$. Hence $||W_0| - 2x|$ and similarly $||W_1| - 2y|$ is $O(2^{k/2})$. meaning that $\Phi(S) = O(2^k) = \Theta(\sqrt{d})$, quasi-randomness at its best.

In conclusion we have shown that G is a triangle free graph with $\alpha(G) \leq n\Theta(\sqrt{d})/d = \Theta(n^{2/3})$. This establishes $ExpR(3, k) \geq ck^{3/2}$.

Can one improve further by making the graph denser? Maybe even up to the vicinity of the truth, $\Theta(n^2/\log^2 n)$? The next subsection shows that to be impossible in very strong sense.

Exercise 6.4 *Prove that the number of C_4 in Alon's (n, d, λ)-graph is asymptotically the same as the expected number of C_4 in the random graph $G(n, n^{-1/3})$.*

6.1.3 Turán property of quasi-random graphs

Exercise 6.5 *Prove that the independence of both of the previous constructions do have independence number $\Theta(n^{2/3})$. Even more, any K_3-free d-regular graph with $d = \Theta(n^{2/3})$ has an independent set of order $\Theta(n^{2/3})$.*

In the next section we will see that Alon's construction is best possible in a much stronger sense. There is not only one triangle in an $(n, d, \Theta(\sqrt{d}))$-graph with $d = \Omega(n^{2/3})$ but also there are so many triangles that one needs to delete half of the edges of the graph to kill of them.

The quasi-randomizable proof of Turán's theorem

There are many proofs of Turán's Theorem available (see the paper of Aigner [?] for six of them). The main difficulty in generalizing these classical arguments that they are all very much tailored to the complete graph K_n. Here we need an approach that uses only the quasi-random property of K_n, that is, that the edges are "distributed sufficiently evenly", a property shared by all (n, d, λ)-graphs with the appropriate parameters.

Our strategy will be the following.

1. We give a new(?) proof of Turán's theorem and identify the quasi-random properties that make it work.

2. Prove that our quasi-random graphs have these properties.

3. Prove that any graph having these properties are t-Turán.