

# Tilings in randomly perturbed graphs: bridging the gap between Hajnal–Szemerédi and Johansson–Kahn–Vu

Jie Han\*, Patrick Morris† and Andrew Treglown‡

July 28, 2020

## Abstract

A perfect  $K_r$ -tiling in a graph  $G$  is a collection of vertex-disjoint copies of  $K_r$  covering all the vertices in  $G$ . In this paper we consider perfect  $K_r$ -tilings in the setting of randomly perturbed graphs; a model introduced by Bohman, Frieze and Martin [7] where one starts with a dense graph and then adds  $m$  random edges to it. Specifically, given any fixed  $0 < \alpha < 1 - 1/r$  we determine how many random edges one must add to an  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq \alpha n$  to ensure that, asymptotically almost surely, the resulting graph contains a perfect  $K_r$ -tiling. As one increases  $\alpha$  we demonstrate that the number of random edges required ‘jumps’ at regular intervals, and within these intervals our result is best-possible. This work therefore closes the gap between the seminal work of Johansson, Kahn and Vu [25] (which resolves the purely random case, i.e.,  $\alpha = 0$ ) and that of Hajnal and Szemerédi [18] (showing that when  $\alpha \geq 1 - 1/r$  the initial graph already houses the desired perfect  $K_r$ -tiling).

## 1 Introduction

A significant facet of both extremal graph theory and random graph theory is the study of embeddings. In the setting of random graphs, one is interested in the threshold for the property that  $G(n, p)$  *asymptotically almost surely* (a.a.s.) contains a fixed (spanning) subgraph  $F$ . Meanwhile, a classical line of inquiry in extremal graph theory is to determine the minimum degree threshold that ensures a graph  $G$  contains a fixed (spanning) subgraph  $F$ . A much studied problem in both the extremal and random settings concerns the case when  $F$  is a so-called *perfect  $H$ -tiling*. In this paper we bridge the gap between the random and extremal models for the problem of perfect clique tilings.

### 1.1 Perfect tilings in graphs

Given two graphs  $H$  and  $G$ , an  *$H$ -tiling* in  $G$  is a collection of vertex-disjoint copies of  $H$  in  $G$ . An  *$H$ -tiling* is called *perfect* if it covers all the vertices of  $G$ . Perfect  *$H$ -tilings* are also referred to as  *$H$ -factors*,  *$H$ -matchings* or *perfect  $H$ -packings*. Note that a perfect  *$H$ -tiling* is a generalisation

---

\*University of Rhode Island, Kingston, RI, USA, [jie.han@uri.edu](mailto:jie.han@uri.edu).

†Freie Universität Berlin, Germany and Berlin Mathematical School, Germany, [pm0041@mi.fu-berlin.de](mailto:pm0041@mi.fu-berlin.de). Research supported by a Leverhulme Trust Study Abroad Studentship (SAS-2017-052\9).

‡University of Birmingham, United Kingdom, [a.c.treglown@bham.ac.uk](mailto:a.c.treglown@bham.ac.uk). Research supported by EPSRC grant EP/M016641/1.

of the notion of a perfect matching; indeed, perfect matchings correspond to the case when  $H$  is a single edge.

One of the cornerstone results in extremal graph theory is the Hajnal–Szemerédi theorem [18] which determines the minimum degree threshold that ensures a graph contains a perfect  $K_r$ -tiling.

**Theorem 1.1** (Hajnal and Szemerédi [18]). *Every graph  $G$  whose order  $n$  is divisible by  $r$  and whose minimum degree satisfies  $\delta(G) \geq (1 - 1/r)n$  contains a perfect  $K_r$ -tiling. Moreover, there is an  $n$ -vertex graph  $G$  with  $\delta(G) = (1 - 1/r)n - 1$  that does not contain a perfect  $K_r$ -tiling.*

Earlier, Corrádi and Hajnal [10] proved Theorem 1.1 in the case when  $r = 3$ . See [28] for a short proof of Theorem 1.1.

Since the proof of Theorem 1.1 there have been many generalisations obtained in several directions. For example, Kühn and Osthus [38] characterised, up to an additive constant, the minimum degree which ensures that a graph  $G$  contains a perfect  $H$ -tiling for an *arbitrary* graph  $H$ . Keevash and Mycroft [27] proved an analogue of the Hajnal–Szemerédi theorem in the setting of  *$r$ -partite graphs*, whilst there are now several generalisations of Theorem 1.1 in the setting of *directed graphs* (see e.g. [11, 12, 48]). See [37] for a survey including many of the results on graph tiling. There has also been significant interest in tiling problems in hypergraphs; the survey of Zhao [49] describes many of the results in the area.

## 1.2 Perfect tilings in random graphs

Recall that the random graph  $G(n, p)$  consists of vertex set  $[n] := \{1, \dots, n\}$  where each edge is present with probability  $p$ , independently of all other choices. In the early 1990s, the problem of determining the threshold for the property that  $G(n, p)$  contains a perfect  $H$ -tiling attracted the attention of Erdős (see the appendix of [1]). Indeed, as well as raising the general problem, Erdős particularly focused on the case when  $H = K_3$  stating that ‘the correct answer will be probably about  $n^{4/3}$  edges’, though cautioned that ‘the lack of analogs to Tutte’s theorem may cause serious trouble’. This caution turned out to be well-founded as for a number of years even the case of triangles remained quite stubborn, despite some partial results towards it [29, 33]. However, in 2008, spectacular work of Johansson, Kahn and Vu [25] not only resolved the problem for perfect  $K_3$ -tilings, but the general problem of perfect  $H$ -tilings for all so-called *strictly balanced* graphs  $H$ . Below we state their result only in the case of perfect clique tilings.

**Theorem 1.2** (Johansson, Kahn and Vu [25]). *Let  $n \in \mathbb{N}$  be divisible by  $r \in \mathbb{N}$  where  $r \geq 3$ .*

- *If  $p = \omega(n^{-2/r}(\log n)^{2/(r^2-r)})$  then a.a.s.  $G(n, p)$  contains a perfect  $K_r$ -tiling.*
- *If  $p = o(n^{-2/r}(\log n)^{2/(r^2-r)})$  then a.a.s.  $G(n, p)$  does not contain a perfect  $K_r$ -tiling.*

## 1.3 The model of randomly perturbed graphs

Bohman, Frieze and Martin [7] introduced a model which provides a connection between the extremal and random graph settings. In their model one starts with a dense graph and then adds  $m$  random edges to it. A natural problem in this setting is to determine *how many* random edges are required to ensure that the resulting graph a.a.s. contains a given graph  $F$  as a spanning subgraph. For example, the main result in [7] states that for every  $\alpha > 0$ , there is a  $c = c(\alpha)$  such that if we start with an arbitrary  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq \alpha n$  and add  $cn$  random edges

to it, then a.a.s. the resulting graph is Hamiltonian. This result characterises how many random edges we require for *every* fixed  $\alpha > 0$ . Indeed, if  $\alpha \geq 1/2$  then Dirac's theorem implies that we do not require any random edges; that is any  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq \alpha n$  is already Hamiltonian. On the other hand, if  $0 < \alpha < 1/2$  then the following example implies that we indeed require a linear number of random edges: Let  $G'$  be the complete bipartite graph with vertex classes of size  $\alpha n, (1 - \alpha)n$ . It is easy to see that if one adds fewer than  $(1 - 2\alpha)n$  (random) edges to  $G'$ , the resulting graph is *not* Hamiltonian.

In recent years, a range of results have been obtained concerning embedding spanning subgraphs into a randomly perturbed graph, as well as other properties of the model; see e.g. [2, 3, 5, 6, 8, 9, 13, 14, 26, 30, 35, 36]. The model has also been investigated in the setting of directed graphs and hypergraphs (see e.g. [4, 23, 34, 41]). Much of this work has focused on the range where the minimum degree of the deterministic graph is linear but with respect to some arbitrarily small constant  $\alpha$ . In this range, one thinks of the deterministic graph as 'helping'  $G(n, p)$  to get a certain spanning structure and the observed phenomenon is usually a decrease in the probability threshold of a logarithmic factor, as is the case for Hamiltonicity as above. Recently, there has been interest in the other extreme, where one starts with a minimum degree slightly less than the extremal minimum degree threshold for a certain spanning structure and requires a small 'sprinkling' of random edges to guarantee the existence of the spanning structure in the resulting graph, see e.g. [14, 45].

Balogh, Treglown and Wagner [3] considered the perfect  $H$ -tiling problem in the setting of randomly perturbed graphs. Indeed, for every fixed graph  $H$  they determined how many random edges one must add to a graph  $G$  of linear minimum degree to ensure that a.a.s.  $G \cup G(n, p)$  contains a perfect  $H$ -tiling. Again we only state their result in the case of perfect clique tilings.

**Theorem 1.3** (Balogh, Treglown and Wagner [3]). *Let  $r \geq 2$ . For every  $\alpha > 0$ , there is a  $C = C(\alpha, r) > 0$  such that if  $p \geq Cn^{-2/r}$  and  $(G_n)_{n \in r\mathbb{N}}$  is a sequence of graphs with  $|G_n| = n$  and minimum degree  $\delta(G_n) \geq \alpha n$  then a.a.s.  $G_n \cup G(n, p)$  contains a perfect  $K_r$ -tiling.*

Theorem 1.3, unlike Theorem 1.2, does not involve a logarithmic term. Thus comparing the randomly perturbed model with the random graph model, we see that starting with a graph of linear minimum degree instead of the empty graph saves a logarithmic factor in terms of how many random edges one needs to ensure the resulting graph a.a.s. contains a perfect  $K_r$ -tiling. Further, Theorem 1.3 is best-possible in the sense that given any  $0 < \alpha < 1/r$ , there is a constant  $c = c(\alpha, r) > 0$  and sequence of graphs  $(G_n)_{n \in r\mathbb{N}}$  where  $G_n$  is an  $n$ -vertex graph with minimum degree at least  $\alpha n$  so that a.a.s.  $G_n \cup G(n, p)$  does not contain a perfect  $K_r$ -tiling when  $p \leq cn^{-2/r}$  (see Section 2.1 in [3]). However, as suggested in [3], this still leaves open the question of how many random edges one requires if  $\alpha > 1/r$ .

In this paper we give a sharp answer to this question. Before we can state our result we introduce some notation.

**Definition 1.4.** [*Perturbed perfect tiling threshold*] *Given some  $0 \leq \alpha \leq 1$ , and a graph  $H$  of order  $h$ , the perturbed perfect tiling threshold  $p(H, \alpha)$  satisfies the following.*

- (i) *If  $p = p(n) = \omega(p(H, \alpha))$ , then for any sequence  $(G_n)_{n \in h\mathbb{N}}$  of  $n$ -vertex graphs with  $\delta(G_n) \geq \alpha n$ , the graph  $G_n \cup G(n, p)$  a.a.s. contains a perfect  $H$ -tiling.*
- (ii) *If  $p = p(n) = o(p(H, \alpha))$ , for some sequence  $(G_n)_{n \in h\mathbb{N}}$  of  $n$ -vertex graphs with  $\delta(G_n) \geq \alpha n$ , the graph  $G_n \cup G(n, p)$  a.a.s. does not contain a perfect  $H$ -tiling.*

If it is the case that every sufficiently large  $n$ -vertex graph of minimum degree at least  $\alpha n$  contains a perfect  $H$ -tiling we define  $p(H, \alpha) := 0$ . We say the threshold  $p(H, \alpha)$  is sharp if there are constants  $C(H, \alpha), D(H, \alpha) > 0$  such that (i) remains valid with  $p \geq Cp(H, \alpha)$  and (ii) is satisfied for any  $p \leq Dp(H, \alpha)$ .

Thus, Theorem 1.1 implies that  $p(K_r, \alpha) = 0$  for all  $\alpha \geq 1 - 1/r$  whilst Theorem 1.2 precisely states that  $p(K_r, 0) = n^{-2/r}(\log n)^{2/(r^2-r)}$  (actually Theorem 2.3 in [25] and Theorem 3.22(ii) in [24] imply this threshold is sharp). Our main result deals with the intermediate cases (i.e. when  $0 < \alpha < 1 - 1/r$ ).

**Theorem 1.5.** *Let  $2 \leq k \leq r$  be integers. Then given any  $1 - \frac{k}{r} < \alpha < 1 - \frac{k-1}{r}$ ,*

$$p(K_r, \alpha) = n^{-2/k}.$$

*Moreover, the threshold  $p(K_r, \alpha)$  is sharp.*

Thus, Theorem 1.5 provides a bridge between the Hajnal–Szemerédi theorem and the Johansson–Kahn–Vu theorem. Notice that the value of  $p(K_r, \alpha)$  demonstrates a ‘jumping’ phenomenon; given a fixed  $k$  the value of  $p(K_r, \alpha)$  is the same for all  $\alpha \in ((r-k)/r, (r-k+1)/r)$ , however if  $\alpha$  is just above this interval the value of  $p(K_r, \alpha)$  is significantly smaller.

Note in the case when  $k = r$ , Theorem 1.5 is implied by the results from [3]; whilst finalising the paper we learned of a very recent result [45] concerning powers of Hamilton cycles in randomly perturbed graphs which implies the case when  $k = 2$  and  $r$  is even. To help provide some intuition for Theorem 1.5, note that  $n^{-2/k}$  is the threshold for the property that  $G(n, p)$  contains a copy of  $K_k$  in every linear sized subset of vertices; this property will be exploited throughout the proof. Our proof uses the absorption method, and in particular the novel ‘absorption reservoir method’ introduced by Montgomery [42, 43], where we use a robust sparse bipartite graph, which we call a template, in order to build an absorbing structure in our graph. We also use ‘reachability’ arguments, introduced by Lo and Markstöm [40], in order to build absorbing structures. We use various probabilistic techniques throughout, such as multi-round exposure, and we appeal to Szemerédi’s regularity lemma in order to obtain an ‘almost tiling’.

The paper is organised as follows. In the next section we introduce some fundamental tools that will be applied in the proof of Theorem 1.5. Section 3 then contains the construction that provides the lower bound on  $p(K_r, \alpha)$  in Theorem 1.5. In Section 4 we give an overview of our proof for the upper bound on  $p(K_r, \alpha)$  in Theorem 1.5, which is given in Section 7 after developing the necessary theory in Section 5 and Section 6. Finally some open problems are raised in the concluding remarks section (Section 8).

**Additional Note:** Since the paper was first submitted there have been some related results proven. Indeed, Knierim and Su [30], expanding on work of Nenadov and Pehova [44], considered the so-called Ramsey–Turán problem for clique tilings. They showed that for any  $\alpha > 1 - \frac{2}{r}$ , there exists an  $\eta > 0$  such that if  $G'$  is a graph with  $\delta(G') \geq \alpha n$  and independence number less than  $\eta n$ , then  $G'$  contains a perfect  $K_r$ -tiling. This implies Theorem 1.5 for  $k = 2$  and all  $r$  as if  $G$  has minimum degree  $\alpha n$  with  $\alpha$  as above and  $p = \omega(n^{-1})$  then  $G(n, p)$  (and hence  $G' = G \cup G(n, p)$ ) has sublinear independence number. In a different direction, Antoniuk, Dudek, Reiher, Ruciński and Schacht [2] explored the appearance of powers of Hamilton cycles in randomly perturbed graphs, building on the work of Nenadov and Trujić [45]. As the  $k^{\text{th}}$  power of a Hamilton cycle is a supergraph of a perfect  $K_{k+1}$ -tiling, their work gives bounds for the existence of clique tilings.

They focus solely on  $G$  with minimum degree  $\alpha n$  where  $\alpha = j/(j+1) + \epsilon$  for some  $j \in \mathbb{N}$ ,  $j \geq 1$  and they obtain tight results in certain cases. The only range where their implied results on clique tilings is tight with regards to the threshold obtained and the minimum degree condition is the case when  $k = 2$  and  $r$  is even in Theorem 1.5, already implied by [45] and [30] as mentioned above.

## 2 Notation and preliminaries

We use standard graph theory notation throughout. In particular we use  $|G|$  to denote the number of vertices of a graph  $G$ . Sometimes we will also write  $v_G$  and  $e_G$  to denote the number of vertices and edges in  $G$  respectively. We write  $N_G(v)$  to denote the neighbourhood of a vertex  $v \in G$ . For a subset of vertices  $V' \subseteq V = V(G)$ ,  $G[V']$  denotes the graph induced by  $G$  on  $V'$  and we use the shorthand  $G \setminus V'$  to denote  $G[V \setminus V']$ . If  $V' = \{x\}$  we simply write  $G \setminus x$ . Further, for disjoint subsets of vertices  $V', V'' \subseteq V$ ,  $G[V', V'']$  denotes the bipartite graph induced by  $G$  on  $V' \cup V''$  by considering only the edges of  $G$  with one endpoint in  $V'$  and the other endpoint in  $V''$ . If  $G'$  is a graph on the same vertex set as  $G$  we write  $G \cup G'$  to denote the graph on vertex set  $V(G)$  with edge set  $E(G) \cup E(G')$ . We write  $G - E(G')$  to be the graph obtained from  $G$  by deleting any edges that also lie in  $G'$ . One key exception to the use of standard notation is our use of  $\overline{H}$ , to denote the complement of  $H$  with respect to a graph which is *not* complete, see Definition 6.2.

We write  $K_{m_1, m_2, \dots, m_r}^r$  to denote the complete  $r$ -partite graph with parts of size  $m_1, \dots, m_r$ . For a graph  $J$  on  $r$  vertices  $\{v_1, \dots, v_r\}$  and  $m_1, \dots, m_r \in \mathbb{N}$ , we define the *blow-up* of  $J$  to be the graph  $J_{m_1, \dots, m_r}$  with vertex set  $P_1 \sqcup P_2 \sqcup \dots \sqcup P_r$ , such that  $|P_i| = m_i$  and for all  $i, j \in [r]$  and  $w \in P_i, w' \in P_j$  we have  $ww' \in E(J_{m_1, \dots, m_r})$  if and only if  $v_i v_j \in E(J)$ . Given a set  $A$  and  $k \in \mathbb{N}$  we denote by  $A^k$  the set of all ordered  $k$ -tuples of elements from  $A$ , while  $\binom{A}{k}$  denotes the set of all (unordered)  $k$ -element subsets of  $A$ . At times we have statements such as the following: ‘‘Choose constants  $0 \ll c_1 \ll c_2 \ll \dots \ll c_k$ ’’. This should be taken to mean that one can choose constants from right to left so that all the subsequent constraints are satisfied. That is, there exist increasing functions  $f_i$  for  $i \in [k]$  such that whenever  $c_i \leq f_{i+1}(c_{i+1})$  for all  $i \in [k-1]$ , all constraints on these constants that are in the proof, are satisfied. Finally, we omit the use of floors and ceilings unless it is necessary, so as not to clutter the arguments.

Throughout, we will deal exclusively with *ordered* embeddings of graphs, which we also refer to as *labelled* embeddings. Thus when we refer to an embedding of  $H$  in  $G$ , we implicitly fix an ordering on  $V(H)$ , say  $V(H) := \{h_1, \dots, h_{v_H}\}$  and say that there is an embedding of  $H$  onto an (ordered) vertex set  $\{v_1, \dots, v_{v_H}\} \subseteq V(G)$  if  $v_i v_j \in E(G)$  for all  $i$  and  $j$  such that  $h_i h_j \in E(H)$ .

In what follows, we introduce the tools that we will use in our proofs. Most of these are well known and so are stated without proof. One exception is Lemma 2.8, which is tailored to our purposes and slightly technical (but follows from well-known techniques nonetheless).

### 2.1 A deterministic tiling result

Let  $\chi_{cr}(H) := (\chi(H) - 1) \frac{|H|}{|H| - \sigma(H)}$  where  $\sigma(H)$  is the smallest size of a colour class over all colourings of  $H$  with  $\chi(H)$  colours. The following result of Komlós [31] is a crucial tool in the proof of Theorem 1.5. It determines the minimum degree threshold for the property of containing an ‘almost’ perfect  $H$ -tiling.

**Theorem 2.1.** *For every graph  $H$  and every  $\alpha > 0$ , there exists  $n_0$  such that if  $G$  is a graph on  $n \geq n_0$  vertices with  $\delta(G) \geq \left(1 - \frac{1}{\chi_{cr}(H)}\right)n$ , then  $G$  contains an  $H$ -tiling which covers all but at most  $\alpha n$  vertices of  $G$ .*

This was later improved to a constant number of uncovered vertices by Shokoufandeh and Zhao [46], but Komlós' result suffices for our purposes. We will apply Komlós' theorem to find an almost perfect  $H$ -tiling in a reduced graph  $R$  of our (deterministic) graph  $G$  from Theorem 1.5; here  $H$  will be a carefully chosen auxiliary graph (not  $K_r$ !). We discuss this further in the proof overview section.

## 2.2 Regularity

We will use the famous regularity lemma due to Szemerédi [47]. The lemma and its consequences appeared in the form we give here, in a survey of Komlós and Simonovits [32], which we also recommend for further details on the subject. First we introduce some necessary terminology. Let  $G$  be a bipartite graph with bipartition  $\{A, B\}$ . For non-empty sets  $X \subseteq A, Y \subseteq B$ , we define the *density* of  $G[X, Y]$  to be  $d_G(X, Y) := e(G[X, Y]) / (|X||Y|)$ . We say that  $G$  is  $\epsilon$ -regular for some  $\epsilon > 0$  if for all sets  $X \subseteq A$  and  $Y \subseteq B$  with  $|X| \geq \epsilon|A|$  and  $|Y| \geq \epsilon|B|$  we have

$$|d_G(A, B) - d_G(X, Y)| < \epsilon.$$

It is also common, when the underlying graph  $G$  is clear, to refer to  $(A, B)$  as an  $\epsilon$ -regular pair.

We will use the following two well-known results in our proof. The so-called slicing lemma shows that regularity is hereditary, with slightly weaker parameters.

**Lemma 2.2** (Slicing lemma [32, Fact 1.5]). *Let  $G$  be  $\epsilon$ -regular on parts  $\{A, B\}$  with density  $d$  and let  $\alpha > \epsilon$ . Let  $A' \subseteq A$  and  $B' \subseteq B$  with  $|A'| \geq \alpha|A|$  and  $|B'| \geq \alpha|B|$ . Then  $G[A', B']$  is  $(2\epsilon/\alpha)$ -regular with density at least  $d - \epsilon$ .*

The next lemma is an extremely useful tool, extending the control on the edge count in regular pairs to be able to count the number of embeddings of small subgraphs.

**Lemma 2.3** (Counting lemma [32, Theorem 2.1]). *Given  $d > \epsilon > 0$ ,  $m \in \mathbb{N}$  and  $H$  some fixed graph on  $r$  vertices, let  $G$  be a graph obtained by replacing every vertex  $x_i$  of  $H$  with an independent set  $V_i$  of size  $m$  and every edge of  $H$  with an  $\epsilon$ -regular pair of density at least  $d$  on the corresponding sets. If  $\epsilon \leq \frac{d^r}{(2+r)2^r} =: d_0$ , then there are at least  $(d_0 m)^r$  embeddings of  $H$  in  $G$  so that each  $x_i$  is embedded into the set  $V_i$ .*

We now turn to the regularity lemma, which tells us that there is a way to partition *any* large enough graph in such a way that the graph induces  $\epsilon$ -regular pairs on almost all of the pairs of parts in this partition. Actually, we apply a variant of the lemma which ensures that, ignoring a small number of edges and a small exceptional set of vertices, *all* such pairs are  $\epsilon$ -regular.

**Lemma 2.4** (Degree form of the regularity lemma [32, Theorem 1.10]). *Let  $0 < \epsilon < 1$  and  $m_0 \in \mathbb{N}$ . Then there is an  $N = N(\epsilon, m_0)$  such that the following holds for every  $0 \leq d < 1$  and for every graph  $G$  on  $n \geq N$  vertices. There exists a partition  $\{V_0, V_1, \dots, V_m\}$  of  $V(G)$  and a spanning subgraph  $G'$  of  $G$  satisfying the following:*

1.  $m_0 \leq m \leq N$ ;



2.  $|V_0| \leq \epsilon n$  and  $|V_1| = \dots = |V_m| =: n' \leq \epsilon n$ ;
3. for each  $v \in V(G)$ ,  $d_{G'}(v) > d_G(v) - (d + \epsilon)n$ ;
4. for all pairs  $V_i, V_j$ , where  $1 \leq i < j \leq k$ , the graph  $G'[V_i, V_j]$  is  $\epsilon$ -regular and has density either 0 or greater than  $d$ .

The sets  $V_1, \dots, V_m$  are called *clusters*,  $V_0$  the *exceptional set* and the vertices in  $V_0$  are *exceptional vertices*.

The degree condition (3.) in Lemma 2.4 guarantees that the majority of the edges of  $G$  lie in  $G'$ . To make this more transparent it is useful to focus on the dense  $\epsilon$ -regular pairs and define the following auxiliary graph. The  $(\epsilon, d)$ -*reduced graph*  $R$  is as follows: The vertex set of  $R$  is the set of clusters  $\{V_i : i \in [m]\}$  and for each  $U, U' \in V(R)$ ,  $UU'$  is an edge of  $R$  if the subgraph  $G'[U, U']$  is  $\epsilon$ -regular and has density greater than  $d$ . The following then follows easily from Lemma 2.4.

**Corollary 2.5.** *Suppose that  $0 < \epsilon \leq d \leq c$  are constants. Let  $G$  be a graph on  $n$  vertices and  $\delta(G) \geq cn$ . Suppose that  $G$  has a partition  $\mathcal{P} = \{V_0, V_1, \dots, V_m\}$  and a subgraph  $G' \subseteq G$  as given by Lemma 2.4 and corresponding  $(\epsilon, d)$ -reduced graph  $R$ . Then  $\delta(R) \geq (c - d - 2\epsilon)m$ .*

### 2.3 Supersaturation

The following phenomenon was first noticed by Erdős and Simonovits in their seminal paper [15]. It states that if there are many copies of a given small subgraph in some host graph, then we can also find many copies of a blow-up in the host graph. It can be proven easily e.g. by induction.

**Lemma 2.6.** *Let  $r, m_1, m_2, \dots, m_r \in \mathbb{N}$ , let  $J$  be some graph on  $r$  vertices  $\{v_1, \dots, v_r\}$  and  $c > 0$ . Then there exists  $c' = c'(r, m_1, m_2, \dots, m_r, c) > 0$  such that the following holds. Suppose  $G$  is a graph on  $n$  vertices with  $n$  sufficiently large such that there are subsets  $V_1, \dots, V_r \subset V(G)$  and  $G$  contains at least  $cn^r$  labelled copies of  $J$  with  $v_i \in V_i$  for  $i = 1, \dots, r$ . Then  $G$  contains at least  $c'n^{m_1 + \dots + m_r}$  labelled copies of  $J_{m_1, m_2, \dots, m_r}$  with parts  $P_1, \dots, P_r$  such that  $P_i \subset V_i$  and  $|P_i| = m_i$ .*

### 2.4 Subgraph counts in random graphs

We first recall Janson's inequality (see e.g. [24, Theorem 2.14]). Let  $\Gamma$  be a finite set and let  $\Gamma_p$  be a random subset of  $\Gamma$  such that each element of  $\Gamma$  is included independently with probability  $p$ . Let  $\mathcal{S}$  be a family of non-empty subsets of  $\Gamma$  and for each  $S \in \mathcal{S}$ , let  $I_S$  be the indicator random variable for the event  $S \subseteq \Gamma_p$ . Thus each  $I_S$  is a Bernoulli random variable  $\text{Be}(p^{|S|})$ . Let  $X := \sum_{S \in \mathcal{S}} I_S$  and  $\lambda := \mathbb{E}(X)$ . Let  $\Delta_X := \sum_{S \cap T \neq \emptyset} \mathbb{E}(I_S I_T)$ , where the sum is over not necessarily distinct ordered pairs  $S, T \in \mathcal{S}$ . Then Janson's inequality states that for any  $0 \leq t \leq \lambda$ ,

$$\mathbb{P}(X \leq \lambda - t) \leq \exp\left(-\frac{t^2}{2\Delta_X}\right). \quad (2.1)$$

Consider the random graph  $G(n, p)$  on an  $n$ -vertex set  $V$ . Note that we can view  $G(n, p)$  as  $\Gamma_p$  with  $\Gamma := \binom{V}{2}$ . Following [24], for a fixed graph  $F$ , we define  $\Phi_F = \Phi_F(n, p) := \min\{n^{v_H} p^{e_H} : H \subseteq F, e_H > 0\}$ . This parameter helps to simplify calculations of  $\Delta_X$  in the context of counting the number of embeddings of the graph  $F$  in  $G(n, p)$ . We will also be interested in the appearance of

graphs in  $G(n, p)$  where we require some subset of vertices to be already fixed in place. Therefore, for a graph  $F$ , and some independent<sup>1</sup> subset of vertices  $W \subset V(F)$ , we define

$$\Phi_{F,W} = \Phi_{F,W}(n, p) := \min\{n^{v_H - v_H[W]} p^{e_H} : H \subseteq F, e_H > 0\}.$$

Note that  $\Phi_F = \Phi_{F, \emptyset}$  and  $\Phi_{F \setminus W} \geq \Phi_{F,W}$  for any  $F$  and independent set  $W \subset V(F)$ . If  $W = \{w\}$  for a single vertex  $w \in V(F)$ , we drop the brackets and simply write  $\Phi_{F,w}$  and  $\Phi_{F \setminus w}$ . Let us collect some more simple observations concerning  $\Phi_F$  and  $\Phi_{F,W}$  which will be useful later.

**Lemma 2.7.** *The following hold:*

1. Let  $C > 1$  be some constant,  $k \in \mathbb{N} \setminus \{1\}$  and  $p = p(n) \geq Cn^{-\frac{2}{k}}$ . Let  $k' \leq k$  and  $F_1 := K_{k'}$ , then we have that  $\Phi_{F_1} \geq Cn$ .
2. As above, let  $C > 1$  be some constant,  $k \in \mathbb{N} \setminus \{1\}$  and  $p = p(n) \geq Cn^{-\frac{2}{k}}$ . Suppose now that  $3 \leq k' \leq k$  and let  $F_2 := K_{k'}^-$  be the complete graph on  $k'$  vertices with one edge missing and let  $w \in V(F_2)$  be one of the endpoints of the missing edge. Then  $\Phi_{F_2} \geq Cn$  and  $\Phi_{F_2,w} \geq \min\{Cn^{1-\frac{2}{k}}, Cn^{\frac{2}{k}}\} \geq Cn^{\frac{1}{k}}$ .
3. Let  $F_3, F_4$  be graphs with vertex subsets  $W_3 \subset V(F_3)$ ,  $W_4 \subset V(F_4)$ , let  $\Phi_3 := \Phi_{F_3, W_3}$  and  $\Phi_4 := \Phi_{F_4, W_4}$  and suppose that  $\Phi_3, \Phi_4 \geq 1$ . Let  $F_5$  be the graph formed by the union of  $F_3$  and  $F_4$  meeting in exactly one vertex  $x \in (V(F_3) \setminus W_3) \cap (V(F_4) \setminus W_4)$ , and let  $F_6$  be the graph obtained by taking a disjoint union of  $F_3$  and  $F_4$ . Then letting  $W_5 := W_3 \sqcup W_4$ , we have that  $\Phi_{F_5, W_5} \geq \min\{\Phi_3, \Phi_4, \Phi_3 \Phi_4 n^{-1}\}$  and  $\Phi_{F_6, W_5} = \min\{\Phi_3, \Phi_4\}$ .

*Proof.* For parts 1 and 2, it suffices to consider the case  $k' = k$ . For part 1, we have a simple calculation. Let  $H$  be a subgraph of  $K_k$  with  $v_H$  vertices and  $e_H$  edges. As  $v_H \leq k$ , we obtain

$$n^{v_H} p^{e_H} \geq n^{v_H} (Cn^{-\frac{2}{k}})^{\frac{v_H(v_H-1)}{2}} \geq Cn^{v_H - (v_H-1)} = Cn.$$

For part 2 first note that as  $F_2 \subseteq K_k = F_1$  we have that  $\Phi_{F_2} \geq \Phi_{F_1} \geq Cn$ . Let  $H$  be a subgraph of  $K_k^-$ . If  $w \notin H$ , the calculation from part 1 gives that  $n^{v_H} p^{e_H} \geq Cn$ . So suppose  $w \in H$ . Now let us distinguish two cases, depending on whether the vertex  $u$  is in  $H$ , where  $u$  is the vertex in  $K_k^-$  such that  $uw$  is a *non*-edge. If  $u \in H$ , we have that

$$n^{v_H-1} p^{e_H} \geq n^{v_H-1} Cn^{-\frac{2}{k}} \left( \frac{v_H(v_H-1)}{2} - 1 \right) \geq Cn^{v_H-1 - (v_H-1 - \frac{2}{k})} \geq Cn^{\frac{2}{k}},$$

again using that  $v_H \leq k$ . Likewise, if  $u \notin H$ , we have that

$$n^{v_H-1} p^{e_H} \geq n^{v_H-1} p^{\binom{v_H}{2}} \geq Cn^{(v_H-1)(1 - \frac{v_H}{k})} \geq Cn^{1 - \frac{2}{k}},$$

where the last inequality follows as  $(v_H - 1) \left(1 - \frac{v_H}{k}\right)$  is minimised in the range  $2 \leq v_H \leq k - 1$  at  $v_H = 2$  and  $v_H = k - 1$ . This shows that  $\Phi_{F_2,w}$  is bounded as desired.

Part 3 also follows from the definition. Indeed, note that one subgraph  $H$  of  $F_6$  that is a minimiser of the term in the definition of  $\Phi_{F_6, W_5}$  must be a subgraph of  $F_3$  or a subgraph of  $F_4$ . This ensures  $\Phi_{F_6, W_5} = \min\{\Phi_3, \Phi_4\}$ . Similarly, one subgraph  $H$  of  $F_5$  that is a minimiser of the term in the definition of  $\Phi_{F_5, W_5}$  must be a subgraph of  $F_3$ , a subgraph of  $F_4$ , or a subgraph of  $F_5$  that contains  $x$ . This ensures  $\Phi_{F_5, W_5} \geq \min\{\Phi_3, \Phi_4, \Phi_3 \Phi_4 n^{-1}\}$ .  $\square$

<sup>1</sup>With respect to  $F$  i.e.  $E(F[W]) = \emptyset$ .



We now apply Janson's inequality in order to give a general result about embedding constant sized graphs into  $G(n, p)$ . The following lemma provides the basis for a greedy process in which we find some larger (linear size) graph in  $G(n, p)$ . We will require that the embedding of our larger graph has certain vertices already prescribed and repeated applications of Lemma 2.8 will then allow us to embed the remaining vertices of the graph in a greedy manner. So it is crucial that we can apply the lemma to *any* subset of  $s$  (remaining) indices while avoiding *any* small enough set of (previously used) vertices from being used.

For future applications, we state and prove the following lemma in the context of  $r$ -uniform hypergraphs ( $r$ -graphs), and the definition of  $\Phi_F$  extends naturally to  $r$ -graphs  $F$  and  $G^{(r)}(n, p)$ . Recall that  $G^{(r)}(n, p)$  is an  $r$ -graph on  $n$  vertices where each  $r$ -tuple of vertices forms an edge with probability  $p$ , independent of all other  $r$ -tuples.

**Lemma 2.8.** *Let  $n, t(n), s(n) \in \mathbb{N}$ ,  $0 < \beta < 1/2$  and  $L, v, w, e, r \in \mathbb{N}$  such that  $r \geq 2$ ,  $Lt, sw \leq \frac{\beta n}{4v}$  and  $\binom{t}{s} \leq 2^n$ . Let  $F_1, \dots, F_t$  be labelled  $r$ -graphs with distinguished vertex subsets  $W_i \subset V(F_i)$  such that  $|W_i| \leq w$ ,  $|V(F_i \setminus W_i)| = v$ ,  $e(F_i) = e$  and  $e(F_i[W_i]) = 0$  for all  $i \in [t]$ . Now let  $V$  be an  $n$ -vertex set and let  $U_1, \dots, U_t \subset V$  be labelled vertex subsets with  $|U_i| = |W_i|$  for all  $i \in [t]$ . Finally, suppose there are families  $\mathcal{F}_1, \dots, \mathcal{F}_t \subset \binom{V}{v}$  of labelled vertex sets such that for each  $i \in [t]$ ,  $|\mathcal{F}_i| \geq \beta n^v$ .*

*Now suppose that  $1 \leq s(n) \leq t(n)$  and  $p = p(n)$  are such that*

$$s \cdot \Phi \geq \left( \frac{2^{v+\tau} v!}{\beta^2} \right) \min\{Lt \log n, n\} \text{ and } \Phi' \geq \left( \frac{2^{v+\tau} v!}{\beta^2} \right) n, \quad (2.2)$$

*where  $\Phi := \min\{\Phi_{F_i, W_i} : i \in [t]\}$  and  $\Phi' := \min\{\Phi_{F_i \setminus W_i} : i \in [t]\}$  with respect to  $p = p(n)$ . Then, a.a.s., for any  $V' \subseteq V$ , with  $|V'| \geq n - Lt$  and any subset  $S \subseteq [t]$  such that  $|S| = s$  and  $U_i \cap U_j = \emptyset$  for  $i \neq j \in [s]$ , there exists some  $i \in S$  such that there is an embedding (which respects labelling) of  $F_i$  in  $G^{(r)}(n, p)$  on  $V$  which maps  $W_i$  to  $U_i$  and  $V(F_i) \setminus W_i$  to a labelled set in  $\mathcal{F}_i$  which lies in  $V'$ .*

Note by 'labelled' here we mean that for all  $j$ , the  $j^{\text{th}}$  vertex in  $W_i$  is mapped to the  $j^{\text{th}}$  vertex in  $U_i$ ; moreover the  $j^{\text{th}}$  vertex in  $V(F_i) \setminus W_i$  is mapped to the  $j^{\text{th}}$  vertex in some labelled set from  $\mathcal{F}_i$ .

*Proof.* Let us fix  $S \subset [t]$  with  $|S| = s$  and a vertex subset  $V' \subset V$  as in the statement of the lemma. Let  $U := \cup_{i \in S} U_i$  and fix  $V'' := V' \setminus U$ . Note that  $(V \setminus V') \cup U$  intersects at most  $\frac{\beta}{2} n^v$  of the elements of  $\mathcal{F}_i$  for each  $i$  and so we can focus on a subset  $\mathcal{F}'_i$  of each  $\mathcal{F}_i$  of at least  $\frac{\beta}{2} n^v$  sets which are all contained in  $V''$ . For each  $i \in S$  and each labelled subset  $X \in \mathcal{F}'_i$ , let  $I_{X,i}$  denote the indicator random variable that  $X \cup U_i$  hosts a labelled copy of  $F_i$  where  $W_i$  is mapped to  $U_i$ . To ease notation sometimes we write  $I_X$  instead of  $I_{X,i}$ . Note that  $Z := \sum\{I_{X,i} : X \in \cup_{i \in S} \mathcal{F}'_i\}$  counts the number of suitable embeddings in  $G^{(r)}(n, p)$ . (So here if  $X$  is in  $a$  of the collections  $\mathcal{F}'_i$ , then there are  $a$  indicator random variables in this sum corresponding to  $X$ .)

An easy calculation (using the first part of (2.2)) gives that  $\mathbb{E}[Z] > 2$  for large enough  $n$ . We will show that

$$\Delta_Z \leq \frac{\mathbb{E}[Z]^2}{16 \min\{Lt \log n, n\}} \quad (2.3)$$

and thus by Janson's inequality (2.1),  $\mathbb{P}[Z \leq \mathbb{E}[Z]/2] \leq \exp(-2 \min\{Lt \log n, n\})$ . If  $Lt \log n \leq n$ , taking a union bound over the (at most  $2^t$ ) possible sets  $S$  and the  $\binom{n}{Lt} \leq \exp(Lt(1 + \log n))$  possible  $V'$ , we have that a.a.s.,  $Z \geq 1$  for all such  $S$  and  $V'$ ; if  $Lt \log n > n$ , we instead bound both the

number of  $V'$  and the number of  $S$  by  $2^n$  and draw the same conclusion. So in both cases  $Z \geq 1$  a.a.s. for all such  $S$  and  $V'$  and we are done.

Now it remains to verify (2.3). Firstly let  $Z_i := \sum_{X \in \mathcal{F}'_i} I_{X,i}$ . Then

$$\mathbb{E}[Z]^2 = \left( \sum_{i \in S} \mathbb{E}[Z_i] \right)^2 = \sum_{i,j \in S} \mathbb{E}[Z_i] \mathbb{E}[Z_j]. \quad (2.4)$$

To ease notation, let  $\mathcal{F} := \cup_{i \in S} \mathcal{F}'_i$  and for  $X, X' \in \mathcal{F}$ , we write  $X \sim X'$  if, assuming  $X \in \mathcal{F}_i$ ,  $X' \in \mathcal{F}_j$ , the labelled copies of  $F_i$  on  $X \cup U_i$  and  $F_j$  on  $X' \cup U_j$  intersect in at least one edge. We split  $\Delta_Z$  as follows:

$$\begin{aligned} \Delta_Z &= \sum_{\{(X,X') \in \mathcal{F}^2: X \sim X'\}} \mathbb{E}[I_X I_{X'}] \\ &= \sum_{i \in S} \Delta_{Z_i} + \sum_{\{(i,j) \in S^2: i \neq j\}} \sum_{\{(X,X') \in \mathcal{F}'_i \times \mathcal{F}'_j: X \sim X'\}} \mathbb{E}[I_X I_{X'}], \end{aligned} \quad (2.5)$$

where  $\Delta_{Z_i}$  is defined analogously to  $\Delta_Z$  for the random variable  $Z_i$ .

For integers  $a$  and  $b$ , write  $(a)_b := a(a-1) \cdots (a-b+1)$ . Fix  $1 \leq k \leq v$ . There are  $\binom{v}{k} (v)_k \leq \binom{v}{k} v!$  ways that two labelled  $v$ -sets share exactly  $k$  vertices. Fixing two such  $v$ -sets, there are at most  $(n)_{2v-k} \leq n^{2v-k}$  ways of mapping their  $2v-k$  vertices into  $V$ . Let  $f_k$  denote the maximum number of edges of a  $k$ -vertex subgraph of  $F_i \setminus W_i$ , taken over all  $i \in [t]$ . As we explain in the next paragraph, we have that for  $i \neq j$ ,

$$\sum_{\{(X,X') \in \mathcal{F}'_i \times \mathcal{F}'_j: X \sim X'\}} \mathbb{E}[I_X I_{X'}] \leq \sum_{k=1}^v \binom{v}{k} v! n^{2v-k} p^{2e-f_k} \leq \frac{2^v v! n^{2v} p^{2e}}{\Phi'} \leq \frac{2^{v+2} v! \mathbb{E}[Z_i] \mathbb{E}[Z_j]}{\beta^2 \Phi'}.$$

Here, we crucially used that any copy of  $F_i$  on  $X \in \mathcal{F}'_i$  does not have edges intersecting  $U_j$  for  $j \neq i$ . Note that the penultimate inequality follows by definition of  $\Phi'$ . The last inequality follows as  $\beta n^v p^e / 2 \leq \mathbb{E}[Z_i]$  for all  $i \in S$ .

Using the above calculation (and the second part of (2.2)) to compare (2.5) and (2.4), we see that the right hand summand of (2.5) is less than  $\mathbb{E}[Z]^2 / (32n)$ . We now estimate the left hand summand of (2.5) in a similar fashion. For a fixed  $i \in S$ , let  $1 \leq k \leq v$ . We let  $g_k$  denote the maximum number of edges of a subgraph of  $F_i$  which has  $k$  vertices disjoint from  $W_i$ . We have, similarly to before, that

$$\Delta_{Z_i} \leq \sum_{k=1}^v \binom{v}{k} (v)_k (n)_{2v-k} p^{2e-g_k} \leq \sum_{k=1}^v \binom{v}{k} v! n^{2v-k} p^{2e-g_k} \leq \frac{2^v v! n^{2v} p^{2e}}{\Phi}.$$

Thus,

$$\sum_{i \in S} \Delta_{Z_i} \leq \frac{s 2^v v! n^{2v} p^{2e}}{\Phi} \stackrel{(2.2)}{\leq} \frac{(s \beta n^v p^e / 2)^2}{32 \min\{Lt \log n, n\}} \leq \frac{(\sum_{i \in S} \mathbb{E}[Z_i])^2}{32 \min\{Lt \log n, n\}} = \frac{\mathbb{E}[Z]^2}{32 \min\{Lt \log n, n\}}.$$

So bringing both summands together, (2.3) holds and we are done.  $\square$

In its full generality, Lemma 2.8 will be a valuable tool in our proof. However, we will also have instances where we do not need to use the full power of the lemma. For instance, setting  $r = 2$ ,  $s = 1$  and  $W_i = U_i = \emptyset$  for all  $i \in [t]$ , we recover a more standard application of Janson's inequality to subgraph containment which we state below for convenience.

**Corollary 2.9.** *Let  $\beta > 0$ ,  $1 \leq t \leq 2^n$  and  $F$  some fixed labelled graph on  $v$  vertices. Then there exists  $C > 0$  such that the following holds. If  $V$  is a set of  $n$  vertices,  $\mathcal{F}_1, \dots, \mathcal{F}_t \subset \binom{V}{v}$  are families of labelled subsets such that  $|\mathcal{F}_i| \geq \beta n^v$  and  $p = p(n)$  is such that  $\Phi_F \geq Cn$ , then a.a.s., for each  $i \in [t]$ , there is an embedding of  $F$  onto a set in  $\mathcal{F}_i$ , which respects labellings.*

### 3 Lower bound construction for the proof of Theorem 1.5

In this section we give a construction that provides the lower bound in the proof of Theorem 1.5. Our construction is a generalisation of that used for the lower bound in Theorem 1.3 (see Section 2.1 of [3]). We will make use of the following result.

**Theorem 3.1** ([24], part of Theorem 4.9). *For every  $k \geq 2$  and for every  $0 < \epsilon < 1$  there is a positive constant  $c = c(r, \epsilon)$  such that if  $p \leq cn^{-2/k}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G(n, p) \text{ contains a } K_k\text{-tiling of size } \epsilon n) = 0.$$

Let  $k$  and  $r$  be in the statement of Theorem 1.5. Consider any  $1 - \frac{k}{r} < \alpha < 1 - \frac{k-1}{r}$  and let  $\gamma > 0$  such that  $(1 - \gamma)(1 - \frac{k-1}{r}) = \alpha$ . Let  $n \in \mathbb{N}$  be divisible by  $r$ . Suppose  $G$  is an  $n$ -vertex graph with vertex classes  $A$  and  $B$  such that  $|B| = (1 - \gamma)(1 - \frac{k-1}{r})n$  and  $|A| = n - |B|$ , where there are all possible edges in  $G$  except that  $A$  is an independent set. So  $\delta(G) \geq \alpha n$ .

Choose  $c' = c'(\gamma, k, r) = c'(\alpha, k, r)$  sufficiently small so that, if  $p = c'n^{-2/k}$  a.a.s.  $G(n, p)[A] \cong G(|A|, p)$  does not contain a  $K_k$ -tiling of size  $\gamma|A|/r \leq \gamma n/r$ . The existence of  $c'$  is guaranteed by Theorem 3.1 since  $A$  has size linear in  $n$ .

Observe that any copy of  $K_r$  in  $G \cup G(n, p)$  either contains a  $K_k$  in  $A$ , or uses at least  $r - (k - 1)$  vertices in  $B$ . Thus, a.a.s., the largest  $K_r$ -tiling in  $G \cup G(n, p)$  has size less than  $|B|/(r - k + 1) + \gamma n/r = n/r$  and we are done.

### 4 Overview of the proof of the upper bound of Theorem 1.5

In this section we sketch some of the ideas in the remainder of our proof of Theorem 1.5. We use the by now well-known absorbing method, which reduces the problem into finding a small absorbing structure on some vertex subset  $A$  and finding a  $K_r$ -tiling that leaves a set  $U$  of  $o(n)$  vertices uncovered. The property of the absorbing structure on  $A$  is that for any small set  $U$  with  $|U| \in r\mathbb{N}$ , one can find a perfect  $K_r$ -tiling in  $(G \cup G(n, p))[A \cup U]$ , which will finish the proof.

Let  $p \geq Cn^{-2/k}$  and let  $G$  be an  $n$ -vertex graph with  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ . Note that it might be true that both  $G$  and  $G(n, p)$  are  $K_r$ -free (a.a.s. for  $G(n, p)$ ). Thus, to build even a single copy of  $K_r$ , we may have to use both deterministic edges (from  $G$ ) and random edges (from  $G(n, p)$ ). We will use the following partition of the edge set of  $K_r$ .

**Definition 4.1.** *For  $r \in \mathbb{N}$  and  $k \in \mathbb{N}$  such that  $2 \leq k \leq r$ , let  $r^*, q \in \mathbb{N}$  be such that  $k(r^* - 1) + q = r$  and  $0 < q \leq k$ . Then  $H_{det} = H_{det}(r, k)$  is the complete  $r^*$ -partite graph with parts  $V_1, V_2, \dots, V_{r^*}$*

such that  $|V_1| = |V_2| = \dots = |V_{r^*-1}| = k$  and  $|V_{r^*}| = q$ , i.e.  $H_{\det} := K_{k, \dots, k, q}^{r^*}$ . We also define  $\overline{H_{\det}}$  to be  $K_r - E(H_{\det})$ , i.e. the complement of  $H_{\det}$  on the same vertex set.

Some examples are given in Figure 1. Note that when  $k = r$ ,  $H_{\det}$  is simply an independent set of size  $k = r$  and  $\overline{H_{\det}}$  is an  $r$ -clique. The motivation for this partition comes from the following observation. We can build a copy of  $K_r$  in  $G \cup G(n, p)$  by taking  $\Omega(n^r)$  copies of  $H_{\det}$  in  $G$  and then applying Janson's inequality to conclude that we can 'fill up' the independent sets in some copy of  $H_{\det}$  by  $K_k$ s and a  $K_q$  and obtain a copy of  $K_r$ . With a few more ideas, one can repeatedly apply this naive idea to greedily obtain an almost perfect  $K_r$ -tiling (see Theorem 5.1).

To build the absorbing set, we use the reachability arguments introduced by Lo and Markström [40]. The main part of the reachability arguments rely on the following notion of reachable paths. Given two vertices  $u, v$ , a set  $P$  of constant size is called a reachable path for  $u, v$  if both  $P \cup \{u\}$  and  $P \cup \{v\}$  contain perfect  $K_r$ -tilings. Then we meet the same problem as above, and thus need to build certain structures by deterministic edges and 'fill up the gaps' by random edges. We need much more involved arguments, including building copies of  $K_r$  in a few different ways and making sure that we can recover the missing edges by  $G(n, p)$ . Moreover, when  $k > r/2$  we cannot prove the reachability between *every* two vertices and have to pursue a weaker property, namely, building a partition of  $V(G)$  such that the reachability can be established within each part.

Once we have established the existence of reachable paths, we piece these together to form what we call 'absorbing gadgets' (Definition 6.14) and then further combine these absorbing gadgets to define our full absorbing structure in  $G \cup G(n, p)$ . We use an idea of Montgomery [42, 43] in order to define our absorbing structure, using an auxiliary 'template' to dictate how we interweave our absorbing gadgets, which will ensure that the resulting absorbing structure has a strong absorbing property, in that it can contribute to a  $K_r$ -tiling in many ways. We will introduce the random edges of  $G(n, p)$  only in the last stage, when proving the existence of the full absorbing structure in  $G \cup G(n, p)$ . Thus, we will first be occupied with finding many reachable paths and absorbing gadgets which use these reachable paths, restricting our attention *only* to the deterministic edges which will contribute to our eventual absorbing structure.

Our analysis splits into three cases depending on the structure of  $H_{\det}$  or equivalently, the values of  $r$  and  $k$ . The cases are as follows:

1.  $H_{\det}$  is balanced i.e.  $q = k$  and so  $r/k \in \mathbb{N}$ ,
2.  $\chi(H_{\det}) = r^* \geq 3$  and  $H_{\det}$  is not balanced i.e.  $k < r/2$  and  $r/k \notin \mathbb{N}$ ,
3.  $\chi(H_{\det}) = r^* = 2$  and  $H_{\det}$  is not balanced i.e.  $r/2 < k < r$ .

Examples of each case can be seen in Figure 1.

## 5 An almost perfect tiling

In this section we study almost perfect tilings and prove Theorem 5.1 below. As is the case throughout, in this almost perfect tiling, the edges of  $G$  which contribute to the copies of  $K_r$  will be copies of  $H_{\det}$  as defined in Definition 4.1. We will rely on  $G(n, p)$  to then 'fill in the gaps', providing the missing edges i.e.  $\overline{H_{\det}}$ , to guarantee that each copy of  $H_{\det}$  is in fact part of a copy of  $K_r$  in  $G \cup G(n, p)$ . Note that  $\chi(H_{\det}) = \lceil r/k \rceil$  and recall the definition of  $\chi_{\text{cr}}$  discussed in Section

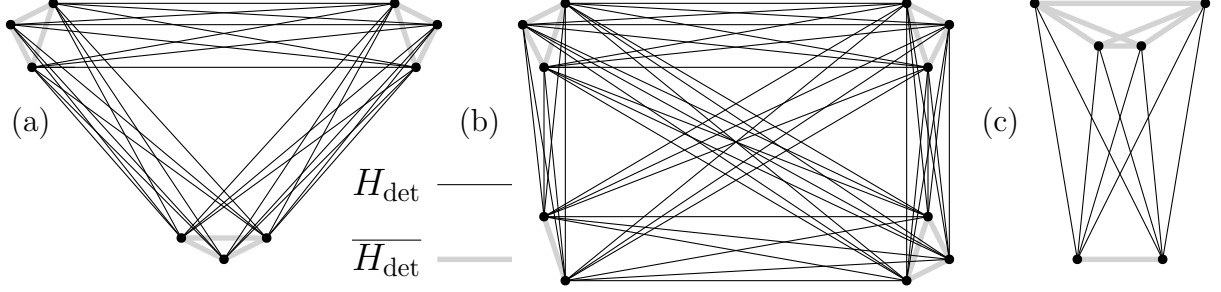


Figure 1: Some examples of the graphs  $H_{\det}$  and  $\overline{H_{\det}}$  from Definition 4.1 for each case:

1. (a)  $r = 9, k = q = r^* = 3,$
2. (b)  $r = 11, k = 3, q = 2, r^* = 4,$
3. (c)  $r = 6, k = 4, q = r^* = 2.$

**2.1.** When  $k$  divides  $r$ , we have  $\chi_{cr}(H_{\det}) = r/k = \chi(H_{\det})$  and when  $k$  does not divide  $r$ , we have  $\chi_{cr}(H_{\det}) = \lfloor r/k \rfloor \frac{r}{r - (r-k)\lfloor r/k \rfloor} = r/k$ .

Thus, the almost perfect tiling result of Komlós, Theorem 2.1, guarantees the existence of an  $H_{\det}$ -tiling in  $G$  which covers almost all the vertices. However, given such a tiling we cannot guarantee that the correct edges appear in  $G(n, p)$  in order to extend each copy of  $H_{\det}$  in the tiling to a copy of  $K_r$ . We aim instead to greedily build a  $K_r$ -tiling and guarantee that at each step there are  $\Omega(n^r)$  copies of  $H_{\det}$ . To achieve this, we use the regularity lemma and apply Theorem 2.1 to the reduced graph of  $G$ . Then by the counting lemma, each copy of  $H_{\det}$  in the reduced graph will provide many copies of  $H_{\det}$  in  $G$ .

**Theorem 5.1.** *Let  $2 \leq k \leq r$  and  $\alpha, \gamma > 0$ . Then there exists  $C = C(\alpha, \gamma, r, k) > 0$  such that if  $p \geq Cn^{-2/k}$  and  $G$  is an  $n$ -vertex graph with  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ , then  $G \cup G(n, p)$  a.a.s. contains a  $K_r$ -tiling covering all but at most  $\alpha n$  vertices.*

*Proof.* Apply Lemma 2.4 to  $G$  with  $0 < \epsilon \ll d \ll \gamma/4, \alpha/4, 1/r$  and  $m_0 \in \mathbb{N}$  large, such that  $d_0 := \frac{(d-\epsilon)^r}{(2+r)2^r} \geq 4\epsilon/\alpha$ . We may assume that  $n$  is sufficiently large. Note that by Corollary 2.5, the resulting  $(\epsilon, d)$ -reduced graph  $R$  has  $m \geq m_0$  vertices and satisfies  $\delta(R) \geq (1 - \frac{k}{r} + \gamma/2)m$ . Let the size of the clusters in the regularity partition be  $n'$  and note that  $n/N(\epsilon, m_0) \leq n' \leq \epsilon n$ . Now by Theorem 2.1, as  $m \geq m_0$  is sufficiently large, there exists an  $H_{\det}$ -tiling  $\mathcal{H}$  covering all but at most  $\alpha m/4$  vertices of  $R$ . Let  $\mathcal{U}_1, \dots, \mathcal{U}_t \in \binom{V(R)}{r}$  such that the  $\mathcal{U}_j$  span disjoint copies of  $H_{\det}$  in  $\mathcal{H}$ .

Next, let  $\mathcal{F}$  be the collection of subsets  $W \subseteq V(G)$  such that there exists some  $j \in [t]$  for which  $W$  intersects each  $U \in \mathcal{U}_j$  in at least  $\alpha n'/2$  elements (and  $W$  contains no vertices of clusters from outside of  $\mathcal{U}_j$ ). Here we say that  $\mathcal{U}_j$  corresponds to  $W$ . Moreover, we call a copy of  $K_r$  in  $W$  *crossing* if it contains precisely one vertex from each cluster in the class  $\mathcal{U}_j$ .

We claim that a.a.s., every  $W$  in  $\mathcal{F}$  contains a crossing copy of  $K_r$  in  $G \cup G(n, p)$ . Indeed, fix some  $W \in \mathcal{F}$  and suppose  $\mathcal{U}_j$  corresponds to  $W$ . Then there are subsets  $W_1, \dots, W_r \subset V(G)$  and clusters  $\{U_1, \dots, U_r\} = \mathcal{U}_j$  such that  $W_i \subseteq U_i$ ,

$$|W_i| = \frac{\alpha n'}{2} \geq \frac{\alpha n}{2N(\epsilon, m_0)},$$

$\cup_{i \in r} W_i \subseteq W$  and  $U_1, U_2, \dots, U_r$  form a copy  $H$  of  $H_{\det}$  in  $R$ . By Lemma 2.2, for every  $U_i U_j \in E(H)$ , we have that  $G[W_i, W_j]$  is a  $(4\epsilon/\alpha)$ -regular pair with density at least  $d - \epsilon$ . Thus by Lemma 2.3  $G[W]$  contains at least  $(d_0 \alpha n'/2)^r = \Omega(n^r)$  copies of  $H_{\det}$  where in each such copy of  $H_{\det}$  precisely one vertex lies in each of  $W_1, \dots, W_r$ ; let  $\mathcal{C}_W$  denote this collection. Now noting that  $F := \overline{H_{\det}} = K_r - E(H_{\det})$  is a collection of disjoint cliques of size at most  $k$ , Lemma 2.7 (part 1 and 3) implies that  $\Phi_F \geq Cn$ . Also, we have that  $|\mathcal{F}| \leq 2^n$ . Thus for  $C > 0$  sufficiently large, Corollary 2.9 gives that for every  $W \in \mathcal{F}$  there is a copy of  $H_{\det}$  from  $\mathcal{C}_W$  which hosts a labelled copy of  $\overline{H_{\det}}$  in  $G(n, p)$ ; thus the claim is satisfied.

One can now use the claim to greedily build the almost perfect  $K_r$ -tiling in  $G \cup G(n, p)$ . Indeed, initially set  $\mathcal{K} := \emptyset$ . At each step we will add a copy of  $K_r$  to  $\mathcal{K}$  whilst ensuring  $\mathcal{K}$  is a  $K_r$ -tiling in  $G \cup G(n, p)$ . Further, at every step we only add a copy  $K$  of  $K_r$  if there is some  $j \in [t]$  such that each vertex in  $K$  lies in a different cluster in  $\mathcal{U}_j$  (recall each  $\mathcal{U}_j$  consists of  $r$  clusters).

Suppose we are at a given step in this process such that there exists some cluster  $U_i \in \mathcal{U}_j$  (for some  $j$ ) that still has at least  $\alpha n'/2$  vertices uncovered by  $\mathcal{K}$ . This in fact implies that every cluster in  $\mathcal{U}_j$  contains at least  $\alpha n'/2$  vertices uncovered by  $\mathcal{K}$ ; these uncovered vertices correspond precisely to a set  $W \in \mathcal{F}$ . Hence by the above claim there is a crossing copy of  $K_r$  in  $(G \cup G(n, p))[W]$ . Add this to  $\mathcal{K}$ . Thus, we can repeat this process, increasing the size of  $\mathcal{K}$  at every step, until we find that for every  $j \in [t]$ , all the clusters in  $\mathcal{U}_j$  have at least  $(1 - \alpha/2)n'$  vertices covered by  $\mathcal{K}$ .

That is, a.a.s. there is a  $K_r$ -tiling in  $G \cup G(n, p)$  covering all but at most

$$(\alpha n'/2 \times m) + (\alpha m/4 \times n') + |V_0| \leq \alpha n$$

vertices, as desired. Note that the first term in the above expression comes from the vertices in clusters from the classes  $\mathcal{U}_j$ ; the second term comes from those vertices in clusters that were uncovered by  $\mathcal{H}$ .  $\square$

Note that one can in fact establish the case  $k = r$  in a much simpler way because the copies of  $K_r$  that we look for can be completely provided by  $G(n, p)$ , see e.g. [24, Theorem 4.9].

## 6 The absorption

The aim of this section is to prove the existence of an *absorbing structure*  $\mathcal{A}$  in  $G' := G \cup G(n, p)$ . The main outcomes are Corollaries 6.22, 6.24 and 6.25, which will be used in next section to prove our main result.

The key component of the absorbing structure will be some absorbing subgraph  $F \subset G'$ . We will define  $F$  so that it can contribute to a  $K_r$ -tiling in many ways. In fact we will define  $F$  so that if we remove  $F$  from  $G'$  and we tile almost all of what remains (Theorem 5.1), then no matter which small set of vertices remains, the properties of  $F$  allow us to complete this tiling to a full tiling of  $G'$ . There are some complications, and the absorbing structure will have different features depending on the exact values of minimum degree and the size of the cliques that we look to tile with.

Our absorbing subgraph will be comprised of two sets of edges, namely the deterministic edges in  $G$  and the random edges in  $G(n, p)$ . Initially, we will be concerned with finding (parts of) the appropriate subgraph in  $G$  (Section 6.1). In fact, we will need to prove the existence of many copies of the deterministic subgraphs we want, as we will rely on there being enough of these to guarantee that one of them will match up with random edges in  $G(n, p)$  (Section 6.2) to give the



desired subgraph. Therefore it is useful throughout to consider, with foresight, the random edges that we will be looking for to complete our desired structure, as this also motivates the form of our deterministic subgraphs.

## 6.1 The absorbing structure - deterministic edges

The smallest building block in our absorbing graph will be  $K_{r+1}^-$ , the complete graph on  $r + 1$  vertices with one edge missing, say between  $w_1$  and  $w_2$ . This is useful for the simple reason that it can contribute to a  $K_r$ -tiling in two ways, namely  $K_{r+1}^- \setminus \{w_i\}$  for  $i = 1, 2$ . We introduce the following notation to keep track of the partition of the edges between the deterministic graph and the random graph.

**Definition 6.1.** *Suppose  $t, r, r_1, r_2, \dots, r_t \in \mathbb{N}$  such that  $\sum_{i=1}^t r_i = r + 1$ . We use the notation*

$$H := (K_{r_1, r_2, \dots, r_t}^t, i, j),$$

for not necessarily distinct  $i, j \in [t]$ , to denote the  $(r + 1)$ -vertex graph  $K_{r_1, r_2, \dots, r_t}^t$  with two distinct distinguished vertices:  $w_1$  in the  $i^{\text{th}}$  part (which has size  $r_i$ ) and  $w_2$  in the  $j^{\text{th}}$  part (which has size  $r_j$ ).

**Definition 6.2.** *Let  $r \in \mathbb{N}$  and consider an  $(r + 1)$ -vertex graph  $F$  with two distinguished vertices  $w_1$  and  $w_2$ . (Typically we will take  $F = H$  as in Definition 6.1.) We then write<sup>2</sup>  $\overline{F}$  to denote the graph on the same vertex set  $V(K_{r+1}^-) = V(F)$  such that  $E(\overline{F}) := E(K_{r+1}^-) \setminus E(F)$ , where we take the non-edge of  $K_{r+1}^-$  to be  $w_1 w_2$ . Thus  $K_{r+1}^- \subseteq F \cup \overline{F}$ .*

We think of  $H, \overline{H}$  and  $K_{r+1}^-$  as all lying on the same vertex set throughout with the two distinguished vertices  $w_1, w_2$  being defined for all three. The following graph gives the paradigm for how we split the edges of  $K_{r+1}^-$  between the deterministic and the random graph.

**Definition 6.3.** *For  $r \in \mathbb{N}$  and  $k \in \mathbb{N}$  such that  $2 \leq k \leq r$ , let  $r^*, q \in \mathbb{N}$  be such that  $k(r^* - 1) + q = r$  and  $0 < q \leq k$ . Then  $H_0 := (K_{k, \dots, k, q+1}^{r^*}, r^*, r^*)$ .*

Some examples of  $H_0$  and  $\overline{H_0}$  can be seen in Figures 3 and 4. Note that if  $w_1$  and  $w_2$  are the distinguished vertices of  $H_0$ , then  $H_0 \setminus w_i$  for  $i = 1, 2$  are both copies of the graph  $H_{\text{det}}$  from Definition 4.1. Also note that  $\overline{H_0}$  is a disjoint union of  $k$ -cliques as well as a disjoint copy of  $K_{q+1}^-$ . Thus, when  $q \leq k - 1$ , it follows from Lemma 2.7 and Corollary 2.9 that the graph  $\overline{H_0}$  is abundant<sup>3</sup> in  $G(n, p)$  when  $p \geq Cn^{-2/k}$  for some large enough  $C$ . Furthermore, as we will see, the minimum degree condition for  $G$  along with Lemma 2.6 will imply that there are  $\Omega(n^{r+1})$  copies of  $H_0$  in  $G$ . This suggests the suitability of this definition as a candidate for how to partition the edge set of  $K_{r+1}^-$  between deterministic and random edges. We remark that the case when  $q = k$  is slightly more subtle and we have to adjust our decomposition accordingly. We will discuss this in more detail in the next section.

<sup>2</sup>Note that our use of the notation  $\overline{F}$  is non-standard here.

<sup>3</sup>Specifically, one can see that any linear sized set in  $V(G(n, p))$  contains a copy of  $\overline{H_0}$  a.a.s..

### 6.1.1 Reachability

In this subsection, we define reachable paths and show that we can find many of these in our deterministic graph  $G$ , when the graphs used to define such paths are chosen appropriately. The main results are Proposition 6.6, Proposition 6.7 and Proposition 6.12 which deal with Case 1, 2 and 3 respectively. We first define a reachable path which is a graph which connects together  $(r + 1)$ -vertex graphs as follows.

**Definition 6.4.** *Let  $t, r \in \mathbb{N}$  and let  $\mathbf{H} = (H^1, H^2, \dots, H^t)$  be a vector of  $(r + 1)$ -vertex graphs  $H^i$  such that each  $H^i$  has two distinguished vertices,  $w_1^i$  and  $w_2^i$ . Then an  $\mathbf{H}$ -path is the graph  $P$  obtained by taking one copy of each  $H^i$  and identifying  $w_2^i$  with  $w_1^{i+1}$ , for  $i \in [t - 1]$ . We call  $w_1^1$  and  $w_2^t$  the endpoints of  $P$ .*

*In the case where  $H^1 = H^2 = \dots = H^t = H$  for some  $(r + 1)$ -vertex graph  $H$ , we use the notation  $\mathbf{H} = (H, t)$  and thus refer to  $(H, t)$ -paths. For  $\mathbf{H} = (H^1, H^2, \dots, H^t)$ , we also define  $\overline{\mathbf{H}} := (\overline{H^1}, \overline{H^2}, \dots, \overline{H^t})$  where  $\overline{H^i}$  is as defined in Definition 6.2.*

We give some explicit examples of  $\mathbf{H}$ -paths later in Figure 6. In the following, as we look to find embeddings of  $\mathbf{H}$ -paths and larger subgraphs in  $G$  and  $G(n, p)$ , we will always be considering *labelled* embeddings. Therefore, implicitly, when we define graphs such as the  $\mathbf{H}$ -paths above, we think of these graphs as having some fixed labelling of their vertices. Again, the motivation for the definition of  $\mathbf{H}$ -paths comes from considering  $K_{r+1}^-$ , with vertices  $w_1, w_2$  such that  $w_1 w_2 \notin E(K_{r+1}^-)$ . Indeed, then a  $(K_{r+1}^-, t)$ -path  $P$  has two  $K_r$ -tilings missing a single vertex; one on the vertices of  $V(P) \setminus w_1^1$ , and one on the vertices of  $V(P) \setminus w_2^t$ . Our first step is to find many  $\mathbf{H}$ -paths in the deterministic graph  $G$ , for an appropriately defined  $\mathbf{H}$ . In particular, we are interested in the images of the endpoints of the paths.

**Definition 6.5.** *Let  $\beta > 0$ ,  $t, r \in \mathbb{N}$  and  $\mathbf{H} = (H^1, \dots, H^t)$  be a vector of  $(r + 1)$ -vertex graphs (each of which is endowed with a tuple of distinguished vertices). We say that two vertices  $x, y \in V(G)$  in an  $n$ -vertex graph  $G$  are  $(\mathbf{H}; \beta)$ -reachable (or  $(H, t; \beta)$ -reachable if  $\mathbf{H} = (H, \dots, H) = (H, t)$ ) if there are at least  $\beta n^{tr-1}$  distinct labelled embeddings of the  $\mathbf{H}$ -path  $P$  in  $G$  such that the endpoints of  $P$  are mapped to  $\{x, y\}$ .*

As discussed before, the graph  $H_0$  from Definition 6.3 will be used to provide deterministic edges for our absorbing structure. That is, we look for  $(H_0, t)$ -paths in  $G$  for some appropriate  $t$ . However, for various reasons there are complications with this approach. Sometimes using a slightly different graph  $H$  will allow more vertices to be reachable to each other. Also, as is the case below when  $r/k \in \mathbb{N}$ , it is possible that  $\overline{H_0}$  is not sufficiently common in the random graph  $G(n, p)$ . Therefore, we have to tweak the graph  $H_0$  in order to accommodate these subtleties. This is the reason for using a vector of graphs  $\mathbf{H}$  as we will see. We will look first at Case 1, when  $r/k \in \mathbb{N}$  and so  $\overline{H_0}$  contains a copy of  $K_{k+1}^-$ . This is too dense to appear in the random graph  $G(n, p)$  with the frequency that we require and thus we define  $H_1$  as in the following proposition.

**Proposition 6.6.** *Let  $\gamma > 0$ ,  $n, r, k \in \mathbb{N}$  such that  $r/k =: r^* \in \mathbb{N}$ ,  $2 \leq k \leq r$  and  $n$  is sufficiently large. Let  $H_1 := J_1 \cap K_{r+1}^-$  where  $J_1 := (K_{k, k, \dots, k, 1}^{(r^*+1)}, r^* + 1, 1)$ , as defined in Definition 6.1, and we consider  $K_{r+1}^-$  to be on the same vertex set as  $J_1$  with a non-edge between the distinguished vertices of  $J_1$ . Likewise, let  $H'_1 := (K_{k, k, \dots, k, 1}^{(r^*+1)}, 1, r^* + 1) \cap K_{r+1}^-$  be the same graph with the labels of the distinguished vertices switched. See Figure 2 for an example of  $H_1$  (and  $H'_1$  which is identical).*

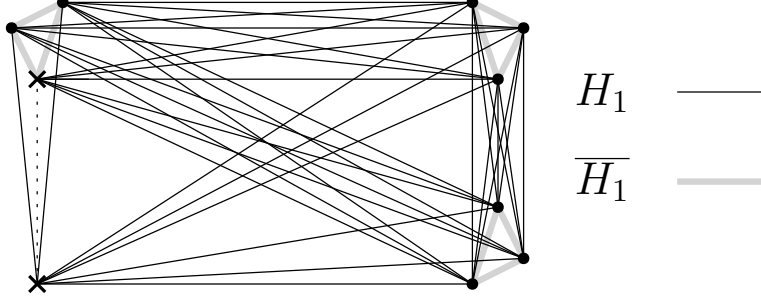


Figure 2: An example of  $H_1$  and  $\overline{H_1}$  for Case 1 (a)  $r = 9, k = q = r^* = 3$  (see Definitions 6.1–6.2 and Proposition 6.6 for the definitions).

Then there exists a  $\beta_1 = \beta_1(r, k, \gamma) > 0$  such that for any  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ , any pair of distinct vertices in  $V(G)$  are  $(\mathbf{H}_1; \beta_1)$ -reachable where  $\mathbf{H}_1 := (H_1, H'_1)$ .

*Proof.* Let  $w_1, w_2$  be the distinguished vertices of  $H_1 = (K_{k, k, \dots, k, 1}^{(r^*+1)}, r^* + 1, 1) \cap K_{r+1}^-$  as defined in Definition 6.1. Fix a pair  $x, y \in V(G)$ .

We will show that for any  $z \in V(G) \setminus \{x, y\}$ , there are at least  $\beta'_1 n^{r-1}$  labelled embeddings of  $H_1$  which map  $w_1$  to  $x$  and  $w_2$  to  $z$ , for some  $\beta'_1 = \beta'_1(r, k, \gamma) > 0$ . Once we have established this property, this implies the proposition. Indeed, by symmetry, we can also find  $\beta'_1 n^{r-1}$  embeddings of  $H_1$  which map  $w_1$  to  $y$  and  $w_2$  to  $z$ . Set  $\beta_1 := \beta_1'^2/2$ . Thus there are at least

$$(n - 2) \times \beta_1' n^{r-1} \times (\beta_1' n^{r-1} - r^2 n^{r-2}) \geq \beta_1 n^{2r-1}$$

distinct embeddings of the  $\mathbf{H}_1$ -path in  $G$  such that the endpoints are mapped to  $\{x, y\}$ , as desired. This follows as there are  $n - 2$  choices for  $z$ ; at least  $\beta_1' n^{r-1}$  choices for the copy of  $H_1$  containing  $x$  and  $z$ ; at least  $(\beta_1' n^{r-1} - r^2 n^{r-2})$  choices for the copy of  $H_1'$  containing  $z$  and  $y$  that are *disjoint* from the choice of  $H_1$  (except for the vertex  $z$ ).

So let us fix  $x, z \in V(G)$ . The proof now follows easily from Lemma 2.6. As  $kr^* = r$ , we can express the minimum degree as  $\delta(G) \geq (1 - \frac{1}{r^*} + \gamma)n$ . Thus any set of at most  $r^*$  vertices has at least  $\gamma n$  common neighbours. Therefore we have at least  $(\gamma n)^{r^*}$  labelled copies  $K$  of  $K_{r^*}$  where  $V(K) = \{x_1, \dots, x_{r^*}\} \subset N_G(x)$  and  $\{x_2, \dots, x_{r^*}\} \subset N_G(x) \cap N_G(z)$ . This follows by first choosing  $\{x_2, \dots, x_{r^*}\}$  and then  $x_1$  with the right adjacencies. Thus, by Lemma 2.6 we have  $\beta_1' n^{r-1}$  labelled embeddings of the blow-up,  $H_1 \setminus \{w_1, w_2\} = K_{k, \dots, k, k-1}^{r^*}$ , of these cliques, crucially within the correct neighbourhoods ( $N_G(x)$  and  $N_G(x) \cap N_G(z)$ ) to ensure that together with  $\{x, z\}$  they give us the required embeddings of  $H_1$ .  $\square$

Note that an  $\overline{\mathbf{H}_1}$ -path  $\overline{P_1}$  has endpoints which are isolated. The other vertices of  $\overline{P_1}$  lie in copies of  $K_k$  and these copies are disjoint from each other except for a single pair of  $K_k$ s that meet at a singular vertex. See Figure 6 for an example. We now turn to Case 2, as described in Section 4. Here we can use the graph  $H_0$  from Definition 6.3. We also use a slight variant of  $H_0$  where we redefine the distinguished vertices.

**Proposition 6.7.** *Suppose  $\gamma > 0$ ,  $n, r, k \in \mathbb{N}$ , such that  $n$  is sufficiently large,  $2 \leq k < r/2$  and  $r/k \notin \mathbb{N}$ . Further, let  $r^*, q$  and  $H_0$  be as defined in Definition 6.3 and let  $H'_0 = (K_{k, \dots, k, q+1}^{r^*}, 1, 2)$  be*

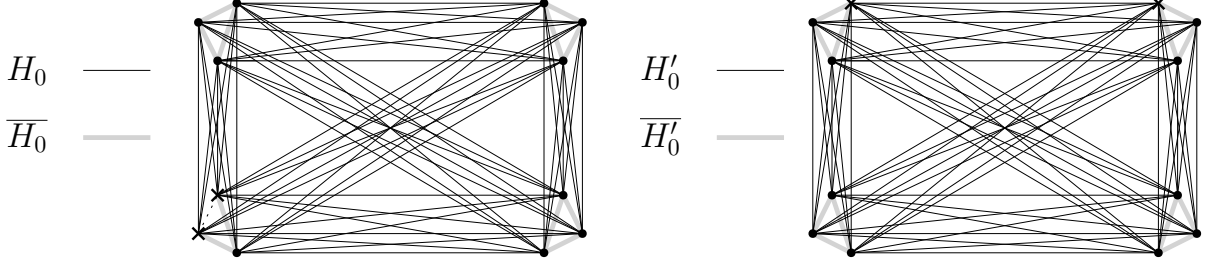


Figure 3: An example of  $H_0$ ,  $\overline{H_0}$ ,  $H'_0$  and  $\overline{H'_0}$  for Case 2 (b)  $r = 11, k = 3, q = 2, r^* = 4$  (see Definitions 6.1–6.3 and Proposition 6.7 for the definitions).

the same graph as  $H_0$  with distinguished vertices in distinct<sup>4</sup> parts of size  $k$  (see Figure 3).

Then there exists  $\beta_2 = \beta_2(r, k, \gamma) > 0$  such that for any  $n$ -vertex  $G$  of minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ , every pair of distinct vertices  $x, y$  in  $V(G)$  are  $(\mathbf{H}_2; \beta_2)$ -reachable where  $\mathbf{H}_2 := (H_0, H'_0, H_0, H_0)$ .

*Proof.* We know that  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n \geq (1 - \frac{1}{r^*-1} + \gamma)n$ . Therefore every set of at most  $r^* - 1$  vertices has a common neighbourhood of size at least  $\gamma n$ . We will appeal to Lemma 2.6 to give us the whole  $\mathbf{H}_2$ -path in one fell swoop. Let  $J$  be a graph with vertex set

$$V(J) = \{x_1, \dots, x_{r^*-1}, z_1, w_1, \dots, w_{r^*-2}, z_2, u_1, \dots, u_{r^*-2}, z_3, y_1, \dots, y_{r^*-1}\},$$

and  $E(J)$  consisting of  $r^*$ -cliques on  $\{x_1, \dots, x_{r^*-1}, z_1\}$ ,  $\{z_1, w_1, \dots, w_{r^*-2}, z_2\}$ ,  $\{z_2, u_1, \dots, u_{r^*-2}, z_3\}$  and  $\{z_3, y_1, \dots, y_{r^*-1}\}$ . We claim that if we can find  $(\gamma n)^{4r^*-3}/2$  copies of  $J$  in  $G$  such that  $x_i \in N_G(x)$  and  $y_i \in N_G(y)$  for  $i = 1, \dots, r^* - 1$ , then we are done. Indeed, consider a blow-up  $J'$  of  $J$  with parts

$$\{X_1^{(k)}, \dots, X_{r^*-1}^{(k)}, Z_1^{(k+q-1)}, W_1^{(q+1)}, W_2^{(k)}, \dots, W_{r^*-2}^{(k)}, Z_2^{(2k-1)}, \\ U_1^{(q+1)}, U_2^{(k)}, \dots, U_{r^*-2}^{(k)}, Z_3^{(k+q-1)}, Y_1^{(k)}, \dots, Y_{r^*-1}^{(k)}\},$$

where the parts correspond to the vertices of  $J$  in the obvious way and the size of each part is indicated by the superscript. Now if we have a copy of  $J'$  in  $G \setminus \{x, y\}$  with  $X_i \subset N_G(x)$  and  $Y_i \subset N_G(y)$  for all  $i = 1, \dots, r^* - 1$ , then this gives us an embedding of an  $\mathbf{H}_2$ -path. Indeed for  $i = 1, 3$ , arbitrarily partition  $Z_i := \{z'_i\} \sqcup Z''_i \sqcup Z'''_i$  with  $|Z''_i| = q - 1$  and  $|Z'''_i| = k - 1$  and partition  $Z_2 := \{z'_2\} \sqcup Z''_2 \sqcup Z'''_2$  with  $|Z''_2| = |Z'''_2| = k - 1$ . Then  $\{x, z'_1\} \cup Z''_1 \cup_{i=1}^{r^*-1} X_i$  and  $\{y, z'_3\} \cup Z''_3 \cup_{i=1}^{r^*-1} Y_i$  both give copies of  $H_0$  whilst  $\{z'_1, z'_2\} \cup Z''_1 \cup Z''_2 \cup_{i=1}^{r^*-2} W_i$  and  $\{z'_2, z'_3\} \cup Z''_3 \cup Z''_2 \cup_{i=1}^{r^*-2} U_i$  both give copies of  $H'_0$  where in all cases the distinguished vertices appear in the first set of the union.

It suffices then, by Lemma 2.6, to find  $(\gamma n)^{4r^*-3}/2$  embeddings of  $J$  in  $G$  with the  $x_i \in N_G(x)$  and  $y_i \in N_G(y)$ . We can do this greedily. Indeed if we choose the  $x_i$  and  $y_i$  first, followed by  $z_1$  and  $z_3$ , then  $z_2 \in N_G(z_1) \cap N_G(z_3)$  and then the remaining vertices, we are always seeking to choose a vertex in  $G$  which has at most  $r^* - 1$  neighbours which have already been chosen. Thus, by our degree condition, we have at least  $\gamma n$  choices for each vertex with the right adjacencies. To ensure that these choices actually give an embedding of  $J$  we then discard any set of choices with repeated vertices, of which there are  $O(n^{4r^*-4})$ , and thus the conclusion holds as  $n$  is sufficiently large.  $\square$

<sup>4</sup>Note that this is possible as we are in the case where the number of parts,  $r^*$  of  $H_0$  is at least 3.

Consider an  $\overline{\mathbf{H}_2}$ -path which we denote  $\overline{P_2}$  (see Figure 6 for an example). It is formed by copies of  $K_k$  and  $K_{q+1}^-$  which intersect in at most one vertex and such that the endpoints of  $\overline{P_2}$  lie in copies of  $K_{q+1}^-$ . Furthermore, note that the endpoints of  $\overline{P_2}$  are in distinct connected components. This will be an important feature when we start to address the random edges of our absorbing structure as it will allow us to use Lemma 2.7 to conclude certain statements about the likelihood of finding our desired random subgraph in  $G(n, p)$ . This motivated the introduction of  $H'_0$  in the previous proposition.

In Case 3, we cannot hope to prove reachability between every pair of vertices. Indeed our minimum degree in this case is  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$  and  $k > r/2$  and so it is possible that  $\delta(G) < n/2$  and  $G$  is disconnected. Thus, as in [19, 22], we use a partition of the vertices into ‘closed’ parts, where we can guarantee that two vertices in the same part are reachable, with some set of parameters. We adopt the following notation which also allows us to consider different possibilities for what vectors we use for reachability.

**Definition 6.8.** Let  $\mathcal{H}$  be a set of vectors, such that the entry of each vector in  $\mathcal{H}$  is an  $(r+1)$ -vertex graph endowed with a tuple of distinguished vertices. We say that two vertices in  $G$  are  $(\mathcal{H}; \beta)$ -reachable if they are  $(\mathbf{H}; \beta)$ -reachable for some  $\mathbf{H} \in \mathcal{H}$ .

We say that a subset  $V$  of vertices in a graph  $G$  is  $(\mathcal{H}; \beta)$ -closed if every pair of vertices<sup>5</sup> in  $V$  is  $(\mathcal{H}; \beta)$ -reachable. We denote<sup>6</sup> by  $N_{\mathcal{H}, \beta}(v)$  the set of vertices in  $G$  that are  $(\mathcal{H}; \beta)$ -reachable to  $v$ .

Thus, in this notation, the conclusion of Proposition 6.6 states that  $V(G)$  is  $(\mathbf{H}_1; \beta_1)$ -closed for all  $G$  satisfying the given hypothesis (and similarly for Proposition 6.7). Notice that if a set  $V$  is  $(\mathcal{H}; \beta)$ -closed in a graph  $G$  it may be the case that two vertices  $x, y \in V$  are  $(\mathbf{H}; \beta)$ -reachable whilst two other vertices  $z, w \in V$  are  $(\mathbf{H}'; \beta)$ -reachable for some distinct  $\mathbf{H}, \mathbf{H}' \in \mathcal{H}$  of different lengths.

It will be useful for us to consider the following notion.

**Definition 6.9.** Let  $\mathcal{H}, \tilde{\mathcal{H}}$  be two sets of vectors as in Definition 6.8. Then

$$\mathcal{H} + \tilde{\mathcal{H}} := \mathcal{H} \cup \tilde{\mathcal{H}} \cup (\mathcal{H} \cdot \tilde{\mathcal{H}}),$$

where  $\mathcal{H} \cdot \tilde{\mathcal{H}}$  is defined to be the set

$$\mathcal{H} \cdot \tilde{\mathcal{H}} := \{(\mathbf{H}, \tilde{\mathbf{H}}) := (H^1, \dots, H^t, \tilde{H}^1, \dots, \tilde{H}^{\tilde{t}}) : \mathbf{H} := (H^1, \dots, H^t) \in \mathcal{H}, \tilde{\mathbf{H}} := (\tilde{H}^1, \dots, \tilde{H}^{\tilde{t}}) \in \tilde{\mathcal{H}}\}.$$

That is,  $\mathcal{H} + \tilde{\mathcal{H}}$  comprises of all vectors that lie in  $\mathcal{H}, \tilde{\mathcal{H}}$ , or that can be obtained by a concatenation of a vector from  $\mathcal{H}$  with a vector from  $\tilde{\mathcal{H}}$ .

As an important example, defining  $\mathcal{H}(H, \leq t) := \{(H, s) : 1 \leq s \leq t\}$ , we have that

$$\mathcal{H}(H, \leq t_1) + \mathcal{H}(H, \leq t_2) = \mathcal{H}(H, \leq t_1 + t_2).$$

In what follows we will apply the following simple lemma repeatedly.

<sup>5</sup>Note that we do not require the vertices of the  $\mathbf{H}$ -paths which give the reachability to lie in  $V$ .

<sup>6</sup>If  $\mathcal{H}$  consists of just one vector  $\mathbf{H}$ , we simply refer to sets being  $(\mathbf{H}; \beta)$ -closed and use  $N_{\mathbf{H}, \beta}(v)$  to denote the closed neighbourhood of a vertex.

**Lemma 6.10.** *Let  $r \in \mathbb{N}$  and let  $\mathcal{H}_x, \mathcal{H}_y$  be two sets of vectors of  $(r+1)$ -vertex graphs, each of which is endowed with a tuple of distinguished vertices and suppose that  $t_x := |\mathcal{H}_x|$  and  $t_y := |\mathcal{H}_y|$  are both finite. Suppose  $G$  is a sufficiently large  $n$ -vertex graph and  $x, y \in V(G)$ . Suppose there exist  $\beta_x, \beta_y, \epsilon > 0$ , and some subset  $U \subseteq V(G)$  with  $|U| \geq \epsilon n$  such that for every  $z \in U$ ,  $x$  and  $z$  are  $(\mathcal{H}_x; \beta_x)$ -reachable and  $z$  and  $y$  are  $(\mathcal{H}_y; \beta_y)$ -reachable. Then  $x$  and  $y$  are  $(\mathcal{H}_x + \mathcal{H}_y; \beta)$ -reachable for  $\beta := \frac{\epsilon \beta_x \beta_y}{2 t_x t_y} > 0$ .*

*Proof.* By the pigeonhole principle, there exists some  $U' \subseteq U$  such that  $|U'| \geq \frac{\epsilon n}{t_x t_y}$  and some  $\mathbf{H}_x \in \mathcal{H}_x, \mathbf{H}_y \in \mathcal{H}_y$  such that for every  $z \in U'$ ,  $z$  and  $x$  are  $(\mathbf{H}_x; \beta_x)$ -reachable and  $z$  and  $y$  are  $(\mathbf{H}_y; \beta_y)$ -reachable. Suppose  $\mathbf{H}_x$  has length  $s_x$  and  $\mathbf{H}_y$  has length  $s_y$ . Thus, fixing  $z \in U'$ , there are at least  $\beta_x \beta_y n^{(s_x + s_y)r - 2}$  pairs of labelled vertex sets  $S_x$  and  $S_y$  in  $G$  such that there is an embedding of an  $\mathbf{H}_x$ -path on  $S_x \cup \{x, z\}$  mapping endpoints to  $\{x, z\}$  and an embedding of a  $\mathbf{H}_y$ -path on the vertices  $S_y \cup \{y, z\}$  which maps the endpoints to  $\{y, z\}$ . Of these pairs, at most

$$s_x s_y r^2 n^{(s_x + s_y)r - 3}$$

are *not* vertex disjoint or they intersect  $\{x, y\}$ . Hence, as  $n$  is sufficiently large we have at least  $\frac{\beta_x \beta_y}{2} n^{(s_x + s_y)r - 2}$  vertex disjoint pairs which together form an embedding of an  $(\mathbf{H}_x, \mathbf{H}_y)$ -path. As we have at least  $\frac{\epsilon}{t_x t_y} n$  choices for  $z$ , this gives that  $x$  and  $y$  are  $((\mathbf{H}_x, \mathbf{H}_y); \beta)$ -reachable and  $(\mathbf{H}_x, \mathbf{H}_y) \in \mathcal{H}_x + \mathcal{H}_y$ .  $\square$

We now turn to proving reachability in Case 3. The following two lemmas together find the partition we will work on. Similar ideas have been used in [19, 22].

**Lemma 6.11.** *Suppose  $\gamma > 0$  and  $n, r, k, q \in \mathbb{N}$  such that  $r/2 < k \leq r - 1$ ,  $r = k + q$  and  $n$  is sufficiently large. Let  $c := \lceil r/q \rceil$  and for  $t \in \mathbb{N}$  define  $\mathcal{H}^t := \mathcal{H}(H_0, \leq 2^t) = \{(H_0, s) : 1 \leq s \leq 2^t\}$ , where  $H_0 = (K_{k, q+1}^2, 2, 2)$  is as defined in Definition 6.3 with distinguished vertices  $w_1$  and  $w_2$ .*

*Then there exists constants  $0 < \beta'_3 = \beta'_3(r, k, \gamma) \ll \alpha = \alpha(r, k, \gamma)$  such that any  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$  can be partitioned into at most  $c - 1$  parts, each of which is  $(\mathcal{H}^c; \beta'_3)$ -closed and of size at least  $\alpha n$ .*

*Proof.* Firstly, observe that there is some  $\eta = \eta(r, k, \gamma) > 0$  such that in every set of at least  $c$  vertices, there are two vertices which are  $((H_0, 1); \eta)$ -reachable. Indeed, fix some arbitrary set of vertices  $S = \{v_1, \dots, v_c\} \subset V(G)$ , and for  $v \in V = V(G)$ , define  $d_S(v) := |\{i \in [c] : vv_i \in E(G)\}|$ . Let  $\bar{d}_S := \sum_{v \in V} d_S(v)/n$  be the average. Then we have that

$$\sum_{v \in V} d_S(v) = \sum_{i \in [c]} d_G(v_i) \geq c \left(1 - \frac{k}{r} + \gamma\right) n \geq (1 + c\gamma)n.$$

Thus  $\sum_{v \in V} \binom{d_S(v)}{2} \geq n \binom{\bar{d}_S}{2} \geq \frac{(c\gamma)^2}{2} n$  by Jensen's inequality. By averaging over all pairs we have that there exists a pair  $i \neq j \in [c]$  so that both  $v_i$  and  $v_j$  are in the neighbourhood of at least  $\gamma^2 n$  vertices. That is,  $|N_G(v_i) \cap N_G(v_j)| \geq \gamma^2 n$ .

Therefore there are at least  $\gamma^3 n^2$  edges in  $G$  with one endpoint in  $N_G(v_i) \cap N_G(v_j)$ . Applying Lemma 2.6 this ensures that there is  $\eta = \eta(r, k, \gamma) > 0$  so that there are  $\eta n^{r-1}$  copies of  $K_{k, q-1}^2$  where the first vertex class lies in  $N_G(v_i) \cap N_G(v_j)$ . Thus together they form copies of  $H_0$  with distinguished vertices  $v_i$  and  $v_j$ ; so  $v_i$  and  $v_j$  are  $((H_0, 1); \eta)$ -reachable in  $G$ .



Note also that there is some fixed  $\alpha' = \alpha'(r, k, \gamma) > 0$  such that  $|N_{(H_0, 1), \alpha'}(v)| \geq \alpha'n$  for every  $v \in V(G)$ . Indeed, this follows as there are at least  $(\gamma n)^2/2$  edges in  $G$  with one endpoint in  $N_G(v)$ . So, by Lemma 2.6, there is a fixed  $\alpha'' = \alpha''(r, k, \gamma) > 0$  such that there are at least  $\alpha''n^r$  embeddings of  $H_0$  which map  $w_1$  to  $v$ . Setting  $\alpha' := \alpha''/3$ , this implies that there are at least  $\alpha'n$  vertices which are  $((H_0, 1); \alpha')$ -reachable to  $v$ .

Now let  $0 < \epsilon \ll \alpha'$ ,  $\eta =: \eta_0$  and  $\eta_i := \frac{\epsilon^4}{2^{2i+1}}\eta_{i-1}$  for  $i = 1, \dots, c$ . Set  $\beta'_3 := \eta_c$ . As in the statement of the lemma, define  $\mathcal{H}^t := \mathcal{H}(H_0, \leq 2^t)$  for values of  $t \leq c$  and note that  $\mathcal{H}^t + \mathcal{H}^t = \mathcal{H}^{t+1}$ . We will be interested in  $(\mathcal{H}^t; \eta_t)$ -reachability and so we will use the shorthand notation  $\tilde{N}_t(v) := N_{\mathcal{H}^t, \eta_t}(v)$ . Let  $\ell$  be the maximal integer such that there exists a set of  $\ell$  vertices,  $v_1, \dots, v_\ell$  with  $v_j$  and  $v_{j'}$  not  $(\mathcal{H}^{c-\ell}; \eta_{c-\ell})$ -reachable for any pair  $j \neq j' \in [\ell]$ .

Suppose  $\ell = 1$ . Then  $V(G)$  is  $(\mathcal{H}^{c-1}; \eta_{c-1})$ -closed. As  $\mathcal{H}^{c-1} \subseteq \mathcal{H}^c$  and  $\eta_{c-1} > \beta'_3$ , the lemma holds in this case.

We also have that  $\ell \leq c - 1$  from our observations above, so we can assume  $2 \leq \ell \leq c - 1$ . Now fix such a set of  $\ell$  vertices,  $v_1, \dots, v_\ell$ . We make the following two observations:

- (i) Any  $v \in V(G) \setminus \{v_1, \dots, v_\ell\}$  is in  $\tilde{N}_{c-\ell-1}(v_j)$  for some  $j \in [\ell]$  from our definition of  $\ell$ , as otherwise  $v$  could be added to give a larger family contradicting the maximality of  $\ell$ . Indeed, this follows because two vertices that are not  $(\mathcal{H}^{c-\ell}; \eta_{c-\ell})$ -reachable are certainly not  $(\mathcal{H}^{c-\ell-1}; \eta_{c-\ell-1})$ -reachable by definition.
- (ii)  $|\tilde{N}_{c-\ell-1}(v_j) \cap \tilde{N}_{c-\ell-1}(v_{j'})| \leq \epsilon n$  for every pair  $j \neq j' \in [\ell]$ . This follows from Lemma 6.10 as otherwise we would have that  $v_j$  and  $v_{j'}$  are  $(\mathcal{H}^{c-\ell}; \eta_{c-\ell})$ -reachable, a contradiction.

We define  $U_j := \left( \tilde{N}_{c-\ell-1}(v_j) \cup \{v_j\} \right) \setminus \left( \bigcup_{j' \in [\ell] \setminus \{j\}} \tilde{N}_{c-\ell-1}(v_{j'}) \right)$  for  $j \in [\ell]$ , and  $U_0 := V(G) \setminus \bigcup_{j \in [\ell]} U_j$ . Now for  $j \in [\ell]$ , we have that  $U_j$  is  $(\mathcal{H}^{c-\ell-1}; \eta_{c-\ell-1})$ -closed. Indeed, if there was a  $j \in [\ell]$  and  $u_1, u_2 \in U_j$  not reachable, then  $\{u_1, u_2\} \cup \{v_1, \dots, v_\ell\} \setminus \{v_j\}$ , is a larger family contradicting the definition of  $\ell$ . Thus, the  $U_j$  almost form the partition we are looking for except that it remains to consider the vertices in  $U_0$ . For these, we greedily add them to the other  $U_j$ : We have that for each  $u \in U_0$ ,

$$|N_{(H_0, 1), \alpha'}(u) \setminus U_0| \geq \alpha'n - |U_0| \geq \alpha'n - \binom{\ell}{2} \epsilon n \geq \ell \epsilon n. \quad (6.1)$$

Here the second inequality holds due to (i), (ii) and the definition of the  $U_j$ ; the final inequality holds by our choice of  $\epsilon$ . Thus, there is a  $j$  such that  $|N_{(H_0, 1), \alpha'}(u) \cap U_j| \geq \epsilon n$ , and we add  $u$  to this  $U_j$ , arbitrarily choosing such a  $j$  if there are multiple choices. Let  $V_1, \dots, V_\ell$  be the resulting partition.

Applications of Lemma 6.10 show that each  $V_j$  is  $(\mathcal{H}^c; \eta_c)$ -closed. Indeed suppose, for example, that  $w_1$  and  $w_2$  are two vertices that lie in  $U_0$  and are added to  $U_j$  in the process of defining  $V_j$ . Then for each  $i = 1, 2$ , taking  $W_i = N_{(H_0, 1), \alpha'}(w_i) \cap U_j$ , an application of Lemma 6.10 with  $U = W_1$  gives that for any  $x \in U_j$ ,  $w_1$  and  $x$  are  $(\mathcal{H}'; \eta')$ -reachable where  $\mathcal{H}' = \mathcal{H}(H_0, \leq 2^{c-\ell-1} + 1)$  and  $\eta' = \epsilon \alpha' \eta_{c-\ell-1} / 2^{c-\ell}$ . Another application of Lemma 6.10, this time with  $U = W_2$  then gives that  $w_1$  and  $w_2$  are  $(\mathcal{H}''; \eta'')$ -reachable with  $\mathcal{H}'' = \mathcal{H}(H_0, \leq 2^{c-\ell-1} + 2) \subseteq \mathcal{H}^c$  and  $\eta'' = \epsilon \eta' \alpha' / (2^{c-\ell} + 2) > \eta_{c-\ell} > \beta'_3$ . Showing other cases of reachability within each  $V_j$  are similar. We are now done since for each  $j \in [\ell]$ ,

$$|V_j| \geq |U_j| \geq |N_{(H_0, 1), \alpha'}(v_j) \setminus (N_{(H_0, 1), \alpha'}(v_j) \cap U_0)| \stackrel{(ii)}{\geq} \alpha'n - \ell \epsilon n = \alpha n,$$

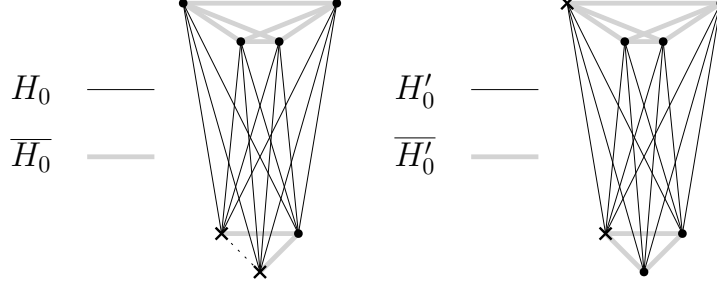


Figure 4: An example of  $H_0$ ,  $\overline{H_0}$ ,  $H'_0$  and  $\overline{H'_0}$  for Case 3 (c)  $r = 6, k = 4, q = r^* = 2$  (see Definitions 6.1–6.3 and Proposition 6.12 for the definitions).

where  $\alpha := \alpha' - \ell\epsilon \gg \eta_{c-\ell} \geq \eta_c = \beta'_3$ . □

The rough idea for how to handle Case 3 is to run the same proof as in the other cases on each *part* of the partition given by Lemma 6.11. The point of Lemma 6.11 is that we recover the reachability within each part, albeit at the expense of allowing a family of possible paths used for reachability. However, in the process, we lose the minimum degree condition within each part. The purpose of the next proposition is to fix this, by adjusting parameters and making the partition coarser. Thus, we recover a minimum degree condition which is not quite as strong as what we had previously but good enough to work with in what follows.

**Proposition 6.12.** *Suppose  $\gamma > 0$  and  $n, r, k, q \in \mathbb{N}$  such that  $r/2 < k \leq r - 1$ ,  $r = k + q$  and  $n$  is sufficiently large. Let  $c := \lceil r/q \rceil$  and let  $H_0 = (K_{k,q+1}^2, 2, 2)$  as defined in Definition 6.3 and  $H'_0 = (K_{k,q+1}^2, 1, 2)$  be the same graph with distinguished vertices in distinct parts of the bipartition<sup>7</sup> (see Figure 4). We define the following family of vectors of  $(r + 1)$ -vertex graphs (endowed with tuples of vertices):*

$$\mathcal{H}_3 := \bigcup_{t=3}^{c(2^{c+1}+1)} \{ \mathbf{H} \in \{H_0, H'_0\}^t : \mathbf{H}[1] = \mathbf{H}[t] = H_0 \text{ and } \mathbf{H}[i] = H'_0 \text{ for some } 2 \leq i \leq t - 1 \},$$

where  $\mathbf{H}[i]$  denotes the  $i^{\text{th}}$  entry of  $\mathbf{H}$ .

Then for all  $\epsilon > 0$ , there exists constants  $0 < \beta_3(r, k, \gamma, \epsilon) \ll \alpha(r, k, \gamma)$  such that for any  $n$ -vertex graph  $G$  with minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$  there is a partition  $\mathcal{P}$  of  $V(G)$  into at most  $c - 1$  parts such that each part  $U \in \mathcal{P}$  satisfies the following:

- (i)  $|U| \geq \alpha n$ ;
- (ii) All but at most  $\epsilon n$  vertices  $v \in U$  satisfy  $|N_G(v) \cap U| \geq (1 - \frac{k}{r} + \frac{\gamma}{2})|U|$ ;
- (iii)  $U$  is  $(\mathcal{H}_3; \beta_3)$ -closed.

*Proof.* This is a simple case of adjusting the partition already obtained after applying Lemma 6.11. Let  $\alpha, \beta'_3$  be defined as in the outcome of Lemma 6.11 and let  $\mathcal{P}'$  be the partition of  $V(G)$  obtained, with vertex parts denoted  $V_1, \dots, V_s$ . Fix  $\mu := \epsilon\alpha\gamma/2c^3$ . We create an auxiliary graph  $J$  on vertex set  $\{V_1, \dots, V_s\}$  where for  $i \neq j \in [s]$  we have an edge  $V_i V_j$  in  $J$  if and only if there are at least  $\mu n^2$

<sup>7</sup>This is analogous to the graph  $H'_0$  defined in Proposition 6.7.

edges in  $G$  with one endpoint in  $V_i$  and one in  $V_j$ . Then our new partition  $\mathcal{P}$  in  $G$  will come from the connected components of  $J$ . That is, if  $C_1, \dots, C_t$  are the components of  $J$ , then for  $i \in [t]$ , we define  $U_i := \cup_{j: V_j \in C_i} V_j$  and let  $\mathcal{P}$  consist of the  $U_i$  with  $i \in [t]$ . Then certainly point (i) of the hypothesis is satisfied for all  $U_i$ . Also (ii) is satisfied. Indeed, suppose there exists  $i \in [t]$ , with  $d_{U_i}(v) := |N_G(v) \cap U_i| < (1 - \frac{k}{r} + \frac{\gamma}{2})|U_i|$  for at least  $\epsilon n$  vertices of  $U_i$ . Thus, for such vertices  $|N_G(v) \cap (V(G) \setminus U_i)| \geq \frac{\gamma\alpha}{2}n$  and by averaging there exists some  $V_{j(v)} \in \mathcal{P}'$  such that  $V_{j(v)} \cap U_i = \emptyset$  and  $d_{V_{j(v)}}(v) \geq \frac{\gamma\alpha}{2c}n$ . We average again to conclude that there is some  $j, j' \in [s]$  such that  $V_j \subset U_i$ ,  $V_{j'} \cap U_i = \emptyset$  and  $V_j$  contains at least  $\epsilon n/c^2$  vertices  $v$  which have degree into  $V_{j'}$   $d_{V_{j'}}(v) \geq \frac{\gamma\alpha}{2c}n$ . This contradicts our definition of  $J$  as then  $V_j V_{j'}$  should be an edge of  $J$  and thus in the same part of  $\mathcal{P}$ .

Thus it only remains to establish reachability. We begin by proving the following claim which is a slight variation of Lemma 6.10.

**Claim 6.13.** *Let  $\mathcal{H}^c$  be as defined in Lemma 6.11. Suppose  $x, y \in V(G)$  and that there exist (not necessarily disjoint) sets  $S_x, S_y \subset V(G)$  such that for any  $z_x \in S_x$ ,  $x$  and  $z_x$  are  $(\mathcal{H}^c; \beta'_3)$ -reachable and for any  $z_y \in S_y$ ,  $y$  and  $z_y$  are  $(\mathcal{H}^c; \beta'_3)$ -reachable. If there exists at least  $\mu n^2$  edges with one endpoint in  $S_x$  and one endpoint in  $S_y$ , then  $x$  and  $y$  are  $(\mathbf{H}; \beta''_3)$ -reachable for some  $\beta''_3 = \beta''_3(\mu, \beta'_3, c) > 0$  and  $\mathbf{H} \in \mathcal{H}_3$  of length at most  $2^{c+1} + 1$ .*

Indeed letting  $w'_1, w'_2$  be the distinguished vertices of  $H'_0$ , we have, by Lemma 2.6, that there are at least  $\mu' n^{r+1}$  embeddings of  $H'_0$  into  $G$  which map  $w'_1$  to  $S_x$  and  $w'_2$  to  $S_y$  for some  $\mu' = \mu'(\mu) > 0$ . By averaging, there exists  $\mathbf{H}_x, \mathbf{H}_y \in \mathcal{H}^c$  such that there are  $\frac{\mu'}{2^{2c}} n^{r+1}$  embeddings of  $H'_0$  such that the image of  $w'_1$  and  $x$  are  $(\mathbf{H}_x; \beta'_3)$ -reachable and the image of  $w'_2$  and  $y$  are  $(\mathbf{H}_y; \beta'_3)$ -reachable. By considering the embeddings of  $\mathbf{H}_x, \mathbf{H}_y$  and  $H'_0$  which join to give an embedding of an  $(\mathbf{H}_x, H'_0, \mathbf{H}_y)$ -path (that is, ignoring choices of embeddings which are not vertex disjoint), we see that  $x$  and  $y$  are  $((\mathbf{H}_x, H'_0, \mathbf{H}_y), \beta''_3)$ -reachable with  $\beta''_3 := \frac{\mu' \beta_3^{r+2}}{2^{2c+1}}$ . This completes the proof of the claim.

Recall the partition  $\mathcal{P}' = \{V_1, V_2, \dots, V_s\}$ . Further consider any part  $U \in \mathcal{P}$ . First suppose  $U = V_j$  for some  $j$ . Now given any  $x, y \in U$ , by Lemma 6.11,  $x$  and  $y$  are already  $(H_0, s)$ -reachable for some  $s \leq 2^c$ . However,  $(H_0, s)$  does not contain a copy of  $H'_0$  and so is not a valid vector in the family  $\mathcal{H}_3$ . We therefore apply Claim 6.13 with  $S_x = S_y = V_j \setminus \{x, y\}$ , to conclude that  $x$  and  $y$  are  $(\mathbf{H}; \beta''_3)$ -reachable for some  $\mathbf{H} \in \mathcal{H}_3$  of length at most  $2^{c+1} + 1$ . Indeed since  $V_j$  sends fewer than  $\mu n^2$  edges out to any other part  $V_i$  of  $\mathcal{P}'$  and  $|V_j| \geq \alpha n$ , the minimum degree condition on  $G$  ensures that there are at least  $2\mu n^2$  edges in  $G[V_j]$  and hence  $\mu n^2$  edges in  $S_x = S_y$  allowing Claim 6.13 to be applied.

Next suppose  $U$  is the union of more than one part from  $\mathcal{P}'$ . If  $x \in V_i \subseteq U$  and  $y \in V_j \subseteq U$ , for  $i \neq j \in [s]$  and  $V_i V_j \in E(J)$  as defined above, we can again apply Claim 6.13 to conclude  $x$  and  $y$  are  $(\mathbf{H}; \beta''_3)$ -reachable for some  $\mathbf{H} \in \mathcal{H}_3$  of length at most  $2^{c+1} + 1$ . Therefore, we just need to establish reachability for vertices  $x, y$  such that  $x \in V_i, y \in V_j$  with  $V_i V_j \notin E(J)$  but such that  $V_i$  and  $V_j$  are in the same component of  $J$ . If  $i \neq j$ , there is a path of (at most  $c$ ) edges from  $V_i$  to  $V_j$  in  $J$ ; if  $i = j$  there is a walk of length  $2 \leq c$  in  $J$  that starts and ends at  $V_i = V_j$  (i.e. traverse a single edge in  $J$ ). In both cases we can repeatedly apply Lemma 6.10 to derive that  $x$  and  $y$  are  $(\mathcal{H}_3; \beta_3)$ -reachable with  $\beta_3 := \left(\frac{\alpha \beta''_3}{2^{(2^{c+1}+2)}}\right)^c$ . It is crucial here that we apply Claim 6.13 in all cases to establish the reachability here (even when  $i = j$ ) in order to guarantee that the vectors witnessing the reachability contain a copy of  $H'_0$  and hence indeed lie in  $\mathcal{H}_3$ .  $\square$

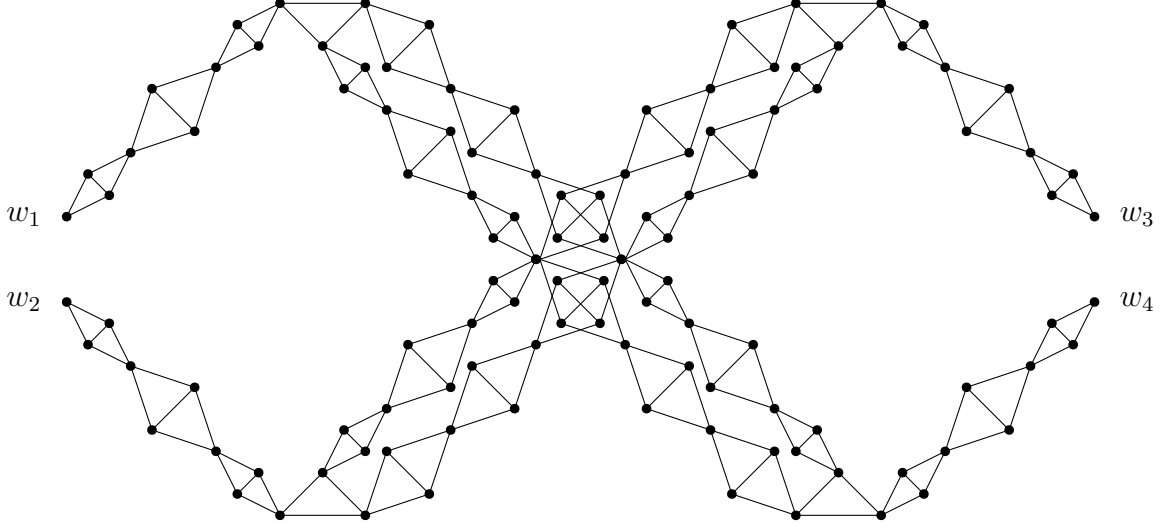


Figure 5: An  $(\underline{H}, K_3)$ -absorbing gadget with  $\underline{H} := \{H_{i,j} : i \in [3], j \in [4]\}$  such that each  $H_{i,j} = (K_4^-, 3)$ . The base set of the absorbing gadget is  $W := \{w_1, w_2, w_3, w_4\}$ .

We remark that the reason for the introduction of  $H'_0$  in Proposition 6.12 is two-fold. Firstly, it allows us to establish reachability between parts from Lemma 6.11 which have many edges between them. Moreover, as in Proposition 6.7, we have that for every  $\mathbf{H} \in \mathcal{H}_3$ , if  $\bar{P}$  is an  $\bar{H}$ -path, then the endpoints of  $\bar{P}$  are in distinct connected components on  $\bar{P}$  (see Figure 6 for an example), which is something that we will require later.

### 6.1.2 Absorbing gadgets

In this section, we will focus on larger subgraphs which we look to embed in our graph and which will be used as part of an absorbing structure. These are formed by piecing together the  $\mathbf{H}$ -paths of the previous section and the aim will be to obtain subgraphs with even more flexibility, in that they will be able to contribute to a tiling in many ways. The key definition is a graph which we call an absorbing gadget.

**Definition 6.14.** *Let  $r, s \in \mathbb{N}$ , let  $H$  be an  $r$ -vertex graph and let  $\underline{H} := \{H_{i,j} : i \in [r], j \in [s]\}$  be a labelled family of vectors of  $(r+1)$ -vertex graphs (with tuples of distinguished vertices). Then an  $(\underline{H}, H)$ -absorbing gadget is a graph obtained by the following procedure. Take disjoint  $H_{i,j}$ -paths for  $1 \leq i \leq r$  and  $1 \leq j \leq s$  and denote their endpoints by  $u_{i,j}$  and  $v_{i,j}$ . Place a copy of  $H$  on  $\{v_{i,j} : i \in [r]\}$  for each  $j \in [s]$ . For  $2 \leq i \leq r$ , identify all vertices  $\{u_{i,j} : 1 \leq j \leq s\}$  and relabel this vertex  $u_i$ . Finally relabel  $u_{1,j}$  as  $w_j$  for  $j \in [s]$  and let  $W := \{w_1, w_2, \dots, w_s\}$ , which we refer to as the base set of vertices for the absorbing gadget.*

An example of an absorbing gadget is given in Figure 5. Recall that we always consider  $K_{r+1}^-$  to have two distinguished vertices which form the only non-edge of the graph. In the previous section we commented on how a  $(K_{r+1}^-, t)$ -path  $P$  with endpoints  $x$  and  $y$  has two  $K_r$ -tilings covering all but one vertex; the first misses  $x$ , the other misses  $y$ . The point of the absorbing gadget is to generalise this property, giving a graph which can use any one of a number of vertices (the base set) in a  $K_r$ -tiling. In more detail, suppose  $s, t^* \in \mathbb{N}$  and  $\mathcal{H} = \mathcal{H}(K_{r+1}^-, \leq t^*) = \{(K_{r+1}^-, t) : 1 \leq t \leq t^*\}$

as defined in the previous subsection. Let  $\underline{\mathbf{H}} := \{\mathbf{H}_{i,j} : i \in [r], j \in [s]\}$  where each  $\mathbf{H}_{i,j}$  is an element from  $\mathcal{H}$ . Then an  $(\underline{\mathbf{H}}, K_r)$ -absorbing gadget  $F$  with base set  $W = \{w_1, \dots, w_s\}$  has the property that for any  $j \in [s]$ , there is a  $K_r$ -tiling covering precisely  $(F \setminus W) \cup \{w_j\}$ . Indeed, we have that for all  $j' \neq j$  and  $i \in [r]$ , there is a  $K_r$ -tiling of the  $\mathbf{H}_{i,j'}$ -path  $P_{i,j'}$  which uses<sup>8</sup>  $v_{i,j'}$  and not the other endpoint of  $P_{i,j'}$ . Then there is a tiling of the  $\mathbf{H}_{1,j}$ -path which uses  $w_j$ , a tiling of the  $\mathbf{H}_{i,j}$ -path for  $2 \leq i \leq r$  which uses  $u_i$  and a copy of  $K_r$  on  $\{v_{i,j} : i \in [r]\}$  which completes the desired  $K_r$ -tiling.

As in the previous subsection, we begin by showing that there are many absorbing gadgets in the deterministic graph. Again, although we are interested in  $(\underline{\mathbf{H}}, K_r)$ -absorbing gadgets for some  $\underline{\mathbf{H}}$  consisting of vectors, all of whose entries are  $K_{r+1}^-$ , we split the edges of our absorbing gadget and rely on the deterministic graph to provide many copies of a subgraph of the gadget. In particular, we will use here our paradigm  $H_{\det}$ , defined in Definition 4.1. The following general proposition allows us to show that we can find many absorbing gadgets if all the vertices which we hope to map the base set to, are reachable to each other.

**Definition 6.15.** *Let  $r, s \in \mathbb{N}$ . Let  $\mathcal{H}$  be a finite set of vectors, such that each entry of each vector in  $\mathcal{H}$  is an  $(r+1)$ -vertex graph with a tuple of distinguished vertices. We write  $\mathcal{H}(r \times s)$  for the collection of all ordered labelled sets  $\underline{\mathbf{H}} := \{\mathbf{H}_{i,j} : i \in [r], j \in [s]\}$  where each  $\mathbf{H}_{i,j}$  is an element from  $\mathcal{H}$ . If  $\mathcal{H}$  consists of a single vector  $\mathbf{H}$  we write  $\mathbf{H}(r \times s) := \mathcal{H}(r \times s)$ . That is,  $\mathbf{H}(r \times s)$  is the ordered labelled (multi)set with each element a copy of  $\mathbf{H}$ .*

**Proposition 6.16.** *Let  $\alpha, \gamma, \beta' > 0, q, k, r \in \mathbb{N}$  and let  $\mathcal{H}$  be a finite set of vectors, such that each entry of each vector in  $\mathcal{H}$  is an  $(r+1)$ -vertex graph with a tuple of distinguished vertices.*

*Then there exists  $\beta = \beta(\alpha, \gamma, \beta', q, k, r, \mathcal{H}) > 0$ , such that for sufficiently large  $n$ , if  $G$  is an  $n$ -vertex graph with vertex subset  $U \subseteq V(G)$  such that  $U$  is  $(\mathcal{H}; \beta')$ -closed,  $|U| \geq \alpha n$  and  $\delta(G[U]) \geq (1 - \frac{k}{r} + \gamma)|U|$ , then for any set  $X = \{x_1, \dots, x_s\} \subset U$  with  $|X| \leq q$ , there exists some  $\underline{\mathbf{H}} \in \mathcal{H}(r \times s)$  and some  $(\underline{\mathbf{H}}, H_{\det})$ -absorbing gadget  $F$  with base set  $W = \{w_1, \dots, w_s\}$  such that there are at least  $\beta n^{v(F)-s}$  embeddings of  $F$  in  $G$  which map  $w_i$  to  $x_i$  for  $i \in [s]$ .*

*Proof.* Firstly notice that for a fixed  $s \leq q$ , there is a finite number (i.e.  $|\mathcal{H}|^{rs}$ ) of  $(\underline{\mathbf{H}}, H_{\det})$ -absorbing gadgets  $F$  such that  $\underline{\mathbf{H}} \in \mathcal{H}(r \times s)$  and  $F$  has a base set of size  $s$ . Let  $\mathcal{F}_s$  be the set of all such absorbing gadgets, let  $f := |\mathcal{F}_s|$  and set  $Q := \max\{|F| - s : F \in \mathcal{F}_s\}$ . We claim that there is some  $\beta'' = \beta''(\alpha, \gamma, \beta', q, k, r, \mathcal{H}) > 0$  such that with  $G$  and  $U$  as in the statement of the proposition and  $X \subset U$  of size  $s$ , there are at least  $\beta'' n^Q$  subsets  $S \subseteq V(G) \setminus X$  of  $Q$  ordered vertices such that there is an embedding of some  $F \in \mathcal{F}_s$  in  $G$  which maps the base set of  $F$  to  $X$  and the other vertices to a subset<sup>9</sup> of  $S$ . Given this claim, the conclusion of the proposition follows easily. Indeed, by averaging we get that there is some  $F \in \mathcal{F}_s$  and at least  $(\beta''/f)n^Q$  ordered subsets  $S$  of  $Q$  vertices in  $V(G)$  as above, that correspond to an embedding of  $F$ . Then setting  $\beta := \beta''/(Q!f)$ , we get that there must be at least  $\beta n^{|F|-s}$  embeddings of  $F$  in  $G$  which map the base set to  $X$ . Indeed for each such embedding  $F'$  of  $F$ , the vertex set  $V(F') \setminus X$  lies in at most  $Q!n^{Q-(|F|-s)}$  different ordered sets of vertices  $S \subseteq V(G)$ .

So it remains to find these  $\beta'' n^Q$  ordered subsets  $S$ . We will show that  $S$  can be generated in a series of steps so that every time we choose some  $a$  vertices, we have  $\Omega(n^a)$  choices. We will use the notation of Definition 6.14. Firstly we select  $r-1$  vertices  $Y = \{y_2, y_3, \dots, y_r\}$  in  $U \setminus X$  which we

<sup>8</sup>We label all vertices in this discussion as in Definition 6.14.

<sup>9</sup>In particular, if  $|F| < Q$  then not all of the vertices of  $S$  are used in this embedding.

can do in  $\binom{|U \setminus X|}{r-1} = \Omega(n^{r-1})$  many ways. Now repeatedly find disjoint copies of  $H_{\det}$  in  $U \setminus (X \cup Y)$  and label these  $\{z_{i,j} : 1 \leq i \leq r, 1 \leq j \leq s\}$  such that  $\{z_{i,j} : 1 \leq i \leq r\}$  comprise a copy of  $H_{\det}$  for each  $j \in [s]$ . In order to do this we repeatedly apply Lemma 2.6 and the degree condition which we can take to be  $\delta(G[U]) \geq (1 - \frac{k}{r} + \frac{\gamma}{2})|U|$  (ignoring any neighbours of vertices that have already been chosen in  $S$ ). Hence there are  $\Omega(n^{rs})$  choices for these copies of  $H_{\det}$ .

Now for  $2 \leq i \leq r$  and  $1 \leq j \leq s$ , we have that  $y_i$  and  $z_{i,j}$  are  $(\mathbf{H}_{i,j}, \beta')$ -reachable for some  $\mathbf{H}_{i,j} \in \mathcal{H}$  of length  $t_{i,j}$  say. Thus there are  $\beta' n^{rt_{i,j}-1}$  embeddings of an  $\mathbf{H}_{i,j}$ -path  $P$  in  $G$  which map the endpoints of  $P$  to  $\{y_i, z_{i,j}\}$ . We ignore those choices of embeddings of  $P$  which use previously chosen vertices of  $S$ , of which there are  $O(n^{rt_{i,j}-2})$ . Similarly, for  $1 \leq j \leq s$ ,  $x_j$  and  $z_{1,j}$  are  $(\mathbf{H}_{1,j}, \beta')$ -reachable for some  $\mathbf{H}_{1,j} \in \mathcal{H}$ , so select an embedding of an  $\mathbf{H}_{1,j}$ -path in  $G$  which maps the endpoints to  $\{x_j, z_{1,j}\}$  and has all other vertices disjoint from previously chosen vertices. This gives an embedding of an  $(\mathbf{H}, H_{\det})$ -absorbing gadget in  $G$  which maps the base set  $W$  to  $X$ ,  $u_{i,j}$  to  $z_{i,j}$  for  $i \in [r], j \in [s]$  and maps  $u_i$  to  $y_i$  for  $i \in [r]$ . Choosing unused vertices arbitrarily until we have a set  $S$  of  $Q$  vertices, the claim and hence the proof of the proposition are settled.  $\square$

## 6.2 The absorbing structure - random edges

In this section, we will introduce the edges of  $G(n, p)$  and show that  $G \cup G(n, p)$  contains the absorbing structure we desire. The absorbing structure will be formed by choosing absorbing gadgets rooted on certain prescribed sets of vertices. The absorbing gadgets will be  $(\mathbf{H}, K_r)$ -absorbing gadgets  $F^*$  for some  $\mathbf{H}$  consisting of vectors whose entries are all  $K_{r+1}^-$ . In order to obtain these absorbing gadgets, we consider the absorbing gadgets of just deterministic edges which we looked at in the previous section and show that with high probability, one of these matches up with random edges to get the required subgraph  $F^*$ . We begin by investigating the absorbing gadgets that we look for in the random graph.

### 6.2.1 Absorbing gadgets in the random graph

Recalling Definitions 4.1 and 6.14, let  $\mathbf{H} := \{\mathbf{H}_{i,j} : i \in [r], j \in [s]\}$  be a labelled family of vectors of  $(r+1)$ -vertex graphs and suppose that there is an embedding  $\varphi$  of an  $(\mathbf{H}, H_{\det})$ -absorbing gadget  $F'$  in  $G$  which maps the base set of the gadget to some  $U \subset V(G)$ , with  $|U| = s$ . Recalling Definition 6.4, define  $\overline{\mathbf{H}} := \{\overline{\mathbf{H}_{i,j}} : i \in [r], j \in [s]\}$ . Now in order to complete this absorbing gadget  $F'$  into one which has the form that we require, we have to find a labelled embedding of an  $(\overline{\mathbf{H}}, \overline{H_{\det}})$ -absorbing gadget  $F$  onto the ordered vertex set  $\varphi(V(F'))$  in  $G(n, p)$ . The following lemma will be used to show that there are sufficiently many embeddings in  $G(n, p)$  of the necessary  $F$ s defined as above. It is worth noting that as  $F$  is uniquely defined by  $F'$ , it is in fact the way that we chose our deterministic absorbing gadgets, that guarantees the following conclusions.

**Lemma 6.17.** *Let  $k, r, s \in \mathbb{N}$  and  $C > 1$ , with  $2 \leq k \leq r$  and suppose  $p = p(n) \geq Cn^{-2/k}$ . Suppose  $\mathbf{H}$  is such that:*

1.  $\mathbf{H} \in \mathbf{H}_1(r \times s)$  if  $r/k \in \mathbb{N}$ , recalling the definition of  $\mathbf{H}_1$  from Proposition 6.6;
2.  $\mathbf{H} \in \mathbf{H}_2(r \times s)$  if  $r/k \notin \mathbb{N}$  and  $k < r/2$ , recalling the definition of  $\mathbf{H}_2$  from Proposition 6.7;
3.  $\mathbf{H} \in \mathbf{H}_3(r \times s)$  if  $k > r/2$ , recalling the definition of  $\mathbf{H}_3$  from Proposition 6.12.

*Then if  $F$  is an  $(\overline{\mathbf{H}}, \overline{H_{\det}})$ -absorbing gadget with base set  $W$  such that  $|W| = s$ , we have that  $\Phi_{F \setminus W} \geq Cn$  and  $\Phi_{F, W} \geq Cn^{1/k}$ .*



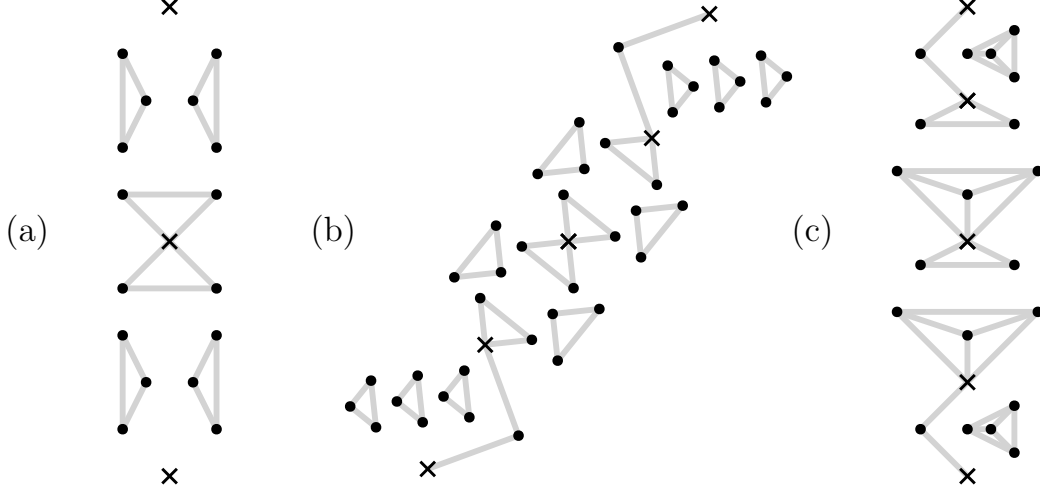


Figure 6: Some examples of  $\mathbf{H}$ -paths for various  $\mathbf{H}$  (see Definition 6.4 and Propositions 6.6, 6.7 and 6.12 for relevant definitions):

1. (a) An  $\overline{\mathbf{H}}_1$  path with  $r = 9$  and  $k = q = r^* = 3$ .
2. (b) An  $\overline{\mathbf{H}}_2$  path with  $r = 11, k = 3, q = 2$  and  $r^* = 4$ .
3. (c) An  $\overline{\mathbf{H}}_3$ -path where  $\mathbf{H}_3 = (H_0, H'_0, H'_0, H_0) \in \mathcal{H}_3$  and we have  $r = 6, k = 4$  and  $q = r^* = 2$ .

*Proof.* We recommend that the reader refers to the examples in Figure 6 to help visualise some of the ideas in this proof. Note that as the endpoints of an  $\overline{\mathbf{H}}_1$ -path are isolated, we have that the base set of an  $(\overline{\mathbf{H}}_1(r \times s), \overline{H}_{\det})$  absorbing gadget  $F$  is also an isolated set of vertices and so  $\Phi_{F,W} = \Phi_{F \setminus W}$ . Defining  $K_k + K_k$  as two copies of  $K_k$  which meet in a singular vertex, we have that  $F \setminus W$  consists of disjoint copies of  $K_k$  and  $r \times s$  disjoint copies of  $K_k + K_k$ , one for each  $\overline{\mathbf{H}}_1$ -path used in  $F$ . Therefore Lemma 2.7 (1) shows that  $\Phi_{K_k} \geq Cn$ , and repeated applications of Lemma 2.7 (3) show that  $\Phi_{K_k + K_k} \geq Cn$  and in turn  $\Phi_{F \setminus W} \geq Cn$  as required.

Case 2 is similar. Here we have that  $q = r - k \lfloor r/k \rfloor < k$  and each of the base vertices  $w$  of  $F$  lie in a copy, say  $F_w$ , of the graph defined as follows. Take a copy of  $K_{q+1}^-$  and a copy of  $K_k$  that meet in exactly one vertex, which is one of the vertices of the nonedge in  $K_{q+1}^-$ . Furthermore, we have that the base vertex  $w$  is the other vertex in the nonedge of this copy of  $K_{q+1}^-$ . We have that each of the  $F_w$  is disconnected from the rest of  $F$  and an application of Lemma 2.7 (1), (2) and (3) gives that  $\Phi_{F_w \setminus w} \geq Cn$  and  $\Phi_{F_w, w} \geq Cn^{1/k}$  if  $q \geq 2$ . If  $q = 1$ , then  $F_w$  is an isolated vertex  $w$  and a copy of  $K_k$  so we have  $\Phi_{F_w, w} = \Phi_{F_w \setminus w} \geq Cn$ . Now note that  $F \setminus (\cup_{w \in W} F_w)$  consists of copies of  $K_k, K_{q+1}^-, K_{q+1}$  and a copy of  $K_q$  (in the copy of  $\overline{H}_{\det}$  in  $F$ ) which intersect each other in at most one vertex. Furthermore, one can view  $F \setminus (\cup_{w \in W} F_w)$  as being ‘built up’ from these copies in the following way: there is an ordering (starting with  $\overline{H}_{\det}$ ) on these copies of  $K_k, K_{q+1}^-, K_{q+1}$  and  $K_q$  such that, starting with the empty graph and adding these copies in this order, each new copy shares at most one vertex with the previous copies already added, and at the end of the process we obtain  $F \setminus (\cup_{w \in W} F_w)$ . Each time we add a copy, we can apply Lemma 2.7 (3) and then again to add in the  $F_w$  (to obtain  $F$ ). This leads us to conclude that  $\Phi_{F \setminus W} \geq Cn$  and  $\Phi_{F,W} \geq Cn^{1/k}$  as required.

In Case 3, let  $q = r - k < k$  and let us fix some  $\underline{H} \in \mathcal{H}_3(r \times s)$  which then defines our  $F$ . For each  $w \in W$ , let  $F_w$  be the connected component of  $F$  which contains  $w$ . Due to the definition of  $\mathcal{H}_3$ , and in particular the fact that each  $\underline{H} \in \mathcal{H}_3$  contains a copy of  $H'_0$  as defined in Proposition 6.12, we have that  $F_w \neq F_{w'}$  for all  $w \neq w' \in W$ . Also, for  $q \geq 2$ , it can be seen that  $F_w$  is a graph obtained by sequentially ‘gluing’ copies of  $K_{q+1}^-$  to vertices of degree  $q - 1$  and that  $w$  is a vertex of degree  $q - 1$  in the resulting graph. Similarly to the previous case, applications of Lemma 2.7 (2) and (3) imply that  $\Phi_{F_w \setminus w} \geq Cn$  and  $\Phi_{F_w, w} \geq Cn^{1/k}$  if  $q \geq 2$  and if  $q = 1$ , we see that  $F_w$  is an isolated vertex, namely  $w$  itself. Also as before, we have that  $F \setminus (\cup_{w \in W} F_w)$  consists of copies of  $K_k, K_{q+1}^-, K_{q+1}$  and a copy of  $K_q$  which intersect each other in at most one vertex. Thus, introducing the ordering of these copies as in Case 2, we can apply Lemma 2.7 repeatedly to obtain the desired conclusion.  $\square$

We will use Lemma 6.17 to prove the existence of our desired absorbing gadgets in  $G' = G \cup G(n, p)$ . Before embarking on this however, we need to know how we wish our absorbing gadgets (in particular their base sets) to intersect in  $G'$ . This is given by the notion of a *template* in the following subsection.

### 6.2.2 Defining an absorbing structure

A *template*  $T$  with *flexibility*  $m \in \mathbb{N}$  is a bipartite graph on  $7m$  vertices with vertex classes  $I$  and  $J_1 \sqcup J_2$ , such that  $|I| = 3m, |J_1| = |J_2| = 2m$ , and for any  $\bar{J} \subset J_1$ , with  $|\bar{J}| = m$ , the induced graph  $T[V(T) \setminus \bar{J}]$  has a perfect matching. We call  $J_1$  the *flexible* set of vertices for the template. Montgomery first introduced the use of such templates when applying the absorbing method in his work on spanning trees in random graphs [42, 43]. There, he used a sparse template of maximum degree 40, which we will also use. It is not difficult to prove the existence of such templates for large enough  $m$  probabilistically; see e.g. [42, Lemma 2.8]. The idea has since been used by various authors in different settings [16, 17, 20, 21, 39, 44]. We will use a template here as an auxiliary graph in order to build an absorbing structure for our purposes.

**Definition 6.18.** *Let  $m, t^* \in \mathbb{N}$  and  $T = (I = \{1, \dots, 3m\}, J_1 \sqcup J_2 = \{1, \dots, 2m\} \sqcup \{2m + 1, \dots, 4m\}, E(T))$  be a bipartite template with maximum degree  $\Delta(T) \leq 40$  and flexibility  $m$  as defined above. Further, let*

$$\mathcal{H} := \mathcal{H}(K_{r+1}^-, \leq t^*) = \{(K_{r+1}^-, t) : 1 \leq t \leq t^*\}$$

be the set of vectors of length at most  $t^*$  whose entries are all  $K_{r+1}^-$ .

A ( $t^*$ -bounded) absorbing structure  $\mathcal{A} = (\Phi, Z, Z_1)$  of flexibility  $m$  in a graph  $G'$  consists of a vertex set  $Z = Z_1 \sqcup Z_2 \subset V(G')$  which we label  $Z_1 := \{z_1, \dots, z_{2m}\}$  and  $Z_2 := \{z_{2m+1}, \dots, z_{4m}\}$  and a set  $\Phi := \{\varphi_1, \dots, \varphi_{3m}\}$  of embeddings of absorbing gadgets into  $G'$ . We require the following properties:

- For  $i \in [3m]$ , setting  $N(i) := \{j : (i, j) \in E(T) \subset I \times J\}$  and  $n(i) := |N(i)|$ , we have that  $\varphi_i$  is an embedding of some  $(\underline{H}, K_r)$ -absorbing gadget  $F_i$  such that  $\underline{H} \in \mathcal{H}(r \times n(i))$  and the base set of  $F_i$ , which we denote  $W_i$ , is mapped to  $\{z_j : j \in N(i)\} \subseteq Z$  by  $\varphi_i$ .
- The embeddings of the absorbing gadgets are vertex disjoint other than the images of the base sets. That is, for all  $i \in [3m]$ ,  $\varphi_i(V(F_i) \setminus W_i) \subseteq V(G') \setminus Z$  and  $\varphi_i(V(F_i) \setminus W_i) \cap \varphi_{i'}(V(F_{i'}) \setminus W_{i'}) = \emptyset$  for all  $i \neq i' \in [3m]$ .

We call  $Z_1$  the flexible set of the absorbing structure.

Thus the absorbing structure is an embedding of a larger graph which is formed of  $3m$  disjoint absorbing gadgets whose base vertices are then identified according to a template of flexibility  $m$ . We will refer to the vertices of  $\mathcal{A}$  which are the vertices which feature in the embedding of this larger graph. That is,

$$V(\mathcal{A}) := \bigsqcup_{i \in [3m]} \varphi(V(F_i) \setminus W_i) \bigsqcup Z.$$

**Remark 6.19.** If  $\mathcal{A}$  is a  $t^*$ -bounded absorbing structure of flexibility  $m$ , then it has less than  $125t^*r^2m$  vertices in total.

In our proof, we will bound  $t^*$  by a constant and look for an absorbing structure on a small linear number of vertices. The key property of the absorbing structure is that it inherits the flexibility of the template that defines it, but in the context of  $K_r$ -tilings, as detailed in the following remark.

**Remark 6.20.** If  $G'$  contains an absorbing structure  $\mathcal{A} = (\Phi, Z, Z_1)$  of flexibility  $m$ , then for any subset of vertices  $\bar{Z} \subset Z_1$  such that  $|\bar{Z}| = m$ , there is a  $K_r$ -tiling in  $G'$  covering precisely  $V(\mathcal{A}) \setminus \bar{Z}$ .

Indeed given such a  $\bar{Z}$ , letting  $\bar{J}$  be the corresponding indices from  $J$ , we have that  $T \setminus \bar{J}$  has a perfect matching. The matching then indicates, for each  $i \in [3m]$ , which vertex  $z_{j_i}$  of  $Z$  to use in a tiling of the corresponding absorbing gadget. That is, for each  $i$ , if  $\varphi_i$  is ‘matched’ to  $z_{j_i}$  by the perfect matching, then we take the  $K_r$ -tiling covering  $\varphi_i(F_i \setminus W_i) \cup \{z_{j_i}\}$  (which exists by the key property of the absorbing gadget mentioned after Definition 6.14) and then take their union.

### 6.2.3 The existence of an absorbing structure

In order to prove the existence of an absorbing structure, we must find embeddings of absorbing gadgets in our graph. In the previous section we found many embeddings of certain absorbing gadgets with deterministic edges and thus it remains to find embeddings of complementary absorbing gadgets, using only random edges. Therefore we will turn to Lemma 2.8, which is a general result regarding embeddings in random graphs. However, there is still some work to do in the application of this lemma and the following proposition shows how we can use Lemma 2.8 repeatedly in order to embed a larger graph. We state the proposition in a more general form than just for showing the existence of absorbing structures as we will also use the result at other points in the proof. As the statement of the proposition is somewhat technical, we recommend that the reader sees how it is applied in Corollaries 6.22, 6.24 and 6.25 to help with digesting it.

**Proposition 6.21.** *Let  $\kappa_d, \kappa_w, \kappa_e, \kappa_v, k, \in \mathbb{N}$  and  $\beta > 0$ . Then there exists  $\eta_0 > 0$  and  $C > 0$  such that the following holds for any  $0 < \eta < \eta_0$ ,  $n \in \mathbb{N}$  and  $t = \eta n \in \mathbb{N}$ .*

*Suppose that  $F_1, \dots, F_t$  are labelled graphs with distinguished base vertex sets  $W_i \subset V(F_i)$  such that  $|W_i| \leq \kappa_w$ ,  $v_i := |V(F_i) \setminus W_i| \leq \kappa_v$ ,  $e(F_i[W_i]) = 0$  and  $e(F_i) \leq \kappa_e$  for all  $i \in [t]$ . Suppose that  $p = p(n)$  such that  $\Phi_{F_i \setminus W_i} = \Phi_{F_i \setminus W_i}(n, p) \geq Cn$  and  $\Phi_{F_i, W_i} = \Phi_{F_i, W_i}(n, p) \geq Cn^{\frac{1}{k}}$  for all  $i \in [t]$ . Let  $V$  be an  $n$ -vertex set, and  $U_1, \dots, U_t \subset V$  be subsets such that  $|U_i| = |W_i|$  for each  $i \in [t]$ , and defining*

$$d(i) := |\{j \in [t] : U_i \cap U_j \neq \emptyset\}|,$$

*we have that  $d(i) \leq \kappa_d$ . Finally, suppose that  $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_t$  are families of vertex sets such that each  $\mathcal{F}_i$  contains  $\beta n^{v_i}$  ordered subsets of  $V$  of size  $v_i$ .*

Then a.a.s. there is a set of embeddings  $\varphi_1, \varphi_2, \dots, \varphi_t$  such that each  $\varphi_i$  embeds a copy of  $F_i$  into  $G(n, p)$  on  $V$  with  $W_i$  being mapped to  $U_i$  and  $V(F_i) \setminus W_i$  being mapped to a set in  $\mathcal{F}_i$  which does not intersect  $\cup_{i \in [t]} U_i$ . Furthermore for  $i \neq i'$ , we have that  $\varphi_i(V(F_i) \setminus W_i) \cap \varphi_{i'}(V(F_{i'}) \setminus W_{i'}) = \emptyset$ .

*Proof.* Fix  $\kappa'_v := \kappa_v + 2\kappa_e$  and let  $\eta_0 < \beta^2 2^{-(\kappa'_v + 2\kappa_e + 9)} (\kappa'_v! (\kappa'_v + \kappa_e + \kappa_w))^{-1}$ . The idea here is to greedily extract the desired embeddings, finding them one at a time in  $G(n, p)$ . To achieve this, we use the multi-round exposure trick, having a constant number of phases such that in each phase we find a collection of embeddings. At the beginning of each phase we ‘reveal’ another copy of  $G(n, p)$  on the same vertex set and focus *only* on the indices for which we have not yet found a suitable embedding, showing that in any sufficiently large subset of these indices there is an index for which we can find a suitable embedding. At each phase, we will apply Lemma 2.8 and so we first need to slightly adjust the sets we are considering in order to be in the setting of that lemma.

Firstly let us adjust each  $F_i$  so that it has  $\kappa'_v$  non-base vertices and  $\kappa_e$  edges. To each  $F_i$  add  $\kappa_e - e(F_i)$  isolated edges. Then add isolated vertices until the resulting graph has  $\kappa'_v + |W_i|$  vertices and redefine  $F_i$  as the resulting graph. Note that if  $p = p(n)$  is such that  $\Phi_{F_i \setminus W_i}(n, p) \geq Cn$  and  $\Phi_{F_i, W_i}(n, p) \geq Cn^{\frac{1}{k}}$  for the original  $F_i$  as in the statement of the proposition, then these conditions are preserved under the above changes to  $F_i$  for each  $i$ , by Lemma 2.7. We also arbitrarily extend each set in each  $\mathcal{F}_i$  to get sets of size  $\kappa'_v$ . As we can extend with any vertices not already in the set, it can be seen that we can have families  $\mathcal{F}_i$  of size at least  $\beta' n^{\kappa'_v}$  for some  $\beta' > \beta / (2^{\kappa'_v})$  which we now fix. Clearly, a set of valid embeddings of these new  $F_i$  (where the new vertices of  $F_i$  are mapped to the new vertices from a set in  $\mathcal{F}_i$ <sup>10</sup>) will also yield a set of embeddings of the original graphs we were interested in.

Now let us turn to the phases of our algorithm. We will generate  $G(n, p)$  in  $k + 1$  rounds so that  $G(n, p) = \cup_{j=1}^{k+1} G_j$  with each  $G_j$  an independent copy of  $G(n, p')$ , where  $p'$  is such that  $(1 - p) = (1 - p')^{k+1}$ . Note that for any graph  $F$ , vertex subset  $W \subset V(F)$ , constant  $c > 0$  and probability  $p$ , one has that  $\Phi_{F \setminus W}(n, cp) = c' \Phi_{F \setminus W}(n, p)$  for some constant  $c' = c'$  between 1 and  $e^{e(F)}$ . Likewise, multiplication of the probability by some constant  $c > 0$  results in multiplication of  $\Phi_{F, W}$  by some constant factor. Hence, choosing  $C > 0$  sufficiently large, we can guarantee that if  $\Phi_{F_i \setminus W_i}(n, p) \geq Cn$  and  $\Phi_{F_i, W_i}(n, p) \geq Cn^{\frac{1}{k}}$  as in the statement of the proposition, then  $\Phi_{F_i \setminus W_i}(n, p') \geq C'n$  and  $\Phi_{F_i, W_i}(n, p') \geq C'n^{\frac{1}{k}}$  with  $C'$  such that  $C' \geq \frac{2^{\kappa'_v + 9} \kappa'_v! \kappa'_v}{\beta'^2}$ . We fix such a  $C > 0$  and for  $j = 1, \dots, k$ , we define  $t_j := \eta n^{1 - \frac{j-1}{k}} ((\kappa_d + 1) \log n)^{j-1}$  and  $s_j := t_j n^{-\frac{1}{k}} \log n = \eta n^{1 - \frac{j}{k}} (\kappa_d + 1)^{j-1} (\log n)^j$ . We also define  $t_{k+1} := (\kappa_d + 1) s_k = \eta ((\kappa_d + 1) \log n)^k$ ,  $s_{k+1} := 1$  and  $t_{k+2} := 0$ .

Now, as discussed, we look to choose embeddings one by one in order to reach the desired conclusion. Therefore, for the sake of brevity, at any point in the argument let us say that an embedding  $\varphi_i$  of  $F_i$  is *valid* if it maps  $W_i$  to  $U_i$  and maps  $V(F_i \setminus W_i)$  to a set in  $\mathcal{F}_i$  which is disjoint from  $U := \cup_{i \in [t]} U_i$  and also disjoint from  $\varphi_{i'}(V(F_{i'} \setminus W_{i'}))$  for all indices  $i' \in [t]$  for which we have already chosen an embedding. Our claim is that a.a.s. (with respect to  $G(n, p)$ ) we can repeatedly choose valid embeddings until we have found embeddings for all  $t$  indices in  $T := \{1, \dots, t\}$ . We therefore need to show that we never get stuck and that this greedy algorithm always finds a valid embedding. In order to do this, we split the algorithm into  $k + 1$  phases and rely on the edges of  $G_j$  in the  $j^{\text{th}}$  phase where we will find  $t_j - t_{j+1}$  valid embeddings. We will show that for all  $j \in [k + 1]$ , conditioned on the fact that the algorithm has succeeded so far, we have that a.a.s (with respect to

<sup>10</sup>This will be guaranteed in applications of Lemma 2.8 as the lemma is concerned with *labelled* embeddings.

$G_j = G(n, p')$ ) the algorithm will succeed for a further phase. The conclusion then follows easily as there are constantly many phases.

So let us analyse the  $j^{\text{th}}$  phase and condition on the fact that the process has been successful so far and so there are  $t_j$  indices that remain for us to find embeddings for. Let us further fix a specific set<sup>11</sup> of  $t_j$  indices  $T_j \subseteq T$  that remain and some set of already chosen valid embeddings  $\{\varphi_i : i \in R_j\}$  where  $R_j := T \setminus T_j$ . By the law of total probability, it suffices to condition on this fixed set of embeddings so far and show that a.a.s. (with respect to  $G_j$ ) we can repeatedly find valid embeddings, each time removing the corresponding index from  $T_j$ , until there are  $t_{j+1}$  indices remaining. So let  $V_j'' := \cup_{i \in R_j} \varphi_i(V(F_i)) \cup U$  and for  $i \notin R_j$ , define  $\mathcal{F}_i^{(j)} := \{S \in \mathcal{F}_i : S \cap V_j'' = \emptyset\}$ . We have that  $|\mathcal{F}_i^{(j)}| \geq \frac{\beta'}{2} n^{\kappa'_v}$  as  $|V_j''| < \frac{\beta'}{2} n$  due to our condition on  $\eta_0$ . We then apply Lemma 2.8 to the sets  $\mathcal{F}_i^{(j)}$  such that  $i \in T_j$ , and where  $t_j, s_j, \frac{\beta'}{2}, \kappa'_v, \kappa'_w, \kappa_e$  and  $p'$  play the role of  $t, s, \beta, L, v, w, e$  and  $p$  respectively. Let us check that the conditions needed for the lemma are satisfied. Indeed, we certainly have that  $\kappa'_v t_j \leq \kappa'_v t \leq \frac{\beta' n}{8\kappa'_v}$ ,  $s_j \kappa_w \leq \frac{\beta' n}{8\kappa'_v}$  and  $\binom{t_j}{s_j} \leq \binom{n}{s_j} \leq 2^n$ . Moreover, when  $1 \leq j \leq k$ ,

$$C' s_j n^{\frac{1}{k}} = C' t_j \log n \geq \left( \frac{2^{k'_v+9} \kappa'_v! \kappa'_v}{\beta'^2} \right) t_j \log n \text{ and } C' n \geq \left( \frac{2^{\kappa'_v+9} \kappa'_v!}{\beta'^2} \right) n,$$

by our definition of  $C'$  whilst for  $j = k + 1$ ,  $C' s_{k+1} n^{\frac{1}{k}} = \omega(t_{k+1} \log n)$ . This verifies the conditions in (2.2) in all cases and so we conclude that a.a.s., given any set  $V_j'$  of at most  $\kappa'_v t_j$  vertices and any set  $S_j$  of  $s_j$  indices in  $T_j$  such that the sets  $U_i$  with  $i \in S_j$  are pairwise disjoint, there is an index  $i^* \in S_j$  and a valid embedding of  $F_{i^*}$  in  $G_j$  which avoids  $V_j'$ . This then implies that the greedy process will succeed throughout this phase. Indeed, we can now initiate with  $V_j' = \emptyset$  and repeatedly find indices  $i \in T_j$  for which we have a valid embedding  $\varphi_i$ . We add this embedding to our chosen embeddings, add the vertices of it to  $V_j'$  and delete the index  $i$  from  $T_j$ . The conclusion that we drew from Lemma 2.8 above asserts that we continue this process until we have  $t_{j+1}$  indices remaining in  $T_j$ , which is precisely what we need. Indeed, for  $1 \leq j \leq k$ , if we have more than  $t_{j+1}$  indices in  $T_j$  left then by the upper bound on  $d(i)$  for  $i$  in  $T_j$  taking a maximal set  $S \subset T_j$  such that  $U_i$  are all pairwise disjoint for  $i \in S$ , we have that  $|S| \geq t_{j+1}/(\kappa_d + 1) \geq s_j$ . In the final phase when  $j = k + 1$  we can simply find embeddings one at a time as  $s_{k+1} = 1$ . This concludes the proof.  $\square$

As corollaries, we can conclude the existence of absorbing structures in  $G \cup G(n, p)$ . We split the cases here as Case 1 and 2 are much simpler.

**Corollary 6.22.** *Let  $k, r \in \mathbb{N}$  such that either  $2 \leq k \leq r/2$  or  $k = r$  and let  $\gamma > 0$ . There exists  $\eta_0 > 0$  and  $C > 0$  such that if  $p \geq Cn^{-2/k}$  and  $G$  is an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ , then for any  $0 < \eta < \eta_0$  and any set of  $2\eta n$  vertices  $X_1 \subseteq V(G)$ , a.a.s. there exists a 4-bounded absorbing structure  $\mathcal{A} = (\Phi, Z, Z_1)$  in  $G' := G \cup G(n, p)$  of flexibility  $m := \eta n$ , which has flexible set  $Z_1 = X_1$ .*

*Proof.* We look to apply Proposition 6.21 and simply need to establish the hypothesis of the proposition. Consider a bipartite template  $T = (I = \{1, \dots, 3m\}, J_1 \sqcup J_2 = \{1, \dots, 2m\} \sqcup \{2m + 1, \dots, 4m\}, E(T))$  as in Definition 6.18; recall such a template exists [42]. Fix  $Z_1 = X_1 = \{z_1, \dots, z_{2m}\}$  and choose an arbitrary set of  $2m$  vertices  $Z_2 \subset V(G) \setminus Z_1$  which we label

<sup>11</sup>Note that when  $j = 1$  we must have that  $T_1 = T$ .

$\{z_{2m+1}, \dots, z_{4m}\}$ . Now towards applying Proposition 6.21, we set  $t := 3m$  and for  $i \in [t]$  we define the sets  $U_i := \{z_j : j \in N(i)\}$  where  $N(i)$  is as in Definition 6.18. Note that we can set  $\kappa_d := 1600$  as we start with a template  $T$  with  $\Delta(T) \leq 40$ , so for any set  $N(i) \subset J$  (of at most 40 vertices), there are at most 1600 indices  $i' \in I = [3m]$  such that  $N(i') \cap N(i) \neq \emptyset$ .

Now, fixing  $i$ , the collection  $\mathcal{F}_i$ , which we will use when applying Proposition 6.21, will be obtained from Proposition 6.16. Indeed, this proposition implies, along with Propositions 6.6 and 6.7, that there is some  $\beta > 0$  such that the following holds with  $a = 1$  if  $r/k \in \mathbb{N}$  (Case 1) and  $a = 2$  otherwise (Case 2).

**Claim 6.23.** *For any set  $U$  of at most 40 vertices, there is an  $(\mathbf{H}_a(r \times |U|), H_{\det})$ -absorbing gadget<sup>12</sup>  $F'$  such that there are at least  $\beta n^{|F'| - |U|}$  embeddings of  $F'$  in  $G$  which map the base set of the absorbing gadget to  $U$ .*

For each  $i$ , apply Claim 6.23 with  $U_i$  playing the role of  $U$  to obtain a collection  $\mathcal{F}_i$  of ordered vertex sets from  $V(G)$  that combined with  $U_i$  each span such an absorbing gadget  $F'_i = F'$ . For each such embedding of  $F'_i$ , if we have an ordered  $(\overline{\mathbf{H}}_a(r \times |U_i|), \overline{H}_{\det})$ -absorbing gadget  $F_i$  (in  $G(n, p)$ ), on the same vertex set, then we obtain the desired embedding  $\varphi_i$  of a  $(\mathbf{K}_a(r \times |U_i|), K_r)$ -absorbing gadget in  $G \cup G(n, p)$ , where  $\mathbf{K}_a$  is a  $(K_{r+1}^-, 2a)$ -path. Applying Proposition 6.21 with small enough  $\eta > 0$  thus gives us the absorbing structure, upon noticing that the conditions on  $\Phi_{F_i, W_i}$  and  $\Phi_{F_i \setminus W_i}$  are satisfied by Lemma 6.17.  $\square$

The third case, when  $r/2 < k \leq r - 1$ , follows the exact same method of proof. The main difference comes from the fact that we do not have many absorbing gadgets for *all* small sets of vertices in the deterministic graph but only for sets which lie in one part of the partition dictated by Lemma 6.12. Therefore we look to find an absorbing structure in each part of the partition. Thus when we apply Proposition 6.21, we do so to find all these absorbing structures at once, in order to guarantee that these absorbing structures are disjoint. The conclusion is as follows.

**Corollary 6.24.** *Let  $r/2 < k \leq r - 1$  be integers, and define  $q := r - k$ ,  $c := \lceil r/q \rceil$  and  $\gamma > 0$ . Then there exists  $\alpha > 0$  such that the following holds for all  $0 < \epsilon < \alpha\gamma/4$ . There exists  $C = C(r, k, \gamma, \epsilon) > 0$  and  $\eta_0 = \eta_0(r, k, \gamma, \epsilon) > 0$  such that if  $p \geq Cn^{-2/k}$  and  $G$  is an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ , then for any  $0 < \eta < \eta_0$  there is a partition  $\mathcal{P} = \{V_1, V_2, \dots, V_\varrho, W\}$  of  $V(G)$  into at most  $c$  parts with the following properties:*

- $|V_i| \geq \alpha n$  for  $i \in [\varrho]$ ;
- $|W| \leq \epsilon n$ ;
- $\delta(G[V_i]) \geq (1 - \frac{k}{r} + \frac{\gamma}{4})|V_i|$  for each  $i \in [\varrho]$ ;
- For any collection of subsets  $X_i \subset V_i$  such that  $1.8\eta|V_i| \leq |X_i| \leq 2\eta|V_i|$  for all  $i \in [\varrho]$ , there a.a.s. exists a set of  $c(2^{c+1} + 1)$ -bounded absorbing structures  $\{\mathcal{A}_i = (\Phi_i, Z_i, Z_{i1}) : i \in [\varrho]\}$  in  $G' := G \cup G(n, p)$  such that each  $\mathcal{A}_i$  has flexibility  $m_i := |X_i|/2$  and has flexible set  $Z_{i1} = X_i$ . Furthermore  $V(\mathcal{A}_i) \cap V(\mathcal{A}_{i'}) = \emptyset$  for all  $i \neq i' \in [\varrho]$ .

<sup>12</sup>Recall here the definition of  $\mathbf{H}_1$  from Proposition 6.6, of  $\mathbf{H}_2$  from Proposition 6.7 and  $H_{\det}$  from Definition 4.1. The notation  $\mathbf{H}_a(r \times |U|)$  is also defined as in Definition 6.15.



*Proof.* We begin by applying Proposition 6.12 to get a vertex partition  $\mathcal{P}$  with at most  $c - 1$  parts and in each part  $U \in \mathcal{P}$  we remove any vertex  $v$  which has internal degree  $d_U(v) = |N_G(v) \cap U| < (1 - \frac{k}{r} + \frac{\gamma}{2})|U|$ , and add  $v$  to  $W$ . The resulting partition is the partition we will use. Choosing  $\epsilon_{6.12}$  in the application of Proposition 6.12 to be less than  $\epsilon/c$ , we have that the first three bullet points are satisfied. Below we show the last bullet point, and to aid readability we temporarily fix  $i = 1$ .

Now given a set of  $X_1 \subset V_1$  we choose a set  $Z_2 \subset V_1 \setminus X_1$  such that  $|Z_2| = 2m_1$ . Further, according to some template  $T = (I = \{1, \dots, 3m_1\}, J_1 \sqcup J_2 = \{1, \dots, 2m_1\} \sqcup \{2m_1 + 1, \dots, 4m_1\}, E(T))$  as in Definition 6.18, we label  $X_1$  according to  $J_1$  and  $Z_2$  according to  $J_2$  and identify sets  $U_{i'} \subseteq X_1$  for each  $i' \in [3m_1]$  according to the neighbourhood of  $i'$  in  $T$ . As in Corollary 6.22, by Propositions 6.12 and 6.16 there exists some  $\beta > 0$  such that for each  $i' \in [3m_1]$ , fixing  $s_{i'} = |U_{i'}|$  the following holds. There is some  $\underline{H}_{i'} \in \mathcal{H}_3(r \times s_{i'})$  and some  $(\underline{H}_{i'}, H_{\det})$ -absorbing gadget  $F'_{i'}$  such that there are at least  $\beta n^{|F'_{i'}| - s_{i'}}$  embeddings of  $F'_{i'}$  in  $G$  which map the base set of  $F'_{i'}$  to  $U_{i'}$ . Each of these embeddings gives a candidate vertex set for which we could embed an  $(\underline{H}_{i'}, \overline{H_{\det}})$ -absorbing gadget, say  $F_{i'}$  to get a copy of a  $(\underline{K}, K_r)$ -absorbing gadget in  $G'$ , with base set  $U_{i'}$ , where  $\underline{K} \in \mathcal{H}(r \times s_{i'})$  and  $\mathcal{H} = \mathcal{H}(K_{r+1}^-, \leq c(2^{c+1} + 1))$ . Using Lemma 6.17, we can now apply Proposition 6.21 (provided  $\eta > 0$  is sufficiently small) to get the desired embeddings of all the  $F_{i'}$  which an absorbing structure  $\mathcal{A}_1$  as in the statement of the corollary. We in fact apply Proposition 6.21 for all  $i \in [\rho]$  at once which gives the collection of absorbing structures as required.  $\square$

Before proving the upper bound in our main result, Theorem 1.5, we give one last consequence of Proposition 6.21 which will be useful for us.

**Corollary 6.25.** *Suppose that  $2 \leq k \leq r$  and  $\gamma, \beta > 0$ . Then there exists  $\alpha = \alpha(r, k, \gamma, \beta) > 0$  and  $C > 0$  such that the following holds. Suppose  $G$  is an  $n$ -vertex graph with disjoint vertex sets  $U, W$  such that  $|U| \leq \alpha n$ ,  $|W| \geq \beta n$  and for all  $v \in U \cup W$ ,  $|N_G(v) \cap W| \geq (1 - \frac{k}{r} + \gamma)|W|$  and  $p = p(n)$  is such that  $p \geq Cn^{-2/k}$ . Then a.a.s. in  $G \cup G(n, p)$  there is a set of  $|U|$  disjoint copies  $K_r$  so that each copy of  $K_r$  contains a vertex of  $U$  and  $r - 1$  vertices of  $W$ .*

*Proof.* Firstly, let  $r^* := \lceil r/k \rceil$ . By the fact that  $|N_G(v) \cap W| \geq (1 - \frac{k}{r} + \gamma)|W|$  for all  $v \in U \cup W$ , we have that each vertex  $u \in U$  is in at least  $(\frac{\beta\gamma}{2}n)^{r^*}$  distinct copies of  $K_{r^*+1}^-$  in  $G$  such that the other vertices of each copy lie in  $W$ , and  $u$  is contained in the nonedge of each  $K_{r^*+1}^-$ . Thus by Lemma 2.6, there exists some  $\beta' > 0$  such that each  $u \in U$  is in  $\beta'n^{r-1}$  copies of  $H_{\det}$  with the other vertices of each copy in  $W$ , and  $u$  in the part of size  $q := r - (r^* - 1)k$  in  $H_{\det}$ . Let  $\mathcal{F}_u$  be the collection of  $(r - 1)$ -sets of vertices in  $W$  that, together with  $u$ , give rise to these copies of  $H_{\det}$  containing  $u$ . Set  $F_u := \overline{H_{\det}} = K_r - E(H_{\det})$  with an identified vertex  $w_u$  in the clique of size  $q$  in  $\overline{H_{\det}}$ . Thus an ordered embedding in  $G(n, p)$  of  $F_u$  which maps  $w_u$  to  $u$  and  $V(F_u) \setminus \{w_u\}$  to an ordered set in  $\mathcal{F}_u$  will give an embedding of  $K_r$  in  $G \cup G(n, p)$  containing  $u$  and vertices of  $W$ . By Lemma 2.7 we have that  $\Phi_{F_u, w_u} \geq Cn^{1/k}$  and  $\Phi_{F_u \setminus w_u} \geq Cn$ . Thus, provided  $\alpha > 0$  is sufficiently small, an application of Proposition 6.21 gives the desired set of embeddings of  $K_r$  in  $G \cup G(n, p)$ .  $\square$

## 7 Proof of the upper bound of Theorem 1.5

In this section we prove the upper bound of Theorem 1.5. Fix some sufficiently large  $n \in r\mathbb{N}$  and let  $G$  be an  $n$ -vertex graph with  $\delta(G) \geq (1 - \frac{k}{r} + \gamma)n$ . We will show that there exists  $C = C(\gamma, k, r) > 0$  such that if  $p \geq Cn^{-2/k}$ , then  $G' := G \cup G(n, p)$  a.a.s. contains a perfect  $K_r$ -tiling. Again, we split the proof according to the parameters. We first treat Cases 1 and 2 together (i.e. when  $2 \leq k \leq r/2$

or  $k = r$ ). Here we avoid many of the technicalities which occur in Case 3 and the main scheme of the proof is clear.

*Proof of Cases 1 and 2.* Suppose  $2 \leq k \leq r/2$  or  $k = r$ , and let  $C, C' > 0$  be chosen such that we can express  $G(n, p) = \cup_{j=1}^4 G_j$  with each  $G_j$  a copy of  $G(n, p')$  where  $p' \geq C'n^{-2/k}$  and  $C' > 0$  is large enough to be able to draw the desired conclusions in what follows. Now fix  $0 < \eta < \min\{\frac{\gamma}{2000r^2}, \eta_0\}$  where  $\eta_0$  is as in Corollary 6.22 and consider  $X' \subseteq V(G)$  to be the subset generated by taking every vertex in  $V(G)$  in  $X'$  with probability  $1.9\eta$ , independently of the other vertices. With high probability, by Chernoff's theorem, we have that  $1.8\eta n \leq |X'| \leq 2\eta n$  and for every vertex  $v \in V(G)$ ,  $|N_G(v) \cap X'| \geq (1 - \frac{k}{r} + \frac{3\gamma}{4})|X'|$ . Take an instance of  $X'$  where this is the case and let  $X := X'$  if  $|X'|$  is even and  $X := X' \cup \{x\}$  for some arbitrary vertex  $x \in V(G) \setminus X'$  if  $|X'|$  is odd. Apply Corollary 6.22 to get a 4-bounded absorbing structure  $\mathcal{A} = (\Phi, Z, Z_1)$  in  $G \cup G_1$  with flexibility  $|X|/2$  and flexible set  $Z_1 = X$ . Remark 6.19 implies  $|V(\mathcal{A})| \leq 500r^2\eta n \leq \gamma n/4$ .

Then letting  $V' := V(G) \setminus V(\mathcal{A})$ , we have that  $\delta(G[V']) \geq (1 - \frac{k}{r} + \frac{\gamma}{2})|V'|$ . Choose  $\alpha := \min\{\alpha_{6.25}, \frac{\gamma\eta}{4r}\}$ , where  $\alpha_{6.25}$  is the constant obtained when applying Corollary 6.25 with  $r, k, \gamma/2, \eta$  playing the role of  $r, k, \gamma, \beta$  respectively.

Apply Theorem 5.1 to obtain a  $K_r$ -tiling  $\mathcal{K}_1$  in  $(G \cup G_2)[V']$  covering all but at most  $\alpha n$  vertices of  $V'$ . Let  $Y$  denote the set of those vertices in  $V'$  uncovered by  $\mathcal{K}_1$ . Apply Corollary 6.25 to obtain a  $K_r$ -tiling  $\mathcal{K}_2$  in  $(G \cup G_3)[X \cup Y]$  which covers  $Y$  and covers precisely  $(r-1)|Y| \leq \gamma\eta n/2 \leq \gamma|X|/2$  vertices of  $X$ . Let  $\tilde{X}$  be the set of those vertices in  $X$  not covered by  $\mathcal{K}_2$ . We have that  $\delta(G[\tilde{X}]) \geq (1 - \frac{k}{r} + \frac{\gamma}{4})|\tilde{X}|$  so we can apply Theorem 5.1 to obtain a  $K_r$ -tiling  $\mathcal{K}'_3$  in  $(G \cup G_4)[\tilde{X}]$  which covers all but at most  $|X|/4$  vertices of  $\tilde{X}$ .

By Remark 6.20 we know that for any subset  $X_1$  of  $X$  of size  $|X|/2$ , there is a  $K_r$ -tiling covering precisely  $V(\mathcal{A}) \setminus X_1$ . Thus,  $|V(\mathcal{A})| - |X|/2$  is divisible by  $r$ . Therefore, as the only vertices in  $V(G)$  uncovered by  $V(\mathcal{K}_1 \cup \mathcal{K}_2)$  are those from  $(V(\mathcal{A}) \setminus X) \cup \tilde{X}$ , there must be a subtiling  $\mathcal{K}_3 \subseteq \mathcal{K}'_3$  which covers all but exactly  $|X|/2$  vertices of  $\tilde{X}$ .

Let  $\bar{X}$  be the set of vertices of  $X$  that are covered by cliques in  $\mathcal{K}_2 \cup \mathcal{K}_3$ . Thus  $|\bar{X}| = |X|/2$  and by Remark 6.20 there is a  $K_r$ -tiling  $\mathcal{K}_4$  in  $G \cup G_1$  covering precisely  $V(\mathcal{A}) \setminus \bar{X}$ . Hence,  $\mathcal{K} := \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3 \cup \mathcal{K}_4$  gives a perfect  $K_r$ -tiling of  $G \cup G(n, p)$  as required.  $\square$

If  $r/2 < k \leq r-1$ , we have to overcome a few technicalities. The idea is to apply Corollary 6.24 and to apply the same approach as above in each of the parts of the resulting partition to find a  $K_r$ -tiling. Of course we also have to incorporate the vertices of the exceptional class  $W$  into copies of cliques in our tiling; this is straightforward using Corollary 6.25. So we cover these vertices first before embarking on tiling the majority of the graph.

More subtle is a problem that arises from divisibility. That is, when we tile each part according to the scheme above, we cannot guarantee that we are left with a subset of the flexible set of the right size to apply the key property of the absorbing structure. Therefore we embed 'crossing' copies of  $K_r$  in our flexible sets in order to resolve this divisibility hurdle at the end of our process. We find these copies in the following manner. Consider the graph  $F := K_{\lceil \frac{r-1}{2} \rceil, \lfloor \frac{r-1}{2} \rfloor}$ . Because of our minimum degree condition and Lemma 2.6, every part  $V_i$  contains at least  $\gamma'n^{r-1}$  copies of  $F$  for some  $\gamma' > 0$ . Now let  $\bar{F}$  be the graph consisting of a copy of  $K_{\lceil \frac{r-1}{2} \rceil + 1}$  and a copy of  $K_{\lfloor \frac{r-1}{2} \rfloor + 1}$  joined at a single vertex  $x$ , say. If we consider  $F$  and  $\bar{F} - x$  to have the same vertex set so that  $\bar{F} - x = K_{r-1} - E(F)$ , then  $F \cup \bar{F}$  is a copy of  $K_r$ . Also note that it follows from Lemma 2.7 that  $\Phi_{\bar{F}} \geq C'n$  for  $p \geq C'n^{-2/k}$ . We will look for embeddings of  $K_r = F \cup \bar{F}$  in  $G \cup G(n, p)$  such that the vertex  $x$  is mapped to one part of the partition and the  $r-1$  other vertices lie in another part

of the partition.

*Proof of Case 3.* Suppose  $r/2 < k \leq r-1$ ,  $q := r-k$  and  $c := \lceil r/q \rceil$ . Now let  $C, C' > 0$  be chosen so that we can express  $G(n, p) = \cup_{j=1}^4 G_j \cup_{i=1}^c (G_{i1} \cup G_{i2})$  with each  $G_j, G_{i1}$  and  $G_{i2}$  a copy of  $G(n, p')$  where  $p' \geq C'n^{-2/k}$  and  $C' > 0$  is large enough to be able to draw the desired conclusions in what follows.

We use our first copy of  $G(n, p')$  to find the crossing copies of  $K_r$  discussed above. Apply Corollary 6.24, letting  $\alpha_1 > 0$  be the outcome of the corollary with input  $r, k, \gamma$ . Choose  $0 < \epsilon < \min\{\alpha_1\gamma/(8r), \alpha_{6.25}\}$ , where  $\alpha_{6.25}$  is the constant obtained when applying Corollary 6.25 with  $r, k, \gamma/4, 1/2$  playing the roles of  $r, k, \gamma, \beta$  respectively. Thus, Corollary 6.24 yields a partition  $V_1, \dots, V_\varrho, W$  of  $V(G)$  where  $\varrho \leq c$  and  $|W| \leq \epsilon n$ . Note  $\delta(G[V_i]) \geq (1 - \frac{k}{r} + \frac{\gamma}{4})|V_i|$  for each  $i \in [\varrho]$ . Thus for each  $i \in [\varrho-1]$  and every subset  $V' \subseteq V(G)$  of at least  $n - c(r-1)r$  vertices, Lemma 2.6 implies there exists some  $\gamma' = \gamma'(r, k, \gamma) > 0$  such that there are at least  $\gamma'n^r$  choices of pairs  $(S, v) \in \binom{V' \cap V_{i+1}}{r-1} \times (V' \cap V_i)$  such that  $S$  hosts a copy of the graph  $F := K_{\lceil \frac{r-1}{2} \rceil, \lfloor \frac{r-1}{2} \rfloor}$  discussed above.

Therefore, using that  $\Phi_{\overline{F}} \geq C'n$  with the graph  $\overline{F}$  as described above, we can apply Corollary 2.9 to conclude that for any subset  $V'$  of vertices of at least  $n - c(r-1)r$  vertices and any  $i \in [\varrho-1]$ , there is a copy of  $K_r$  in  $G \cup G_1$  which has  $r-1$  vertices in  $V_{i+1}$  and one vertex in  $V_i$ . Therefore, we can greedily choose copies of  $K_r$  so that we have a set  $\mathcal{R} := \cup_{i \in [\varrho-1]} \mathcal{R}_i$  of disjoint copies of  $K_r$  in  $G \cup G_1$  such that  $\mathcal{R}_i$  contains  $r-1$  copies of  $K_r$  with one vertex in  $V_i$  and  $r-1$  vertices of  $V_{i+1}$ . Let  $\mathcal{R}_\varrho := \emptyset$  and  $R_i := V(\mathcal{R}) \cap V_i$  for  $i \in [\varrho]$ , where  $V(\mathcal{R})$  denotes the vertices which feature in cliques in  $\mathcal{R}$ . Note that  $|R_1| = r-1$ ,  $|R_2| = |R_3| = \dots = |R_{\varrho-1}| = r(r-1)$  and  $|R_\varrho| = (r-1)^2$ . We will incorporate these  $R_i$  into our flexible sets in order to use the copies of  $K_r$  that they define to fix divisibility issues that arise in the final stages of the argument.

Now fix  $0 < \eta < \min\{\frac{\gamma}{4000c^2r^2}, \eta_0\}$  where  $\eta_0$  is as in Corollary 6.24 and for each  $i \in [\varrho]$  consider  $X'_i \subseteq V(G)$  to be a subset selected by taking every vertex in  $V_i \setminus R_i$  with probability  $1.9\eta$ , independently of the other vertices. With high probability, by Chernoff's theorem, we have that  $1.8\eta|V_i| \leq |X'_i| \leq 2\eta|V_i| - r(r-1)$  and for every vertex  $v \in V_i$ ,  $|N_G(v) \cap X'_i| \geq (1 - \frac{k}{r} + \frac{\gamma}{8})|X'_i|$ . Therefore, for each  $i$ , take an instance of  $X'_i$  where this is the case and let  $X_i := X'_i \cup R_i$  if  $|X'_i| + |R_i|$  is even and  $X_i := X'_i \cup R_i \cup \{x\}$  for some arbitrary vertex  $x \in V_i \setminus (X'_i \cup R_i)$  if  $|X'_i| + |R_i|$  is odd. Apply Corollary 6.24 to get a collection  $\{\mathcal{A}_i = (\Phi_i, Z_i, Z_{i1}) : i \in [\varrho]\}$  of absorbing structures in  $G \cup G_2$  such that each  $\mathcal{A}_i$  has flexibility  $|X_i|/2$  and flexible set  $Z_{i1} = X_i$ . By Remark 6.19 we have that  $A := \cup_{i \in [\varrho]} V(\mathcal{A}_i)$  is such that  $|A| \leq 125c^2c^{+2}r^2\eta n \leq \gamma n/8$ .

Therefore, setting  $V' := V(G) \setminus (W \cup A)$ , we have that for every  $w \in W \cup V'$ ,  $|N_G(w) \cap V'| \geq (1 - \frac{k}{r} + \frac{\gamma}{4})|V'|$  and so an application of Corollary 6.25 yields a  $K_r$ -tiling  $\mathcal{K}_1$  in  $G \cup G_3$  of  $|W|$  cliques, each using one vertex of  $W$  and  $r-1$  vertices of  $V'$ . Setting  $V'' := V(G) \setminus (A \cup V(\mathcal{K}_1))$ , we have that  $\delta(G[V'']) \geq (1 - \frac{k}{r} + \frac{\gamma}{8})|V''|$ . So, as in the previous proof, we let  $\alpha_2 := \min\{\alpha_{6.25}, \frac{\gamma\eta}{16r}\}$ , where  $\alpha_{6.25}$  is obtained from Corollary 6.25 (where  $\gamma/8$  and  $\eta$  play the roles of  $\gamma$  and  $\beta$  respectively), and we apply Theorem 5.1 to obtain a  $K_r$ -tiling  $\mathcal{K}_2$  in  $(G \cup G_4)[V'']$  covering all but at most  $\alpha_2 n$  vertices of  $V''$ . Let  $Y$  be the set of vertices from  $V''$  uncovered by  $\mathcal{K}_2$  and set  $Y_i := Y \cap V_i$  for each  $i \in [\varrho]$ .

Now for each  $i \in [\varrho]$  a simple application of Corollary 6.25 yields a  $K_r$ -tiling  $\mathcal{K}_{i1}$  in  $G \cup G_{i1}$  which covers  $Y_i$  and uses precisely  $(r-1)|Y_i| \leq \gamma\eta n/16 \leq \gamma|X'_i|/16$  vertices of  $X'_i$ . Note that we do not use any vertices of  $R = \cup_{i \in [\varrho]} R_i$  in these cliques. For each  $i \in [\varrho]$  let  $\tilde{X}_i$  be the vertices of  $X_i \setminus R_i$  not involved in copies of  $K_r$  in  $\mathcal{K}_{i1}$ . As  $\delta(G[\tilde{X}_i]) \geq (1 - \frac{k}{r} + \frac{\gamma}{16})|\tilde{X}_i|$ , we can apply Theorem 5.1 to

obtain a  $K_r$ -tiling  $\mathcal{K}'_i$  in  $(G \cup G_{i2})[\tilde{X}_i]$  which covers all but at most  $|X_i|/4$  vertices of  $\tilde{X}_i$ , for each  $i \in [\varrho]$ .

Note that we will not use the full tilings  $\mathcal{K}'_i$  in our final tiling. So (ignoring for now the tilings  $\mathcal{K}'_i$ ), it remains to cover the vertices in  $(V(\mathcal{A}_i) \setminus X_i) \cup R_i \cup \tilde{X}_i$  for each  $i \in [\varrho]$ . We do so by means of the following algorithm. We initiate with the  $\mathcal{K}'_i, \mathcal{R}_i$  as above and set  $\bar{Z}_i := V(\mathcal{K}_{i1}) \cap X_i$  and  $\mathcal{K}_{i2} := \emptyset$  for all  $i \in [\varrho]$  and  $i' = 1$ . Now whilst  $|\bar{Z}_{i'}| \leq |X_{i'}|/2 - r + 1$ , remove a clique from  $\mathcal{K}'_{i'}$ , add it to  $\mathcal{K}_{i'2}$  and add its vertices to  $\bar{Z}_{i'}$ . Once this process stops, add  $|X_{i'}|/2 - |\bar{Z}_{i'}|$  copies of  $K_r$  in  $\mathcal{R}_{i'}$  to  $\mathcal{K}_{i'2}$ , and add all their vertices in  $X_j$  to  $\bar{Z}_j$  for  $j = i', i' + 1$ . If  $i' \leq \varrho - 1$ , repeat this process, setting  $i' = i' + 1$ . Note that when  $i' = \varrho$ ,  $\mathcal{R}_{i'} = \emptyset$  and there are no cliques which we could add in this process. However, setting  $\mathcal{K}_0 = \mathcal{K}_1 \cup \mathcal{K}_2 \cup_{i \in [\varrho]} (\mathcal{K}_{i1} \cup \mathcal{K}_{i2})$  we have that  $|V(\mathcal{K}_0)|$ ,  $n$ , and  $|V(\mathcal{A}_i)| - |X_i|/2$  are divisible by  $r$  for each  $i$ , so we can deduce that the algorithm takes no cliques from  $\mathcal{R}_\varrho$  and terminates with  $|\bar{Z}_i| = |X_i|/2$  for all  $i \in [\varrho]$ .

Finally, by the key property of the absorbing structure (Remark 6.20), we have that for each  $i \in [\varrho]$ , there is a  $K_r$ -tiling  $\mathcal{K}_{i3}$  in  $G \cup G_2$  covering  $V(\mathcal{A}_i) \setminus \bar{Z}_i$  and thus  $\cup_{i \in [\varrho]} \mathcal{K}_{i3} \cup \mathcal{K}_0$  is the desired perfect  $K_r$ -tiling in  $G \cup G(n, p)$ .  $\square$

## 8 Concluding remarks

In this paper we have almost completely resolved the perfect  $K_r$ -tiling problem for randomly perturbed graphs. The only cases that Theorem 1.5 does not resolve is when  $\alpha = 1/r, 2/r, \dots, (r-2)/r$ . Note however for  $\alpha := 1 - k/r$  where  $2 \leq k \leq r - 1$  we know that

$$n^{-2/k} \leq p(K_r, \alpha) \leq n^{-2/(k+1)}.$$

In fact, one can slightly improve the lower bound, giving that  $p(K_r, \alpha) \geq p(K_k, 0) = n^{-2/k} (\log n)^{\frac{2}{k^2-k}}$ . Indeed, as in our lower bound construction (Section 3), take  $G$  to be complete graph on  $n$  vertices with a clique of size  $kn/r$  removed. Letting  $A$  be the resulting independent set of vertices, if  $p = o(p(K_k, 0))$  then a.a.s. we have that the number of copies of  $K_{k+1}$  in  $G(n, p)$  is less than, say,  $(\log n)^3$  by Markov's inequality, whilst the number of vertices in  $A$  which do not lie in copies of  $K_k$  is at least  $n^{1-o(1)}$ , as can be seen by a second moment calculation (see e.g. [24, Theorem 3.22]). This precludes the existence of a perfect  $K_r$ -tiling in  $G \cup G(n, p)$ , as the average intersection of a clique in such a tiling with the vertex set  $A$  would be  $k$  and we cannot tile  $G(n, p)[A]$  with a family of cliques whose average size is  $k$  given the restrictions above. This leaves a gap between the upper and lower bounds and it would be very interesting to resolve the problem for these ‘boundary’ cases.

It is also of interest to consider the analogous problem for perfect  $H$ -tilings for arbitrary graphs  $H$ . Note that whilst the main result from [3] determines  $p(H, \alpha)$  for all graphs  $H$  and  $0 < \alpha < 1/|H|$ , the problem is still wide open for larger values of  $\alpha$ . The methods from our paper are likely to be useful for the general problem, though we suspect how  $p(H, \alpha)$  ‘jumps’ as  $\alpha$  increases will depend heavily on the structure of  $H$ . Thus we believe it would be a significant challenge to prove such a general result.

## Acknowledgments

Much of the research in this paper was undertaken whilst the authors were visiting the Instituto Nacional de Matemática Pura e Aplicada in Rio de Janeiro as part of the Graphs@IMPA thematic

program in 2018. We are grateful to IMPA for the nice working environment. We are also extremely grateful to Louis DeBiasio and Shagnik Das for many useful conversations and to the referees for their careful reviews and their suggestions. In particular, we thank one of the referees who brought our attention to the improvement on the lower bound in the concluding remarks.

## References

- [1] N. Alon and J. H. Spencer, *The probabilistic method*, 1st ed., John Wiley & Sons, 1992. [↑1.2](#)
- [2] S. Antoniuk, A. Dudek, C. Reiher, A. Ruciński, and M. Schacht, *High powers of Hamiltonian cycles in randomly augmented graphs*, arXiv preprint arXiv:2002.05816 (2020). [↑1.3](#), [1.3](#)
- [3] J. Balogh, A. Treglown, and A. Z. Wagner, *Tilings in randomly perturbed dense graphs*, *Combinatorics, Probability and Computing* **28** (2019), 159–176. [↑1.3](#), [1.3](#), [1.3](#), [3](#), [8](#)
- [4] W. Bedenknecht, J. Han, Y. Kohayakawa, and G. O. Mota, *Powers of tight Hamilton cycles in randomly perturbed hypergraphs*, *Random Structures & Algorithms* **55** (2019), no. 4, 795–807. [↑1.3](#)
- [5] P. Bennett, A. Dudek, and A. Frieze, *Adding random edges to create the square of a Hamilton cycle*, arXiv preprint arXiv:1710.02716 (2017). [↑1.3](#)
- [6] T. Bohman, A. Frieze, M. Krivelevich, and R. Martin, *Adding random edges to dense graphs*, *Random Structures & Algorithms* **24** (2004), no. 2, 105–117. [↑1.3](#)
- [7] T. Bohman, A. Frieze, and R. Martin, *How many random edges make a dense graph Hamiltonian?*, *Random Structures & Algorithms* **22** (2003), no. 1, 33–42. [↑\(document\)](#), [1.3](#)
- [8] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk, and Y. Person, *Universality for bounded degree spanning trees in randomly perturbed graphs*, *Random Structures & Algorithms* **55** (2019), no. 4, 854–864. [↑1.3](#)
- [9] J. Böttcher, R. Montgomery, O. Parczyk, and Y. Person, *Embedding spanning bounded degree graphs in randomly perturbed graphs*, *Mathematika* **66** (2020), no. 2, 422–447. [↑1.3](#)
- [10] K. Corrádi and A. Hajnal, *On the maximal number of independent circuits in a graph*, *Acta Mathematica Hungarica* **14** (1963), no. 3–4, 423–439. [↑1.1](#)
- [11] A. Czygrinow, L. DeBiasio, H. A. Kierstead, and T. Molla, *An extension of the Hajnal–Szemerédi theorem to directed graphs*, *Combinatorics, Probability and Computing* **24** (2015), no. 5, 754–773. [↑1.1](#)
- [12] A. Czygrinow, L. DeBiasio, T. Molla, and A. Treglown, *Tiling directed graphs with tournaments*, *Forum of Mathematics, Sigma* **6** (2018), e2. [↑1.1](#)
- [13] S. Das and A. Treglown, *Ramsey properties of randomly perturbed graphs: cliques and cycles*, *Combinatorics, Probability and Computing*, to appear (2020). [↑1.3](#)
- [14] A. Dudek, C. Reiher, A. Ruciński, and M. Schacht, *Powers of Hamiltonian cycles in randomly augmented graphs*, *Random Structures & Algorithms* **56** (2020), no. 1, 122–141. [↑1.3](#)
- [15] P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, *Combinatorica* **3** (1983), no. 2, 181–192. [↑2.3](#)
- [16] A. Ferber, G. Kronenberg, and K. Luh, *Optimal threshold for a random graph to be 2-universal*, *Transactions of the American Mathematical Society* **372** (2019), no. 6, 4239–4262. [↑6.2.2](#)
- [17] A. Ferber and R. Nenadov, *Spanning universality in random graphs*, *Random Structures & Algorithms* **53** (2018), no. 4, 604–637. [↑6.2.2](#)
- [18] A. Hajnal and E. Szemerédi, *Proof of a conjecture of P. Erdős*, *Combinatorial theory and its applications* **2** (1970), no. 4, 601–623. [↑\(document\)](#), [1.1](#), [1.1](#)
- [19] J. Han, *Decision problem for perfect matchings in dense  $k$ -uniform hypergraphs*, *Transactions of the American Mathematical Society* (2017), 5197–5218. [↑6.1.1](#), [6.1.1](#)
- [20] J. Han, Y. Kohayakawa, P. Morris, and Y. Person, *Clique-factors in sparse pseudorandom graphs*, *European Journal of Combinatorics* **82** (2019), 102999. [↑6.2.2](#)
- [21] ———, *Finding any given 2-factor in sparse pseudorandom graphs efficiently*, *Journal of Graph Theory* (2020), 1–22. [↑6.2.2](#)



- [22] J. Han and A. Treglown, *The complexity of perfect matchings and packings in dense hypergraphs*, Journal of Combinatorial Theory, Series B **141** (2020), 72–104. [↑6.1.1](#), [6.1.1](#)
- [23] J. Han and Y. Zhao, *Hamiltonicity in randomly perturbed hypergraphs*, Journal of Combinatorial Theory, Series B, to appear (2020). [↑1.3](#)
- [24] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, John Wiley & Sons, 2010. [↑1.3](#), [2.4](#), [2.4](#), [3.1](#), [5](#), [8](#)
- [25] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs*, Random Structures & Algorithms **33** (2008), no. 1, 1–28. [↑\(document\)](#), [1.2](#), [1.2](#), [1.3](#)
- [26] F. Joos and J. Kim, *Spanning trees in randomly perturbed graphs*, Random Structures & Algorithms **56** (2020), no. 1, 169–219. [↑1.3](#)
- [27] P. Keevash and R. Mycroft, *A multipartite Hajnal–Szemerédi theorem*, Journal of Combinatorial Theory, Series B **114** (2015), 187–236. [↑1.1](#)
- [28] H. A. Kierstead and A. V. Kostochka, *A short proof of the Hajnal–Szemerédi theorem on equitable colouring*, Combinatorics, Probability and Computing **17** (2008), no. 2, 265–270. [↑1.1](#)
- [29] J. H. Kim, *Perfect matchings in random uniform hypergraphs*, Random Structures & Algorithms **23** (2003), no. 2, 111–132. [↑1.2](#)
- [30] C. Knierim and P. Su,  *$K_r$ -factors in graphs with low independence number*, arXiv preprint arXiv:1912.00230 (2019). [↑1.3](#), [1.3](#)
- [31] J. Komlós, *Tiling Turán theorems*, Combinatorica **20** (2000), no. 2, 203–218. [↑2.1](#)
- [32] J. Komlós and M. Simonovits, *Szemerédi’s regularity lemma and its applications in graph theory*, 1996. [↑2.2](#), [2.2](#), [2.3](#), [2.4](#)
- [33] M. Krivelevich, *Triangle factors in random graphs*, Combinatorics, Probability and Computing **6** (1997), no. 3, 337–347. [↑1.2](#)
- [34] M. Krivelevich, M. Kwan, and B. Sudakov, *Cycles and matchings in randomly perturbed digraphs and hypergraphs*, Combinatorics, Probability and Computing **25** (2016), no. 6, 909–927. [↑1.3](#)
- [35] ———, *Bounded-degree spanning trees in randomly perturbed graphs*, SIAM Journal on Discrete Mathematics **31** (2017), no. 1, 155–171. [↑1.3](#)
- [36] M. Krivelevich, B. Sudakov, and P. Tetali, *On smoothed analysis in dense graphs and formulas*, Random Structures & Algorithms **29** (2006), 180–193. [↑1.3](#)
- [37] D. Kühn and D. Osthus, *Embedding large subgraphs into dense graphs*, Surveys in combinatorics, 2009, pp. 137–167. [↑1.1](#)
- [38] ———, *The minimum degree threshold for perfect graph packings*, Combinatorica **29** (2009), no. 1, 65–107. [↑1.1](#)
- [39] M. Kwan, *Almost all Steiner triple systems have perfect matchings*, arXiv preprint arXiv:1611.02246 (2016). [↑6.2.2](#)
- [40] A. Lo and K. Markström, *F-factors in hypergraphs via absorption*, Graphs and Combinatorics **31** (2015), no. 3, 679–712. [↑1.3](#), [4](#)
- [41] A. McDowell and R. Mycroft, *Hamilton  $\ell$ -cycles in randomly-perturbed hypergraphs*, Electronic Journal of Combinatorics **25** (2018), P4.36. [↑1.3](#)
- [42] R. Montgomery, *Embedding bounded degree spanning trees in random graphs.*, arXiv preprint arXiv:1405.6559 (2014). [↑1.3](#), [4](#), [6.2.2](#), [6.2.3](#)
- [43] ———, *Spanning trees in random graphs*, Adv. Math. **356** (2019), 106793. [↑1.3](#), [4](#), [6.2.2](#)
- [44] R. Nenadov and Y. Pehova, *On a Ramsey–Turán variant of the Hajnal–Szemerédi theorem*, SIAM Journal on Discrete Mathematics **34** (2020), no. 2, 1001–1010. [↑1.3](#), [6.2.2](#)
- [45] R. Nenadov and M. Trujić, *Sprinkling a few random edges doubles the power*, 2018. arXiv preprint arXiv:1811.09209. [↑1.3](#), [1.3](#)
- [46] A. Shokoufandeh and Y. Zhao, *Proof of a tiling conjecture of Komlós*, Random Structures & Algorithms **23** (2003), no. 2, 180–205. [↑2.1](#)
- [47] E. Szemerédi, *Regular partitions of graphs*, Problèmes Combinatoires et Théorie des Graphes Colloques Internationaux CNRS 260, 1978, pp. 399–401. [↑2.2](#)



- [48] A. Treglown, *On directed versions of the Hajnal–Szemerédi theorem*, *Combinatorics, Probability and Computing* **24** (2015), no. 6, 873–928. [↑1.1](#)
- [49] Y. Zhao, *Recent advances on Dirac-type problems for hypergraphs*, *Recent trends in combinatorics*, 2015, pp. 145–165. [↑1.1](#)