GRAPH EMBEDDING LECTURE NOTE

JAEHOON KIM

Abstract. In this lecture, we aim to learn several techniques to find sufficient conditions on a dense graph G to contain a sparse graph H as a subgraph. In particular, for a given graph H, we investigate conditions on $e(G)$ or $\delta(G)$ to guarantee an embedding of H into G.

1. Basic probabilistic methods

Materials in this lecture note come from several sources which are listed in the reference at the end of the lecture note.

For each $n \in \mathbb{N}$, we define $[n] := \{1, 2, \dots, n\}$. We write $a = b \pm c$ if we have $b - c \leq a \leq b + c$. We define $\binom{X}{k} := \{A \subseteq X : |A| = k\}$ and $\binom{X}{\leq k} := \{A \subseteq X : |A| \leq k\}$. If a set A has size k, then we say that A is a *k-set*.

A tuple $H = (V, E)$ is a *k-uniform hypergraph* or *k-graph* if $V = V(H)$ is a finite set and $E = E(H) \subseteq \binom{V}{k}$. In particular, if $G = (V, E)$ is a 2-graph, then we call it a *graph*. For k-graph H, we let $v(H) := |V(H)|$ and $e(H) := |E(H)|$. For graph G and $u, v \in V$, we denote the edges $\{u, v\} \in E(G)$ as uv or vu. If $uv \in E(G)$, we say that u and v are *adjacent* and v is a neighbor of u. We say that a set $U \subseteq V(G)$ is an *independent set* if there are no edges uv of G with $\{u, v\} \subseteq U$. We say that a set U is a *clique* of G if every pair uv in U is an edge of G.

For a graph G and two vertices u and v, we let $\mathrm{dist}_G(u, v)$ be the minimum length of paths between u and v, where the length of a path is the number of edges in the path. For sets $U = \{u_1, \dots, u_t\} \subseteq V(G)$, $A \subseteq V(G)$ and $k \in \mathbb{N}$, we write

$$
\begin{aligned}
N_G(U; A) &:= N_G(u_1, \dots, u_k; A) := A \cap \bigcap_{v \in U} N_G(v) \\
N_G^k(U) &:= \{v \in V(G) : \mathrm{dist}(v, u) = k \text{ for some } u \in U\} \\
N_G^{\leq k}(U) &:= \bigcup_{i=0}^{k} N_G^k(U) \\
N_G(U) &:= N_G(u_1, \dots, u_k) := N_G(U; V(G)) \text{ and } d_G(U; A) := |N_G(U; A)|.
\end{aligned}
$$

Sometimes, we omit the subscripts if it's clear from the context. Note that $N_G(U)$ denotes the set of vertices adjacent to all vertices in U while $N_G^1(U)$ denotes the set of vertices adjacent to at least one vertex in U.

We denote *the maximum degree* of G as $\Delta(G) = \max_{v \in V(G)} d_G(v)$ and *the minimum degree* of G as $\delta(G) = \min_{v \in V(G)} d_G(v)$ and *the average degree* of G as $d(G) = \frac{1}{v(G)} \sum_{v \in V(G)} d_G(v) = \frac{2e(G)}{n}$. We say that a graph G is an *r-partite graph* or an *r-chromatic graph* if $V(G)$ can be partitioned into r independent sets. We say that a partition $V_1, \dots, V_r$ of $V(G)$ is an *r-partition* if each V_i is an independent set. $\chi(G)$ denote the minimum natural number r where G is r-partite. We say that G is a *complete r-partite graph* with vertex partition $V_1, \dots, V_r$ if uv is an edge for all $u \in V_i, v \in V_j$ and $i \neq j$.

Definition 1.1. *We say that H embeds into G, or equivalently G contains H as a subgraph if there exists an injective function $\phi : V(H) \to V(G)$ satisfying $\phi(u)\phi(v) \in E(G)$ for each*

$uv \in E(H)$. We say ϕ is an embedding of H into G. If such an embedding exists, we write $H \subseteq G$.

We will frequently use the probabilistic method in this lecture. The probabilistic method is a powerful tool to prove an existence of a structure with certain properties. The basic philosophy for the probabilistic method is the following.

Define an appropriate probability space of structures. Show that the desired properties hold in this space with positive probability. Then this proves the existence of a desired structure.

To illustrate this method, we prove a theorem regarding Ramsey number. The Ramsey number $r(H)$ is the smallest n such that in any 2 edge-coloring of a complete graph K_n on n vertices by red and blue, either H embeds into the red graph or H embeds into the blue graph. Erdős first used this method to prove the following simple theorem.

Theorem 1.2 (Erdős, 1947). *For each $r \in \mathbb{N}$, we have $r(K_r) > 2^{r/2}$.*

Proof. Consider a complete graph K_n on vertex set V with $|V| = n = 2^{r/2}$ and a 2-edge-coloring of K_n by red and blue as follows. For each edge of K_n, we choose its color as red or blue with probability $1/2$ independently at random. For each $U \in \binom{V}{r}$, the set U induces a monochromatic K_r with probability $2 \cdot 2^{-\binom{r}{2}}$. The probability that K_n with the edge-coloring contains a monochromatic K_r is at most

$$\sum_{U \in \binom{V}{r}} 2 \cdot 2^{-\binom{r}{2}} \leq \binom{n}{r} 2^{1-\binom{r}{2}} < 1.$$

Here, it is easy to check the last inequality. This implies that, with positive probability, the edge-coloring contains no monochromatic K_r. Thus there exists an edge-coloring of K_n with no monochromatic K_r. □

Note that in the above, we use the union bound, which says that at least one of the events $E_1, \dots, E_k$ holds with probability at most $\sum_{i=1}^{k} \mathbf{Pr}[E_i]$.

A random variable is a function $X : \mathcal{S} \to \mathbb{R}$ from a probability space $\mathcal{S}$ to real numbers. For a random variable X, we write $\mathbb{E}[X]$ for the expectation of X. For two random variables X and Y, we have $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ and this is called *linearity of expectation*. Note that if $\mathbb{E}[X] = \mu$, then there exists an element $S \in \mathcal{S}$ with $X(S) \geq \mu$. To see this, imagine a set of numbers with average μ, then at least one number in the set must be at least as large as μ.

Another useful inequality we will use is Markov's inequality. Which says that for a non-negative random variable X and $t \geq 1$, we have $\mathbf{Pr}[X \geq t] \leq \frac{\mathbb{E}[X]}{t}$.

By using linearity of expectation, we can prove the following theorem.

Theorem 1.3 (Caro 1979, Wei 1981). *Let $n, r \in \mathbb{N}$ with $n \geq r$ and $n = qr + k$ with $0 \leq k < r$. Any n-vertex graph G with less than $k\binom{q+1}{2} + (r-k)\binom{q}{2}$ edges contains an independent set of size at least $r + 1$.*

Proof. Let m be the number of edges of G. We choose an ordering $\sigma = (v_1, \dots, v_n)$ of $V(G)$ uniformly at random. In other words, each ordering is chosen with the same probability $\frac{1}{n!}$. In this ordering, we choose the set $U(\sigma)$ of all vertices which comes before all their neighbors. Then $U(\sigma)$ is an independent set of G. Let X be the random variable such that $X(\sigma) = |U(\sigma)|$. For each $v \in V(G)$, let I_v be the indicator random variable such that

$$I_v(\sigma) := \begin{cases} 1 & \text{if } v \in U(\sigma), \\ 0 & \text{if } v \notin U(\sigma). \end{cases}$$

We have $X = \sum_{v\in V(G)} I_v$. Since the ordering is chosen uniformly at random, the probability that v appears before all of its neighbors is $\mathbf{Pr}[I_v = 1] = (d(v)+1)^{-1}$.

$$\mathbb{E}[X] = \sum_{v\in V(G)} \mathbb{E}[I_v] = \sum_{v\in V(G)} \frac{1}{d(v)+1}.$$

As $\sum_{v\in V(G)}(d(v)+1) = 2e(G)+n$ is fixed, the final expression above is minimized when degrees of vertices in $V(G)$ are as even as possible. Note that $2(k\binom{q+1}{2} + (r-k)\binom{q}{2}) + n > 2m+n$ is a sum of $(r-k)q$ terms of q and $k(q+1)$ terms of $q+1$, total n terms. Thus we have

$$\mathbb{E}[X] > (r-k)q\cdot\frac{1}{q} + k(q+1)\cdot\frac{1}{q+1} = r.$$

This implies that there exist an ordering σ such that $|U(\sigma)| \geq r+1$, meaning that G contains an independent set of size $r+1$. □

By analyzing the above proof, we can characterize all extremal graphs. In other words, we can characterize all graphs with minimum number of edges containing no independent set of size $r+1$.

Proposition 1.4. *Let $n, r \in \mathbb{N}$ with $n \geq r$ and $n = qr+k$ with $0 \leq k < r$. Let G be an n-vertex graph with $k\binom{q+1}{2} + (r-k)\binom{q}{2}$ edges whose largest independent set has size exactly r. Then G is a vertex-disjoint union of k complete graphs on $q+1$ veritces and $r-k$ complete graphs on q vertices.*

Proof. Note that we have

$$\mathbb{E}[X] = \sum_{v\in V(G)} \mathbb{E}[I_v] = \sum_{v\in V(G)} \frac{1}{d(v)+1} \geq r. \tag{1.1}$$

On the other hand, as G has no independent set of size $r+1$, we must have $\mathbb{E}[X] \leq r$, thus we have $\mathbb{E}[X] = r$. Hence, we have the equality in (1.1). The final equality in (1.1) implies that $d(v)$ must be as close together as possible. Moreover, X must be a constant random variable. (Otherwise, thre exists a choice σ such that $X(\sigma) > \mathbb{E}[X]$.) Suppose that G is not union of cliques. Then there exists $x, y, z \in V$ with $xy, xz \in E(G)$ and $yz \notin E(G)$. Consider orderings $\sigma = (x, y, z, \dots)$ and $\sigma' = (y, z, x, \dots)$ in such a way that they are identical on $V(G)\setminus\{x, y, z\}$. Then $U(\sigma)$ and $U(\sigma')$ are the same except $x \in U(\sigma)$, $y, z \notin U(\sigma)$ and $x \notin U(\sigma')$, $y, z \in U(\sigma')$. Hence, we conclude that $X = |U(\sigma)|$ is not constant, this implies that G is the union of cliques. As degrees of G are as even as possible, we prove the proposition. □

The main theme of this lecture series is to answer the following question.

In order to guarantee an embedding of H into G, how dense G must be?

First natural choice of the graph H is a complete graph. Let $T_r(n)$ be the n-vertex complete r-partite graph such that each part has size either $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$. Let $t_r(n)$ be the number of edges in $T_r(n)$. It is easy to see that $T_r(n)$ does not contain K_{r+1} as a subgraph. Note that by considering a complement of a graph G (which is a graph with vertex set $V(G)$ and edge set $\binom{V(G)}{2}\setminus E(G)$), Theorem 1.3 implies the following theorem.

Theorem 1.5 (Turán, 1941)**.** *Let $n, r \in \mathbb{N}$ with $n \geq r+1$. If G is an n-vertex graph with more than $t_r(n)$ edges, then G contains K_{r+1} as a subgraph.*

Note that $t_r(n) \le (1-\frac{1}{r})\frac{n^2}{2}$. This means that if G is dense enough, then we can embed K_r into G. Moreover, $T_r(n)$ is the only graph having the maximum number of edges containing no copy of K_{r+1}. In other words, we can characterize all K_r-free graphs with the maximum number of edges.

By Proposition 1.4, the following holds.

Theorem 1.6. *Let $n, r \in \mathbb{N}$ with $n \ge r+1$. If G is an n-vertex K_{r+1}-free graph with exactly $t_r(n)$ edges, then G must be isomorphic to $T_r(n)$.*

We will explore more for graphs other than complete graphs.

The following is another useful probabilistic method. When a random variable $X = X_1 + X_2 + \cdots + X_n$ is a sum of mutually independent random variables $X_1, \dots, X_n$, then it's often possible to prove that the value of X is close to its expectation with high probability. To see this, imagine that tossing a fair coin n times with large n, then the number of times we get head is almost surely very close to $n/2$. The following lemma describes this phenomenon.

Lemma 1.7 (Chernoff's inequality, see [2])**.** *Suppose $X_1, \dots, X_n$ are independent random variables such that $\mathbf{Pr}[X_i = 0] = p_i$ and $\mathbf{Pr}[X_i = 1] = 1 - p_i$ for all $i \in [n]$. Let $X := X_1 + \cdots + X_n$. Then for all $t > 0$, $\mathbf{Pr}[|X - \mathbb{E}[X]| \ge t] \le 2e^{-t^2/(2n)}$.*

To illustrate that this lemma is useful, we give a proof of the following easy proposition.

Proposition 1.8. *There exists $n_0 \in \mathbb{N}$ satisfying the following for all $n \ge n_0$. Let G be a graph then there exists a set $U \subseteq V(G)$ with $|U| = n/2 \pm n^{2/3}$ such that for each $v \in V(G)$,*

$$|N_G(v) \cap U| \ge \frac{1}{2} d_G(v) \pm n^{2/3}.$$

Proof. For each $v \in V(G)$, we add v to U independently at random with probability $1/2$. Then for a vertex $v \in V(G)$, let I_v be the indicator function such that

$$I_v := \begin{cases} 1 & \text{if } v \in U, \\ 0 & \text{if } v \notin U. \end{cases}$$

Then, for a vertex $v \in V(G)$, the random variable $|N_G(v) \cap U| = \sum_{u \in N_G(v)} I_u$ is a sum of independent random variables with $\mathbf{Pr}[I_u = 1] = 1/2$ and $\mathbf{Pr}[I_u = 0] = 1/2$. Then we have

$$\mathbb{E}\big[|N_G(v) \cap U|\big] = \sum_{u \in N_G(v)} \mathbf{Pr}[u \in U] = \frac{1}{2} d_G(v).$$

Thus Chernoff inequality (Lemma 1.7) with the fact that $d_G(v) \le n$ implies that

$$\mathbf{Pr}\big[|N_G(v) \cap U| \ne \frac{1}{2} d_G(v) \pm n^{2/3}\big] \le 2e^{-n^{4/3}/(2d_G(v))} \le 2e^{-n^{1/3}/2}.$$

By union bound, the probability that $|N_G(v) \cap U| = \frac{1}{2} d_G(v) \pm n^{2/3}$ does not hold for some vertex $v \in V(G)$ is at most $2ne^{-n^{1/3}/2}$. If we take n_0 large enough, this probability is strictly smaller than 1. Thus there exists a desired set U. $\square$

Exercises with asterisk (*) will be used later in the lecture. I recommend you to think about those exercises. You must at least understand the statement of those exercises with asterisk.

Exercise 1.1 (*)**.** *Prove that for each $0 < \varepsilon, \delta < 1$ with $\varepsilon < \delta/2$, there exists n_0 such that the following holds for all $n \ge n_0$. Let G be an n-vertex graph with average degree $d(G) = \delta n$. Then G contains a subgraph G' on n' with $\delta(G') \ge (\delta - \varepsilon)n'$ and $n' \ge \varepsilon n/4$. (Hint. if there's a low degree vertex, delete. Repeat and see how many vertices are left)*

Exercise 1.2. *Suppose that G is an n-vertex graph with m edges having no triangles.*

(a) *Prove that if $m > n^2/4$, then G contains a vertex v such that $G - v$ has more than $(n-1)^2/4$ edges.*
(b) *Prove that $m \leq n^2/4$.*

Exercise 1.3. [25]*Suppose that G is an n-vertex K_{r+1}-free graph with maximum number of edges.*

(a) *Prove that G contains a copy of K_r.*
(b) *Prove that G contains at most $t_r(n)$ edges, and characterize all graphs achieving the bound. (Hint. Use induction on n.)*

Exercise 1.4. [21] *Prove that every graph G with m edges has a bipartite subgraph with at least $m/2$ edges. (Hint. Consider a random partition of vertices.)*

Exercise 1.5. [21] *On n vertices, we select each pair of vertices to be an edge, independently at random, with probability $1/2$. Let the resulting graph be G. Show that there exists $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. With probability at least 0.99, G does not contain a bipartite subgraph with more than $n^2/8 + n^{2/3}$ edges. This shows that the results in Exercise 1.4 is close to best possible.*

Note that $G(n,p)$ is called a *random graph* which is obtained by selecting each pair of vertices to be an edge, independently at random, with probability p. (To be precise, it is a probability space rather than just a graph.)

Exercise 1.6 (*)**.** *Prove that for given $\varepsilon > 0$, there exists $n_0, C \in \mathbb{N}$ satisfying the following for all $n \geq n_0$. If $p \geq \frac{C}{n}$, with probability at least 0.99, the number of edges of random graph $G(n,p)$ is $(1 \pm \varepsilon)\binom{n}{2}p$.*

2. Extremal number

2.1. **Erdős-Stone-Simonovits Theorem.** Theorem 1.5 implies that the maximum number of edges in an n-vertex graph without K_{r+1} is exactly $t_r(n)$. We can ask the same question for more general graphs than just a complete graph. For given graph H, we define

$$ex(n, H) := \max\{e(G) : v(G) = n \text{ and } H \nsubseteq G\}.$$

This is called the *extremal number* of H or the *Turán number* of H. The definition implies that, if an n-vertex graph G has more than $ex(n, H)$ edges, then we can embed H into G while there exists an n-vertex graph G' with exactly $ex(n, H)$ edges such that H does not embed into G'.

The following Erdős-Stone-Simonovits Theorem provides an understanding of extremal number of a graph H.

Theorem 2.1 (Erdős-Stone 1946, Erdős-Simonovits 1966)**.** *Let H be a graph with $\chi(H) = r$. For given $\delta > 0$ and H, there exists $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$.*

$$ex(n, H) \leq (1 - \frac{1}{r-1} + \delta)\frac{n^2}{2}.$$

Note that the graph $T_{r-1}(n)$ does not contain any subgraph H with $\chi(H) = r$. This theorem implies that

$$\lim_{n\to\infty} \frac{ex(n, H)}{\binom{n}{2}} = 1 - \frac{1}{r-1}.$$

Note that $ex(n, H)/\binom{n}{2}$ is the maximum density of an n-vertex graph having no H as a subgraph. Thus, extremal number of a graph H does not depend very much on what H is, but rather to the chromatic number $\chi(H)$.

Proof of Theorem 2.1. If $\chi(H) = r$, then H is contained in a large complete r-partite graph. The idea of the proof is to find a large complete r-partite graph using induction on r.

First, we prove the following claim.

Claim 1. *For $r, h \in \mathbb{N}$ and $\delta > 0$, there exists $n_0' = n_0'(r, h, \delta)$ such that the following holds for all $n' > n_0'$. If G' is an n'-vertex graph with $\delta(G') \geq (1 - \frac{1}{r-1} + \delta/2)n'$, then G' contains a complete r-partite graph with each part having size at least h as a subgraph.*

Proof. We use induction on r. For $r = 1$, clearly $n_0'(r, h, \delta) = h$ would work. Assuem $r \geq 2$ and we will show that $n_0'(r, h, \delta) = n_0'(r-1, t, \delta) \cdot t\binom{t}{h}^{r-1}$ would work where $t := \lceil \frac{4h}{\delta} \rceil$.

Let $n_0' = n_0'(r-1, t, \delta) \cdot t\binom{t}{h}^{r-1}$, then the induction hypothesis gives $r-1$ disjoint sets $A_1, \dots, A_{r-1}$ of size t such that $G'[A_i, A_j]$ forms a complete bipartite graph for all $i \neq j \in [r-1]$. Let $U = V(G') \setminus \bigcup_{i\in[r-1]} A_i$. We will count the vertices in U which has at least h neighbors in each A_i. Let

$$U' := \{u \in U : d_{G'}(u; A_i) \geq h \text{ for each } i \in [r-1]\}.$$

To estimate the size of U', we define for each $u \in U$,

$$W_u := \big\{(v_1, \dots, v_{r-1}, u) : v_i \in N_{G'}(u; A_i) \text{ for each } i \in [r-1]\big\} \text{ and } W := \bigcup_{u\in U} W_u.$$

Then, for each $u \in U \setminus U'$, we have that

$$|W_u| < ht^{r-2}.$$

This implies that

$$|W| \leq \sum_{u\in U'} t^{r-1} + \sum_{u\in U\setminus U'} ht^{r-2} \leq t^{r-1}|U'| + n'ht^{r-2}. \tag{2.1}$$

On the other hand, to obtain a lower bound on $|W|$, we observe that every $(v_1, \dots, v_{r-1}) \in A_1 \times \dots \times A_{r-1}$, we have $d_{G'}(v_1, \dots, v_{r-1}) \geq (r-1)\delta(G') - (r-2)n' \geq r\delta n'/2$. Hence

$$|W| \geq \sum_{(v_1,\dots,v_{r-1}) \in A_1 \times \dots \times A_{r-1}} r\delta n'/2 \geq rt^{r-1}\delta n'/2.$$

This with (2.1) implies that $|U'| \geq r\delta n'/4$. For each $u \in U'$, we consider a tuple $f(u) := (V_1^u, \dots, V_{r-1}^u)$ such that $V_i^u \subseteq N_{G'}(u; A_i)$ and $|V_i^u| = h$ for each $i \in [r-1]$. Then by pigeonhole principle, there exists a tuple $(V_1, \dots, V_{r-1}) \in \binom{A_1}{h} \times \dots \times \binom{A_{r-1}}{h}$, such that

$$|f^{-1}(V_1, \dots, V_{r-1})| = |\{u \in U' : f(u) = (V_1, \dots, V_{r-1})\}| \geq \frac{|U'|}{\binom{t}{h}^{r-1}} \geq h.$$

Here we get the final inequality by the definition of n_0'. Hence, such a choice of disjoint vertex sets $V_1, \dots, V_{r-1}, f^{-1}(V_1, \dots, V_{r-1})$ form a complete r-partite graph with each part having size at least h, this proves the Claim. □

Now we prove the theorem. Let $h := v(H)$ and $n_0 = 4\delta^{-1}n_0'(r, h, \delta)$. Assume that $n \geq n_0$ and G is an n-vertex graph with at least $(1 - \frac{1}{r-1} + \delta)\frac{n^2}{2}$ edges. By Exercise 1.1, there exists a subgraph G' of G with $v(G') = n' \geq \delta n/4 \geq n_0'(r, h, \delta)$ and $\delta(G') \geq (1 - \frac{1}{r-1} + \delta/2)n'$. By Claim 1, the graph G' contains a complete r-partite graph each having size h. Since H is an r-chromatic graph with h vertices, it is easy to see that H embeds into the complete r-partite graph with each part size h. This prove the theorem. □

2.2. **Complete bipartite graphs.** The Erdős-Stone-Simonovits theorem approximately determines extremal number of any graph H. However, if H is a bipartite graph, the theorem only tells us that $ex(n, H)$ is subquadratic. (Note that $1 - \frac{1}{2-1} = 0$.) In other words, for any $\delta > 0$, there exists n_0 such that for $n \geq n_0$, any n-vertex graph with at least δn^2 edges contains H as a subgraph. Thus, we say that bipartite graphs are 'degenerate cases' for extremal problems. For bipartite graphs, we deal with each class of bipartite graphs in different ways. We introduce some results.

Theorem 2.2 (Kövari-Sós-Turán). *Let $s, t \in \mathbb{N}$ with $s \leq t$. Then we have*

$$ex(n, K_{s,t}) \leq tn^{2-1/s}.$$

Proof. Let G be an n-vertex graph with at least $tn^{2-1/s}$ edges and let $V = V(G)$.

$$\sum_{v \in V} |\binom{N(v)}{s}| = \sum_{v \in V} \binom{d(v)}{s} \geq n\binom{\frac{1}{n}\sum_v d(v)}{s} \geq n\binom{2tn^{1-1/s}}{s} \geq \frac{(tn)^s}{s!} \geq t\binom{n}{s}.$$

Here, we obtain the first inequality from the convexity of function $x \to \binom{x}{s}$. By pigeonhole principle, there exists a set $S \subseteq V(G)$ of size s which belongs to $N(v)$ for at least t distinct vertices $v \in V(G)$. Thus we obtain a copy of $K_{t,s}$. □

This result gives an upper bound on $ex(n, K_{s,t})$. As every bipartite graph H is a subgraph of $K_{s,t}$ for some s, t, this gives an upper bound of extremal numbers of all bipartite graphs. However, most of the cases the bound is not sharp. It is not known whether the above result is sharp up to constant. In other words, we don't know whether there exists a $K_{s,t}$-free graph G with at least $cn^{2-1/s}$ edges. However, if t is much larger than s, then we know the above bound is tight up to constant. Kollár, Rónyai and Szabó [11] proved that the bound is tight if $t \geq s! + 1$ and Alon, Rónyai and Szabó improved this to $t \geq (s-1)! + 1$.

By using a simple probabilistic method, we can prove the following weaker lower bound.

Theorem 2.3. *Let $s \in \mathbb{N}$. Then, there exists n_0 such that the following holds for all $n \geq n_0$. There exists an n-vertex graph G with $e(G) \geq \frac{1}{100} n^{2-2/s}$ which does not have $K_{s,s}$ as a subgraph.*

Proof. Consider a random graph $G(n,p)$ on vertex set V with $p = \frac{1}{10} n^{-2/s}$. For a fixed pair (A,B) of disjoint subsets of V with $|A| = |B| = s$, the probability that (A,B) induces a complete bipartite graph is p^{s^2}. Let X be the number of distinct copies of $K_{s,s}$ in $G(n,p)$. By linearity of expectation, we have

$$\mathbb{E}[X] = \sum_{A,B \in \binom{V}{s}, A \cap B = \emptyset} p^{s^2} \leq \binom{n}{s}\binom{n-s}{s} p^{s^2} \leq \frac{n^{2s} p^{s^2}}{2s^{2s}} \leq 1/10.$$

Hence, by Markov's inequality, $X \geq 1$ with probability less than $1/10$. On the other hand, Exercise 1.6 implies that $G(n,p)$ has less than $n^{2-2/s}/100$ edges with probability less than $1/2$. There union bound gives that there exists a graph G with $e(G) \geq n^{2-2/s}/100$ which does not have $K_{s,s}$ as a subgraph. □

Exercise 2.1. *Prove that for given $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$.*

$$ex(n, K_{s,t}) \leq (1+\delta)(t-1)^{1/s} n^{2-1/s}.$$

Exercise 2.2.

(a) (*) Consider $G(n,p)$ with $p = n^{-2/(s+1)}$. Compute the expected number of copies of $K_{s,s}$ in $G(n,p)$.
(b) Prove that for large n, there exists an n-vertex graph G with $e(G) \geq n^2 p/3$ which has at most $n^2 p/4$ copies of $K_{s,s}$.
(c) By using the graph G in (b), prove that $ex(n, K_{s,s}) \geq n^{2-2/(s+1)}/100$ for sufficiently large n.

Exercise 2.3.

(a) (*) Prove that if d is a positive integer and an n-vertex graph G has dn edges, then G contains a subgraph G' with $\delta(G') \geq d+1$.
(b) Prove that for a tree T with $d+1$ vertices, we have $ex(n,T) \leq (t-1)n$.
(c) Prove that for a tree T with $d+1$ vertices and $n \in \mathbb{N}$ divisible by d, we have $ex(n,T) \geq (d-1)n/2$.

2.3. **Even cycles.** Another important graph other than complete graph and complete bipartite graph are cycles. As odd cycles are not bipartite, their extremal number is approximately determined by Theorem 2.1. However, this still leaves cycles with even length. In this section we study the extremal number of even cycles.

We call a cycle with length at least $2k$ with a chord (an edge between two non-adjacent vertices in a cycle) Θ_k-*graph*. The following is an easy lemma we can prove.

Lemma 2.4. *For $k \geq 2$, any graph H with the average degree at least $4k$ contains a Θ_k-graph.*

Proof. By Exercise 2.3 (a), H contains a subgraph H' with $\delta(H') \geq 2k$.

Take a longest path $P = (x_1 \dots x_m)$ in H'. The vertex x_1 has at least $2k$ neighbors in H. As P is a longest path, all neighbors of x_1 are in P, say $x_{i_1}, \dots, x_{i_{2k}}$ are neighbors of x_1 with $1 < i_1 < \cdots < i_{2k}$. As $i_{2k} > 2k$ the path $(x_1, \dots, x_{i_{2k}})$ with the edge $x_1 x_{i_2}$ and $x_1 x_{i_{2k}}$ forms a Θ_k-graph. □

We consider this Θ_k-graph because it satisfies the following good property.

Lemma 2.5. *Let $k \geq 2$ and F be a Θ_k-graph. Let $A \cup B$ be a partition of $V(F)$ into two non-empty sets such that at least one of A and B is not an independent set in F. Then there exists a path of any length between 1 to $v(F) - 1$ from A to B.*

Proof. Let $V(F) = \mathbb{Z}_n$ such that i is adjacent to $i-1$ and $i+1$ and assume that $\ell \in [n-1]$ is the an integer such that there is no path of length ℓ from A to B. Let the chord be between 0 and r. Let $c(i) = 1$ if $i \in A$ and $c(i) = 2$ if $i \in B$. Let

$$P := \{m \in \mathbb{Z}_n : \text{ for all } i \in \mathbb{Z}_n, c(i) = c(i+m)\}.$$

If $\ell \notin P$, then we are done. Choose a smallest positive integer $m \in P$, then it is easy to see that $P = \{mi : i \in \mathbb{N}\}$ and m divides both ℓ and n. Thus we have $\ell \leq n - m$.

If $m = 1$, then either A or B is the emptyset. If $m = 2$ and $c(r) \neq c(0)$, then both A and B are independent sets, a contradiction. If $m = 2$ and $c(r) = c(0)$, then it is easy to find a path between A and B of length ℓ containing the chord.

Hence, we may assume that $m > 2$. At least one of r or $n - r$ are not congruent to 1 modulo m, we assume that r is not congruent to 1 modulo m. As $r - 1 \notin P$, there exists $j \in \mathbb{Z}_n$ such that $c(j) \neq c(j+r-1)$. As $c(j) = c(j+mi)$ and $c(j+r-1) = c(j+r-1+mi)$ for all $i \in \mathbb{Z}$, we may assume that $-m < j \leq 0$. We consider the following two cases.
Case 1. $1 < r \leq m$. Note that $j + \ell + r - 1 \leq j + n - m + r - 1 < n + j$ as $\ell \leq n - m$. Then the path $(j, j+1, \dots, 0, r, r+1, \dots, j+\ell+r-1)$ is a path between A and B with length ℓ, a contradiction.
Case 2. $m < r \leq n - m$. For each $-m < i < 0$, consider paths $(i, i+1, \dots, 0, r, r-1, \dots, r-i-m+1)$ and $(m+i, m+i-1, \dots, 0, r, r+1, \dots, r-i-1)$. Also, consider paths $(0, r, r-1, \dots, r-m+1)$ and $(0, r, r+1, \dots, r+m-1)$. As $c(i) = c(i+m)$, we have $c(r-i+1) = c(r-i-m+1) = c(r-i-1)$ and $c(r-m+1) = c(0) = c(r+m-1)$. These facts with the fact that $m \in P$ implies that $c(i+2) = c(i)$ for all $i \in \mathbb{Z}_n$, thus $m \in \{1, 2\}$, a contradiction. $\square$

Theorem 2.6 (Bondy and Simonovits). *For $k \geq 2$, there exists a constant c such that $ex(n, C_{2k}) \leq cn^{1+1/k}$.*

Proof. Our proof is a simplified version of the proof by Pikhurko [22]. Assume that G is an n'-vertex graph with $100kn'^{1+1/k}$ edges, having no C_{2k} as a subgraph. By using Exercise 1.4 and 2.3, we can obtain an n-vertex bipartite subgraph G' with $\delta(G') \geq 20kn^{1/k}$ and $n \leq n'$.

Note that our goal is to find a copy of C_{2k} in G' (and hence in G) to derive a contradiction. If our goal was to find any even cycles of length at most $2k$, then the situation is easier. We choose one vertex x and for each $i \leq k$ consider $V_i := N^i_{G'}(x)$, the set of all vertices of G' that is distance i from x in G'. If we don't find any cycle of length at most $2k$, then it is easy to see two vertices $u \neq v \in N^i_{G'}(x)$ does not have same neighbor in V_{i+1}. Thus $|V_{i+1}| \geq (\delta(G') - 1)|V_i|$, hence $|V_k| \geq (\delta(G') - 1)^k > n = |V(G)|$ which is a contradiction. However, as we only want to find a copy of C_{2k}, not a smaller cycle, our situation is more complicated than this. However, we will still prove that $|V_{i+1}|$ is significantly bigger than $|V_i|$ to derive a contradiction.

Fix an arbitrary vertex x of G'. For each $i \in [k]$, let $V_i := N^i_{G'}(x)$ consist of those vertices of G' that distance i from x in G'. Hence $V_0 = \{x\}$ and $V_1 = N_{G'}(x)$. For each $i \in [k-1]$, let $H_i := G'[V_i, V_{i+1}]$.

Claim 2. *For each $i \in [k-1]$, H_i does not contain a Θ_k-graph.*

Proof. Suppose that H_i contains a Θ_k-graph F. As H_i is a bipartite graph, F is also a bipartite graph. Moreover, F has a unique bipartition (A', B') of F into two independent sets. Assume that $A' \subseteq V_i$.

Let $T \subseteq G'$ be a breadth-first-search tree in G' with the root x. Let y be the vertex farthest from x such that every vertex of A' is a descendant of $y \in V_j$ in this tree T with $j < i$. Let z be a child of y in T, such that some (but not all) vertices in A' are descendants of z. Let $A \subseteq A'$ be the set of vertices of A' which are also descendants of z in T, and let $B = B' \cup (A' \setminus A)$.

Note that (A, B) is a partition of $V(F)$ with two sets and B is not an independent set. By Lemma 2.5, there exists a path P of length exactly $2k - 2(i-j) < 2k$ starting from $w \in A$ and ending at $w' \in B$. As F is a bipartite graph and P has an even length, $w' \in A' \setminus A$. Hence by definition of y and z, there exists a path in $T \subseteq G'$ of length $(i-j-1)$ from w and z and path in $T \subseteq G'$ of length $(i-j-1)$ from w' and $z' \neq z$, where z' is a child of y These paths with edges yz and yz' gives us a cycle of length $2k$, a contradiction. □

Note that we have the following for all $0 \leq i \leq k-1$.

$$e(H_i) < 5k(|V_i| + |V_{i+1}|). \tag{2.2}$$

Indeed, if $e(H_{i+1}) \geq 5k(|V_{i+1}| + |V_{i+2}|) = 5k|V(H_{i+1})|$, thus Lemma 2.4 implies that H_{i+1} contains a Θ_k-graph, a contradiction to Claim 2. So, we have (2.2).

Now we prove that the following holds for all $0 \leq i \leq k-1$

$$e(H_i) \geq 10kn^{1/k}|V_i| \text{ and } |V_{i+1}| > n^{1/k}|V_i|. \tag{2.3}$$

We use induction on i to prove (2.3). Note that (2.3) holds for $i = 0$. Assume that (2.3) holds for some $i \in [k-1]$. As G' is a bipartite graph, we know that $G'[V_{i+1}]$ is an independent set. Hence, we have

$$\begin{aligned} e(H_{i+1}) &\geq \sum_{v \in V_{i+1}} (d_{G'}(v) - d_{H_i}(v)) \geq 20kn^{1/k}|V_{i+1}| - e(H_i) \geq 20kn^{1/k}|V_{i+1}| - 5k(|V_i| + |V_{i+1}|) \\ &\geq 10kn^{1/k}|V_{i+1}|. \end{aligned}$$

Here, we obtain the final inequality because $|V_{i+1}| > 5kn^{1/k}|V_i|$. Hence, this with (2.2) imply that

$$|V_{i+2}| \geq \frac{1}{5k}e(H_{i+1}) - |V_{i+1}| \geq \frac{10kn^{1/k}|V_{i+1}|}{5k} - |V_{i+1}| \geq n^{1/k}|V_{i+1}|.$$

Thus (2.3) holds for $i+1$, this completes the induction. However, this implies that $|V_k| > n^{1/k}|V_{k-1}| > \cdots > n|V_0| \geq n$, a contradiction. Hence, every graph G with at least $100kn^{1+1/k}$ edges contain a copy of C_{2k}. □

Note that the above argument with more careful calculations show that $ex(n, C_{2k}) \leq (k-1)n^{1+1/k} + 16(k-1)n$. See [22]. Bukh and Jiang [4] recently improved this into $ex(n, C_{2k}) \leq 80\sqrt{k}(\log k)n^{1+1/k} + O(n)$.

For a lower bound, it is known that $ex(n, C_{2k}) \geq cn^{1+1/k}$ for $k \in \{2, 3, 5\}$.

Theorem 2.7. *For infinitely many values of n, there exists a graph on n vertices and $cn^{3/2}$ edges containing no C_4.*

Proof. Let p be a prime and consider $\mathbb{F}_p^3 \setminus \{(0,0,0)\}$. We say that two points (a_1, a_2, a_3) and (b_1, b_2, b_3) are equivalent if and only if there exists $k \in \mathbb{F}_p$ such that $b_i = ka_i$ for each $i \in [3]$. For a point (a_1, a_2, a_3), let $[a_1, a_2, a_3]$ be the equivalence class containing (a_1, a_2, a_3).

Let

$$V := \{[a_1, a_2, a_3] : (a_1, a_2, a_3) \in \mathbb{F}_p^3 \setminus \{(0,0,0)\}\}.$$

Let G be a graph over the vertex set V where $[a_1, a_2, a_3]$ and $[b_1, b_2, b_3]$ are adjacent if and only if $a_1b_1 + a_2b_2 + a_3b_3 = 0$. For two distinct vertices $u = [a_1, a_2, a_3]$ and $v = [b_1, b_2, b_3]$ of G, if $[x, y, z]$ is adjacent to both u and v only if

$$a_1x + a_2y + a_3z = 0 \text{ and } b_1x + b_2y + b_3z = 0.$$

As u and v are distinct, the set of points in $\mathbb{F}_p^3$ satisfying both equations forms a line in $\mathbb{F}_p^3$. Hence, the solutions define a single equivalence class. Hence, there exists unique common neighbor of u and v, implying that G does not contain any C_4.

Note that G contains $n = p^2 + p + 1 = \frac{p^3-1}{p-1}$ vertices. It is easy to see that every vertex of G is adjacent to $\frac{p^2-1}{p-1} = p + 1 \geq \sqrt{n} - 1$ other vertices. □

Exercise 2.4. *Prove that for $k \geq 2$, there exists a constant c such that $ex(n, C_{2k}) \geq cn^{1+1/(2k-1)}$. (Hint. consider a random graph.)*

3. Dependent random choice

This section is based on [9]. Consider a bipartite graph H with a vertex partition $X \cup Y$. A very natural way to embed a graph H into G is the following. We order the vertices in X into $(x_1, \dots, x_h)$ and we embed x_i into a vertex $\phi(x_i) \in V(G)$ one by one. While doing this, we make sure that for all $y \in Y$, the common neighborhood of vertices in $\phi(N_H(y))$ is large. As X is an independent set, this is a partial embedding, and there are many choices of vertices in G for y to embed. Hence, we can embed vertices in Y one by one into different vertices. Hence, the ideal situation is when we have a set $U \subseteq V(G)$ of vertices which satisfies the following property for some large r, m. Then we can freely embed vertices in X into U.

Every r vertices of U have at least m common neighbors.

If $r, m > h$, then we can arbitrarily embed each x_i into U to obtain an embedding of H into G. How can we obtain such a set U? We can use such a set by choosing vertices in a clever random way.

Lemma 3.1. *Let $a, d, m, n, r \in \mathbb{N}$. Let G be an n-vertex graph with $d(G) = d$. If there exists $t \in \mathbb{N}$ satisfying*

$$\frac{d^t}{n^{t-1}} - \binom{n}{r}(\frac{m}{n})^t \geq a,$$

then there exists a subset $U \subseteq V(G)$ with $|U| \geq a$ such that every r vertices in U have at least m common neighbors in G.

Proof. We will randomly choose vertices, and prove that chosen vertices have the desired property with positive probability.

Assume that we have for two sets W_1 and W_2 of size r with $d_G(W_1) > d_G(W_2)$. In order to increase the probability of obtaining a desired set, we want W_1 to be more likely to be included in U than W_2 (as our goal is to make all r-subsets of U to have many common neighbors). For this, we can randomly choose some vertex, say v, and we let U be the neighborhood $N_G(v)$. In this way, we can ensure that W_1 is more likely to be included in U than W_2 as

$$\mathbf{Pr}[W_1 \subseteq U] = \mathbf{Pr}[v \in N_G(W_1)] = \frac{d_G(W_1)}{n} > \frac{d_G(W_2)}{n} = \mathbf{Pr}[v \in N_G(W_2)] = \mathbf{Pr}[W_3 \subseteq U].$$

Now we start the proof. For each $i \in [t]$, we choose a random vertex $v_i \in V(G)$ independently uniformly at random. Note that two vertex v_i and v_j may be the same as we choose independently. Let $A = N_G(v_1, \dots, v_t)$ and $X = |A|$ be the random variable denoting the size of A. Linearity of expectation implies

$$\mathbb{E}[X] = \sum_{v \in V(G)} (\frac{|N_G(v)|}{n})^t = n^{-t} \sum_{v \in V(G)} d_G(v)^t \geq n^{1-t}(\frac{1}{n} \sum_{v \in V(G)} d_G(v))^t \geq \frac{d^t}{n^{t-1}}.$$

Here, we obtain the penultimate inequality by the convexity of the function $z \to z^t$.

Let

$$\mathcal{R} := \{R \in \binom{V(G)}{r} : |N_G(R)| \leq m\}.$$

Let Y be the random variable counting the number of subsets $R \subseteq A$ of size r such that $|N_G(R)| \leq m$. For given $R \in \mathcal{R}$, we have $\mathbf{Pr}[R \subseteq A] = (\frac{|N(R)|}{n})^t$. Thus

$$\mathbb{E}[Y] < \sum_{R \in \mathcal{R}} (\frac{|N(R)|}{n})^t \leq \binom{n}{r}(\frac{m}{n})^t.$$

By linearity of expectation,

$$\mathbb{E}[X-Y] \geq \frac{d^t}{n^{t-1}} - \binom{n}{r}(\frac{m}{n})^t \geq a.$$

This implies that there exists a choice of $x_1, \dots, x_t$ which yields a set A satisfying $X - Y \geq a$. Consider such a set A and delete one vertex from each subset $R \in \mathcal{R}$ lying inside A. Let U be the set of remaining vertices, then U has size at least $X - Y \geq a$, and U is our desired subset. □

Observe that we choose U in a dependent way. By choosing something else, and the choice of U depends from the earlier choice. Hence we call it *dependent random choice*. By this, we achieve that certain sets will be included in U more likely that some other sets. By using this lemma, we can extend the Theorem 2.2 into the following.

Theorem 3.2. *Let $s \in \mathbb{N}$ and let H be a bipartite graph with vetex partition A, B such that all vertices in B has degree at most s. Then there exists $c = c(H)$ such that $ex(n, H) \leq cn^{2-1/s}$.*

Proof. Let $a := |A|, b := |B|, m := a + b, d := 2cn^{1-1/s}$ and $c \geq 3ma$. Suppose that G is an n-vertex graph with $e(G) \geq cn^{2-1/s}$, hence $d(G) \geq d$. We have

$$\frac{d^s}{n^{s-1}} - \binom{n}{s}(\frac{m}{n})^s \geq 2c^s - \frac{n^s}{s!}(\frac{m}{n})^s \geq c^s \geq a.$$

Thus Lemma 3.1 implies that there exists a set U with $|U| = a$ such that any s vertices in U has at least m common neighbors.

We take an arbitrary injective map $\phi : A \to U$. Let $B = \{x_1, \dots, x_b\}$. We embed $x_1, \dots, x_b$ one by one in order. Right before we embed x_i, we have $|N_G(\phi(N_H(x_i)))| \geq m = a + b$, thus there exists a vertex $v_i \in N_G(\phi(N_H(x_i))$ which is not an image of any vertices in $A \cup \{x_1, \dots, x_{i-1}\}$. We embed x_i to v_i. By repeating this, we obtain an embedding of H into G. This prove the theorem. □

So far, we try to embed graph H with fixed size into a graph G with n vertices. While we consider a large number n, the graph H has been always fixed. What if we want to embed large graph H which contains at least cn vertices for some $c > 0$ into an n-vertex graph G? By using dependent random choice, we can prove several results on embedding H with cn vertices into an n-vertex graph G.

In the above theorem, we require that all s-sets in U have large common neighborhood. However, even if some s-sets in U have small common neighborhood, it is ok if we can somehow avoid using the sets. By using this idea, we can prove the following result regarding embedding linear size H into G.

Theorem 3.3. *Let $\varepsilon > 0$ and $\Delta \in \mathbb{N}$. Let G be an n-vertex graph with $e(G) \geq \varepsilon\binom{n}{2}$. If H is an $\frac{1}{8}\Delta^{-1}\varepsilon^{\Delta}n$-vertex bipartite graph with $\Delta(H) \leq \Delta$, then H embeds into G.*

To prove this theorem, we prove the following lemma using dependent random choice. Let $n' := \frac{1}{8}\Delta^{-1}\varepsilon^{\Delta}n$.

Lemma 3.4. *Let $\varepsilon > 0, \Delta, n, n' \in \mathbb{N}$ such that $n' := \frac{1}{8}\Delta^{-1}\varepsilon^{\Delta}n$ and $\Delta \leq n'$. If G is an n-vertex graph with at most $\varepsilon\binom{n}{2}$ edges, then there exists a set $U \subseteq V(G)$ with $|U| \geq 2n'$ such that less than $(2\Delta)^{-\Delta}\binom{|U|}{\Delta}$ sets $S \in \binom{U}{\Delta}$ satisfies $|N_G(S)| < n'$.*

Exercise 3.1. *Let $\varepsilon, \Delta, n, n'$ as in Lemma 3.4. Choose Δ vertices $v_1, \dots, v_\Delta$ in $V(G)$ independently uniformly at random with repetitions, and let $U = N_G(v_1, \dots, v_\Delta)$. Let X, Y be random variables with $X := |U|$ and $Y := |\{S \in \binom{U}{\Delta} : |N(S)| < n'\}|$.*

(a) Compute $\mu = \mathbb{E}[X]$ and $\nu = \mathbb{E}[Y]$.

(b) *By using the fact* $\mathbb{E}[X^\Delta] \geq \mu^\Delta$*, prove that there's a choice of* $v_1, \dots, v_\Delta$ *which make* $|U|^\Delta \geq \frac{1}{2}\mu^\Delta$ *and* $Y \leq 2|U|^\Delta \nu \mu^{-\Delta}$*. (Hint. Fist prove* $\mathbb{E}[X^\Delta - \frac{\mu^\Delta}{2} - \frac{\mu^\Delta}{2\nu} Y] \geq 0$*.)*

(c) *Prove Lemma 3.4.*

Proof of Theorem 3.3. Assume that H is a bipartite graph with vertex partition $A \cup B$ and $\Delta(H) \leq \Delta$. By Lemma 3.4, there exists a set U with $|U| \geq 2n'$ such that less than $(2\Delta)^{-\Delta}\binom{|U|}{\Delta}$ sets $S \in \binom{U}{\Delta}$ satisfies $|N(S)| < n'$. We call a subset $S \in \binom{U}{\Delta}$ *good* if $|N(S)| \geq n'$, and we call a subset $S' \subseteq U$ *nice* if it has size at most $\Delta - 1$ and S is contained in more than $(1-(2\Delta)^{|S'|-\Delta})\binom{|U|}{\Delta - |S'|}$ good sets.

Exercise 3.2. *For a nice set* S'*, there are at least* $(1 - 1/(2\Delta))|U|$ *vertices* u *in* $U \setminus S'$ *such that* $S' \cup \{u\}$ *is either nice or good.*

It is easy to see that for every nice set S', we have $d_G(S') \geq n'$.

Our aim is to find an injective function $f : A \to U$ such that $f(N_H(y))$ is nice or good for every $y \in B$. Then, for each $y \in B$, the common neighborhood of images of its neighbors, $N_G(f(N_H(y)))$ is large, so we can embed y into it. To obtain f, we order A into $x_1, \dots, x_t$, and embed vertices one by one.

Let $L_i := \{x_1, \dots, x_i\}$ for each $i \in [t]$. Assume we have defined $f : L_i \to U$ such that $f(N_H(y) \cap L_i)$ is nice or good for all $y \in B$. Now we will determine $f(x_{i+1})$, where we should embed x_{i+1}. For each $y \in N_H(x_{i+1})$, the set $f(N_H(y) \cap L_i)$ is nice. Thus Exercise 3.2 implies that there are at least $|U|/2 \geq n'$ vertices $w \in U$ such that $f((N_H(y) \cap L_i)) \cup \{w\}$ is nice or good for every $y \in N_H(x_{i+1})$. Since $t < n' = |V(H)|$, there exists a vertex w' among those n' vertices such that $w' \notin f(L_i)$. Let $f(x_{i+1}) := w'$.

Then for all $y \in N_H(x_{i+1})$, the set $f(N_H(y) \cap L_{i+1}) = f((N_H(y) \cap L_i) \cup \{x_{i+1}\})$ is nice or good. For all $y \in B \setminus N_H(x_{i+1})$, the set $f(N_H(y) \cap L_{i+1}) = f(N_H(y) \cap L_i)$ is also nice or good. By repeating this, we obtain an injective function $f : A \to U$ such that $f(N_H(y))$ is nice or good for each $y \in B$.

Now, order the vertices in B into $y_1, \dots, y_{t'}$. For each $i \in [t']$, we embed y_i into $N_G(f(N_H(y_i)))$ in such a way that y_i is not embedded to the same vertex as y_j with $1 \leq j \leq i-1$. It is possible because $f(N_H(y_i))$ is nice or good, thus $|N_G(f(N_H(y_i)))| \geq n' > t' + |A|$. This proves the theorem. □

As a consequence of this theorem, we can prove the following.

Exercise 3.3. *Let* H *be a bipartite graph with* r *vertices and maximum degree* Δ*. Then the Ramsey number* $r(H)$ *satisfies* $r(H) \leq \Delta 2^{\Delta+3} r$*.*

Note that this shows that Ramsey number of H is linear in terms of $|V(H)|$ when H is sparse bipartite graph, contrast to Theorem 1.2 which shows the Ramsey number of some dense graphs is exponential in $|V(H)|$. See [9] for more applications of dependent random choice.

Exercise 3.4. *Suppose* G *is a graph with* n *vertices with* $e(G) \geq \varepsilon n^2$*. Prove that* G *contains a* 1*-subdivision of a complete graph with* $\varepsilon^{3/2} n^{1/2}$ *vertices.*

Exercise 3.5 ([24])**.** *Let* $\varepsilon > 0$ *and suppose* G *is an* n*-vertex graph with* $e(G) \geq n^{1-\varepsilon}$*. If* n *is sufficiently large compared to* $1/\varepsilon$*, prove that either* G *contains* K_4 *or an independent set of size* $n^{1-\varepsilon}$*.*

4. Regularity

Imagine we have disjoint vertex sets $V_1, \dots, V_r$ of size n. For all $i \neq j \in [r]$ and $(u, v) \in V_i \times V_j$, we add an edge uv to a graph G independently at random with probability $p > 0$. If n is sufficiently large compare to $1/p$, then with high probability, the resulting graph contains many copies of any small r-partite graph. In other words, it is easy to embed any small r-partite graph H into such a random graph G. What makes such a graph G to have this 'good' property?

Indeed, the reason why we can easily embed H into G is because such a graph G satisfies some 'pseudo-random' condition, called *ε-regularity.* For a graph G and disjoint vertex sets A and B, let $e_G(A, B)$ be the number of edges of G between A and B and let

$$\mathrm{den}_G(A, B) := \frac{e_G(A, B)}{|A||B|}$$

be the density of G between A and B.

Definition 4.1. *A pair (A, B) of disjoint sets of vertices is (ε, d)-regular in a graph G if the following holds. For any $A' \subseteq A, B' \subseteq B$ of sizes $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$,*

$$|\mathrm{den}_G(A', B') - d| \leq \varepsilon.$$

We say (A, B) is ε-regular if it is (ε, d)-regular for some real number d. We say (A, B) is $(\varepsilon, d+)$-regular if it is (ε, d')-regular for some real number $d' \geq d$.

Note that a complete bipartite graph is $(0, 1)$-regular. Moreover, a random graph satisfies ε-regularity condition as below.

Exercise 4.1. *Consider two disjoint vertex sets U and V of size n and a real number $0 \leq p \leq 1$ and $\varepsilon > 0$. For each pair $u \in U$ and $v \in V$, we include an edge uv into G independently at random with probability p. Show that if n is sufficiently larger than $1/p$ and $1/\varepsilon$, then (U, V) is (ε, p)-regular pair in G with probability at least $1 - 1/n$.*

The condition of ε-regularity ensures that local density of the given graph at any part is close to its global density. This guarantees some 'uniformity' of the graph. This uniformity further gives the following property, which says that almost all vertices in G has "almost correct" degrees.

Lemma 4.2. *If (A, B) is (ε, d)-regular pair in a graph G and $B' \subseteq B$ with $|B'| \geq \varepsilon|B|$, then*

$$|\{a \in A : d_G(a; B') \neq (d \pm \varepsilon)|B'|\}| < 2\varepsilon|A|.$$

Proof. Let

$$A^- = \{a \in A : d_G(a; B') < (d - \varepsilon)|B'|\} \text{ and } A^+ = \{a \in A : d_G(a; B') > (d + \varepsilon)|B'|\}.$$

If $|A^-| \geq \varepsilon|A|$, then

$$\mathrm{den}_G(A^-, B') = \frac{\sum_{a \in A^-} d_G(a; B')}{|A^-||B'|} < \frac{\sum_{a \in A^-} (d - \varepsilon)|B'|}{|A^-||B'|} \leq d - \varepsilon.$$

Thus $|\mathrm{den}_G(A^-, B') - d| > \varepsilon$. It is a contradiction to the definition of ε-regularity because $|A^-| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$. Thus we conclude $|A^-| < \varepsilon|A|$. Similarly, we can show that $|A^+| < \varepsilon|A|$. $\square$

The following lemma shows that this notion of ε-regularity is useful for embedding a graph H into a graph G which admits an appropriate partition with ε-regularity.

Lemma 4.3 (Embedding lemma)**.** *Let $\Delta, n \in \mathbb{N}$, $d > 0$ and $\varepsilon < \frac{1}{\Delta+2}(d-\varepsilon)^{\Delta}$. Suppose $V(G) = V_1 \cup \cdots \cup V_r$ with $|V_i| = n$ for all $i \in [r]$ and (V_i, V_j) is $(\varepsilon, d+)$-regular for all $i \neq j \in [r]$. Let H be an r-partite graph with partition $X_1 \cup \cdots \cup X_r$ with maximum degree at most Δ and $|X_i| \leq \varepsilon n$ for each $i \in [r]$. Then there exists an embedding ϕ of H into G with $\phi(X_i) \subseteq V_i$.*

Proof. Let $x_1, \dots, x_h$ be the vertices of H, and let $\psi(i) \in [r]$ be the index such that $x_i \in X_{\psi(i)}$. We will construct an embedding ϕ by selecting images $v_1 = \phi(x_1), v_2 = \phi(x_2), \dots$ one by one. We will denote $C_i(j)$ to be the set of possible candidates of v_j after we determined $v_1, \dots, v_{i-1}$.

We want to embed vertices in X_i into V_i. Hence, we initially have $C_1(j) = V_{\psi(j)}$. Assume we have already determined $v_1, \dots, v_{i-1}$, and we have $|C_i(j)| \geq (\Delta+1)\varepsilon n$ for all $j \geq i$. We want to select v_i from $C_i(i)$. Consider the set A of neighbors of x_i in H, where

$$A = \{x_j \in N_H(x_i) : j > i\} = \{x_{s_1}, \dots, x_{s_p}\}.$$

Since $(V_{\psi(i)}, V_{\psi(s_\ell)})$ is $(\varepsilon, d+)$-regular for all $\ell \in [p]$, Lemma 4.2 implies that all but at most εn vertices in $C_i(i)$ have at least $(d-\varepsilon)|C_i(s_\ell)|$ neighbors in $C_i(s_\ell)$. Hence, at least $|C_i(i)| - \Delta\varepsilon n \geq \varepsilon n$ vertices in $C_i(i)$ has at least $(d-\varepsilon)|C_i(s_\ell)|$ neighbors in $C_i(s_\ell)$ for all $\ell \in [p]$. Among them, at most $\varepsilon n - 1$ vertices are already in $\{v_1, \dots, v_{i-1}\}$. Since $|C_i(i)| - \Delta\varepsilon n - (\varepsilon - 1) \geq (\Delta+1)\varepsilon n - \Delta\varepsilon n - \varepsilon n + 1 \geq 1$. Thus we can choose one of such vertex as $v_i \notin \{v_1, \dots, v_{i-1}\}$ which has at least $(d-\varepsilon)|C_i(s_\ell)|$ neighbors in $C_i(s_\ell)$ for all $\ell \in [p]$. Let $\phi(x_i) := v_i$. After choosing v_i, we update

$$C_{i+1}(j) := \begin{cases} C_i(j) \cap N_G(v_i) & \text{if } x_j \in N_H(x_i) \\ C_i(j) \setminus \{v_i\} & \text{if } x_j \notin N_H(x_i) \end{cases}$$

This will shrink $C_i(j)$ by the factor at least $(d-\varepsilon)$. Note that x_j has at most Δ neighbors, thus $|C_i(j)| \geq (d-\varepsilon)^{\Delta} n - \varepsilon n \geq (\Delta+1)\varepsilon n$. Thus we can choose v_i as long as $i \leq h$. By repeating this, we can select $v_1, \dots, v_h$ and it is easy to see that $\phi(x_i) = v_i$ is an embedding of H into G. $\square$

There is another reason why this notion of ε-regularity is so useful. The following theorem asserts that any graph can be partitioned into bounded number of vertex sets, so that most of pairs of vertex sets are ε-regular.

Theorem 4.4 (Szemerédi's regularity lemma)**.** *For every $\varepsilon > 0$ and $t \in \mathbb{N}$, there exist integers N and T such that the following holds for every $n \geq N$. Every n-vertex graph G admits a partition $V_0 \cup V_1 \cup \cdots \cup V_r$ satisfying the following.*

(R1) *$t \leq r \leq T$.*
(R2) *$|V_i| = |V_j|$ for all $i, j \in [r]$,*
(R3) *$|V_0| \leq \varepsilon n$,*
(R4) *for each $i \in [r]$, the pairs (V_i, V_j) are ε-regular in G for all $j \in [r]$ except at most εr indices.*

We call such a partition satisfying (R2)–(R4) a *ε-regular partition.* Note that T depends on the value of ε. (It's at most tower of twos, $2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$ with height in polynomial of $1/\varepsilon$.) However, it does not depend on n.

For a given ε-regular partition $V_0 \cup \cdots \cup V_r$ of G, let R be a graph on vertex set $[r]$ such that $ij \in E(R)$ if and only if (V_i, V_j) is an $(\varepsilon, \delta+)$-regular pair. We say such R is *(ε, δ)-reduced graph.* For appropriate δ, this graph R contains a lot of information about the original graph G. For example, the density of G is somewhat inherited into R as below.

Exercise 4.2 (*)**.** *Suppose $0 < 1/n \ll \varepsilon, 1/r \ll \delta < 0$. Let $V_0 \cup V_1 \cup \cdots \cup V_r$ be an ε-regular partition of an n-vertex graph G and R is an (ε, δ)-reduced graph for this ε-regular partition. Show that if $e(G) \geq (d + 2\delta)\binom{n}{2}$, then $e(R) \geq d\binom{r}{2}$.*

By using regularity lemma, we can provide an easy proof of Theorem 2.1.

Proof of Theorem 2.1. Suppose $1/n_0 \ll \varepsilon \ll \delta$. Assume that G is an n-vertex graph with $e(G) \geq (1 - \frac{1}{k-1} + \delta)\frac{n^2}{2}$. Apply the regularity lemma to G with $t = 1/\varepsilon$, then we obtain an ε-regular partition $V_0 \cup V_1 \cup \cdots \cup V_r$ of G, and let R be a $(\varepsilon, \delta/3)$-reduced graph for this ε-regular partition. By Exercise 4.2, we have $e(R) \geq (1 - \frac{1}{r-1} + \delta/3)r^2/2$. Turán's theorem implies that R contains a copy of K_k. Without loss of generality, assume that $R[\{1, 2, \dots, k\}] \simeq K_k$. By embedding lemma (Lemma 4.3), we can find a copy of H in $G[V_1 \cup \cdots \cup V_k]$. □

There are many applications of regularity lemma. For any graph H, the Ramsey number $r(H)$ is the minimum number n such that for any edge-coloring of the complete graph K_n contains a monochromatic copy of H. For example, it is well known that $r(H)$ exists for any graph H. For general graphs H, $r(H)$ is exponential in terms of $|V(H)|$. For example, we know $2^{d/2} \leq r(K_d) \leq 4^d$. However, the following result of Chvátal, Rödl, Szemerédi and Trotter [6] says that $r(H)$ is only linear in $|V(H)|$ if H has bounded maximum degree.

Theorem 4.5. *Let $\Delta > 0$ and let H be a graph of maximum degree at most Δ. Then there exists a constant $C = C(\Delta)$ such that $r(H) \leq C|V(H)|$.*

Proof. Let $k = r(K_{\Delta+1})$ be the Ramsey number of $K_{\Delta+1}$. In other words, every 2-edge-coloring of K_k contains a monochromatic $K_{\Delta+1}$. Take $\varepsilon = 1/(2^{\Delta+1}k)$, and $t = k + 1$. There exists $N = N(\varepsilon, t)$ and $T = T(\varepsilon, t)$ as stated in the regularity lemma. Let $C = C(\Delta) = 3\varepsilon^{-1}TN$. Note that C only depends on Δ.

Take $n \geq C|V(H)| = 3\varepsilon^{-1}TN|V(H)|$ and we want to show that any 2-edge-coloring of K_n contains a monochromatic copy of H. For a red/blue 2-edge coloring of K_n, let G be the red graph. By applying regularity lemma to G with parameters ε and t, we have ε-regular partition $V_0 \cup V_1 \cup \cdots \cup V_r$ for G with $t \leq r \leq T$ and $|V_i| \geq 2\varepsilon^{-1}|V(H)|$ for $i \in [r]$. Consider a $(\varepsilon, 0)$-reduced graph R on vertex set $[r]$ such that $ij \in E(R)$ if and only if (V_i, V_j) is ε-regular. Since $V_0 \cup \dots V_r$ is ε-regular partition, we obtain $E(R) > (1-\varepsilon)\binom{r}{2} > (1 - \frac{1}{k-1})\binom{r}{2}$. Thus Turán's theorem implies that R contains K_k.

Consider the copy of K_k in R, and color the edge ij of K_k red if (V_i, V_j) is $(\varepsilon, 1/2+)$-regular and otherwise color it blue. By the definition of k, there is a monochromatic copy of $K_{\Delta+1}$ in K_k. By re-indexing, assume that $\{1, 2, \dots, \Delta + 1\}$ forms such monochromatic copy of $K_{\Delta+1}$ in R. If it is red-monochromatic, then let $G' = G$, otherwise let $G' = \overline{G}$. Note that G' is a monochromatic subgraph of K_n.

For each $i, j \in [\Delta + 1]$ with $i \neq j$, (V_i, V_j) is $(\varepsilon, 1/2+)$-regular in G' (check this as an exercise). Also H is $(\Delta + 1)$-partite graph with each partition class has size at most $|V(H)| \leq \varepsilon|V_i|/2$. Thus we may apply lemma 4.3 with $d = 1/2$ and $m = |V(H)|$ to find a copy of H in G'. This gives us a monochromatic copy of G in K_n and it concludes that $r(H) \leq 3\varepsilon^{-1}TN|V(H)|$ where ε, T, N are all determined by Δ. □

The above proof gives tower type large constant $C(\Delta)$. A stronger approach to the problem using Dependent Random Choice gives a better constant $C(\Delta) \leq 2^{C'\Delta\log(\Delta)}$ for some constant C'.(see [8]). In [19], Lee showed that for d-degenerate graph G, there exists $C(d)$ such that $r(G) \leq C(d)|G|$ holds.

Exercise 4.3 (*). *Suppose* $0 \le \varepsilon < \alpha \le 1$. *If* (A, B) *is* (ε, d)*-regular pair in* G *and* $|A'| \ge \alpha|A|$ *and* $|B'| \ge \alpha|B|$, *then prove that* (A', B') *is* $(\frac{\varepsilon}{\alpha}, d)$*-regular pair in* G.

Exercise 4.4 (*). *Suppose* $0 < 1/n \ll \varepsilon \ll d \le 1$. *Suppose that* (A, B) *is* (ε, d)*-regular pair in* G. *Sets* A, B, A', B' *are four pairwise disjoint sets in* $V(G)$ *such that* $|A| = |B| = n$, $|A'| \le \varepsilon|A|$ *and* $|B'| \le \varepsilon|B|$. *Prove that* $(A \cup A', B \cup B')$ *is* $(4\varepsilon^{1/2}, d)$*-regular pair in* G.

Exercise 4.5. *Let* $V(G)$ *be a* 3*-partite graph with partition* $V_1 \cup V_2 \cup V_3$ *such that* (V_i, V_j) *is* (ε, d)*-regular pair in* G *for all* $1 \le i < j \le 3$. *How many triangles does* G *have? Find an upper bound and a lower bound on the number of triangles in* G.

Exercise 4.6 (*). *Suppose* $0 < 1/n \ll \varepsilon, 1/r \ll \delta$. *Let* $V_0 \cup V_1 \cup \cdots \cup V_r$ *be an* ε*-regular partition of an* n*-vertex graph* G. *Let* R *be a* (ε, δ)-reduced graph of the partition $V_0 \cup V_1 \cup \cdots \cup V_r$. Show that if $\delta(G) \ge (d + 2\delta)n$, then $\delta(R) \ge dr$.

Exercise 4.7 (*). *Let* $\Delta, n \in \mathbb{N}$ *with* $0 < 1/n \ll \varepsilon \ll d, 1/\Delta$. *Suppose* $V(G) = V_1 \cup \cdots \cup V_r$ *with* $n_i := |V_i| = (1 \pm \varepsilon)n$ *for all* $i \in [r]$ *and* R *is a* (ε, d)*-reduced graph for the partition* $V_1 \cup \cdots \cup V_r$. *Let* H *be an* r*-partite graph with partition* $X_1 \cup \cdots \cup X_r$ *with maximum degree at most* Δ *and* $|X_i| = n_i$ *for each* $i \in [r]$, *and* $H[X_i, X_j]$ *induces no edges for* $ij \notin E(R)$. *For each* $i \in [r]$, *we have a set* $Y_i \subseteq X_i$ *with* $|Y_i| \le \varepsilon n$ *and let* $Y = \bigcup_{i \in [r]} Y_i$. *Prove that there exists an embedding* $\phi : H[Y] \to G$ *such that the following hold.*

- *(a)* $\phi(Y_i) \subseteq V_i$.
- *(b) For all* $i \in [r]$ *and* $x \in X_i \setminus Y_i$, *we have* $|\bigcap_{z \in N_H(x;Y)} N_G(\phi(z))| \ge (d - 2\varepsilon)^{\Delta} n$.

5. Embedding large graphs

So far, we have proved that if an n-vertex graph G is dense enough, then it contains a small graph H as a subgraph. For this, n is always sufficiently large compared to $|V(H)|$. What if we want to find a subgraph H which has n vertices, in other words, $V(H)$ has the same size as $|V(G)|$. In this case, density condition on G alone is not sufficient.

To see this, consider a sparse n-vertex graph H without isolated vertex, say an n-vertex cycle, and an n-vertex graph G containing disjoint union of K_{n-1} and one isolated vertex. Even if G is extremely dense, the graph G does not contain H as a subgraph. To avoid this condition, we usually seek for minimum degree conditions on G.

The simplest spanning subgraph we can consider is a *perfect matching* of G, which is a collection of disjoint edges covering all vertices of G. The following is a useful theorem for finding a perfect matching. We omit the proof as it is well-known.

Theorem 5.1 (Hall's theorem)**.** *Suppose that G is a bipartite graph with vertex partition (A, B) such that for each $A' \subseteq A$, we have $|N^1_G(A')| \geq |A'|$. Then G contains a matching covering all vertices of A.*

Exercise 5.1. *Suppose $0 \leq \varepsilon < d/2$. Suppose that G is a bipartite graph on vertex partition (A, B) with $|A| = |B| = n$. Prove that if (A, B) is an (ε, d)-regular pair in G and $\delta(G) \geq dn/2$, then G contains a perfect matching.*

As mentioned before, we seek for a minimum degree condition on n-vertex graph G to contain an n-vertex graph H as a subgraph. One interesting graph H is when it is a cycle. We say that a cycle is *Hamilton cycle* of G if it contains every vertex of G.

Theorem 5.2 (Dirac's theorem)**.** *If an n-vertex graph G satisfies $\delta(G) \geq n/2$, then G contains a Hamilton cycle.*

Proof. It is easy to see that G is connected. Let $P = x_1 \dots x_s$ be a longest path in G. Then every vertices in $N_G(x_1)$ and $N_G(x_s)$ are in $V(P)$, otherwise we obtain a longer path. As $|N_G(x_1)|, |N_G(x_s)| \geq n/2$, we have

$$|\{i \in [s-1] : x_{i+1} \in N_G(x_1)\} \cap \{i \in [s-1] : x_i \in N_G(x_s)\}| \geq n/2 + n/2 - (s-1) \geq 1.$$

There exists $k \in [s-1]$ such that $x_1 x_{k+1} \in E(G), x_k x_s \in E(G)$. Consider a cycle

$$C = x_1 x_{k+1} x_{k+2} \dots x_s x_k x_{k-1} \dots x_2 x_1.$$

If C is not Hamilton cycle, as G is connected, there exists a vertex $y \in V(G) \setminus V(C)$ and $j \in [s]$ such that $yx_j \in E(G)$. However, C with and edge $x_j y$ contains a path longer than P, a contradiction. Hence C is a Hamilton cycle of G, so G contains a Hamilton cycle. □

Exercise 5.2. *For given $n \in \mathbb{N}$, find an n-vertex graph G with $\delta(G) \geq \frac{n}{2} - 1$ such that G does not contain a Hamilton cycle.*

As we have seen in the small graph embedding results, embedding a complete graph is an important case. So, it is natural to consider embedding vertex disjoint union of K_r covering (almost) all vertices of G. The following theorem provides a sufficient condition on the minimum degree of G to contain many vertex-disjoint copies of K_r. For our convenience, we will consider a complement $\overline{G}$ of G, and we will try to find vertex-disjoint independent sets of size r in $\overline{G}$. We say that a coloring $f : V(G) \to [k]$ is an *equitable k-coloring* if it is a proper coloring such that the sizes of the color classes differ by at most one. The following proof is due to Kierstead and Kostochka [10].

Theorem 5.3 (Hajnal-Szemerédi)**.** *If an n-vertex graph G has maximum degree less than k, then there exists an equitable k-coloring of G. In particular, if G' is an n-vertex graph with $\delta(G') \geq (1-\frac{1}{r})n$, then G' contains $\lfloor \frac{n}{r} \rfloor$ copies of K_r.*

Proof. It is easy to see that G has k-equitable coloring if and only if $G \cup K_{k'}$ has k-equitable coloring for $0 \leq k' < k$ with $n \equiv -k' \pmod{k}$. Hence we may assume that $n = kr$. We use induction on $m = |E(G)|$, it is easy to see that the theorem holds when $m = 0$. We say that a coloring $f : V(G) \to [k]$ is a *near-equitable k-coloring* if it is a proper coloring with color class sizes $r-1, r, \dots, r, r+1$.

Assume a minimum counterexample G which has no equitable k-coloring, with $|E(G)| > 0$.

Exercise 5.3. *Such a graph G has an near-equitable k-coloring.*

For a given near-equitable k-coloring f, let $V_1, \dots, V_k$ be color classes with $|V_1| = r-1$ and $|V_k| = r+1$. We consider a digraph D^f with vertices $V_1, \dots, V_k$ such that $\overrightarrow{V_iV_j} \in E(D^f)$ if there exists a vertex $v \in V_i$ having no neighbor in V_j. A vertex V_i of D^f is *accessible* if D^f has a path from V_i to V_1. If V_k is accessible, then G has an equitable k-coloring, by shifting a vertex of G from each class on the path to the next class on the path. We say that V_i *blocks* V_j if $D^f - V_i$ does not contain a path from V_j to V_1, and an accessible class is *free* if it blocks no other accessible classes.

Among all near-equitable k-colorings of G, choose one coloring f with the fewest inaccessible vertices in D^f. WLOG we let $V_1, \dots, V_a$ be accessible vertex classes and $V_{a+1}, \dots, V_k$ be inaccessible vertex classes with $|V_1| = r-1$ and $|V_k| = r+1$, and assume that $V_1, \dots, V_{a-a'}$ are not free and $V_{a-a'+1}, \dots, V_a$ are free. Let $b := k - a$. Let

$$A = \bigcup_{i=1}^{a} V_i, \ \ A' := \bigcup_{i=a-a'+1}^{a} V_i, \ \text{ and } \ B := \bigcup_{i=a+1}^{k} V_i.$$

For $x \in V_i \subseteq A'$, we say that x is *movable* if there exist $j \in [a] \setminus \{i\}$ such that x has no neighbor in V_j. And an edge xy with $x \in V_i \subseteq A'$ and $y \in B$ is a *solo edge* if y has only one neighbor in V_i which is x.

Claim 3. *If xy is a solo edge with $x \in V_i \subseteq A'$ and $y \in V_\ell \subseteq B$ and x is movable, then G has an equitable k-coloring.*

Proof. There exists $j \in [a] \setminus \{i\}$ such that x has no neighbor in V_j. We move x to V_j and y into V_i, then we obtain a near-equitable a-coloring f' of $G[A \cup \{y\}]$ and a near equitable b-coloring g of $G[B - \{y\}]$.

Since V_i was free in the original coloring, V_j remains accessible under f'. Hence $D^{f'}$ has a path from V_j to V_1, thus $G[A \cup \{y\}]$ has an equitable a-coloring.

For $z \in V_s \subseteq B$, as V_s is inaccessible, we know $d_G(z; A) \geq a$, hence $d_G(z; B) < k - a = b$. Then $G[B - \{y\}]$ has maximum degree less than b, the induction hypothesis implies that $G[B - \{y\}]$ has an equitable b-coloring. Together, these two coloring combined gives an equitable k-coloring of G. □

Exercise 5.4. *If xy and xy' are solo edges with $x \in V_i \subseteq A'$, but x is not movable and $yy' \notin E(G)$, then G has a near-equitable k-coloring with more accessible vertex classes than a. (Hint. Obtain a new equitable coloring g' of $G[B - \{y\}]$ and add x to one of the color class of g'. Replace V_i with $V_i - x + y$. Check it contains more accessible classes.)*

By using Claim 3 and Exercise 5.4, it suffices to find a solo edge with its endpoint in A' is movable, or two solo edges having a common endpoint in A' whose the other endpoints are non-adjacent. We consider two cases.

Case 1. $a' \leq b$. Assume that each of $V_2, \dots, V_{a-a'}$ has a lower-indexed neighbor in D^f. The set $V_{a-a'}$ blocks some class V_j with $a - a' < j \leq a$, then for each $x \in V_j$, we have $d_G(x; A) \geq a - a' - 1$ as x is not movable to any V_i with $i < a - a'$. Hence $d_G(x; B) \leq k - (a - a') \leq b + a' \leq 2b$. Let

$$U := \{x \in V_j : x \text{ is incident to a solo edge}\} \text{ and } U' := V_j \setminus U.$$

Let m be the number of edges joining U' to B, so $m \leq 2b|U'|$. On the other hand, as classes in B are inaccessible, every vertex of B has a neighbor in V_j. The vertices in $B \setminus N_G(U; B)$ have at least two such neighbors in U', so $m \geq 2(|B| - |N^1_G(U) \cap B|)$. We have

$$\begin{aligned}\sum_{x \in U} d_G(x; A) &\leq (k-1)|U| - |N^1_G(U) \cap B| = (a-1)|U| + br - b|U'| - |N^1_G(U) \cap B| \\ &\leq (a-1)|U| + (|B| - 1) - m/2 - |N^1_G(U) \cap B| \\ &\leq (a-1)|U| + |B| - 1 - (|B| - |N^1_G(U) \cap B|) - |N^1_G(U) \cap B| = (a-1)|U| - 1.\end{aligned}$$

Hence there exists a vertex $x \in U$ such that $d_G(x; A) < a - 1$, hence x is movable. Since $x \in U$, we have a solo edge incident to x.

Case 2. $b \leq a'$. Let I be a maximum independent subset of B, then we have $|I| \geq r$. For each $y \in B$, let $\sigma(y)$ be the number of solo edges incident to y, then we have $d_G(y; A) \geq a + (a' - \sigma(y))$, thus we have $\sigma(y) \geq a' + a - d_G(y; A) \geq a' - b + d_G(y; B) + 1$. As I is maximal, we have

$$\sum_{y \in I} (d_G(y; B) + 1) \geq |B| = br + 1.$$

Thus

$$\begin{aligned}\sum_{y \in I} \sigma(y) &\geq \sum_{y \in I} (a' - b + d_G(y; B) + 1) \geq |I|(a' - b) + \sum_{y \in I} (d_G(y; B) + 1) \\ &\geq r(a' - b) + |B| \geq a'r + 1 > |A'|.\end{aligned}$$

Hence, a vertex in A' is incident to at least two solo edges, a contradiction. □

What if H is more general graphs, other than cycle or clique-factors? What would be a minimum degree condition on G guarantees a subgraph H in G? The following theorem by Sauer and Spencer gives us a bound on this question.

Theorem 5.4. *Suppose that G and H are n-vertex graph and $\delta(G) > (1 - \frac{1}{2\Delta(H)})n - 1$. Then G contains H as a subgraph.*

Proof. Note that we have $\Delta(H)\Delta(\overline{G}) < n/2$. We may assume that $n \geq 3$. Consider a bijection $\phi : V(H) \to V(G)$ which maps the most number of edges of H into edge in G. Let $V(H) = \{x_1, \dots, x_n\}, V(G) = \{v_1, \dots, v_n\}$ and $\phi(x_i) = v_i$ for each $i \in [n]$. We call an edge e of H is *good* for ϕ if $\phi(e) \in E(G)$, and we call it *bad* for ϕ otherwise. If every edges of H are good for ϕ, then we are done. Otherwise, we may assume that $x_1x_n \in E(H)$ and $v_1v_n \notin E(G)$.

We try to find $k \in [n-1] \setminus \{1\}$ such that we can change the image of x_k and x_n of ϕ to obtain a better bijection ϕ_k such that $\phi_k(x_k) = v_n, \phi_k(x_n) = v_k$ and $\phi_k(x_i) = v_i$ for each $i \in [n] \setminus \{k, n\}$.

If for some $j \in [n-1]$, we have $x_jx_n \in E(H)$ and $v_jv_k \notin E(G)$, then ϕ_k maps an edge x_jx_n into non-edge in G, so x_jx_n is bad for ϕ_k. As it could be a new bad edge, we want to avoid this, and there are at most $\Delta(H)\Delta(\overline{G}) - 1$ such choices of (j, k) other than $(1, n)$.

Similarly, if for some $j \in [n-1]$, we have $x_jx_k \in E(H)$ and $v_jv_n \notin E(G)$, then we could introduce a new bad edge. Again there are at most $\Delta(H)\Delta(\overline{G}) - 1$ such choices of (j, k) other than $(1, n)$.

For each such (j,k) we exclude k. As there are at most $2\Delta(H)\Delta(\overline{G}) - 2 < n-2$ such choices of (j,k) other than $(1,n)$, there exists $k \in [n-1] \setminus \{1\}$ which is not excluded. It is easy to check that every good edge for ϕ is still good for ϕ_k. Moreover, x_1x_n is a bad edge for ϕ but good edge for ϕ_k, a contradiction to the definition of ϕ. Thus ϕ is an embedding of H into G, thus G contains H as a subgraph. □

Conjecture 5.5 (Bollobás-Eldridge 1978, Catlin 1976)**.** *If G and H are n-vertex graphs with $\delta(G) \geq (1 - \frac{1}{\Delta(H)+1})n$, then H embeds into G.*

Kaul, Kostochka and Yu proved that if $\delta(G) \geq (1 - \frac{3}{5(\Delta(H)+1)})n$, then H embeds into G.

Exercise 5.5. *Let H and G be two n-vertex graphs with $|E(H)||E(\overline{G})| < \binom{n}{2}$. Prove that H embeds into G. (Hint. Count the number of bijections mapping an edge into a non-edge.)*

Exercise 5.6. *Prove that if an n-vertex graph G satisfies $|E(\overline{G})| < n/2$, then G contains a copy of every n-vertex tree.*

Exercise 5.7. *Suppose that H and G are $2n$-vertex graphs such that each of them has an equitable 2-coloring. Prove that if $\delta(G) \geq (1 - \frac{1}{2\Delta(H)})n$, then H embeds into G.*

6. The blow-up lemma

Lemma 4.3 together with regularity lemma helps us to find an embedding of a small graph H with bounded maximum degree into a dense graph G. Can we prove a similar statement for graphs H which has the same number of vertices with G? Let's first consider the simplest case, when H is a perfect matching.

Does (ε, d)-bipartite graph on vertex partition (A, B) with $|A| = |B|$ has a perfect matching? The answer is no. As ε-regularity does not restrict the local behavior of the graph, there could be isolated vertices in $A \cup B$. Thus, we need a stronger notion to ensure a perfect matching as follows.

Definition 6.1. *Let G be a graph. A disjoint pair (A, B) of sets of vertices is (ε, d)-super-regular if it is $(\varepsilon, d+)$-regular and*

$$d_G(a; B) \geq (d - \varepsilon)|B| \text{ and } d_G(b; A) \geq (d - \varepsilon)|A|$$

for all $a \in A$ and $b \in B$.

This notion of super-regularity is sufficient for having a perfect matching.

Exercise 6.1. *Let $0 < 2\varepsilon < d \leq 1$ and $n \geq 2/\varepsilon$. Let G be a graph. If (A, B) is (ε, d)-super-regular pair in G with $|A| = |B| = n$, then $G[A, B]$ contains a perfect matching.*

Indeed, we can extend Lemma 4.3 into the following blow-up lemma proved by Komlós, Sárközy and Szemerédi.

Theorem 6.2 (The blow-up lemma)**.** [12] *Suppose $0 < 1/n \ll \varepsilon \ll 1/r, d, 1/\Delta \leq 1$. Let H be a vertex graph with vertex partition $X_1 \cup \cdots \cup X_r$ with $\Delta(H) \leq \Delta$. Let G be a graph with vertex partition $V_1 \cup \cdots \cup V_r$ such that $|V_i| = |X_i| = n$ and (V_i, V_j) is $(\varepsilon, d+)$-super-regular for all $i \neq j \in [r]$. Then there exists an embedding ϕ of H into G such that $\phi(X_i) = V_i$.*

This is a simplified version. Actually we can do more than this.

Remark on blow-up lemma.

1. $X_1 \cup \cdots \cup X_r$ does not have to be equitable partition. Some sets can have size bigger than n.
2. In the above, the '(ε, d)-reduced graph' for G is a complete r-vertex graph K_r. However, we can find the embedding even when the reduced graph R is not complete, if H has a structures compatible to R. (If $ij \notin E(R)$, then $H[X_i, X_j]$ must have no edges.)
3. We want to use the blow-up lemma after applying regularity lemma. However, regularity lemma provides an ε-regular partition $V_0 \cup \cdots \cup V_r$ with $1/r \ll \varepsilon$. On the other hand, in the above blow-up lemma, we assumed $\varepsilon \ll 1/r$ for complete reduced graph R. This causes a problem.

   However, the blow-up lemma works as long as R has maximum degree Δ_R and $\varepsilon \ll 1/\Delta_R$. For example, if R is a cycle of length r, then blow-up lemma still works even if we have $1/r \ll \varepsilon$.
4. For each $i \in [r]$ and vertex sets $X'_i \subseteq X_i$ of size at most εn, we can roughly specify where $x \in X'_i$ should be embedded. Specifically, for all $x \in X'_i$, we can specify a set $V'_x \subseteq V_i$ such that $|V'_x| \geq d'|V_i|$ for some d' with $\varepsilon \ll d'$. Then we can find an embedding ϕ of H into G such that $\phi(x) \in V'_x$.

Our goal is to apply the regularity lemma to obtain an ε-regular partition G, and use it to apply the blow-up lemma with the partition and H. For this, we need to take care of two problems.

- ε-regular partition may not give ε-super-regularity.
- Reduced graphs usually are big and have large maximum degree.

We can overcome the first problem by deleting (or moving) some vertices from the ε-regular partition. See Exercise 4.4 and the following two exercises.

Exercise 6.2 (*)**.** *Suppose $0 < \varepsilon \ll d \leq 1$. If (A, B) is (ε, d)-regular pair in G, then there exists $A' \subseteq B, A' \subseteq B$ such that $|A'| \geq (1-\varepsilon)|A|$, $|B'| \geq (1-\varepsilon)|B|$ and (A', B') is $(2\varepsilon, d)$-super-regular pair in G*

Exercise 6.3 (*)**.** *Suppose $0 < \varepsilon \ll d \leq 1$. If (A, B) is (ε, d)-super-regular pair in G, then for any $A' \subseteq A, B' \subseteq B$ with $|A'| \geq (1-\varepsilon)|A|$, $|B'| \geq (1-\varepsilon)|B|$, (A', B') is $(2\varepsilon, d)$-super-regular pair in G.*

Exercise 6.4 (*)**.** *Supose $0 < 1/m \ll \varepsilon \ll \delta, 1/\Delta \leq 1$. Let $V_0 \cup V_1 \cup \cdots \cup V_r$ be an ε-regular partition of G with $|V_i| = (1 \pm \varepsilon)m$ for each $i \in [r]$, and let R be an (ε, δ)-reduced graph of the ε-regular partition with $\Delta(R) \leq \Delta$. Let m' be an integer such that $(1 - 3\Delta\varepsilon)m \leq m' \leq (1 - 2\Delta\varepsilon)m$. Prove that there exists subsets $U_i \subseteq V_i$ for each $i \in [r]$ such that $|U_i| = m'$ and (U_i, U_j) is a $(6\Delta\varepsilon, \delta+)$-super-regular pair in G for all $ij \in E(R)$.*

The following theorem is known as Alon-Yuster conjecture. It is proved by Komlós, Sárközy and Szemerédi. Here, *F-factor* means vertex-disjoint copies of F which covers all vertices in another graph G.

Theorem 6.3. [13] *For given k-partite f-vertex graph F, there exist n_0 and C such that the following holds for $n \geq n_0$. If an n-vertex graph G satisfies*

$$\delta(G) \geq (1 - \frac{1}{k})n + C$$

and n is multiple of f, then G contains an F-factor.

The above bound $\delta(G) \geq (1-\frac{1}{k})n+C$ is sharp as there exists a k-chromatic graph F and an graph G with minimum degree at least $(1 - \frac{1}{k})n - C$ such that F does not embed into G. However, there are some k-chromatic graphs F such that a minimum degree bounds on G lower than $(1 - \frac{1}{k})n$ suffices to guarantees an embedding of F into G. In [**?**], Kühn and Osthus determined the sharp (up to additive constant) minimum degree threshold on G for guaranteeing an embedding of F into G for each graph F.

Here we introduce an application of blow-up lemma to prove the following approximate version of Alon-Yuster conjecture.

Theorem 6.4. *For given k-colorable h-vertex graph F and real number $\alpha > 0$, there exist n_0 such that the following holds for $n \geq n_0$. If an n-vertex graph G satisfies*

$$\delta(G) \geq (1 - \frac{1}{k} + \alpha)n,$$

then G contains at least $(1-\alpha)n/|V(F)|$ vertex-disjoint copies of F.

Proof. Take F' which is a disjoint union of k copies of F. Then it is easy to check that F' has a vertex partition $X_1 \cup \cdots \cup X_k$ such that $|X_i| = f$ for all $i \in [k]$.

Consider $\varepsilon, t = \varepsilon^{-2}$ such that $0 < 1/n_0 \ll \varepsilon \ll \delta \ll \alpha, 1/k \leq 1$. Apply regularity lemma to G with parameters ε, t, and obtain ε-regular partition $V_0 \cup \cdots \cup V_r$ with $1/n_0 \ll 1/r < 1/t$ and consider (ε, δ)-reduced graph R on vertex set $[r]$ of the partition. Let $m := |V_i|$ for each $i \in [r]$. By moving vertices in $V_{r-k+2}, \ldots, V_r$ to V_0 if necessary, we may assume that r is multiple of k and $|V_0| \leq 2\varepsilon n$. (Note that $m = (1 \pm 2\varepsilon)n/r$.)

In G, we delete all edges incident to V_0, and all edges between V_i and V_j with $ij \notin E(R)$ (i.e. all edges between non-regular pairs, and all edges between pairs of density at most δ) and all edges inside each V_i. Let G' be the resulting graph after this deletion.

Since $\alpha \geq 2\delta$, by using Exercise 4.6, we have $\delta(R) \geq (1 - \frac{1}{k})r$. Thus Hajnal-Szemerédi Theorem implies that we can find r/k vertex disjoint copies of K_k in R.

Consider $R' \subseteq R$ which is union of r/k vertex-disjoint copies of K_ks in R. By applying Exercise 6.4 with R' and $V_1, \dots, V_r$, we can find $U_i \subseteq V_i$ such that $m' := |U_i| = (1-3k\varepsilon)m'$ and (U_i, U_j) is $(6k\varepsilon, \delta^+)$-super-regular for all $ij \in E(R')$. We put all vertices in $V_i \setminus U_i$ into V_0, then $|V_0| \leq 3k\varepsilon n$.

We let F'' be $\lfloor \frac{m'}{|V(F)|} \rfloor$ disjoint union of the copies of F', which has equitable vertex k-partition. Let K be a copy of K_k in R', and let $V_K := \bigcup_{i \in K} U_i$. By using blow-up lemma, we can embed F'' into $G'[V_K]$ for each K which is a copy of K_k in R'. After doing this, the only unpacked vertices are vertices in V_0 and vertices in V_i' which is not packed. Note that $|V_0| \leq 2k\varepsilon n$, since the number of deleted vertices are at most $r \times \frac{2k\varepsilon n}{r} \leq 3k\varepsilon n$, and the number of vertices in V_K which is not packed is at most $k \times m' - |V(F')| \times \frac{m'}{|V(F)|} \leq r|V(F')|$. Thus the number of unpacked vertices is at most

$$2k\varepsilon n + r|V(F)| \leq \alpha n.$$

Thus we found $(1 - \alpha)n/|V(F)|$ vertex disjoint copies of F in $G' \subseteq G$. □

Exercise 6.5. *For given $\alpha > 0$, show that there exists n_0 such that the following holds for $n \geq n_0$. If G is an n-vertex graph with $\delta(G) \geq (\frac{2}{3} + \alpha)n$ and H is an $(1 - \alpha)n$-vertex graph with $\Delta(H) \leq 2$, then H embeds into G. (Hint. Can we partition $V(G)$ into U_1 and U_2 such that $\delta(G[U_i]) \geq (\frac{2}{3} + \alpha/2)|U_i|$ for each $i \in [2]$?.)*

The proof of Theorem 6.4 illustrates how we can use the blow-up lemma to prove graph embedding problems. However, the situation was simple because H contains less vertices than n, and H is very disconnected. What about more general graph, spanning and connected? In order to obtain a spanning copy of H in G, we also need to use vertices in the exceptional set V_0.

For $k \geq 2$, we defined G^k, **k-th power of** G such that

$$V(G^k) = V(G) \text{ and } E(G^k) = \{uv : \text{ distance between } u \text{ and } v \text{ in } G \text{ is at most } k\}.$$

For a graph G, we call G^2 **the square of** G. A classical results of Dirac asserts that any n-vertex graph with minimum degree at least $n/2$ contains a Hamilton cycle. As a generalization of this Pósa and Seymour conjectured the followings.

Conjecture 6.5 (Pósa, 1962)**.** *Let G be an n-vetex graph with $\delta(G) \geq 2n/3$, then G contains the square of a Hamilton cycle.*

Conjecture 6.6 (Seymour, 1974)**.** *Let G be an n-vetex graph with $\delta(G) \geq kn/(k + 1)$, then G contains the k-th power of a Hamilton cycle.*

Komlós, Sárközy and Szemerédi [14] proved Pósa's conjecture for sufficiently large n. In [16] Komlós, Sárközy and Szemerédi proved that for large n the Seymour's conjecture holds. Both proofs of the theorems use the regularity lemma and the blow-up lemma. In fact, there is a more general theorem than this, which is called 'bandwidth theorem'. We will not define what bandwidth is but we will define the following more general concept of 'separability'. (For bounded degree graphs, low separability and low bandwidth are roughly equivalent. So we will use the concept of separability in this lecture.)

Definition 6.7. *An n-vertex graph H is η-separable if it contains a set $S \subseteq V(H)$ with $|S| \leq \eta n$ such that every component of $H - S$ has size at most ηn.*

It is easy to see that n-vertex cycle or k-th power of an n-vertex cycle is η-separable for any η, assuming n is sufficiently large.

Theorem 6.8 (Böttcher, Schacht and Taraz [3])**.** *For given $k, \Delta \in \mathbb{N}$ and $\alpha > 0$, there exists $\eta > 0$ and n_0 such that the following holds for all $n \geq n_0$. If H is an n-vertex k-chromatic graph with $\Delta(H) \leq \Delta$ which is η-separable. If G is an n-vertex graph with $\delta(G) \geq (1 - \frac{1}{k} + \alpha)n$, then H embeds into G.*

This theorem is originally stated using the concept of 'bandwidth', thus it is called the bandwidth theorem. Again, the proof utilizes the regularity lemma and the blow-up lemma. Now we prove the theorem for bipartite graphs as follows. The following proof is a simplified version of the proof in [5]. The theorem in [5] deals with finding many edge-disjoint copies of H in G, hence much more complicated than the proof here.

Theorem 6.9. *For given $\Delta \in \mathbb{N}$ and $\alpha > 0$, there exists $\eta > 0$ and n_0 such that the following holds for all $n \geq n_0$. If H is an n-vertex bipartite graph with $\Delta(H) \leq \Delta$ which is η-separable. If G is an n-vertex graph with $\delta(G) \geq (\frac{1}{2} + \alpha)n$, then H embeds into G.*

Proof. We choose constants $n_0, \eta, M, \varepsilon, d$ as follows.

$$0 < 1/n_0 \ll \eta \ll 1/M \ll \varepsilon \ll d \ll \alpha, 1/\Delta.$$

Note that η is very small, so that $\eta \ll 1/r$ holds when r is the size of the ε-regular partition we obtain from the regularity lemma applied with the parameter ε, M.

First, we apply the regularity lemma to G with parameters ε, M, and obtain ε-regular partition $V_0 \cup \dots \cup V_r$ with $1/n_0 \ll \eta \ll 1/r < 1/M \ll \varepsilon$ and consider (ε, d)-reduced graph R on vertex set $[r]$ of the partition. Let $m := |V_i|$ for each $i \in [r]$. By moving vertices in V_r to V_0 if necessary, we may assume that r is even and $|V_0| \leq 2\varepsilon n$. (Note that $m = (1 \pm 2\varepsilon)n/r$.) Since $\alpha \geq 2d$, by using Exercise 4.6, we have

$$\delta(R) \geq (\frac{1}{2} + \alpha/2)r. \tag{6.1}$$

By Dirac's theorem, R contains a perfect matching. By permuting indices, assume that the perfect matching is $R' = \{12, 34, \dots (r-1)r\}$. For each $i \in [r]$, let i^* be the unique number in $[r]$ such that $ii^* \in R'$.

By applying Exercise 6.4 with R' and $V_1, \dots, V_r$, we can find $U_i \subseteq V_i$ such that $m' := |U_i| = (1 - 4\varepsilon)n/r$ and (U_i, U_{i^*}) is $(6\varepsilon, d+)$-super-regular for all $i \in [r]$. We move all vertices in $V_i \setminus U_i$ into V_0, then $|V_0| \leq 5\varepsilon n$.

For each $v \in V_0$, we have $d_G(v) \geq (1/2 + \alpha)n$, thus there exists at least $(1/2 + \alpha/2)r$ indices $i \in [r]$ such that $d_G(v; U_i) \geq dm'$. Observe that, by Exercise 4.4, if we add v to U_{i^*}, then (U_i, U_{i^*}) is still super-regular with slightly worse parameter. With that purpose in mind, we distribute vertices in V_0 into sets $U'_1, \dots, U'_r$ such that the following holds.

(U′1) For each $i \in [r]$ and $v \in U'_i$, we have $d_G(v'; U_{i^*}) \geq dm'$.
(U′2) For each $i \in [r]$, we have $|U'_i| \leq \frac{10|V_0|}{r} \leq 100\varepsilon m'$.
(U′3) For each $ii^* \in E(R')$, we have $|U'_i| = |U'_{i^*}| \pm 1$.

Exercise 6.6. *Prove that such distribution is possible. (Hint. For each $v \in V(G)$, there are at least $\alpha r/2$ indices $i \in [r]$ such that $d_G(v; V_i) \geq dm'$ and $d_G(v; V_{i^*}) \geq dm'$.)*

Define

$$U^*_i := U_i \cup U'_i \text{ and } n_i := |U^*_i|.$$

Then by Exercise 4.4, we have the following.

(U*1) For each $ii^* \in E(R')$, the graph $G[U_i^*, U_{i^*}^*]$ is $(\varepsilon^{1/2}, d+)$-super-regular with $n_i = n_{i^*} \pm 1$.
(U*2) $n_i = (1 \pm \varepsilon^{1/2})n/r$.

Now we have a desired partition of G, and we aim to partition the vertices of H into appropriate vertex sets with size $n_1, \dots, n_r$ and apply the blow-up lemma.

In order to prepare H, first we partition H by using its separability. Let $A_1 \cup A_2$ be the vertex partition of $V(H)$ into two independent sets. As H is η-separable, we can partition H into $W_0, W_1, \dots, W_t$ such that the following holds.

(W1) For each $s \in [t] \cup \{0\}$, we have $|W_s| \leq \eta n$.
(W2) For all $s \neq s' \in [t]$, H has no edges between W_s and $W_{s'}$.
(W3) $\eta^{-1} \leq t \leq 10\eta^{-1}$.

All edges of H are either within each W_s or between W_0 and W_s. For each $s \in [t] \cup \{0\}$, we define

$$W_{s,1} := W_s \cap A_1 \text{ and } W_{s,2} := W_s \cap A_2$$

be the corresponding bipartition.

We will first partition $\{W_{0,1}, W_{0,2}, W_{1,1}, W_{1,2}, W_{2,1}, \dots, W_{t,2}\}$ into collections $\mathcal{W}_1, \dots, \mathcal{W}_r$ as follows. Later when we apply the blow-up lemma, we aim to embed the most of vertices in $\bigcup_{W \in \mathcal{W}_i} W$ into U_i^*.

($\mathcal{W}$1) For all $ii^* \in E(R')$ and $s \in [t] \cup \{0\}$ and $\ell \in [2]$, if $W_{s,\ell} \in \mathcal{W}_i$ then $W_{s,3-\ell} \in \mathcal{W}_{i^*}$.
($\mathcal{W}$2) For all $i \in [r]$, we have $\sum_{W \in \mathcal{W}_i} |W| = (1 \pm \eta^{1/2})n_i$.

Indeed, we can find such a partition. For each $s \in [t]$, we independently at random choose a number $i \in [r/2]$ such that $i \in [r/2]$ is chosen with probability $(n_{2i-1} + n_{2i})/n$, and choose $\ell \in [2]$ with probability $1/2$. If $\ell = 1$, then we add $W_{s,1}$ and $W_{s,2}$ to $\mathcal{W}_{2i-1}$ and $\mathcal{W}_{2i}$, respectively and if $\ell = 2$, then we $W_{s,2}$ and $W_{s,1}$ to $\mathcal{W}_{2i-1}$ and $\mathcal{W}_{2i}$, respectively. By using Azuma's inequality, we may prove the following.

Exercise 6.7. *Prove that in the above random experiment, both properties ($\mathcal{W}$1) and ($\mathcal{W}$2) holds with probability at least 0.99.*

By permuting indices again, assume that $W_{0,1} \in \mathcal{W}_1$ and $W_{0,2} \in \mathcal{W}_2$. For each $i \in [r]$, let $W_i^* := \bigcup_{W \in \mathcal{W}_i} W$. If later we embed W_i^* into U_i^*, then the edges inside each W_s embeds into $G[U_i^*, U_j^*]$ which is super-regular, which is what we want. However, the edges between W_0 and W_s can be problematic. Hence, we want to further modify the partition.

Assign vertices in $W_{0,1}$ to Y_1 and $W_{0,2}$ to Y_2. For each $s \in [t]$ and $i \in [r]$, assume that $W_{s,1}$ belongs to $\mathcal{W}_i$ and $W_{s,2}$ belongs to $\mathcal{W}_{i^*}$. By using (6.1), we know $d_R(2, i^*) \geq dr$, thus we can find the following index $q(s)$.

(Q1) $q(s) \in N_R(2, i^*)$.

Now, we assign vertices in $W_{s,1} \cup W_{s,2}$ which is close to $W_{0,1} \cup W_{0,2}$ as follows.

- Assign vertices in $N_H^1(W_{0,1}) \cap W_{s,2}$ to Y_2.
- Assign vertices in $N_H^1(W_{0,2}) \cap W_{s,1}$ to $Y_{q(s)}$.
- Assign vertices in $N_H^2(W_{0,1}) \cap W_{s,1}$ to $Y_{q(s)}$.
- Assign the rest of vertices of $W_{s,1}$ to X_i and the rest of the vertices of $W_{s,2}$ to X_{i^*}.

By (Q1),

If there exists an edge between Y_i and $Y_j \cup X_j$, then $ij \in E(R)$ and if there exists an edge between X_i and X_j, then $ij \in E(R')$.

Note that every vertices in Y_i is distance at most two away from W_0. As $\Delta(H) \leq \Delta$ and $|W_0| \leq \eta n$, we have $|Y_i| \leq 2\Delta^2 \eta n$. Thus, for each $i \in [r]$, by (U*2) and ($\mathcal{W}$2), we have

$$n_i' := |X_i \cup Y_i| = (1 \pm \eta^{1/3})n_i. \tag{6.2}$$

Now we have sets $X_i \cup Y_i$ which has 'almost correct' sizes. However, we want to obtain sets with 'exactly correct' sizes. For this, we will not move vertices in H anymore, but instead, we will move some vertices in U_i^* to some U_j^*, so that the resulting partition $\widetilde{U}_1, \dots, \widetilde{U}_r$ have size exactly $n'_1, \dots, n'_r$ while $G[\widetilde{U}_i, \widetilde{U}_j]$ is still super-regular with slightly worse parameters. For this, we consider the following directed graph D such that $\overrightarrow{ij} \in D$ implies that we can move some in U_i^* to U_j^*.

Claim 4. *There exists a multi-directed graph D on vertex set $[r]$ satisfying the following.*

(D1) $|E(D)| \le \eta^{1/5} n$.
(D2) *For each $i \in [r]$, we have $d_D^+(i) - d_D^-(i) = n_i - n'_i$.*
(D3) *For each $\overrightarrow{ij} \in E(D)$, we have $ij^* \in E(R)$.*

Proof. Let D' be a directed graph on the vertex set $[r]$ such that $\overrightarrow{ij} \in E(D')$ if and only if $ij^* \in E(R)$.

Exercise 6.8. *Use (6.1) to prove that D' is strongly connected. (For any $i \in [r]$ and $j \in [r]$, there exists a directed path from i to j)*

Note that $\sum_{i\in[r]} n_i = \sum_{i\in[r]} n'_i = n$. Let

$$I^+ := \{i \in [r] : n_i > n'_i\} \text{ and } I^- := \{i \in [r] : n_i < n'_i\}.$$

Then $\sum_{i\in I^+} n_i - n'_i = \sum_{i\in I^-} n'_i - n_i$. This shows that we can find a collection of (not necessarily distinct) pairs $\mathcal{P} := \{(i_1, j_1), \dots, (j_a, j_a)\}$ such that each pair is in $I^+ \times I^-$ such that the following holds.

(P1) $|\mathcal{P}| = \sum_{i\in I^+} n_i - n'_i = \sum_{i\in I^-} n'_i - n_i \overset{(6.2)}{\le} r\eta^{1/3} n$.
(P2) For each $i \in I^+$, we have $|\{b \in [a] : i_b = i\}| = n_i - n'_i$.
(P3) For each $i \in I^-$, we have $|\{b \in [a] : j_b = i\}| = n'_i - n_i$.

By Exercise 6.8, for each $(i, j) \in \mathcal{P}$, we can find a path $P_{i,j}$ of D' with length at most r. We take a disjoint union of all paths of P_{i_b,j_b} for each $b \in [a]$ to obtain a path. In other words, we consider a multi-directed graph D on vertex set $[r]$ such that for each $(i, j) \in [r] \times [r]$, the multi-digraph D contains exactly

$$|\{b \in [a] : \overrightarrow{ij} \in P_{i_b,j_b}\}|$$

directed edges from i to j.

As each path has length at most r, (P1) implies that $|E(D)| \le r^2\eta^{1/3} n \le \eta^{1/5} n$, thus (D1) holds. For each $i \in [r]$, we have

$$d_D^+(i) - d_D^-(i) = |\{b \in [a] : i \text{ is in } P_{i_b,j_b} - \{j_b\}\}| - |\{b \in [a] : i \text{ is in } P_{i_b,j_b} - \{i_b\}\}| \overset{(P2),(P3)}{=} n_i - n'_i.$$

We obtain the final equality by considering two cases of $i \in I^+$ and $i \in I^-$ separately. Thus (D2) holds. It is easy to see that the definition of D and D' implies (D3). This proves the claim. □

Now we use Claim 4 to modify the partition $U_1^*, \dots, U_r^*$ into the following partition. For each $\overrightarrow{e} \in E(D)$ from i to j, we know that $ij^* \in E(R)$. As U_i^* is a slightly modification of V_i, we know that $G[U_i^*, U_{j^*}^*]$ is still $(\varepsilon^{1/3}, d+)$-regular. By Exercise 6.2, at least $(1-\varepsilon^{1/3})n_i$ vertices in U_i^* has at least $(d - \varepsilon^{1/3})n/r$ neighbors in U_k^*. We move one such vertex to U_j^*. We do this for all $\overrightarrow{ij} \in E(D)$, and we indicate the resulting vertex set $\widetilde{U}_i$.

By (D1), for $ij \in E(R')$, we have moved at most $\eta^{1/5} n < \varepsilon n/(2r)$ vertices, we obtain $\widetilde{U}_i$ from U_i^* by removing at most $\varepsilon|U_i^*|$ vertices and adding at most $\varepsilon|U_i^*|$ new vertices which all has at least $(d - 2\varepsilon^{1/3})n/r - \varepsilon|U_j^*|$ neighbors in $\widetilde{U}_j$. Hence Exercise 4.3, Exercise 4.4 and the definition of super-regularity together imply that

($\tilde{\mathrm{U}}$1) For each $i \in [r]$, $G[\tilde{U}_i, \tilde{U}_{i^*}]$ is $(\varepsilon^{1/5}, d/2)$-super-regular.
($\tilde{\mathrm{U}}$2) For each $ij \in E(R)$, $G[\tilde{U}_i, \tilde{U}_j]$ is $(\varepsilon^{1/5}, d+)$-regular.
($\tilde{\mathrm{U}}$3) For each $i \in [r]$, we have $|\tilde{U}_i| \overset{\text{(D2)}}{=} n'_i$.

Now, we use Exercise 4.7 to embed all vertices in Y_i into $\tilde{U}_i$ by an embedding ϕ in such a way that the following holds where $Y := \bigcup_{i\in[r]} Y_i$.

$$\text{For all } i \in [r] \text{ and } x \in X_i, \text{ we have } |A_x| \geq (d - 2\varepsilon^{1/5})^{\Delta} n'_i, \text{ where } A_x := \bigcap_{z\in N_H(x;Y)} N_G(\phi(z); \tilde{U}_i) \tag{6.3}$$

Now, we want to embed X_i into $\tilde{U}_i \setminus \phi(Y_i)$ by using the blow-up lemma. As all edges in $H[\bigcup_{i'\in[r]} X_{i'}]$ are between X_i and X_j with $ij \in E(R')$, and $|X_i| = n_i - |Y_i| = |\tilde{U}_i \setminus \phi(Y_i)| = (1 \pm \varepsilon^{1/10})n/r$, we can use the blow-up lemma(Theorem 6.2). Furthermore, by using the Remark 4 after the blow-up lemma, we can ensure that $x \in X_i$ embeds into A_x. Recall that every vertices in Y is distance at most two away from W_0. By (W1), there are at most $\Delta|Y| \leq 2\Delta^3\eta n < \varepsilon|X_i|$ vertices in X_i which has a neighbor in Y, thus those vertices we need to specify its 'target set' is small. By (6.3), such a target set has not too small size. Further if x embeds into A_x, then this gives us an embedding of the entire graph H into G. This finishes the proof of the Theorem. □

Exercise 6.9. *In the proof of Theorem 6.9, replace the Dirac's theorem with Hajnal-Szemeredi's theorem to prove the following weakening of the Theorem 6.8. (Hint. replace $q(s)$ with an appropriate sequence of numbers and change the definition of D' in an appropriate way.)*

For given $k, \Delta \in \mathbb{N}$ and $\alpha > 0$, there exists $\eta > 0$ and n_0 such that the following holds for all $n \geq n_0$. If H is an n-vertex k-chromatic graph with $\Delta(H) \leq \Delta$ which is η-separable. If G is an n-vertex graph with $\delta(G) \geq (1 - \frac{1}{2(k-1)} + \alpha)n$, then H embeds into G.

Exercise 6.10. *Read Section 3.1 in [16] and prove Theorem 6.8. (Hint. Again replace $q(s)$ with a sequence of numbers as in [16]. In this case, D' might not be strongly connected. Can we still find a specific index $i \in [r]$ such that there exists a direct path from i to j for every $j \in [r]$? How can we change the above proof to complete the task?)*

7. The absorbing method

One downside of using the regularity lemma and the blow-up lemma is that it requires the number of vertices in the host graph G to be huge. To overcome this disadvantage, the absorbing method has been introduced. The absorbing method aims to do the following.

- Build an absorber with good properties.
- Build and approximate structures.
- Use absorber to convert the approximate structure into an optimal structure. (The absorber 'absorbs' the remaining vertices into the approximate structure.)

Now we will prove the following approximate version of Pósa's conjecture by using absorbing method. This will illustrate the absorbing method. This proof is a simplified version of the proof in [20]. Note that the proof in [20] proves Theorem 7.1 with $\alpha = 0$, verifying Pósa's conjecture for large n.

We say that for given a square $(u_1, u_2, \ldots, u_{\ell-1}, u_\ell)$ of path P, we say that (u_1, u_2) and $(u_{\ell-1}, u_\ell)$ are *endpairs* of P.

Theorem 7.1. *For given $\alpha > 0$, there exists n_0 such that the following holds for all $n \geq n_0$. If G is an n-vertex graph with $\delta(G) \geq (2/3 + \alpha)n$, then G contains a square of Hamilton cycle.*

To prove this theorem, we first prove the following lemmas.

Lemma 7.2 (Connecting lemma)**.** *For given $0 < \beta < 1/100$, there exists an n_0 such that the following holds for all $n \geq n_0$. If H is an n-vertex graph with $\delta(H) \geq (2/3 + \beta)n$, then for two disjoint edges $ab, cd \in E(H)$, there exists a square of path $(a, b, u_1, \ldots, u_j, c, d)$ in H with $j \leq 20\beta^{-2}$.*

Proof. For a, b, we consider the following sets $A_0, A_1, \ldots$ and bipartite graph $G_1, \ldots$ such that G_i is a bipartite graph on $A_{i-1} \cup A_i$. Let $A_0 := \{b\}$ and $A_1 = N_H(a, b)$, and let G_1 be a complete bipartite graph between A_0 and A_1, note that $|A_1| \geq (1/3 + 2\beta)n$. Let

$$A'_2 := \bigcup_{x \in A_1} N_H(b, x) \text{ and } G'_2 := \{xy \in A_1 \times A'_2 : y \in N_H(b, x)\}.$$

For each $x \in A_1$, we have $d_{G'_2}(x) \geq (1/3 + \beta)n$. Let

$$A_2 := \{y \in A'_2 : d_{G'_2}(y) > \beta^2 n\} \text{ and } G_2 := G'_2[A_1 \cup A_2].$$

Note that $y \in A_2$ has many choices of x such that $yxba$ forms a square of path. It is easy to verify that $|A_2| \geq n/3$ holds.

Assume that we have constructed $A_0, \ldots, A_j$ and $G_1, \ldots, G_j$ with $j \geq 2$, and we construct A_{j+1} and graph G_{j+1} between A_j and A_{j+1}.

We further aim that for any $yz \in E(G_{j+1})$ with $y \in A_j$, we want to make sure that there are many choices of vertex $x \in A_{j-1}$ (moreover $xy \in E(G_j)$) such that $x \in N_H(y, z)$. In this way, we can 'climb back' to $A_j, A_{j-1}, \ldots A_0$ to form a square of path $zyx \ldots ba$. To this purpose, we first define B^j_y and A'_{i+1}, G'_{i+1} as follows.

For each $y \in A_j$, let

$$\begin{aligned} E(B^j_y) &:= \{xz : x \in N_{G_j}(y), z \in V(H) \text{ and } z \in N_H(x, y)\}, \\ A'_{j+1} &:= \{z : \exists y \in A_j \text{ such that } d_{B^j_y}(z) \geq \beta^4 n\} \text{ and} \\ G'_{j+1} &:= \{yz \in E(H) : y \in A_j, d_{B^j_y}(z) \geq \beta^4 n\}. \end{aligned}$$

In other words, $yz \in G'_+j1$ if there are many vertices x in $N_{G_j}(y)$ such that xyz forms a triangle. Note that $d_{B^j_y}(z) \subseteq N_{G_j}(y)$. Let

$$A_{j+1} := \{z \in A'_{j+1} : d_{G'_{j+1}}(z) > \beta^2 n\} \text{ and } G_{j+1} := G'_{j+1}[A_j \cup A_{j+1}].$$

Note that A_{j+1} is obtained by deleting vertices z which does not have many choices of $y \in A_j$ to 'climb up'. So, for every vertex $z \in A_{j+1}$, there are many choices of $y \in A_j$ and $x \in A_{j-1}$ to 'climb up'.

Note that the sets $A_0, A_1, A_2, \dots$ may not be disjoint. We ensured the 'many choices' because we want to avoid using an already-used vertices again when we 'climb up' to extend the square of the path.

Our strategy is to consider the above sets and graphs and use two vertices c, d analogously to consider another sets $C_0, C_1, \dots$ and bipartite graphs $F_1, \dots$. This will give us two squares of paths, but the problem is how to merge them into a one square of path. For this we first prove the following Claim.

Claim 5. *For all $j \geq 2$ and $y \in A_j$, we have $d_{G'_{j+1}}(y) \geq (1/3 + \beta)n$.*

Proof. Let s be the number of vertices $z \in V(H)$ with $d_{B^j_y}(z) < \beta^4 n$, then $n - s = d_{G'_{j+1}}(y)$. Then we have

$$s\beta^4 n + (n-s)|N_{G_j}(y)| \geq |E(B^j_y)| = \sum_{x \in N_{G_j}(y)} d_H(y,x) \geq |N_{G_j}(y)|(1/3 + 2\beta)n.$$

Note that $|N_{G_j}(y)| > \beta^2 n$ and $s \leq n$, hence we obtain

$$n - s \geq (1/3 + 2\beta)n - \frac{s\beta^4 n}{|N_{G_j}(y)|} \geq (1/3 + \beta)n.$$

This proves the claim. □

We say that a vertex $y \in A_j$ is *j-heavy* if $d_{G_j}(y) \geq n/3$. The purpose of Claim 5 is to find a heavy vertex. We consider this definition of heavy vertex for the following reason.

If y is j-heavy, then there are many choices of x so that we can obtain a square of path $yx \dots ba$. One advantage of y being a heavy vertex is that $d_{G_j}(y) \geq n/3$ and $\delta(H) \geq (2/3 + \beta)n$, we can freely choose a vertex z, then $|N_{G_j}(y) \cap N_H(z)| \geq \beta n$. Hence, there are still many choices of x such that there is a square of path $zyx \dots ba$.

Hence, we hope to find heavy vertex y_a for a and b, and another heavy vertex y_c for d and c. And choose appropriate vertices z_a, z_c, u (recall that we have much freedom to choose these vertices as y_a, y_c are heavy vertices) so that $y_a z_a u z_c y_c$ can 'merge' two square of paths. The following claim provides the existence of heavy vertices.

Claim 6. *There exists $j \leq 5\beta^{-2}$ such that A_j contains a j-heavy vertices.*

Proof. Suppose not. We will show that

$$|A_j| \geq (1 + \beta^2)|A_{j-1}| \tag{7.1}$$

holds for all $j \leq 5\beta^{-2}$ to derive a contradiction. Choose smallest $j > 2$ such that (7.1) does not hold. By the choice, $|A_{j-1}| \geq n/3$ as $|A_2| \geq n/3$.

For each $j \leq 5\beta^{-2}$, Claim 5 implies that

$$|E(G_j)| = \sum_{y \in A_{j-1}} d_{G'_j}(y) - \sum_{z \in A'_j \setminus A_j} d_{G'_j}(z) \geq (1/3 + \beta)n|A_{j-1}| - \beta^2 n^2.$$

On the other hand, as there are no heavy vertices, we have

$$|E(G_j)| \leq \sum_{z \in A_j, \text{ not heavy}} d_{G_{j+1}}(z) \leq \sum_{z \in A_j} n/3 \leq (n/3)|A_j|.$$

From this and the fact that $|A_{j-1}| \geq n/3$, we conclude that

$$|A_j| \geq (1 + \beta^2)|A_{j-1}|.$$

this proves (7.1). On the other hand, then for $j_0 := \lfloor 5\beta^{-2} \rfloor$, we have $|A_{j_0}| \geq (1 + \beta^2)^{j_0 - 2}|A_2| > n$, a contradiction. This proves the claim. □

By Claim 6, we have $j(1) \leq 5\beta^{-2}$ and sets $A_0, A_1, \dots$ and graphs $G_1, \dots$ such that $A_{j(1)}$ contains a heavy vertex y_a. Similarly, for and edge cd, we can construct sets $C_0, C_1, \dots,$ and graphs $F_1, \dots$ and $j(2) \leq 5\beta^{-2}$ such that $C_{j(2)}$ contains a heavy vertex y_c.

Choose a vertex $u \notin \{a, b, c, d\}$ which is a common neighbor of y_a and y_c. As y_a and y_c are heavy, both $N_{G_{j(1)}}(y_a)$ and $N_{G_{j(2)}}(y_c)$ has size at least $n/3$. As $\delta(H) \geq (2/3 + \beta)n$, every vertex $z \in N_H(y_a)$ has at least βn neighbors in $N_{G_{j(1)}}(y_a)$, thus z is also a neighbor of y_a in $G'_{j(1)+1}$, by the definition of $G'_{j(1)+1}$. Hence

$$d_{G'_{j(1)+1}}(y_a) \geq 2n/3 \text{ and } d_{G'_{j(1)+1}}(y_c) \geq 2n/3.$$

Thus we can find an edge $z_a z_c$ of H between two sets $N_{G'_{j(1)+1}}(y_a) \cap N_H(u)$ and $N_{G'_{j(1)+1}}(y_c) \cap N_H(u)$ of size at least $n/3$ (because $\delta(H) \geq (2/3 + \beta)n$). Then note that $y_a z_a u z_c y_c$ is a part of square of a path. By the definition of $A_0, A_1, \dots$ and C_0, C_1, we can extend this into a square of a path of a form $(a, b, u_1, \dots, u_j, c, d)$ with $j \leq 20\beta^{-2}$. This proves the lemma. □

Next, we consider the following Absorbing lemma.

Lemma 7.3 (Absorbing lemma)**.** *For given $0 < \beta < 1/100$, there exists an n_0 such that the following holds for all $n \geq n_0$. If G is an n-vertex graph with $\delta(G) \geq (2/3 + \beta)n$, there is a square-path P_A in G with at most $\beta^{10}n$ vertices such that for every subset $U \subseteq V(G) \setminus V(P_A)$ of size at most $\beta^{20}n$, there exists a square-path $P_{A,U}$ in G such that $V(P_{A,U}) = V(P_A) \cup U$ and $P_{A,U}$ has the same endpairs as P_A.*

Proof. We say that an (ordered) 5-tuple (x, a, b, c, d) of vertices *absorbs* v, if all five vertices are neighbors of v and all a, b, c, d are neighbors of x and a, b, c, d forms a path in G.

Claim 7. *For each $v \in V(G)$, there are at least $2\beta^5 n^5$ many 5-tuples absorbing v.*

Proof. For given v, we can choose x in at least $(2/3 + \beta)n$ different ways. Consider the vertex set $U = N_G(v, x)$, which has size at least $(1/3 + 2\beta)n$. By the minimum degree condition, we have $\delta(G[U]) \geq 3\beta n$. We can choose $a \in U$ in $|U|$ different ways and choose $b \in N_{G[U]}(a)$, $c \in N_{G[U]}(b)$ and $d \in N_{G[U]}(c)$ in each $3\beta n$ different ways. Hence, there are at least

$$(2/3 + \beta)n|U|(3\beta n)^3 \geq 2\beta^5 n^5$$

5-tuples absorbing v. This proves the claim. □

Let $\mathcal{A}_v$ be the family of all 5-tuples absorbing v. Then we have $|\mathcal{A}_v| \geq 2\beta^5 n^5$ for all $v \in V(G)$. Now we prove the following claim.

Claim 8. *There exists a family $\mathcal{F}$ of at most $2\beta^{14}n$ disjoint 5-tuples of vertices of G such that for every $v \in V(G)$, we have $|\mathcal{A}_v \cap \mathcal{F}| \geq \beta^{20}n$.*

Proof. We select a family $\mathcal{F}'$ of 5-tuples of vertices at random by including each of $n(n-1)(n-2)(n-3)(n-4)$ tuples independently with probability $\beta^{14}n^{-4}$.

Exercise 7.1. *Use Chernoff's inequality, show that we have the following with probability at least* 0.99*.*

(F1) *$|\mathcal{F}'| \leq 2\beta^{14}n$ and*
(F2) *for each $v \in V(G)$, we have $|\mathcal{A}_v \cap \mathcal{F}'| \geq \beta^{19}n$.*

Furthermore, the expected number of intersecting pairs of 5-tuples in $\mathcal{F}'$ is at most

$$n^5 \times 5 \times 5 \times n^4 \times (\beta^{14}n^{-4})^2 = 25\beta^{28}n.$$

By Markov's inequality, with probability at least $1/2$,

(F3) there are at most $50\beta^{28}n$ pairs of intersecting 5-tuples in $\mathcal{F}'$.

Thus, with probability at least $0.99 + 1/2 - 1 \geq 0.49$, a random family $\mathcal{F}'$ satisfies all (F1)–(F3). Hence, there exists a family $\mathcal{F}'$ satisfying all three. From this, we delete all 5-tuples that intersect other 5-tuples and all 5-tuples which does not absorb any vertices, and let $\mathcal{F}$ be the resulting collection. By using (F2), for each $v \in V(G)$, we have

$$|\mathcal{A}_v \cap \mathcal{F}| \geq \beta^{19} n - 50\beta^{28} n \geq \beta^{20} n.$$

It is easy to check that $\mathcal{F}$ has the correct size. This proves the claim. $\square$

Now, we consider a collection $\mathcal{F} = \{F_1, \dots, F_f\}$ as Claim 8. For each end pairs of (a, b) of F_i and (c, d) of F_{j+1}, as $|\mathcal{F}| \leq 2\beta^{14} n$, for any set $U' \subseteq V(G)$ with $|U'| \leq \beta n/2$, the graph

$$G[(V(G) \setminus (\bigcup_{i \in [f]} V(F_i) \cup U)) \cup \{a, b, c, d\}]$$

has minimum degree at least $(2/3 + \beta/3)n$. Hence, we can use the Connecting Lemma to find vertices C_i such that F_i, C_i and F_{i+1} forms a square of path, and $|C_i| \leq 20\beta^{-2}$. By include C_i to U and repeating this, we may obtain a path P_A, which has length at most

$$(20\beta^{-2} + 5) \times f \leq 2\beta^{14} \times 30\beta^{-2} n \leq \beta^{10} n.$$

It is easy to check that this is the desired path. This proves the lemma. $\square$

Exercise 7.2 (Reservoir Lemma)**.** *For given $0 < \beta < 1/100$, there exists n_0 such that the following holds for all $n \geq n_0$. If G is an n-vertex graph with $\delta(G) \geq (2/3 + \beta)n$, then there exists a subset $R \subseteq G$ with $|R| = \beta^{20} n/2 \pm n^{2/3}$ such that for every $v \in V(G)$, we have*

$$d_G(v; R) \geq (2/3 + \beta/2)|R|.$$

Proof of Theorem 7.1. Assume that $\alpha < 1/100$ and $\beta = \alpha/3$, and $\varepsilon := \alpha^{100}$. Assume that n_0 is large enough so that three previous lemmas all holds whenever $n \geq n_0/2$.

By applying the Absorbing Lemma 7.3, we find an absorbing square-path P_A with $|P_A| \leq \alpha^{10} n$.

Note that $G \setminus V(P_A)$ has minimum degree at least $(2/3 + \beta)|G \setminus V(P_A)|$. Thus we can apply exercise 7.2 to $G \setminus V(P_A)$ to obtain a set R satisfying the following.

(R1) $|R| = (1 \pm 1/2)\beta^{20} n/2$ and every vertex $v \in V(G)$ satisfies $d_G(v; R) \geq (2/3 + \beta/2)|R|$.

Note that for any set $|U| \leq \beta|R|/10$ and four vertices $a, b, c, d \in G \setminus R$, the graph $G' = G[(R \setminus U) \cup \{a, b, c, d\}]$ has minimum degree at least $(2/3 + \beta)|V(G')|$ (note that $|V(P_A) \cup R| \leq 2\beta n$), thus the Connecting Lemma implies that we can connect the pairs ab and cd through the vertices in R.

Now, we will extend P_A into a longer path covering all vertices in $V(G) \setminus R$ except at most $\alpha^{20} n/2$ vertices and some vertices in R. By repeatedly taking common neighbors of an end pair of the current path, we can extend P_A into a square-path P which has length at least $n/3$ and shares an end pair with P_A.

Let $P' := P \setminus P_A$ and $P' = (x_1, \dots, x_\ell)$. Let $T := V(G) \setminus (V(P) \cup R)$. If $|T| < \varepsilon n$, then we are done.

For a given vertex $a \in T$, we define a string $I_a = (i_1, \dots, i_\ell) \in \{0, 1\}^\ell$ such that $i_j = 1$ if and only if x_j is a neighbor of a. We may suppose that I_a contains no four consecutive ones, otherwise we can extend P' by inserting a in the middle of its four neighbors.

We call a zero followed by a maximal consecutive ones a *j-block* if the maximum consecutive ones has size j. We call a zero followed by another a 0-block. Then I_a consists of disjoint j-blocks for $j \in \{0, 1, 2, 3\}$. We call a substring of I_a an *interval* if it starts with a

3-block and it ends right before the next 3-block. Then $I_a = I_a^1 \dots I_a^{\ell'}$ such that each I_a^j is an interval. We call an interval *heavy* if all blocks are 2-blocks except the leading 3-block.

Assume that the following holds.

(S1) There exists $i \in [\ell]$ such that $S_i := \{a \in T : I_a \text{ has a heavy interval starting at } x_1\}$ and the minimum length of the heavy intervals beginning at x_i is less than $3|S_i|-2$.

If (S1) holds, then we define the operation SWAP. Let $a \in S_i$ be a vertex whose heavy internal beginning at x_i is of minimum length, say $3k+1$. Take a subpath $Q' \subseteq P'$ of length $3k+4$ consisting of vertices right after x_i,

$$Q' = (y_1, y_2, y_3, z_1, y_4, y_5, z_2, y_6, y_7, \dots, z_{k-1}, y_{2k}, y_{2k+1}, z_k, y_{2k+2}, y_{2k+3}, y_{2k+4}).$$

We have $y_i \in N_G(a)$ and $z_i \notin N_G(a)$, and the substring of I_a corresponding to Q' is

$$(1, 1, 1, 0, 1, 1, 0, 1, 1, \dots, 0, 1, 1, 0, 1, 1, 1).$$

In fact, by the minimality of a, for every $b \in S_i$, the substring of I_b corresponding to Q' is the same except the last one. As $|S_i| \geq k+1$, we can choose $b_1, \dots, b_k \in S_i$. Then we can replace Q' in P with the following

$$Q'' := (y_1, y_2, b_1, y_3, y_4, b_2, y_5, y_6, b_3, \dots, y_{2k}, b_k, y_{2k+1}, y_{2k+2}, a, y_{2k+3}, y_{2k+4}).$$

Then we obtain a path longer than P'. By using this operation, we can prove the following claim.

Claim 9. *If there exists a subset $U \subseteq T$ of size εn such that for all $a \in U$, we have $d_G(a; V(P')) \geq (2/3+\varepsilon)\ell$, then we can extend P.*

Proof. For $a \in U$, assume I_a contains t heavy intervals, such that the size of the union of the heavy intervals is s. Then we have

$$d_G(a; V(P')) \leq \frac{2}{3}(s-t) + t + \frac{2}{3}(\ell - s) \leq 2\ell/3 + t/3.$$

Thus

$$t \geq 3(d_G(a; V(P') - 2\ell/3) \geq 3\varepsilon\ell.$$

Hence, for every $a \in U$, the string I_a contains at least $\varepsilon\ell/3$ heavy intervals. Let a heavy interval *short* if it is of length less than $10\varepsilon^{-1}$. Then for each $a \in U$, there are at least $\varepsilon\ell/2$ short heavy interval in I_a, otherwise, the size of the union of heavy interval is longer than

$$(\varepsilon\ell - \varepsilon\ell/2) \cdot 10\varepsilon^{-1} \geq 5\ell > n,$$

a contradiction as $\ell \geq n/3$. By pigeonhole principle, there exists $i \in [\ell]$ such that short heavy interval begins at x_i for at least $(|U| \cdot \varepsilon\ell/2)/\ell \geq \varepsilon^2 n/2$ distinct vertices in T. Let S_i be those vertices, then as $|S_i| \geq \varepsilon^2 n/2 > 3 \cdot 10\varepsilon^{-1} - 2$, we can perform the operation SWAP at i. This proves the claim. □

Let $U := \{a \in T : d_G(a; V(P')) \geq (2/3+\varepsilon)\ell\}$. Note that by Claim 9, we can assume that $|U| \leq \varepsilon n$. Then, for $a \in T \setminus U$, we have

$$d_G(a; T \setminus U) \geq (\frac{2}{3} + \alpha)n - (\frac{2}{3} + \varepsilon)\ell - |R| - |P_A| \geq \frac{2}{3}|T| + \alpha n/3.$$

Thus, this implies that any pair of vertices in $T \setminus U$ has at least $\alpha n/3$ common neighbors. This guarantees a square-path in $T \setminus U$ of length $\alpha n/3$. Let (a, b) be an end pair of this path. Use Connecting Lemma to find a square-path between (b, a) and $(x_\ell, x_{\ell-1})$ of length at most $20\beta^{-2}$ in $G[R \cup \{a, b, x_\ell, x_{\ell-1}\}]$, then this extends P' by adding at least $\alpha n/3$ more vertices.

Note that by repeating the above procedure, we use Connecting Lemma at most $3\alpha^{-1}$ times, as P' can not be longer than n. Hence, we only use at most $3\alpha^{-1} \times 20\beta^{-2} \leq \alpha^{-4}$, we can keep using the Connecting Lemma until the path P covers all but at most εn vertices in $V(G) \setminus R$.

By repeating this we can obtain a path P covering all but εn vertices in $V(G) \setminus R$ and at most α^{-4} vertices in R. Let ab and cd be the endpairs of P. We use Connecting Lemma to find a square-path between ab and cd through $R \setminus V(P)$, then we obtain a square-cycle C. Note that P_A is a subpath of C. Since $|V(G) \setminus V(C)| < \alpha^{20}n$, by Absorbing Lemma, there exists a path $P_{A,V(G)\setminus V(C)}$ sharing both endpairs with P_A covering all vertices of $V(P_A) \cup (V(G) \setminus V(C))$. This gives a square of Hamilton cycle.

□

This absorbing method was successful in proving other graph embedding results and hypergraph embedding results. See [17] or [23] for more results using the absorbing method.

References

[1] N. Alon, L. Rónyai and T. Szabó, *Norm-graphs: variations and applications*, J. Combin. Theor. Ser. B 76, (1999), 280–290.

[2] N. Alon and J. H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[3] J. Böttcher, M. Schacht and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Math. Ann. **343** (2009), 175–205.

[4] B. Bukh and Z. Jiang, *A bound on the number of edges in graphs without an even cycle*, Combin. Probab. Comput. 26(1), (2017), 1–15.

[5] P, Condon, J. Kim, D. Kühn and D. Osthus, *A bandwidth theorem for approximate decompositions*, arXiv:1712.04562, (2018).

[6] V. Chvátal, V. Rödl, E. Szemerédi, W. Trotter, *The Ramsey number of a graph with bounded maximum degree*, J. Combin. Theory Ser. B **34** (1983), 239–243.

[7] D. Conlon, *Notes for a course on extremal graph theory*, http://people.maths.ox.ac.uk/conlond/.

[8] D. Conlon, J. Fox, and B. Sudakov, *On two problems in graph Ramsey theory*, Combinatorica **32** (2012), 513–535.

[9] J. Fox and B. Sudakov, *Dependent random choice*, Random Structures Algorithm 38, (2011), 1–32.

[10] H. A. Kierstead and A. V. Kostochka, *A short proof of the Hajnal-Szemerédi Theorem on equitable coloring*, Combin. Prob. and Comput. 17, (2008), 265–270.

[11] J. Kollár, L. Rónyai and T. Szabó, *Norm-graphs and bipartite Turán numbers*, Combinatorica 16, (1996), 399–406.

[12] J. Komlós, G.N. Sárközy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[13] J. Komlós, G.N. Sárközy and E. Szemerédi, *Proof of the Alon-Yuster conjecture*, Discrete Math. **235** (2001), 255–269

[14] J. Komlós, G. N. Sárközy, and E. Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Random Structures Algorithms **9** (1996) 193–211.

[15] J. Komlós, G. N. Sárközy and E. Szemerédi, *On the Pósa-Seymour conjecture*, J. Graph Theory, **29** (1998) 167–176

[16] J. Komlós, G. N. Sárközy and E. Szemerédi, *Proof of the Seymour Conjecture for Large Graphs*, Ann. Comb. **2** (1998) 43–60

[17] M. Krivelevich, *Triangle Factors in Random Graphs*, Combin. Prob. and Comput. 6 (1997), 337–347.

[18] C. Lee, *Lecture note of Extremal Combinatorics*, http://math.mit.edu/ cb_lee/teaching.html, 2015.

[19] C. Lee, *Ramsey numbers of degenerate graphs*, Ann. of Math. 185 (2017), 791–829.

[20] I. Levitt, G. N. Sárközy and E. Szemerédi, *How to avoid using the Regularity Lemma: Pósa's conjecture revisited*, Disc. Math. 310, (2010) 630–641.

[21] M. Molloy and B. Reed, *Graph colouring and the probabilistic method*, Springer-Verlag Berlin Heidelberg, 2002

[22] O. Pikhurko, *A Note on the Turán Function of Even Cycles*, Proc Amer. Math. Soc. 140, (2012) 3687–3992.

[23] V. Rödl, A .Ruciński and E. Szemerédi, *A Dirac-Type Theorem for* 3*-Uniform Hypergraphs*, Combin. Prob. and Comput. 15, (2006) 229–251.

[24] B. Sudakov, *A few remarks on the Ramsey-Turán-type problems*, J. Combin. Theor. Ser. B 88, (2003), 99–106.

[25] P. Turán, *On an extremal problem in graph theory*, Matematikai és Fizikai Lapok (in Hungarian) 48, (1941), 436–452.

[26] D. West, *Combinatorial mathematics*, unpublished.