On Embedding Degree Sequences

Béla Csaba
Bolyai Institute, Interdisciplinary Excellence Centre, University of Szeged, Hungary
E-mail: csaba@math.u-szeged.hu

Bálint Vásárhelyi
University of Szeged, Hungary
E-mail: mesti@math.u-szeged.hu

Keywords: graph theory, degree sequence, embedding

Received: October 30, 2018

Assume that we are given two graphic sequences, π_1 and π_2. We consider conditions for π_1 and π_2 which guarantee that there exists a simple graph G_2 realizing π_2 such that G_2 is the subgraph of any simple graph G_1 that realizes π_1.

Povzetek: Recimo, da imamo grafni zaporedji π_1 in π_2. V prispevku preučujemo pogoje za zaporedji, ki zagotavljajo, da obstaja preprost graf G_2, ki je realizacija π_2 in je podgraf grafa G_1, ki je realizacija zaporedja π_2.

1 Introduction

All graphs considered in this paper are simple. We use standard graph theory notation, see for example [16]. Let us provide a short list of a few perhaps not so common notions, notations. Given a bipartite graph $G(A, B)$ we call it *balanced* if $|A| = |B|$. This notion naturally generalizes for r-partite graphs with $r \in \mathbb{N}$, $r \geq 2$.

If $S \subset V$ for some graph $G = (V, E)$ then the subgraph spanned by S is denoted by $G[S]$. Moreover, let $Q \subset V$ so that $S \cap Q = \emptyset$, then $G[S, Q]$ denotes the bipartite subgraph of G on vertex classes S and Q, having every edge of G that connects a vertex of S with a vertex of Q. The number of vertices in G is denoted by $v(G)$, the number of its edges is denoted by $e(G)$. The degree of a vertex $x \in V(G)$ is denoted by $deg_G(x)$, or if G is clear from the context, by $deg(x)$. The number of neighbors of x in a subset $S \subset V(G)$ is denoted by $deg_G(x, S)$, and $\delta(G)$ and $\Delta(G)$ denote the minimum and maximum degree of G, respectively. The complete graph on n vertices is denoted by K_n, the complete bipartite graph with vertex class sizes n and m is denoted by $K_{n,m}$.

A finite sequence of natural numbers $\pi = (d_1, \ldots, d_n)$ is a *graphic sequence* or *degree sequence* if there exists a graph G such that π is the (not necessarily) monotone degree sequence of G. Such a graph G *realizes* π. For example, the degree sequence $\pi = (2, 2, \ldots, 2)$ can be realized only by vertex-disjoint union of cycles.

The largest value of π is denoted by $\Delta(\pi)$. Often the positions of π will be identified with the elements of a vertex set V. In this case, we write $\pi(v)$ $(v \in V)$ for the corresponding component of π.

The degree sequence $\pi = (a_1, \ldots, a_k; b_1, \ldots, b_l)$ is a *bigraphic* sequence if there exists a simple bipartite graph $G = G(A, B)$ with $|A| = k$, $|B| = l$ realizing π such that the degrees of vertices in A are $a_1, \ldots, a_k$, and the degrees of the vertices of B are $b_1, \ldots, b_l$.

Let G and H be two graphs on n vertices. We say that H is a subgraph of G, if we can delete edges from G so that we obtain an isomorphic copy of H. We denote this relation by $H \subset G$. In the literature the equivalent complementary formulation can be found as well: we say that H and $\overline{G}$ *pack* if there exist edge-disjoint copies of H and $\overline{G}$ in K_n. Here $\overline{G}$ denotes the *complement* of G.

It is an old an well-understood problem in graph theory to tell whether a given sequence of natural numbers is a degree sequence or not. We consider a generalization of it, which is remotely related to the so-called discrete tomography[1] (or degree sequence packing) problem (see e.g. [5]) as well. The question whether a sequence of n numbers π is a degree sequence can also be formulated as follows: Does K_n have a subgraph H such that the degree sequence of H is π? The question becomes more general if K_n is replaced by some (simple) graph G on n vertices. If the answer is yes, we say that π *can be embedded into* G, or equivalently, π packs with $\overline{G}$.

Let us mention two classical results in extremal graph theory.

Theorem 1 (Dirac, [6]). *Every graph G with $n \geq 3$ vertices and minimum degree $\delta(G) \geq \frac{n}{2}$ has a Hamilton cycle.*

[1]In the discrete tomography problem we are given two degree sequences of length n, π_1 and π_2, and the questions is whether there exists a graph G on n vertices with a red-blue edge coloration so that the following holds: for every vertex v the red degree of v is $\pi_1(v)$ and the blue degree of v is $\pi_2(v)$.

Theorem 2 (Corrádi-Hajnal, [3]). *Let $k \geq 1$, $n \geq 3k$, and let H be an n-vertex graph with $\delta(H) \geq 2k$. Then H contains k vertex-disjoint cycles.*

Observe, that Dirac's theorem implies that given a constant 2 degree sequence π of length n and any graph G on n vertices having minimum degree $\delta(G) \geq n/2$, π can be embedded into G. One can interprete the Corrádi-Hajnal theorem similarly, but here one may require more on the structure of the graph that realizes π and in exchange a larger minimum degree of G is needed.

One of our main results is the following.

Theorem 3. *For every $\eta > 0$ and $D \in \mathbb{N}$ there exists an $n_0 = n_0(\eta, D)$ such that for all $n > n_0$ if G is a graph on n vertices with $\delta(G) \geq \left(\frac{1}{2} + \eta\right) n$ and π is a degree sequence of length n with $\Delta(\pi) \leq D$, then π is embeddable into G.*

It is easy to see that Theorem 3 is sharp up to the ηn additive term. For that let n be an even number, and suppose that every element of π is 1. Then the only graph that realizes π is the union of $n/2$ vertex disjoint edges. Let $G = K_{n/2-1,n/2+1}$ be the complete bipartite graph with vertex class sizes $n/2 - 1$ and $n/2 + 1$. Clearly G does not have $n/2$ vertex disjoint edges.

In order to state the other main result of the paper we introduce a new notion.

Definition 4. *Let $q \geq 1$ be an integer. A bipartite graph H with vertex classes S and T is q-unbalanced, if $q|S| \leq |T|$. The degree sequence π is q-unbalanced, if it can be realized by a q-unbalanced bipartite graph.*

Theorem 5. *Let $q \geq 1$ be an integer. For every $\eta > 0$ and $D \in \mathbb{N}$ there exist an $n_0 = n_0(\eta, q)$ and an $M = M(\eta, D, q)$ such that if $n \geq n_0$, π is a q-unbalanced degree sequence of length $n - M$ with $\Delta(\pi) \leq D$, G is a graph on n vertices with $\delta(G) \geq (\frac{1}{q+1} + \eta)n$, then π can be embedded into G.*

Hence, if π is unbalanced, the minimum degree requirement of Theorem 3 can be substantially decreased, what we pay for this is that the length of π has to be slightly smaller than the number of vertices in the host graph.

2 Proof of Theorem 3

Proof. First, we find a suitable realization H of π, our H will consists of components of bounded size. Second, we embed H into G using a theorem by Chvátal and Szemerédi and a result on embedding so called well-separable graphs. The details are as follows.

We construct H in several steps. At the beginning, let H be the empty graph and let all degrees in π be *active*. While we can find $2i$ active degrees of π with value i (for some $1 \leq i \leq \Delta(\pi)$) we realize them with a $K_{i,i}$ (that is, we add this complete bipartite graph to H, and the $2i$ degrees are "inactivated"). When we stop we have at most $\sum_{i=1}^{\Delta(\pi)}(2i-1)$ active degrees. This way we obtain several components, each being a balanced complete bipartite graph. These are the *type 1 gadgets*. Observe that if a vertex v belongs to some type 1 gadget, then its degree is exactly $\pi(v)$. Observe further that if there are no active degrees in π at this point then the graph H we have just found is a realization of π.

Assume that there are active degrees left in π. Let $R = R_{odd} \cup R_{even}$ be the vertex set that is identified with the active vertices ($v \in R_{odd}$ if and only if the assigned active degree is odd). Since $\sum_{v \in R} d(v)$ must be an even number we have that $|R_{odd}|$ is even. Add a perfect matching on R_{odd} to H. With this we achieved that every vertex of R misses an even number of edges.

Next we construct the *type 2 gadgets* using the following algorithm. In the beginning every type 1 gadget is *unmarked*. Suppose that $v \in R$ is an active vertex. Take a type 1 gadget K, *mark* it, and let M_K denote an arbitrarily chosen perfect matching in K (M_K exists since K is a balanced complete bipartite graph). Let xy be an arbitrary edge in M_K. Delete the xy edge and add the new edges vx and vy. While v is missing edges repeat the above procedure with edges of M_K, until M_K becomes empty. If M_K becomes empty, take a new unmarked type 1 gadget L, and repeat the method with L. It is easy to see that in $\pi(v)/2$ steps v reaches its desired degree and gets inactivated. Clearly, the degrees of vertices in the marked type 1 gadgets have not changed.

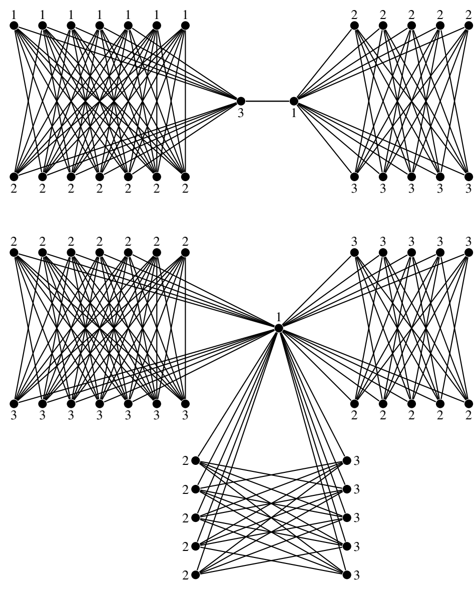

Figure 1: Type 2 gadgets of H with a 3-coloring

Figure 1 shows examples of type 2 gadgets. In the upper one two vertices of R_{odd} were first connected by an edge and then two type 1 gadgets were used so that they could reach their desired degree, while in the lower one we used three type 1 gadgets for a vertex of R. The numbers at the vertices indicate the colors in the 3-coloring of H.

Let $A \subset V(H)$ denote the set of vertices containing the union of all type 2 gadgets. Observe that type 2 gadgets are 3-colorable and all have less than $5\Delta^2(\pi)$ vertices. Let us summarize our knowledge about H for later reference.

Claim 6. *(1)* $|A| \le 5\Delta^3(\pi)$,

(2) the components of $H[V-A]$ *are balanced complete bipartite graphs, each having size at most* $2\Delta(\pi)$,

(3) $\chi(H[A]) \le 3$, *and*

(4) $e(H[A, V-A]) = 0$.

We are going to show that $H \subset G$. For that we first embed the possibly 3-chromatic part $H[A]$ using the following strengthening of the Erdős–Stone theorem proved by Chvátal and Szemerédi [2].

Theorem 7. *Let* $\varphi > 0$ *and assume that* G *is a graph on* n *vertices where* n *is sufficiently large. Let* $r \in \mathbb{N}$, $r \ge 2$. *If*

$$e(G) \ge \left(\frac{r-2}{2(r-1)} + \varphi\right) n^2,$$

then G *contains a* $K_r(t)$, *i.e. a complete* r*-partite graph with* t *vertices in each class, such that*

$$t > \frac{\log n}{500 \log \frac{1}{\varphi}}. \qquad (1)$$

Since $\delta(G) \ge (1/2+\eta)n$, the conditions of Theorem 7 are satisfied with $r = 3$ and $\varphi = \eta/2$, hence, G contains a balanced complete tripartite subgraph T on $\Omega(\log n)$ vertices. Using Claim 6 and the 3-colorability of $H[A]$ this implies that $H[A] \subset T$.

Observe that after embedding $H[A]$ into G every uncovered vertex of G still has at least $\delta(G) - v(F) > (1/2 + \eta/2)n$ uncovered neighbors. Denoting the subgraph of the uncovered vertices of G by G' we obtain that $\delta(G') > (1/2 + \eta/2)n$.

In order to prove that $H[V-A] \subset G'$ we first need a definition.

Definition 8. *A graph* F *on* n *vertices is* well-separable *if it has a subset* $S \subset V(F)$ *of size* $o(n)$ *such that all components of* $F - S$ *are of size* $o(n)$.

We need the following theorem.

Theorem 9 ([4]). *For every* $\gamma > 0$ *and positive integer* D *there exists an* n_0 *such that for all* $n > n_0$ *if* F *is a bipartite well-separable graph on* n *vertices,* $\Delta(F) \le D$ *and* $\delta(G) \ge \left(\frac{1}{2} + \gamma\right) n$ *for a graph* G *of order* n, *then* $F \subset G$.

Since $H[V-A]$ has bounded size components by Claim 6, we can apply Theorem 9 for $H[V-A]$ and G', with parameter $\gamma = \eta/2$. With this we finished proving what was desired.

□

3 Further tools for Theorem 5

When proving Theorem 3, we used the Regularity Lemma of Szemerédi, but implicitly, via the result on embedding well-separable graphs. When proving Theorem 5 we will apply this very powerful result explicitly, hence, below we give a very brief introduction to the area. The interested reader may consult with the original paper by Szemerédi [15] or e.g. with the survey paper [10].

3.1 Regularity lemma

The *density* between disjoint sets X and Y is defined as:

$$d(X,Y) = \frac{e(X,Y)}{|X||Y|}.$$

We will need the following definition to state the Regularity Lemma.

Definition 10 (Regularity condition)**.** *Let* $\varepsilon > 0$. *A pair* (A, B) *of disjoint vertex-sets in* G *is* ε*-regular if for every* $X \subset A$ *and* $Y \subset B$, *satisfying*

$$|X| > \varepsilon|A|, \ |Y| > \varepsilon|B|$$

we have

$$|d(X,Y) - d(A,B)| < \varepsilon.$$

This definition implies that regular pairs are highly uniform bipartite graphs; namely, the density of any reasonably large subgraph is almost the same as the density of the regular pair.

We will use the following form of the Regularity Lemma:

Lemma 11 (Degree Form)**.** *For every* $\varepsilon > 0$ *there is an* $M = M(\varepsilon)$ *such that if* $G = (W, E)$ *is any graph and* $d \in [0,1]$ *is any real number, then there is a partition of the vertex set* V *into* $\ell + 1$ *clusters* $W_0, W_1, \ldots, W_\ell$, *and there is a subgraph* G' *of* G *with the following properties:*

- $\ell \le M$,
- $|W_0| \le \varepsilon|W|$,
- *all clusters* W_i, $i \ge 1$, *are of the same size* $m \left(\le \left\lfloor \frac{|W|}{\ell} \right\rfloor < \varepsilon|W|\right)$,
- $deg_{G'}(v) > deg_G(v) - (d+\varepsilon)|W|$ *for all* $v \in W$,
- $G'|_{W_i} = \emptyset$ *(*W_i *is an independent set in* G'*) for all* $i \ge 1$,
- *all pairs* (W_i, W_j), $1 \le i < j \le \ell$, *are* ε*-regular, each with density either 0 or greater than* d *in* G'.

We call W_0 the *exceptional cluster*, $W_1, \ldots, W_\ell$ are the *non-exceptional clusters*. In the rest of the paper we will assume that $0 < \varepsilon \ll d \ll 1$. Here $a \ll b$ means that a *is sufficiently smaller than* b.

Definition 12 (Reduced graph)**.** *Apply Lemma 11 to the graph $G = (W, E)$ with parameters ε and d, and denote the clusters of the resulting partition by $W_0, W_1, \ldots, W_\ell$ (W_0 being the exceptional cluster). We construct a new graph G_r, the reduced graph of G' in the following way: The non-exceptional clusters of G' are the vertices of the reduced graph G_r (hence $v(G_r) = \ell$). We connect two vertices of G_r by an edge if the corresponding two clusters form an ε-regular pair with density at least d.*

The following corollary is immediate:

Corollary 13. *Apply Lemma 11 with parameters ε and d to the graph $G = (W, E)$ satisfying $\delta(G) \geq \gamma n$ $(v(G) = n)$ for some $\gamma > 0$. Denote G_r the reduced graph of G'. Then $\delta(G_r) \geq (\gamma - \theta)\ell$, where $\theta = 2\varepsilon + d$.*

The (fairly easy) proof of the lemma below can be found in [10].

Lemma 14. *Let (A, B) be an ε-regular–pair with density d for some $\epsilon > 0$. Let $c > 0$ be a constant such that $\varepsilon \ll c$. We arbitrarily divide A and B into two parts, obtaining the non-empty subsets A', A'' and B', B'', respectively. Assume that $|A'|, |A''| \geq c|A|$ and $|B'|, |B''| \geq c|B|$. Then the pairs (A', B'), (A', B''), (A'', B') and (A'', B'') are all ε/c–regular pairs with density at least $d - \varepsilon/c$.*

3.2 Blow-up lemma

Let H and G be two graphs on n vertices. Assume that we want to find an isomorphic copy of H in G. In order to achieve this one can apply a very powerful tool, the Blow-up Lemma of Komlós, Sárközy and Szemerédi [8, 9]. For stating it we need a new notion, a stronger one-sided property of regular pairs.

Definition 15 (Super-Regularity condition)**.** *Given a graph G and two disjoint subsets of its vertices A and B, the pair (A, B) is (ε, δ)*-super-regular*, if it is ε-regular and furthermore,*

$$deg(a) > \delta|B|, \text{ for all } a \in A,$$

and

$$deg(b) > \delta|A|, \text{ for all } b \in B.$$

Theorem 16 (Blow-up Lemma)**.** *Given a graph R of order r and positive integers δ, Δ, there exists a positive $\varepsilon = \varepsilon(\delta, \Delta, r)$ such that the following holds: Let $n_1, n_2, \ldots, n_r$ be arbitrary positive parameters and let us replace the vertices $v_1, v_2, \ldots, v_r$ of R with pairwise disjoint sets $W_1, W_2, \ldots, W_r$ of sizes $n_1, n_2, \ldots, n_r$ (blowing up R). We construct two graphs on the same vertex set $V = \cup_i W_i$. The first graph F is obtained by replacing each edge $v_i v_j \in E(R)$ with the complete bipartite graph between W_i and W_j. A sparser graph G is constructed by replacing each edge $v_i v_j$ arbitrarily with an (ε, δ)-super-regular pair between W_i and W_j. If a graph H with $\Delta(H) \leq \Delta$ is embeddable into F then it is already embeddable into G.*

4 Proof of Theorem 5

Let us give a brief sketch first. Recall, that π is a q-unbalanced and bounded degree sequence with $\Delta(\pi) \leq D$. In the proof we first show that there exists a q-unbalanced bipartite graph H that realizes π such that H is the vertex disjoint union of the graphs $H_1, \ldots, H_k$, where each H_i graph is a bipartite q-unbalanced graph having bounded size. We will apply the Regularity lemma to G, and find a special substructure (a decomposition into vertex-disjoint stars) in the reduced graph of G. This substructure can then be used to embed the union of the H_i graphs, for the majority of them we use the Blow-up lemma.

4.1 Finding H

The goal of this subsection is to prove the lemma below.

Lemma 17. *Let π be a q-unbalanced degree sequence of positive integers with $\Delta(\pi) \leq D$. Then π can be realized by a q-unbalanced bipartite graph H which is the vertex disjoint union of the graphs $H_1, \ldots, H_k$, such that for every i we have that H_i is q-unbalanced, moreover, $v(H_i) \leq 4D^2$.*

Before starting the proof of Lemma 17, we list a few necessary notions and results.

We call a finite sequence of integers a *zero-sum sequence* if the sum of its elements is zero. The following result of Sahs, Sissokho and Torf plays an important role in the proof of Lemma 17.

Proposition 18. *[14] Assume that K is a positive integer. Then any zero-sum sequence on $\{-K, \ldots, K\}$ having length at least $2K$ contains a proper nonempty zero-sum subsequence.*

The following result, formulated by Gale [7] and Ryser [13], will also be useful. We present it in the form as discussed in Lovász [11].

Lemma 19. *[11] Let $G = (A, B; E(G))$ be a bipartite graph and f be a nonnegative integer function on $A \cup B$ with $f(A) = f(B)$. Then G has a subgraph $F = (A, B; E(F))$ such that $d_F(x) = f(x)$ for all $x \in A \cup B$ if and only if*

$$f(X) \leq e(X, Y) + f(\overline{Y}) \qquad (2)$$

for any $X \subseteq A$ and $Y \subseteq B$, where $\overline{Y} = B - Y$.

We remark that such a subgraph F is also called an f-factor of G.

Lemma 20. *If $f = (a_1, \dots, a_s; b_1, \dots, b_t)$ is a sequence of positive integers with $s, t \geq 2\Delta^2$, where Δ is the maximum of f, and $f(A) = f(B)$ with $A = \{a_1, \dots, a_s\}$ and $B = \{b_1, \dots, b_t\}$ then f is bigraphic.*

Proof. All we have to check is whether the conditions of Lemma 19 are met if $G = K_{s,t}$.

Suppose indirectly that there is an (X, Y) pair for which (2) does not hold. Choose such a pair with minimal $|X| + |Y|$. Then $X = \emptyset$ or $Y = \emptyset$ are impossible, as in those cases (2) trivially holds. Hence, $|X|, |Y| \geq 1$. Assuming that (2) does not hold, we have that

$$f(X) \geq e(X, Y) + f(\overline{Y}) + 1, \quad (3)$$

which is equivalent to

$$f(X) \geq |X||Y| + f(\overline{Y}) + 1, \quad (4)$$

as G is a complete bipartite graph. Furthermore, using the minimality of $|X| + |Y|$, we know that

$$f(X - a) \leq |X - a||Y| + f(\overline{Y}) \quad (5)$$

for any $a \in X$. (5) is equivalent to

$$f(X) - f(a) \leq |X||Y| - |Y| + f(\overline{Y}). \quad (6)$$

From (4) and (6) we have

$$f(a) - 1 \geq |Y| \quad (7)$$

for any $a \in X$, which implies

$$\Delta > |Y|. \quad (8)$$

The same reasoning also implies that $\Delta > |X|$ whenever (X, Y) is a counterexample. Therefore we only have to verify that (2) holds in case $|X| < \Delta$ and $|Y| < \Delta$. Recall that $f(B) \geq t$, as all elements of f are positive. Hence, $f(X) \leq \Delta|X| \leq \Delta^2$, and $f(\overline{Y}) = f(B) - f(Y) \geq t - \Delta^2$, and we get that

$$f(X) \leq \Delta^2 \leq t - \Delta^2 \leq f(\overline{Y}) \leq f(\overline{Y}) + e_G(X, Y) \quad (9)$$

holds, since $t \geq 2\Delta^2$. □

Proof. (Lemma 17) Assume that $J = \big(S, T; E(J)\big)$ is a q-unbalanced bipartite graph realizing π. Hence, $q|S| \leq |T|$. Moreover, $|T| \leq D|S|$, since $\Delta(\pi) \leq D$. We form vertex disjoint tuples of the form $(s; t_1, \dots, t_h)$, such that $s \in S$, $t_i \in T$, $q \leq h \leq D$, and the collection of these tuples contains every vertex of $S \cup T$ exactly once. We define the *bias* of the tuple as

$$\zeta = \pi(t_1) + \cdots + \pi(t_h) - \pi(s).$$

Obviously, $-D \leq \zeta \leq D^2$. The conditions of Proposition 18 are clearly met with $K = D^2$. Hence, we can form groups of size at most $2D^2$ in which the sums of biases are zero. This way we obtain a partition of (S, T) into q-unbalanced set pairs which have zero bias. While these sets may be small, we can combine them so that each combined set is of size at least $2D^2$ and has zero bias. By Lemma 20 these are bigraphic sequences. The realizations of these small sequences give the graphs $H_1, \dots, H_k$. It is easy to see that $v(H_i) \leq 4D^2$ for every $1 \leq i \leq k$. Finally, we let $H = \cup_i H_i$. □

4.2 Decomposing G_r

Let us apply the Regularity lemma with parameters $0 < \varepsilon \ll d \ll \eta$. By Corollary 13 we have that $\delta(G_r) \geq \ell/(q+1) + \eta\ell/2$.

Let $h \geq 1$ be an integer. An h-star is a $K_{1,h}$. The *center* of an h-star is the vertex of degree h, the other vertices are the *leaves*. In case $h = 1$ we pick one of the vertices of the 1-star arbitrarily to be the center.

Lemma 21. *The reduced graph G_r has a decomposition $\mathcal{S}$ into vertex disjoint stars such that each star has at most q leaves.*

Proof. Take a partial star-decomposition of G_r that is as large as possible. Assume that there are uncovered vertices in G_r. Let U denote those vertices that are covered (so we assume that U has maximal cardinality), and let v be an uncovered vertex. Observe that v has neighbours only in U, otherwise, if $uv \in E(G_r)$ with $u \notin U$, then we can simply add uv to the star-decomposition, contradicting to the maximality of U. See Figure 2 for the possible neighbors of v.

a) If v is connected to a 1-star, then we can replace it with a 2-star.

b) If v is connected to the center u of an h-star, where $h < q$, then we can replace this star with an $h+1$-star by adding the edge uv to the h-star.

c) If v is connected to a leaf u of an h-star, where $2 \leq h \leq q$, then replace the star with the edge uv and an $(h-1)$-star (i.e., delete u from it).

We have not yet considered one more case: when v is connected to the center of a q-star. However, simple calculation shows that for every vertex v at least one of the above three cases must hold, using the minimum degree condition of G_r. Hence we can increase the number of covered vertices. We arrived at a contradiction, G_r has the desired star-decomposition. □

4.3 Preparing G for the embedding

Consider the q-star-decomposition $\mathcal{S}$ of G_R as in Lemma 21. Let ℓ_i denote the number of $(i-1)$-stars in

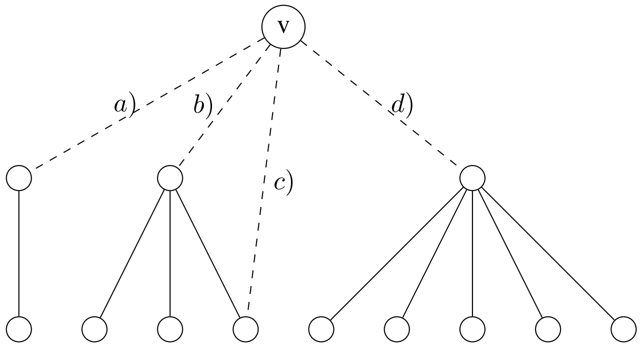

Figure 2: An illustration for Lemma 21

the decomposition for every $2 \le i \le q+1$. It is easy to see that

$$\sum_{i=2}^{q+1} i\ell_i = \ell.$$

First we will make every ε-regular pair in $\mathcal{S}$ super-regular by discarding a few vertices from the non-exceptional clusters. Let for example $\mathcal{C}$ be a star in the decomposition of G_r with center cluster A and leaves $B_1, \dots, B_k$, where $1 \le k \le q$. Recall that the (A, B_i) pairs has density at least d. We repeat the following for every $1 \le i \le k$: if $v \in A$ such that v has at most $2dm/3$ neighbors in B_i then discard v from A, put it into W_0. Similarly, if $w \in B_i$ has at most $2dm/3$ neighbors in A, then discard w from B_i, put it into W_0. Repeat this process for every star in $\mathcal{S}$. We have the following:

Claim 22. *We do not discard more than $q\varepsilon m$ vertices from any non-exceptional cluster.*

Proof. Given a star $\mathcal{C}$ in the decomposition $\mathcal{S}$ assume that its center cluster is A and let B be one of its leaves. Since the pair (A, B) is ε-regular with density at least d, neither A, nor B can have more than εm vertices that have at most $2dm/3$ neighbors in the opposite cluster. Hence, during the above process we may discard up to $q\varepsilon m$ vertices from A. Next, we may discard vertices from the leaves, but since no leaf B had more than εm vertices with less than $(d - \varepsilon)m$ neighbors in A, even after discarding at most $q\varepsilon m$ vertices of A, there can be at most εm vertices in B that have less than $(d - (q + 1)\varepsilon)m$ neighbors in A. Using that $\varepsilon \ll d$, we have that $(d - (q + 1)\varepsilon) > 2d/3$. We obtained what was desired. □

By the above claim we can make every ε-regular pair in $\mathcal{S}$ a $(2\varepsilon, 2d/3)$-super-regular pair so that we discard only relatively few vertices. Notice that we only have an upper bound for the number of discarded vertices, there can be clusters from which we have not put any points into W_0. We repeat the following for every non-exceptional cluster: if s vertices were discarded from it with $s < q\varepsilon m$ then we take $q\varepsilon m - s$ arbitrary vertices of it, and place them into W_0. This way every non-exceptional cluster will have the same number of points, precisely $m - q\varepsilon m$. For simpler notation, we will use the letter m for this new cluster size. Observe that W_0 has increased by $q\varepsilon m\ell$ vertices, but we still have $|W_0| \le 3dn$ since $\varepsilon \ll d$ and $\ell m \le n$. Since $q\varepsilon m \ll d$, in the resulting pairs the minimum degree will be at least $dm/2$.

Summarizing, we obtained the following:

Lemma 23. *By discarding a total of at most $q\varepsilon n$ vertices from the non-exceptional clusters we get that every edge in $\mathcal{S}$ represents a $(2\varepsilon, d/2)$-super-regular pair, and all non-exceptional clusters have the same cardinality, which is denoted by m. Moreover, $|W_0| \le 3dn$.*

Since $v(G) - v(H)$ is bounded above by a constant, when embedding H we need almost every vertex of G, in particular those in the exceptional cluster W_0. For this reason we will assign the vertices of W_0 to the stars in $\mathcal{S}$. This is not done in an arbitrary way.

Definition 24. *Let $v \in W_0$ be a vertex and (Q, T) be an ε-regular pair. We say that $v \in T$* has large degree *to Q if v has at least $\eta|Q|/4$ neighbors in Q. Let $S = (A, B_1, \dots, B_k)$ be a star in $\mathcal{S}$ where A is the center of S and $B_1, \dots, B_k$ are the leaves, here $1 \le k \le q$. If v has large degree to any of $B_1, \dots, B_k$, then v* can be assigned *to A. If $k < q$ and v has large degree to A, then v* can be assigned to *any of the B_i leaves.*

Observation 25. *If we assign new vertices to a q-star, then we necessarily assign them to the center. Since before assigning, the number of vertices in the leaf-clusters is exactly q times the number of vertices in the center cluster, after assigning new vertices to the star, q times the cardinality of the center will be larger than the total number of vertices in the leaf-clusters. If $S \in \mathcal{S}$ is a k-star with $1 \le k < q$, and we assign up to cm vertices to any of its clusters, where $0 < c \ll 1$, then even after assigning new vertices we will have that q times the cardinality of the center is larger than the total number of vertices in the leaf-clusters.*

The following lemma plays a crucial role in the embedding algorithm.

Lemma 26. *Every vertex of W_0 can be assigned to at least $\eta\ell/4$ non-exceptional clusters.*

Proof. Suppose that there exists a vertex $w \in W_0$ that can be assigned to less than $\eta\ell/4$ clusters. If w cannot be assigned to any cluster of some k-star S_k with $k < q$, then the total degree of w into the clusters of S_k is at most $k\eta m/4$. If w cannot be assigned to any cluster of some q-star S_q, then the total degree of w into the clusters of S_q is at most $m + q\eta m/4$, since every vertex of the center cluster could be adjacent to w. Considering that w can be assigned to at most $\eta\ell/4 - 1$ clusters and that $d(w, W - W_0) \ge n/(q+1) + \eta n/2$, we obtain the following inequality:

$$\frac{n}{q+1}+\frac{\eta n}{2}\le d(v,W-W_0)\le \eta\frac{\ell m}{4}+ \sum_{k=1}^{q-1}(k+1)\eta\frac{\ell_{k+1}m}{4}+q\eta\frac{\ell_{q+1}m}{4}+\ell_{q+1}m.$$

Using $m\ell \le n$ and $\sum_{k=1}^{q}(k+1)\ell_{k+1}=\ell$, we get

$$\frac{m\ell}{q+1}+\frac{\eta m\ell}{2}\le \eta\frac{\ell m}{4}+(\ell-\ell_{q+1})\frac{\eta m}{4}+ q\eta\frac{\ell_{q+1}m}{4}+\ell_{q+1}m.$$

Dividing both sides by m and cancellations give

$$\frac{\ell}{q+1}\le q\frac{\eta\ell_{q+1}}{4}+(1-\frac{\eta}{4})\ell_{q+1}.$$

Noting that $(q+1)\ell_{q+1}\le \ell$, one can easily see that we arrived at a contradiction. Hence every vertex of W_0 can be assigned to several non-exceptional clusters. □

Lemma 26 implies the following:

Lemma 27. *One can assign the vertices of W_0 so that at most $\sqrt{d}m$ vertices are assigned to non-exceptional clusters.*

Proof. Since we have at least $\eta\ell/4$ choices for every vertex, the bound follows from the inequality $\frac{4|W_0|}{\eta\ell}\le\sqrt{d}m$, where we used $d\ll\eta$ and $|W_0|\le 3dn$. □

Observation 28. *A key fact is that the number of newly assigned vertices to a cluster is much smaller than their degree into the opposite cluster of the regular pair since $\sqrt{d}m\ll\eta m/4$.*

4.4 The embedding algorithm

The embedding is done in two phases. In the first phase we cover every vertex that belonged to W_0, together with some other vertices of the non-exceptional clusters. In the second phase we are left with super-regular pairs into which we embed what is left from H using the Blow-up lemma.

4.4.1 The first phase

Let (A,B) be an ε-regular cluster-edge in the h-star $\mathcal{C}\in\mathcal{S}$. We begin with partitioning A and B randomly, obtaining $A=A'\cup A''$ and $B=B'\cup B''$ with $A'\cap A''=B'\cap B''=\emptyset$. For every $w\in A$ (except those that came from W_0) flip a coin. If it is heads, we put w into A', otherwise we put it into A''. Similarly, we flip a coin for every $w\in B$ (except those that came from W_0) and depending on the outcome, we either put the vertex into B' or into B''. The proof of the following lemma is standard, uses Chernoff's bound (see in [1]), we omit it.

Lemma 29. *With high probability, that is, with probability at least $1-1/n$, we have the following:*

- $\big||A'|-|A''|\big|=o(n)$ *and* $\big||B'|-|B''|\big|=o(n)$
- $deg(w,A'),deg(w,A'')>deg(w,A)/3$ *for every* $w\in B$
- $deg(w,B'),deg(w,B'')>deg(w,B)/3$ *for every* $w\in A$
- *the density* $d(A',B')\ge d/2$

It is easy to see that Lemma 29 implies that (A',B') is a $(5\varepsilon,d/6)$-super-regular pair having density at least $d/2$ with high probability.

Assume that v was an element of W_0 before we assigned it to the cluster A, and assume further that $deg(v,B)\ge\eta m/4$. Since (A,B) is an edge of the star-decomposition, either A or B must be the center of $\mathcal{C}$.

Let H_i be one of the q-unbalanced bipartite subgraphs of H that has not been embedded yet. We will use H_i to cover v. Denote S_i and T_i the vertex classes of H_i, where $|S_i|\ge q|T_i|$. Let $S_i=\{x_1,\dots,x_s\}$ and $T_i=\{y_1,\dots,y_t\}$.

If A is the center of $\mathcal{C}$ then the vertices of T_i will cover vertices of A', and the vertices of S_i will cover vertices of B'. If B is the center, S_i and T_i will switch roles. The embedding of H_i is essentially identical in both cases, so we will only discuss the case when A is the center.[2]

In order to cover v we will essentially use a well-known method called Key lemma in [10]. We will heavily use the fact that

$$0<\varepsilon\ll d\ll\eta.$$

The details are as follows. We construct an edge-preserving injective mapping $\varphi: S_i\cup T_i\longrightarrow A'\cup B'$. In particular, we will have $\varphi(S_i)\subset B'$ and $\varphi(T_i)-v\subset A'$. First we let $\varphi(y_1)=v$. Set $N_1=N(v)\cap B'$. Using Lemma 29 we have that $|N_1|\ge\eta m/12\gg\varepsilon m$.

Next we find $\varphi(y_2)$. Since $|N_1|\gg\varepsilon m$, by 5ε-regularity the majority of the vacant vertices of A' will have at least $d|N_1|/3$ neighbors in N_1. Pick any of these, denote it by v_2 and let $\varphi(y_2)=v_2$. Also, set $N_2=N_1\cap N(v_2)$.

In general, assume that we have already found the vertices $v_2,v_3,\dots,v_i$, their common neighborhood in B' is N_i, and

$$|N_i|\ge\frac{\eta d^{i-1}}{3^{i-2}\cdot 36}m\gg\varepsilon m.$$

By 5ε-regularity, this implies that the majority of the vacant vertices of A' has large degree into N_i, at least $d\cdot|N_i|/3$, and this, as above, can be used to find v_{i+1}. Then we set $\varphi(y_{i+1})=v_{i+1}$. Since η and d is large compared to ε, even into the last set N_{t-1} many vacant vertices will have large degrees.

[2] Recall that if $h<q$ then we may assigned v to a leaf, so in such a case B could be the center.

As soon as we have $\varphi(y_1), \ldots, \varphi(y_t)$, it is easy to find the images for $x_1, \ldots, x_t$. Since $|N_t| \gg \varepsilon m \gg s = |S_i|$, we can arbitrarily choose s vacant points from N_t for the $\varphi(x_j)$ images.

Note that we use less than $v(H_i) \leq 4D^2$ vertices from A' and B' during this process. We can repeat it for every vertex that were assigned to A, and still at most $\sqrt{d}2D^2m$ vertices will be covered from A' and from B'.

Another observation is that every h-star in the decomposition before this embedding phase was h-unbalanced, now, since we were careful, these have become h'-balanced with $h' \leq h$.

Of course, the above method will be repeated for every (A, B) edge of the decomposition for which we have assigned vertices of W_0.

4.4.2 The second phase

In the second phase we first unite all the randomly partitioned clusters. For example, assume that after covering the vertices that were coming from W_0 the set of vacant vertices of A' is denoted by A'_v. Then we let $A_v = A'_v \cup A''$, and using analogous notation, let $B_v = B'_v \cup B''$.

Claim 30. *All the (A_v, B_v) pairs are $(3\varepsilon, d/6)$-super-regular with density at least $d/2$.*

Proof. The 3ε-regularity of these pairs is easy to see, like the lower bound for the density, since we have only covered relatively few vertices of the clusters. For the large minimum degrees note that by Lemma 29 every vertex of A had at least $dm/6$ neighbors in B'', hence, in B_v as well, and analogous bound holds for vertices of B. □

At this point we want to apply the Blow-up lemma for every star of $\mathcal{S}$ individually. For that we first have to assign those subgraphs of H to stars that were not embedded yet. We need a lemma.

Lemma 31. *Let $K_{a,b}$ be a complete bipartite graph with vertex classes A and B, where $|A| = a$ and $|B| = b$. Assume that $a \leq b = ha$, where $1 \leq h \leq q$. Let H' be the vertex disjoint union of q-unbalanced bipartite graphs:*

$$H' = \bigcup_{j=1}^{t} H_j,$$

such that $v(H_j) \leq 2D^2$ for every j. If $v(H') \leq a + b - 4(2q+1)D^2$, then $H' \subset K_{a,b}$.

Observe that if we have Lemma 31, we can distribute the H_i subgraphs among the stars of $\mathcal{S}$, and then apply the Blow-up lemma. Hence, we are done with proving Theorem 5 if we prove Lemma 31 above.

Proof. The proof is an assigning algorithm and its analysis. We assign the vertex classes of the H_j subgraphs to A and B, one-by-one. Before assigning the jth subgraph H_j the number of vacant vertices of A is denoted by a_j and the number of vacant vertices of B is denoted by b_j.

Assume that we want to assign H_k. If $ha_k - b_k > 0$, then the larger vertex class of H_k is assigned to A, the smaller is assigned to B. Otherwise, if $ha_k - b_k \leq 0$, then we assign the larger vertex class to B and the smaller one to A. Then we update the number of vacant vertices of A and B. Observe that using this assigning method we always have $a_k \leq b_k$.

The question is whether we have enough room for H_k. If $ha \geq 4hD^2$, then we must have enough room, since $b_k \geq a_k$ and every H_j has at most $2D^2$ vertices. Hence, if the algorithm stops, we must have $a_k < 4D^2$. Since $b_k - ha_k \leq 2D^2$ must hold, we have $b_k < (2h+1)2D^2 < (2q+1)2D^2$. From this the lemma follows. □

5 Remarks

One can prove a very similar result to Theorem 5, in fact the result below follows easily from it. For stating it we need the notion of graph edit distance which is detailed e.g. in [12]: the edit distance between two graphs on the same labeled vertex set is defined to be the size of the symmetric difference of the edge sets

Theorem 32. *Let $q \geq 1$ be an integer. For every $\eta > 0$ and $D \in \mathbb{N}$ there exists an $n_0 = n_0(\eta, q)$ and a $K = K(\eta, D, q)$ such that if $n \geq n_0$, π is a q-unbalanced degree sequence of length n with $\Delta(\pi) \leq D$, G is a graph on n vertices with $\delta(G) \geq (\frac{1}{q+1} + \eta)n$, then there exists a graph G' on n vertices such that the edit distance of G and G' is at most K, and π can be embedded into G'.*

Here is an example showing that Theorem 5 and 32 are essentially best possible.

Example 33. *Assume that π has only odd numbers and G has at least one odd sized component. Then the embedding is impossible. Indeed, any realization of π has only even sized components, hence, G cannot contain it as a spanning subgraph.*

Note that this example does not work in case G is connected. In Theorem 3 the minimum degree $\delta(G) \geq (1/2 + \eta)n$, hence, G is connected, and in this case we can embed π into G.

Acknowledgement

Partially supported by Ministry of Human Capacities, Hungary, grant 20391-3/2018/FEKUSTRAT, by ERC-AdG. 321400 and by the National Research, Development and Innovation Office - NKFIH Fund No. SNN-117879 and KH 129597.

Supported by TÁMOP-4.2.2.B-15/1/KONV-2015-0006.

References

[1] N. Alon, J. Spencer. *The probabilistic method.* Third edition, John Wiley & Sons, Inc., 2008. https://doi.org/10.1002/9780470277331

[2] V. Chvátal and E. Szemerédi. On the Erdős–Stone Theorem. *Journal of the London Mathematical Society*, s2-23 (2):207–214, 1981.

[3] K. Corrádi and A. Hajnal. On the maximal number of independent circuits in a graph. *Acta Mathematica Academiae Scientiarum Hungaricae*, 14:423–439, 1963. https://doi.org/10.1007/BF01895727

[4] B. Csaba. On embedding well-separable graphs. *Discrete Mathematics*, 308(19):4322–4331, 2008. https://doi.org/10.1016/j.disc.2007.08.015

[5] J. Diemunsch, M. Ferrara, S. Jahanbekam, and J. Shook. Extremal theorems for degree sequence packing and the 2-color discrete tomography problem. *SIAM Journal of Discrete Mathematics*, 29(4):2088–2099, 2015. https://doi.org/10.1137/140987912

[6] G. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, s3-2(1):69–81, 1952. https://doi.org/10.1112/plms/s3-2.1.69

[7] D. Gale. A theorem on flows in networks. *Pacific Journal of Mathematics*, 7(2):1073–1082, 1957. https://doi.org/10.2140/pjm.1957.7.1073

[8] J. Komlós, G. Sárközy, and E. Szemerédi. Blow-up lemma. *Combinatorica*, 17:109–123, 1997. https://doi.org/10.1007/BF01196135

[9] J. Komlós, G. Sárközy, and E. Szemerédi. An algorithmic version of the blow-up lemma. *Random Structures and Algorithms*, 12:297–312, 1998. https://doi.org/10.1002/(SICI)1098-2418(199805)12:3<297::AID-RSA5>3.0.CO;2-Q

[10] J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. *Combinatorics: Paul Erdős is eighty*, 2:295–352, 1993.

[11] L. Lovász. *Combinatorial problems and exercises*. Corrected reprint of the 1993 second edition, AMS Chelsea Publishing, Providence, RI, 2007.

[12] R. Martin. The edit distance in graphs: Methods, results, and generalizations. In A. Beveridge, J. Griggs, L. Hogben, G. Musiker, and P. Tetali, editors, *Recent Trends in Combinatorics*, pages 31–62. Springer International Publishing, Cham, 2016.

[13] H. Ryser. Combinatorial properties of matrices of zeros and ones. *Canadian Journal of Mathematics*, 9:371–377, 1957. https://doi.org/10.4153/CJM-1957-044-3

[14] M. Sahs, P. Sissokho, and J. Torf. A zero-sum theorem over $\mathbb{Z}$. *Integers*, 13:#A70, 2013.

[15] E. Szemerédi. Regular partitions of graphs. *Colloques Internationaux C.N.R.S* ***260*** *- Problèmes Combinatoires et Théorie des Graphes*, pages 399–401, 1976.

[16] D. West. *Introduction to graph theory*. Prentice Hall, Second edition, 2001.