A Survey on Graphs Decomposable into Induced Matchings

Juncheng Wan, 120033910148, jorgenwan@gamil.com

January 10th, 2022

1 Introduction

A graph G an (r, t)-Ruzsa-Szemerédi graph graph ((r, t)-RS graph for short) if its edge set can be partitioned into t edge-disjoint induced matchings, each of size r. Such graphs consisting of amounts of large pairwise edge-disjoint induced matchings have found several applications in combinatorics, complexity theory, and information theory.

In this work, I briefly introduce some results about graphs decomposable into induced matchings. Besides, I will also present some of my thoughts.

2 Definition

In this section, I clarify basic concepts and definitions.[1]

Definition 1 ((r, t)-Ruzsa-Szemerédi graph)**.** *A graph G an (r, t)-Ruzsa-Szemerédi graph graph ((r, t)-RS graph for short) if its edge set can be partitioned into t edge-disjoint induced matchings, each of size r.*

The problem of the decomposing graph into subgraphs is closely related to the problem of set packing.

Definition 2 ((v, k, t)-packing)**.** *Let X be a v-element set, $X = \{1, 2, \ldots, v\}$. A $P \subset \binom{X}{k}$ is called a (v, k, t')-packing if $|P \cap P'| < t$ holds for every pair $P, P' \in P$.*

Note that in the paper [8], a (v, k, t')-packing is also called "r-sparse".

Definition 3 ($(v, k, \mathcal{H})$-packing)**.** *Let $\mathcal{H}$ be a family of t'-sets on $\{1, 2, \ldots, k\}$. A family $\mathcal{F}$ of k-sets on v elements is called a $(v, k, \mathcal{H})$-packing if for all $F \in \mathcal{F}$ there is a copy of $\mathcal{H}, \mathcal{H}_F$ such that the t-sets of F corresponding to $\mathcal{H}_F$ are covered only by F.*

By counting the number of t-sets, we have $|\mathcal{F}| \le \begin{pmatrix} v \\ t' \end{pmatrix} / |\mathcal{H}|$.

From my understanding, the problem of (r, t)-RS graph is a special case of the $(v, k, \mathcal{H})$-packing problem, where $v = N$, $t' = 2$, $k = 2r$, and $\mathcal{H} = \{\{2k-1, 2k\} : 1 \le k \le r\}$.

[1] Though some concepts are fundamental for students of mathematics, they are fresh for me. Thus, I also write them down.

3 Main Results

In this section, I collect some results. My work is more like a research porter. From my perspective, there are three problems people care for (r,t)-RS graph:

1. A general question: what kind of values (r,t) make (r,t)-RS graph exists?
2. Is it possible for the graph to be dense and r large in the polynomial?
3. For r linear in N, what is the maximum value of r?

3.1 Values (r,t) Making (r,t)-RS Exist

For the first problem, there is a negative answer.

Theorem 1 (Ruzsa and Szemerédi, 1978 [9]). *There is no N-vertex (r,t)-RS graph for r, t both linear in N.*

They proved that this result implies the celebrated theorem of Roth [10], that a subset S of $[N]=\{1,\cdots,N\}$ without nontrivial 3 -term arithmetic progressions have size at most $o(N)$.

However, they prove that there is (r,t)-RS graph with sufficient large r and t.

Theorem 2 (Ruzsa and Szemerédi, 1978 [9]). *There is N-vertex (r,t)-RS graph for $r=N/e^{O(\sqrt{\log N})}$, $t=N/3$.*

Though the construction in the proof of this theorem provides rather dense graphs, but still ones in which the number of edges is $o\left(N^2\right)$. I want to see this paper but can't find it on the Internet.

For the first problem, there is a positive answer, which is useful for applications in complexity theory and information theory. It is from $(v,k,\mathcal{H})$-packing problem.

Theorem 3 (Rödl, 1985 [8]). *For every fixed k,t, whenever $v\longrightarrow\infty$, we have:*

$$\max\{|P|: P \textit{ is a}(v,k,t)\textit{-packing }\}=(1-o(1))\binom{v}{t}/\binom{k}{t} \tag{1}$$

Theorem 4 (Frankl and Füredi, 1987 [6]). *For every fixed k and $\mathcal{H}$, the size of the largest $\mathcal{H}$-packing is $(1-o(1))\binom{v}{t'}/|\mathcal{H}|$, whenever $v\to\infty$.*

From Theorem 4 and the fact that the problem of (r,t)-RS graph is a special case of the $(v,k,\mathcal{H})$-packing problem, we have the following result with $v=N$, $t'=2$, $k=2r$, and $\mathcal{H}=\{\{2k-1,2k\}:1\le k\le r\}$.

Theorem 5. *For any fixed r, there are (r,t)-RS graphs on N vertices with $rt=(1-o(1))\binom{N}{2}$ edges.*

To prove that induced matchings of size larger than $\Theta(\log N)$, the techniques in the original paper don't work. In 2002, Fischer et al [4] construct graphs whose matchings are of linear size.

There is another construction results for problem 1.

Theorem 6 (Birk, Linial and Meshulam, 1993 [3]). *There is N-vertex (r,t)-RS graph, in which $r = (\log N)^{\Omega\left(\log\log N/(\log\log\log N)^2\right)}$, $t = N^2/24r$.*

The construction in this theorem is for the application in [3]. It is crucial to obtain graphs with positive density. Thus, their number of edges is about $N^2/24$. The method here relies on a construction of a low degree representation of the OR function, due to Barrington, Beigel and Rudich [2]. The application in [3] is in information theory, the graphs are applied to design an efficient deterministic scheduling scheme for communicating over a shared directional multichannel.

I want to understand the proof, but the paper is hard with communication theory. I leave it for future learning.

3.2 Dense (r,t)-RS with r Polynomial in N

The problem 2 people care about is for some applications-especially ones in which there is a tradeoff between the number of missing edges and the number of induced matchings needed to cover the graph. The question is they want to find a (r,t)-RS graph with the following two properties:

1. Have positive density rt/N^2 asymptotically;
2. t is polynomially large in N.

None of these constructions satisfy the above properties and solve problem 2.

Meshulam conjectured that there were no such graphs. In 2013, Alon, Moitra and Sudakov [1] disprove this conjecture in the strongest possible sense. They construct graphs with density $1-o(1)$ and yet r is nearly linear in N.

Theorem 7 (Alon, Moitra and Sudakov, 2013 [1]). *There are (r,t)-RS graphs on N vertices with $rt = (1-o(1))\binom{N}{2}$, and $r = N^{1-o(1)}$.*

I think it is exciting. From this construction, not only can we have graphs with positive edge density which are edge-disjoint unions of induced matchings of size $N^{\Omega(1)}$, but in fact, we can have edge density $1-o(1)$, where the size of each matching is $N^{1-o(1)}$.

3.3 Maximum t when r linear in N

For problem 3, to find the maximum value of t when r linear in N, there are two works. The first one is by Fischer et al [4].

Theorem 8 (Fischer et al, 2002 [4]). *There is N-vertex (r,t)-RS graph for $r = N/3 - o(N)$, $t = N^{\Omega(1/\log\log N)}$.*

This theorem also contains a construction. The matchings here are of linear size, but their number is much smaller than in the original construction of Ruzsa and Szemerédi. The construction here is combinatorial, and Fischer et al. use these graphs to establish an $N^{\Omega(1/\log\log N)}$ lower bound for testing monotonicity in general posets.

I am interested in this construction and present the proof in the following.

Proof of Theorem 8. Let m, n be two integers where n is divisible by 3 and $n = o(m)$. The vertex set of U is $X = Y = [m]^n$, thus $N = m^n$. We define a family of (partial) matchings on the vertices of U and take the edge-set of the graph to be the union of the edge-sets of these matchings. The matchings are indexed by a family of $\frac{n}{3}$-subsets of $[n]$. Let $T \subseteq [n], |T| = \frac{n}{3}$. Let $p = \frac{n}{3}$.

Definition of a matching M_T. Color the points in the two copies of $[m]^n$ by blue, red and white. The color of a point x is determined by $\sum_{i \in T} x_i$. First, partition the vertex set into levels, where the level L_s is the set $\left\{x : \sum_{i \in T} x_i = s\right\}$. Then combine levels into strips, where for an integer $k = 1 \ldots m$, the strip $S_k = L_{kp} \cup \ldots \cup L_{(k+1)p-1}$. Color the strips S_k with $k \equiv 0(\text{mod}3)$ blue, the strips with $k \equiv 1(\text{mod}3)$ red, and the remaining strips white. The matching M_T is defined by matching blue points in X to red points in Y as follows: If a blue point b in X has all its T-coordinates greater than 2 , match it to a point $r = b - 2 \cdot 1_T$ in Y. The vector 1_T is the characteristic vector of T; it is 1 on T and 0 outside T. Note that r is necessarily red. M_T is clearly a matching. Our next step is to show that it is large.

Claim. $|M_T| \geq N/3 - o(N)$

Proof of the claim. Consider the "projected" matching M on the vertices of the bipartite graph $U^T = \left([m]^T, [m]^T\right)$, which is defined by T. Namely, partition the points of $[m]^T$ as described above, coloring them by blue, red and white, and match a blue point in one copy of $[m]^T$ to a red one in another, by subtracting $2 \cdot 1_T$. Since M_T is determined by the coordinates in T, it is enough to show that $|M| \geq P/3 - o(P)$, where $P = m^p$.

Let $B, R, W \subseteq [m]^T$ be the sets of the blue, red and white points, respectively. Then $P = |B| + |R| + |W|$.

First, we claim that $|W| \leq |R| + |\{x : \exists i, x_i = 1\}|$. Indeed, consider a new matching between W and R defined by matching $w \in W$ to $w - 1_T$. Assume that $m \equiv 0(\text{mod}3)$. Then the only unmatched points in W are contained in the set $\{x : \exists i, x_i = 1\}$, proving this claim. Similarly $|W| \leq |B| + |\{x : \exists i, x_i = m\}|$. Next, observe that the only blue and red points (in the corresponding copies of $[m]^T$) unmatched by M are these which have a coordinate whose value is in $\{1, 2, m-1, m\}$. It follows that $|M| > (|R| + |B|)/2 - |\{x : \exists i, x_i \in \{1, 2, m-1, m\}\}| > P/3 - \left(|\{x : \exists i, x_i \in \{1, 2, m-1, m\}\}| + |\{x : \exists i, x_i = 1, m\}|\right) \geq P/3 - \frac{6p}{m} \cdot P$. Since $p = o(m)$, the claim holds.

Now, let T, T_1 be two $\frac{n}{3}$-sets in $[n]$, such that $|T \cap T_1| \leq n/7$. We claim that no edge of M_T is induced by M_{T_1}. Indeed, let b be matched to r by M_T, in particular $b - r = 2 \cdot 1_T$. If the edge (b, r) is induced by M_{T_1}, then b is colored blue and r is colored red in the coloring defined by T_1. By the definition of the coloring, since $\sum_{i=1}^n b_i > \sum_{i=1}^n r_i, b$ is located in a blue level separated by a white level from the red level of r. This implies that $\left|\sum_{i \in T_1} b_i - \sum_{i \in T_1} r_i\right| \geq \frac{n}{3}$. On the other hand, $\left|\sum_{i \in T_1} b_i - \sum_{i \in T_1} r_i\right| = \left|\sum_{i \in T_1}^i (b_i - r_i)\right| = \left|\sum_{i \in T_1} (2 \cdot 1_T)_i\right| = 2 \cdot |T \cap T_1| \leq \frac{2n}{7} < \frac{n}{3}$, reaching a contradiction.

We would like to have a large family $\mathcal{F}$ of $\frac{n}{3}$-subsets of $[n]$, such that the intersection between any two of them is of size at most $\frac{n}{7}$, or, equivalently, such that the Hamming distance between any two of them is at least $\frac{2n}{3} - \frac{2n}{7} = \frac{8n}{21}$. So we need a lower bound on the size of a constant weight binary error-correcting code $\mathcal{F}$ with the following parameters: block length n, weight $w = \frac{n}{3}$, distance $d = \frac{8n}{21}$. Applying the Gilbert-Varshamov bound for constant weight codes [7], we get $\frac{1}{n} \log |\mathcal{F}| \geq H(1/3) - 1/3 \cdot H(4/7) - 2/3 \cdot H(2/7) - o(1) = 0.014 - o(1)$. Choose $m = n^2$ and define the edge-set $E(U)$ of U by $E(U) = \bigcup_{T \in \mathcal{F}} M_T$. By the preceding discussion, U is a graph on $N = n^{2n}$ vertices, whose edge-set is a disjoint union of $2^{\Omega(n)} = N^{\Omega\left(\frac{1}{\log \log N}\right)}$ induced matchings of size $N/3 - o(N)$. □

The above construction also gives $(cN, N^{\Omega(1/\log \log N)})$-RS graphs for $c \leq 1/4$. Alon notices that when $c > 1/4$, t can only be constant. This is also obvious by the following Theorem 9.

The followings are works from [5]. Fox, Huang and Sudakov discuss several situations of the constant c, for $c > 1/4$, $c = 1/4$ and $c < 1/4$ respectively.

Honestly, I only understand parts of the paper, from the start to Theorem 3.4. Thus, I mainly discuss these parts I understand.

3.3.1 $c > 1/4$ for $r = cn$

Theorem 9. *Suppose G is an (r,t)-RS graph on N vertices, then*

$$r \leq \begin{cases} \frac{N}{4}\left(1+\frac{1}{t}\right) & \text{if } 2 \nmid t \\ \frac{N}{4}\left(1+\frac{1}{t+1}\right) & \text{if } 2 \mid t \end{cases} \tag{2}$$

Moreover, these bounds are tight for every positive integer t and infinitely many N.

Proof of Theorem 9. Suppose the edge set of G can be partitioned into induced matchings $M_1, \cdots, M_t$, each containing exactly r edges. Denote by V_i the set of vertices contained in the edges of M_i. Then $|V_i| = 2r$. Moreover, each of the r edges of M_i intersects V_j in at most one vertex, since otherwise V_i and V_j must span a common edge of G. This implies that $|V_i \cap V_j| \leq r$. Let $v_i \in \{0,1\}^N$ be the characteristic vector of V_i. Then for all $1 \leq i < j \leq t$, the Hamming distance satisfies

$$\operatorname{dist}(v_i, v_j) = |V_i| + |V_j| - 2|V_i \cap V_j| \geq 2r + 2r - 2r = 2r.$$

This is already enough to show that t is constant, using bounds from coding theory. But one can do slightly better. Let v_0 be the all-zero vector. To get a tight upper bound, notice that the above inequality can be extended to all $0 \leq i < j \leq t$ since, for $1 \leq i \leq t, |V_i| = 2r$. Denote by a_i (resp. b_i) the number of vectors v_j equal to 0 (resp. 1) in the i-th coordinate, then $a_i + b_i = t + 1$. By double counting,

$$\begin{aligned} 2r\binom{t+1}{2} &\leq \sum_{0 \leq i < j \leq t} \operatorname{dist}(v_i, v_j) \\ &= \sum_{i=1}^{N} a_i b_i \\ &\leq \begin{cases} N(t+1)^2/4 & \text{if } 2 \nmid t \\ Nt(t+2)/4 & \text{if } 2 \mid t \end{cases} \end{aligned}$$

The last inequality follows from that $a_i b_i$ is maximized when $a_i = b_i = (t+1)/2$ for odd t, and $\{a_i, b_i\} = \{(t+2)/2, t/2\}$ for even t. By simplifying the inequality we immediately obtain Equation 2.

To show that the bound is tight, it suffices to consider the case $t = 2k+1$. Let H be $KG(2k+1, k)$, the Kneser graph whose vertices correspond to all the k-subsets of a set of $2k+1$ elements, and where two vertices are adjacent if and only if the two corresponding sets are disjoint. We define the matchings $M_1, \cdots, M_{2k+1}$ in the following way: the edge (A, B) belongs to M_i if and only if $A \cap B = \emptyset$ and $A \cup B = [2k+1] \backslash \{i\}$. It is easy to see that B is determined after fixing A and i, which implies that M_i forms a matching. In order to show that every matching is induced in H, we take (A, B) and (C, D) both from M_i with $A \neq C, D$, then $A \cup B = C \cup D = [2k+1] \backslash \{i\}$, it is not hard to check that $A \cap C, A \cap D$ are both nonempty and therefore (A, C) and (A, D) are not contained in any M_j. By calculation, $N = \binom{2k+1}{k}$, while $r = \frac{1}{2}\binom{2k}{k} = \frac{N}{4}\left(1+\frac{1}{2k+1}\right)$. Hence H is a $\left(\frac{N}{4}\left(1+\frac{1}{2k+1}\right), 2k+1\right)$-RS graph on N vertices. □

During the proof, the sentence "This is already enough to show that is constant, using bounds from coding theory" is not obvious for me. Thus, I read the proof following the sentence, where the most important is the double counting and the conditions for quadratic function taking the maximum.

For the construction of the upper bound, it utilizes the seperating property of the Kneser graph. Note that $KG(2k+1,k)$ has $\begin{pmatrix} 2k+1 \\ k \end{pmatrix}$ vertices and $\begin{pmatrix} k+1 \\ k \end{pmatrix} = k+1$ edges.

3.3.2 $c = 1/4$ for $r = cn$

Proposition 1. *There exists $(N/4, 2\log_2 N)$-RS graph for every integer N that is a power of 2.*

The following proof is based on hypercube graph. It utilizes the parity of the vector representation of vertices in hypercube.

Proof. Let $k = \log_2 N$ and consider the k-dimensional hypercube graph H with vertex set $\{0,1\}^k$, where two vectors are adjacent if their Hamming distance is 1 . We first partition the vertices into odd and even vectors, according to the parity of the sum of their coordinates. For $1 \leq i \leq k$, we let the i-th matching M_i consist of edges between vectors $\vec{v}$ and $\vec{v} + \vec{e}_i$, such that $\vec{v}$ is even and its i-th coordinate is 0 ; and the $(k+i)$-th matching M_{k+i} consist of edges between an odd vector $\vec{v}$ whose i-th coordinate equals 0 and the vector $\vec{v} + \vec{e}_i$. This construction gives $2k = 2\log_2 N$ matchings, and obviously each matching involves exactly half of the vertices. In order to verify that the matchings are induced, we consider two distinct edges from M_i, which are $(\vec{u}, \vec{u} + \vec{e}_i)$ and $(\vec{v}, \vec{v} + \vec{e}_i)$, such that both $\vec{u}$ and $\vec{v}$ are even and their i-th coordinates are 0 . Clearly the pairs $(\vec{u}, \vec{v})$ and $(\vec{u} + \vec{e}_i, \vec{v} + \vec{e}_i)$ cannot form edges of H since they have the same parity. Moreover, $\vec{u}$ and $\vec{v} + \vec{e}_i$ (similarly, $\vec{v}$ and $\vec{u} + \vec{e}_i$) differ in at least two coordinates. Therefore for all $1 \leq i \leq k$, the matchings M_i we defined are induced. Using a similar argument, we can also show that the matchings M_{i+k} are induced. □

The above construction can be improved When $\log_2 n$ is an even integer. We can add two additional induced matchings. The first one consists of $(\vec{u}, \vec{v})$ where $\vec{u} + \vec{v} = \overrightarrow{1}$ and both are even vectors. The second matching contains all edges $(\vec{u}, \vec{v})$ where $\vec{u} + \vec{v} = \overrightarrow{1}$ and both are odd vectors. This gives the following corollary.

Corollary 1. *There exists $(N/4, 2(\log_2 N + 1)) - RS$ graphs on N vertices for every N that is an even power of 2 .*

Lemma 1. *If G is an $(r,t) - RS$ graph on N vertices, then its bipartite double cover $G \times K_2$ is an $(2r,t) - RS$ graph on $2N$ vertices.*

Proof. Denote by $G' = G \times K_2$ the bipartite double cover of the graph G. The vertices of G' are (v,i) with $v \in V(G)$ and $i \in \{0,1\}$. Two vertices $(u,0)$ and $(v,1)$ are adjacent whenever u and v form an edge in G. Note that an induced matching $M_i = \{(u_j, v_j)\}_{j=1}^{r}$ in G corresponds to a matching $M' = \{((u_j,0),(v_j,1))\}_{j=1}^{r} \cup \{((u_j,1),(v_j,0))\}_{j=1}^{r}$ in G', which is of size $2r$. It is also straightforward to check by definition that M' is an induced matching. Therefore G' is an $(2r,t)$-RS graph on $2N$ vertices, □

Theorem 10. *If G is an $\left(\frac{N}{4}, t\right)$-RS graph on N vertices, then $t \leq 8(\log_2 N + 1)$.*

Proof. From Lemma 1, it suffices to show that for all N-vertex bipartite graphs G whose edges can be decomposed into induced matchings $M_1, \cdots, M_t$, each of size $N/4, t$ is at most $8\log_2 N$.

Denote by d_v the degree of vertex v in G. We consider the subgraph H of G, with edges being the pair of vertices u, v such that $d_u + d_v \geq t$ and $(u,v) \in E(G)$. Note that $e(G) = \frac{Nt}{4}$. By the Cauchy-Schwarz inequality,

$$\sum_{(u,v)\in E(G)} (d_u + d_v - t) = \left(\sum_{v\in V(G)} d_v^2 \right) - t \cdot e(G) \geq N \left(\frac{\sum_{v\in V(G)} d_v}{N} \right)^2 - \frac{Nt^2}{4}$$
$$= N \left(\frac{Nt/2}{N} \right)^2 - \frac{Nt^2}{4} = 0.$$

For any edge $(u,v) \in E(G)$, all the edges incident to either u or v must belong to different matchings. Therefore $d_u + d_v \leq t+1$. If we denote by E_i the number of edges (u,v) such that $d_u + d_v = t + i$, we have $E_1 + E_0 + E_{-1} + \cdots + E_{-t} = Nt/4$, while inequality (3) implies that $E_1 - \sum_{j=1}^{t} jE_{-j} \geq 0$. Summing these two inequalities gives $2E_1 + E_0 \geq Nt/4$ and so $E_1 + E_0 \geq Nt/8$. So H is a N-vertex graph with at least $Nt/8$ edges and thus its average degree is at least $t/4$. Hence, H has a subgraph F of minimum degree at least $t/8$.

Set $s = t/8$. For each vertex v of G, let A_v denote the set of induced matchings containing v. Clearly $|A_v| = d_v$. We claim that if v and u are at distance k in F, then when k is odd, $|A_u \cap A_v| \leq k$; and when k is even, $|A_u \cap A_v^c| \leq k$. This statement can be proved using induction. The base cases when $k = 0$ and 1 are obvious. Now we assume that it is true for all $k \leq i$. For $k = i+1$, suppose u and v are at distance k. Let w be a vertex at distance 1 from v and $i = k-1$ from u. When i is odd, from the inductive hypothesis we have $|A_v^c \cap A_w^c| = t - |A_v \cup A_w| = t - |A_v| - |A_w| + |A_v \cap A_w| \leq 1$ and $|A_w \cap A_u| \leq i$. Therefore,

$$|A_u \cap A_v^c| \leq |A_u \cap A_w| + |A_w^c \cap A_v^c| \leq i + 1.$$

Similarly, when i is even, we have $|A_v \cap A_w| \leq 1$ and $|A_w^c \cap A_u| \leq i$, and hence

$$|A_u \cap A_v| \leq |A_v \cap A_w| + |A_w^c \cap A_u| \leq i + 1.$$

Now choose an arbitrary vertex v in F, the degree of v in F is at least $s = t/8$. For every integer $i \geq 0$, let N_i be the set of vertices at distance i from v in graph F. By the assumption that G is bipartite, each N_i induces an independent set. We denote by e_i the number of edges of F that are between N_i and N_{i+1} and contained in matchings in A_v (resp. A_v^c) when i is odd (resp. even). For odd i, we estimate the number of edges of F between N_i and N_{i+1} that are contained in matchings in A_v^c in two different ways. Since every vertex $u \in N_i$ is contained in at least $s - |A_u \cap A_v| \geq s - i$ such edges, and every vertex $w \in N_{i+1}$ is contained in no more than $|A_w \cap A_v^c| \leq i+1$ such edges, we have

$$(s-i)\,|N_i| - e_{i-1} \leq (i+1)\,|N_{i+1}| - e_{i+1}.$$

Similarly, when N is even, by bounding the number of edges between N_i and N_{i+1} that belong to matchings in A_v, we obtain the same inequality as above. Summing up the inequalities for $i = 0, \cdots, k$, we have

$$\sum_{i=0}^{k} (s-i)\,|N_i| - \sum_{i=0}^{k-1} e_i \leq \sum_{i=1}^{k+1} i\,|N_i| - \sum_{i=1}^{k+1} e_i.$$

Simplifying this inequality gives

$$(k+1)\,|N_{k+1}| \geq \sum_{i=0}^{k} (s-2i)\,|N_i| - e_0 + e_{k+1} + e_k \geq \sum_{i=0}^{k} (s-2i)\,|N_i|\,.$$

The second inequality follows from the observation that $e_0 = 0$ since all edges between N_0 and N_1 are in A_v.

In the next step, we prove by induction that $|N_i| \geq \binom{s}{i}$. For $i = 0$ and 1 this is obvious. Now, assuming it is true for all $i \leq k$, we have

$$
\begin{aligned}
|N_{k+1}| &\geq \frac{1}{k+1}\sum_{i=0}^{k}(s-2i)\,|N_i| \geq \frac{1}{k+1}\sum_{i=0}^{k}(s-2i)\binom{s}{i} \\
&= \frac{1}{k+1}\left(\sum_{i=0}^{k}(s-i)\binom{s}{i} - \sum_{i=0}^{k} i\binom{s}{i}\right) \\
&= \frac{1}{k+1}\left(\sum_{i=0}^{k} s\binom{s-1}{i} - \sum_{i=0}^{k} s\binom{s-1}{i-1}\right) \\
&= \frac{s}{k+1}\binom{s-1}{k} = \binom{s}{k+1}
\end{aligned}
$$

Note that the number of vertices in $N_0 \cup N_1 \cup \cdots \cup N_s$ is at most N. We therefore have

$$
N \geq \sum_{k=0}^{s} |N_i| \geq \sum_{k=0}^{s}\binom{s}{k} = 2^s,
$$

Solving this inequality gives $s \leq \log_2 N$ and hence $t \leq 8\log_2 N$. □

The best result the paper gets for $c = 1/4$ is the following theorem. As I don't fully understand it, I just paste it here.

Theorem 11. *If an N-vertex graph G is an $(N/4, t) - RS$ graph, then $t \leq (6 + o(1))\log_2 N$.*

3.3.3 $c < 1/4$ for $r = cN$

Similarly, as I don't read through the case $c < 1/4$, I just paste the results of the paper here.

Theorem 12. *For every $\varepsilon > 0$, if G is an (r,t)-RS graph on N vertices with $r = cN$ for $1/5 + \varepsilon \leq c < 1/4$, then $t = O(N/\log N)$.*

Theorem 13. *There exists an absolute constant $b > 0$, such that for $r \geq (1/4 - b)N$, if G is an $(r,t) - RS$ graph G on N vertices, then $t = N/\left((\log N)2^{\Omega(\log^* N)}\right) = o(N\log N)$.*

4 Thoughts for Conjecture 1

There are several conjectures in [5]. I mainly think about the following one.

Conjecture 1 When the size of matchings is close to $n/4$, there are two quite different constructions of Ruzsa-Szemerédi graphs. One is the Kneser graph $KG(2k+1,k)$ with $k \sim \frac{1}{2}\log_2 n$, which is an $(n/4 + \Theta(n/\log n), (1+o(1))\log_2 n)$-RS graph. The other is the hypercube $\{0,1\}^{\log_2 n}$, which is an $(n/4, 2\log_2 n)$-RS graph. Can we find a family of (r,t)-RS graphs that bridge between these two examples, say with $r = c\log_2 n$ for some $c \in [1,2]$, and $t - n/4 = \Omega(\log n)$?

Thoughts for Conjecture 1 What if I expand the definition of the Kneser graph? In Kneser graph $KG(2k+1,k)$, vertices correspond to all the k-subsets of a set of $2k+1$ elements, and where two vertices are adjacent if and only if the two corresponding sets are disjoint.

Define $KG(2k+1,k,g,h)$, vertices correspond to all the $(k+g)$-subsets of a set of $2k+1$ elements, and where two vertices are adjacent if and only if the two corresponding sets have the intersection of exact h elements. But this construction can not meet the following properties:

1. When $g=k+1$, $h=2k$, $KG(2k+1,k,k+1,2k)$ is exact the hypercube of dimension $(k+1)$.
2. When $g=0$, $h=0$, $KG(2k+1,k,0,0)$ is degenrated to the Kneser graph $KG(2k+1,k)$.

Thus, $KG(2k+1,k,g,h)$ is not a transition or interpolation from kneser to hypercube.

Then, I calculate the density of the two graphs. Assume that n is the dimension of the hypercube. I want to fix the n for hypercube and Kneser. Because both of their vertices can be treated as a kind of coding in the space $\{0,1\}^n$. I want n fixed during the transition. Then, we have the ratio:

1. #edge/#vertex of hypercube is $\frac{n2^n/2}{2^n}=\frac{n}{2}$;
2. #edge/#vertex of Kneser is $\frac{\binom{n}{k}k/2}{\binom{n}{k}}=\frac{n-1}{4}$.

Thus, I want to make the density transits from $\frac{n}{2}$ to $\frac{n-1}{4}$ by deleting edges in hypercube or adding edges in Kneser. First, I ask the following question:

Question 1: What happen if I forbidden the edge generated by the difference in the first dimension connection rule in hypercude? For example, when $n=3$, vertices like $(0,1,0)$ and $(1,1,0)$ will not connect.

In this way, the hypercube becomes a $(n-1)$-regular graph. As the construction of matching in the Proposition 1, we know that M_0 and M_k (k is n here) are no longer matching. Because within these induced subgraphs, there are no edges. Thus, t goes from $2\log_2 n$ to $(2\log_2 n-2)$. In this way, for M_i, we don't need to restrict all $\vec{v}$ to be odd (or even). We just need to restrict that all $\vec{v}_{\geq 2}$ to be even (or odd), where $\vec{v}_{\geq 2}$ means the part of $\vec{v}$ from dimension 2. However, in this way, there are also $t=n/4$ without any increase. I think the construction is based on parity, which is not suitable for the "forbidding the edge generated by $\leq i$-th dimension". I think it is due to the construction is very specific for the hypercube.

Then, I ask the next question:

Question 2: Can we paste the Kneser to the hypercube?

Here, "canonically" pasting is no use. "canonically" means treating the vertex vector (k elements are 1 and the rest are 0) in the Kneser as the vertex in hypercube (with dimension $n=2k+1$). It is obvious that Kneser and hypercube have no edge intersection and the edges of Kneser will destroy the construction of the hypercube.

Other pasting methods, such as pasting partial of Kneser to hypercube, seem no use. It is hard.

5 Conclusion

In this survey, I do nothing but collect others' results about induced matchings, read their proofs, and make my failed attempts. I think I need to read more constructions to get more familiar to the (r, t)-RS graph.

References

[1] Noga Alon, Ankur Moitra, and Benjamin Sudakov. "Nearly complete graphs decomposable into large induced matchings and their applications". In: *Journal of the European Mathematical Society* 15.5 (2013), pp. 1575–1596.

[2] David A Mix Barrington, Richard Beigel, and Steven Rudich. "Representing Boolean functions as polynomials modulo composite numbers". In: *Computational Complexity* 4.4 (1994), pp. 367–382.

[3] Yitzhak Birk, Nathan Linial, and Roy Meshulam. "On the uniform-traffic capacity of single-hop interconnections employing shared directional multichannels". In: *IEEE Transactions on Information Theory* 39.1 (1993), pp. 186–191.

[4] Eldar Fischer et al. "Monotonicity testing over general poset domains". In: *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*. 2002, pp. 474–483.

[5] Jacob Fox, Hao Huang, and Benny Sudakov. "On graphs decomposable into induced matchings of linear sizes". In: *Bulletin of the London Mathematical Society* 49.1 (2017), pp. 45–57.

[6] P Frankl and Z Füredi. "Colored packing of sets". In: *North-Holland Mathematics Studies*. Vol. 149. Elsevier, 1987, pp. 165–177.

[7] VI Levenšteın. "Upper bounds for codes with a fixed weight of vectors". In: *Problemy Peredaci Informacii* 7 (1971), pp. 3–12.

[8] Vojtěch Rödl. "On a packing and covering problem". In: *European Journal of Combinatorics* 6.1 (1985), pp. 69–78.

[9] Imre Z Ruzsa and Endre Szemerédi. "Triple systems with no six points carrying three triangles". In: *Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai* 18 (1978), pp. 939–945.

[10] Tom Sanders. "On Roth's theorem on progressions". In: *Annals of Mathematics* (2011), pp. 619–636.