PERFECTLY PACKING GRAPHS WITH BOUNDED DEGENERACY AND MANY LEAVES

PETER ALLEN, JULIA BÖTTCHER, DENNIS CLEMENS, AND ANUSCH TARAZ

Abstract. We prove that one can perfectly pack degenerate graphs into complete or dense n-vertex quasirandom graphs, provided that all the degenerate graphs have maximum degree $o\big(\frac{n}{\log n}\big)$, and in addition $\Omega(n)$ of them have at most $\big(1-\Omega(1)\big)n$ vertices and $\Omega(n)$ leaves. This proves Ringel's conjecture and the Gyárfás Tree Packing Conjecture for all but an exponentially small fraction of trees (or sequences of trees, respectively).

1. Introduction

Let $\mathcal{G} = \{G_1, G_2, \dots, G_s\}$ be a collection of graphs, and H be a graph. We say that $\mathcal{G}$ *packs into* H if we can find pairwise edge disjoint copies in H of the graphs $G_1, \dots, G_s$. If in addition we have $\sum_{i\in[s]} e(G_i) = e(H)$, we call the packing *perfect*: in this case, each edge of H is used in a copy of exactly one G_i.

The study of perfect packings in graphs has a long history, beginning with Plücker [24], who in 1835 showed that for certain values of n there is a perfect packing of copies of K_3 into K_n. Steiner [28] in 1853 asked, more generally, when one can perfectly pack the n-vertex k-uniform complete hypergraph with cliques on r vertices. He phrased the question as a problem in set theory, and gave some obvious divisibility-based necessary conditions on n; today such perfect packings are called *combinatorial designs*. In 1846 Kirkman [18] asked for a strengthening of Plücker's ideas: when can one have a perfect packing of spanning K_3-factors (that is, $\frac{n}{3}$ vertex disjoint copies of K_3) into K_n? Again, he showed that for specific values of n such a thing is possible. Generalising this in the direction of Steiner one obtains the concept of a *resolvable design*; again, it is easy to find divisibility-based necessary conditions on n.

Despite their simple statement, these problems turned out to be difficult. Kirkman gave explicit constructions showing that one can perfectly pack K_n with copies of K_3 if and only if n is congruent to 1 or 3 modulo 6. But it took more than a century until, in 1975 Wilson [29] proved the (much harder) statement that the necessary divisibility conditions are also sufficient for cliques of any fixed size in large enough (2-uniform) graphs. Ray-Chaudhuri and Wilson [25] in 1971 solved Kirkman's problem. There was then another pause, till 2014—up to which time, despite significant work, not a single example of a non-trivial hypergraph perfect packing for uniformity at least 6 was discovered—when Keevash [14], in a major breakthrough, proved that the necessary divisibility conditions are also sufficient for any fixed clique size and hypergraph uniformity, provided n is large enough. The problem was re-solved, by a rather different method, by Glock, Kühn, Lo and Osthus [9], who also solved the problem of perfect packings with general fixed hypergraphs in place of cliques. In [15], Keevash made the

Date: May 4, 2021.
PA was partially supported by the EPSRC, grant number EP/P032125/1.
JB was partially supported by the EPSRC, grant number EP/R00532X/1.
PA and JB were partially supported by a STICERD grant.

beautiful observation that a resolvable design is equivalent to a perfect packing in an auxiliary well-structured hypergraph, and established the existence of such a packing. Hence, he proved that resolvable designs exist whenever the obvious necessary conditions are satisfied and n is large enough.

When one moves away from packings with fixed-size objects (or statements which can be reduced to such packings), the first positive result is due to Walecki in the 1800s (see [20]), who proved that K_n can be perfectly packed with Hamilton cycles whenever n is odd. In the 1960s and 70s, interest in this area was renewed, in particular due to conjectures of Ringel [26] and Gyárfás [10] on packings of trees. These conjectures state, respectively, that K_{2n-1} can be packed with $2n-1$ copies of any given n-vertex tree, and that if $T_2, \dots, T_n$ is any sequence of trees such that $v(T_i) = i$, then $\{T_2, \dots, T_n\}$ packs into K_n. In both cases, the packing is necessarily perfect, which makes these conjectures difficult. It is not too hard to prove either conjecture for stars or paths, and a considerable amount of effort was put into solving special cases of both cases (for the former, see the survey of Gallian [8]). However until rather recently, there were no proofs of either conjecture for any reasonably large family of trees. Then Joos, Kim, Kühn and Osthus [13] proved (among other things) that both conjectures hold when the trees have constant maximum degree Δ and n is large enough. The proof of this result is very hard, using a variety of powerful techniques from modern extremal graph theory.

Broadly, the recent solutions to perfect packing conjectures (both, in the case of combinatorial designs and in the case of tree packing) depend on two advances: *randomised packing methods*, and the *absorbing method*. The idea is that, rather than deterministically specifying how to pack, one gives a randomised packing algorithm and argues that it is likely to succeed. Here 'succeed' means packing almost all (not all) of the graphs, and there will be some edges remaining. This leftover is dealt with by the absorption method: one should begin by cleverly choosing an 'absorbing packing' of the first few graphs which has the property that whatever the remaining edges from the randomised algorithm turn out to be, one can modify the absorbing packing in order to incorporate the leftover to a perfect packing. In the work of Keevash, roughly this template is followed (though there are some mild conditions on the leftover), and an intricate algebraic structure is used to obtain the absorbing packing. In the work of Joos et al., the *iterative absorption* method (originating in [19]) is used: here one packs in a way that uses all the edges adjacent to most vertices and almost no edges among the remaining few vertices, and then iterates this process, until all the difficulty has been pushed into a set of vertices so tiny that a relatively simple absorber suffices.

The idea of randomised packing dates back to Rödl's celebrated *nibble method* [27] in which he solved the Erdős-Hanani problem, of showing that if n is large enough then one can pack most of the edges of the complete k-uniform n-vertex hypergraph with cliques of size r, solving Steiner's problem approximately. Note that for this problem there is no divisibility restriction on n. The nibble method was brought to tree packing by Böttcher, Hladký, Piguet and Taraz [2], who showed that one can pack most of the edges of K_n with bounded degree trees, provided the trees are not too close to spanning. This was the trigger for a sequence of generalisations: Messuti, Rödl and Schacht [21] showed that one can replace trees with graphs from any non-trivial minor-closed family; Ferber, Lee and Mousset [5] showed that one can additionally allow spanning graphs; Kim, Kühn, Osthus and Tyomkyn [17] discarded the structural assumption entirely, packing most of the edges of K_n with arbitrary bounded degree graphs. All these results work in more generality than just for packings in K_n. In particular, we should note that the results of [17] work in the Szemerédi regularity setting, which was necessary for the proof strategy of [13].

All the results mentioned so far deal with bounded degree graphs; the first result to handle growing degrees is due to Ferber and Samotij [6], who showed that one can pack most of the edges of K_n with trees of maximum degree $O\big(\frac{n}{\log n}\big)$ (for almost-spanning trees) or $O\big(\frac{n}{\log n}\big)^{1/6}$ (for spanning trees). In [1] it was proved that one can pack most of the edges of K_n with arbitrary D-degenerate graphs with maximum degree $O\big(\frac{n}{\log n}\big)$. Finally, recently Montgomery, Pokrovskiy and Sudakov [23] were able to deal with trees of unbounded degree, at least in the setting of Ringel's conjecture: they proved an approximate version of Ringel's conjecture, proving that K_{2n-1} can be packed with $2n-1$ copies of any tree T with $n-o(n)$ vertices.[1]

One might be tempted to think that, while a randomised strategy is very good for packing most of the edges, one cannot hope for a perfect packing: After all, at some point one has to pack the last few graphs, or at least somehow use the last few edges; at this point the packing is very constrained and any mistake will cause the packing to fail (and there cannot be many choices left, so that one cannot hope for strong concentration bounds), but a randomised algorithm will probably make a mistake (at least, unless it does a good deal of 'looking ahead' which will be hard to analyse). But this is not the case, as we show in this paper. A rather natural, simple randomised algorithm can succeed in giving a perfect packing. Using this algorithm, we prove the following.

1.1. Main result.

Given a graph G, an ordering of the vertices is a *D-degeneracy order* if each vertex has at most D neighbours preceding it in the order. We say G is *D-degenerate* if it has a D-degeneracy order. Trees, for example, are 1-degenerate.

Definition 1 ((μ, n)-sequence)**.** *We say that a sequence $(G_i)_{i\in[m]}$ of graphs is a D-*degenerate *(μ, n)-*graph sequence with maximum degree *Δ if*

- $(G1)$ *G_i is D-degenerate and $\Delta(G_i) \leq \Delta$ for each $i \in [m]$,*
- $(G2)$ *$v(G_i) = n$ for each $1 \leq i \leq m - \lfloor \mu n \rfloor$, and*
- $(G3)$ *$v(G_i) = n - \lfloor \mu n \rfloor$ and G_i has at least $\lfloor \mu n \rfloor$ leaves for each i with $m - \lfloor \mu n \rfloor < i \leq m$.*

We also call the G_i with $m - \lfloor \mu n \rfloor < i \leq m$ the special *graphs of the sequence.*

An n-vertex graph H is *(α, k)-quasirandom* with density p if $e(H) = p\binom{n}{2}$ and for every $\ell \in [k]$ and every set $\{v_1, \ldots, v_\ell\}$ of vertices of H we have

$$\big|N_H(v_1, \ldots, v_\ell)\big| = (1 \pm \alpha)p^\ell n\,.$$

Our main result states that a D-degenerate (μ, n)-sequence $(G_i)_{i\in[t]}$ of guest graphs with maximum degree of order at most $\frac{n}{\log n}$ can be perfectly packed into a sufficiently quasirandom graph $\widehat{H}$.

Theorem 2 (main result)**.** *For every D and $\mu, \hat{p}_0 > 0$ there are n_0 and $\xi, c > 0$ such that for every $\hat{p} \geq \hat{p}_0$, every $n \geq n_0$, and every m, the following holds for every n-vertex graph $\widehat{H}$ which is $(\xi, 4D+7)$-quasirandom with density $\hat{p}$. Every D-degenerate (μ, n)-graph sequence $(G_i)_{i\in[m]}$ with maximum degree $\frac{cn}{\log n}$ such that $\sum_{i\in m} e(G_i) \leq e(\widehat{H})$ packs into $\widehat{H}$.*

It is easy to see (and proved for completeness in Proposition 8) that if T is a uniform random labelled n-vertex tree, then for each $c > 0$, with probability $1 - e^{-\Omega(n)}$ the tree T will have

[1]Very recently, after we completed this paper, Montgomery, Pokrovskiy and Sudakov [22] announced a proof of Ringel's conjecture for large trees. This result also follows from even more recent work of Keevash and Staden [16].

at least $n/100$ leaves and maximum degree at most $\frac{cn}{\log n}$. In particular, we have the following corollary to Theorem 2, proving almost all cases of Ringel's conjecture and the Gyárfás Tree Packing Conjecture.

Corollary 3. *Let T be a uniform random n-vertex tree. With probability $1-e^{-\Omega(n)}$, there is a packing of $2n-1$ copies of T into K_{2n-1}.*

Let $T_2,\dots,T_n$ be chosen independently and uniformly at random such that T_i is an i-vertex tree for each $2\le i\le n$. With probability at least $1-e^{-\Omega(n)}$, there is a packing of $\{T_2,\dots,T_n\}$ into K_n. □

We should briefly compare these results to the earlier result of Joos, Kim, Kühn and Osthus [13]. On the one hand, we cannot handle trees with few leaves, so our result does not contain theirs. Furthermore, as far as Corollary 3 goes, packing bounded degree trees 'almost' covers a typical uniform random tree, whose maximum degree is likely to be $\Theta\big(\frac{\log n}{\log\log n}\big)$, and perhaps the approach of [13] could be pushed to allow for a few vertices of logarithmic degree. On the other hand, the method of [13] heavily relies on the structure of trees, and in particular that one can embed them effectively in a Szemerédi partition; handling general degenerate graphs with high maximum degree would be rather challenging with their approach.

Finally, we discuss which conditions in Theorem 2 are needed. It is easy to see that a typical graph H on n vertices with density $\frac{1}{2}$ will be quasirandom. However such a graph will typically not contain any $\frac{1}{10}\log n$-set S of vertices such that each other vertex has a neighbour in S. In particular, if G is an n-vertex graph which is the vertex disjoint union of $\frac{1}{10}\log n$ stars, each with the same number of leaves (up to an error 1), then G has maximum degree less than $\frac{20n}{\log n}$ but does not embed into H. Thus the maximum degree bound in our theorem is optimal up to a constant factor.

One can allow D to grow with n. Examination of our proof shows it can grow roughly as $\log\log\log n$, but this is presumably not optimal. On the other hand, D cannot be as big as $10\log n$, since a typical random graph is unlikely to contain any given graph with $9n\log n$ edges.

We cannot allow all graphs to be spanning in Theorem 2, as an example in [2, Section 9.1] shows. However we expect one can do better than needing linearly many graphs to be linearly far from spanning.

We do not believe that it is necessary to have many graphs with many leaves. We should note that one cannot simply omit this condition, because for example no collection of cycles can perfectly pack K_{2n}, due to a parity obstruction: cycles use an even number of edges at each vertex, but K_{2n} has odd degree vertices. However for the case $D=1$ (i.e. forests) we believe one can omit the condition entirely (as the leaves should allow for parity correction). Work on this problem is work in progress with Hladký and Piguet.

1.2. **Proof outline.**

This paper builds on [1], so we begin by outlining the randomised algorithm *PackingProcess* described there (and repeated here later). In *PackingProcess*, we embed graphs one-by-one into $\widehat{H}$. To embed a given graph, we take the vertices in the degeneracy order, and one by one embed them: at each step, we choose uniformly at random from the set of all vertices to embed to which do not immediately break our packing (either by re-using a vertex already used in the current embedding or by re-using an edge already used for a previous graph). We do this until almost the entire graph is embedded; then we choose arbitrarily a way of completing the embedding to a spanning embedding; we argue that such a completion exists with high probability. (This is a slight simplification, but the simplification does not affect the point.) Note that here we certainly do not 'look ahead' in any way at what we will embed in the future, and the algorithm is essentially purely random. This gives the following packing result from [1].

Theorem 4. *For every $\gamma > 0$ and each $D \in \mathbb{N}$ there exist $c, \xi > 0$ and n_0 such that the following holds for each integer $n \geq n_0$. Suppose that $\left(G_t\right)_{t\in[t^*]}$ 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 that H is a $(\xi, 4D+7)$-quasirandom n-vertex graph. Suppose further that the total number of edges of $\left(G_t\right)_{t\in[t^*]}$ is at most $e(H) - \gamma n^2$. Then $\left(G_t\right)_{t\in[t^*]}$ packs into H.*

To prove our main theorem, we will begin by removing some of the leaves from each special graph in a given (μ, n)-graph sequence. We will then use *PackingProcess* to pack all the non-special graphs and all the special graphs minus the removed leaves into $\widehat{H}$. It remains to embed these removed leaves into the graph H consisting of the unused edges of $\widehat{H}$. We say a removed leaf is *dangling* at a vertex $v \in V(H)$ if its parent is embedded to v. We will show that at each vertex of H, it is likely that there are about twice as many edges as dangling leaves. In order to decide where to embed the dangling leaves, we first orient the edges of H randomly, then 'correct' this orientation (by reversing a few carefully chosen directed paths of length 2) such that the out-degree of each vertex is equal to the number of dangling leaves. This is the only step in our algorithm where we 'look ahead' and prepare for the future.

We then complete the packing by going through H vertex-by-vertex, and for each vertex choosing a uniform random assignment of the dangling leaves to the out-neighbours which preserves having a packing. We should note that this last step has some similarity to the approach of [13], where the authors also complete their perfect packing by assigning dangling leaves to out-neighbours, but in a small set of vertices. However in their setting, they only need to assign one dangling leaf per tree, and no other vertices of that tree are embedded to the small vertex set. As they already did all the hard work to reach this point, it is not hard for them to make such an assignment. In our setting, we need to embed linearly many dangling leaves per tree, which dangle on many different vertices, and the previously embedded images of these trees can cover most of the vertices to which we want to embed dangling leaves. It is already non-trivial that we can even assign the dangling leaves at the first vertex of H, and this assignment affects what we can do at later vertices.

In order to understand how it can be that this random process succeeds in obtaining a perfect packing, one should note that when we embed the dangling leaves at the first vertex of H, we have no choice over the set of edges we use (these are fixed as the out-neighbours) but the set of assignments, from which we choose uniformly, is rather large. This property is preserved right through to the last vertex of H—even in the last step, we have not one but many possible assignments to choose from, so that even in the last steps we have quite a lot of randomness.

A remark concerning convention in our proofs is in place. We will rather often want to talk about the guest graph to which a given vertex belongs. For this to make sense, we will throughout consider the guest graphs G_s to be pairwise vertex disjoint, even though in our proof each of them is defined on the vertex set $[n]$. This allows us to write statements referring for example to the guest graph G_s containing the given vertex x, but also when G_s is clear from the context to talk about a vertex x such that $i < x < j$ for given numbers i and j.

1.3. **Organisation.** In Section 2 we fix notation and collect some concentration inequalities and facts about degenerate graphs. In Section 3 we state our main technical theorem, show that it implies Theorem 2, and formalise our random packing process. In Section 4 we fix the constants we will use throughout our proofs. In Section 5 we provide our main lemmas which analyse what happens in the different phases of our packing process: the almost perfect packing lemma, the orientation lemma, and the matching lemma. In Section 6 we show that these lemmas imply our main technical theorem. Section 7 proves the orientation lemma, Section B the matching lemma, and Section 8 the almost perfect packing lemma. The latter takes up the (technical) bulk of the paper. Concluding remarks are given in Section 9.

2. Preliminaries

2.1. **Notation.** For a graph G we write $V(G)$ for the vertices of G, and $E(G)$ for its edges, $v(G)$ for the number of vertices in G, and $e(G)$ for the number of edges. For disjoint vertex sets $X, Y \subseteq V(G)$ we write $G[X]$ for the subgraph of G induced by X, and $G[X,Y]$ for the bipartite subgraph of G on vertex set $X \cup Y$ and with all edges of G with one end in X and the other in Y. For a set S of vertices of G, we write $N_G(S)$ for the *common neighbourhood* $\{u \in V(G) : su \in E(G) \text{ for each } s \in S\}$. We write $\deg_G(S) := \big|N_G(S)\big|$ for the *common degree* of S in G. When $S = \{v_1, \dots, v_\ell\}$ we will omit the set braces and simply write $N_G(v_1, \dots, v_\ell)$ and $\deg_G(v_1, \dots, v_\ell)$. We will not use joint neighbourhoods of sets of vertices in this paper.

Given a graph G and a set of vertices X, if $X \subseteq V(G)$ we write $G - X$ for the graph obtained by removing the vertices X from $V(G)$, i.e. $G\big[V(G) \setminus X\big]$. If X is disjoint from $V(G)$, we write $G + X$ for the graph obtained by adding X as a set of isolated vertices, i.e. the graph on vertex set $V(G) \cup X$ whose edge set is $E(G)$. Given graphs G_1 and G_2 with $V(G_2) \subseteq V(G_1)$, we write $G_1 - G_2$ for the graph obtained by removing the edges of G_2 from G_1, i.e. the graph on vertex set $V(G_1)$ whose edge set is $E(G_1) \setminus E(G_2)$.

Given an ordering $V(G) = \{v_1, \dots, v_n\}$ of the vertices of a graph G, we write $N^-_G(v_i)$ for the *left-neighbourhood of* v_i, i.e. the set

$$N^-_G(v_i) := N_G(v_i) \cap \{v_k : \; k \in [i-1]\} .$$

We write $\deg^-_G(v_i) := \big|N^-_G(v_i)\big|$ for the *left-degree of* v_i. Thus the order is a D-degeneracy order if for each $i \in [n]$ we have $\deg^-_G(v_i) \le D$.

An *orientation* of a graph $H = (V, E)$ is an oriented graph on V which contains, for each undirected edge $uv \in E$, exactly one directed edge, either $\vec{uv}$ or $\vec{vu}$. The *outdegree* $\deg^+_{\vec{H}}(v)$ of a vertex v in an oriented graph $\vec{H}$ is the number of vertices u in $\vec{H}$ such that $\vec{vu}$ is an edge of $\vec{H}$; the set of these vertices u is the *outneighbourhood* $N^+_{\vec{H}}(v)$ of v.

Let Ω be a finite probability space. A *filtration* $\mathcal{F}_0$, $\mathcal{F}_1$,…, $\mathcal{F}_n$ is a sequence of partitions of Ω such that $\mathcal{F}_i$ refines $\mathcal{F}_{i-1}$ for all $i \in [n]$. In our application, the partition $\mathcal{F}_i$ is given by all possible histories of the run of one of our algorithms up to time i. (For more explanation see [1].) We say that a function $f : \Omega \to \mathbb{R}$ is $\mathcal{F}_i$*-measurable* if f is constant on each part of $\mathcal{F}_i$. Further, for any random variable $Y \colon \Omega \to \mathbb{R}$ the *conditional expectation* $\mathbb{E}(Y|\mathcal{F}_i) \colon \Omega \to \mathbb{R}$ and the *conditional variance* $\mathrm{Var}(Y|\mathcal{F}_i) \colon \Omega \to \mathbb{R}$ of Y with respect to $\mathcal{F}_i$ are defined by

$$\begin{aligned} \mathbb{E}(Y|\mathcal{F})(x) &= \mathbb{E}(Y|X), \\ \mathrm{Var}(Y|\mathcal{F})(x) &= \mathrm{Var}(Y|X), \end{aligned} \qquad \text{where } X \in \mathcal{F} \text{ is such that } x \in X .$$

Suppose that we have an algorithm which proceeds in m rounds using a new source of randomness Ω_i in each round i. Then the probability space underlying the run of the algorithm is $\prod_{i=1}^m \Omega_i$. By *history up to time* t we mean a set of the form $\{\omega_1\} \times \dots \times \{\omega_t\} \times \Omega_{t+1} \times \cdots \Omega_m$, where $\omega_i \in \Omega_i$. We shall use the symbol $\mathscr{H}_t$ to denote any particular history of such a form.

By a *history ensemble up to time t* we mean any union of histories up to time t; we shall use the symbol $\mathscr{L}$ to denote any one such. Observe that there are natural filtrations associated to such a probability space: given times $t_1 < t_2 < \dots$ we let $\mathcal{F}_{t_i}$ denote the partition of Ω into the histories up to time t_i.

2.2. Probabilistic tools.

Theorem 5 (Chernoff bounds, [11, Theorem 2.10])**.** *Suppose X is a random variable which is the sum of a collection of independent Bernoulli random variables. Then we have for $\delta \in (0, 3/2)$*

$$\mathbb{P}\big[X > (1+\delta)\mathbb{E}X\big] < e^{-\delta^2\mathbb{E}X/3} \quad \textit{and} \quad \mathbb{P}\big[X < (1-\delta)\mathbb{E}X\big] < e^{-\delta^2\mathbb{E}X/3}\,.$$

We use the following consequence of Freedman's inequality [7], derived in [1], for analysing our random embedding algorithms.

Lemma 6 (Freedman's inequality on a good event)**.** *Let Ω be a finite probability space, and $(\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_n)$ be a filtration. Suppose that we have $R > 0$, and for each $1 \le i \le n$ we have an $\mathcal{F}_i$-measurable non-negative random variable Y_i, nonnegative real numbers $\tilde{\mu}$, $\tilde{\nu}$ and $\tilde{\sigma}$, and an event $\mathcal{E}$. Suppose that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big] = \tilde{\mu} \pm \tilde{\nu}$, and $\sum_{i=1}^n \mathrm{Var}\,\big[Y_i\big|\mathcal{F}_{i-1}\big] \le \tilde{\sigma}^2$, and $0 \le Y_i \le R$ for each $1 \le i \le n$. Then for each $\tilde{\varrho} > 0$ we have*

$$\mathbb{P}\left[\mathcal{E} \textit{ and } \sum_{i=1}^n Y_i \neq \tilde{\mu} \pm (\tilde{\nu} + \tilde{\varrho})\right] \le 2\exp\Big(-\frac{\tilde{\varrho}^2}{2\tilde{\sigma}^2 + 2R\tilde{\varrho}}\Big)\,.$$

Furthermore, if we assume only that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big] \le \tilde{\mu} + \tilde{\nu}$, and $\sum_{i=1}^n \mathrm{Var}\,\big[Y_i\big|\mathcal{F}_{i-1}\big] \le \tilde{\sigma}^2$, and $0 \le Y_i \le R$ for each $1 \le i \le n$, then for each $\tilde{\varrho} > 0$ we have

$$\mathbb{P}\left[\mathcal{E} \textit{ and } \sum_{i=1}^n Y_i > \tilde{\mu} + \tilde{\nu} + \tilde{\varrho}\right] \le \exp\Big(-\frac{\tilde{\varrho}^2}{2\tilde{\sigma}^2 + 2R\tilde{\varrho}}\Big)\,.$$

We should stress that here quantities such as $\mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big]$ are random variables on Ω; when we say that a random variable satisfies a given statement involving ranges or inequalities, we mean this statement is true pointwise for each $\omega \in \Omega$. Thus 'either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big] = \tilde{\mu} \pm \tilde{\nu}$' means that for each $\omega \in \Omega$, either $\omega \notin \mathcal{E}$ or the sum of real numbers $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big](\omega)$ is in the range $\tilde{\mu} \pm \tilde{\nu}$.

A special case is the following corollary.

Corollary 7. *Let Ω be a finite probability space, and $(\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_n)$ be a filtration. Suppose that we have $R > 0$, and for each $1 \le i \le n$ we have an $\mathcal{F}_i$-measurable non-negative random variable Y_i, nonnegative real numbers $\tilde{\mu}, \tilde{\nu}$ and an event $\mathcal{E}$.*

(*a*) *Suppose that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big] \le \tilde{\mu}$, and $0 \le Y_i \le R$ for each $1 \le i \le n$. Then*

$$\mathbb{P}\left[\mathcal{E} \textit{ and } \sum_{i=1}^n Y_i > 2\tilde{\mu}\right] \le \exp\Big(-\frac{\tilde{\mu}}{4R}\Big)\,.$$

(*b*) *Suppose that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^n \mathbb{E}\big[Y_i\big|\mathcal{F}_{i-1}\big] = \tilde{\mu} \pm \tilde{\nu}$, and $0 \le Y_i \le R$ for each $1 \le i \le n$. Then for each $\tilde{\varrho} > 0$ we have*

$$\mathbb{P}\left[\mathcal{E} \textit{ and } \sum_{i=1}^n Y_i \neq \tilde{\mu} \pm (\tilde{\nu} + \tilde{\varrho})\right] \le 2\exp\Big(-\frac{\tilde{\varrho}^2}{2R(\tilde{\mu} + \tilde{\nu} + \tilde{\varrho})}\Big)\,.$$

In particular, if $\tilde{\nu} = \tilde{\varrho} = \tilde{\mu}\tilde{\eta} > 0$ and $\tilde{\eta} \leq \frac{1}{2}$, then

$$\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i \neq \tilde{\mu}(1 \pm 2\tilde{\eta})\right] \leq 2\exp\Big(-\frac{\tilde{\mu}\tilde{\eta}^2}{4R}\Big).$$

Proof. Both parts follow from Lemma 6 with $\tilde{\sigma}^2 = R(\tilde{\mu} + \tilde{\nu})$; for the first part we also set $\tilde{\nu} = 0$ and $\tilde{\varrho} = \tilde{\mu}$. Observe that

$$\operatorname{Var}\big[Y_i \big| \mathcal{F}_{i-1}\big] \leq \mathbb{E}\big[Y_i^2 \big| \mathcal{F}_{i-1}\big] \leq R \cdot \mathbb{E}\big[Y_i \big| \mathcal{F}_{i-1}\big],$$

so that

$$\sum_{i=1}^{n} \operatorname{Var}\big[Y_i \big| \mathcal{F}_{i-1}\big] \leq R \sum_{i=1}^{n} \mathbb{E}\big[Y_i \big| \mathcal{F}_{i-1}\big] \leq R(\tilde{\mu} + \tilde{\nu})$$

when $\mathcal{E}$ holds, justifying the choice of $\tilde{\sigma}^2$. □

We conclude this subsection by giving maximum degree and leaf statistics for random labelled trees, whose proof we leave to Appendix A.

Proposition 8. *Let T_n be a tree chosen uniformly at random from the set of n-vertex labelled trees. Then*

(i) With probability at most $\exp\big(-\frac{n}{500}\big)$ the number of leaves in T_n is less than $\frac{n}{100}$.

(ii) Given $c > 0$, if n is sufficiently large then with probability at most $e^{-cn/2}$ there is a vertex in T_n with degree greater than $\frac{cn}{\log n}$.

We should point out that much more precise statistics are known; we give these rough and simple bounds for completeness.

2.3. **Degenerate graphs.** It is easy to show that degenerate graphs contain large independent sets all of whose vertices have the same degree, and moving such a set to the end of the degeneracy order, we obtain the following lemma. The (standard) proof is provided in Appendix A.

Lemma 9. *Let G be a D-degenerate n-vertex graph. Then there exists an integer $0 \leq d \leq 2D$ and a $2D$-degeneracy order of $V(G)$ such that the last $\lceil (2D+1)^{-3} n \rceil$ vertices in this order form an independent set and all have degree d.*

Further, we shall use the following auxiliary lemma, which given an arbitrary family of graphs we want to pack produces a family with at most $\frac{3}{2}n$ members and the same bound on maximum degree and degeneracy by combining graphs with many isolated vertices or leaves. Obtaining such a family with at most $\frac{3}{2}n$ members needs some argument; while obtaining a family with at most $2n$ members instead is straightforward. In [1] we only used the latter, and the reason why we use the smaller family here is that it allows us to stay consistent with the constants used in [1]. More precisely, the constant α_{2n} that will be defined in (4) is not small enough for our analysis here, while $\alpha_{7n/4}$ is small enough.

Lemma 10 (compression lemma)**.** *Let $(G_i)_{i\in[m]}$ be a family of D-degenerate graphs with maximum degree at most Δ, with $\sum_{i=1}^{m} e(G_i) \leq \binom{n}{2}$ and $v(G_i) \leq n$ for all $i \in [m]$. Then there is a family of graphs $(\check{G}_i)_{i\in[\check{m}]}$ with $\check{m} \leq \frac{3}{2}n$, such that for each $i \in [\check{m}]$ we have $v(\check{G}_i) \leq n$, $\Delta(\check{G}_i) \leq \max\{2, \Delta\}$, and $\check{G}_i$ is $\max\{2, D\}$-degenerate, and such that $(\check{G}_i)$ is a packing of (G_i).*

The proof of this lemma is given in Appendix A.

3. Main technical theorem and the packing algorithm

In this section we detail our packing algorithm, introduce the definitions necessary for this algorithm, and outline the proof of why this algorithm succeeds. We deduce Theorem 2 from the following technical result.

Theorem 11. *For every D and $\mu, \hat{p}_0 > 0$ with $\mu \leq \frac{1}{4}$ there are n_0 and $\xi, c > 0$ such that for every $\hat{p} \geq \hat{p}_0$ and every $n \geq n_0$ the following holds. Suppose that $\widehat{H}$ is a $(\xi, 2D+3)$-quasirandom graph with n vertices and density $\hat{p}$. Suppose that $s^* \leq \frac{7}{4}n$ and that the graph sequence $(G_s)_{s\in[s^*]}$ is a D-degenerate (μ, n)-graph sequence with maximum degree at most $\Delta = \frac{cn}{\log n}$, such that for each $s \in [s^*]$ there is a D-degeneracy order of G_s such that the last $\lceil (D+1)^{-3}n \rceil$ vertices form an independent set in G_s, and all have the same degree d_s in G_s. Suppose further that $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$. Then $(G_s)_{s\in[s^*]}$ packs into $\widehat{H}$.*

Before sketching the proof of Theorem 11, we show that it implies Theorem 2.

Proof of Theorem 2. Given D, μ, $\hat{p}_0$, let n_0', ξ, c be as given by Theorem 11 for input $D' = 2\max\{2, D\}$, $\mu' = \min\{\mu, \frac{1}{4}\}$, and $\hat{p}_0$. Choose $n_0 = \max\{n_0', 10c^{-2}\}$. Next, let $\hat{p}$ and n as well as the graphs $\widehat{H}$ and $(G_i)_{i\in[m]}$ be given.

Now we first add new graphs G_i with $i > m$ consisting of single edges to our graph sequence until $\sum e(G_i) = e(\widehat{H})$. Assume that the resulting sequence has m' graphs and reorder the sequence so that the $\lfloor \mu n \rfloor$ special graphs come last. In a second step we apply the compression lemma, Lemma 10, to the non-special graphs $(G_i)_{i\in[m'-\lfloor\mu' n\rfloor]}$ to obtain a family $(\check{G}_i)_{i\in[\check{m}]}$ with $\check{m} \leq \frac{3}{2}n$ that is a packing of $(G_i)_{i\in[m'-\lfloor\mu' n\rfloor]}$. In a third step, we add the remaining special graphs to this compressed family, that is, for $1 \leq i \leq \lfloor \mu' n \rfloor$ we let $\check{G}_{\check{m}+i} = G_{m'-\lfloor\mu' n\rfloor+i}$. We obtain a family $(\check{G}_s)_{s\in[s^*]}$ of $\max\{2, D\}$-degenerate graphs with maximum degree at most $cn/\log n$, where $s^* \leq \frac{3}{2}n + \lfloor \mu' n \rfloor \leq \frac{7}{4}n$. In a fourth step, we apply Lemma 9 to obtain a D'-degeneracy order of each $\check{G}_s$ such that the last $\lceil (D+1)^{-3}n \rceil \geq \lceil (D'+1)^{-3}n \rceil$ vertices form an independent set in $\check{G}_s$, and all have the same degree d_s in $\check{G}_s$. Hence the family $(\check{G}_s)_{s\in[s^*]}$ satisfies all conditions required by Theorem 11 with the above chosen constants. Since $\widehat{H}$ is $(\xi, 4D+7)$-quasirandom, it is also $(\xi, 2D'+3)$-quasirandom as required for Theorem 11. Applying this theorem, we obtain a perfect packing of $(\check{G})_{s\in[s^*]}$ into $\widehat{H}$, which gives a packing of $(G_i)_{i\in[m]}$ since $(\check{G})_{s\in[s^*]}$ is a packing of $(G_i)_{i\in[m]}$ (plus possibly some additional edges). $\square$

We now sketch the proof of Theorem 11. We start by creating an almost perfect packing, which omits linearly many leaves in the linearly many special graphs G_s, by packing only the following subgraph sequence omitting $\ell = \lfloor \nu n \rfloor$ leaves.

Definition 12 (corresponding subgraph sequence)**.** *For a D-degenerate (μ, n)-graph sequence $(G_s)_{s\in[s^*]}$ with maximum degree Δ, we say that $(G_s')_{s\in[s^*]}$ is a* corresponding subgraph sequence omitting ℓ leaves *if*

- ($G1'$) *for each $s \leq s^* - \lfloor \mu n \rfloor$ we have $G_s' = G_s$, and*
- ($G2'$) *for each $s > s^* - \lfloor \mu n \rfloor$ we have $G_s' = G_s - V_s + I_s$ for an independent set V_s of leaves in G_s with $|V_s| = \ell$, and a set I_s of new and independent vertices with $|I_s| = \ell$.*

We remark that the addition of the independent set I_s in ($G2'$) is purely for technical reasons: it guarantees that the special G_s' have $n - \lfloor \mu n \rfloor$ vertices, which makes the statement of some of our later lemmas easier (in particular Lemma 18). The restriction that the set V_s is independent simply says that if there is a component of G_s which contains exactly one edge, both endpoints are leaves but only one may be in V_s.

The subgraph sequence (G'_s) is packed with the help of the following *PackingProcess*, which uses an algorithm *RandomEmbedding* that we shall describe thereafter. This *PackingProcess* was introduced and analysed in [1], and it requires n-vertex graphs with a degeneracy ordering whose last vertices form an independent set as input. To that end, for each $s \in [s^*]$ with $s > s^* - \lfloor \mu n \rfloor$, we do the following. We let I'_s be a set of $n - v(G'_s)$ new isolated vertices and we obtain G''_s by adding I'_s to G'_s. Each non-special graph G'_s, with $s \le s^* - \lfloor \mu n \rfloor$, already has n vertices and so we simply set $G''_s = G'_s$. For the special graphs G'_s, with $s > s^* - \lfloor \mu n \rfloor$, we fix a D-degeneracy order of G''_s such that the $\lfloor \mu n \rfloor > \delta n$ isolated vertices in I'_s come last. We then relabel vertices, so that again $V(G''_s) = [n]$ and the fixed D-degeneracy order is the natural order on $[n]$.

Briefly, the idea of the following algorithm is that we split the host graph $\widehat{H}$ randomly into a large part H_0 and a small part H^*_0, and then we pack the graphs G''_s one by one. We use the large part to pack all but the last δn vertices of each G''_s, and the small part to complete the embedding of each G''_s. Note that for the special graphs, the 'completion' will consist of simply embedding δn isolated vertices, which is trivial. For the non-special graphs this completion will use some edges of H^*_0 and it is not automatic that such a completion need exist, but [1] proves it is very likely. It turns out we do not need any special properties of the completion (because δ is so tiny that the few edges affected will not destroy quasirandomness properties) and so we simply choose one arbitrarily. Note that having removed linearly many leaves from linearly many graphs, the total number of edges of the G''_s is smaller by $\Theta(n^2)$ than the number of edges of H, so (as stated in Theorem 4) the analysis given in [1] proves that the packing of the G''_s succeeds. We however need to know a good deal more than simply that the packing succeeds; this is a substantial part of the work of this paper.

Algorithm 1: *PackingProcess*

Input : graphs $G''_1, \dots, G''_{s^*}$, with G''_s on vertex set $[n]$ such that the last δn vertices of G''_s form an independent set; a graph $\widehat{H}$ on n vertices

Output: a packing $(\phi^*_s)_{s\in[s^*]}$ of $(G''_s)_{s\in[s^*]}$ into $\widehat{H}$ and a left-over graph H

choose H^*_0 by picking edges of $\widehat{H}$ independently with probability $\gamma\binom{n}{2}/e(\widehat{H})$;
let $H_0 = \widehat{H} - H^*_0$;
for $s = 1$ **to** s^* **do**
    run *RandomEmbedding*(G''_s, H_{s-1}) to get an embedding ϕ''_s of $G''_s[[n-\delta n]]$ into H_{s-1};
    let H_s be the graph obtained from H_{s-1} by removing the edges of $\phi''_s\big(G''_s[[n-\delta n]]\big)$;
    choose an arbitrary extension ϕ^*_s of ϕ''_s embedding all of G''_s and embedding the edges of $G''_s - G''_s[[n-\delta n]]$ into H^*_{s-1} if one exists, otherwise halt with failure;
    let H^*_s be the graph obtained from H^*_{s-1} by removing the edges of $\phi^*_s\big(G''_s - G''_s[[n-\delta n]]\big)$;
end
return $(\phi^*_s)_{s\in[s^*]}$ and $H = H_{s^*} + H^*_{s^*}$;

For describing *RandomEmbedding* we need the following definitions. We shall use the symbol $\hookrightarrow$ to denote embeddings produced by *RandomEmbedding*. We write $G \hookrightarrow H$ to indicate that the graph G is to be embedded into H. Also, if $t \in V(G)$, $v \in V(H)$ and $A \subseteq V(H)$ then $t \hookrightarrow v$ means that t is embedded on v, and $t \hookrightarrow A$ means that t is embedded on a vertex of A.

Definition 13 (partial embedding, candidate set)**.** *Let G be a graph with vertex set $[v(G)]$, and H be a graph with $v(H) \geq v(G)$. Further, assume $\psi_j \colon [j] \to V(H)$ is a* partial embedding *of G into H for $j \in [v(G)]$, that is, ψ_j is a graph embedding of $G\big[[j]\big]$ into H. Finally, let $t \in [v(G)]$ be such that $N^-_G(t) \subseteq [j]$. Then the* candidate set of t *(with respect to ψ_j) is the common neighbourhood in H of the already embedded neighbours of t, that is,*

$$C^j_{G\hookrightarrow H}(t) = N_H\Big(\psi_j\big(N^-_G(t)\big)\Big) .$$

RandomEmbedding (see Algorithm 2) randomly embeds most of a guest graph G into a host graph H. The algorithm is simple: we iteratively embed the first $(1-\delta)n$ vertices of G randomly to one of the vertices of their candidate set which was not used for embedding another vertex already.

Algorithm 2: *RandomEmbedding*

Input : graphs G and H, with $V(G) = [v(G)]$ and $v(H) = n$
Output: an embedding ψ_{t^*} of $G\big[[n-\delta n]\big]$ into H
$\psi_0 := \emptyset$;
$t^* := (1-\delta)n$;
for $t = 1$ **to** t^* **do**
  if $C^{t-1}_{G\hookrightarrow H}(t) \setminus \mathrm{im}(\psi_{t-1}) = \emptyset$ **then** halt with failure;
  choose $v \in C^{t-1}_{G\hookrightarrow H}(t) \setminus \mathrm{im}(\psi_{t-1})$ uniformly at random;
  $\psi_t := \psi_{t-1} \cup \{t \hookrightarrow v\}$;
end
return ψ_{t^*};

If successful, *PackingProcess* returns the packing $(\phi^*_s)_{s\in[s^*]}$ of $(G''_s)_{s\in[s^*]}$ and a leftover graph H. For each $s \in [s^*]$, we obtain an embedding ϕ'_s of G'_s into $\widehat{H}$ from the embedding ϕ^*_s of G''_s into $\widehat{H}$ by ignoring the vertices of G''_s which are not in G'_s. Recall that all these vertices are isolated vertices. It follows that the $(\phi'_s)_{s\in[s^*]}$ give a packing of $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$ which leaves unused exactly the edges of H. In [1] it was shown that *PackingProcess* is indeed a.a.s. successful. We shall use the techniques developed there to show in Lemma 18 that moreover *PackingProcess* returns a packing of $(G'_s)_{s\in[s^*]}$ and a leftover graph with suitable properties for the following steps.

It remains to pack all the leaves we omitted from $(G_s)_{s\in[s^*]}$. For this we shall proceed vertex by vertex of the remaining host graph H, and when considering $r \in V(H)$ we shall randomly embed all leaves *dangling at* r, that is, the leaves of all guest graphs such that the neighbour of the leaf is already embedded to r. For describing this process in more detail, we will need the following definitions.

Definition 14 (weights)**.** *Let $(G_s)_{s\in[s^*]}$ be a (μ,n)-graph sequence, and $(G'_s)_{s\in[s^*]}$ be a corresponding subgraph sequence, H be an n-vertex graph, and $\phi'_s \colon V(G'_s) \to V(H)$ be an injection for each $s \in [s^*]$. For $s^* - \lfloor \mu n \rfloor < s \leq s^*$ we define for each $x \in V(G_s)$ the weight*

$$w_s(x) = \big|\{y \in N_{G_s}(x) \colon y \text{ is a leaf of } G_s \text{ in } G_s - G'_s\}\big| ,$$

and for each $v \in V(H)$ the weight

$$w_s(v) = w_s\big({\phi'}_s^{-1}(v)\big)\,.$$

Further, for each $v \in V(H)$ we define

$$w(v) = \sum_{s^*-\lfloor \mu n\rfloor < s \leq s^*} w_s(v)\,.$$

Note that since each set V_s of omitted leaves is an independent set in G_s, the weight of an omitted leaf is 0. Thus the entire weight of G_s (which is ℓ, the number of omitted leaves) is on the vertices in G'_s embedded by ϕ'_s. We next choose an orientation $\vec{H}$ of H such that $N^+_{\vec{H}}(r) = w(r)$ for each $r \in V(H)$. We shall show in Lemma 19 that we can choose an orientation with this property which is moreover random-like (in the sense that it suitably inherits the properties guaranteed by Lemma 18). The idea now is to embed the remaining leaves dangling at r by using only edges directed away from r. We define the following auxiliary graphs, which encode the ways in which we can embed the dangling leaves. Recall that the G_s are considered to be pairwise vertex disjoint, so that if $x \in V(G_s)$ then s is unique.

Definition 15 (leaf matching graphs)**.** *Given $r \in V(\vec{H})$, we define the* leaves at r *to be the set*

$$L_r := \Big\{x\colon \exists s \text{ such that } x \in V(G_s) \setminus V(G'_s) \text{ and } x{\phi'}_s^{-1}(r) \in E(G_s)\Big\}$$

Let the leaf matching graph F_r *be the bipartite graph with parts L_r and $N^+_{\vec{H}}(r)$, and edges xu with $x \in L_r$ and $u \in N^+_{\vec{H}}(r)$ whenever $u \notin \operatorname{im}\phi'_s$ for the s such that $x \in V(G'_s)$.*

Observe that a perfect matching in F_r defines an assignment of the leaves at (all preimages of) r to $N^+_{\vec{H}}(r)$ which extends the packing of $(G'_s)_{s\in[s^*]}$. We will see that each F_r is a graph whose parts have size roughly $\frac{1}{2}pn$ and whose density is roughly μ. If we simply chose a perfect matching in each F_r to embed all the leaves $\bigcup_r L_r$, then we would almost have a perfect packing—each edge of K_n would be used exactly once—but it could be the case that multiple leaves of some G_s (not in the same L_r) are embedded to a single $u \in V(H)$. To avoid this, we find perfect matchings in each F_r one at a time and update the leaf matching graphs by removing edges which are no longer useable. In order that not too many edges are removed from any one vertex in any F_r, we choose perfect matchings uniformly at random. Making this precise, assume $V(\vec{H}) = \{1,\dots,n\}$, and set $F^{(0)}_r := F_r$ for each $r \in V(\vec{H})$. We use the following algorithm.

We shall show that, throughout, the graphs $F^{(r)}_k$ satisfy a certain degree-codegree condition. We shall show in Lemma 20 that under this degree-codegree condition we can find a perfect matching in $F^{(r-1)}_r$. Further, the same lemma asserts that a perfect matching σ_r chosen uniformly at random in $F^{(r-1)}_r$ uses edges almost uniformly, which is important for maintaining the degree-codegree condition.

We will then, for each $s \in [s^*]$ and each $x \in V(G_s)$ set

$$\phi_s(x) = \begin{cases} \phi'_s(x) & \text{if } x \in \operatorname{dom}(\phi'_s) \\ \sigma_r(x) & \text{if } x \in L_r\,. \end{cases} \tag{1}$$

This is a perfect packing of $(G_s)_{s\in[s^*]}$ into $\widehat{H}$ since $(\phi'_s)_{s\in[s^*]}$ is a packing of the subgraph sequence $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$ and we chose matchings in the leaf matching graphs $F^{(r-1)}_r$ to embed the remaining leaves and updated the subsequent leaf matching graphs accordingly.

Algorithm 3: *MatchLeaves*

Input : a (μ, n)-graph sequence $(G_s)_{s\in[s^*]}$, a corresponding subgraph sequence $(G'_s)_{s\in[s^*]}$ omitting $\lfloor \nu n \rfloor$ leaves, and associated leaf matching graphs $F_1^{(0)}, \dots, F_n^{(0)}$
Output: matchings $(\sigma_r)_{r\in[n]}$ of the omitted leaves to feasible image vertices as given by the leaf matching graphs
for $r = 1$ **to** n **do**
  if $F_r^{(r-1)}$ has no perfect matching, halt with failure;
  let σ_r be a uniform random perfect matching in $F_r^{(r-1)}$;
  for $k = r + 1$ **to** n **do**
    let $B_k := \{xu \in E(F_k^{(r-1)}) \colon \exists s \text{ s.t. } x \in V(G_s) \setminus V(G'_s) \text{ and } \sigma_r^{-1}(u) \in V(G_s)\}$;
    let $F_k^{(r)} := F_k^{(r-1)} - B_k$;
  end
end
return $(\sigma_r)_{r\in[n]}$;

Summing up, our packing algorithm proceeds as described in Algorithm 4.

Algorithm 4: *PerfectPacking*

Input : graphs $G_1, \dots, G_{s^*}$ that form a (μ, n)-graph sequence such that the last $(D + 1)^{-3}n$ vertices of G_s form an independent set; a graph $\widehat{H}$ on n vertices
Output: A packing $(\phi_s)_{s\in[s^*]}$ of $(G_s)_{s\in[s^*]}$ into $\widehat{H}$
let $(G'_s)_{s\in[s^*]}$ be a subgraph sequence corresponding to $(G_s)_{s\in[s^*]}$ omitting $\lfloor \nu n \rfloor$ leaves;
for $s = s^* - \lfloor \mu n \rfloor + 1$ **to** s^* **do**
  let I'_s be a set of $n - v(G'_s)$ (new) isolated vertices;
  $G''_s := G'_s + I'_s$, where we place I'_s at the end of the degeneracy order;
end
foreach $s \in [s^*]$ **do** assume that $V(G''_s) = [n]$, with the natural degeneracy order;
run *PackingProcess* to obtain embeddings $(\phi^*_s)_{s\in[s^*]}$ of $(G''_s)_{s\in[s^*]}$ into $\widehat{H}$ with leftover H;
obtain embeddings $(\phi'_s)_{s\in[s^*]}$ of $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$ from $(\phi^*_s)_{s\in[s^*]}$ by ignoring the I'_s;
construct a random-like orientation $\vec{H}$ of H with $N^+_{\vec{H}}(r) = w(r)$ for all $r \in V(\vec{H})$;
foreach $r \in V(\vec{H})$ **do** let $F_r^{(0)}$ be the leaf matching graph F_r;
run *MatchLeaves* to obtain embeddings $(\sigma_r)_{r\in[n]}$ of the leaves at r;
for $s = 1$ **to** s^* *and* **for each** $x \in V(G_s)$ **do** set $\phi_s(x)$ as in (1);
return $(\phi_s)_{s\in[s^*]}$;

3.1. **Graphs and maps used in the algorithm.** As described above, a number of different (auxiliary) graphs and maps are used in our packing procedure. For the convenience of the reader we collect these in the following table.

G_s are the given n-vertex guest graphs, forming a D-degenerate (μ, n)-graph sequence, whose last $\lfloor (D+1)^{-3}n \rfloor$ vertices form an independent set.

G'_s is in the subgraph sequence corresponding to $(G_s)_{s\in[s^*]}$ omitting $\lfloor \nu n \rfloor$ leaves; the special G'_s have $n - \lfloor \mu n \rfloor$ vertices, the others n.

G''_s is obtained from G'_s by adding isolated vertices to the end of the degeneracy order until we have n vertices.

$\widehat{H}$ is the given n-vertex $(\xi, 2D+3)$-quasirandom host graph.

H_{s-1} is the part of $\widehat{H}$ used by *RandomEmbedding* to embed $G''_s\big[[n-\delta n]\big]$.

H^*_{s-1} is the part of $\widehat{H}$ used in *PackingProcess* to complete the embedding of G''_s.

H is the leftover host graph after running *PackingProcess*.

$\vec{H}$ is a random-like orientation of H with as many outgoing edges for each vertex r as there are leaves dangling at r.

$F_k^{(r)}$ is what remains of the leaf matching graph F_k after round r of *MatchLeaves*.

ϕ''_s embeds $G''_s\big[[n-\delta n]\big]$ into H_{s-1} and is constructed by *RandomEmbedding*.

ϕ^*_s is an extension of ϕ''_s, embedding G''_s into $H_{s-1} \cup H^*_{s-1}$ constructed in *PackingProcess*.

ϕ'_s is an embedding of G'_s into $\widehat{H}$ obtained from ϕ^*_s by ignoring the added isolated vertices.

ϕ_s is an embedding of G_s into $\widehat{H}$ obtained from ϕ'_s and the σ_r in *PerfectPacking*.

ψ_t is the partial embedding obtained in round t of *RandomEmbedding*.

σ_r is the matching in the leaf matching graph $F_r^{(r-1)}$ found by *MatchLeaves*.

4. Constants

In this section we set values for the various constants we need throughout our proofs (including those used in the algorithms above), which are the following.

α_x is the error in the quasirandomness of H_x.

α is a quasirandomness error used in auxiliary lemmas; we always assume $\alpha_0 \le \alpha \le \alpha_{2n}$.

β_t is the error in the diet-condition (see Definition 30) for round t of *RandomEmbedding*.

c is the constant in the maximum degree bound of the G_s.

C appears in the error term for the probability of embedding a fixed vertex of G''_s on a fixed vertex of H_{s-1}.

C' appears in the error term for the fraction of vertices of certain sets that get covered by embedding one graph G''_s

D is the degeneracy bound of the guest graphs G_s.

n_0 is the lower bound on the number n of vertices.

$\hat{p}$ is the density of the host graph $\widehat{H}$.

$\hat{p}_0$ is the lower bound on $\hat{p}$.

p is the density of the leftover host graph H after running *PackingProcess* to embed the subgraph sequence.

p_s is the density of the graph H_s.

δ is the proportion of vertices in G''_s formed by the independent set at the end of the degeneracy order as required by *PackingProcess*.

ε gives the length εn of intervals in $V(G_s)$ used in the cover condition (see Definition 30); it also appears in the error term of the cover condition.

η is the error in the quasirandomness of H_0^*.

γ is the proportion of host graph edges used by *PackingProcess* to complete almost spanning embeddings to spanning embeddings.

γ' determines the error bound in our analysis here of *PackingProcess*.

μ specifies the fraction of special G_s, how far they are from spanning and how many leaves they have.

ν specifies the fraction of leaves omitted in the subgraph sequence.

ξ is the error in the quasirandomness of $\widehat{H}$.

The constants D, $\hat{p}$ and μ are provided as input to our main technical theorem; the other constants are chosen to satisfy

$$0 \ll \frac{1}{n_0}, c \ll \varepsilon \le \xi \ll \alpha_0 \le \alpha_{2n} \le \frac{1}{C'} \ll \frac{1}{C} \ll \delta \ll \eta \ll \gamma \ll \gamma' \ll \nu \ll \mu, \hat{p}_0, \frac{1}{D}\,.$$

Here $a \ll b$ means that we choose a sufficiently small in terms of b, so that our calculations work. In other words, there is a monotone increasing function $f\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ with $f(b) \le b$ such that we choose $a = f(b)$.

For the constants ν, γ', γ, and n_0 we do not provide explicit dependencies (mainly because Lemma 50, which we take from elsewhere, does not provide explicit dependencies), but merely state that we can choose them suitably with the above relations.

The various host graph densities satisfy the following relations. We have $\hat{p} \ge \hat{p}_0$. Given $n \ge n_0$, the density p is determined by

$$p = \lfloor \mu n \rfloor \lfloor \nu n \rfloor \binom{n}{2}^{-1}\,. \tag{2}$$

Moreover, in our later proofs we will assume that $e(H_0^*) = (1 \pm \frac{1}{10})\gamma\binom{n}{2}$, which can be seen to hold with probability larger than $1 - e^{-n}$ by an application of Theorem 5. Then, since the density of $\widehat{H}$ is $\hat{p}$, the density of H_0 is $p_0 \ge \hat{p} - 1.1\gamma$, and therefore, because *PackingProcess* embeds $\sum_{s\in[s^*]} e(G_s) - \lfloor \mu n \rfloor \lfloor \nu n \rfloor \le \hat{p}\binom{n}{2} - \lfloor \mu n \rfloor \lfloor \nu n \rfloor$ edges we have

$$p_s \ge \hat{p} - 1.1\gamma - \frac{\hat{p}\binom{n}{2} - \lfloor \mu n \rfloor \lfloor \nu n \rfloor}{\binom{n}{2}} \ge \nu\mu - 1.1\gamma \ge \gamma \qquad \text{for all } s \in [s^*]\,. \tag{3}$$

The remaining constants are only used in the proof of the most technically involved of our main lemmas, Lemma 18. These are defined precisely in the same way as in [1] apart from C', which is added here.[2] This is important, because much of our proof of Lemma 18 builds on tools developed in [1], and the relation of the constants involved is somewhat more intricate.

[2]The density of the host graph $\widehat{H}$ is denoted by p in [1]; we denote it by $\hat{p}$ here.

Setting 16. *Let $D, n \in \mathbb{N}$ and $\hat{p}, \gamma > 0$ be given. We define*

$$\begin{aligned} &\eta = \frac{\gamma^D}{200D}\,, \quad \delta = \frac{\gamma^{10D}\eta}{10^6 D^4}\,, \quad C = 40D\exp\left(1000D\delta^{-2}\gamma^{-2D-10}\right), \quad C' = 10^4 C\delta^{-1}\,, \\ &\alpha_x = \frac{\delta}{10^8 CD}\exp\left(\frac{10^8 CD^3\delta^{-1}(x-2n)}{n}\right) \qquad \text{for each } x \in \mathbb{R}, \\ &\varepsilon = \alpha_0\delta^2\gamma^{10D}/1000CD\,, \quad c = D^{-4}\varepsilon^4/100 \quad \text{and} \quad \xi = \alpha_0/100\,. \end{aligned} \tag{4}$$

Moreover, given $\alpha > 0$ we use the following constants $\beta_t(\alpha)$, which are chosen such that $\beta_0(\alpha) = \alpha$ and such that $\beta_{v(G)}(\alpha)/\beta_0(\alpha)$ is bounded by a constant which does not depend on α (though it does depend on D, γ and δ). We define

$$\beta_t(\alpha) = 2\alpha\exp\left(\tfrac{1000D\delta^{-2}\gamma^{-2D-10}t}{n}\right)\,. \tag{5}$$

Remark 17. When using the constants α_x, β_t, we will mainly take x and t integer in the range $[0, 2n]$, but it is convenient to allow them to be any real number.

Note that we call α_x and β_t 'constant' even though n appears in their definition. It is easy to check though that α_x is strictly increasing in x and β_t is strictly increasing in t and that neither α_0, β_0 nor α_{2n}, β_{2n} depends on n. Further, for each $t \geq 0$, we have

$$\begin{aligned} &\tfrac{1}{n}\int_{i=0}^{t} 1000D\delta^{-2}\gamma^{-2D-10}\beta_i \,\mathrm{d}i \\ \leq &2\alpha\int_{i=-\infty}^{t} \frac{1000D\delta^{-2}\gamma^{-2D-10}}{n}\exp\left(\tfrac{1000D\delta^{-2}\gamma^{-2D-10}i}{n}\right)\mathrm{d}i = \beta_t\,. \end{aligned} \tag{6}$$

5. Main lemmas

In this section we collect the main lemmas we need for the proof of our main technical theorem. Our first lemma states that the randomised algorithm *PackingProcess* generates an almost perfect packing of the corresponding subgraph sequence of our guest graphs such that this packing and the leftover H of the host graph satisfy certain properties. We prove this lemma in Section 8. The fact that *PackingProcess* produces a packing of this type such that the leftover H is quasirandom is the main result of [1]. Here, we need to establish additional properties for completing this to a perfect packing.

Lemma 18 (almost perfect packing lemma)**.** *Assume $0 \ll c \ll \xi \ll \delta \ll \gamma \ll \gamma' \ll \nu \ll \mu, \hat{p}, \frac{1}{D}$. Let $\widehat{H}$ be a $(\xi, 2D+3)$-quasirandom graph with n vertices and density $\hat{p}$. Let $s^* \leq \frac{7}{4}n$, let $(G_s)_{s\in[s^*]}$ be a D-degenerate (μ, n)-graph sequence with maximum degree $\frac{cn}{\log n}$ and $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$, and let $(G'_s)_{s\in[s^*]}$ be a corresponding subgraph sequence omitting $\lfloor \nu n\rfloor$ leaves. Then* PackingProcess *(applied with constants γ and δ to the graphs $(G''_s)_{s\in[s^*]}$ obtained in* PerfectPacking *from $(G'_s)_{s\in[s^*]}$ by adding isolated vertices) a.a.s. provides a packing $(\phi'_s)_{s\in[s^*]}$ of $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$ with leftover H such that for $p = \lfloor \mu n\rfloor\lfloor \nu n\rfloor\binom{n}{2}^{-1}$ we have*

(P1) H is $(\gamma'^3, 2D+3)$-quasirandom and has density p,

for all $v \in V(H)$ and $s^* - \lfloor \mu n\rfloor < s, s' \leq s^*$ we have

(P2) $w(v) = (1 \pm \gamma'^3)\frac{pn}{2}$,

(P3) $\left|N_H(v) \setminus \operatorname{im}\phi'_s\right| = (1 \pm \gamma'^3)\mu pn$,

(P4) $\left|N_H(v) \setminus (\operatorname{im}\phi'_s \cup \operatorname{im}\phi'_{s'})\right| = (1 \pm \gamma'^3)\mu^2 pn$ if $s \neq s'$,

for all $u, v \in V(H)$ with $u \neq v$ we have,

(P5) $\sum_s w_s(v) \mathbb{1}_{u \notin \operatorname{im} \phi'_s} = (1 \pm \gamma'^3)\mu \frac{pn}{2}$,

and for all $u \in V(H)$ and $s^* - \lfloor \mu n \rfloor < s \leq s^*$ we have

(P6) *If $u \notin \operatorname{im} \phi'_s$ then* $\sum_{v\colon vu \in E(H)} w_s(v) < \frac{10p^2 n}{\mu}$.

Our second lemma states that there is an orientation of H suitable for completing the perfect packing by embedding the leaves with the help of the algorithm *MatchLeaves*. A *random orientation* of a graph $H = (V, E)$ is an orientation of H in which the orientation of each edge $\{u, v\} \in E$ is chosen independently and uniformly at random. We prove this lemma in Section 7.

Lemma 19 (orientation lemma)**.** *Let H be a $(\gamma'^3, 2)$-quasirandom graph of density p with vertex weights $w\colon V(H) \to \mathbb{N}_0$ such that $w(v) = (1 \pm \gamma'^3)\frac{pn}{2}$ for all $v \in V(H)$ and such that $\sum_{v \in V} w(v) = e(H)$. If $\vec{H}_0$ is a random orientation of H, then a.a.s. there is an orientation $\vec{H}$ of H such that for all $v \in V(H)$*

(O1) $\deg^+_{\vec{H}}(v) = w(v)$, *and*

(O2) $\left|\{uv \in E(H)\colon uv \text{ is oriented differently in } \vec{H} \text{ and } \vec{H}_0\}\right| \leq \gamma'^2 n$.

Our last lemma states that if in a graph F satisfying a certain degree-codegree condition, we remove a few edges and then choose a perfect matching uniformly at random, then each edge is roughly equally likely to appear in the matching. In the proof of our main theorem, we shall show that the leaf matching graphs F_v satisfy these conditions, and hence *MatchLeaves* can find a perfect matching in F_v, using edges almost uniformly.

Lemma 20 (matching lemma)**.** *Assume $0 \ll \frac{1}{m} \ll p \ll \mu \ll 1$. Let $F = F[U, W]$ be a bipartite graph with $|U| = |W| = (1 \pm p)m$ such that*

(M1) $\deg_F(x) = (1 \pm p)\mu m$ *for all $x \in U \cup W$, and*

(M2) $\deg_F(u, u') = (1 \pm p)\mu^2 m$ *for all but at most $\frac{m^2}{\log m}$ pairs $\{u, u'\} \in \binom{U}{2}$,*

and let $F' = F'[U, W]$ be a spanning subgraph of $F[U, W]$ such that

(M3) $\deg_F(x) - \deg_{F'}(x) < \frac{100pm}{\mu^2}$ *for all $x \in U \cup W$.*

Then F' has a perfect matching and for a perfect matching σ chosen uniformly at random among all perfect matchings in F' and for all $uw \in E(F')$ we have

$$\mathbb{P}[\sigma(u) = w] \leq \frac{2}{\mu m}\,.$$

This lemma is a straightforward consequence of a lemma (Lemma 50) on random matchings in super-regular pairs by Felix Joos (see [17]) and the degree-codegree characterisation of super-regular pairs (Lemma 49) provided by Duke, Lefmann, and Rödl in [4]. For completeness, we provide the deduction in Appendix B.

6. Proof of the main technical theorem

To prove Theorem 11 we shall run the algorithm *PerfectPacking* (Algorithm 4), which uses *PackingProcess* to pack the G'_s. The resulting graph H of unused edges is likely to satisfy the conclusions of the almost perfect packing lemma (Lemma 18). *PerfectPacking* then chooses a random orientation $\vec{H}_0$ of H and modifies this orientation slightly to obtain

$\vec{H}$, which satisfies the conclusions of the orientation lemma (Lemma 19) and also oriented versions of properties $(P3)$ and $(P4)$ of Lemma 18. Finally, *PerfectPacking* runs *MatchLeaves* to complete the packing. To show that *MatchLeaves* succeeds, we will verify that with high probability for each r the graphs F_r and $F_r^{(r-1)}$ satisfy the conditions of the matching lemma (Lemma 20). For this we use Corollary 7 and the union bound.

Proof of Theorem 11. We use constants with relations as given in Section 4, that is

$$0 \ll c \ll \xi \ll \delta \ll \gamma \ll \gamma' \ll \nu \ll \mu, \hat{p}_0, \frac{1}{D},$$

and $\hat{p} \geq \hat{p}_0$. Suppose that $\widehat{H}$ is an $(\xi, 2D+3)$-quasirandom graph with n vertices and density $\hat{p}$. Suppose that $s^* \leq \frac{7}{4}n$ and that the graph sequence $(G_s)_{s\in[s^*]}$ is a D-degenerate (μ, n)-graph sequence, with maximum degree $\Delta \leq \frac{cn}{\log n}$, such that the last $\lceil (D+1)^{-3}n \rceil$ vertices in the degeneracy order form an independent set in G_s, and all have the same degree d_s in G_s. Suppose further that $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$. We use *PerfectPacking* (Algorithm 4) for packing $(G_s)_{s\in[s^*]}$ into $\widehat{H}$ and argue in the following that it succeeds a.a.s.

As $\lfloor \nu n \rfloor < \mu n$, *PerfectPacking* can choose a corresponding subgraph sequence $(G'_s)_{s\in[s^*]}$ omitting $\lfloor \nu n \rfloor$ leaves. Next it creates for each $s \in [s^*]$ a graph G''_s. For the non-special graphs ($s \leq s^* - \lfloor \mu n \rfloor$) it sets $G''_s := G'_s$. For the special graphs ($s > s^* - \lfloor \mu n \rfloor$) it obtains G''_s by adding the set I'_s of $n - v(G'_s)$ isolated vertices, which we place at the end of the D-degeneracy order. Note that for each G''_s the last δn vertices of G''_s in the degeneracy order are an independent set all of whose vertices have degree d_s. Indeed, if $s \leq s^* - \lfloor \mu n \rfloor$ then this holds by assumption on G_s and because $\delta < (D+1)^{-3}$, and if $s > s^* - \lfloor \mu n \rfloor$ then this holds because $n - v(G'_s) = n - \lfloor \mu n \rfloor$ and $\delta < \mu$ (and in this case $d_s = 0$).

PerfectPacking next runs *PackingProcess* with input $(G''_s)_{s\in[s^*]}$ and $\widehat{H}$. By the almost perfect packing lemma (Lemma 18), *PackingProcess* a.a.s. returns a packing $(\phi^*_s)_{s\in[s^*]}$ of $(G''_s)_{s\in[s^*]}$ into $\widehat{H}$, and a graph H consisting of all the edges not used in the packing, which satisfies the conclusions $(P1)$–$(P6)$ of Lemma 18. As described in *PerfectPacking*, we let for each $s \in [s^*]$ the map ϕ'_s be the embedding of G'_s into $\widehat{H}$ induced by ϕ^*_s. By construction of the $(G''_s)_{s\in[s^*]}$, the $(\phi'_s)_{s\in[s^*]}$ form a packing of the $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$, with H being the graph formed by the unused edges. The total number of unused edges is by construction $\lfloor \mu n \rfloor \lfloor \nu n \rfloor = p\binom{n}{2}$, so H has density p.

PerfectPacking next chooses a random-like orientation of H. More precisely, we want to use an orientation $\vec{H}$ of H such that $w(v) = \deg^+_{\vec{H}}(v)$ for each $v \in V(H)$, which in addition inherits oriented versions of $(P3)$ and $(P4)$. The next claim states that such an orientation exists.

Claim 21. *For all sufficiently large n there exists an orientation $\vec{H}$ of H such that $w(v) = \deg^+_{\vec{H}}(v)$ for each $v \in V(H)$, and in addition for each $s^* - \lfloor \mu n \rfloor < s, s' \leq s^*$ we have*

(*P'3*) $\left| N^+_{\vec{H}}(v) \setminus \operatorname{im} \phi'_s \right| = (1 \pm \gamma') \frac{\mu p n}{2}$, *and*

(*P'4*) $\left| N^+_{\vec{H}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'}) \right| = (1 \pm \gamma') \frac{\mu^2 p n}{2}$ *if $s \neq s'$.*

Proof. Recall that $\sum_s e(G_s) = e(\widehat{H})$, and hence $\sum_s e(G_s) - e(G'_s) = e(H)$. By definition, each edge of $e(G_s) - e(G'_s)$ contributes weight one to $w_s(x)$, where $x \in G'_s$, and hence weight one to $w(v)$ where v is the vertex of H to which x is embedded. We conclude $\sum_{v\in V(H)} w(v) = e(H)$.

By $(P1)$, in particular H is $(\gamma'^3, 2)$-quasirandom and of density p, and by $(P2)$ we have $w(v) = (1 \pm \gamma'^3)\frac{pn}{2}$ for all $v \in V(H)$. This verifies that H satisfies the conditions of the orientation lemma (Lemma 19).

Let $\vec{H}_0$ be a random orientation of H. Given $v \in V(H)$ and $s^* - \lfloor \mu n \rfloor < s \leq s^*$, by $(P3)$ and Theorem 5, with probability at least $1 - \exp\big(-\frac{\gamma'^6 \mu pn}{12}\big)$ we have

$$\big|N^+_{\vec{H}_0}(v) \setminus \operatorname{im}\phi'_s\big| = (1 \pm 3\gamma'^3)\tfrac{\mu pn}{2}\,.$$

Similarly, given $v \in V(H)$ and $s^* - \lfloor \mu n \rfloor < s < s' \leq s^*$, by $(P4)$ and Theorem 5, with probability at least $1 - \exp\big(-\frac{\gamma'^6 \mu^2 pn}{12}\big)$ we have

$$\big|N^+_{\vec{H}_0}(v) \setminus (\operatorname{im}\phi'_s \cup \operatorname{im}\phi'_{s'})\big| = (1 \pm 3\gamma'^3)\tfrac{\mu^2 pn}{2}\,.$$

Taking the union bound, and by Lemma 19, with probability at least $1 - 2n^3 \exp\big(-\frac{\gamma'^6 \mu^2 pn}{12}\big) - o(1)$ each of the above good events holds for each $v \in V(H)$ and each $s^* - \lfloor \mu n \rfloor < s, s' \leq s^*$, and in addition there is an orientation $\vec{H}$ of H satisfying conclusions $(O1)$ and $(O2)$ of Lemma 19.

For sufficiently large n we have $1 - 2n^3 \exp\big(-\frac{\gamma'^6 \mu^2 pn}{12}\big) - o(1) > 0$, so we fix $\vec{H}_0$ and $\vec{H}$ satisfying all these properties. By $(O1)$ the orientation $\vec{H}$ satisfies $\deg^+_{\vec{H}}(v) = w(v)$ for each $v \in V(H)$, as desired. Given $v \in V(H)$ and $s^* - \lfloor \mu n \rfloor < s \leq s^*$, by $(O2)$ we have

$$\big|N^+_{\vec{H}}(v) \setminus \operatorname{im}\phi'_s\big| = \big|N^+_{\vec{H}_0}(v) \setminus \operatorname{im}\phi'_s\big| \pm \gamma'^2 n = (1 \pm 3\gamma'^3)\tfrac{\mu pn}{2} \pm \gamma'^2 n = (1 \pm \gamma')\tfrac{\mu pn}{2}\,,$$

where the final inequality is by choice of γ'. This verifies $(P'3)$. Similarly, given $v \in V(H)$ and $s^* - \lfloor \mu n \rfloor < s < s' \leq s^*$, we have

$$\begin{aligned}\big|N^+_{\vec{H}}(v) \setminus (\operatorname{im}\phi'_s \cup \operatorname{im}\phi'_{s'})\big| &= \big|N^+_{\vec{H}_0}(v) \setminus (\operatorname{im}\phi'_s \cup \operatorname{im}\phi'_{s'})\big| \pm \gamma'^2 n = (1 \pm 3\gamma'^3)\tfrac{\mu^2 pn}{2} \pm \gamma'^2 n \\ &= (1 \pm \gamma')\tfrac{\mu^2 pn}{2}\,,\end{aligned}$$

giving $(P'4)$. □

This orientation is now used to embed the remaining dangling leaves. *PerfectPacking* runs *MatchLeaves* (Algorithm 3) for this purpose. Recall that, for a vertex $v \in V(\vec{H})$, the leaf matching graph F_v (see Definition 15) is a bipartite graph with parts consisting of the leaves L_v which we need to embed at v (which will be in many different G_s) and the out-neighbours $N^+_{\vec{H}}(v)$ of v in $\vec{H}$ to which we will embed these leaves, with an edge from a leaf in some G_s to an out-neighbour u of v if $u \notin \operatorname{im}\phi'_s$. Recall that for convenience we assume $V(\vec{H}) = [n]$. *MatchLeaves* starts with $F_v^{(0)} := F_v$ for each $v \in [n]$, and then for each $r \in [n]$ in succession takes a random perfect matching σ_r in $F_r^{(r-1)}$ and for each $k > r$ removes some edges from $F_k^{(r-1)}$ to form $F_k^{(r)}$. As explained in Section 3, it is enough to show that with positive probability *MatchLeaves* does not halt with failure. To analyse the running of *MatchLeaves*, we aim to show that for each r the graphs $F_r^{(0)}$ and $F_r^{(r-1)}$ satisfy the conditions of Lemma 20 with $m := \frac{pn}{2}$, with $F = F_r^{(0)}$ and $F' = F_r^{(r-1)}$, and with $U = L_r$ and $W = N^+_{\vec{H}}(r)$. We shall then use Lemma 20 to conclude that the matching σ_r we choose in $F_r^{(r-1)}$ does not use any given edge with exceptionally high probability, which in turn will allow us to show that *MatchLeaves* is successful.

Property $(M1)$: Given $x \in V(F_r^{(0)})$, we separate two cases. If $x \in L_r$ is in the graph G_s, then by $(P'3)$ we have $\deg_{F_r^{(0)}}(x) = \left|N^+_{\vec{H}}(r) \setminus \operatorname{im}\phi'_s\right| = (1 \pm \gamma')\frac{\mu pn}{2}$. If $x \in N^+_{\vec{H}}(r)$, then by $(P5)$ we have $\deg_{F_r^{(0)}}(x) = \sum_s w_s(r)\mathbb{1}_{x \notin \operatorname{im}\phi'_s} = (1 \pm {\gamma'}^3)\frac{\mu pn}{2}$. In either case, since $p > \gamma'$ this verifies $(M1)$ for $F = F_r^{(0)}$, $F' = F_r^{(r-1)}$ and every $r \in [n]$.

Property $(M2)$: Given $u, u' \in L_r$, if $u \in V(G_s)$ and $u' \in V(G_{s'})$, where $s \neq s'$, then by $(P'4)$ we have $\deg_{F_r^{(0)}}(u,u') = \left|N^+_{\vec{H}}(r) \setminus (\operatorname{im}\phi'_s \cup \operatorname{im}\phi'_{s'})\right| = (1 \pm \gamma')\frac{\mu^2 pn}{2}$. Again since $\gamma' < p$ this is as required by $(M2)$, and we only need to show that the number of $u, u' \in L_r$ which are both in G_s for some $s \in [s^*]$ is at most $\frac{p^2n^2}{4\log(pn/2)}$. But any given G_s has at most $w_s(r) \leq \Delta = \frac{cn}{\log n}$ vertices in L_r, so that for a given u there are at most $\frac{cn}{\log n}$ choices of u' with $u, u' \in V(G_s)$ for some $s \in [s^*]$. Since $|L_r| \leq n$ we conclude that there are at most $\frac{cn^2}{\log n} < \frac{p^2n^2}{4\log(pn/2)}$ pairs $u, u' \in L_r$ such that $u, u' \in V(G_s)$ for some $s \in [s^*]$. This completes the verification of $(M2)$ for $F = F_r^{(0)}$, $F' = F_r^{(r-1)}$ and every $r \in [n]$.

Property $(M3)$: This property does not hold deterministically, but we shall show that it holds for all r with high probability. For this purpose we define the following events. For each $r \in [n]$ let $\mathcal{E}_r$ be the event that for each $y \in V(F_r^{(0)})$ we have

$$\deg_{F_r^{(0)}}(y) - \deg_{F_r^{(r-1)}}(y) \leq 50p^2n\mu^{-2}\,, \tag{7}$$

that is, $\mathcal{E}_r$ is the event that $(M3)$ holds for $F = F_r^{(0)}$ and $F' = F_r^{(r-1)}$. What we want to do is show that it is likely each $\mathcal{E}_r$ holds, since we then obtain the following claim.

Claim 22. *For $r \in [n]$, for $u \in N^+_{\vec{H}}(r)$ and for $s^* - \lfloor \mu n \rfloor < s \leq s^*$ the following holds. Either $\mathcal{E}_r$ does not occur, or $F_r^{(r-1)}$ has a perfect matching, and a random perfect matching σ_r in $F_r^{(r-1)}$ satisfies*

$$\mathbb{P}\Big[\sigma_r^{-1}(u) \in V(G_s) | \mathscr{H}_{r-1}\Big] \leq \tfrac{4w_s(r)}{\mu pn}\,,$$

where $\mathscr{H}_{r-1}$ denotes the collection of perfect matchings $\sigma_1, \ldots, \sigma_{r-1}$.

Proof. If $\mathcal{E}_r$ occurs, then all properties $(M1)$–$(M3)$ from Lemma 20 are satisfied with $F = F_r^{(0)}$ and $F' = F_r^{(r-1)}$. Thus a perfect matching in $F_r^{(r-1)}$ exists, and furthermore a random matching σ_r in F' satisfies for any given edge $xu \in E\big(F_r^{(r-1)}\big)$

$$\mathbb{P}\big[xu \in \sigma_r \big| \mathscr{H}_{r-1}\big] \leq \tfrac{4}{\mu pn}.$$

By the union bound over the $w_s(r)$ choices of $x \in L_r$ which are in G_s, the claim follows. □

To prove the main theorem, we only need to know that each perfect matching exists; however to analyse the probability of $\mathcal{E}_r$ we need the statement about random perfect matchings in addition. At this point, the following claim completes the proof of our main theorem.

Claim 23. *With probability at least $1-n^{-1}$, the event $\mathcal{E}_r$ holds for each $r \in [n]$ simultaneously.*

Before giving the proof of this claim, we spell out the details of how it implies the main theorem. If $\mathcal{E}_r$ holds then by Claim 22 the perfect matching σ_r exists for each r, and so Algorithm 3 does not halt with failure in round r.

Hence, assuming Claim 23, we get that Algorithm 3 does not halt with failure at all with probability at least $1 - n^{-1}$ and provides matchings $\sigma_1, \dots, \sigma_r$. *PerfectPacking* uses these matchings to define for each $s \in [s^*]$ the map $\phi_s : V(G_s) \to V(\widehat{H})$ by setting

$$\phi_s(x) = \begin{cases} \phi_s'(x) & \text{if } x \in V(G_s) \cap \operatorname{dom}(\phi_s) \\ \sigma_r(x) & \text{if } x \in L_r \,. \end{cases}$$

Recall that for each s, the map ϕ_s' is an embedding of G_s' into $\widehat{H}$. All the edges of G_s which are not in G_s' have one end in the removed leaves V_s and the other end in $V(G_s')$. Consider those leaves of G_s which are adjacent to $x \in V(G_s')$. By definition, these are in $L_{\phi_s(x)}$ and by construction of $F_{\phi_s(x)}^{(\phi_s(x)-1)}$, they are embedded to distinct vertices of $\widehat{H}$ which are adjacent in H to $\phi_s(x)$ and which are neither in $\operatorname{im}\phi_s$, nor are of the form $\sigma_i(y)$ for some $i < \phi_s(x)$ and $y \in V_s$. It follows that ϕ_s is indeed an embedding of G_s into $\widehat{H}$ for each $s \in [s^*]$.

We now check that these embeddings together form a packing. The maps $(\phi_s')_{s\in[s^*]}$ pack the graphs $(G_s')_{s\in[s^*]}$ into $\widehat{H}$, leaving exactly the edges of H unused. By construction $\vec{H}$ is an orientation of H, so for $\vec{vu} \in E(\vec{H})$, the edge $uv \in E(H)$ is used in the embedding of G_s, where $\sigma_v^{-1}(u) \in V(G_s)$. It follows that each edge of $\widehat{H}$ is used in the maps $(\phi_s)_{s\in[s^*]}$ at least once, and since $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$ each edge must be used exactly once. This justifies that the maps $(\phi_s)_{s\in[s^*]}$ perfectly pack the graphs $(G_s)_{s\in[s^*]}$ into $\widehat{H}$, as desired. This proves Theorem 11, assuming Claim 23.

So it remains to verify Claim 23. We shall first argue that the claimed probability bound follows from a probability bound, given in (8) below, which is of the right form to use Corollary 7. Indeed, let $\mathcal{A}_r$ be the event that $\mathcal{E}_i$ holds for each $1 \le i < r$ but $\mathcal{E}_r$ does not hold. Observe that if for each r the event $\mathcal{A}_r$ does not hold, then $\mathcal{E}_r$ holds for each $r \in [n]$. In particular, by the union bound over $r \in [n]$, if we establish that for each fixed $r \in [n]$ we have $\mathbb{P}[\mathcal{A}_r] \le n^{-2}$, then we conclude $\mathbb{P}[\bigcap_r \mathcal{E}_r] \ge 1 - n^{-1}$, which is the statement of Claim 23. Further, by another union bound over the at most $v(F_r^{(0)}) = 2w(r) \le 2n$ different $y \in V(F_r^{(0)})$ and since $\mathcal{A}_r \subseteq \bigcap_{1\le i\le r-1} \mathcal{E}_i$ it is enough to show that for any fixed $y \in V(F_r^{(0)})$ we have

$$\mathbb{P}\Big[\bigcap_{1\le i\le r-1} \mathcal{E}_i \quad \text{and} \quad \deg_{F_r^{(0)}}(y) - \deg_{F_r^{(r-1)}}(y) > 50p^2n\mu^{-2}\Big] \le \tfrac{1}{2}n^{-3}\,, \tag{8}$$

where we used the definition of $\mathcal{E}_i$ (see (7)). The remainder of this proof is devoted to establishing this bound. We will use Corollary 7 for this purpose, with the good event $\bigcap_{1\le i\le r-1} \mathcal{E}_i$. To that end, define for each $1 \le i \le r-1$ the random variable

$$Y_i := \deg_{F_r^{(i-1)}}(y) - \deg_{F_r^{(i)}}(y)$$

and observe that

$$\deg_{F_r^{(0)}}(y) - \deg_{F_r^{(r-1)}}(y) = \sum_{i=1}^{r-1} Y_i\,.$$

To apply Corollary 7 we need to find the range of each Y_i and the expectation of each Y_i, conditioned on the history $\mathscr{H}_{i-1}$ which consists of the collection of matchings $\sigma_1, \dots, \sigma_{i-1}$. This is encapsulated in Claim 24.

Claim 24. *For each $1 \le i \le r-1$, we have $0 \le Y_i \le \Delta$. Furthermore, either some $\mathcal{E}_i$ with $1 \le i \le r-1$ does not occur, or we have $\sum_{i=1}^{r-1} \mathbb{E}[Y_i|\mathscr{H}_{i-1}] \le 25p^2n\mu^{-2}$.*

Proof. We first show $0 \le Y_i \le \Delta$. There are two cases to consider. First, if $y \in L_r$, then y is in G_s for some $s \in [s^*]$. An edge yu of $F_r^{(i-1)}$ is removed to form $F_r^{(i)}$ only if u is assigned a leaf of G_s in σ_i. Since there are at most $w_s(i) \le \Delta$ such leaves, we have $Y_i \le \Delta$ in this case. Second, if $y \in N^+_{\vec{H}}(r)$, and y is assigned a leaf of G_s in σ_i, then we remove all edges of $F_r^{(i-1)}$ from y to leaves of G_s to form $F_r^{(i)}$. Since σ_i is a matching, this happens for at most one $s \in [s^*]$. There are at most $w_s(r) \le \Delta$ such leaves of G_s, so also in this case we have $Y_i \le \Delta$.

We now give an upper bound on the sum of conditional expectations. Again, there are two cases to consider. First, if $y \in L_r$, then let s be such that $y \in V(G_s)$. Suppose that $\mathscr{H}_{i-1}$ is a history up to and including σ_{i-1} such that $\mathcal{E}_i$ holds. Recall that Y_i is defined to be the change in degree of y in round i; that is, it is the number of $u \in N_{F_r^{(i-1)}}(y)$ which are not in $N_{F_r^{(i)}}(y)$. By linearity of expectation, $\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big]$ is the sum over $u \in N_{F_r^{(i-1)}}$ of the probability that $u \notin N_{F_r^{(i)}}(y)$ conditional on $\mathscr{H}_{i-1}$. Now a given u contributes to this sum exactly when a leaf of G_s dangling at i is matched to u by σ_i; this can only occur if $\vec{iu} \in E(\vec{H})$. So we can restrict to summing over $u \in N_{F_r^{(i-1)}}$ such that $\vec{iu} \in E(\vec{H})$, and for such a u the probability of its contributing to the sum is $\mathbb{P}\big[\sigma_i^{-1}(u) \in V(G_s)\big|\mathscr{H}_{i-1}\big]$. Putting this together, we obtain

$$\begin{aligned}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] &= \sum_{\substack{u \in N_{F_r^{(i-1)}}(y)\\ \vec{iu}\in E(\vec{H})}} \mathbb{P}\big[\sigma_i^{-1}(u) \in V(G_s)\big|\mathscr{H}_{i-1}\big]\\ &\le \sum_{\substack{u \in N_{F_r^{(i-1)}}(y)\\ \vec{iu}\in E(\vec{H})}} w_s(i)\tfrac{4}{\mu p n} \le \sum_{u \in N_H(r,i)} w_s(i)\tfrac{4}{\mu p n}\,,\end{aligned}$$

where the first inequality is by Claim 22 and the second holds since $\vec{iu} \in E(\vec{H})$ implies $iu \in E(H)$ and since $u \in N_{F_r^{(i-1)}}(y)$ implies $ru \in E(H)$. Summing over i, either some $\mathcal{E}_i$ with $i \in [r-1]$ does not hold, or we have

$$\sum_{i=1}^{r-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] \le \sum_{i=1}^{r-1}\sum_{u\in N_H(r,i)} w_s(i)\tfrac{4}{\mu p n} \le \sum_{i=1}^{n}\big|N_H(r,i)\big|\cdot w_s(i)\tfrac{4}{\mu p n} \le 2p^2 n\cdot\tfrac{4}{\mu p n}\sum_{i=1}^{n} w_s(i)\,,$$

where the final inequality is by $(P1)$. Recall that we defined $p = \lfloor \mu n\rfloor\lfloor \nu n\rfloor\binom{n}{2}^{-1}$, so in particular $\nu n \le \frac{pn}{\mu}$. Since $\sum_{i=1}^{n} w_s(i) = \lfloor \nu n\rfloor \le \frac{pn}{\mu}$ counts the number of leaves removed from G_s to form G'_s, we obtain that either some $\mathcal{E}_i$ with $i \in [r-1]$ does not hold, or

$$\sum_{i=1}^{r-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] \le \tfrac{8p^2 n}{\mu^2}\,,$$

as desired.

Finally, we consider the case $y \in N^+_{\vec{H}}(r)$. If a leaf of G_s is assigned to y by σ_i, it follows that y is adjacent to $w_s(r)$ leaves of G_s in $F_r^{(i-1)}$ and the edges to these leaves are exactly the edges at y removed from $F_r^{(i-1)}$ to form $F_r^{(i)}$. Suppose that $\mathscr{H}_{i-1}$ is a history up to and including σ_{i-1} such that $\mathcal{E}_i$ holds. Since a leaf of G_s can only be assigned to y by σ_i if $\vec{iy} \in E(\vec{H})$, and by linearity of expectation, we have

$$\begin{aligned}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] &= \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*} \mathbb{1}_{y\in N^+_{\vec{H}}(i)}\mathbb{P}\big[\sigma_i^{-1}(y)\in V(G_s)\big|\mathscr{H}_{i-1}\big]\cdot w_s(r)\\ &\le \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*} \mathbb{1}_{y\in N^+_{\vec{H}}(i)} w_s(r)w_s(i)\tfrac{4}{\mu pn}\,,\end{aligned}$$

where the second line follows by Claim 22. Summing over i, either some $\mathcal{E}_i$ with $i\in[r-1]$ does not hold, or we have

$$\begin{aligned}\sum_{i=1}^{r-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] &\le \sum_{i=1}^{r-1}\sum_{s^*-\lfloor\mu n\rfloor<s\le s^*} \mathbb{1}_{y\in N^+_{\vec{H}}(i)} w_s(r)w_s(i)\tfrac{4}{\mu pn}\\ &\le \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*}\sum_{i=1}^{n} \mathbb{1}_{y\in N^+_{\vec{H}}(i)} w_s(r)w_s(i)\tfrac{4}{\mu pn}\\ &= \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*}\sum_{v:\vec{vy}\in E(\vec{H})}^{n} w_s(r)w_s(v)\tfrac{4}{\mu pn}\\ &\le \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*}\tfrac{4w_s(r)}{\mu pn}\sum_{v:vy\in E(H)} w_s(v)\\ &\le \sum_{s^*-\lfloor\mu n\rfloor<s\le s^*}\tfrac{4w_s(r)}{\mu pn}\cdot\tfrac{10p^2n}{\mu} = \tfrac{40p}{\mu^2}\sum_{s^*-\lfloor\mu n\rfloor<s\le s^*} w_s(r)\,,\end{aligned}$$

where the last inequality is by $(P6)$. By definition of $w(r)$, by $(P2)$ and by choice of γ' we have $\sum_{s^*-\lfloor\mu n\rfloor<s\le s^*} w_s(r)=w(r)\le\frac{5}{8}pn$, so we conclude that either some $\mathcal{E}_i$ with $i\in[r-1]$ does not hold, or we have

$$\sum_{i=1}^{r-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big]\le\tfrac{40p}{\mu^2}\cdot\frac{5}{8}pn=\tfrac{25p^2n}{\mu^2}\,,$$

as desired. □

Using Claim 24, we are now in a position to apply Corollary 7, with $R=\Delta=\frac{cn}{\log n}$, with $\tilde{\mu}=25p^2n\mu^{-2}$, and with the event $\mathcal{E}=\bigcap_{i=1}^{r-1}\mathcal{E}_i$, which gives

$$\mathbb{P}\Big[\bigcap_{1\le i\le r-1}\mathcal{E}_i\quad\text{and}\quad\sum_{i=1}^{r-1}Y_i>50p^2n\mu^{-2}\Big]\le\exp\big(-\tfrac{\tilde{\mu}}{4R}\big)=\exp(-6.25c^{-1}p^2\mu^{-2}\log n)<\tfrac{1}{2}n^{-3}\,,$$

where the final inequality is by choice of c. This establishes (8). □

7. Proof of the orientation lemma

In this section we prove Lemma 19.

Proof of Lemma 19. By the given quasirandomness of H we know that $\deg_H(v)=(1\pm\gamma'^3)pn$ and $|N_H(v)\cap N_H(w)|=(1\pm\gamma'^3)p^2n$ for every $v\ne w\in V(H)$. Applying a standard Chernoff argument, i.e. using Theorem 5, we obtain that a.a.s. for every $v\ne w\in V(H)$ we have

$$\deg^+_{\vec{H}_0}(v)=(1\pm2\gamma'^3)\frac{pn}{2}\text{ and }|N^+_{\vec{H}_0}(v)\cap N^-_{\vec{H}_0}(w)|=(1\pm2\gamma'^3)\frac{p^2n}{4}\,.$$

From now on fix an arbitrary orientation $\vec{H}_0$ satisfying these two properties. Starting with $\vec{H}_0$ we aim to switch the orientations of some edges until we find an oriented graph $\vec{H}$ as desired. In order to do so, we will successively switch the orientations of pairs of edges, thus producing a sequence of oriented graphs $(\vec{H}_i)_{0\leq i\leq t}$ that eventually end up with $\vec{H}_t = \vec{H}$. For any such oriented graph $\vec{H}_i$ and every vertex $v \in V(H)$ we define the potential $\phi_i(v) := \deg^+_{\vec{H}_i}(v) - w(v)$ and

$$\phi(\vec{H}_i) := \sum_{v\in V(H)} |\phi_i(v)| \ .$$

Initially we have $|\phi_0(v)| \leq 3\gamma'^3 pn$ for every $v \in V(H)$. Note that for each i we have

$$\sum_{v\in V(H)} \phi_i(v) = e(H) - \sum_{v\in V(H)} w(v) = 0$$

where the second equality is by assumption of Lemma 19. In particular if there is a vertex with positive potential there is also a vertex with negative potential.

The algorithm *OrientationSwitch* describes how orientations are switched. In every iteration of this algorithm, the central idea is to change the potential of two vertices $x, y \in V(H)$ with $\phi_i(x) > 0$ and $\phi_i(y) < 0$ in the following way: We choose a vertex $m \in N^+_{\vec{H}_i}(x) \cap N^-_{\vec{H}_i}(y)$ uniformly at random. We then *switch* (the orientation of) the directed edge xm, that is we replace xm with mx, and we also switch the edge my. Switching these two edges creates a new orientation $\vec{H}_{i+1}$ of H. The vertex m will be called the *middle vertex* of the switching, while x and y are called the *end vertices*. In case that

$$\big|\{uv \in E(H) \colon uv \text{ is oriented differently in } \vec{H}_i \text{ and } \vec{H}_0\}\big|$$

gets too large in some round i and for some vertex v, we let the algorithm halt with failure. However, we will see in the following that this happens with probability tending to 0.

Algorithm 5: *OrientationSwitch*

let $t := \phi(\vec{H}_0)/2$;
for $i = 0$ **to** $t-1$ **do**
  if $\exists\ v$ *with* $\big|\{uv \in E(H) \colon uv$ *is oriented differently in* $\vec{H}_i$ *and* $\vec{H}_0\}\big| > 100\gamma'^3 n$
    then halt with failure;
  choose vertices $x, y \in V(H)$ with $\phi_i(x) > 0$ and $\phi_i(y) < 0$;
  choose a vertex $m \in N^+_{\vec{H}_i}(x) \cap N^-_{\vec{H}_i}(y)$ uniformly at random;
  create the new oriented graph $\vec{H}_{i+1}$ by switching the orientations of xm and my;
end
return H_t ;

We start with some easy observations.

Observation 25. *As long as the algorithm does not halt with failure we have*

$$\phi(\vec{H}_{i+1}) = \phi(\vec{H}_i) - 2\,.$$

Observation 26. *For every vertex $v \in V(H)$ with $\phi_0(v) > 0$ (or $\phi_0(v) < 0$) it holds that $\phi_{i-1}(v) \geq \phi_i(v) \geq 0$ (or $\phi_{i-1}(v) \leq \phi_i(v) \leq 0$) for all $i \in [t]$.*

Indeed, both observations hold since the switching of the orientations of xm and my ensures that $\phi_{i+1}(x) = \phi_i(x) - 1$ and $\phi_{i+1}(y) = \phi_i(y) + 1$, while the potentials of all the other vertices do not change.

Claim 27. *A.a.s. throughout the algorithm every vertex $v \in V(H)$ is chosen at most $40\gamma'^3 n$ times as the middle vertex of a switching.*

Proof. Every vertex $v \in V(H)$ can become a middle vertex only if $v \in N^+_{\vec{H}_i}(x) \cap N^-_{\vec{H}_i}(y)$ for some $0 \leq i \leq t-1$ and $x, y \in V(H)$ with $\phi_i(x) > 0$ and $\phi_i(y) < 0$. Now, v has at most $(1+\gamma'^3)pn < 2pn$ neighbours $x \in V(H)$ and every such vertex with positive potential participates in a switching as an end vertex in at most $|\phi_0(x)| \leq 3\gamma'^3 pn$ rounds. Thus, there are at most $6\gamma'^3 p^2 n^2$ rounds which may consider v as a suitable middle vertex. In each such round, the middle vertex is chosen uniformly at random from a set $N^+_{\vec{H}_i}(x) \cap N^-_{\vec{H}_i}(y)$. As long as the algorithm does not halt with failure we have

$$|N^+_{\vec{H}_i}(v) \cap N^-_{\vec{H}_i}(w)| = |N^+_{\vec{H}_0}(v) \cap N^-_{\vec{H}_0}(w)| \pm 2 \cdot 100\gamma'^3 n = (1 \pm \gamma'^2)\frac{p^2 n}{4}\ .$$

Thus, when v is suitable for being a middle vertex, the probability that v is chosen is bounded from above by $\frac{5}{p^2 n}$. Now, applying a Chernoff-type argument the claim follows. □

With the above statements in hand, we can show that a.a.s. *OrientationSwitch* does not halt with failure and that the resulting oriented graph $\vec{H} = \vec{H}_t$ satisfies the properties $(O1)$ and $(O2)$. Indeed, let $v \in V(H)$ be any vertex. In some round, we change the orientation of exactly one edge incident with v if and only if v is an end vertex of the switching in this round. As such a switching decreases $|\phi_i(v)|$ by 1 and since $|\phi_i(v)|$ never increases according to Observation 26, this happens at most $|\phi_0(v)| \leq 3\gamma' pn$ times. Moreover, we change the orientation of exactly two edges incident with v if and only if v is a middle vertex of a switching. By the above claim a.a.s. this happens at most $40\gamma'^3 n$ times. Thus, as long as the algorithm runs, we a.a.s. switch the orientations of at most

$$3\gamma' pn + 2 \cdot 40\gamma'^3 n < 100\gamma'^3 n < \gamma'^2 n$$

edges incident with v. It follows that the algorithm runs without failures, and also that property $(O2)$ holds. By Observation 25 and since $t = \phi(\vec{H}_0)/2$ we obtain that $\phi(\vec{H}_t) = 0$, meaning that $(O1)$ holds for $\vec{H} = \vec{H}_t$. □

8. Proof of the almost perfect packing lemma

In this section we prove Lemma 18. This is the technical part of this paper, which requires some stamina.

We start this section by explaining the setup which we use throughout. Then, in Section 8.1 we define some auxiliary properties that our random packing process preserves. In Section 8.2 we analyse the behaviour of the algorithm *RandomEmbedding*, and in Section 8.3 the behaviour of *PackingProcess*. In Section 8.4, finally, we use the obtained results to show Lemma 18.

In the results in this section we shall use the following setup.

Setting 28. *We use the constants defined in Setting 16.*

Let $(G''_s)_{s\in[s^*]}$ (for some $s^* \leq \frac{7}{4}n$) be graphs on $[n]$, such that for each s and $x \in V(G''_s)$ we have $\deg^-_{G''_s}(x) \leq D$, such that $\Delta(G''_s) \leq cn/\log n$, and such that the final δn vertices of G''_s all have degree d_s and form an independent set.

Let $\widehat{H}$ be a $(\xi, 2D+3)$-quasirandom graph with n vertices and density $\hat{p}$. Recall that PackingProcess *chooses H_0^* as a subgraph of $\widehat{H}$ by picking edges of $\widehat{H}$ independently with probability $\gamma\binom{n}{2}/e(\widehat{H})$. We will assume that $e(H_0^*) \leq 1.1\gamma\binom{n}{2}$.*

We note at this point that we assume $e(H_0^*) \leq 1.1\gamma\binom{n}{2}$ in order to make use of (3). This inequality holds with probability at least $1 - e^{-n}$ and hence this assumption does not affect the proof of Lemma 18, since we will see that, if this inequality holds, each of the properties $(P1)$ – $(P6)$ occurs with probability at least $1 - n^{-4}$.

8.1. **Coquasirandomness, diet, codiet, and cover.** The following properties coined in [1] are preserved throughout the run of our random packing process. Firstly, for our analysis of *PackingProcess*, we need the concept of coquasirandomness. This controls the intersections of vertex neighbourhoods in two edge disjoint graphs on the same vertex set.

Definition 29 (coquasirandom)**.** *For $\alpha > 0$ and $L \in \mathbb{N}$, we say that a pair of graphs (F, F^*), both on the same vertex set V of order n and with densities p and p^*, respectively, is (α, L)-*coquasirandom *if for every set $S \subseteq V$ of at most L vertices and every subset $R \subseteq S$ we have*

$$|N_F(R) \cap N_{F^*}(S \setminus R)| = (1 \pm \alpha)p^{|R|}(p^*)^{|S \setminus R|}n\,.$$

For the analysis of one run of *RandomEmbedding* we further need the following concepts.

Definition 30 (diet condition, codiet condition, cover condition)**.** *Let H be a graph with n vertices and $p\binom{n}{2}$ edges, and let $X \subseteq V(H)$ be any vertex set. We say that the pair (H, X) satisfies the (β, L)-*diet condition *if for every set $S \subseteq V(H)$ of at most L vertices we have*

$$|N_H(S) \setminus X| = (1 \pm \beta)p^{|S|}(n - |X|)\,.$$

*Given further H^* on the same vertex set as H, which has no edges in common with H and which has $p^*\binom{n}{2}$ edges, we say that the triple (H, H^*, X) satisfies the (β, L)-*codiet condition *if for every set $S \subseteq V(H)$ of at most L vertices, and for every $R \subseteq S$, we have*

$$\big|\big(N_H(R) \cap N_{H^*}(S \setminus R)\big) \setminus X\big| = (1 \pm \beta)p^{|R|}(p^*)^{|S \setminus R|}(n - |X|)\,.$$

Further, let G be a graph with vertex set $[n]$. Given $\varepsilon > 0$, $i \in [n - \varepsilon n]$, and $d \in \mathbb{N}$, we define

$$X_{i,d} := \{x \in V(G) \colon i \leq x < i + \varepsilon n, |N_G^-(x)| = d\}\,.$$

*We say that a partial embedding ψ of G into H, which embeds $N_G^-(x)$ for each $i \leq x < i + \varepsilon n$, satisfies the (ε, β, i)-*cover condition *if for each $v \in V(H)$, and for each $d \in \mathbb{N}$, we have*

$$\big|\big\{x \in X_{i,d} : v \in N_H\big(\psi(N_G^-(x))\big)\big\}\big| = (1 \pm \beta)p^d|X_{i,d}| \pm \varepsilon^2 n\,.$$

Following [1], we use Definition 30 to define key events $\mathsf{DietE}(\cdot;\cdot)$, $\mathsf{CoverE}(\cdot;\cdot)$, $\mathsf{CoDietE}(\cdot)$ on the probability space $\Omega^{G \hookrightarrow H}$ underlying the run of *RandomEmbedding* which attempts to embed G into H. (For a formal definition of this probability space, see [1, Section 4.1].)

Suppose that D, δ and ε are as in Setting 16. Suppose that $\lambda > 0$. Suppose that we have graphs G and H as in Algorithm 2. Suppose that we run *RandomEmbedding* to partially embed G into H. Let $(\psi_i)_{i \in [t_*]}$ be the partial embeddings of $G\big[[i]\big]$ into H, where $t_* = n - \delta n$ if *RandomEmbedding* succeeded, and otherwise $t_* + 1$ is the step in which *RandomEmbedding* halted with failure.

- For each $t \in [n - \delta n]$, let $\mathsf{DietE}(\lambda; t) \subseteq \Omega^{G \hookrightarrow H}$ correspond to executions of *RandomEmbedding* for which $t_* \geq t$ and the pair $(H, \operatorname{im}\psi_t)$ satisfies the $(\lambda, 2D+3)$-diet condition.

- For each $t \in [n-\delta n]$, let $\mathsf{CoverE}(\lambda;t) \subseteq \Omega^{G\hookrightarrow H}$ correspond to executions of *RandomEmbedding* for which $t_* \geq t+\varepsilon n$ and the embedding ψ_{t^*} of G into H satisfies the (ε,λ,t)-cover condition.
- Suppose further that we have a graph H_0^* with $V(H)=V(H_0^*)$. For each $t \in [n-\delta n]$, let $\mathsf{CoDietE}(t) \subseteq \Omega^{G\hookrightarrow H}$ correspond to executions of *RandomEmbedding* for which $t_* \geq t$ and the triple $(H, H_0^*, \operatorname{im}\psi_t)$ satisfies the $(2\eta, 2D+3)$-codiet condition.

8.2. **Properties of *RandomEmbedding*.** In this section we collect properties that are preserved during a run of *RandomEmbedding*. The constants we use are as in Setting 16. However, since we are only concerned with a single run of *RandomEmbedding* here, we only consider a single guest graph G, and a single host graph H with the following properties.

Setting 31. *Let G be a graph on vertex set $[n]$ such that $\deg^-_G(x) \leq D$ for each $x \in V(G)$ and $\Delta(G) \leq cn/\log n$. Let H be an $(\alpha, 2D+3)$-quasirandom graph with n vertices and $p\binom{n}{2}$ edges, with $p \geq \gamma$, and suppose that H_0^* is a graph on $V(H)$ such that (H, H_0^*) is $(\eta, 2D+3)$-coquasirandom.*

The following lemma comes from [1, Lemma 24] and the deduction of [1, Lemma 18] which comes immediately after. Specifically, (*a*) is the deduction of [1, Lemma 18] and (*d*) is explicitly in [1, Lemma 24], while (*b*) and (*c*) differ only from the statements of [1, Lemma 24] in that the error bound we give here is in terms of β_t whereas in [1, Lemma 24] a (larger) error bound $C\alpha$ is given. In the proof of [1, Lemma 24], the stronger error bounds we claim here are explicitly obtained. The cover conditions asserted are otherwise identical, despite being written slightly differently.

Lemma 32. *Given $D \in \mathbb{N}$ and $\gamma > 0$, let $\delta, \alpha_0, \alpha_{2n}, C, \varepsilon$ be as in Setting 16. Then the following holds for any $\alpha_0 \leq \alpha \leq \alpha_{2n}$ and all sufficiently large n. Let G, H and H_0^* be as in Setting 31. Let $\beta_t = \beta_t(\alpha)$ be as in Setting 16. If we run* RandomEmbedding *to embed $G[[n-\delta n]]$ into H, then with probability at least $1-2n^{-9}$*

(*a*) RandomEmbedding *succeeds in constructing partial embeddings $(\psi_i)_{i\in[n-\delta n]}$,*
(*b*) *$(H, \operatorname{im}\psi_t)$ satisfies the $(\beta_t, 2D+3)$-diet condition (i.e. $\mathsf{DietE}(\beta_t;t)$ occurs) for each $t \in [n-\delta n]$,*
(*c*) *ψ_t has the $(\varepsilon, 20D\beta_{t-\varepsilon n+2}, t-\varepsilon n+2)$-cover condition (i.e. $\mathsf{CoverE}(20D\beta_{t-\varepsilon n+2}, t-\varepsilon n+2)$ occurs) for each $t \in [\varepsilon n-1, n-\delta n]$.*
(*d*) *$(H, H_0^*, \operatorname{im}\psi_t)$ satisfies the $(2\eta, 2D+3)$-codiet condition (i.e. $\mathsf{CoDietE}(t)$ occurs) for each $t \in [n-\delta n]$.* □

The next lemma is proven as part of [1, Lemma 26] (it can be found in [1, Claim 26.1]).

Lemma 33. *Given $D \in \mathbb{N}$ and $\gamma > 0$, let $\delta, \alpha_0, \alpha_{2n}, C, \varepsilon$ be as in Setting 16. Then the following holds for any $\alpha_0 \leq \alpha \leq \alpha_{2n}$ and all sufficiently large n. Let G and H be as in Setting 31 and let $1 \leq j \leq t+1-\varepsilon n$ for $t \leq (1-\delta)n$. Let $\beta_j = \beta_j(\alpha)$ be as in Setting 16. Assume we run* RandomEmbedding *to embed $G[[n-\delta n]]$ into H, that it produces a partial embedding ψ_j such that $(H, \operatorname{im}\psi_j)$ has the $(\beta_j, 2D+3)$-diet condition, and let $T \subseteq V(H)\setminus\operatorname{im}\psi_j$ with $|T| \geq \frac{1}{2}\gamma^{2D+3}\delta n$. Then with probability at least $1-2n^{-2D-19}$, one of the following occurs.*

(*a*) *ψ_t does not have the $(\varepsilon, 20D\beta_j, j)$-cover condition (i.e. $\mathsf{CoverE}(20D\beta_j, j)$ does not occur) , or*
(*b*) $\big|\{x\colon j \leq x < j+\varepsilon n, \psi_{t-1}(x) \in T\}\big| = (1 \pm 40D\beta_j)\frac{|T|\varepsilon n}{n-j}$.

In [1, Lemma 28] we estimated the probability that, when running *RandomEmbedding*, a given vertex $x \in V(H)$ is not used in the embedding of the first t_1 vertices of G.

Lemma 34 (Lemma 28 in [1]). *Given $D \in \mathbb{N}$ and $\gamma > 0$, let $\delta, \alpha_0, \alpha_{2n}, C, \varepsilon$ be as in Setting 16. Then the following holds for any $\alpha_0 \le \alpha \le \alpha_{2n}$ and all sufficiently large n. Let G and H be as in Setting 31. Let $0 \le t_0 < t_1 \le n - \delta n$. Let $\mathscr{L}$ be a history ensemble of* RandomEmbedding *up to time t_0, and suppose that $\mathbb{P}[\mathscr{L}] \ge n^{-4}$. Then the following hold for any distinct vertices $u, v \in V(H)$.*

(*a*) *If $v \notin \operatorname{im} \psi_{t_0}$ then we have*
$$\mathbb{P}\big[v \notin \operatorname{im} \psi_{t_1} | \mathscr{L}\big] = (1 \pm 100C\alpha\delta^{-1}) \tfrac{n-1-t_1}{n-t_0}\,.$$

(*b*) *If $u, v \notin \operatorname{im} \psi_{t_0}$ then we have*
$$\mathbb{P}\big[u, v \notin \operatorname{im} \psi_{t_1} | \mathscr{L}\big] = (1 \pm 100C\alpha\delta^{-1}) (\tfrac{n-1-t_1}{n-t_0})^2\,. \qquad \square$$

In addition we estimated the probability that a given edge of G is embedded to a given edge of H. The following lemma is [1, Lemma 29], together with equation (6.10) of that paper which is established in the proof.

Lemma 35 (Lemma 29 in [1]). *Given $D \in \mathbb{N}$, and $\gamma > 0$, let constants $\delta, \varepsilon, C, \alpha_0, \alpha_{2n}$ be as in Setting 16. Then the following holds for any $\alpha_0 \le \alpha \le \alpha_{2n}$ and all sufficiently large n. Let G and H be as in Setting 31. Let uv be an edge of H, and let xy be an edge of G. When* RandomEmbedding *is run to embed $G[[n-\delta n]]$ into H, we have*
$$\mathbb{P}\big[x \hookrightarrow u, y \hookrightarrow v\big] = \big(1 \pm 500C\alpha\delta^{-1}\big)^{4D+2} \cdot p^{-1} n^{-2}\,,$$
and furthermore the probability that some edge of G is embedded to uv is
$$\big(1 \pm 500C\alpha\delta^{-1}\big)^{4D+2} p^{-1} n^{-2} \cdot 2e(G)\,. \qquad \square$$

We can use these two lemmas for estimating the probability that a given vertex of G is embedded on a given vertex of H.

Lemma 36 (embedding a vertex on a given vertex). *Given $D \in \mathbb{N}$, $\gamma > 0$, let δ, ε, C, α_0, α_{2n} be as in Setting 16 and let $p \ge \gamma$. Let $\alpha_0 < \alpha \le \alpha_{2n}$ and let n be sufficiently large. Let G and H be as in Setting 31. Let $x \in V(G)$ with $x \le (1-\delta)n$ and $u \in V(H)$. When we run* RandomEmbedding *to embed $G[[n-\delta n]]$ into H, then*
$$\mathbb{P}\big[x \hookrightarrow u\big] = (1 \pm 10^4 C\alpha D\delta^{-1})\frac{1}{n}\,.$$

Proof. While we could prove this lemma directly following the methods of [1], it is convenient to deduce it from the results of [1]. We separate two cases.

If x is an isolated vertex in G, then we embed x to u if and only if the first $x-1$ vertices of G are not embedded to u, and then among the $n-x+1$ vertices of H to which we could embed x, we choose u. Using Lemma 34(a) to estimate the probability of the first event occurring, with $t_0 = 0$ and $t_1 = x-1$ (and so $\mathscr{L}$ is trivial) we have
$$\begin{aligned}\mathbb{P}\big[x \hookrightarrow u\big] &= \mathbb{P}\big[u \notin \psi_{x-1}\big]\mathbb{P}\big[x \hookrightarrow u \big| u \notin \psi_{x-1}\big] = (1 \pm 100C\alpha\delta^{-1})\frac{n-1-x+1}{n} \cdot \frac{1}{n-x+1} \\ &= (1 \pm 200C\alpha\delta^{-1})\frac{1}{n}\,.\end{aligned}$$

If, on the other hand, there is y such that $xy \in E(H)$, then we embed x to u if and only if we embed x to u and y to some neighbour v of u in H. Since these events are disjoint as v ranges over the neighbours of u, the probability that one of them occurs is exactly the sum of their individual probabilities, and the latter are estimated by Lemma 35. Since by the $(\alpha, 2D+3)$-quasirandomness of H, the vertex u has $(1\pm\alpha)pn$ neighbours, we obtain

$$\begin{aligned}\mathbb{P}\big[x \hookrightarrow u\big] &= \sum_{v\in N_H(u)} \mathbb{P}\big[x \hookrightarrow u, y \hookrightarrow v\big] = (1\pm\alpha)pn\cdot(1\pm 500C\alpha\delta^{-1})^{4D+2}\frac{1}{pn^2}\\ &= (1\pm 10^4 C\alpha D\delta^{-1})\frac{1}{n}\,.\end{aligned}$$

In either case, we conclude the desired bound. □

We further need the following lemma, estimating the probability that a given vertex of G is embedded to a given vertex of H and another given vertex of H is not used in the embedding of the first $n-\lfloor\mu n\rfloor$ vertices of G. We will be interested in this when G is a special graph; so the remaining vertices of G (which *RandomEmbedding* also embeds) are isolated vertices. The proof of this lemma is rather similar to the proof of [1, Lemma 29].

Lemma 37 (embedding a vertex on a given vertex and not using another vertex)**.** *Given $D\in\mathbb{N}$, $\gamma>0$, let δ, ε, C, α_0, α_{2n} be as in Setting 16 and let $p\ge\gamma$. Let $\alpha_0<\alpha\le\alpha_{2n}$ and let n be sufficiently large. Let G and H be as in Setting 31. Let $x\in V(G[[n-\mu n]])$ and $u,v\in V(H)$ with $u\neq v$. When we run* RandomEmbedding *to construct an embedding $\psi_{n-\lfloor\mu n\rfloor}$ of the first $n-\lfloor\mu n\rfloor$ vertices of G into H, then*

$$\mathbb{P}\big[x\hookrightarrow v \text{ and } u\notin \operatorname{im}\psi_{n-\lfloor\mu n\rfloor}\big] = (1\pm 10^3 C\alpha D\delta^{-1})\frac{\mu}{n}\,.$$

Proof. Let $y_1,\dots,y_d$ with $d\le D$ be the neighbours of x in $N^-_G(x)$ in degeneracy order, and (for convenience) define $y_0=0$. We define a collection of events. Let $\mathscr{L}'_0$ be the almost sure event. For each $1\le i\le d$, let $\mathscr{L}_i$ be the intersection of $\mathscr{L}'_{i-1}$ and the event that neither u nor v is in the image of ψ_{y_i-1}, and let $\mathscr{L}'_i$ be the intersection of $\mathscr{L}_i$ and the event that y_i is embedded to a vertex of $N_H(v)\setminus\{u\}$. Let $\mathscr{L}_{d+1}$ be the intersection of $\mathscr{L}'_d$ and the event that neither u nor v is in the image of ψ_{x-1}. Let $\mathscr{L}'_{d+1}$ be the intersection of $\mathscr{L}_{d+1}$ and the event $x\hookrightarrow v$. And finally let $\mathscr{L}_{d+2}$ be the intersection of $\mathscr{L}'_{d+1}$ and the event that $u\notin\operatorname{im}\psi_{n-\lfloor\mu n\rfloor}$. Note that all of these events are history ensembles up to some given time.

Now what we want to do is estimate $\mathbb{P}[\mathscr{L}_{d+2}]$, and the reason for giving this collection of events is that we can estimate each of the successive conditional probabilities. We can estimate $\mathbb{P}[\mathscr{L}_i|\mathscr{L}'_{i-1}]$ for each $1\le i\le d+2$ using Lemma 34 (using part (b) for $1\le i\le d+1$ and part (a) for the final part). And we can estimate $\mathbb{P}[\mathscr{L}'_i|\mathscr{L}_i]$ using the diet condition for each $1\le i\le d+1$; the probability that the diet condition fails is tiny. To justify both of these steps we need to know $\mathbb{P}[\mathscr{L}_i],\mathbb{P}[\mathscr{L}'_i]>n^{-4}$; this is (by induction) valid since the final $\mathscr{L}_{d+2}$ is the smallest event and we will argue its probability satisfies this bound. Assuming this bound for a moment, by Lemma 34, for each $1\le i\le d$ we have

$$\mathbb{P}[\mathscr{L}_i|\mathscr{L}'_{i-1}] = (1\pm 100C\alpha\delta^{-1})\big(\tfrac{n-y_i}{n-y_{i-1}}\big)^2\,,\quad \mathbb{P}[\mathscr{L}_{d+1}|\mathscr{L}'_d] = (1\pm 100C\alpha\delta^{-1})\big(\tfrac{n-x}{n-y_d}\big)^2$$
$$\text{and}\quad \mathbb{P}[\mathscr{L}_{d+2}|\mathscr{L}'_{d+1}] = (1\pm 100C\alpha\delta^{-1})\tfrac{\lfloor\mu n\rfloor-1}{n-x}\,.$$

For each $1\le i\le d$, we have

$$\mathbb{P}[\mathscr{L}'_i|\mathscr{L}_i] = \frac{(1\pm C\alpha)p^{\deg^-_G(y_i)+1}(n-y_i+1)\pm 1}{(1\pm C\alpha)p^{\deg^-_G(y_i)}(n-y_i+1)} \pm 4n^{-5} = (1\pm 4C\alpha)p\,.$$

The fraction in the first term assumes the $(C\alpha, 2D+3)$-diet condition, for the vertices $\psi_{y_i-1}\big(N_G^-(y_i)\big) \cup \{v\}$ in the numerator and $\psi_{y_i-1}\big(N_G^-(y_i)\big)$ in the denominator, to estimate respectively the number of neighbours of v in the candidate set of y_i which are not in $\operatorname{im}\psi_{y_i-1}$ and the number of vertices in the candidate set of y_i which are not covered by $\operatorname{im}\psi_{y_i-1}$. The ± 1 term in the numerator covers the possibility $u \in N_H(v)$. The $4n^{-5}$ error term covers the possibility of failure of the diet condition: By Lemma 32 the probability that the diet condition fails is at most $2n^{-9}$, hence since $\mathbb{P}[\mathscr{L}_i] > n^{-4}$ the probability that the diet condition fails conditioned on $\mathscr{L}_i$ is at most $2n^{-5}$. By similar logic, we have

$$\mathbb{P}[\mathscr{L}'_{d+1}|\mathscr{L}_{d+1}] = \frac{1}{(1 \pm C\alpha)p^d(n-x+1)} \pm 4n^{-5} = (1 \pm 4C\alpha)\tfrac{1}{p^d(n-x+1)}\,.$$

Multiplying together all these conditional probabilities, many terms cancel and we obtain

$$\begin{aligned}\mathbb{P}[\mathscr{L}_{d+2}] &= (1 \pm 100C\alpha\delta^{-1})^{d+2}(1 \pm 4C\alpha)^{d+1}\frac{(n-x)(\lfloor \mu n\rfloor - 1)}{n^2(n-x+1)} \\ &= (1 \pm 100C\alpha\delta^{-1})^{2D+4} \cdot \frac{\mu}{n}\,,\end{aligned}$$

which since $\alpha \le \alpha_{2n}$ and by choice of α_{2n} implies the desired bound. □

8.3. **Properties of *PackingProcess*.** The following lemma summarises some facts we obtain in the course of proving [1, Theorem 11].

Lemma 38 (*PackingProcess* lemma)**.** *Given D, $\hat{p}$, γ, let $(\alpha_s)_{s\in[s^*]}$, η and the graphs $(G''_s)_{s\in[s^*]}$, $\widehat{H}$ be as in Setting 28. When* PackingProcess *is run with input $(G''_s)_{s\in[s^*]}$ and $\widehat{H}$, with probability at least $1-2n^{-5}$, the following holds.*

(*a*) PackingProcess *succeeds in packing $(G''_s)_{s\in[s^*]}$ into $\widehat{H}$.*
(*b*) *For each $s \in [s^*]$ the pair (H_s, H_0^*) is $(\alpha_s, 2D+3)$-coquasirandom.*
(*c*) *The leftover graph H is $(\eta, 2D+3)$-quasirandom.*
(*d*) *H_0^* has maximum degree at most $2\gamma n$.*

Proof. (*a*) is obtained by summing the failure probabilities of all exceptional events in [1, Proof of Theorem 11].

(*b*) holding is implied by the exceptional event (ii) of that proof not occurring.

(*c*) is implied by exceptional event (v) of [1, Proof of Theorem 11] not occurring. Again, event (v) not occurring states that $(H_{s^*}, H^*_{s^*})$ is $(\eta, 2D+3)$-quasirandom. We would like to know that this implies $H = H_{s^*} \cup H^*_{s^*}$ is $(\eta, 2D+3)$-quasirandom. Since H_{s^*} and $H^*_{s^*}$ are edge disjoint, given any vertex set S of size at most $2D+3$, the neighbours $N_H(S)$ are partitioned into parts indexed by the subsets R of S, where a vertex v is in the part indexed by R if it is adjacent in H_{s^*} to the vertices R and in $H^*_{s^*}$ to the vertices $S \setminus R$. Now $(\eta, 2D+3)$-coquasirandomness gives bounds on these part sizes with a $(1\pm\eta)$ relative error, and summing the bounds we obtain the desired $(\eta, 2D+3)$-quasirandomness of H. Indeed, by the argument above we obtain

$$\begin{aligned}N_H(S) &= \sum_{R\subseteq S}(1\pm\eta)(p_{s*})^{|R|}(p^*_{s*})^{|S\setminus R|}n \\ &= (1\pm\eta)n\sum_{r=0}^{|S|}\binom{s}{r}(p_{s*})^r(p^*_{s*})^{|S|-r} = (1\pm\eta)(p_{s*}+p^*_{s*})^{|S|}n\end{aligned}$$

for every S of size at most $2D+3$.

(d) is implied by exceptional event (i) not occurring: this event in particular implies that H_0^* is $(\frac{1}{4}\alpha_0, 2D+3)$-quasirandom, which together with the fact $e(H_0^*) = (1 \pm \alpha_0)\gamma\binom{n}{2}$ from [1, Lemma 16] implies the claimed maximum degree. □

We further need the following two lemmas. The first states that, while running *Packing-Process*, chosen subsets T of neighbourhoods of vertices shrink roughly as expected. We will use this with T being a vertex neighbourhood with the embedded image of one or two of the G_i'' removed. Recall that p_s denotes the density of H_s.

Lemma 39. *Assume Setting 28 and let $s^* - \lfloor \mu n \rfloor < s < s' \le s^*$. Consider the following experiment. Run* PackingProcess *with input $(G''_{s''})_{s'' \in [s^*]}$ and $\widehat{H}$ up to and including the embedding of G''_s. Then fix $T \subseteq N_{H_s}(v)$ with $|T| \ge \frac{1}{2}p\mu^2 n$, and continue* PackingProcess *to perform the embedding of $G''_{s+1}, \dots, G''_{s'}$.*

The probability that PackingProcess *fails before embedding $G''_{s'}$, or H_i fails to be $(\alpha_i, 2D{+}3)$-quasirandom for some $1 \le i \le s'$, or we have*

$$\big|T \cap N_{H_{s'}}(v)\big| = (1 \pm \gamma^{-1}\alpha_{s'})\frac{p_{s'}}{p_s}|T|\,,$$

is at least $1 - n^{-C}$.

Proof. For $s \le i \le s'$, we define the event $\mathcal{E}_i$ that *PackingProcess* does not fail before embedding G_i'', and H_j is $(\alpha_j, 2D+3)$-quasirandom for each $1 \le j \le i$, and $|T \cap N_{H_j}(v)| = (1 \pm \gamma^{-1}\alpha_j)\frac{p_j}{p_s}|T|$ for each $s \le j \le i$. If the event in the lemma statement fails to occur, then there must exist some $s \le i < s'$ such that $\mathcal{E}_i$ occurs and

$$|T \cap N_{H_{i+1}}(v)| \neq (1 \pm \gamma^{-1}\alpha_{i+1})\frac{p_{i+1}}{p_s}|T|\,.$$

It suffices to show that each of these bad events occurs with probability at most n^{-C-1}, since then the union bound over the at most μn choices of i gives the lemma statement. This is an estimate we can obtain using Corollary 7. We now fix $s \le i < s'$ and prove the desired estimate.

Suppose $s \le j \le i$, and let $Y_j := |N_{H_j}(v) \cap T \setminus N_{H_{j+1}}(v)|$ be the number of edges from v to T used for the embedding G_{j+1}. Then we have $\big|T \cap N_{H_{i+1}}(v)\big| = |T| - \sum_{j=s}^{i} Y_j$, and what we want to do is argue that the sum of random variables is concentrated. To that end, suppose $\mathscr{H}$ is a history of *PackingProcess* up to time j such that H_j is $(\alpha_j, 2D+3)$-quasirandom and $|T \cap N_{H_j}(v)| = (1 \pm \gamma^{-1}\alpha_j)\frac{p_j}{p_s}|T|$. Then we have

$$\mathbb{E}\big[Y_j \,\big|\, \mathscr{H}\big] = \frac{2e(G''_{j+1}) \cdot (1 \pm 500C\alpha_j\delta^{-1})^{4D+2}}{p_j n^2} \cdot (1 \pm \gamma^{-1}\alpha_j)\frac{p_j}{p_s}|T|$$

where we use linearity of expectation: the first factor is by Lemma 35 the probability that a given edge from v to T in H_j is used in the embedding of G_j'', and the second factor is the number of such edges. Note that the p_j terms cancel, so we obtain

$$\begin{aligned}\mathbb{E}\big[Y_j \,\big|\, \mathscr{H}\big] &= \frac{2e(G''_{j+1}) \cdot (1 \pm 500C\alpha_j\delta^{-1})^{4D+2}}{p_s n^2} \cdot (1 \pm \gamma^{-1}\alpha_j)|T| \\ &= \frac{2e(G''_{j+1})|T|}{p_s n(n-1)} \pm \frac{10^5\delta^{-1}CD^2|T|}{p_s n}\alpha_j\,,\end{aligned}$$

where for the error term we use the upper bound $e(G''_{j+1}) \le Dn$ and our choice $\delta^{-1} > \gamma^{-1}$. Let

$$\tilde{\mu} := \sum_{j=s}^{i} \frac{2e(G''_{j+1})|T|}{p_s n(n-1)} \quad \text{and} \quad \tilde{\nu} := \sum_{j=s}^{i} \frac{10^5\delta^{-1}CD^2|T|}{p_s n}\alpha_j$$

and observe that $\tilde{\mu} \le |T| \le n$ and $\tilde{\nu} \le \frac{10^5\delta^{-1}CD^2|T|}{p_s}\alpha_i < \frac{|T|}{10^3}$ since $p_s \ge \gamma$ and by the definition of α_j.

We trivially have $0 \le Y_j \le \Delta(G''_{j+1}) \le cn/\log n$. So what Corollary 7($b$), with $\tilde{\varrho} = \varepsilon n$, gives us is that

$$\mathbb{P}\left[\mathcal{E}_i \text{ and } \sum_{j=s}^{i} Y_i \ne \tilde{\mu} \pm (\tilde{\nu} + \varepsilon n)\right] < 2\exp\left(-\tfrac{\varepsilon^2 n^2}{4cn^2/\log n}\right) < n^{-C-1},$$

where we use the upper bound $\tilde{\mu} + \tilde{\nu} + \tilde{\varrho} \le 2n$ for the first inequality and the choice of c as well as $\varepsilon < \frac{1}{C}$ for the second. This is the probability bound we wanted. We now simply need to show that if

$$\sum_{j=s}^{i} Y_i = \tilde{\mu} \pm (\tilde{\nu} + \varepsilon n)$$

then we have

$$|T \cap N_{H_{i+1}}(v)| = (1 \pm \gamma^{-1}\alpha_{i+1})\frac{p_{i+1}}{p_s}|T|.$$

Since

$$|T| - \tilde{\mu} = |T|\Big(1 - \tfrac{\sum_{j=s}^{i} e(G''_{j+1})}{p_s\binom{n}{2}}\Big) = |T|\Big(1 - \tfrac{(p_s - p_{i+1})\binom{n}{2}}{p_s\binom{n}{2}}\Big) = \frac{p_{i+1}}{p_s}|T|,$$

what remains is to argue $\tilde{\nu} + \varepsilon n < \gamma^{-1}\alpha_{i+1}\frac{p_{i+1}}{p_s}|T|$. Since $\alpha_j = \frac{\delta}{10^8 CD}\exp\left(\frac{10^8CD^3\delta^{-1}(j-2n)}{n}\right)$ is increasing in j, we have

$$\begin{aligned}\sum_{j=s}^{i}\alpha_j &\le \int_s^{i+1}\alpha_j\,\mathrm{d}j \le \int_{-\infty}^{i+1}\alpha_j\,\mathrm{d}j \\ &= \Big[\frac{\delta}{10^8CD}\cdot\frac{n}{10^8CD^3\delta^{-1}}\cdot\exp\Big(\frac{10^8CD^3\delta^{-1}(j-2n)}{n}\Big)\Big]_{j=-\infty}^{i+1} = \frac{\delta n}{10^8CD^3}\alpha_{i+1}.\end{aligned} \tag{9}$$

It follows that

$$\tilde{\nu} + \varepsilon n \le \frac{10^5\delta^{-1}CD^2|T|}{p_s n}\cdot\frac{\delta n}{10^8CD^3}\alpha_{i+1} + \varepsilon n \le \tfrac{\alpha_{i+1}}{1000D}\cdot\tfrac{1}{p_s}|T| + \varepsilon n.$$

Finally, since $p_{i+1}, p \ge \gamma$, by choice of ε, since $\delta \le \mu$ and because $|T| \ge \frac{1}{2}p\mu^2 n$, we conclude $\tilde{\nu} + \varepsilon n \le \gamma^{-1}\alpha_{i+1}\frac{p_{i+1}}{p_s}|T|$ as desired. □

The second lemma states that for a set S of host graph vertices fixed before the embedding of G''_s, it is likely that the embedding of G''_s (which has $n - \lfloor\mu n\rfloor$ vertices) uses about $(1-\mu)|S|$ vertices of S. To prove it, we repeatedly apply Lemma 33, which tells us that it is likely that each successive εn vertices of G''_s embedded cover about the expected fraction of S.

Lemma 40. *Assume Setting 28 and let $s^* - \lfloor\mu n\rfloor < s \le s^*$. Run* PackingProcess *with input $(G''_{s''})_{s''\in[s^*]}$ and $\widehat{H}$ up to just before the embedding of G''_s. Then fix any $S \subseteq V(H_{s-1})$ with*

$|S| \geq \frac{1}{2}p\mu^2 n$, *and let* PackingProcess *perform the embedding of* G''_s. *With probability at least* $1-3n^{-9}$ *either* H_{s-1} *is not* $(\alpha_{s-1}, 2D+3)$*-quasirandom or*

$$|S \setminus \operatorname{im}\phi'_s| = (1 \pm C'\alpha_s)\mu|S|\,.$$

Proof. Fix s such that $s^* - \lfloor \mu n \rfloor < s \leq s^*$, and condition on H_{s-1}. If H_{s-1} is not $(\alpha_{s-1}, 2D+3)$-quasirandom, then the bad event of this lemma cannot occur. So it suffices to show that if H_{s-1} is $(\alpha_{s-1}, 2D+3)$-quasirandom, then the probability of the event $|S \setminus \operatorname{im}\phi'_s| \neq (1 \pm C'\alpha_s)\mu|S|$, conditioned on H_{s-1}, is at most $3n^{-9}$. This is what we will now do, so we suppose that H_{s-1} is $(\alpha_{s-1}, 2D+3)$-quasirandom. Consider the run of *RandomEmbedding* which embeds $G''_s[[n-\delta n]]$.

Recall that the embedding ϕ'_s of G'_s is given by letting *RandomEmbedding* perform the embedding of $G''_s[[n-\delta n]]$, constructing the partial embeddings ψ_t for $0 \leq t \leq (1-\delta)n$. More precisely, ϕ'_s is given by ignoring the embedding of all vertices not in G'_s, that is, by $\psi_{n-\lfloor \mu n \rfloor}$.

Define $S_0 = S$, and for $i = 1, \ldots, \tau$ with $\tau = \lceil \frac{(1-\mu)}{\varepsilon} \rceil$ set $S_i = S_{i-1} \setminus \operatorname{im}\psi_{i\varepsilon n}$. Since $S_\tau \subseteq S \setminus \operatorname{im}\phi'_s \subseteq S_{\tau-1}$, it is enough to show both $|S_{\tau-1}|$ and $|S_\tau|$ are likely to be in the claimed range. Since the two quantities differ by at most εn, we will focus on estimating $|S_\tau|$. In this proof we will always use $\alpha = \alpha_s$, and hence will often omit the parameter α in $\beta_t = \beta_t(\alpha)$. By Lemma 33 (applied with $t = j + \varepsilon n + 1$), with probability at least $1 - n^{-2D-18}$, either for some $j \leq n - \mu n - \varepsilon n$

(*a*) *RandomEmbedding* failed to construct ψ_j, or

(*b*) the partial embedding $\psi_{j+\varepsilon n+1}$ of $G''_s[[j+\varepsilon n+1]]$ into H_{s-1} does not have the $(\varepsilon, 20D\beta_j, j)$-cover condition,

or we have that for every $j \leq n - \mu n - \varepsilon n$

(*c*) $|\{x :\ j \leq x < j + \varepsilon n :\ \psi_{j+\varepsilon n}(x) \in S \setminus \operatorname{im}\psi_j\}| = (1 \pm 40D\beta_j)\frac{|S\setminus \operatorname{im}\psi_j|\varepsilon n}{n-j}$.

By Lemma 32, with probability at least $1 - 2n^{-9}$, the first two options do not hold, so with probability at least $1 - 3n^{-9}$ we have that (*c*) holds for every $j \leq n - \mu n - \varepsilon n$. Applying (*c*) with $j = (i-1)\varepsilon n$ we conclude

$$|S_i| = |S_{i-1}| - (1 \pm 40D\beta_{(i-1)\varepsilon n})\frac{|S_{i-1}|\varepsilon n}{n-(i-1)\varepsilon n}$$

for all $i \geq 1$.

Assuming this is the case, we get

$$|S_i| = |S_{i-1}|\Big(1 - \frac{(1 \pm 40D\beta_{(i-1)\varepsilon n})\varepsilon}{1-(i-1)\varepsilon}\Big)\,,$$

and hence

$$|S_\tau| = |S|\prod_{i=1}^{\tau}\Big(1 - \frac{(1 \pm 40D\beta_{(i-1)\varepsilon n})\varepsilon}{1-(i-1)\varepsilon}\Big)\,.$$

In order to evaluate this product, observe that

$$1 - \frac{(1 \pm 40D\beta_{i\varepsilon n})\varepsilon}{1-i\varepsilon} = \frac{1-(i+1)\varepsilon}{1-i\varepsilon} \pm \frac{40D\beta_{i\varepsilon n}\varepsilon}{1-i\varepsilon} = \frac{1-i\varepsilon-\varepsilon}{1-i\varepsilon}\Big(1 \pm \frac{40D\beta_{i\varepsilon n}\varepsilon}{1-(i+1)\varepsilon}\Big)\,,$$

and therefore

$$|S_\tau| = |S|\prod_{i=0}^{\tau-1}\frac{1-i\varepsilon-\varepsilon}{1-i\varepsilon}\cdot\Big(1 \pm \frac{40D\beta_{i\varepsilon n}\varepsilon}{1-(i+1)\varepsilon}\Big) = |S|(1-\tau\varepsilon)\prod_{i=0}^{\tau-1}\Big(1 \pm \frac{40D\beta_{i\varepsilon n}\varepsilon}{1-(i+1)\varepsilon}\Big)\,.$$

By the definition of τ we have $\frac{(1-\mu)}{\varepsilon} \leq \tau \leq \frac{(1-\mu)}{\varepsilon} + 1$ and hence $(1-\tau\varepsilon) = \mu(1 \pm \frac{\varepsilon}{\mu})$. Moreover, we obtain that

$$\begin{aligned}
\sum_{i=0}^{\tau-1} \frac{40D\beta_{i\varepsilon n}\varepsilon}{1-(i+1)\varepsilon} &\leq \frac{40D\varepsilon}{1-\tau\varepsilon}\sum_{i=0}^{\tau-1}\beta_{i\varepsilon n} \leq \frac{80D\varepsilon}{\mu}\sum_{i=0}^{\tau-1}\beta_{i\varepsilon n} \\
&\leq \frac{80D}{\mu n}\int_0^{\tau} \varepsilon n \beta_{i\varepsilon n}\,\mathrm{d}i \leq \frac{80D}{\mu n}\int_0^{\tau\varepsilon n}\beta_x\,\mathrm{d}x \\
&\overset{(6)}{\leq} \frac{80D}{\mu \cdot 1000D\delta^{-2}\gamma^{-2D-10}}\beta_{\tau\varepsilon n} \leq \beta_{(1-\mu+\varepsilon)n} \leq \frac{1}{2}C'\alpha = \frac{1}{2}C'\alpha_s
\end{aligned}$$

since $\beta_{(1-\mu+\varepsilon)n} = \beta_{(1-\mu+\varepsilon)n}(\alpha) = 2\alpha\exp(1000D\delta^{-2}\gamma^{-2D-10}(1-\mu+\varepsilon)) < 2\alpha$ and

$$C' = 10^4 \cdot \frac{40D}{\delta}\exp(1000D\delta^{-2}\gamma^{-2D-10})\,.$$

So, since $\prod_i(1 \pm x_i) = 1 \pm 2\sum_i x_i$ as long as $\sum_i x_i < \frac{1}{100}$ and since $\frac{1}{2}C'\alpha_s < \frac{1}{100}$, we get

$$\begin{aligned}
|S_\tau| &= |S|(1-\tau\varepsilon)\Big(1 \pm 2\sum_{i=0}^{\tau-1}\frac{40D\beta_{i\varepsilon n}\varepsilon}{1-(i+1)\varepsilon}\Big) = |S|(1-\tau\varepsilon)\Big(1 \pm \frac{80D\varepsilon}{1-\tau\varepsilon}\sum_{i=0}^{\tau-1}\beta_{i\varepsilon n}\Big) \\
&= |S|\Big(1-\tau\varepsilon \pm 80D\varepsilon\sum_{i=0}^{\tau-1}\beta_{i\varepsilon n}\Big) = |S|\mu\Big(1 \pm \frac{\varepsilon}{\mu} \pm \frac{80D\varepsilon}{\mu}\sum_{i=0}^{\tau-1}\beta_{i\varepsilon n}\Big) \\
&= |S|\mu\Big(1 \pm \frac{1}{2}\alpha_s \pm \frac{1}{2}C'\alpha_s\Big)\,,
\end{aligned}$$

where for the last equation we use that $\varepsilon \leq \alpha_0\delta^2\gamma \leq \frac{1}{2}\alpha_s\mu$. It follows that

$$|S \setminus \operatorname{im}\phi'_s| = |S_\tau| \pm \varepsilon n = |S|\mu\big(1 \pm \tfrac{1}{2}\alpha_s \pm \tfrac{1}{2}C'\alpha_s\big) \pm \varepsilon n = \big(1 \pm C'\alpha_s\big)\mu|S|\,,$$

as desired. □

8.4. **Proof of Lemma 18.** We now have all tools at hand to prove the almost perfect packing lemma.

Proof of Lemma 18. For $0 \leq s < s^*$ let $\mathcal{E}_s$ be the event that H_s is $(\alpha_s, 2D+3)$-quasirandom. By Lemma 38(*b*) we have

$$\mathbb{P}\Big[\bigcap_s \mathcal{E}_s\Big] \geq 1 - 2n^{-5}\,. \tag{10}$$

Let $\mathscr{H}_s$ be an embedding of $G''_1, \ldots, G''_s$ by *PackingProcess* such that $\mathcal{E}_s$ holds.

Recall that we may assume that $e(H^*_0) \leq 1.1\gamma\binom{n}{2}$ holds, which is fine as the probability of this inequality not being satisfied is at most e^{-n}. So, from now on we always condition on this assumption, and we shall show in the following that then each of the properties $(P1)$–$(P6)$ holds with probability at least $1 - n^{-4}$, which gives the lemma.

$(P1)$: H is $(\gamma'^3, 2D+3)$-quasirandom and has density p.

By Lemma 38(*c*), the leftover graph H is $(\eta, 2D+3)$-quasirandom with probability at least $1-2n^{-5}$. By (4) and since $\gamma \ll \gamma'$ we have $\eta \leq \gamma'^3$, which gives $(P1)$.

$(P2)$: $w(v) = (1 \pm \gamma'^3)\frac{pn}{2}$.

Fix $v \in V(H)$ and let $Y_s = w_s(x)\mathbb{1}_{x\hookrightarrow v}$. We have $Y_s \le \Delta$ and

$$w(v) = \sum_s w_s(v) = \sum_{s, x\in V(G_s)} Y_s\,.$$

We want to apply Corollary 7. By Lemma 36 we have

$$\sum_{s\in[s^*]} \mathbb{E}[Y_s|\mathscr{H}_{s-1}] = \sum_{\substack{s\in[s^*]\\ x\in V(G_s)}} w_s(x)\mathbb{P}[x\hookrightarrow v|\mathscr{H}_{s-1}] = \sum_{\substack{s\in[s^*]\\ x\in V(G_s)}} w_s(x)(1\pm 10^4 C\alpha_s D\delta^{-1})\frac{1}{n}\,.$$

It follows that

$$\sum_{s\in[s^*]} \mathbb{E}[Y_s|\mathscr{H}_{s-1}] = p\binom{n}{2}(1\pm 10^4 C\alpha_{s^*} D\delta^{-1})\frac{1}{n} = \frac{pn}{2}(1\pm 2\cdot 10^4 C\alpha_{s^*} D\delta^{-1})\,.$$

By the second part of Corollary 7(*b*) applied with $\mathcal{E} = \bigcap_s \mathcal{E}_s$, $\tilde\mu = \frac{pn}{2}$, $\tilde\eta = 2\cdot 10^4 C\alpha_{s^*} D\delta^{-1}$ we obtain

$$\begin{aligned}\mathbb{P}\big[\mathcal{E} \text{ and } \sum_s Y_s \ne \frac{pn}{2}\cdot(1\pm 4\cdot 10^4 C\alpha_{s^*}D\delta^{-1})\big] &\le 2\exp\Big(-\frac{\tilde\mu\cdot 4\cdot 10^8 C^2\alpha_{s^*}^2 D^2\delta^{-2}}{4\Delta}\Big)\\ &\le 2\exp(-10^{10}\log n)\,,\end{aligned} \tag{11}$$

where the last inequality uses $\Delta \le cn/\log n$, $c \le 10^{-10}\gamma^{10D}\alpha_0^4$, $\alpha_0 \le \alpha_{s^*}$, $p \ge \mu\nu$ and $\gamma \ll \nu \le \mu$.

We have $s^* \le \frac{7}{4}n$ and hence by the definition of α_x and of C in (4) we get

$$\begin{aligned}4\cdot 10^4 C\alpha_{s^*}D\delta^{-1} &\le 4\cdot 10^4 C\alpha_{\frac{7}{4}n} D\delta^{-1}\\ &= 4\cdot 10^4 C\cdot\frac{\delta}{10^8 CD}\exp\big(-10^8 CD^3\delta^{-1}\cdot\tfrac14\big)\cdot D\delta^{-1}\\ &\le \exp\big(-10^7 CD^3\delta^{-1}\big) \le \exp\big(-10^7\cdot 40D\exp(1000D\delta^{-2}\gamma^{-2D-10})\big)\\ &\le \exp\big(-\exp(\gamma^{-2D-10})\big) \le \gamma^3 \le {\gamma'}^3\,.\end{aligned} \tag{12}$$

Combining this with (11) and (10) and a union bound over v, we conclude that $(P2)$ fails with probability at most $2n^{-5} + n\cdot n^{-10} \le n^{-4}$.

<u>$(P3)$: $\big|N_H(v)\setminus \operatorname{im}\phi'_s\big| = (1\pm{\gamma'}^3)\mu pn$ and</u>
<u>$(P4)$: $\big|N_H(v)\setminus(\operatorname{im}\phi'_s\cup\operatorname{im}\phi'_{s'})\big| = (1\pm{\gamma'}^3)\mu^2 pn$ if $s\ne s'$.</u>

We prove these together. Fix $v\in V(H)$ and s, s' with $s^* - \lfloor \mu n\rfloor < s < s' \le s^*+1$. The artificial case $s' = s^*+1$ will be used to prove $(P3)$.

We first run *PackingProcess* up to time $s-1$ and consider the embedding of G''_s. We want to apply Lemma 40 to estimate what happens in this first stage. We set $S = N_{H_{s-1}}(v)$, so if $\mathcal{E}_{s-1}$ holds then $|S| = (1\pm\alpha_{s-1})p_{s-1}n \ge \frac12 pn \ge \frac12 p\mu^2 n$. Hence we can apply Lemma 40 with S and conclude that with probability at least $1-3n^{-9}$ either $\mathcal{E}_{s-1}$ does not hold or

$$|N_{H_{s-1}}(v)\setminus \operatorname{im}\phi'_s| = (1\pm C'\alpha_s)\mu|S| = (1\pm 3C'\alpha_s)p_{s-1}\mu n\,. \tag{13}$$

Further, we have $N_{H_s}(v)\setminus\operatorname{im}\phi'_s = N_{H_{s-1}}(v)\setminus\operatorname{im}\phi'_s$.

Now let *PackingProcess* perform the embeddings of $G''_{s+1},\dots,G''_{s'-1}$. We want to apply Lemma 39 to estimate what happens in this second stage. Set $T = N_{H_s}(v)\setminus\operatorname{im}\phi'_s$ and observe that $T\cap N_{H_{s'-1}}(v) = N_{H_{s'-1}}(v)\setminus\operatorname{im}\phi'_s$. If (13) holds, then $|T| \ge \frac12 p\mu^2 n$ because by (4) we have $C'\alpha_s \leq 10^{-4}$. So by Lemma 39 applied with T we get that with probability at least $1 - n^{-C}$ either $\bigcap_i \mathcal{E}_i$ fails or we have

(14)
$$\begin{aligned} |N_{H_{s'-1}}(v) \setminus \operatorname{im} \phi'_s| &= (1 \pm \gamma^{-1}\alpha_{s'-1})\frac{p_{s'-1}}{p_s}|T| \overset{(13)}{=} (1 \pm \gamma^{-1}\alpha_{s'-1})\frac{p_{s'-1}}{p_s}(1 \pm 3C'\alpha_s)p_{s-1}\mu n \\ &= (1 \pm 5C'\alpha_{s'-1})\mu p_{s'-1} n \end{aligned}$$

where the last equality follows from $\frac{p_{s-1}}{p_s} = 1 + o(1)$ and since $\gamma^{-1} < C'$. For the case $s' = s^* + 1$ this immediately implies $(P3)$. Indeed, in this case (14) gets

$$|N_{H_{s^*}}(v) \setminus \operatorname{im} \phi'_s| = (1 \pm 5C'\alpha_{s^*})\mu p_{s^*} n \ .$$

As long as $\Delta(H^*_{s^*}) \leq \Delta(H^*_0) \leq 2\gamma n$, which holds with probability at least $1 - 2n^{-5}$ according to Lemma 38(d), we have that $|N_H(v)| - |N_{H_{s^*}}(v)| \leq 2\gamma n$ and $p_{s^*} = p \pm 2\gamma = \left(1 \pm \frac{2\gamma}{p}\right)p$ from which we conclude that

$$|N_H(v) \setminus \operatorname{im} \phi'_s| = (1 \pm 5C'\alpha_{s^*})\left(1 \pm \frac{2\gamma}{p}\right)\mu p n \pm 2\gamma n = (1 \pm {\gamma'}^3)\mu p n$$

since $C'\alpha_{s^*} < \frac{1}{100}{\gamma'}^3$, since $\gamma \ll \gamma' \ll \nu \ll \mu$ and $p \geq \mu\nu$. Hence, in total, taking a union bound over v and s and using (10), the probability that $(P3)$ fails is at most $4n^{-5} + n^2(3n^{-9} + n^{-C}) \leq n^{-4}$.

For proving $(P4)$, assume that $s' \leq s^*$ and consider next the embedding of $G''_{s'}$ by *PackingProcess*. We again want to apply Lemma 40, this time with $S = N_{H_{s'-1}}(v) \setminus \operatorname{im} \phi'_s$. If (14) holds, then $|S| = (1 \pm 5C'\alpha_{s-1})\mu p_{s-1} n \geq \frac{1}{2}p\mu^2 n$. Hence we can apply Lemma 40 with S and with s' in place of s to conclude that with probability at least $1 - 3n^{-9}$ either (14) fails, or $\mathcal{E}_{s'-1}$ fails or

$$\begin{aligned} |N_{H_{s'}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'})| &= |N_{H_{s'-1}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'})| \\ &= (1 \pm C'\alpha_{s'})\mu|S| = (1 \pm 7C'\alpha_{s'})p_{s'-1}\mu^2 n \,. \end{aligned} \tag{15}$$

In a last stage, consider the embedding of $G''_{s'+1}, \ldots, G''_{s^*}$ by *PackingProcess*. We apply Lemma 39 with $T = N_{H_{s'}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'}) \subseteq N_{H_{s'}}(v)$ and with s' replaced by s^*, which is possible if (15) holds since then $|T| \geq \frac{1}{2}p\mu^2 n$. In this case, because $T \cap N_{H_{s^*}}(v) = N_{H_{s^*}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'})$, we conclude that with probability at least $1 - n^{-C}$ we have

$$\begin{aligned} |N_{H_{s*}}(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'})| &= (1 \pm \gamma^{-1}\alpha_{s^*})\frac{p_{s^*}}{p_{s'}}|T| \\ &\overset{(15)}{=} (1 \pm \gamma^{-1}\alpha_{s^*})\frac{p_{s^*}}{p_{s'}}(1 \pm 7C'\alpha_{s'})p_{s'-1}\mu^2 n = (1 \pm 9C'\alpha_{s^*})\mu^2 p_{s^*} n \end{aligned}$$

from which we obtain

$$|N_H(v) \setminus (\operatorname{im} \phi'_s \cup \operatorname{im} \phi'_{s'})| = (1 \pm {\gamma'}^3)\mu^2 p n$$

analogously to the discussion of $(P3)$ and as long as $\Delta(H_0) \leq 2\gamma n$. We conclude, using a union bound over v, s and s' and again (10) and Lemma 38(d), that $(P4)$ fails with probability at most $4n^{-5} + n^3(2 \cdot 3n^{-9} + 2 \cdot n^{-C}) \leq n^{-4}$.

<u>$(P5)$: $\sum_s w_s(v)\mathbb{1}_{u \notin \operatorname{im} \phi'_s} = (1 \pm {\gamma'}^3)\mu\frac{pn}{2}$</u>.

Fix u and $v \neq u$ and define

$$Y_s = w_s(v)\mathbb{1}_{u \notin \operatorname{im} \phi'_s} \,,$$

and observe that $Y_s \leq w_s(v) \leq \Delta$. Again, we want to apply Corollary 7. We have

$$\mathbb{E}[Y_s|\mathscr{H}_{s-1}] = \sum_{x\in V(G_s)} w_s(x) \cdot \mathbb{P}[x \hookrightarrow v, u \notin \operatorname{im}\phi'_s|\mathscr{H}_{s-1}] .$$

By Lemma 37 we obtain

$$\mathbb{E}[Y_s|\mathscr{H}_{s-1}] = \sum_{x\in V(G_s)} w_s(x) \cdot (1 \pm 10^3 C\alpha_{s-1} D\delta^{-1})\frac{\mu}{n} = \lfloor \nu n \rfloor \cdot (1 \pm 10^3 C\alpha_{s-1} D\delta^{-1})\frac{\mu}{n} .$$

This implies

$$\sum_s \mathbb{E}[Y_s|\mathscr{H}_{s-1}] = \lfloor \mu n \rfloor \lfloor \nu n \rfloor \cdot (1 \pm 10^3 C\alpha_{s^*} D\delta^{-1})\frac{\mu}{n} = \frac{\mu p n}{2} \cdot (1 \pm 2 \cdot 10^3 C\alpha_{s^*} D\delta^{-1}) .$$

We apply the second part of Corollary 7(b) with

$$\mathcal{E} = \bigcap_s \mathcal{E}_s , \quad R = \Delta , \quad \tilde{\mu} = \frac{\mu p n}{2} , \quad \tilde{\eta} = 2 \cdot 10^3 C\alpha_{s^*} D\delta^{-1}$$

and use $\tilde{\eta} \leq \frac{1}{2}$, which holds by definition of α_{s^*}, to conclude that

$$\begin{aligned}\mathbb{P}\big[\mathcal{E} \text{ and } \sum_s Y_s \neq \frac{\mu p n}{2} \cdot (1 \pm 4 \cdot 10^3 C\alpha_{s^*} D\delta^{-1})\big] &\leq 2\exp\Big(- \frac{\tilde{\mu} \cdot 4 \cdot 10^6 C^2 \alpha_{s^*}^2 D^2 \delta^{-2}}{4\Delta}\Big) \\ &\leq 2\exp(-10^{10} \log n) ,\end{aligned}$$

where the last inequality uses $\Delta \leq cn/\log n$, $c \leq 10^{-10}\gamma^{10D}\alpha_0^4$, $\alpha_0 \leq \alpha_{s^*}$, $p \geq \mu\nu$ and $\gamma \ll \nu \ll \mu$. Combining this with (12) and (10) and using a union bound over all u, v, we conclude that $(P5)$ fails with probability at most $2n^{-5} + n^2 \cdot n^{-10} \leq n^{-4}$.

$(P6)$: If $u \notin \operatorname{im}\phi'_s$ then we have $\sum_{v\colon vu\in E(H)} w_s(v) < \frac{10p^2n}{\mu}$.

The verification of this statement is the most complicated part of this proof. We fix $u \in V(H)$ and s with $s^* - \lfloor \mu n \rfloor < s \leq s^*$. We shall show that either an unlikely event occurs, or the desired property holds when G'_s is embedded, and then continues to hold while the remaining guest graphs are embedded. The embeddings of these guest graphs $G'_{s'}$ is performed in the graphs $H_{s'}$ and we shall show that $\sum_{v\colon vu\in E(H_{s'})} w_s(v)$ stays concentrated. But since $(P6)$ concerns the whole graph H, we additionally need to control the contribution of edges vu in $H^*_{s'}$, for which we can only provide an upper bound. More precisely, we shall establish the following claim. We will then, at the end of this proof, argue that this implies $(P6)$,

Claim 41. *Suppose $u \notin \operatorname{im}\phi'_s$. Then with probability at least $1 - 4n^{-19}$ either (H_i, H^*_0) is not $(\alpha_i, 2D+3)$-coquasirandom for some $i \in [s^*]$, or $\big(H_{i-1}, \phi'_i([t])\big)$ does not satisfy the $(C\alpha_{i-1}, 2D+3)$-diet condition for some $i \in [s^*]$ and $t \in [n-\delta n]$, or $\big(H_{i-1}, H^*_0, \phi'_i([t])\big)$ does not satisfy the $(2\eta, 2D+3)$-codiet condition for some $i \in [s^*]$ and $t \in [n-\delta n]$, or for each $s \leq s' \leq s^*$ we have*

$$\sum_{v:vu\in E(H_{s'})} w_s(v) = \big(1 \pm 10Cp^{-1}\alpha_{s'}\big)\tfrac{p_{s'}pn}{2\mu} , \tag{16}$$

and

$$\sum_{v:vu\in E(H^*_0)} w_s(v) \leq \tfrac{\gamma p n}{\mu} . \tag{17}$$

We will prove this claim in two steps. First (in Claim 42), we establish that it is very likely that (16) holds for $s'=s$ and that (17) holds. Then, based on Claims 43 to 46, we show that it is unlikely that any given $s'>s$ is the smallest s' for which (16) fails. Taking the union bound over s' will complete the proof of the claim.

Recall again that the embedding ϕ'_s of G'_s is given by letting *RandomEmbedding* perform the embedding of $G''_s[[n-\delta n]]$, thus constructing partial embeddings ψ_t for $0\le t\le(1-\delta)n$, and then ignoring the vertices that do not belong to G'_s, i.e. the last $\mu n-\delta n$ ones.

Claim 42. *Suppose $u\notin \operatorname{im}\phi'_s$. Then with probability at least $1-n^{-20}$ the pair (H_{s-1},H^*_0) is not $(\alpha_{s-1},2D+3)$-coquasirandom, or $\big(H_{s-1},\phi'_s\big([t]\big)\big)$ does not satisfy the $(C\alpha_{s-1},2D+3)$-diet condition for some $t\in[n-\lfloor\mu n\rfloor]$, or $\big(H_{s-1},H^*_0,\phi'_s\big([t]\big)\big)$ does not satisfy the $(2\eta,2D+3)$-codiet condition for some $t\in[n-\lfloor\mu n\rfloor]$, or we have*

$$\sum_{v:vu\in E(H_s)} w_s(v)=\big(1\pm 10Cp^{-1}\alpha_s\big)\tfrac{p_spn}{2\mu}\quad\textit{and}\quad \sum_{v:vu\in E(H^*_0)} w_s(v)\le \tfrac{\gamma pn}{\mu}\,.$$

Proof. We begin by proving the concentration of $\sum_{v:vu\in E(H_s)}w_s(v)$. For every $t\in[n]$, let x_t be the t-th vertex of G''_s, let $\mathcal{E}'_t$ be the event that H_{s-1} is $(\alpha_{s-1},2D+3)$-quasirandom and $\big(H_{s-1},\phi'_s\big([t]\big)\big)$ satisfies the $(C\alpha_{s-1},2D+3)$-diet condition, and let $\mathscr{H}'_t$ be a history up to and including the embedding of x_t which satisfies $\mathcal{E}'_t$. When *RandomEmbedding* is run, for $t\in[n-\lfloor\mu n\rfloor]$ we obtain

$$\mathbb{P}[x_t\hookrightarrow N_{H_{s-1}}(u)|\mathscr{H}_{t-1}]=\frac{(1\pm C\alpha_{s-1})p_{s-1}^{1+\deg^-_G(x_t)}(n-t+1)}{(1\pm C\alpha_{s-1})p_{s-1}^{\deg^-_G(x_t)}(n-t+1)}=(1\pm 3C\alpha_{s-1})p_{s-1}$$

where the first equality holds, since under assumption of the diet condition for $(H_{s-1},\phi'_s([t-1]))$ the candidate set for x_t is of size $(1\pm C\alpha_{s-1})p_{s-1}^{\deg^-_G(x_t)}(n-t+1)$, and since $u\notin\operatorname{im}\phi'_s$ and thus there exist $(1\pm C\alpha_{s-1})p_{s-1}^{1+\deg^-_G(x_t)}(n-t+1)$ candidates among $N_{H_{s-1}}(u)$. Now, set

$$X_t:=w_s(x_t)\cdot\mathbb{1}_{x_t\hookrightarrow N_{H_{s-1}}(u)}$$

so that

$$\sum_{v:vu\in E(H_s)}w_s(v)=\sum_{v:vu\in E(H_{s-1})}w_s(v)=\sum_{t\in[n]}X_t\ ,$$

where the first equation holds because of $u\notin\operatorname{im}\phi'_s$. In order to apply Corollary 7(b) observe that $0\le X_t\le\Delta$. Moreover,

$$\begin{aligned}\sum_{t\in[n]}\mathbb{E}[X_t|\mathscr{H}'_{t-1}]&=\sum_{t\in[n-\lfloor\mu n\rfloor]}\mathbb{E}[X_t|\mathscr{H}'_{t-1}]=\sum_{t\in[n-\lfloor\mu n\rfloor]}w_s(x_t)\mathbb{P}[x_t\hookrightarrow N_{H_{s-1}}(u)|\mathscr{H}'_{t-1}]\\&=(1\pm 3C\alpha_{s-1})p_{s-1}\sum_{t\in[n-\lfloor\mu n\rfloor]}w_s(x_t)=(1\pm 4C\alpha_{s-1})\frac{p_spn}{2\mu}\end{aligned}$$

since $p_s=(1-o(1))p_{s-1}$, and $\sum_{t\in[n-\lfloor\mu n\rfloor]}w_s(x_t)=\lfloor\nu n\rfloor$, and by definition of p. So, Corollary 7(b) with $\tilde\mu=\frac{p_spn}{2\mu}$, $\tilde\nu=4C\alpha_{s-1}\frac{p_spn}{2\mu}$ and $\tilde\varrho=C\alpha_{s-1}\frac{p_spn}{2\mu}$ yields

$$\mathbb{P}\left[\mathcal{E}'_t\text{ and }\sum_{t\in[n]}X_t\neq(1\pm 5C\alpha_{s-1})\frac{p_spn}{2\mu}\right]\le 2\exp\left(-\frac{C^2\alpha_{s-1}^2p_spn}{4\Delta\mu(1+5C\alpha_{s-1})}\right)\le n^{-21}$$

where the last inequality holds, since $\Delta \leq \frac{cn}{\log n}$ and by choice of c. This gives the first part of the claim, as $5C\alpha_{s-1} \leq 10Cp^{-1}\alpha_s$.

The second part of the claim, concerning H_0^*, is very similar, and we only sketch the proof. We define $\mathcal{E}_t''$ to be the event that (H_{s-1}, H_0^*) is $(\alpha_{s-1}, 2D+3)$-coquasirandom and $\big(H_{s-1}, H_0^*, \phi_s'([t])\big)$ satisfies the $(2\eta, 2D+3)$-codiet condition, and let $\mathscr{H}_t''$ be a history up to and including the embedding of x_t which satisfies $\mathcal{E}_{t''}''$. By a similar calculation as before, with $p^* \leq 1.1\gamma$ being the density of H_0^*, we see that when *RandomEmbedding* is run, we have

$$\mathbb{P}[x_t \hookrightarrow N_{H_0^*}(u) | \mathscr{H}_{t-1}] = \big(1 \pm 6\eta\big)p^* \leq \tfrac{3}{2}\gamma,$$

as the codiet condition for (H_{s-1}, H_0^*) makes sure that the candidate set for x_t is of size $(1 \pm 2\eta)p_{s-1}^{\deg_G^-(x_t)}(n-t+1)$, while among these candidates $(1 \pm 2\eta)p_{s-1}^{\deg_G^-(x_t)}p^*(n-t+1)$ vertices belong to $N_{H_0^*}(u)$.

Having that, we can again define $X_t' := w_s(x_t) \cdot \mathbb{1}_{x_t \hookrightarrow N_{H_0^*}(u)}$, and as before we obtain

$$\sum_{t\in[n]} \mathbb{E}[X_t' | \mathscr{H}_{t-1}''] = \sum_{t\in[n-\mu n]} \mathbb{E}[X_t' | \mathscr{H}_{t-1}''] \leq \tfrac{5}{3} \cdot \frac{\gamma p n}{2\mu}\,.$$

Applying Corollary 7(a), we get

$$\mathbb{P}\left[\mathcal{E}_t'' \text{ and } \sum_{t\in[n]} X_t' > \frac{\gamma p n}{\mu}\right] \leq n^{-21}\,.$$

This is the second part of the claim; the total failure probability is at most $2n^{-21} < n^{-20}$. □

We now need to show that it is unlikely that a given $s' > s$ is the first s' for which (16) fails. To that end, fix s' with $s^* - \lfloor \mu n \rfloor \leq s < s' \leq s^*$. For $s < i \leq s^*$ we define

$$Y_i := \sum_{v \in N_{H_{i-1}}(u) \setminus N_{H_i}(u)} w_s(v)\,.$$

We have

$$\sum_{v:\, vu \in E(H_{s'})} w_s(v) = \sum_{v:\, vu\in E(H_s)} w_s(v) - \sum_{i=s+1}^{s'} Y_i\,,$$

and so the missing piece to establishing Claim 41 is to show that the sum of the Y_i is likely to stay close to its expectation. We start by determining this expectation.

Claim 43. *Suppose that H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom, and suppose that* (16) *holds for $s' = i-1$. Then when* RandomEmbedding *is run to embed $G_i''[[n-\delta n]]$ into H_{i-1}, we have*

$$\mathbb{E}[Y_i | H_{i-1}] = \big(1 \pm 10^4 CD\alpha_{i-1}\delta^{-1}\big) \cdot \tfrac{pe(G_i'')}{\mu n}\,.$$

Proof. By definition of Y_i we obtain

$$Y_i = \sum_{v\in N_{H_{i-1}}(u)} w_s(v) \cdot \mathbb{1}_{uv \text{ is used when embedding } G_i''}$$

and therefore

$$\mathbb{E}[Y_i | H_{i-1}] = \sum_{v \in N_{H_{i-1}}(u)} w_s(v) \cdot \mathbb{P}[uv \text{ is used when embedding } G_i'' | H_{i-1}]\,.$$

Under assumption that H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom, Lemma 35 yields

$$\mathbb{E}[Y_i|H_{i-1}] = \sum_{v\in N_{H_{i-1}}(u)} w_s(v)\cdot(1\pm 500C\alpha_{i-1}\delta^{-1})^{4D+2}\frac{2e(G_i'')}{p_{i-1}n^2}\,.$$

Applying (16) for $s'=i-1$ finally leads to

$$\begin{aligned}\mathbb{E}[Y_i|H_{i-1}] &= (1\pm 10Cp^{-1}\alpha_{i-1})\frac{p_{i-1}pn}{2\mu}\cdot(1\pm 500C\alpha_{i-1}\delta^{-1})^{4D+2}\frac{2e(G_i'')}{p_{i-1}n^2}\\ &= (1\pm 10^4C\alpha_{i-1}D\delta^{-1})\frac{pe(G_i'')}{\mu n}\end{aligned}$$

where the last equality holds since $p\geq\nu\mu$ and $\delta\ll\nu\ll\mu$. □

What we would like to do now is apply Corollary 7 (or Lemma 6) to show that the sum of the Y_i is likely to be close to the sum of the observed expectations we just calculated. But unfortunately this approach fails, because the range of the Y_i is too large; it is possible that there are as few as $O(\log n)$ vertices which contain all the weight of w_s in $N_{H_s}(v)$, and we might use all the edges to these vertices in embedding a single G_i''. This is the reason for defining the random variables

$$Z_i := \max\{Y_i - K'\Delta, 0\}\quad\text{with}\quad K' = 10^{10}CD^3\delta^{-1}\,.$$

Trivially the 'capped' random variable

$$Y_i' := Y_i - Z_i$$

does not have an excessively large range (it cannot exceed $K'\Delta$), and we shall see (in the proof of Claim 41) that we can apply Corollary 7 to argue that the sum of the Y_i' is concentrated. In order to show that this implies that also the sum of the Y_i is concentrated, we need to argue that the 'error' caused by the Z_i is not too large, which we establish in Claim 46. As preparation for this, we will analyze the behaviour of the variables Z_i more in detail (in Claim 44) and bound their expectation (in Claim 45; we will need this bound when we show that the sum of the Y_i' is concentrated).

Let us now try to understand the behaviour of Z_i. Consider the embedding of $G_i''[[n-\delta n]]$ into H_{i-1} by *RandomEmbedding*. Observe that Z_i is determined by the vertex x_t that is embedded to u and by the embedding of neighbours of x_t. Until we embed x_t to u at time t, we have used no edges of H_{i-1} leaving u. On embedding a vertex to u, we have

$$\sum_{y\in N^-_{G_i''}(x_t)} w_s\big(\phi_i'(y)\big)\leq D\Delta\,,$$

because x_t has at most D neighbours preceding it in the degeneracy order. Consider now the successive embedding of the forward neighbours $y_1,\dots,y_\ell$ of x_t by *RandomEmbedding*. In order for $Z_i>0$ to occur, we have to embed the next j forwards neighbours of x_t (for some j) to vertices such that $\sum_{k=1}^{j} w_s\big(\phi_i'(y_k)\big)\geq(K'-D-1)\Delta$. We say that the embedding of G_i'' *goes near the cap* at the first time when we embed a y_j such that this inequality holds. We write $\mathsf{CapE}(i,y)$ for the event that the embedding of G_i'' goes near the cap at the time when we embed y (note that these events are pairwise disjoint as y ranges over $V(G_i'')$), and we write $\mathsf{CapE}(i)$ for their union, i.e. the event that the embedding of G_i'' goes near the cap at some time. If $\mathsf{CapE}(i,y_j)$ occurs, we have the inequality $Z_i\leq\sum_{k=j+1}^{\ell} w_s\big(\phi_i'(y_k)\big)$; it is important to note that the right hand side depends only on embeddings after the event of going near the cap is decided. Our next aim is to show that, conditioned on the embedding up to the time when x_t is embedded to u, it is unlikely that the embedding goes near the cap.

Claim 44. *Suppose that H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom. Suppose furthermore that ψ_t is a partial embedding of G_i'' to H_{i-1} generated by* RandomEmbedding *which embeds x_t to u (and embeds no vertices after x_t). Suppose that ψ_t is such that the probability, conditioned on H_{i-1} and ψ_t, of $(H_{i-1}, \operatorname{im}\phi_i')$ failing to have the $(C\alpha_{i-1}, 2D+3)$-diet condition is at most n^{-5}. Then we have*

$$\mathbb{P}\big[\mathsf{CapE}(i)\,\big|\,H_{i-1}, \psi_t\big] \le 3e^{-K'/8}\,.$$

Proof. With the notation from above, set

$$X_k = w_s(\phi_i'(y_k))$$

for every forward neighbour y_k of x_t, and observe that $0 \le X_k \le \Delta$. Let $\mathscr{H}_{k-1}'$ be a history up to and including the embedding ψ_r of the vertex x_r which comes immediately before y_k in the ordering of G_i''. Let $\tilde{\mathcal{E}}_r$ be the event that $(H_{i-1}, \operatorname{im}\psi_r)$ satisfies the $(C\alpha_{i-1}, 2D+3)$-diet condition.

Then, if $\tilde{\mathcal{E}}_r$ holds, we have

$$\mathbb{E}\left[X_k|\mathscr{H}_{k-1}'\right] \le \frac{\lfloor \nu n\rfloor}{\frac{1}{2}p^D\mu n} \le 2p^{-D}\mu^{-1}\nu$$

since the sum over all weights from G_i'' is $\lfloor \nu n\rfloor$, while the diet-condition ensures that the candidate set for y_k is of size at least

$$(1-C\alpha_{i-1})p^D\lfloor\mu n\rfloor \ge \frac{1}{2}p^D\mu n\ .$$

In particular,

$$\sum_{k=1}^{\ell}\mathbb{E}[X_k|\mathscr{H}_{k-1}'] \le 2p^{-D}\mu^{-1}\nu\Delta\ .$$

Applying the first part of Corollary 7(b) with $\mathcal{E} = \bigcup_r \tilde{\mathcal{E}}_r$, $\tilde{\mu} = \tilde{\nu} = p^{-D}\mu^{-1}\nu\Delta$, $\tilde{\varrho} = (K'-D-1-2p^{-D}\mu^{-1}\nu)\Delta$ and $R=\Delta$ we then obtain that

$$\begin{aligned}\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{k=1}^{\ell} X_k \ge (K'-D-1)\Delta\right] &\le 2\exp\left(-\frac{(K'-D-1-2p^{-D}\mu^{-1}\nu)^2}{2(K'-D-1)}\right)\\ &\le 2\exp\left(-\frac{\left(\frac{1}{2}K'\right)^2}{2K'}\right) = 2\exp\left(-\frac{K'}{8}\right)\ .\end{aligned}$$

Since by assumption the probability of $\mathcal{E}$ not occurring is at most n^{-5}, the claim follows. □

Now we can use this, and Lemma 36, to estimate the expectation of Z_i conditioned on H_{i-1} which is quasirandom.

Claim 45. *Suppose that H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom. Then we have*

$$\mathbb{E}\big[Z_i\big|H_{i-1}\big] \le 13\tfrac{e(G_i'')}{n}e^{-K'/8}\cdot 2\nu\mu^{-1}p_{s^*}^{-D}\,.$$

Proof. We have

$$\mathbb{E}[Z_i|H_{i-1}] = \sum_{x\in V(G_i'')}\mathbb{P}[x\hookrightarrow u|H_{i-1}]\cdot\mathbb{E}[Z_i|x\hookrightarrow u, H_{i-1}]\,. \tag{18}$$

Assuming that H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom, we know by Lemma 36 that

$$\mathbb{P}[x \hookrightarrow u | H_{i-1}] = (1 + 10^4 C\alpha_{s-1} D\delta^{-1})\frac{1}{n} = \big(1 \pm \tfrac{1}{2}\big)\frac{1}{n}\,. \tag{19}$$

For estimating $\mathbb{E}[Z_i | x \hookrightarrow u, H_{i-1}]$, we let $\mathcal{E}$ be the event that $(H_{i-1}, \operatorname{im}\phi_i')$ satisfies the $(C\alpha_{i-1}, 2D+3)$-diet condition. Then by linearity of expectation we have

$$\begin{aligned}\mathbb{E}[Z_i | x \hookrightarrow u, H_{i-1}] &= \mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] + \mathbb{E}[Z_i \mathbb{1}_{\bar{\mathcal{E}}} | x \hookrightarrow u, H_{i-1}] \\ &\le \mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] + n \cdot \tfrac{2n^{-9}}{2/n} = \mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] + 4n^{-7}\,,\end{aligned} \tag{20}$$

where the estimate for the second term is from Lemma 32 bounding the probability of $\bar{\mathcal{E}}$ and (19) lower bounding the probability of $x \hookrightarrow u$, and since trivially $Z_i \le Y_i \le n$. To estimate the first term, we observe that since outside $\mathsf{CapE}(i)$ we have $Z_i = 0$, it follows that

$$\begin{aligned}&\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] \\ &\qquad = \sum_{z \in V(G_i'')} \mathbb{P}[\mathsf{CapE}(i,z) | x \hookrightarrow u, H_{i-1}] \cdot \mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)]\,.\end{aligned} \tag{21}$$

Note that the only terms of the sum in which the probability is positive are those with z a forwards neighbour of x, so fix such a z. Recall that if $\mathsf{CapE}(i,z)$ occurs then we have

$$Z_i \le \sum_{\substack{y \in N_{G_i''}(x) \\ y \text{ comes after } z}} w_s(\phi_i'(y))\,, \quad \text{and so} \quad Z_i \mathbb{1}_{\mathcal{E}} \le \sum_{\substack{y \in N_{G_i''}(x) \\ y \text{ comes after } z}} w_s(\phi_i'(y)) \mathbb{1}_{\mathcal{E}}\,.$$

For bounding $\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)]$, for any forwards neighbour y of x which comes after z in the degeneracy order, let $\mathcal{H}'_{<y}$ denote any history up to and including the embedding of the vertex which comes immediately before y that is consistent with $x \hookrightarrow u$ and is contained in $\mathsf{CapE}(i,z)$. Then we have

$$\begin{aligned}&\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)] \\ &\quad \le \sum_{\substack{y \in N_{G_i''}(x) \\ y \text{ comes after } z}} \sum_{\mathcal{H}'_{<y}} \mathbb{E}\Big[w_s\big(\phi_i'(y)\big) \mathbb{1}_{\mathcal{E}} \Big| \mathcal{H}'_{<y}, H_{i-1}\Big] \cdot \mathbb{P}[\mathcal{H}'_{<y} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)]\,.\end{aligned} \tag{22}$$

Let y be a forwards neighbour of x which comes after z. Then y is not isolated, so it is in the first $n - \mu n$ vertices of G_i''. We want to calculate $\mathbb{E}\Big[w_s\big(\phi_i'(y)\big)\mathbb{1}_{\mathcal{E}}\Big|\mathcal{H}'_{<y}, H_{i-1}\Big]$. There are two cases to consider. First, if $\mathcal{E}$ occurs, then since y is in the first $n - \mu n$ vertices of G_i'', it has a candidate set of size at least

$$(1 - C\alpha_{i-1}) p_{i-1}^D \lfloor \mu n \rfloor \ge \frac{1}{2} p_{s^*}^D \mu n\,.$$

Hence we embed y uniformly to a set of size at least $\frac{1}{2}p_{s^*}^D \mu n$, so (because the total weight of all vertices in G_s'' is $\lfloor \nu n \rfloor$) the expectation of $w_s\big(\phi_i'(y)\big)$ conditioned on $\mathcal{H}'_{<y}$ and H_{i-1} is at most $\frac{\nu n}{p_{s^*}^D \mu n/2}$. Second, if y is chosen from a candidate set of size less than $\frac{1}{2}p_{s^*}^D \mu n$, then the event $\mathcal{E}$ does not occur, and so the conditional expectation we want to calculate is zero. In either case, we obtain

$$\mathbb{E}\Big[w_s\big(\phi_i'(y)\big)\mathbb{1}_{\mathcal{E}}\Big|\mathcal{H}'_{<y}, H_{i-1}\Big] \le 2\nu\mu^{-1}p_{s^*}^{-D}\,.$$

Plugging this into (22) gives

$$\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)] \le d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D}\,,$$

since the sum over $\mathcal{H}'_{<y}$ of $\mathbb{P}[\mathcal{H}'_{<y} | x \hookrightarrow u, H_{i-1}, \mathsf{CapE}(i,z)]$ is trivially 1, and $d_{G''_i}(x)$ is at least as big as the number of forward neighbours of x which come after y. Now putting this into (21) we obtain

$$\begin{aligned}\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] &\le \sum_{z \in V(G''_i)} \mathbb{P}[\mathsf{CapE}(i,z) | x \hookrightarrow u, H_{i-1}] \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D} \\ &= \mathbb{P}[\mathsf{CapE}(i) | x \hookrightarrow u, H_{i-1}] \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D}\,.\end{aligned}$$

We finally use Claim 44 to estimate $\mathbb{P}[\mathsf{CapE}(i) | x \hookrightarrow u, H_{i-1}]$. By (19), we have $\mathbb{P}[x \hookrightarrow u | H_{i-1}] \ge \frac{1}{2n}$. By Lemma 32, the probability that (H_{i-1}, ϕ'_i) fails to have the $(C\alpha_{i-1}, 2D+3)$-diet condition, conditioned on H_{i-1}, is at most $2n^{-9}$. Consequently, summing up $\mathbb{P}[\psi_x | x \hookrightarrow u, H_{i-1}]$ over partial embeddings ψ_x which embed the vertices up to and including x of G''_i, and embed x to u, but which fail the condition of Claim 44 (i.e. the probability that (H_{i-1}, ϕ'_i) fails to have the $(C\alpha_{i-1}, 2D+3)$-diet condition, conditioned on H_{i-1} and ψ_x, exceeds n^{-5}), we obtain at most $4n^{-3}$. For any ψ_x which does satisfy the condition of Claim 44, we have $\mathbb{P}[\mathsf{CapE}(i) | H_{i-1}, \psi_x] \le 3e^{-K'/8}$. Putting these together, we have

$$\begin{aligned}\mathbb{P}[\mathsf{CapE}(i) | x \hookrightarrow u, H_{i-1}] &= \sum_{\psi_x} \mathbb{P}[\psi_x | x \hookrightarrow u, H_{i-1}] \cdot \mathbb{P}[\mathsf{CapE}(i) | \psi_x, x \hookrightarrow u, H_{i-1}] \\ &\le 4n^{-3} \cdot 1 + 1 \cdot 3e^{-K'/8} = 3e^{-K'/8} + 4n^{-3}\,.\end{aligned}$$

At last, we obtain

$$\mathbb{E}[Z_i \mathbb{1}_{\mathcal{E}} | x \hookrightarrow u, H_{i-1}] \le (3e^{-\frac{K'}{8}} + 4n^{-3}) \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D}\,.$$

Thus, using (20), we have

$$\begin{aligned}\mathbb{E}[Z_i | x \hookrightarrow u, H_{i-1}] &\le (3e^{-\frac{K'}{8}} + 4n^{-3}) \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D} + 4n^{-7} \\ &\le 3e^{-\frac{K'}{8}} \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D} + n^{-2}\,,\end{aligned}$$

and so by (18) and (19) we get

$$\begin{aligned}\mathbb{E}[Z_i | H_{i-1}] &\le \sum_{x \in V(G''_i)} \frac{2}{n} \cdot \Big(3e^{-\frac{K'}{8}} \cdot d_{G''_i}(x) \cdot 2\nu\mu^{-1} p_{s^*}^{-D} + n^{-2}\Big) \\ &= \frac{2\sum_{x \in V(G''_i)} d_{G''_i}(x)}{n} \cdot 3e^{-\frac{K'}{8}} \cdot 2\nu\mu^{-1} p_{s^*}^{-D} + 2n^{-2} < 13\frac{e(G''_i)}{n} \cdot e^{-\frac{K'}{8}} \cdot 2\nu\mu^{-1} p_{s^*}^{-D}\,.\end{aligned}$$

□

This expectation is tiny, because the term $e^{-K'/8}$ is very small. Thus we see that the expectations of Y'_i and Y_i (conditioning on any H_{i-1} which is quasirandom) are very close. The final thing we have to do before we complete the proof of Claim 41 is to show that the sum of the Z_i is likely to be very small.

Claim 46. *With probability at least* $1-2n^{-20}$*, the following event occurs when* PackingProcess *is run. Either* H_i *is not* $(\alpha_i, 2D+3)$*-quasirandom for some* $i \in [s^*]$*, or* $(H_{i-1}, \phi'_i([n-\mu n]))$

does not satisfy the $(C\alpha_{i-1}, 2D+3)$*-diet condition for some* $i \in [s^*]$*, or we have*

$$\sum_{i=s+1}^{s'} Z_i \le \tfrac{\alpha_s pn}{1000\mu}\,.$$

Proof. Let $\mathcal{E}$ denote the event that H_i is $(\alpha_i, 2D+3)$-quasirandom for each $i \in [s^*]$, and $(H_{i-1}, \phi'_i([n-\mu n]))$ satisfies the $(C\alpha_{i-1}, 2D+3)$-diet condition for each $i \in [s^*]$. So we want to show that it is likely that either $\mathcal{E}$ fails or $\sum_{i=s+1}^{s'} Z_i < \frac{\alpha_s pn}{1000\mu}$.

In order to prove this claim, we need to reinterpret $\sum_{i=s+1}^{s'} Z_i$. The random variables Z_i can be very large, so that Corollary 7 does not help us.

What we do is to use our earlier observation that we can understand Z_i as follows. We watch *RandomEmbedding* as it embeds $G''_i[[n-\delta n]]$, until it embeds some x_t to u, and then embeds the forwards neighbours of x_t until it goes near the cap (if one or the other event does not occur, then $Z_i = 0$). Then Z_i is at most the sum of $w_s\big(\phi'_i(y)\big)$ taken over forwards neighbours y of x_t which are embedded after reaching the cap. We refer to these vertices y as *after-cap* vertices. We then use the inequality

$$\sum_{i=s+1}^{s'} Z_i \le \sum_{i=s+1}^{s'} \sum_{\substack{v=\phi'_i(y)\\ \text{for } y \text{ after-cap in } G''_i}} w_s(v)\,,$$

where the right hand side sum runs over all after-cap vertices in all graphs $G''_{s+1}, \dots, G''_{s'}$. For a given after-cap vertex $y \in V(G''_i)$ we know y is an after-cap vertex before we embed it. Now when we embed y, provided the $(C\alpha_{i-1}, 2D+3)$-diet condition holds for $(H_{i-1}, \phi([n-\mu n]))$, we embed it uniformly into a set S of size at least $\frac{1}{2}p_{i-1}^D \mu n$ (because y, since it is not isolated, must be one of the first $n-\mu n$ vertices of G''_i). The sum of $w_s(z)$ over the vertices z of S is at most $\lfloor \nu n \rfloor$. So the expected value of $w_s\big(\phi'_i(y)\big)$, conditioned on the history up to the time $y-1$ immediately before embedding y and on the $(C\alpha_{i-1}, 2D+3)$-diet condition holding for $(H_{i-1}, \phi'_i([y-1]))$, is at most $2\lfloor \nu n \rfloor p_{i-1}^{-D} \mu^{-1} n^{-1} \le 2pp_{s^*}^{-D}\mu^{-2}$, where the inequality uses the equation $\lfloor \nu n \rfloor \lfloor \mu n \rfloor = p\binom{n}{2}$.

Let

$$L := 40Dne^{-K'/8}\,,$$

and define a random variable X_j for $1 \le j \le L$ by $X_j = w_s\big(\phi'_i(y)\big)$, where the jth after-cap vertex in a run of *PackingProcess* is $y \in V(G''_i)$ (y and hence i depend on the run of *PackingProcess*). If there is no such after-cap vertex, we let $X_j := 0$. Observe that we have $0 \le X_j \le \Delta$ for each j, and what we just calculated is that, letting $\mathscr{H}_j$ denote the history of *PackingProcess* up to immediately before embedding the jth after-cap vertex $y \in V(G''_i)$, if the $(C\alpha_{i-1}, 2D+3)$-diet condition holds for $(H_{i-1}, \phi'_i([y-1]))$, then $\mathbb{E}[X_j|\mathscr{H}_j] \le 2pp_{s^*}^{-D}\mu^{-2}$. So we can apply Corollary 7(a), with $\tilde{\mu} = 2pp_{s^*}^{-D}\mu^{-2}L$, to obtain

$$\mathbb{P}\Big[\mathcal{E} \text{ and } \sum_{j=1}^{L} X_j > 4pp_{s^*}^{-D}\mu^{-2}L\Big] < \exp\Big(-\frac{2pp_{s^*}^{-D}\mu^{-2}L}{4\Delta}\Big) \le n^{-20}\,,$$

where the final inequality follows from $\Delta = \frac{cn}{\log n}$ and by choice of c and K'.

Let T denote the total number of after-cap vertices encountered during the entire run of *PackingProcess*. Since by choice of K' and (4) we have $4pp_{s^*}^{-D}\mu^{-2}L < \frac{\alpha_s pn}{1000\mu}$, what we have just argued is that

$$\mathbb{P}\Big[\mathcal{E} \text{ and } T \le L \text{ and } \sum_{i=s+1}^{s'} Z_i > \tfrac{\alpha_s pn}{1000\mu}\Big] \le n^{-20}\,.$$

What we now want to do is estimate the probability of the event that $\mathcal{E}$ occurs and that $T > L$.

To that end, for each $s+1 \le i \le s'$, we define X'_i to be the number of after-cap vertices embedded from G''_i in a given run of *PackingProcess*. By definition we have $T = \sum_{i=s+1}^{s'} X'_i$. Now, if H_{i-1} is $(\alpha_{i-1}, 2D+3)$-quasirandom, we can estimate $\mathbb{E}[X'_i|H_{i-1}]$ as follows. First, observe X'_i can only be positive if some $x_t \in V(G''_i)$ is embedded to u, and then G''_i goes near the cap, and then the remaining neighbours of x_t will be the after-cap vertices counted by X'_i. So we have

$$X'_i \le \sum_{x_t \in V(G''_i)} \mathbb{1}_{x_t \hookrightarrow u} \mathbb{1}_{\mathsf{CapE}(i)} \cdot d_{G''_i}(x_t)\,.$$

It follows that

$$\mathbb{E}\big[X'_i\big|H_{i-1}\big] \le \sum_{x_t \in V(G''_i)} d_{G''_i}(x_t)\mathbb{P}\big[x_t \hookrightarrow u \mid H_{i-1}\big] \cdot \mathbb{P}\big[\mathsf{CapE}(i)\big|H_{i-1}, \psi_t\big]\,,$$

where ψ_t is a partial embedding of the first t vertices of G''_i into H_{i-1} generated by *RandomEmbedding* which embeds x_t to u. By respectively Lemma 36 and Claim 44, we have

$$\mathbb{E}\big[X'_i\big|H_{i-1}\big] \le \sum_{x_t \in V(G''_i)} d_{G''_i}(x_t) \cdot \big(\tfrac{2}{n} \cdot 3e^{-K'/8} + 2n^{-4}\big) \le 20De^{-K'/8}\,,$$

where the first inequality uses the observation that, by Lemma 32, there is at most $2n^{-4}$ probability of generating ψ_t such that the $(C\alpha_{i-1}, 2D+3)$-diet condition holding for $(H_{i-1}, \phi([n-\mu n]))$ has more than n^{-5} chance of failing (when embedding the remaining vertices). The second inequality uses the fact that G''_i has at most Dn edges and so the sum of its degrees is at most $2Dn$.

Since $0 \le X'_i \le \Delta$ for each i, we can apply Corollary $7(a)$, with $\tilde{\mu} = 20Dne^{-K'/8}$, to obtain

$$\mathbb{P}\big[\mathcal{E} \text{ and } \sum_{i=s+1}^{s'} X'_i > 40Dne^{-K'/8}\big] < \exp\big(-\tfrac{20Dne^{-K'/8}}{4\Delta}\big) \le n^{-20}\,,$$

where the second inequality comes from $\Delta = \frac{cn}{\log n}$ and choice of c and K'. Since $\sum_{i=s+1}^{s'} X'_i = T$, this proves as desired that it is unlikely that $\mathcal{E}$ occurs and $T > L$.

Putting these two pieces together, we conclude that with probability at most $2n^{-20}$, the event $\mathcal{E}$ occurs and we have $\sum_{i=s+1}^{s'} Z_i > \frac{\alpha_s pn}{1000\mu}$. This completes the proof of the claim. □

The reader might at this point wonder why we cannot simply estimate the sum of the Y_i by modifying the above method. The point is that it is not easy to obtain an accurate estimate of the quantity $\mathbb{E}[X_j|\mathscr{H}_j]$ in the above proof (the upper bound we obtain above is off from the truth by a rather large factor, compensated for by the unlikeliness of going near the cap), and we would need such an accurate estimate for Claim 41.

Finally, we are in a position to prove Claim 41.

Proof of Claim 41. Firstly, by Claim 42 we have that either (H_i, H_0^*) is not $(\alpha_i, 2D+3)$-coquasirandom for some $i \in [s^*]$, or $\big(H_{i-1}, \phi'_i\big([t]\big)\big)$ does not satisfy the $(C\alpha_{i-1}, 2D+3)$-diet condition for some $i \in [s^*]$ and $t \in [n-\delta n]$, or $\big(H_{i-1}, H_0^*, \phi'_i\big([t]\big)\big)$ does not satisfy the $(2\eta, 2D+3)$-codiet condition for some $i \in [s^*]$ and $t \in [n-\delta n]$, or that (17) holds and (16) holds for the case

$s' = s$ with probability at least $1 - n^{-20}$. Now, let $s < s' \le s^*$. We aim to show that with probability at most $3n^{-20}$ we have that (16) continues to hold for s'. Taking a union bound over the choices of s' then completes the proof of Claim 41.

More precisely, let $\mathcal{E}$ denote the event that $u \notin \operatorname{im}\phi'_s$, and (H_i, H_0^*) is $(\alpha_i, 2D+3)$-coquasirandom for each $i \in [s^*]$, and $\big(H_{i-1}, \phi'_i([t])\big)$ satisfies the $(C\alpha_{i-1}, 2D+3)$-diet condition for each $i \in [s^*]$ and $t \in [n - \delta n]$, and $\big(H_{i-1}, H_0^*, \phi'_i([t])\big)$ satisfies the $(2\eta, 2D+3)$-codiet condition for each $i \in [s^*]$ and $t \in [n - \delta n]$, and (16) holds for each $s \le i < s'$. Our goal is to show that $\mathcal{E}$ occurs and (16) fails for s' with probability at most $3n^{-20}$.

By Claim 42, with probability at least $1 - n^{-20}$, either we witness a failure of $\mathcal{E}$ before beginning to embed G''_{s+1}, or we have $\sum_{v\colon vu \in E(H_s)} w_s(v) = \big(1 \pm 10Cp^{-1}\alpha_s\big)\frac{p_s pn}{2\mu}$. Suppose that this likely event occurs, and that we do not witness a failure of $\mathcal{E}$ before beginning to embed G''_{s+1}.

Since we have

$$\sum_{v\colon vu \in E(H_{s'})} w_s(v) = \sum_{v\colon vu \in E(H_s)} w_s(v) - \sum_{i=s+1}^{s'} Y_i\,,$$

and we want to conclude that it is unlikely that $\mathcal{E}$ occurs and $\sum_{v\colon vu\in E(H_{s'})} w_s(v) \ne \big(1 \pm 10Cp^{-1}\alpha_{s'}\big)\frac{p_{s'}pn}{2\mu}$, it is enough to estimate the probability, conditioned on H_s, that $\mathcal{E}$ occurs and

$$\sum_{i=s+1}^{s'} Y_i \ne \big(1 \pm 10Cp^{-1}\alpha_s\big)\tfrac{p_s pn}{2\mu} - \big(1 \pm 10Cp^{-1}\alpha_{s'}\big)\tfrac{p_{s'}pn}{2\mu} = \tfrac{(p_s - p_{s'})pn}{2\mu} \pm 20C\alpha_{s'}\tfrac{p_s n}{2\mu}\,. \tag{23}$$

We have $Y_i = Y'_i + Z_i$ for each i, and so $\sum_{i=s+1}^{s'} Y_i = \sum_{i=s+1}^{s'} Y'_i + \sum_{i=s+1}^{s'} Z_i$. For showing that (23) is unlikely to occur, we will use Corollary 7 to argue that $\sum Y'_i$ is concentrated and Claim 46 to bound the contribution of $\sum Z_i$. Accordingly, we shall first calculate the expectation of $\sum Y'_i$.

By Claim 43, provided H_{i-1} does not witness that $\mathcal{E}$ fails, we have $\mathbb{E}[Y_i|H_{i-1}] = \big(1 \pm 10^4 CD\alpha_{i-1}\delta^{-1}\big) \cdot \frac{pe(G''_i)}{\mu n}$. By Claim 45, again provided H_{i-1} does not witness that $\mathcal{E}$ fails, we have $\mathbb{E}[Z_i|H_{i-1}] \le 13\frac{e(G''_i)}{n}e^{-K'/8} \cdot 2\nu\mu^{-1}p_{s^*}^{-D}$. By linearity, we conclude

$$\begin{aligned}\mathbb{E}[Y'_i|H_{i-1}] &= \big(1 \pm 10^4 CD\alpha_{i-1}\delta^{-1}\big) \cdot \tfrac{pe(G''_i)}{\mu n} \pm 13\tfrac{e(G''_i)}{n}e^{-K'/8} \cdot 2\nu\mu^{-1}p_{s^*}^{-D}\\ &= \big(1 \pm 10^5 CD\alpha_{i-1}\delta^{-1}\big) \cdot \tfrac{pe(G''_i)}{\mu n}\,,\end{aligned}$$

where for the second inequality we use our choice of K'. Summing this up, we see that either $\mathcal{E}$ fails or we have

$$\begin{aligned}\sum_{i=s+1}^{s'} \mathbb{E}[Y_i'|H_{i-1}] &= \sum_{i=s+1}^{s'} \left(1 \pm 10^5 CD\alpha_{i-1}\delta^{-1}\right) \cdot \tfrac{pe(G_i'')}{\mu n} \\ &= \sum_{i=s+1}^{s'} \tfrac{pe(G_i'')}{\mu n} \pm \sum_{i=s+1}^{s'} 10^5 CD\alpha_{i-1}\delta^{-1} \cdot \tfrac{pDn}{\mu n} \\ &= \tfrac{p}{\mu n}\left(p_s - p_{s'}\right)\binom{n}{2} \pm 10^5 CD^2\delta^{-1}\mu^{-1}p\int_{i=-\infty}^{s'} \alpha_i \mathrm{d}i \\ &\overset{(9)}{=} \tfrac{pn}{2\mu}\left(p_s - p_{s'}\right) \pm \tfrac{1}{\mu} \pm 10^5 CD^2\delta^{-1}\mu^{-1}p \cdot \frac{\delta n}{10^8 CD^3}\alpha_{s'} \\ &= \tfrac{pn}{2\mu}\left(p_s - p_{s'}\right) \pm \frac{p\alpha_{s'}n}{100\mu} = \tfrac{(p_s - p_{s'})pn}{2\mu} \pm C\alpha_{s'}\tfrac{p_s n}{2\mu}\,,\end{aligned}$$

where the final inequality is by choice of C and since $p \leq p_s + 2\gamma$ according to (3). Now applying the first part of Corollary 7(b), with $\tilde{\mu} = \frac{(p_s - p_{s'})pn}{2\mu}$ and $\tilde{\nu} = \tilde{\varrho} = C\alpha_{s'}\frac{p_s n}{2\mu}$, and using the fact $0 \leq Y_i' \leq K'\Delta$, we obtain

$$\mathbb{P}\Big[\mathcal{E} \text{ and } \sum_{i=s+1}^{s'} Y_i' \neq \tfrac{(p_s - p_{s'})pn}{2\mu} \pm 2C\alpha_{s'}\tfrac{p_s n}{2\mu}\Big] < 2\exp\left(-\tfrac{\tilde{\varrho}^2}{2K'\Delta(\tilde{\mu}+\tilde{\nu}+\tilde{\varrho})}\right) \leq n^{-20}\,,$$

where the final inequality uses $\Delta = \frac{cn}{\log n}$ and the choice of c and K'.

Putting this estimate together with Claim 46, where we show that with probability at least $1 - n^{-20}$ either $\mathcal{E}$ does not occur, or we have $\sum_{i=s+1}^{s'} Z_i \leq \frac{\alpha_s pn}{1000\mu}$, we conclude the following. With probability at least $1 - 3n^{-20}$, either $\mathcal{E}$ does not occur, or we have

$$\sum_{i=s+1}^{s'} Y_i = \tfrac{(p_s - p_{s'})pn}{2\mu} \pm 2C\alpha_{s'}\tfrac{p_s n}{2\mu} \pm \tfrac{\alpha_s pn}{1000\mu} = \tfrac{(p_s - p_{s'})pn}{2\mu} \pm 3C\alpha_{s'}\tfrac{p_s n}{2\mu}\,.$$

If this holds (23) does not occur. With this we finally proved that with probability at most $3n^{-20}$ the event $\mathcal{E}$ occurs and (16) holds for each $s \leq i < s'$ but fails for s'. □

Finally, we argue that Claim 41 implies ($P6$) holds with high probability. It is straightforward to check that $10Cp^{-1}\alpha_{s^*} < 1$, and $p_{s^*} \leq p$. Since $E(H) \subseteq E(H_{s^*}) \cup E(H_0^*)$, provided (16) with $s' = s^*$ and (17) hold, we have

$$\begin{aligned}\sum_{v:vu\in E(H)} w_s(v) &\leq \sum_{v:vu\in E(H_{s^*})} w_s(v) + \sum_{v:vu\in E(H_0^*)} w_s(v) \\ &\leq 2\tfrac{p_{s^*}pn}{2\mu} + \tfrac{\gamma pn}{\mu} < \tfrac{10p^2 n}{\mu}\,,\end{aligned}$$

where the last inequality follows since $p_{s^*}, \gamma < p$. Thus by Claim 41, with probability at least $1 - 4n^{-19}$, either the stated coquasirandomness, diet or codiet conditions fail, or ($P6$) holds for fixed u and s. So it is enough to check that it is unlikely that either the stated coquasirandomness, diet or codiet conditions fail. By respectively Lemma 38(b), and Lemma 32(b) and (d) (and the union bound over the at most $2n$ runs of *RandomEmbedding*), the probability that either of these occur is at most $2n^{-5} + 4n^{-8}$. For the latter, note that $\beta_t(\alpha_{i-1}) \leq C\alpha_{i-1}$

for each i, t. We finally conclude, using a union bound over u and s, that $(P6)$ holds with probability at least $1 - n^2 \cdot 4n^{-19} - 2n^{-5} - 4n^{-8} > 1 - 3n^{-5}$. □

9. Concluding remarks

Once one knows that a given collection of graphs $\mathcal{G}$ can be packed into a host graph $\widehat{H}$, it is natural to ask whether there is an efficient algorithm, randomised or not, which will exhibit such a packing. For $\mathcal{G}$ as in Theorem 2 (with the various parameters taken as fixed while n is large) the obvious answer is simply to run our packing algorithm. Recall that one can find a degeneracy order of any given graph simply by iteratively removing vertices of smallest degree. Lemma 47 guarantees that there is a large independent set of vertices all of which have the same small degree, and the proof in [1] consists of showing that a simple polynomial-time algorithm to find the set succeeds, and hence Lemma 9 is algorithmic. Similarly it is clear from the proof that Lemma 10 is algorithmic. Since the reduction of Theorem 2 to Theorem 11 requires these last two lemmas, it suffices to argue that Theorem 11 is algorithmic.

Most of the steps in our packing algorithm simply consist of uniform random samples from sets which are of linear size and trivial to compute. In addition the completion step of *PackingProcess* requires finding a perfect matching in a linear-sized and easily computed auxiliary bipartite graph; this is well known to be solvable in polynomial time using the augmenting paths algorithm. Finally, the completion step of *MatchLeaves* requires sampling uniformly from the set of perfect matchings of a dense bipartite graph (which is linear-sized and easy to compute).

If one assumes that it is possible to sample in polynomial time from these various distributions, then our algorithm clearly is polynomial time. However, if the source of randomness is an unbiased bit string (which is the natural and usual assumption) then one cannot sample exactly uniformly from arbitrary distributions. It is standard in the literature to ignore this problem (because to sample approximately uniformly is possible and this suffices), but for completeness we give the details.

For the random sampling in *PackingProcess*, it is easy to sample approximately uniformly: using k bits of randomness one can approximately sample any probability p Bernoulli random variable up to an error 2^{-k} by viewing the bits as an integer in $[2^k]$ and returning 1 if this integer is at most $2^k p$. One can similarly select approximately uniformly from a set, by partitioning $[2^k]$ into intervals of approximately equal size corresponding to the set elements. For all the analysis here and in [1], it is easy to check that using n random bits per sample, the sampling error is tiny compared to the probabilities we want to estimate and is absorbed by our error terms (in fact, $O(\log n)$ bits would suffice).

However sampling a perfect matching approximately uniformly, even from a dense bipartite graph, is not so obviously possible. We actually do not need a uniform random perfect matching: what we need is any distribution on perfect matchings which satisfies the conclusion of Lemma 20, i.e. that any given edge is in the matching with probability not too much greater (by at most a factor $\frac{3}{2}$ would suffice) than the average, which is trivially at least n^{-1}. In particular, this is the case if we sample from any distribution on perfect matchings whose total variation distance from the uniform distribution is at most n^{-2}. A result of Jerrum, Sinclair and Vigoda [12, Lemma 2.1] states that for any $2n$-vertex balanced bipartite graph G which has a perfect matching, there is an algorithm which samples from perfect matchings on G with a distribution whose total variation distance from the uniform distribution is at most δ, whose running time is polynomial in n and $\log \delta^{-1}$.

In conclusion, one can actually simulate the randomised algorithm of [1] and this paper in polynomial time. Following the (somewhat) general belief that RP $\neq$ NP, this suggests that the packing problem for the graphs we pack in this paper should not be NP-complete (in contrast to the general packing problem, which is known to be NP-complete [3]). We suspect the problem is in P, but we do not know how to derandomise our algorithm, or otherwise provide a deterministic polynomial time algorithm for the packing.

10. Acknowledgements

Part of the work leading to this paper was done while PA and JB visited Hamburg. PA and JB would like to thank TU Hamburg and the University of Hamburg for their hospitality, the Suntory and Toyota International Centres for Economics and Related Disciplines and TU Hamburg for financial support, and Heike Böttcher for help with childcare arrangements.

We thank the anonymous referee for extremely careful reading and very helpful comments.

References

[1] P. Allen, J. Böttcher, J. Hladký, and D. Piguet, *Packing degenerate graphs*, Adv. Math. **354** (2019), 106739.

[2] J. Böttcher, J. Hladký, D. Piguet, and A. Taraz, *An approximate version of the tree packing conjecture*, Israel J. Math. **211** (2016), no. 1, 391–446.

[3] D. Dor and M. Tarsi, *Graph decomposition is NP-complete: a complete proof of Holyer's conjecture*, SIAM J. Comput. **26** (1997), no. 4, 1166–1187.

[4] R. A. Duke, H. Lefmann, and V. Rödl, *A fast approximation algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. **24** (1995), no. 3, 598–620.

[5] A. Ferber, C. Lee, and F. Mousset, *Packing spanning graphs from separable families*, Israel J. Math. **219** (2017), no. 2, 959–982.

[6] A. Ferber and W. Samotij, *Packing trees of unbounded degrees in random graphs*, J. Lond. Math. Soc. (2) **99** (2019), no. 3, 653–677.

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

[8] J. A. Gallian, *A dynamic survey of graph labeling*, Electron. J. Combin. **5** (1998), Dynamic Survey 6.

[9] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption: hypergraph f-designs for arbitrary f*, Memoirs of the American Mathematical Society, accepted, arXiv:1611.06827.

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

[11] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.

[12] M. Jerrum, A. Sinclair, and E. Vigoda, *A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries*, J. ACM **51** (2004), no. 4, 671–697.

[13] F. Joos, J. Kim, D. Kühn, and D. Osthus, *Optimal packings of bounded degree trees*, J. Eur. Math. Soc. (JEMS) **21** (2019), no. 12, 3573–3647.

[14] P. Keevash, *The existence of designs*, arXiv:1401.3665.

[15] ———, *The existence of designs II*, arXiv:1802.05900.

[16] P. Keevash and K. Staden, *Ringel's tree packing conjecture in quasirandom graphs*, arXiv:2004.09947.

[17] J. Kim, D. Kühn, D. Osthus, and M. Tyomkyn, *A blow-up lemma for approximate decompositions*, Trans. Amer. Math. Soc. **371** (2019), no. 7, 4655–4742.

[18] T. P. Kirkman, *On a problem in combinations*, Cambridge and Dublin Math. J. **2** (1847), 191–204.

[19] F. Knox, D. Kühn, and D. Osthus, *Edge-disjoint Hamilton cycles in random graphs*, Random Structures Algorithms **46** (2015), no. 3, 397–445.

[20] E. Lucas, *Récréations mathématiques*, 2ième éd., nouveau tirage, Librairie Scientifique et Technique Albert Blanchard, Paris, 1960.

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

[22] R. Montgomery, A. Pokrovskiy, and S. B., *A proof of Ringel's conjecture*, arXiv:2001.02665.

[23] R. Montgomery, A. Pokrovskiy, and B. Sudakov, *Embedding rainbow trees with applications to graph labelling and decomposition*, J. Eur. Math. Soc. (JEMS) **22** (2020), no. 10, 3101–3132.
[24] J. Plücker, *System der analytischen Geometrie, auf neue Betrachtungsweisen gegründet, und insbesondere eine ausführliche Theorie der Curven dritter Ordnung enthaltend*, Duncker und Humboldt, Berlin, 1835.
[25] D. K. Ray-Chaudhuri and R. M. Wilson, *Solution of Kirkman's schoolgirl problem*, Combinatorics (Proc. Sympos. Pure Math., Vol. XIX, Univ. California, Los Angeles, Calif., 1968), Amer. Math. Soc., Providence, R.I., 1971, pp. 187–203.
[26] G. Ringel, *Problem 25*, Theory of Graphs and its Applications (Proc. Int. Symp. Smolenice 1963), Czech. Acad. Sci., Prague, 1963.
[27] V. Rödl, *On a packing and covering problem*, European J. Combin. **6** (1985), no. 1, 69–78.
[28] J. Steiner, *Combinatorische Aufgaben*, J. Reine Angew. Math. **45** (1853), 181–182.
[29] R. M. Wilson, *An existence theory for pairwise balanced designs. III. Proof of the existence conjectures*, J. Combin. Theory Ser. A **18** (1975), 71–79.

Appendix A. Standard probability and graph theory

In this section we prove for completeness some easy and fairly standard results stated in the main body of the paper.

Proof of Proposition 8. To prove both probability bounds, we use the well-known Prüfer code bijection between labelled n-vertex trees and sequences of $n-2$ vertex labels.

For (i), we note that a vertex is a leaf if and only if its label does not appear in the corresponding Prüfer code, and hence the Prüfer code of a tree with less than $\frac{n}{100}$ leaves has at least $\frac{49n}{50}$ distinct labels. Consider generating the first $\frac{n}{2}$ terms of a Prüfer code. If there are less than $\frac{n}{4}$ distinct labels, then the full code has less than $\frac{3n}{4}$ distinct labels and hence corresponds to a tree with at least $\frac{n}{100}$ leaves. Otherwise, there are at least $\frac{n}{4}$ distinct labels among the first $\frac{n}{2}$ terms. We now count the number of times that these labels are used in the subsequent $\frac{n}{2}-2$ terms. Each term is chosen uniformly at random from the set of all n vertex labels, hence has probability at least $\frac{1}{4}$ of repeating a label used in the first $\frac{n}{2}$ terms. Thus the expected number of repeated labels is at least $\frac{1}{4}(\frac{n}{2}-2)\geq\frac{n}{16}$. By the Chernoff bound, Theorem 5, with $\delta=\frac{1}{2}$ the probability that less than $\frac{n}{32}$ repeated labels occur is at most $\exp\big(-\frac{n}{500}\big)$.

For (ii), we note that a vertex has degree equal to one plus the number of its appearances in the Prüfer code. Thus a vertex has degree exceeding $\frac{cn}{\log n}$ only if its label appears $\frac{cn}{\log n}$ times in the Prüfer code. For a given vertex label and choice of $\frac{cn}{\log n}$ terms of the Prüfer code, the probability that each of the chosen terms is equal to the given label is $n^{-\frac{cn}{\log n}}=e^{-cn}$. Taking the union bound over the choices of vertex label and terms of the code, the probability that some vertex label appears at least $\frac{cn}{\log n}$ times is at most

$$n\cdot\binom{n-2}{\frac{cn}{\log n}}\cdot e^{-cn}\leq n\cdot\big(\tfrac{n\log n}{cn}\big)^{\frac{cn}{\log n}}e^{-cn}=\exp\Big(\log n+\tfrac{cn\log(c^{-1}\log n)}{\log n}-cn\Big)\leq e^{-cn/2}\,,$$

where the final inequality is valid for all sufficiently large n. □

It is well known that a degenerate graph contains a large independent set of vertices with the same small degree.

Lemma 47 (Lemma 8 of [1])**.** *Let G be a D-degenerate n-vertex graph. Then there exists an integer $0\leq d\leq 2D$ and a set $I\subseteq V(G)$ with $|I|\geq(2D+1)^{-3}n$ which is independent, and all of whose vertices have the same degree d in G.*

In [1] this result was used to show that one can modify a degeneracy order slightly to move such an independent set to the end of the order while not increasing the degeneracy by much. We repeat the straightforward argument here for completeness.

Proof of Lemma 9. By Lemma 47 there is an independent set I in G of $\lceil (2D+1)^{-3}n \rceil$ vertices, each of which has degree d. Now pick a D-degeneracy order of G and then modify this order by moving all vertices of I to the end (in an arbitrary order). Since all vertices in I have degree $d \leq 2D$ the resulting order is a $2D$-degeneracy order. $\square$

Finally, we provide a proof of the compression lemma.

Proof of Lemma 10. Given the family (G_i), repeatedly perform the following operation, packing two members of the family into one graph. If the current family contains two graphs G, G' which have at most $\frac{2}{3}n$ vertices of degree at least 1 and at most $\frac{1}{3}n$ vertices of degree at least 2 then pack G and G' into a graph G'' as follows and then remove G, G' from the family and add G'' instead.

To define an embedding ϕ of G' into the complement $\bar{G}$ of G, let $A, B, C \subseteq V(G)$ be a partition of $V(G)$ into sets of size either $\lfloor \frac{n}{3} \rfloor$ or $\lceil \frac{n}{3} \rceil$, such that $|A| = |C|$, and such that $\deg_G(x) = 0$ for all $x \in A$ and $\deg_G(x) \leq 1$ for all $x \in B$, which is possible by our assumptions on G. Analogously, let $A', B', C' \subseteq V(G')$ be a partition of $V(G')$ into sets of size either $\lfloor \frac{n}{3} \rfloor$ or $\lceil \frac{n}{3} \rceil$, such that $|A'| = |C'|$ and such that $\deg_{G'}(x) = 0$ for all $x \in A'$ and $\deg_{G'}(x) \leq 1$ for all $x \in B'$. We construct ϕ by first finding a packing of $G[B]$ and $G'[B']$ (which is easy since these two graphs are matchings and $|B| \geq 3$) and then extending this by arbitrarily mapping A' to C and C' to A. Note that by construction $|A| = |A'| = |C| = |C'|$. Clearly, this indeed gives a packing of G and G' since vertices in A have degree 0 in G and vertices in A' have degree 0 in G'. Moreover, $\Delta(G'') \leq \max\{2, \Delta\}$, and G'' is $\max\{2, D\}$-degenerate by construction.

We stop combining graphs in this way when at most one graph with at most $\frac{2}{3}n$ vertices of degree at least 1 and at most $\frac{1}{3}n$ vertices of degree at least 2 remains; we call the resulting family $(\check{G}_i)_{i \in [\check{m}]}$. In this family, all graphs $\check{G}_i$ but possibly one graph satisfy at least one of the following conditions:

- $\check{G}_i$ has more than $\frac{2}{3}n$ vertices of degree at least 1,
- $\check{G}_i$ has more than $\frac{1}{3}n$ vertices of degree at least 2.

In either case $e(\check{G}_i) \geq \frac{1}{3}n$, and therefore we conclude from $\sum_{i=1}^{\check{m}} e(\check{G}_i) = \sum_{i=1}^{m} e(G_i) \leq \binom{n}{2}$ that

$$\check{m} \leq 1 + \frac{\binom{n}{2}}{\frac{1}{3}n} \leq \frac{3}{2}n\,.$$

$\square$

Appendix B. Proof of the matching lemma

In this section we provide the proof of Lemma 20. The proof of this lemma is the only place in this paper where we use the concept of a regular pair.

Definition 48 (density, (ε, d)-regular, (ε, d)-super-regular)**.** *Let G be a graph and $U, W \subseteq V(G)$ be disjoint vertex sets. The* density *of the bipartite graph $G[U, W]$ is*

$$d_G(U, W) = \frac{e(G[U, W])}{|U||W|}\,.$$

We say that $G[U,W]$ *is* (ε,d)-regular *if for all* $U'\subseteq U$ *and* $W'\subseteq W$ *with* $|U'|\geq\varepsilon|U|$ *and* $|W'|\geq\varepsilon|W|$ *we have*

$$d_G(U',W')=d\pm\varepsilon\,.$$

The graph $G[U,W]$ *is* (ε,d)-super-regular *if it is* (ε,d)*-regular and for all* $u\in U$ *and for all* $w\in W$ *we have*

$$\deg_{G[U,W]}(u)=(d\pm\varepsilon)|W|,\qquad \textit{and}\qquad \deg_{G[U,W]}(w)=(d\pm\varepsilon)|U|\,.$$

It is well-known that regular pairs are forced by a degree-codegree condition; we use the following formulation due to Duke, Lefmann, and Rödl in [4].

Lemma 49 (degree-codegree condition [4])**.** *Assume* $0<\varepsilon<2^{-200}$ *and let* $G[U,W]$ *be a bipartite graph with parts* U *and* W *of size* $|U|=|W|=n$ *and density* $d=d_{G[U,W]}(U,W)$*. If*

(*i*) $\deg_{G[U,W]}(u)>(d-\varepsilon)|W|$ *for all* $u\in U$*, and*

(*ii*) $\deg_{G[U,W]}(u,u')<(d+\varepsilon)^2|W|$ *for all but at most* $2\varepsilon|U|^2$ *pairs* $\{u,u'\}\in\binom{U}{2}$*,*

then $G[U,W]$ *is* $(\varepsilon^{\frac{1}{6}},d)$*-regular.* □

If we choose a perfect matching uniformly at random in a super-regular pair then each edge is roughly equally likely to appear in the matching, as was shown by Joos (see [17]).

Lemma 50 (perfect matchings in super-regular pairs [17, Theorem 4.3])**.** *Assume* $0\ll\frac{1}{m'}\ll\varepsilon'\ll d\ll 1$*. Let* $G[U,W]$ *be an* (ε',d)*-super-regular graph with* $|U|=|W|=m'$*. Then* $G[U,W]$ *contains a perfect matching. Moreover, for a perfect matching* σ *chosen uniformly at random among all perfect matchings in* $G[U,W]$ *and for all* $uw\in E(G)$ *we have*

$$\mathbb{P}[\sigma(u)=w]=(1\pm(\varepsilon')^{\frac{1}{20}})\frac{1}{dm'}\,. \qquad \square$$

The proof of the matching lemma simply combines these two lemmas.

Proof of Lemma 20. By $(M1)$ and $(M3)$, for all $x\in U\cup W$ we have

$$\deg_{F'}(x)=\big(\mu\pm\tfrac{200p}{\mu^2}\big)m\,. \tag{24}$$

By $(M2)$ and $(M3)$, for all but at most $\frac{m^2}{\log m}$ pairs $\{u,u'\}\in\binom{U}{2}$ we have

$$\deg_{F'}(u,u')=\big(\mu^2\pm\tfrac{300p}{\mu^2}\big)m\,. \tag{25}$$

We want to apply Lemma 49 with $d=\mu$ and $\varepsilon=400p/\mu^3$ to conclude that $F'[U,W]$ is super-regular, and now check the conditions of this lemma. By (24), for $u\in U$ we have

$$\deg_{F'}(u)=\big(d\pm\tfrac{200p}{\mu^2}\big)m=\big(d\pm\tfrac{200p}{\mu^2}\big)\frac{|W|}{1\pm p}=\big(d\pm\tfrac{400p}{\mu^2}\big)|W|>(d-\varepsilon)|W|\,,$$

and similarly for $w\in W$ we have $\deg_{F'}(w)=\big(d\pm\tfrac{400p}{\mu^2}\big)|U|>(d-\varepsilon)|U|$. By (25), for all but at most $\frac{m^2}{\log m}$ pairs $\{u,u'\}\in\binom{U}{2}$ we have

$$\begin{aligned}\deg_{F'}(u,u')&\leq\big(d^2+\tfrac{300p}{\mu^2}\big)m\leq\big(d^2+\tfrac{300p}{\mu^2}\big)\frac{|W|}{1-p}\leq\big(d^2+\tfrac{400p}{\mu^2}\big)|W|<(d^2+2\varepsilon d+\varepsilon^2)|W|\\&=(d+\varepsilon)^2|W|\,,\end{aligned}$$

where the last inequality uses $d = \mu$ and $\varepsilon = 400p/\mu^3$. We conclude that, if $\frac{m^2}{\log m} \leq 2\varepsilon|U|^2$ which holds for $\log m > 1/\varepsilon$, then F' is $\big((400\frac{p}{\mu^3})^{\frac{1}{6}}, \mu\big)$-regular by Lemma 49. Since $\deg_{F'}(x) = \big(d \pm \frac{400p}{\mu^2}\big)|U|$ for all $x \in U \cup W$, it follows that F' is $\big((400\frac{p}{\mu^3})^{\frac{1}{6}}, \mu\big)$-super-regular.

Hence we can apply Lemma 50 to F' with

$$m' = |U| = (1 \pm p)m\,, \quad \varepsilon' = \Big(400\frac{p}{\mu^3}\Big)^{\frac{1}{6}}\,, \quad \text{and} \quad d = \mu\,,$$

and conclude that F' has a perfect matching and that for a perfect matching σ chosen uniformly at random among all perfect matchings of F' and for all $uw \in E(F')$ we have

$$\mathbb{P}[\sigma(u) = w] = \Big(1 \pm \Big(400\frac{p}{\mu^3}\Big)^{\frac{1}{120}}\Big)\frac{1}{\mu(1 \pm p)m} \leq \frac{2}{\mu m}\,,$$

where the inequality holds if p is small and $\big(400\frac{p}{\mu^3}\big)^{\frac{1}{120}} \leq \frac{1}{100}$. $\square$

(PA) London School of Economics, Department of Mathematics, Houghton Street, London WC2A 2AE, UK
Email address: `p.d.allen@lse.ac.uk`

(JB) London School of Economics, Department of Mathematics, Houghton Street, London WC2A 2AE, UK
Email address: `j.boettcher@lse.ac.uk`

(DC) Technische Universität Hamburg, Institut für Mathematik, Am Schwarzenberg-Campus 3, 21073 Hamburg, Germany
Email address: `dennis.clemens@tuhh.de`

(AT) Technische Universität Hamburg, Institut für Mathematik, Am Schwarzenberg-Campus 3, 21073 Hamburg, Germany
Email address: `taraz@tuhh.de`