AN IMPROVED UPPER BOUND ON THE DENSITY OF UNIVERSAL RANDOM GRAPHS

DOMINGOS DELLAMONICA JR.[1]

YOSHIHARU KOHAYAKAWA[2]

VOJTĚCH RÖDL[3]

ANDRZEJ RUCIŃSKI[4]

[1,2,3,4]*Department of Mathematics and Computer Science*
Emory University, Atlanta, GA 30322, USA

[2]*Instituto de Matemática e Estatística, Universidade de São Paulo*
Rua do Matão 1010, 05508-090 São Paulo, Brazil

[4]*Department of Discrete Mathematics*
Adam Mickiewicz University, 61-614 Poznań, Poland

Abstract. We give a polynomial time randomized algorithm that, on receiving as input a pair (H, G) of n-vertex graphs, searches for an embedding of H into G. If H has bounded maximum degree and G is suitably dense and pseudorandom, then the algorithm succeeds with high probability. Our algorithm proves that, for every integer $d \geq 3$ and a large enough constant $C = C_d$, as $n \to \infty$, asymptotically almost all graphs with n vertices and at least $Cn^{2-1/d} \log^{1/d} n$ edges contain as subgraphs all graphs with n vertices and maximum degree at most d.

Date: 2014/03/18, 1:04pm.

A preliminary version of this work [13] has appeared in the Proceedings of LATIN 2012.

[1]Supported by a CAPES-Fulbright scholarship.

[2]Partially supported by FAPESP (2013/03447-6, 2013/07699-0), CNPq (310974/2013-5 and 477203/2012-4), NSF (DMS 1102086) and NUMEC/USP (Project MaCLinC/USP).

[3]Supported by the NSF grants DMS 0800070 and DMS 1102086.

[4]Supported by the Polish NSC grant N201 604940 and the NSF grant DMS 1102086.

1. Introduction

Given graphs H and G, an *embedding* of H into G is an injective edge-preserving map $f\colon V(H) \to V(G)$, that is, an injective map such that for every $e = \{u, v\} \in E(H)$, we have $f(e) = \{f(u), f(v)\} \in E(G)$. We shall say that a graph H *is contained in* G *as a subgraph* if there is an embedding of H into G. Given a family of graphs $\mathcal{H}$, we say that G is *universal with respect to* $\mathcal{H}$, or $\mathcal{H}$*-universal*, if every $H \in \mathcal{H}$ is contained in G as a subgraph.

The construction of sparse universal graphs for various families of graphs received a considerable amount of attention; see, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 10, 11] and the references therein. Here, we are particularly interested in (almost) tight $\mathcal{H}$-universal graphs, i.e., graphs whose number of vertices is (almost) equal to $\max_{H\in\mathcal{H}} |V(H)|$.

Let $d \in \mathbb{N}$ be a fixed constant and let $\mathcal{H}(n, d) = \{H \subset K_n : \Delta(H) \le d\}$ denote the class of (pairwise non-isomorphic) n-vertex graphs with maximum degree bounded by d and $\mathcal{H}(n, n; d) = \{H \subset K_{n,n} : \Delta(H) \le d\}$ be the corresponding class for balanced bipartite graphs.

By counting all unlabeled d-regular graphs on n vertices one can easily show that every $\mathcal{H}(n, d)$-universal graph must have

$$\Omega\left(n^{2-2/d}\right) \tag{1}$$

edges (see [3] for details). This lower bound was almost matched by a construction from [4], which was subsequently improved in [2] and [1]. Those constructions were designed to achieve a nearly optimal bound and as such they did not resemble a "typical" graph with the same number of edges. To pursue this direction, in [3], the $\mathcal{H}(n, d)$-universality of random graphs was also investigated.

For random graphs a slightly better lower bound than (1) is known. Indeed, any $\mathcal{H}(n, d)$-universal graph must contain as a subgraph a union of $\lfloor \frac{n}{d+1} \rfloor$ vertex-disjoint copies of K_{d+1}, and, in particular, all but at most d vertices must each belong to a copy of K_{d+1}. Therefore, recalling the threshold for the latter property (see [17, Theorem 3.22 (i)]), we conclude that the expected number of edges needed for the $\mathcal{H}(n, d)$-universality of $G_{n,p}$ must be

$$\Omega\left(n^{2-2/(d+1)} (\log n)^{1/\binom{d+1}{2}}\right), \tag{2}$$

a quantity bigger than (1).

We say that $G_{n,p}$ possesses a property $\mathcal{P}$ *asymptotically almost surely* (**a.a.s.**) if $\mathbf{P}[G_{n,p} \in \mathcal{P}] = 1 - o(1)$. In [3], it was proved that for a sufficiently large constant C:

- (almost tight universality) $G_{(1+\varepsilon)n,p}$ is **a.a.s.** $\mathcal{H}(n, d)$-universal if $p = Cn^{-1/d} \log^{1/d} n$;
- (bipartite tight universality) $G_{n,n,p}$ is **a.a.s.** $\mathcal{H}(n, n, d)$-universal if $p = Cn^{-1/(2d)} \log^{1/(2d)} n$.

Note that the first result above deals with embeddings of n-vertex graphs into random graphs with larger vertex sets, which makes the embedding somewhat easier. On the other hand, the second result deals with tight universality at the cost of requiring the graphs to be bipartite and with a less satisfactory bound.

Those results were improved by the authors in [12, 14], where it was shown that $G_{n,n,p}$ is **a.a.s.** $\mathcal{H}(n,n,d)$-universal if $p = Cn^{-1/d}\log^{1/d} n$, and $G_{n,p}$ is **a.a.s.** $\mathcal{H}(n,d)$-universal if $p = Cn^{-1/(2d)}\log^{1/(2d)} n$ (for a sufficiently large constant $C > 0$). In this paper, we improve the latter result, by establishing a density threshold for $\mathcal{H}(n,d)$-universality of $G_{n,p}$ which matches the best previous bounds for both, the bipartite tight universality and almost tight universality in general.

Theorem 1.1. *Let $d \geq 3$ be fixed and $p = p(n) = C\, n^{-1/d}\log^{1/d} n$ for some sufficiently large constant C. Then the random graph $G_{n,p}$ is* ***a.a.s.*** *$\mathcal{H}(n,d)$-universal.*

Observe that there is still a gap between the lower bound (2) and the upper bound given by Theorem 1.1.

Remark 1.2. In Theorem 1.1 we assume that $d \geq 3$ since for $d = 2$ our proof would require a few modifications. On the other hand, we feel that for $d = 2$ the true bound is much lower. Possibly as low as $p = n^{-2/3}(\log n)^{1/3}$, which is the threshold for the appearance of a triangle-factor in $G(n,p)$, as proved by Johansson, Kahn, and Vu [19]. We plan to address the case $d = 2$ in a separate paper.

Remark 1.3. An interesting notion of 'almost universality' has been introduced by Frieze and Krivelevich [15]. Given a family of graphs $\mathcal{H}$ and a probability distribution μ on $\mathcal{H}$, a graph Γ is said to be *μ-almost universal for $\mathcal{H}$* if Γ contains a copy of a random graph H sampled from $\mathcal{H}$ according to the distribution μ with high probability. In [15], the case in which $H = G(n, c/n)$ and $\Gamma = G(n,p)$ is investigated. Furthermore, explicit constructions for sparse n-vertex graphs Γ are given in [9] for $H = G(n, c/n)$.

This paper is organized as follows. In the next section we describe a randomized embedding procedure that attempts to find, for any graph $H \in \mathcal{H}(n,d)$ and a graph G on n vertices, an embedding $f\colon V(H) \to V(G)$.

In Section 3 we show that the random graph $G_{n,p}$ with $p \geq C\, n^{-1/d}\log^{1/d} n$ **a.a.s.** satisfies certain properties (conditions (**I**)–(**V**) of Lemma 3.1).

Finally, in Sections 4 and 5 we show that if G satisfies conditions (**I**)–(**V**) of Lemma 3.1 then, for any $H \in \mathcal{H}(n,d)$, the randomized embedding procedure is **a.a.s.** successful (and thus H is embeddable in G) (Lemma 4.1). In particular, any G satisfying (**I**)–(**V**) is $\mathcal{H}(n,d)$-universal and thus Theorem 1.1 follows by combining Lemmas 3.1 and 4.1 (see the end of Section 4). The proof of a technical lemma (Lemma 4.5) is deferred to Section 5, while a probabilistic inequality used therein is established in the appendix.

Throughout the paper we will use the following notation.

- For $v \in V$, let
$$G(v) = \{u \in V \ : \ \{u,v\} \in G\}$$
denote the neighborhood of the vertex v in G.
- For $T \subset V$, let
$$G(T) = \{v \in V \setminus T \ : \ G(v) \cap T \neq \emptyset\} = \bigcup_{u \in T} G(u) \setminus T$$
denote the neighborhood of the set T in G.
- For $T \subset V$, let $G[T]$ denote the subgraph of G induced by T.
- For $U, W \subset V$, $U \cap W = \emptyset$, we denote by $e_G(U,W) = e(U,W)$ the set all of edges of G with one endpoint in U and one in W.
- For a sequence of probability spaces indexed by n, we say that an event occurs **a.a.s.** if the probability of the event is $1 - o(1)$ as $n \to \infty$.

We will also make use of the following definition.

Definition 1.4. For $t \in \mathbb{N}$ and G a graph, a set of vertices $S \subset V(G)$ is called *t-independent* if every pair of distinct vertices in S is at distance at least $t+1$ in G. A 1-independent set is simply called *independent* (and this definition coincides with the usual concept of independence in graph theory).

The following values will be used throughout the paper and are presented here for easy reference:
$$\varepsilon = \varepsilon(d) = \frac{1}{100d^4}, \quad \tau = 2\varepsilon = \frac{1}{50d^4}, \quad t = \lfloor \tau n \rfloor, \quad \omega = C_{\mathrm{L3.1}} \log n, \tag{3}$$
where $C_{\mathrm{L3.1}} = C_{\mathrm{L3.1}}(\delta)$ is the constant of Lemma 3.1.

2. The embedding of H into G

Let $d \geq 3$,
$$\varepsilon = \varepsilon(d) = \frac{1}{100d^4}, \tag{4}$$
and $n_0 = n_0(d)$ be a sufficiently large integer. Let G be a given n-vertex graph, $n \geq n_0$, and $H \in \mathcal{H}(n,d)$. For our analysis, it will be important to have a fixed partition of $V = V(G)$:
$$V = V_0 \cup R_1 \cup \cdots \cup R_{d^2+2}, \text{ where } |R_i| = \lfloor \varepsilon n \rfloor \text{ for all } i = 1, \ldots, d^2+2. \tag{5}$$
(The role of the *buffer* sets R_i will be explained shortly.)

Without loss of generality, we will assume that H is a maximal graph from $\mathcal{H}(n,d)$ in the sense that $|V(H)| = n$, and adding any edge to H increases its maximum degree beyond d. Since in such a graph the vertices with degrees smaller than d must form a clique, there are at most d of them.

We set $X := V(H)$, and fix
$$t = \lfloor \tau n \rfloor, \quad \text{where} \quad \tau = 2\varepsilon = \frac{1}{50d^4}. \tag{6}$$

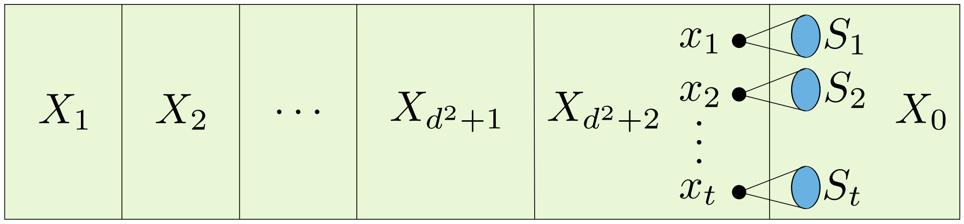

FIGURE 1. The partition of $V(H)$.

In the embedding algorithm we will use the following procedure for pre-processing H.

THE PRE-PROCESSING OF H: Select vertices $x_1, \ldots, x_t \in X$ in such a way that they all have degree d and form a 3-*independent set* in H (recall Def. 1.4). (Owing to our choice of t, we may find these t vertices by a simple greedy algorithm.) Let $S_i = H(x_i)$ for all $i = 1, \ldots, t$, and set

$$X_0 := \bigcup_{j=1}^{t} S_j.$$

Note that by the 3-independence condition, for all $i \neq j$, not only $S_i \cap S_j = \emptyset$, but also there is no edge between S_i and S_j in H, that is, $e_H(S_i, S_j) = 0$.

Next, consider the square H^2 of the graph H, that is, the graph obtained from H by adding edges between all pairs of vertices at distance two. Since the maximum degree of H^2 is at most d^2, by the Hajnal–Szemerédi Theorem (see [20] for a recent algorithmic version) applied to H^2, there is a partition

$$X = X'_1 \cup X'_2 \cup \cdots \cup X'_{d^2+1},$$

such that

- $\big||X'_i| - |X'_j|\big| \leq 1$ for all i, j;
- each set X'_i, $1 \leq i \leq d^2 + 1$, is independent in H^2, and thus, 2-independent in H.

Finally, set

$$X_i = X'_i \setminus \{x_1, \ldots, x_t\} \setminus X_0, \quad i = 1, \ldots, d^2 + 1,$$

and $X_{d^2+2} = \{x_1, \ldots, x_t\}$. Hence, we obtain a partition

$$X = X_0 \cup X_1 \cup \cdots \cup X_{d^2+2}, \tag{7}$$

where, for $i = 1, \ldots, d^2 + 1$, the sets X_i are 2-independent and

$$|X_i| \geq \frac{n}{d^2+1} - 1 - t(d+1) \geq \frac{n}{2d^2}, \tag{8}$$

while X_{d^2+2} is 3-independent, $|X_{d^2+2}| = t$, and X_0 is a (disjoint) union of the d-element neighborhoods of the vertices in X_{d^2+2}. (See Figure 1 for an illustration of this partition.) The numbering of the sets $X_0, \ldots, X_{d^2+2}$

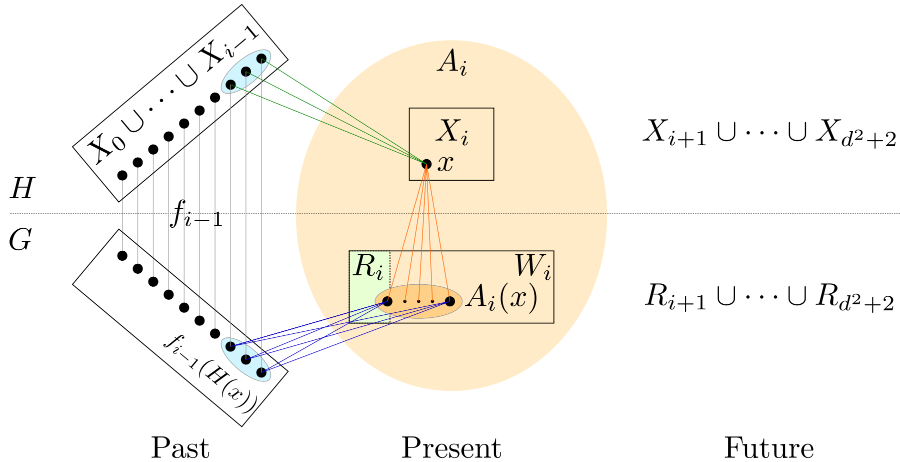

FIGURE 2. An illustration of the graphs G, H, and A_i.

corresponds to the order in which these sets will be embedded into a graph G by the embedding algorithm.

Another building block of our embedding algorithm is a procedure which, given a partial embedding f_{i-1} of $H[X_0 \cup \cdots \cup X_{i-1}]$ into G, constructs an auxiliary graph A_i. The edges of A_i correspond to valid extensions of the embedding f_{i-1}.

THE AUXILIARY GRAPH A_i: For $i = 1, \ldots, d^2+2$ and a partial embedding

$$f_{i-1} \colon X_0 \cup \cdots \cup X_{i-1} \to V \setminus \bigcup_{j=i}^{d^2+2} R_j, \tag{9}$$

let A_i be a bipartite graph with classes X_i and W_i, where,

$$W_i := V \setminus \mathrm{im}(f_{i-1}) \setminus \bigcup_{j=i+1}^{d^2+2} R_j \tag{10}$$

and the edge set is given by

$$E(A_i) = \big\{(x, v) \in X_i \times W_i \,:\, f_{i-1}\big(H(x) \cap (X_0 \cup \cdots \cup X_{i-1})\big) \subset G(v)\big\}. \tag{11}$$

Observe that $A_i(x)$, the neighborhood of x in A_i, is the set of all vertices $v \in W_i$ for which $x \mapsto v$ is a valid extension of the embedding f_{i-1}, while $A_i(v)$ is the set of all vertices $x \in X_i$ for which v is a valid image. See Figure 2 for an illustration of the graph A_i.

Since the set X_i is independent, any matching in A_i saturating X_i corresponds to a valid extension of the embedding f_{i-1}. Hence our objective will be to find such a matching. (The 2-independence of the X_i's will only be used in the analysis of the algorithm for random-like graphs as inputs.)

The embedding will be done in $d^2 + 2$ rounds split into three phases:

- **Phase 1**: The sets $S_1, \dots, S_t$ are mapped *randomly* onto disjoint cliques of $G[V_0]$.
- **Phase 2**: The sets X_i, $i = 1, \dots, d^2 + 1$, are embedded, one by one, into the sets W_i defined above.
- **Phase 3**: The set X_{d^2+2} is embedded onto the set W_{d^2+2} of t remaining vertices of G.

A potential problem for our proposed embedding scheme is that the candidate set for a given vertex $x \in X = V(H)$ may be depleted before we have a chance to embed x. If that happens, there is no way to complete the embedding. Similarly, a vertex $v \in V = V(G)$ may lose all of its neighbors in the auxiliary graph as a result of an unfortunate sequence of extensions. In other words, v can be excluded from all candidate sets and thus cannot be used in the embedding. Since we have to use all vertices of $v \in V$ in the embedding, we must prevent this event as well. Our algorithm incorporates two devices that help to address these problems.

Buffer vertices in G (used in Phases 2 and 3). We will make sure that $\mathrm{im}(f_i) \cap R_{i+1} = \emptyset$ for each $i = 0, \dots, d^2+1$. Indeed, from the definition of W_i in (10),

$$\mathrm{im}(f_i) \subset \mathrm{im}(f_{i-1}) \cup W_i = V \setminus \bigcup_{j=i+1}^{d^2+2} R_j \tag{12}$$

(see also line 5 of Algorithm 1). In particular, the vertices of R_{i+1} can only appear in the image of f_{i+1} or an extension of f_{i+1} (i.e., they are not used by the partial embeddings $f_0, f_1, \dots, f_i$). This way the vertices of R_{i+1} will be reserved as a *buffer* to help embed the set X_{i+1}, provided the sets R_{i+1} will satisfy certain properties in G—see Section 3. Figure 2 shows that R_i may be used in the image of f_i while $R_{i+1} \cup \dots \cup R_{d^2+2}$ is reserved for future use (see (12)).

Buffer vertices in H (used in Phase 3). Since the neighborhoods S_j of the vertices x_j from X_{d^2+2} are embedded during Phase 1, the sets $A_i(v) \cap X_{d^2+2}$, $v \in V$, remain the same throughout Phase 2. This will help to ensure the existence of a perfect matching in A_{d^2+2} in Phase 3, provided the random choices of $f(S_j)$ satisfy certain properties—see Lemma 4.5.

Now we present our embedding algorithm.

This algorithm finds a desired embedding of H into G as long as it is successful in lines 2, 6, and 9. The sets S_i are embedded into V_0 by uniformly sampling a sequence of pairwise disjoint d-subsets $\kappa_1, \dots, \kappa_t \subset V_0$ such that every set κ_i induces a clique in G. Thus, one (trivial) necessary condition for the success of the algorithm is that G contains at least t disjoint cliques K_d. Notice that the map f_0 is an embedding, since the edges within S_i are clearly preserved ($G[\kappa_i]$ is a clique), while $e_H(S_i, S_j) = 0$ holds for all $j \neq i$ by construction.

Algorithm 1: The embedding algorithm

Input : A graph H with n vertices and $\Delta(H) \leq d$ and a graph G together with a vertex partition $V = V_0 \cup R_1 \cup \cdots \cup R_{d^2+2}$ with $|R_i| = \lfloor \varepsilon n \rfloor$ for all $i = 1, \ldots, d^2 + 2$ (see (5)).

Output: An embedding $f\colon V(H) \to V(G)$ (or the algorithm fails).

`// Phase 1`

1 Pre-process H, obtaining a partition $X = X_0 \cup \cdots \cup X_{d^2+2}$ as in (7), where $X_{d^2+2} = \{x_1, \ldots, x_t\}$, $H(x_j) = S_j$ for $j = 1, \ldots, t$, and $X_0 = S_1 \cup \cdots \cup S_t$.

2 Select a sequence of pairwise disjoint d-element sets κ_i $(1 \leq i \leq t)$ so that $G[\kappa_i]$ is a clique for each $i = 1, \ldots, t$: choose κ_1 uniformly at random from all the possibilities and, having chosen $\kappa_1, \ldots, \kappa_j$ $(j < t)$, choose κ_{j+1} uniformly at random from all the possibilities. Stop with failure if this process is unsuccessful.

3 Define a map $f_0\colon X_0 \to \bigcup_{i=1}^{t} \kappa_i$ in such a way that $f_0(S_i) = \kappa_i$ for each $i = 1, \ldots, t$.

`// Phase 2`

4 **for** $i = 1$ *to* $i = d^2 + 1$ **do**

5   Set $W_i = V \setminus \mathrm{im}(f_{i-1}) \setminus \bigcup_{j=i+1}^{d^2+2} R_j$;

6   Construct the auxiliary bipartite graph A_i between the sets X_i and W_i, and find therein a matching M_i of size $|M_i| = |X_i|$. Stop with failure if such a matching does not exist.

7   Define the extension f_i of f_{i-1} by setting $f_i(x) = v$ for all $x \in X_i$, where $(x, v) \in M_i$, and $f_i(x) = f_{i-1}(x)$ for all $x \in X_0 \cup \cdots \cup X_{i-1}$.

`// Phase 3`

8 Set $W_{d^2+2} = V \setminus \mathrm{im}(f_{d^2+1})$ $(\supset R_{d^2+2})$.

9 Construct the auxiliary bipartite graph A_{d^2+2} between the sets X_{d^2+2} and W_{d^2+2}, and find therein a perfect matching M_{d^2+2}. Stop with failure if such a matching does not exist.

10 Define the output embedding f by setting $f(x) = v$ for all $x \in X_{d^2+2}$, where $(x, v) \in M_{d^2+2}$, and $f(x) = f_{d^2+1}(x)$ for all $x \in X \setminus X_{d^2+2}$.

Two more demanding conditions are that the auxiliary bipartite graphs A_i from lines 6 and 9 do possess the required matchings. Superficially, we could have combined the last two phases by including round $d^2 + 2$ into the loop, however we chose not to do so, because of the much more involved analysis of the last round. Indeed, it is a lot harder to prove the existence of a perfect matching in A_{d^2+2} than the existence of a matching saturating one side of A_i when the other side is larger (we show in equation (30) below that $|W_i| \geq |X_i| + \varepsilon n$ for $1 \leq i \leq d^2 + 1$).

It is worth pointing out that the success of Phase 3 relies entirely on the (random) outcome of Phase 1. The algorithm's goal in Phase 3 is to find a perfect matching in the auxiliary bipartite graph A_{d^2+2} (which has classes X_{d^2+2} and W_{d^2+2}). Recall that the neighborhoods $S_j = H(x_j)$ of the vertices $x_j \in X_{d^2+2}$ are completely embedded in Phase 1. Since f_{d^2+1} is an extension of f_0, for each $x_j \in X_{d^2+2}$ we have $f_{d^2+1}(S_j) = f_0(S_j)$. Consequently, by (11),

$$E(A_{d^2+2}) = \big\{(x,v) \in X_{d^2+2} \times W_{d^2+2} \,:\, f_0\big(H(x)\big) \subset G(v)\big\}. \tag{13}$$

This observation is utilized in the analysis of Algorithm 1 in Section 4.

3. Some properties of $G_{n,p}$

In this section we show that a random graph $G_{n,p}$ with $p = p(n)$ as in Theorem 1.1 **a.a.s.** satisfies several properties with respect to the distribution of edges and cliques. These properties are selected in order to jointly guarantee $\mathcal{H}(n,d)$-universality. More specifically, in Section 4 we will show that Algorithm 1 is **a.a.s.** successful on all pairs of input graphs (H,G), where $H \in \mathcal{H}(n,d)$ and G satisfies all these properties.

First we will introduce a few more pieces of notation.

- Given a graph G, $V(G) = V$, and a subset of vertices $U \subset V$, denote by
$$\binom{U}{K_d}$$
the family of all d-element sets $T \subset U$ such that the subgraph of G induced by T is complete, that is, $G[T] \cong K_d$.
- Given a family $\mathcal{X} = \{J_1, \dots, J_r\}$ of pairwise disjoint k-subsets of V and a set $U \subset V$, let $B = B(\mathcal{X}, U)$ be the bipartite graph with vertex classes $\mathcal{X}$ and $U_{\mathcal{X}} := U \setminus \bigcup_{i=1}^r J_i$, where an edge (J_i, v) is included whenever $G(v) \supset J_i$. Furthermore, let
$$\alpha(\mathcal{X}, U) = \big|\{v \in U_{\mathcal{X}} : \deg_B(v) \geq 1\}\big|. \tag{14}$$
If all sets J_i are singletons (i.e., $k = 1$), then we write $B(Y,U)$ instead of $B(\mathcal{X}, U)$, where $Y = \bigcup_{i=1}^r J_i$.
- We write $a = (1 \pm \delta)b$ whenever $(1-\delta)b \leq a \leq (1+\delta)b$.
- For $C = C(\delta)$ defined in Lemma 3.1 below, set
$$\omega = C \log n. \tag{15}$$

Let $\varepsilon = \varepsilon(d) > 0$ be as in (4). Set $V = [n]$ and fix a partition

$$V = V_0 \cup R_1 \cup \cdots \cup R_{d^2+2}$$

satisfying (5). By (4),

$$|V_0| \geq n - (d^2+2)\varepsilon n \geq \frac{3n}{4}. \tag{16}$$

Lemma 3.1. *For every $\delta > 0$, there exists $C > 0$ such that the random graph $G = G_{n,p}$ with $p \geq Cn^{-1/d} \log^{1/d} n$* ***a.a.s.*** *satisfies Properties* (**I**)–(**V**) *below.*

(**I**) (a) *For all $v \in V$,*

$$|G(v) \cap V_0| = (1 + o(1))p|V_0|.$$

(b) *For all $v \neq v' \in V$,*

$$|G(v) \cap G(v') \cap V_0| = (1 + o(1))p^2|V_0|.$$

(c) *For all $v \neq v' \in V$,*

$$|G(v) \cap G(v')| = (1 + o(1))p^2 n.$$

(**II**) (a) *For all $Y \subset V$,*

$$|G(Y) \cap V_0| \geq (1 - 2\delta)p \min\left(|Y|, \delta p^{-1}\right) |V_0|. \tag{17}$$

(b) *For all $Y \subset V$ with $|Y| \geq \omega p^{-1}$ and $U \subset V \setminus Y$ with $|U| \geq \omega p^{-1}$,*

$$|E(B(Y, U))| = (1 \pm \delta)p\, |Y|\, |U|. \tag{18}$$

(**III**) (a) *For all $1 \leq k \leq d$, $r \geq 1$, every family $\mathcal{X} = \{J_1, \ldots, J_r\}$ of pairwise disjoint k-subsets of V, and $U \in \{V_0, R_1, \ldots, R_{d^2+2}, V\}$, we have*

$$\alpha(\mathcal{X}, U) \geq (1 - 2\delta)p^k \min(r, \delta p^{-k})\, |U|. \tag{19}$$

(b) *For all $1 \leq k \leq d$, $r \geq \omega p^{-k}$, every family $\mathcal{X} = \{J_1, \ldots, J_r\}$ of pairwise disjoint k-subsets of V, and $U \subset V \setminus \bigcup_{i=1}^r J_i$ with $|U| \geq \omega p^{-k}$,*

$$|E(B(\mathcal{X}, U))| = (1 \pm \delta)p^k r\, |U|. \tag{20}$$

(**IV**) *We have*

$$\left|\binom{U}{K_d}\right| = (1 \pm \delta)p^{\binom{d}{2}} \binom{|U|}{d} \tag{21}$$

for all $U \subset V$ satisfying at least one of the following conditions:
(a) *$U \subset G(v)$ for some $v \in V$ and $|U| \geq pn/3$, or*
(b) *$U = G(u) \cap G(v)$ for some distinct $u, v \in V$, or*
(c) *$|U| \geq |V|/4$.*

(**V**) *For all $v \in V_0$, the number of d-cliques in $G[V_0]$ containing v is*

$$(1 \pm \delta)p^{\binom{d}{2}} \frac{d}{|V_0|} \binom{|V_0|}{d}.$$

Proof. (**I**)(a), (b) and (c): These properties easily follow from the Chernoff bound (see, e.g., [17], Theorem 2.1, page 26).

(**II**)(a) and (b): These are immediate consequences of (**III**) with $k = 1$. However, in part (a) one needs to choose first an arbitrary $Y' \subseteq Y$ of size $|Y'| = \min\left(|Y|, \delta p^{-1}\right)$.

(**III**)(a): Without loss of generality we assume that $r \le \delta p^{-k}$. Let $Y = \bigcup_{i=1}^{r} J_i$ and note that $B = B(\mathcal{X}, U)$ is a bipartite random graph with vertex classes $\mathcal{X}$ and $U \setminus Y$ and edge probability p^k. We will establish Property (**III**)(a) by counting how many vertices of $U \setminus Y$ are not isolated in $B(\mathcal{X}, U)$.

For each $v \in U \setminus Y$, let $\mathbb{I}_v$ denote the indicator random variable of the event $\deg_B(v) \ge 1$ (that is, some $J_i \subset G(v)$). Notice that $\mathbb{I}_v$ is a Bernoulli random variable. Let q denote the expectation of $\mathbb{I}_v$. By the union bound over the events $J_i \subset G(v)$, $1 \le i \le r$, we have $q \le rp^k$. Using the assumption that $rp^k \le \delta$, and bounds $1 + x \le e^x$ (for all $x \in \mathbb{R}$), $1 - e^{-x} \ge x/(x+1)$ (for $x < 1$), we conclude that

$$q = 1 - (1 - p^k)^r \ge 1 - e^{-rp^k} \ge \frac{rp^k}{1 + rp^k} \ge \frac{rp^k}{1+\delta} > (1-\delta) rp^k.$$

Thus $q = (1 \pm \delta) rp^k$.

Also notice that the variables $\{\mathbb{I}_v : v \in U \setminus Y\}$ are mutually independent. Therefore the distribution of

$$\mathbb{X} := \left| \left\{ v \in U \setminus Y \ : \ \deg_B(v) \ge 1 \right\} \right|$$

is binomial with parameters $|U \setminus Y| = (1 + o(1))|U|$ and q. The expectation of $\mathbb{X}$ is therefore

$$(1 + o(1))(1 \pm \delta) rp^k \, |U|.$$

By the Chernoff bound, we thus have $\mathbb{X} \ge (1 - 2\delta) rp^k \, |U|$ with probability at least

$$1 - \exp\{-cnrp^k\}$$

for some $c = c(\delta) > 0$ (recall that $|U| = \Omega(n)$).

On the other hand, the number of choices of the set Y is less than n^{kr}. Consequently, the probability Property (**III**)(a) fails for $G_{n,p}$ is at most

$$\sum_{r=1}^{\delta p^{-k}} n^{kr} \exp\{-cnrp^k\} = o(1)$$

because $np^k \ge np^d = C^d \log n$ and C is sufficiently large.

(**III**)(b): Here we are just counting the edges of the bipartite graph $B(\mathcal{X}, U)$ defined above. Setting $u = |U|$, the expected number of edges in B is $p^k ru$. Hence, again by the Chernoff bound, the probability that Property (**III**)(b) fails for $G_{n,p}$ is at most

$$\sum_{r \ge \omega p^{-k}} \sum_{u \ge \omega p^{-k}} n^{kr+u} \exp\{-cp^k ru\} = o(1)$$

for $C > 0$ large enough, because $rp^k \ge \omega$ and $up^k \ge \omega$.

(**IV**): Let $\mathbb{X} := \mathbb{X}(d, m, p)$ be a random variable counting the number of copies of K_d in $G_{m,p}$ for some $m \le n$. Let $\delta > 0$ be a fixed small constant.

From the results of [16] and [18, Corollary 1.7], it follows that

$$\mathbf{P}\left[|\mathbb{X}-\mathbf{E}\mathbb{X}|\geq \delta\,\mathbf{E}\mathbb{X}\right]\leq \exp\left\{-c(\delta,d)\,m^2p^{d-1}\right\}, \tag{22}$$

provided

$$m\geq p^{(1-d)/2}=C^{(1-d)/2}(n/\log n)^{\frac{1}{2}-\frac{1}{2d}}. \tag{23}$$

(a): For $v\in V$, expose the random neighborhood $G(v)$. Let us condition on $|G(v)|\leq 1.01pn$ (which is an event occurring with probability at least $1-e^{-\Theta(pn)}$). For any $U\subset G(v)$, $m=|U|\geq pn/3$, the graph $G[U]$ is an instance of $G_{m,p}$. In particular, the assumption (23) on m is satisfied and the bound (22) applies to the random variable $\mathbb{X}=\binom{U}{K_d}$. Moreover, there are fewer than $n\,2^{1.01pn}<e^{2pn}$ choices for v and the set $U\subset G(v)$. In view of (22) and the fact that $pn=o(m^2p^{d-1})$, the union bound yields that with probability

$$1-e^{-\Theta(pn)}-e^{2pn}\exp\left\{-c(\delta,d)\,m^2p^{d-1}\right\}=1-o(1)$$

the equation (21) holds for all $v\in V$ and all $U\subset G(v)$, $m=|U|\geq pn/3$.

(b): For distinct $u,v\in V$, expose the random common neighborhood $U=G(u)\cap G(v)\subset V$. Since **a.a.s.** $|U|=(1+o(1))p^2n$, we condition on $m=|U|>0.99p^2n$. As $d\geq 3$, m satisfies the assumption (23) and therefore we may apply (22) to the random variable $\mathbb{X}=\binom{U}{K_d}$. It follows by the union bound that for all choices of distinct u,v, the set $U=G(u)\cap G(v)$ satisfies (21).

(c): This can be established by the union bound over all large subsets $U\subset V$ using the exponential bound given by (22).

(**V**): By (**I**)(a), **a.a.s.** every $v\in V$ is such that $m_v:=|G(v)\cap V_0|=(1+o(1))p\,|V_0|$. Similarly as before, the results of [16] and [18, Corollary 1.7] applied to the variable $\mathbb{X}=\mathbb{X}(d-1,m_v,p)$ yield

$$\mathbf{P}\left[|\mathbb{X}-\mathbf{E}\mathbb{X}|\geq \frac{\delta}{2}\,\mathbf{E}\mathbb{X}\right]\leq \exp\left\{-c\left(\frac{\delta}{2},d-1\right)m_v^2p^{d-2}\right\},$$

since $m_v>pn/2\gg p^{(2-d)/2}$. There exists a constant $c'>0$ such that for any fixed vertex v, with probability $1-\exp\{-c'p^{d-1}n\}$, we have

$$\begin{aligned}\binom{G(v)\cap V_0}{K_{d-1}}&=(1\pm\delta/2)p^{\binom{d-1}{2}}\binom{(1+o(1))p\,|V_0|}{d-1}\\&=(1\pm\delta)p^{\binom{d}{2}}\frac{d}{|V_0|}\binom{|V_0|}{d}.\end{aligned}$$

Since $\exp\{-c'p^{d-1}n\}=o(1/n)$, Property (**V**) follows from the union bound over all $v\in V$. $\square$

We close this section with a consequence of Properties (**I**)(a) and (**II**)(a) which will be used only in Section 5.

Claim 3.2. *Suppose $W \subset V_0$ satisfies $|W| \le \delta n/4$, where $\delta < 1/48$. Then*

$$\big|\{v \in V \setminus W \ : \ |G(v) \cap W| \ge pn/3\}\big| \le \frac{4}{pn}|W|.$$

Proof. Let $U = \{v \in V \setminus W \ : \ |G(v) \cap W| \ge pn/3\}$ and let $\tilde{U} \subset U$ be an arbitrary subset with

$$|\tilde{U}| = \min\{|U|, \delta/p\}. \tag{24}$$

Further, set

$$T = \{w \in W \ : \ |G(w) \cap \tilde{U}| \ge 2\}.$$

We will show that $e(\tilde{U}, T)$ is very small. Consequently, since the vertices in $W \setminus T$ can each absorb at most one edge coming from $\tilde{U}$ and there are many such edges, the set $W \setminus T$ has to be significantly larger than $\tilde{U}$. However, W itself is not very large, and hence $\tilde{U}$ must be small. In fact, we will show that $|\tilde{U}| < \delta/p$, and thus by (24), that $\tilde{U} = U$.

We have

$$\begin{aligned} |G(\tilde{U}) \cap V_0| &\le |T| + e(\tilde{U}, V_0 \setminus T) \\ &= |T| + e(\tilde{U}, V_0)V_0| - e(\tilde{U}, T) \\ &\overset{(\mathbf{I})(a)}{=} |T| + (1+o(1))p\,|\tilde{U}|\,|V_0| - e(\tilde{U}, T). \end{aligned} \tag{25}$$

By the definition of the set T,

$$e(\tilde{U}, T) \ge 2\,|T|,$$

and consequently,

$$|G(\tilde{U}) \cap V_0| \le (1+o(1))p\,|\tilde{U}|\,|V_0| - \frac{1}{2}e(\tilde{U}, T).$$

Since by (24) we have $|\tilde{U}| \le \delta/p$, Property $(\mathbf{II})(a)$ implies that the left-hand side above is at least $(1-2\delta)p\,|\tilde{U}|\,|V_0|$ and therefore

$$e(\tilde{U}, T) \le (4\delta + o(1))p\,|\tilde{U}|\,|V_0| < 4\delta pn\,|\tilde{U}|.$$

By the definition of the set U, every vertex $v \in \tilde{U} \subseteq U$ satisfies $|G(v) \cap W| \ge pn/3$ and therefore

$$e(\tilde{U}, W \setminus T) = e(\tilde{U}, W) - e(\tilde{U}, T) \ge \Big(\frac{pn}{3} - 4\delta pn\Big)|\tilde{U}|.$$

Given the definition of T, no vertex in $W \setminus T$ has more than one neighbor in $\tilde{U}$, hence the left-hand side of the inequality above is at most $|W \setminus T|$. Since $\delta < 1/48$, it follows that

$$|W| \ge |W \setminus T| \ge \Big(\frac{pn}{3} - 4\delta pn\Big)|\tilde{U}| > \frac{pn}{4}|\tilde{U}|, \tag{26}$$

and consequently

$$|\tilde{U}| < \frac{4}{pn}|W| \le \frac{\delta}{p}.$$

From the definition of $\tilde{U}$ (see (24)) we must have $\tilde{U} = U$ and thus also

$$|U| \le \frac{4}{pn}|W|,$$

as required. $\square$

4. The analysis of Algorithm 1

In this and the next section we show that Algorithm 1 with an input G satisfying the properties established in Lemma 3.1, with $\delta = 0.01$, is **a.a.s.** successful (see Lemma 4.1 below). Consequently, Lemmas 3.1 and 4.1 will together imply Theorem 1.1 (this formal derivation of Theorem 1.1 is given at the end of this section; also, see Figure 3 for the overall structure of the proof of Theorem 1.1). The probability space in Lemma 4.1 is the uniform space of all initial embedding f_0 and corresponds to Step 2 of the algorithm, the only randomized step therein.

Lemma 4.1. *Let ε and τ be as in* (4) *and* (6)*, respectively, and let*

$$\delta = 0.01.$$

Suppose that G is a graph with vertex set $V = [n]$ partitioned as $V = V_0 \cup R_1 \cup \cdots \cup R_{d^2+2}$ as in (5)*, and that $p \ge Cn^{-1/d}\log^{1/d} n$ for a sufficiently large constant C.*

If G and p satisfy Properties (**I**)*–*(**V**) *from Lemma 3.1, then Algorithm 1 with input G is **a.a.s.** successful, that is, for every $H \in \mathcal{H}(n, d)$ it **a.a.s.** outputs an embedding of H into G.*

In order to prove Lemma 4.1, observe that Algorithm 1 is successful if it does not terminate at lines 2, 6, or 9, namely if the following three statements are satisfied.

- (S2) any sequence of pairwise disjoint d-element sets $\kappa_1, \dots, \kappa_j \subset V_0$ with $j < t$ is such that $G\big[V_0 \setminus \bigcup_{1\le i\le j} \kappa_i\big]$ contains a d-clique (line 2);
- (S6) for each $i = 1, \dots, d^2 + 1$ there is a matching in A_i saturating X_i (line 6);
- (S9) there is a perfect matching in A_{d^2+2} (line 9).

We are now going to prove the three statements (S2), (S6) and (S9) one by one (Claims 4.2–4.6 below). The following diagram exhibits the proof flow of Theorem 1.1.

Claim 4.2. *Statement (S2) is true.*

Proof. First note that $|V_0| > 3n/4$ and that, by Property (**IV**)(c), any subset $U \subset V$ with $|U| \ge n/4$ contains a d-clique (in fact, it contains many cliques). Let $j < t$ and suppose j disjoint d-sets $\kappa_1, \dots, \kappa_j$ are given. Let $U = V_0 \setminus \bigcup_{i=1}^{j} \kappa_i$ and note that $|U| = |V_0| - jd > |V_0| - td > n/4$. This guarantees the existence of a d-clique in U. $\square$

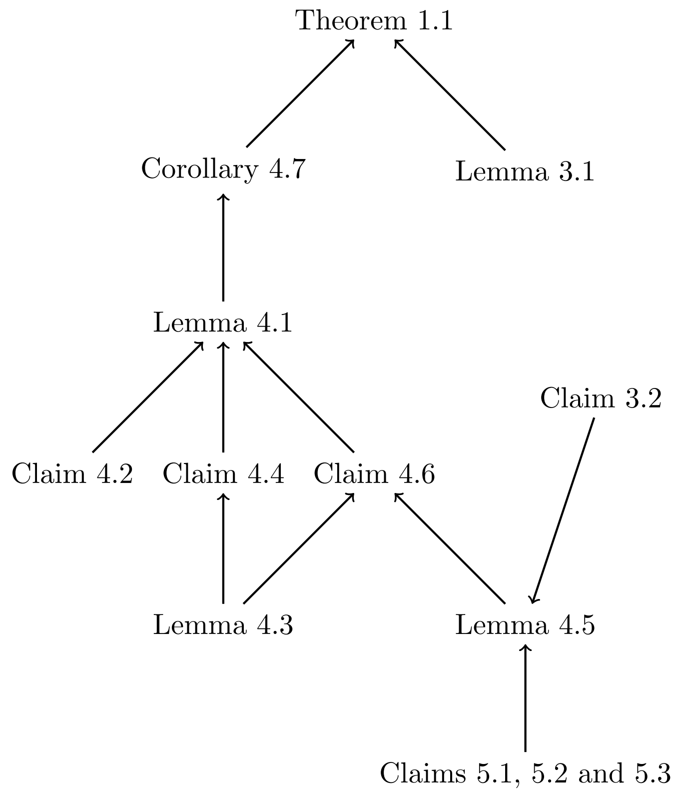

FIGURE 3. The structure of the proof of Theorem 1.1

Statement (S6) will follow from the next, deterministic lemma. We implicitly assume that a fixed graph G satisfies Properties (**I**)–(**V**) from Lemma 3.1, and that (4)–(6) hold.

Lemma 4.3. *For $i = 1, \dots, d^2 + 2$ and for every $Q \subset X_i$ we have*

$$|A_i(Q)| \geq \min\{|Q|, |W_i| - \omega p^{-d}\}. \tag{27}$$

In particular, if $|W_i| \geq |X_i| + \omega p^{-d}$ then

$$|A_i(Q)| \geq |Q|$$

for all sets $Q \subset X_i$.

Proof. Let $i \in \{1, \dots, d^2 + 1\}$ be fixed. We will now prove that (27) holds for any $Q \subset X_i$ regardless of the particular partial embedding f_{i-1} (in fact, we only need f_{i-1} to be a one-to-one map for this proof). For each $k = 0, 1, \dots, d$, let

$$Q_k = \big\{x \in Q \,:\, \big|f_{i-1}\big(H(x)\big)\big| = k\big\}.$$

Clearly $Q = Q_0 \cup \cdots \cup Q_d$ is a partition of Q.

Note that if $Q_0 \neq \emptyset$ then, by (11), $A_i(Q) \supset A_i(Q_0) = W_i$ and thus (27) holds. Hence, assume that $Q_0 = \emptyset$ and let $1 \le k \le d$ be such that $|Q_k| \ge |Q|/d$.

The proof is split into two cases according to whether Q_k is small ($|Q_k| \le \omega p^{-k}$) or large ($|Q_k| > \omega p^{-k}$). First consider the case when Q_k is small. Then,

$$q := \min\bigl\{\delta p^{-k}, |Q_k|\bigr\} \ge \frac{\delta\,|Q_k|}{\omega} \ge \frac{\delta\,|Q|}{\omega d}. \tag{28}$$

Further, notice that

$$\begin{aligned} |A_i(Q)| \;&\ge\; |A_i(Q) \cap R_i| \\ &\overset{(11)}{=} \bigl|\bigl\{w \in R_i \,:\, G(w) \supset f_{i-1}\bigl(H(x)\bigr) \text{ for some } x \in Q\bigr\}\bigr| \\ &\overset{(14)}{=} \alpha(\mathcal{X}, R_i), \end{aligned} \tag{29}$$

for $\mathcal{X} = \{f_{i-1}(H(x)) \,:\, x \in Q\}$. (The k-sets in the family $\mathcal{X}$ are pairwise disjoint because $Q \subset X_i$ is 2-independent in H; they are also disjoint from R_i since $R_i \cap \mathrm{im}(f_{i-1}) = \emptyset$.)

Applying Property (**III**)(a) with $U = R_i$ yields

$$\alpha(\mathcal{X}, R_i) \ge (1-2\delta)p^k\,|R_i|\,q \overset{(5)}{\ge} (1-3\delta)p^k(\varepsilon n)\,q.$$

In particular, for C large enough, we have

$$|A_i(Q)| \ge |A_i(Q)| \ge (1-3\delta)\varepsilon p^k n\,q \ge \frac{\varepsilon}{2}C^d \log n\,q \ge \delta^{-1}\omega d\,q \overset{(28)}{\ge} |Q|.$$

Consequently, (27) holds when Q_k is small.

Now we consider the case when Q_k is large, that is, $|Q_k| > \omega p^{-k}$. Here we will prove that $|A_i(Q)| \ge |W_i| - \omega p^{-d}$ and thus establish that (27) holds when Q_k is large. Suppose for the sake of a contradiction that $|A_i(Q)| < |W_i| - \omega p^{-d}$ or, equivalently, $|W_i \setminus A_i(Q)| > \omega p^{-d}$.

Set $U = W_i \setminus A_i(Q_k)$ and observe that $U \supset W_i \setminus A_i(Q)$, which by assumption means that $|U| > \omega p^{-d}$. Also note that $W_i \cap \mathrm{im}(f_{i-1}) = \emptyset$ and thus $U \subset W_i$ does not intersect any set in $\mathcal{X} = \{f_{i-1}(H(x)) \,:\, x \in Q_k\}$; in other words, $U \subset V \setminus \bigcup_{J \in \mathcal{X}} J$. Applying Property (**III**)(b) yields that $B(\mathcal{X}, U)$ is not empty, namely, there is $x \in Q_k$ and $v \in U$ such that $f_{i-1}(H(x)) \subset G(v)$. Hence, (x, v) is an edge in A_i between Q_k and U, contradicting the definition of $U = W_i \setminus A_i(Q_k)$. □

Now we are ready to prove statement (S6).

Claim 4.4. *Statement (S6) is true. That is, for each $i = 1, \ldots, d^2+1$, the graph A_i has a matching saturating X_i.*

Proof. Fix $1 \le i \le d^2+1$ and recall the definition of W_i in (10):

$$W_i = V \setminus \mathrm{im}(f_{i-1}) \setminus \bigcup_{j=i+1}^{d^2+2} R_j.$$

Note that because $i \le d^2+1$ and $n = |X_0| + \cdots + |X_{d^2+2}|$,

$$\begin{aligned} |W_i| &= n - \sum_{j<i} |X_j| - \sum_{j>i} |R_j| = |X_i| + \sum_{j>i} (|X_j| - |R_j|) \\ &\overset{(8)}{\ge} |X_i| + \sum_{j>i} (t - |R_j|) = |X_i| + (d^2+2-i)(t - \varepsilon n) \qquad (30) \\ &\ge |X_i| + t - \varepsilon n \overset{(6)}{=} |X_i| + \varepsilon n. \end{aligned}$$

For C sufficiently large, we have

$$\varepsilon n \ge C^{1-d} n = \omega p^{-d}.$$

Thus, $|W_i| \ge |X_i| + \omega p^{-d}$, which, by Lemma 4.3, implies that $|A_i(Q)| \ge |Q|$ for all $Q \subset X_i$. Consequently, by Hall's theorem, there is a matching in A_i covering X_i. □

For the proof of Statement (S9), besides Lemma 4.3, we will also need the following probabilistic result.

Lemma 4.5. *The random embedding f_0 of the sets S_i, $i = 1, \ldots, t$, is such that* ***a.a.s.****, for every set $Y \subset V$ with $|Y| \le \delta(4p)^{-d}$, where $\delta = 0.01$, we have*

$$\Big| \big\{ x \in X_{d^2+2} \,:\, f_0\big(H(x)\big) \subset G(v) \text{ for some } v \in Y \big\} \Big| \ge \frac{1}{2} \Big(\frac{p}{5}\Big)^d t\, |Y|. \qquad (31)$$

Since the proof of Lemma 4.5 is quite long, we defer it to Section 5. Meanwhile, we prove the last of our three statements and thus complete the proof of Lemma 4.1.

Claim 4.6. *Statement (S9) is true. That is,* ***a.a.s.*** *the random map f_0 is such that the graph A_{d^2+2} contains a perfect matching.*

Proof. Set $h = d^2 + 2$ for convenience. To prove that A_h has a perfect matching **a.a.s.**, recall that, as a consequence of (13), for every $Y \subset W_h$,

$$A_h(Y) = \big\{ x \in X_{d^2+2} \,:\, f_0\big(H(x)\big) \subset G(v) \text{ for some } v \in Y \big\}.$$

Therefore, by Lemma 4.5, **a.a.s.**, for every $Y \subset W_h$ with $|Y| \le \delta(4p)^{-d}$, we have (see (31)),

$$|A_h(Y)| \ge \frac{1}{2}\Big(\frac{p}{5}\Big)^d t\, |Y| \ge \delta^{-1} 4^d \omega\, |Y|, \qquad (32)$$

provided C is large enough. We claim that the condition above ensures that there is a perfect matching in A_h. Recall that $|X_h| = |W_h| = t$. Let $Q \subset X_h$. If $|Q| \le t - \omega p^{-d}$ then Lemma 4.3 implies that $|A_h(Q)| \ge |Q|$. Assume then that

$$|Q| \ge t - \omega p^{-d} + 1 \qquad (33)$$

(for simplicity, we assume that ωp^{-d} is an integer), and suppose, for the sake of contradiction, that $|A_h(Q)| \le |Q| - 1$, or, equivalently, that

$$|W_h \setminus A_h(Q)| \ge t - |Q| + 1. \qquad (34)$$

Since $A_h(W_h \setminus A_h(Q)) \subset X_h \setminus Q$, it follows that $|A_h(W_h \setminus A_h(Q))| \le t - |Q|$. Next we will contradict this inequality and therefore prove that $|A_h(Q)| \ge |Q|$.

To obtain the desired contradiction we invoke inequality (32) for a set $Y \subset W_h \setminus A_h(Q)$ satisfying $|Y| = \min\{|W_h \setminus A_h(Q)|, \delta(4p)^{-d}\}$. We now argue that

$$\begin{aligned} |A_h(Y)| &\overset{(32)}{\ge} \delta^{-1}4^d\omega\,|Y| \\ &= \delta^{-1}4^d\omega \times \min\{|W_h \setminus A_h(Q)|, \delta(4p)^{-d}\} \\ &\ge \min\{|W_h \setminus A_h(Q)|, \omega p^{-d}\} \\ &\ge t - |Q| + 1. \end{aligned} \tag{35}$$

The third inequality follows from (33) and (34). Clearly, (35) establishes the desired contradiction and thus proves the claim. □

Having proved Lemma 4.1 (except for the proof of Lemma 4.5, deferred to the next section), we conclude this section with the proof of Theorem 1.1. It will be convenient to state first a corollary of Lemma 4.1.

Corollary 4.7. *Let G be a graph as in Lemma 4.1. Then G is $\mathcal{H}(n,d)$-universal.*

Proof. By Lemma 4.1, for every $H \in \mathcal{H}(n,d)$, Algorithm 1 with input G, outputs an embedding of H into G with positive probability, and thus such an embedding exists. □

We finally give the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $\delta = 1/100$ and let $C = C(\delta)$ be large enough, as required by Lemmas 3.1 and 4.1. Let $|V| = n$ and let $V = V_0 \cup R_1 \cup \cdots \cup R_{d^2+2}$ be a partition as in (5) and $p \ge Cn^{-1/d}\log^{1/d} n$. By Lemma 3.1, a random graph $G \in G_{n,p}$, where $V(G) = V$, **a.a.s.** satisfies Properties (**I**)–(**V**). On the other hand, by Corollary 4.7 every such graph is $\mathcal{H}(n,d)$-universal. □

5. Proof of Lemma 4.5

Our goal is to prove that **a.a.s.** the random embedding f_0 satisfies (31) for all $Y \subset V$ with $|Y| \le \delta(4p)^{-d}$. Recall that the images $f_0(S_i)$ are created by randomly selecting from V_0 pairwise disjoint d-sets $\kappa_1, \dots, \kappa_t$, each inducing a clique in G, and then f_0 is defined in any way so that $f_0(S_i) = \kappa_i$ for all i. Let Ω be the space of all such sequences $\boldsymbol{\kappa} = \big(\kappa_1, \dots, \kappa_t\big)$. A sequence $\boldsymbol{\kappa}$ is sampled from Ω by first selecting a d-set κ_1 uniformly from $\binom{V_0}{K_d}$, and then selecting each subsequent κ_i, $i = 2, \dots, t$, uniformly from

$$\binom{V_0 \setminus \bigcup_{j=1}^{i-1} \kappa_j}{K_d}.$$

Fix an integer

$$y \le \delta(4p)^{-d} = o(n), \text{ where, we recall, } \delta = 0.01. \tag{36}$$

Notice that, by Property $(\mathbf{IV})(c)$, for every $i = 1, \dots, t$, we have

$$(1-\delta)p^{\binom{d}{2}}\binom{|V_0| - td}{d} \leq \left|\binom{V_0 \setminus \bigcup_{j=1}^{i-1} \kappa_j}{K_d}\right| \leq (1+\delta)p^{\binom{d}{2}}\binom{|V_0|}{d},$$

From now on we will focus on a fixed set

$$Y \subset V \text{ with } |Y| = y, \tag{37}$$

and define a random variable corresponding to the left-hand side of (31):

$$\begin{aligned} \mathbb{A} = \mathbb{A}_Y &:= \left|\left\{x_i \in X_{d^2+2} \;:\; f_0(H(x_i)) \subset G(v) \text{ for some } v \in Y\right\}\right| \\ &= \left|\left\{i \in [t] \;:\; \kappa_i \subset G(v) \text{ for some } v \in Y\right\}\right|. \end{aligned} \tag{38}$$

We will ultimately show that in the random model described above, the inequality

$$\mathbb{A} \geq \frac{1}{2}\left(\frac{p}{5}\right)^d ty \tag{39}$$

fails with such a small probability that the union bound can be applied over all possible choices for Y still yielding a $o(1)$ failure probability. Consequently, **a.a.s.** (31) will hold for all choices of Y and thus Lemma 4.5 will follow.

In view of (39), we are interested in estimating how many d-sets κ_i are contained in at least one of the neighborhoods $G(v)$ for $v \in Y$. To this end, for each $i = 0, \dots, t-1$, given disjoint d-cliques $\kappa_1, \dots, \kappa_i$, define

$$\mathcal{A}(\kappa_1, \dots, \kappa_i) = \bigcup_{v \in Y} \binom{(G(v) \cap V_0) \setminus \bigcup_{j=1}^{i} \kappa_j}{K_d}. \tag{40}$$

Let

$$\mathbb{A}_i = \mathbf{1}[\kappa_i \in \mathcal{A}(\kappa_1, \dots, \kappa_{i-1})]. \tag{41}$$

Note that

$$\mathbb{A} = \sum_{i=1}^{t} \mathbb{A}_i. \tag{42}$$

Let

$$Z = V_0 \cap \bigcup_{v \in Y} G(v) \tag{43}$$

and let $z = |Z|$. Set also

$$q_1 = q_1(y) := y\left(\frac{p}{5}\right)^d \overset{(36)}{\leq} \delta 20^{-d}.$$

Claim 5.1.

$$q_1 n \leq z \leq pny.$$

Proof. By Property $(\mathbf{I})(a)$, for every $v \in Y$

$$|G(v) \cap V_0| \leq (1 + o(1))p\,|V_0| < pn,$$

and thus

$$z = \left|V_0 \cap \bigcup_{v \in Y} G(v)\right| < pny.$$

For the lower bound on z, first consider the case when $y = |Y| \le \omega p^{-1}$. Then we have $\min\{y, \delta/p\} \ge \delta y/\omega$ and, by Property (**II**)(a),

$$z \ \ge\ |G(Y) \cap V_0| \ge (1-2\delta)p\,|V_0|\,\min\{y, \delta/p\} \ge \frac{\delta pny}{2\omega} > q_1 n.$$

Now suppose that $y = |Y| \ge \omega p^{-1}$ and let $U = V_0 \setminus (G(Y) \cup Y)$. As $B(Y, U) = \emptyset$, by Property (**II**)(b), we must have $|U| < \omega p^{-1} = o(n)$. Since $|U| \ge |V_0| - |Z| - |Y|$, by (36),

$$z = |Z| \ge |V_0| - o(n) > n/2 > q_1 n,$$

as required. $\square$

In order to estimate the rate at which the families $\mathcal{A}(\kappa_1, \ldots, \kappa_i)$ shrink, we introduce another random variable $\mathbb{B}$ which helps to keep track of how many vertices of Z are "consumed" by the sequence $\boldsymbol{\kappa}$.

Let

$$\mathbb{B}_i = \mathbf{1}[\kappa_i \cap Z \neq \emptyset] \tag{44}$$

and

$$\mathbb{B} = \sum_{i=1}^{t} \mathbb{B}_i.$$

Claim 5.2. $\mathbf{P}[\mathbb{B} \ge 3dzt/n] \le \exp\{-c_2 dzt/n\} \le \exp\{-c_3 t q_1\}$.

Proof. Observe that, by Property (**V**), the number of d-cliques in $G[V_0]$ containing a given vertex $v \in Z$ can be bounded above by

$$(1+\delta)p^{\binom{d}{2}} \frac{d}{|V_0|} \binom{|V_0|}{d}.$$

Moreover, by our choice of t in (6), using the Bernoulli inequality (which states that $(1+x)^a \ge 1 + ax$ for all $a \in \mathbb{N}$ and $x \ge -1$), we may ensure that

$$\left(1 - \frac{td}{|V_0|}\right)^d \ge 1 - \frac{td^2}{|V_0|} \ge 1 - \frac{2}{75d^2} \ge 0.99.$$

Thus, it follows that, for any i,

$$\begin{aligned}
\mathbf{P}[\mathbb{B}_i = 1 \mid \kappa_1, \ldots, \kappa_{i-1}] &\le z(1+\delta)p^{\binom{d}{2}} \frac{d}{|V_0|} \binom{|V_0|}{d} \left| \binom{V_0 \setminus \bigcup_{j=1}^{i-1} \kappa_j}{K_d} \right|^{-1} \\
&\overset{(\mathbf{IV})(c)}{\le} \frac{1+\delta}{1-\delta} \frac{zd}{|V_0|} \frac{(|V_0|)_d}{(|V_0| - (t-1)d)_d} \le \frac{(1+3\delta)zd\,|V_0|^d}{|V_0|(|V_0| - td)^d} \\
&= \quad (1+3\delta)\frac{zd}{|V_0|}\left(1 - \frac{td}{|V_0|}\right)^{-d} \le (1+3\delta)\frac{4zd}{3n}\frac{1}{0.99} < \frac{2zd}{n} := q_2.
\end{aligned} \tag{45}$$

We now apply Proposition A.1 from the appendix, setting the $\mathbb{X}_i$ and the $\mathbb{K}_i$ in that proposition to be the $\mathbb{B}_i$ and the κ_i, respectively, and letting $\gamma = 1/2$. We have just shown in (45) that the hypothesis of (b) in Proposition A.1 holds with $q = q_2$ and $\Pi = 0$. Inequality (61) and Claim 5.1 imply that

$$\mathbf{P}[\mathbb{B} \ge 3dzt/n] \le \exp\{-c_2 dzt/n\} \le \exp\{-c_3 t q_1\},$$

for some constants c_2 and $c_3 > 0$. □

Recall that we have fixed a set $Y \subset V$ with $|Y| = y \le \delta(4p)^{-d}$, and defined $Z = V_0 \cap \bigcup_{v\in Y} G(v)$ and $z = |Z|$ (see (43)). Our last claim asserts that if $\mathbb{B}$ is small, then the families $\mathcal{A}(\kappa_1, \dots, \kappa_i)$ remain large throughout the entire process of selecting t random disjoint cliques. Recall that $t = \lfloor \tau n \rfloor$ (see (6)).

Claim 5.3. *For a sequence* $(\kappa_1, \dots, \kappa_t)$ *satisfying* $\mathbb{B} = \mathbb{B}(\kappa_1, \dots, \kappa_t) \le 3dz\tau$, *we have*

$$|\mathcal{A}(\kappa_1, \dots, \kappa_t)| \ge y p^{\binom{d}{2}} \binom{pn/4}{d}. \tag{46}$$

Proof. Let

$$W = Z \cap \bigcup_{1\le i\le t} \kappa_i \tag{47}$$

be the set of all vertices of Z "hit" by some clique κ_i, and let

$$Y' = \{v \in Y \ : \ |G(v) \cap W| \ge pn/3\}.$$

Observe that $|W| \le \mathbb{B}d$. By Claim 3.2 with $U := Y'$, we thus have

$$|Y'| \le \frac{4}{pn}|W| \le \frac{12d^2\tau}{pn} z \le 12d^2\tau y. \tag{48}$$

For every $v \in Y$, we have $G(v) \cap V_0 \subset Z$ (recall (43)). Recalling (47), we see that that, for every $v \in Y$, we have

$$(G(v) \cap V_0) \setminus \bigcup_{1\le i\le t} \kappa_i = (G(v) \cap V_0) \setminus W. \tag{49}$$

Therefore, the definition of $\mathcal{A}(\kappa_1, \dots, \kappa_t)$ (see (40)) and Bonferroni's inequality give that

$$\begin{aligned}
|\mathcal{A}(\kappa_1, \dots, \kappa_t)| &= \left| \bigcup_{v\in Y} \binom{(G(v) \cap V_0) \setminus W}{K_d} \right| \\
&\ge \sum_{v\in Y} \left| \binom{(G(v) \cap V_0) \setminus W}{K_d} \right| - \sum_{v\ne v'\in Y} \left| \binom{(G(v) \cap G(v') \cap V_0) \setminus W}{K_d} \right| \\
&\ge \sum_{v\in Y\setminus Y'} \left| \binom{(G(v) \cap V_0) \setminus W}{K_d} \right| - \sum_{v\ne v'\in Y} \left| \binom{G(v) \cap G(v') \cap V_0}{K_d} \right|.
\end{aligned} \tag{50}$$

Recall that $|V_0| \ge 3n/4$. For $v \in Y \setminus Y'$, Property $(\mathbf{I})(a)$ yields that

$$\begin{aligned}
|(G(v) \cap V_0) \setminus W| &= |G(v) \cap V_0| - |G(v) \cap W| \\
&\ge (1 + o(1))p\,|V_0| - pn/3 > pn/3.
\end{aligned}$$

Hence, the first sum of the last line in (50) may be bounded as follows:

$$\sum_{v\in Y\setminus Y'} \left| \binom{(G(v) \cap V_0) \setminus W}{K_d} \right| \overset{(\mathbf{IV})(a)}{\ge} |Y \setminus Y'|\,(1-\delta)p^{\binom{d}{2}} \binom{pn/3}{d}.$$

Moreover, by (48) and the definition of τ in (6),

$$|Y \setminus Y'| \geq (1 - 12d^2\tau)y \geq \frac{1}{2}y,$$

and thus

$$\sum_{v \in Y \setminus Y'} \left| \binom{(G(v) \cap V_0) \setminus W}{K_d} \right| \geq (1-\delta)\frac{y}{2}p^{\binom{d}{2}}\binom{pn/3}{d}. \tag{51}$$

On the other hand, for $v \neq v' \in Y$, Property $(\mathbf{I})(c)$ tells us that

$$|G(v) \cap G(v') \cap V_0| \leq |G(v) \cap G(v')| = (1 + o(1))p^2 n.$$

Hence, the second sum of the last line in (50) may be bounded, for every large enough n, as follows:

$$\begin{aligned} \sum_{v \neq v' \in Y} \left| \binom{G(v) \cap G(v') \cap V_0}{K_d} \right| &\leq \sum_{v \neq v' \in Y} \left| \binom{G(v) \cap G(v')}{K_d} \right| \\ &\overset{(\mathbf{IV})(b)}{\leq} \binom{y}{2}(1+\delta)p^{\binom{d}{2}}\binom{(1+\delta)p^2 n}{d}. \end{aligned} \tag{52}$$

Consequently, by (50), (51), and (52) we obtain

$$\begin{aligned} |\mathcal{A}(\kappa_1, \ldots, \kappa_t)| &\geq (1-\delta)\frac{y}{2}p^{\binom{d}{2}}\binom{pn/3}{d} - (1+\delta)\binom{y}{2}p^{\binom{d}{2}}\binom{(1+\delta)p^2 n}{d} \\ &\geq \frac{yp^{\binom{d}{2}}}{2d!}\Big\{(1-\delta)(pn/3)_d - (1+\delta)(yp^d)\big((1+\delta)pn\big)^d\Big\}. \end{aligned} \tag{53}$$

From (36) we conclude that $(yp^d)(pn)^d \leq \delta(pn/4)^d$. Using that $d \geq 3$ and that $\delta = 0.01$, we see after a simple calculation that

$$|\mathcal{A}(\kappa_1, \ldots, \kappa_t)| \geq yp^{\binom{d}{2}}\binom{pn/4}{d},$$

which establishes the claim. $\square$

Claims 5.2 and 5.3, and the fact that

$$\mathcal{A}(\emptyset) \supset \mathcal{A}(\kappa_1) \supset \mathcal{A}(\kappa_1, \kappa_2) \supset \cdots \supset \mathcal{A}(\kappa_1, \ldots, \kappa_t),$$

imply that, with probability at least $1 - \exp\{-c_3 t q_1\}$, for all $i = 1, \ldots, t$, the subsequence $(\kappa_1, \ldots, \kappa_{i-1})$ satisfies

$$|\mathcal{A}(\kappa_1, \ldots, \kappa_{i-1})| \geq |\mathcal{A}(\kappa_1, \ldots, \kappa_t)| \geq yp^{\binom{d}{2}}\binom{pn/4}{d}.$$

Hence, with probability at least $1 - \exp\{-c_3 t q_1\}$, for all $i = 1, \ldots, t$,

$$\mathbf{P}[\mathbb{A}_i = 1 \mid \kappa_1, \ldots, \kappa_{i-1}] = \frac{|\mathcal{A}(\kappa_1, \ldots, \kappa_{i-1})|}{\left|\binom{V_0 \setminus \bigcup_{j=1}^{i-1}\kappa_j}{K_d}\right|} \overset{(\mathbf{IV})(c)}{\geq} \frac{y\binom{pn/4}{d}}{(1+\delta)\binom{n}{d}} > q_1.$$

We now apply Proposition A.1, setting the $\mathbb{X}_i$ and the $\mathbb{K}_i$ in that proposition to be the $\mathbb{A}_i$ and the κ_i, respectively, and letting $\gamma = 1/2$. We have just shown that the hypothesis of (a) in Proposition A.1 holds with $q = q_1$ and $\Pi = \exp\{-c_3 t q_1\}$. Inequality (59) then tells us that

$$\mathbf{P}[\mathbb{A} \le tq_1/2] \le \exp\{-c_1 t q_1\}, \tag{54}$$

for some constant $c_1 > 0$. Note that

$$\frac{tq_1}{2} = \frac{1}{2}\Big(\frac{p}{5}\Big)^d ty.$$

In other words, with probability at least $1 - \exp\{-c_1 t q_1\}$ the random embedding f_0 satisfies (31) for a fixed set Y. We will now finish the proof of Lemma 4.5 by using the union bound.

The probability that there is some $Y \subset V$ with $|Y| \le \delta(4p)^{-d}$ that fails to satisfy (31) is, in view of (6) and (54), at most

$$\begin{aligned}\sum_{y=1}^{\delta(4p)^{-d}} \binom{n}{y} \exp\big\{-c_1 t q_1\big\} &\le \sum_y \exp\big\{y \log n - c_1 \tau n (p/5)^d y\big\} \\ &\le \sum_y \exp\big\{y \log n \big(1 - (c_1 \tau 5^{-d}) \cdot C^d\big)\big\} \\ &\le \sum_y n^{-y} = o(1),\end{aligned} \tag{55}$$

for C large enough. Hence, the probability that (31) fails for some Y is at most $o(1)$. This completes the proof of Lemma 4.5.

Acknowledgements: We are very thankful to Matas Šileikis for his suggestions leading to a simplification of the proof of Proposition A.1, as well as to anonymous referees for their numerous comments the implementation of which has improved the readability of the paper.

Appendix A.

Here we prove a concentration result used in the proofs of Lemma 4.5 and Claim 5.2. (For related results, see McDiarmid [21].) Let $\Omega = \mathcal{K}_1 \times \cdots \times \mathcal{K}_t$, where each $\mathcal{K}_i$ is a finite set, and suppose that $\mathbf{P} = \mathbf{P}_\Omega$ is a probability distribution defined on Ω. Let us write $(\mathbb{K}_1, \dots, \mathbb{K}_t)$ for an element of Ω drawn according to $\mathbf{P}$.

For each $1 \le i \le t$, let $f_i \colon \mathcal{K}_1 \times \cdots \times \mathcal{K}_i \to \{0, 1\}$ be given. We are interested in the concentration of the sum $\mathbb{X} = \sum_{1 \le i \le t} \mathbb{X}_i$ of the Bernoulli r.vs $\mathbb{X}_i$ given by

$$\mathbb{X}_i(\kappa_1, \dots, \kappa_t) = f_i(\kappa_1, \dots, \kappa_i) \tag{56}$$

for all $\kappa_j \in \mathcal{K}_j$ $(1 \le j \le t)$ and $1 \le i \le t$. We shall work under hypotheses controlling the conditional expectation of $\mathbb{X}_i$ with respect to the $\mathbb{K}_j$ $(1 \le j < i)$, that is, controlling $\mathbf{E}[\mathbb{X}_i \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}] = \mathbf{P}[\mathbb{X}_i = 1 \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}]$, on ‘most’ of Ω.

Proposition A.1. *Let* Ω, $\mathbf{P}$, $\mathbb{X}_1, \dots, \mathbb{X}_t$ *and* $\mathbb{X} = \sum_{1\le i\le t} \mathbb{X}_i$ *be as above. For every* $1 \le i \le t$, *let* $\mathbb{P}_i$ *be the random variable*

$$\mathbb{P}_i = \mathbf{P}[\mathbb{X}_i = 1 \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}]. \tag{57}$$

Then, for any $\gamma > 0$, *there exists a constant* $c = c(\gamma) > 0$ *for which the following hold.*

(a) *If*

$$\mathbf{P}[\mathbb{P}_i \ge q \text{ for all } i = 1, \dots, t] \ge 1 - \Pi, \tag{58}$$

then

$$\mathbf{P}[\mathbb{X} \le (1-\gamma)tq] \le \exp\{-ctq\} + \Pi. \tag{59}$$

(b) *If*

$$\mathbf{P}[\mathbb{P}_i \le q \text{ for all } i = 1, \dots, t] \ge 1 - \Pi, \tag{60}$$

then

$$\mathbf{P}[\mathbb{X} \ge (1+\gamma)tq] \le \exp\{-ctq\} + \Pi. \tag{61}$$

Proof. We first prove (a). We give a coupling type argument. Consider the uniform distribution on $\Omega' = [0,1]^t$, and write $(\mathbb{U}_i)_{1\le i\le t}$ for a random element of Ω'. Thus, the $\mathbb{U}_i$ ($1 \le i \le t$) form a sequence of independent uniform r.vs, each taking values on the unit interval $[0,1]$. Let us consider the product probability space $\widetilde{\Omega} = \Omega \times \Omega'$, with probability measure $\mathbf{P}_{\widetilde{\Omega}} = \mathbf{P}_\Omega \times \mathbf{P}_{\Omega'}$. We shall define a sequence of r.vs $\mathbb{Z}_i$ on $\widetilde{\Omega}$ ($1 \le i \le t$) in such a way that

(i) the $\mathbb{Z}_i$ ($1 \le i \le t$) are independent Bernoulli r.vs with mean q each.

We shall also define a certain 'bad' event $B \subset \Omega$ in such a way that, setting $\widetilde{B} = B \times \Omega' \subset \widetilde{\Omega}$, we have

(ii) $\mathbf{P}_{\widetilde{\Omega}}[\widetilde{B}] \le \Pi$ and, outside $\widetilde{B}$, we have $\mathbb{X}_i \ge \mathbb{Z}_i$ for all $1 \le i \le t$.

With the $\mathbb{Z}_i$ and $\widetilde{B}$ at hand, we may derive part (a) of our proposition as follows. Let $\mathbb{Z} = \sum_{1\le i\le t} \mathbb{Z}_i$ and note that, on $\widetilde{\Omega} \setminus \widetilde{B}$, we have $\mathbb{Z} \le \mathbb{X}$. Now observe that, for any $\gamma > 0$,

$$\begin{aligned}\mathbf{P}_\Omega[\mathbb{X} \le (1-\gamma)tq] &= \mathbf{P}_{\widetilde{\Omega}}[\mathbb{X} \le (1-\gamma)tq]\\ &\le \mathbf{P}_{\widetilde{\Omega}}[\mathbb{X} \le (1-\gamma)tq \text{ and } \widetilde{B} \text{ fails}] + \mathbf{P}_{\widetilde{\Omega}}[\widetilde{B}]\\ &\le \mathbf{P}_{\widetilde{\Omega}}[\mathbb{Z} \le (1-\gamma)tq] + \Pi,\end{aligned}$$

which, by Chernoff's inequality applied to the binomial random variable $\mathbb{Z}$ (see, e.g., [17], Theorem 2.1, page 26), implies (59).

It remains to construct the $\mathbb{Z}_i$ and B. We proceed as follows. Recall that each $\mathbb{K}_j$ takes values in some finite set $\mathcal{K}_j$. Let $S = \bigcup_{0\le i\le t} \prod_{1\le j\le i} \mathcal{K}_j$. Thus, the i-tuple $(\mathbb{K}_1, \dots, \mathbb{K}_i)$ takes values in S, for all $1 \le i \le t$. One may think of S as the node set of a rooted tree, with each $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_i) \in S$ ($1 \le i \le t$) having as its parent the node $(\kappa_1, \dots, \kappa_{i-1})$. The root of the tree is the empty sequence, which we denote by λ. The points of Ω appear as leaves in this tree. For each $\boldsymbol{\kappa} = (\kappa_j)_{1\le j<i} \in S$ $(1 \le i \le t)$, let

$$\begin{aligned} p(\boldsymbol{\kappa}) &= \mathbf{P}_\Omega[\mathbb{X}_i = 1 \mid \mathbb{K}_j = \kappa_j \text{ for all } 1 \le j < i] \\ &= \mathbf{P}_\Omega[f_i(\mathbb{K}_1, \dots, \mathbb{K}_i) = 1 \mid \mathbb{K}_j = \kappa_j \text{ for all } 1 \le j < i]. \end{aligned} \tag{62}$$

Note that, in particular, $p(\lambda) = \mathbf{P}_\Omega[\mathbb{X}_1 = 1] = \mathbf{P}_\Omega[f_1(\mathbb{K}_1) = 1]$.

We first define the event $\widetilde{B} \subset \widetilde{\Omega}$. Given $\boldsymbol{\kappa} = (\kappa_i)_{1\le i\le t} \in \Omega \subset S$, we say that $\boldsymbol{\kappa}$ is *bad* if, for some $1 \le i \le t$, we have $p(\kappa_1, \dots, \kappa_{i-1}) < q$. Let $B = \{\boldsymbol{\kappa} \colon \boldsymbol{\kappa} \text{ is bad}\}$ and let

$$\widetilde{B} = B \times \Omega'. \tag{63}$$

By the definition of B, we have

$$\widetilde{B} = \{\mathbf{P}_{\widetilde{\Omega}}[\mathbb{X}_i = 1 \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}] < q \text{ for some } i = 1, \dots, t\}. \tag{64}$$

We now define the $\mathbb{Z}_i$ $(1 \le i \le t)$. For every $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_{i-1}) \in S$ with $1 \le i \le t$, let

$$\mathbb{B}_{\boldsymbol{\kappa}} = \begin{cases} \mathbb{1}\{\mathbb{U}_i \le q/p(\boldsymbol{\kappa})\} & \text{if } q \le p(\boldsymbol{\kappa}) \\ \mathbb{1}\{\mathbb{U}_i \le q\} & \text{otherwise.} \end{cases} \tag{65}$$

Conditional on $(\mathbb{K}_1, \dots, \mathbb{K}_{i-1}) = \boldsymbol{\kappa}$, we let the value of $\mathbb{Z}_i$ be given by

$$\mathbb{Z}_i = \begin{cases} \mathbb{X}_i \mathbb{B}_{\boldsymbol{\kappa}} & \text{if } q \le p(\boldsymbol{\kappa}) \\ \mathbb{B}_{\boldsymbol{\kappa}} & \text{otherwise.} \end{cases} \tag{66}$$

We now check conditions (i) and (ii) that are required of the $\mathbb{Z}_i$ and $\widetilde{B}$, as specified in the beginning of the proof. We first prove (i). We have to show that

$$\mathbf{E}_{\widetilde{\Omega}}[\mathbb{Z}_i \mid \mathbb{Z}_1, \dots, \mathbb{Z}_{i-1}] = q \tag{67}$$

for all $1 \le i \le t$. Let us show that

$$\mathbf{E}_{\widetilde{\Omega}}[\mathbb{Z}_i \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}, \mathbb{U}_1, \dots, \mathbb{U}_{i-1}] = q \tag{68}$$

for all $1 \le i \le t$. Fix $1 \le i \le t$, $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_{i-1}) \in S$ and $\mathbf{u} = (u_1, \dots, u_{i-1}) \in [0,1]^{i-1}$. Let us condition on $(\mathbb{K}_1, \dots, \mathbb{K}_{i-1}) = \boldsymbol{\kappa}$ and $(\mathbb{U}_1, \dots, \mathbb{U}_{i-1}) = \mathbf{u}$. Suppose first that $q > p(\boldsymbol{\kappa})$. Then $\mathbb{Z}_i = \mathbb{B}_{\boldsymbol{\kappa}} = \mathbb{1}\{\mathbb{U}_i \le q\}$ (see (65) and (66)), and hence $\mathbb{Z}_i$ is a Bernoulli r.v. with mean q, independent of the $\mathbb{K}_j$ $(1 \le j < i)$ and of the $\mathbb{U}_j$ $(1 \le j < i)$, and hence (68) follows. Suppose now that $q \le p(\boldsymbol{\kappa})$. Then, by the independence of $\mathbb{K}_i$ and $\mathbb{U}_i$, we have

$$\begin{aligned} &\mathbf{E}_{\widetilde{\Omega}}[\mathbb{Z}_i \mid \mathbb{K}_j = \kappa_j \text{ and } \mathbb{U}_j = u_j\ (j < i)] \\ &\qquad = \mathbf{E}_{\widetilde{\Omega}}[\mathbb{X}_i \mathbb{B}_{\boldsymbol{\kappa}} \mid \mathbb{K}_j = \kappa_j \text{ and } \mathbb{U}_j = u_j\ (j < i)] \\ &\qquad = \mathbf{E}_{\widetilde{\Omega}}[\mathbb{X}_i \mid \mathbb{K}_j = \kappa_j \text{ and } \mathbb{U}_j = u_j\ (j < i)] \\ &\qquad\qquad \times \mathbf{E}_{\widetilde{\Omega}}[\mathbb{B}_{\boldsymbol{\kappa}} \mid \mathbb{K}_j = \kappa_j \text{ and } \mathbb{U}_j = u_j\ (j < i)] \\ &\qquad = p(\boldsymbol{\kappa})(q/p(\boldsymbol{\kappa})) = q. \end{aligned} \tag{69}$$

We now derive (67) from (68). Recall that the r.vs $\mathbb{K}_1, \dots, \mathbb{K}_{i-1}$ determine $\mathbb{X}_1, \dots, \mathbb{X}_{i-1}$. Therefore, $\mathbb{K}_1, \dots, \mathbb{K}_{i-1}$, together with $\mathbb{U}_1, \dots, \mathbb{U}_{i-1}$, determine $\mathbb{Z}_1, \dots, \mathbb{Z}_{i-1}$. It follows that

$$\begin{aligned}\mathbf{E}[\mathbb{Z}_i \mid \mathbb{Z}_1, \dots, \mathbb{Z}_{i-1}] &= \mathbf{E}\big[\mathbf{E}[\mathbb{Z}_i \mid \mathbb{K}_1, \dots, \mathbb{K}_{i-1}, \mathbb{U}_1, \dots, \mathbb{U}_{i-1}] \mid \mathbb{Z}_1, \dots, \mathbb{Z}_{i-1}\big] \\ &= \mathbf{E}[q \mid \mathbb{Z}_1, \dots, \mathbb{Z}_{i-1}] \\ &= q.\end{aligned}$$

Therefore, requirement (i) does follow. Let us now check (ii). Fix $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_t) \in \Omega \setminus B$. Note that, for every $1 \le i \le t$, by (63) and (66), we have

$$\mathbb{Z}_i(\boldsymbol{\kappa}) = \mathbb{X}_i(\boldsymbol{\kappa})\mathbb{B}_{(\kappa_1, \dots, \kappa_{i-1})} \le \mathbb{X}_i(\boldsymbol{\kappa}),$$

and hence, $\mathbb{Z}_i \le \mathbb{X}_i$ holds outside $\widetilde{B}$. Finally, by (58) and (64), we have $\mathbf{P}_{\widetilde{\Omega}}[\widetilde{B}] \le \Pi$, as required. This concludes the proof of (a) of our proposition.

We now sketch the proof of (b). We proceed similarly as above, except that we now define the r.vs $\mathbb{B}_{\boldsymbol{\kappa}}$ and $\mathbb{Z}_i$ as follows. For every $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_{i-1}) \in S$ with $1 \le i \le t$, let

$$\mathbb{B}_{\boldsymbol{\kappa}} = \begin{cases} \mathbb{1}\{\mathbb{U}_i \le (1-q)/(1-p(\boldsymbol{\kappa}))\} & \text{if } q \ge p(\boldsymbol{\kappa}) \\ \mathbb{1}\{\mathbb{U}_i \le 1-q\} & \text{otherwise.} \end{cases} \tag{70}$$

Conditional on $(\mathbb{K}_1, \dots, \mathbb{K}_{i-1}) = \boldsymbol{\kappa}$, we let the value of $\mathbb{Z}_i$ be given by

$$1 - \mathbb{Z}_i = \begin{cases} (1-\mathbb{X}_i)\mathbb{B}_{\boldsymbol{\kappa}} & \text{if } q \ge p(\boldsymbol{\kappa}) \\ \mathbb{B}_{\boldsymbol{\kappa}} & \text{otherwise.} \end{cases} \tag{71}$$

One may then check that, with an appropriately defined $\widetilde{B}$, we have

(i) the $\mathbb{Z}_i$ $(1 \le i \le t)$ are independent Bernoulli r.vs with mean q each.
(ii) $\mathbf{P}_{\widetilde{\Omega}}[\widetilde{B}] \le \Pi$ and, outside $\widetilde{B}$, we have $\mathbb{X}_i \le \mathbb{Z}_i$ for all $1 \le i \le t$.

The proof of (b) follows (we omit the details). □

References

1. Noga Alon and Michael Capalbo, *Sparse universal graphs for bounded-degree graphs*, Random Structures Algorithms **31** (2007), no. 2, 123–133. MR MR2343715 (2008e:05104)
2. ______, *Optimal universal graphs with deterministic embedding*, Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (New York), ACM, 2008, pp. 373–378. MR MR2485323
3. Noga Alon, Michael Capalbo, Yoshiharu Kohayakawa, Vojtěch Rödl, Andrzej Ruciński, and Endre Szemerédi, *Universality and tolerance (extended abstract)*, 41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000), IEEE Comput. Soc. Press, Los Alamitos, CA, 2000, pp. 14–21. MR MR1931800
4. ______, *Near-optimum universal graphs for graphs with bounded degrees (extended abstract)*, Approximation, randomization, and combinatorial optimization (Berkeley, CA, 2001), Lecture Notes in Comput. Sci., vol. 2129, Springer, Berlin, 2001, pp. 170–180. MR MR1910361

5. Noga Alon, Michael Krivelevich, and Benny Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), no. 6, 629–644. MR MR2384408 (2009d:05110)
6. Stephen Alstrup and Theis Rauhe, *Small induced-universal graphs and compact implicit graph representations*, 43st Annual Symposium on Foundations of Computer Science (Vancouver, BC, 2002), IEEE Comput. Soc. Press, 2002, pp. 53–62.
7. József Balogh, Béla Csaba, Martin Pei, and Wojciech Samotij, *Large bounded degree trees in expanding graphs*, Electron. J. Combin. **17** (2010), no. 1, Research Paper 6, 9. MR MR2578901
8. Sandeep N. Bhatt, F. R. K. Chung, F. T. Leighton, and Arnold L. Rosenberg, *Universal graphs for bounded-degree trees and planar graphs*, SIAM J. Discrete Math. **2** (1989), no. 2, 145–155. MR MR990447 (90b:05071)
9. Michael Capalbo, *Explicit sparse almost-universal graphs for $\mathcal{G}(n, \frac{k}{n})$*, Random Structures Algorithms **37** (2010), no. 4, 437–454. MR 2760357 (2011k:05218)
10. Michael R. Capalbo and S. Rao Kosaraju, *Small universal graphs*, Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999), ACM, New York, 1999, pp. 741–749 (electronic). MR MR1798099 (2001i:05139)
11. Domingos Dellamonica, Jr. and Yoshiharu Kohayakawa, *An algorithmic Friedman-Pippenger theorem on tree embeddings and applications*, Electron. J. Combin. **15** (2008), no. 1, Research Paper 127, 14. MR MR2448877 (2009i:05059)
12. Domingos Dellamonica, Jr., Yoshiharu Kohayakawa, Vojtěch Rödl, and Andrzej Ruciński, *Universality of random graphs*, Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (New York), ACM, 2008, pp. 782–788. MR 2487648
13. ______, *An improved upper bound on the density of universal random graphs*, LATIN 2012: Theoretical informatics (Arequipa, 2012) (David Fernández-Baca, ed.), Springer, 2012, pp. 231–242.
14. ______, *Universality of random graphs*, SIAM J. Discrete Math. **26** (2012), no. 1, 353–374. MR 2902650
15. Alan Frieze and Michael Krivelevich, *Almost universal graphs*, Random Structures Algorithms **28** (2006), no. 4, 499–510. MR 2225704 (2007a:05126)
16. Svante Janson, *Poisson approximation for large deviations*, Random Structures Algorithms **1** (1990), no. 2, 221–229. MR 1138428 (93a:60041)
17. Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR MR1782847 (2001k:05180)
18. Svante Janson, Krzysztof Oleszkiewicz, and Andrzej Ruciński, *Upper tails for subgraph counts in random graphs*, Israel J. Math. **142** (2004), 61–92. MR 2085711 (2005e:05126)
19. Anders Johansson, Jeff Kahn, and Van Vu, *Factors in random graphs*, Random Structures Algorithms **33** (2008), no. 1, 1–28. MR 2428975 (2009f:05243)
20. H. A. Kierstead and A. V. Kostochka, *A short proof of the Hajnal-Szemerédi theorem on equitable colouring*, Combin. Probab. Comput. **17** (2008), no. 2, 265–270. MR 2396352 (2009a:05071)
21. Colin McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248. MR 1678578 (2000d:60032)