PACKING DEGENERATE GRAPHS GREEDILY

PETER ALLEN, JULIA BÖTTCHER, JAN HLADKÝ, AND DIANA PIGUET

ABSTRACT. We prove that if $\mathcal{G}$ is a family of graphs with at most n vertices each, with constant degeneracy, with maximum degree at most $O(n/\log n)$, and with total number of edges at most $(1-o(1))\binom{n}{2}$, then $\mathcal{G}$ packs into the complete graph K_n. This strengthens recent results of Böttcher–Hladký–Piguet–Taraz, Messuti–Rödl–Schacht, Ferber–Lee–Mousset, Kim–Kühn–Osthus–Tyomkyn, and Ferber–Samotij related to the Tree Packing Conjecture.

In this extended abstract we describe the main steps of our proof.

1. INTRODUCTION

A *packing* of a family $\mathcal{G} = \{G_1, \ldots, G_k\}$ of graphs into a host graph H is a colouring of the edges of H with the colours $0, 1, \ldots, k$ such that the edges of colour i form an isomorphic copy of G_i for each $1 \le i \le k$. Graph packing problems can be considered as a common generalisation of a number of the most important lines of investigation in Extremal Graph Theory: Turán-type problems, Dirac-type problems, and Ramsey-type problems.

The focus of this research announcement is on packings of large connected graphs that either exhaust all (so-called *perfect packings*) or almost all the edges of the host graph H (so-called *near-perfect packings)*. Historically the first and still the most famous problems in this direction concern the packing of trees. In 1963 Ringel [9] conjectured that if T is any $n+1$-vertex tree, then $2n+1$ copies of T pack into K_n, and in 1976 Gyárfás [5] proposed the *Tree Packing Conjecture*, stating that, if T_i is an i-vertex tree for each $1 \le i \le n$, then $\{T_1, \ldots, T_n\}$ packs into K_n. Since we have $(2n+1) \cdot e(T) = \binom{n}{2}$ and $\sum e(T_i) = \binom{n}{2}$, both conjectures ask for perfect packings. Despite many partial results (which mostly deal with very restricted classes of trees) both these problems were wide open until quite recently.

The first near-perfect packing result in the direction of these packing conjectures for trees was obtained by Böttcher, Hladký, Piguet and Taraz [1], who showed that one can pack into K_n any family of trees whose maximum degree is at most Δ, whose order is at most $(1-\delta)n$, and whose total number of edges is at most $(1-\delta)\binom{n}{2}$, provided that n is sufficiently large given the constants $\Delta \in \mathbb{N}$ and $\delta > 0$. This approximately answers both Ringel's Conjecture and the Tree Packing Conjecture for bounded degree trees. Various generalisations of this result were obtained in quick succession. Messuti, Rödl and Schacht [8] showed that one can replace trees with graphs from any nontrivial minor-closed family (satisfying all other conditions). Then, Ferber, Lee and Mousset [2] improved on this result by allowing the graphs to be packed to be spanning (but still only providing a near-perfect packing). Kim, Kühn, Osthus and Tyomkyn [7] proved a near-perfect packing result for families of graphs with bounded maximum degree which are otherwise unrestricted. Joos, Kim, Kühn and Osthus [6] solved both Ringel's conjecture and the Tree Packing conjecture for trees of bounded maximum degree. Relaxing the restriction on the maximum degree, Ferber and Samotij [3] gave two near-perfect packing results for trees, one for spanning trees, and one for almost spanning trees. In these results, they allow the maximum degrees to be as big $O\big(n^{1/6}/\log^6 n\big)$, and $O\big(n/\log n\big)$, respectively.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 628974.

Hladký and Piguet were supported by the Czech Science Foundation, grant number GJ16-07822Y.

Our new result is a near-perfect packing theorem in the complete graph for spanning graphs with bounded degeneracy and maximum degrees up to $O\big(n/\log n\big)$, extending the mentioned packing results in [1, 2, 3, 7, 8].[1] For a graph G, a linear ordering of its vertex set $V(G)$ is *D-degenerate* if every vertex has at most D neighbours preceding it (we call these the *left-neighbours*). The graph G is D-degenerate if $V(G)$ has a D-degenerate ordering. We remark that such graphs G may be expanding and may have very high maximum degree.

Theorem 1. *For each $\gamma > 0$ and each $D \in \mathbb{N}$ there exist $c > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for each $n > n_0$. Suppose that $(G_t)_{t\in[t^*]}$ is a family of D-degenerate graphs, each of which has at most n vertices and maximum degree at most $\frac{cn}{\log n}$. Suppose further that the total number of edges of $(G_t)_{t\in[t^*]}$ is at most $(1-\gamma)\binom{n}{2}$. Then $(G_t)_{t\in[t^*]}$ packs into K_n.*

In the remainder we sketch our proof of Theorem 1. In order to make the presentation of this sketch more accessible we shall simplify the description of our proof. In particular, the auxiliary results required to prove Theorem 1 are technically more involved than the ones stated here.

2. Outline of the main idea

We shall call the host graph H; it turns out that it is convenient not to restrict ourselves to $H = K_n$ but to a more general setting of quasirandom graphs (defined below).

The main idea of our proof is as follows. We shall first select almost spanning subgraphs G'_i of the graphs G_i in the given family. Then we shall use a random embedding process to embed the graphs G'_i one by one edge-disjointly into the host graph H (deleting any edges of H that we used). We shall show that this random embedding process with high probability preserves three invariants: quasirandomness of the host graph, and the diet and cover conditions (defined in Section 3). Further, we prove that, as long as these invariants are satisfied, the random embedding process can successfully continue. Finally, we complete the packing of the almost spanning G'_i to a packing of the G_i by using a matching argument.

An n-vertex graph with $p\binom{n}{2}$ edges is *(ε, Δ)-quasirandom* if the common neighbourhood $\mathsf{N}(S)$ of any set S of at most Δ vertices has size $(1 \pm \varepsilon)p^{|S|}n$. We remark that this notion of quasirandomness is slightly stronger than the standard notion, in which the neighbourhoods of most, rather than all, sets S are controlled.

We now state formally under what conditions we can pack the almost-spanning graphs $\big(G'_t\big)$. For the promised completion to $\big(G_t\big)$ we actually need a slightly stronger statement, which we will sketch later; however the proof of Theorem 2 already contains all the difficulties.

Theorem 2. *For each $\nu > 0$ and each $D \in \mathbb{N}$ there exist $c, \varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for each $n > n_0$. Suppose that $(G'_t)_{t\in[t^*]}$ is a family of $t^* \leq 2n$ many D-degenerate graphs, each of which has at most $(1-\nu)n$ vertices and maximum degree at most $\frac{cn}{\log n}$. Suppose that H is a $(\varepsilon, 2D+3)$-quasirandom graph of order n. Suppose further that the total number of edges of $(G'_t)_{t\in[t^*]}$ is at most $e(H) - \nu\binom{n}{2}$. Then $(G'_t)_{t\in[t^*]}$ packs into H.*

We outline the proof of Theorem 2 in Section 3. Let us now explain why in Theorem 2 we need to pack not into K_n but into any quasirandom subgraph of K_n. Suppose that $(G_t)_{t\in[t^*]}$ is a family as in Theorem 1. We first deal with the possibility $t^* > 2n$ by modifying the family. We remove any isolated vertices from all graphs in the family, and so we obtain $v(G_t) \leq 2e(G_t)$ for each $t \in [t^*]$. Now, given G_{t_1} and G_{t_2} both of which have at most $n/4$ edges and hence at most $n/2$ vertices, we merge them into a single graph $G_{t_1} \sqcup G_{t_2}$ with at most n vertices. Repeating this procedure until no further merging is possible, we end up with t^* graphs each having at least $n/4$ edges; since the total number of edges in the family is at most $\binom{n}{2}$ we have $t^* \leq 2n$, as is required in Theorem 2. Any

[1]While our result extends these results in the setting of complete host graphs, the main focus of [3] is on packing into random graphs, and [7] provides a general packing result in the setting of the Regularity lemma.

packing of the modified family (which we still call $(G_t)_{t\in[t^*]}$) trivially gives a packing of the original family.

Next, we want to find a subgraph $G'_t \subseteq G_t$ for each $t \in [t^*]$ of order at most $(1-\nu)n$. Theorem 2 gives us a packing of the $\big(G'_t\big)$; we want to choose G'_t in order to make it easy to extend this packing to a packing of the $\big(G_t\big)$. It turns out to be convenient to find an independent set I_t in G_t, all of whose vertices have the same degree, and that degree should be at most $2D$, and set $G'_t = G_t - I_t$. We obtain I_t with the following simple lemma.

Lemma 3. *Let G be a D-degenerate n-vertex graph. Then there exists an integer $0 \le d \le 2D$ and a set $I \subseteq V(G)$ with $|I| \ge (2D+1)^{-3}n$ which is independent, and all of whose vertices have the same degree d in G.*

Now Theorem 2 gives a packing of the (G'_t) into any n-vertex sufficiently quasirandom graph H with nearly $\binom{n}{2}$ edges. To complete the derivation of Theorem 1 we need to explain how we choose H inside K_n and how we complete the packing of the (G'_t) to a packing of the (G_t).

We choose H by taking away a random subgraph from K_n, and we let $H^* = K_n - H$. We choose the number of edges in H large enough that Theorem 2 applies, but small enough that H^* contains much more than $\sum_{t\in[t^*]} e(G_t) - e(G'_t)$ edges. We apply Theorem 2 to pack the (G'_t) into H, and then for each $t \in [t^*]$ we find a way to complete the copy of G'_t in H to a copy of G_t in K_n using edges of H^*. The vertices of G_t remaining to embed are an independent set I_t. Each vertex $x \in I_t$ has $d \le 2D$ neighbours $y_1, \dots, y_d$ in G_t, which are all in G'_t and hence already embedded to vertices $v_1, \dots, v_d$ of K_n. Now we complete the embeddings of the G_t, starting with $t = 1$. For $t = 1$, we only allow embedding x to vertices in the candidate set

$$C(x) := \big\{u \in K_n \colon u \notin \mathrm{im}(G'_t), uv_1, \dots, uv_d \in H^*\big\}\,,$$

and we simply need to match the vertices of I_t to the vertices of K_n such that each x is matched to a vertex of $C(x)$. To see that this matching exists, we need to verify Hall's condition. Part of the strengthening of Theorem 2 that we need to do this roughly states that the sets $C(x)$ are distributed in a random-like fashion. It is straightforward to argue from this that Hall's condition holds. Since all vertices of I_t have d neighbours, all the sets $C(x)$ have about the same size, which makes this argument easier.

For $t \ge 2$, of course when we want to complete the embedding of G_t we should not use edges of H^* which were used to complete any of $G_1, \dots, G_{t-1}$, and the definition of $C(x)$ must change accordingly. The other strengthening of Theorem 2 that we require is that the vertices adjacent to those in I_t are embedded to sets distributed in a random-like fashion. This means that during the entire packing process we will use only a few edges of H^* at each vertex, and the verification of Hall's condition is robust enough to allow for such a change.

3. Proof of Theorem 2

In this section we outline the proof of Theorem 2. We will not explain how to obtain the strengthenings sketched in the previous section that we need for Theorem 1. However, this strengthening turns out not to require any fundamentally different ideas.

The vertices of the graphs G'_t will be always the first $v(G'_t)$ natural numbers, in a degeneracy order. We proceed by packing the graphs $G'_1, \dots, G'_{t^*}$ one by one in this order and call the randomised algorithm which embeds the graph G'_t *RandomEmbedding*. The graphs $H =: H_0 \supset H_1 \supset \dots \supset H_{t^*}$ record the host graph edges remaining throughout the process. At a given stage $t = 1, \dots, t^*$, we proceed as follows. We need to embed the graph G'_t into H_{t-1}. We embed the vertex 1 into H_{t-1} uniformly at random. Having embedded vertices $1, \dots, j-1$ of G'_t to H_{t-1}, we need to embed the vertex j. We simply pick a valid choice uniformly at random. In other words, we choose uniformly an image for j from the set of vertices $x \in V(H_{t-1})$ to which we have not embedded any vertex $1, \dots, j$ of G'_t, and which are adjacent to all of the embedded left-neighbours of j. If this set is ever empty then *RandomEmbedding* fails; if for each stage $t \in [t^*]$ and $j \in V(G'_t)$ it is not empty, then the sequence of *RandomEmbedding*s gives an embedding of each G'_t into H_{t-1}, hence a packing of the $\big(G'_t\big)$ into H. Therefore we need to analyse the evolution of $(H_t)_{t\in[t^*]}$ and the run of *RandomEmbedding* at each stage t. In order to analyse the run of *RandomEmbedding* at stage t, we need H_{t-1} to be very quasirandom; on the other hand, the graph H_t will be a little less quasirandom than H_{t-1}. We set

$$\alpha_x = C^{-1}\exp\Big(\frac{C(x-2n)}{n}\Big)$$

for some large constant C. The required quasirandomness for $H = H_0$ is α_0; note that this quantity does not depend on n. Our strategy is to prove that with high probability the sequence of *RandomEmbedding*s does not fail and each of the graphs H_i is $(\alpha_i, 2D+3)$-quasirandom. The following two lemmas are key to the analysis.

Lemma 4. *Suppose that an n-vertex graph H is $(\alpha, 2D+3)$-quasirandom with $p\binom{n}{2}$ edges for some small α and $p \gg \alpha$. The probability that* RandomEmbedding *fails when embedding a D-degenerate graph G of order at most $(1-\nu)n$ into H is $o(1/n)$.*

Lemma 5. *Suppose that we are in the setting described above Lemma 4. As the graphs $G'_1, G'_2, \dots$ are embedded one by one, for each t the following holds. Provided that H_i is $(\alpha_i, 2D+3)$-quasirandom for each $0 \le i < t$ and that* RandomEmbedding *does not fail before the end of stage t, the probability that H_{t-1} fails to be $(\alpha_t, 2D+3)$-quasirandom is $o(1/n)$.*

These two lemmas imply Theorem 2. Indeed, if some *RandomEmbedding* fails, then there must be a first time t when either *RandomEmbedding* fails although H_{t-1} is quasirandom, or *RandomEmbedding* succeeds but the resulting H_t s not quasirandom. Lemmas 4 and 5 respectively state that these two events have probability $o(1/n)$; taking the union bound over times t, the probability of failure is $o(1)$.

3.1. **Sketch of the proof of Lemma 4.** Recall that *RandomEmbedding* fails when it comes to a vertex j and there is no valid choice of an image for j. Since H is quasirandom, if j has d left-neighbours then about $p^d n$ vertices in H are adjacent to all these embedded left-neighbours, so failure can only occur if these vertices have been eaten up by the previous embeddings. We show that this is unlikely; in fact, we show that the following stronger *diet condition* for each $t \in V(H)$ is likely to hold:

(3.1) For each $S \subseteq V(H)$ of size at most $2D+3$, we have that $|\mathsf{N}(S)\setminus \mathrm{im}(G[1,\dots,t])| \approx p^{|S|}(n-t)$.

We fix S and aim to show that S is very unlikely to be a set which witnesses the diet condition failing at the first time (since such an S must exist if the diet condition ever fails). In other words, assuming the diet condition holds up to time $t-1$, we want to show that the sum

$$\sum_{i=1}^{t} \mathbb{1}\big(i \text{ is embedded to } \mathsf{N}(S)\big)$$

is likely to be about $p^{|S|}t$. If these Bernoulli random variables were independent, Hoeffding's inequality would tell us that the sum is very likely to be close to its expectation. They are not independent, but nevertheless a martingale version of Hoeffding's inequality shows that the sum is likely to be close to the sum of conditional expectations

$$\sum_{i=1}^{t} \mathbb{E}\big[\mathbb{1}\big(i \text{ is embedded to } \mathsf{N}(S)\big)\big|\mathscr{H}_{i-1}\big] = \sum_{i=1}^{t}\sum_{w\in\mathsf{N}(S)} \mathbb{P}\big(i \text{ is embedded to } w\big|\mathscr{H}_{i-1}\big) \tag{3.2}$$

where $\mathscr{H}_{i-1}$ denotes the history, that is, the choices for embedding vertices $1,\dots,i-1$, and the equality is by linearity of expectation. This sum is itself a random variable, but it turns out to be easier to control. To avoid a technical complication, let us pretend that each i has exactly d left-neighbours. Letting κ be a very small constant, for i in the interval $j+1,\dots,j+\kappa n$ there is a chance to embed i to w each time w is in the candidate set of i; that is, each time that w is adjacent to all the embedded neighbours of i. The following *cover condition* states that this happens about as often as one would expect:

(3.3) For each $j \in V(G)$ and $w \in V(H)$, there are about $p^d\kappa n$ vertices among $[j+1, j+\kappa n] \subseteq V(G)$ which contain w in their candidate set.

For each i in the interval $j+1, \dots, j+\kappa n$ whose candidate set contains $w \notin \mathrm{im}(G[1, \dots, i-1])$, the vertex i is embedded to a set of size about $p^d(n-i+1) \approx p^d(n-j)$ by the diet condition, where the approximation is since κ is very small. So the probability of embedding i to w is about $p^{-d}(n-j)^{-1}$. On the other hand, by the diet condition we have $\big|\mathsf{N}(S) \setminus \mathrm{im}(G[1, \dots, i-1])\big| \approx p^{|S|}(n-j)$, which gives the number of vertices w contributing to the sum (3.2). Summing up, if the cover condition holds then the interval $j+1, \dots, j+\kappa n$ contributes about $p^{|S|}\kappa n$ to the sum (3.2). So the cover condition holding implies that the whole sum (3.2) comes to about $p^{|S|}t$, as desired. This means that, provided the cover condition did not yet fail, the diet condition is unlikely to fail.

We sketch why the cover condition is likely to hold provided the diet condition has not yet failed. When we embed a vertex i, provided the diet condition did not yet fail, the probability of embedding it to a neighbour of w is about p. Now a similar application of a martingale Hoeffding inequality shows that the probability of a given w and j witnessing the failure of the cover condition, given that the diet condition did not yet fail, is very small.

Consider the first time at which one of the cover and diet conditions fails. Before this time both hold, so the probability that t is that first time is by the above argument very small. Taking the union bound over t we conclude that with high probability no such first time exists, and therefore *RandomEmbedding* succeeds, as desired.

3.2. Idea of the proof of Lemma 5.

For lack of space, we will not explain any details of the proof of Lemma 5. We use similar ideas of showing that sums of dependent random variables concentrate using martingale inequalities. The interesting feature is that, because we allow the G'_t to have vertices of very high degree but (because the G'_t are D-degenerate) these must be very few, this time the random variables we are summing (such as the number of edges removed at a vertex of H_{t-1} to form H_t) have maximum values vastly larger than the expected value. In this situation Hoeffding-type inequalities perform very poorly. We use Freedman's inequality [4] to obtain the desired concentration: this inequality performs better when the variance of each random variable is much smaller than its maximum value, and gives us the concentration we need.

References

[1] J. Böttcher, J. Hladký, D. Piguet, and A. Taraz. An approximate version of the tree packing conjecture. *Israel J. Math.*, 211(1):391–446, 2016.

[2] A. Ferber, C. Lee, and F. Mousset. Packing spanning graphs from separable families. arXiv:1512.08701, to appear in Israel J. Math.

[3] A. Ferber and W. Samotij. Packing trees of unbounded degrees in random graphs. arXiv:1607.07342.

[4] D. A. Freedman. On tail probabilities for martingales. *Ann. Probability*, 3:100–118, 1975.

[5] A. Gyárfás and J. Lehel. Packing trees of different order into K_n. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 463–469. North-Holland, Amsterdam, 1978.

[6] F. Joos, J. Kim, D. Kühn, and D. Osthus. Optimal packings of bounded degree trees. arXiv:1606.03953.

[7] J. Kim, D. Kühn, D. Osthus, and M. Tyomkyn. A blow-up lemma for approximate decompositions. arXiv:1604.07282.

[8] S. Messuti, V. Rödl, and M. Schacht. Packing minor-closed families of graphs into complete graphs. *J. Combin. Theory Ser. B*, 119:245–265, 2016.

[9] G. Ringel. Problem 25. In *Theory of Graphs and its Applications (Proc. Int. Symp. Smolenice 1963)*. Czech. Acad. Sci., Prague, 1963.

Department of Mathematics, London School of Economics, Houghton Street, London, WC2A 2AE, UK
E-mail address: {p.d.allen, j.boettcher}@lse.ac.uk

Fachbereich Mathematik, Institut für Geometrie, TU Dresden, 01062 Dresden, Germany.
E-mail address: honzahladky@gmail.com

Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic. With institutional support RVO:67985807.
E-mail address: piguet@cs.cas.cz