

Very fast construction of bounded-degree spanning graphs via the semi-random graph process

Omri Ben-Eliezer ^{*} Lior Gishboliner [†] Dan Hefetz [‡] Michael Krivelevich [§]

September 13, 2020

Abstract

Semi-random processes involve an adaptive decision-maker, whose goal is to achieve some predetermined objective in an online randomized environment. They have algorithmic implications in various areas of computer science, as well as connections to biological processes involving decision making. In this paper, we consider a recently proposed semi-random graph process, defined as follows: we start with an empty graph on n vertices, and in each round, the decision-maker, called **Builder**, receives a uniformly random vertex v , and must immediately (in an online manner) choose another vertex u , adding the edge $\{u, v\}$ to the graph. **Builder's** end goal is to make the constructed graph satisfy some predetermined monotone graph property. There are also natural offline and non-adaptive modifications of this setting.

We consider the property $\mathcal{P}_H$ of containing a spanning graph H as a subgraph. It was asked by N. Alon whether for every bounded-degree H , **Builder** can construct a graph satisfying $\mathcal{P}_H$ with high probability in $O(n)$ rounds. We answer this question positively in a strong sense, showing that any graph with maximum degree Δ can be constructed with high probability in $(3\Delta/2 + o(\Delta))n$ rounds, where the $o(\Delta)$ term tends to zero as $\Delta \rightarrow \infty$. This is tight (even for the offline case) up to a multiplicative factor of $3 + o_\Delta(1)$. Furthermore, for the special case where H is a forest of maximum degree Δ , we show that H can be constructed with high probability in $O(\log \Delta)n$ rounds. This is tight up to a multiplicative constant, even for the offline setting. Finally, we show a separation between *adaptive* and *non-adaptive* strategies, proving a lower bound of $\Omega(n\sqrt{\log n})$ on the number of rounds necessary to eliminate all isolated vertices w.h.p. using a non-adaptive strategy. This bound is tight, and in fact there are non-adaptive strategies for constructing a Hamilton cycle or a K_r -factor, which are successful w.h.p. within $O(n\sqrt{\log n})$ rounds.

^{*}Blavatnik School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel. Email: omrib@mail.tau.ac.il.

[†]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel. Email: liorgis1@mail.tau.ac.il. Supported in part by ERC Starting Grant 633509.

[‡]Department of Computer Science, Ariel University, Ariel 40700, Israel. Email: danhe@ariel.ac.il. Research supported by ISF grant 822/18.

[§]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel. Email: krivelev@tauex.tau.ac.il. Partially supported by USA-Israel BSF grants 2014361 and 2018267, and by ISF grant 1261/17.

1 Introduction

Recently, the following *semi-random graph process* was proposed by Peleg Michaeli, and analyzed by Ben-Eliezer, Hefetz, Kronenberg, Parczyk, Shikhelman, and Stojaković [3]. A single adaptive player, called **Builder**, starts with an empty graph G on a set V of n vertices. The process then proceeds in *rounds*, where in each round **Builder** is offered a uniformly random vertex v , and chooses an edge of the form $\{v, u\}$ to add to the graph G . **Builder**’s objective is typically to construct a graph that satisfies some predetermined monotone graph property; for example, to make G an expander with certain parameters, or to have G contain a Hamilton cycle. The natural question arising in this context is the following:

*Given a monotone graph property $\mathcal{P}$, how many rounds of the semi-random graph process are required for **Builder** to construct (with high probability¹) a graph which satisfies $\mathcal{P}$?*

Semi-random problems of this type, involving both randomness and intelligent choices made by a “decision-maker”, have been widely studied in the algorithmic literature. One of the first (and most famous) results on such processes, established by Azar et al. [2], concerns sequential allocation of n balls into n bins, where the goal is to minimize the number of balls in the fullest bin. It is well-known that if each ball is simply assigned to a bin uniformly at random, then w.h.p., the fullest bin will contain $\Theta(\ln n / \ln \ln n)$ balls at the end of the process. However, as was shown in [2], very limited “intelligent intervention” substantially improves the above bound: if, instead of the random assignment, for any ball we are given *two (random) choices* of bins to pick from, then the trivial strategy of always choosing the least loaded bin out of the two offered, results w.h.p. in the maximum bin load dropping to $\Theta(\ln \ln n)$ – an exponential improvement. This idea has inspired many subsequent theoretical and practical results in various contexts within computer science, see e.g. [8, 22, 25] for a small sample of these.

In a sense, semi-random processes can be viewed as settings where an *online algorithm* aims to achieve a predetermined objective in a *randomized environment*. As opposed to the “standard” setting where online algorithms are measured in terms of their worst-case performance, in the semi-random setting the task is to design online algorithms that achieve their goal with high (or at least constant) probability, and require as few rounds as possible. Further discussion of related semi-random graph models, such as the so-called Achlioptas model, can be found in Section 1.1.

In this paper we continue the investigation into the semi-random graph process. The first work on this topic [3] proved upper and lower bounds on the number of rounds required to w.h.p. satisfy various properties of interest. Among the upper bounds were a $O(n^{1-\varepsilon})$ bound for the property of containing a copy of any fixed graph H (here ε depends on H), a $O(n)$ bound for containing a perfect matching or a Hamilton cycle, and a $O(\Delta n)$ bound for the property of having minimum degree Δ , as well as for the property of Δ -vertex-connectivity. These results prompted Noga Alon to ask whether it is the case that every given (spanning) graph of bounded maximum degree can be constructed w.h.p. in $O(n)$ rounds in this model. Formally, define $\mathcal{P}_H$ as the property of containing an (unlabeled) copy of H , i.e., as the property that there exists an injection $\varphi: V(H) \rightarrow V(G)$ (where G is the graph constructed by **Builder**), so that $\{\varphi(i), \varphi(j)\} \in E(G)$ for every $\{i, j\} \in E(H)$.

Question 1.1 (Alon, Question 6.2 in [3]). *Consider the semi-random graph process over n vertices. Is it true that for every graph H on n vertices with bounded maximum degree, there exists a strategy which enables **Builder** to w.h.p. construct a copy of H in $O(n)$ rounds?*

The main result of this paper gives a positive answer to the above question (see Theorem 1.3).

The offline setting Before addressing Question 1.1, let us consider the easier *offline* setting, where **Builder** is provided *in advance* with the full sequence of vertices offered throughout the process. In other words, in

¹With probability that tends to one as $n \rightarrow \infty$; abbreviated w.h.p. henceforth.

this setting **Builder** does not need to make his decisions online, but can rather choose all edges at once after seeing the sequence of random vertices.

As an example, consider the case where **Builder**'s goal is to construct a triangle-factor². Observe that if at some point of the process, at least $\frac{2n}{3}$ different vertices have been offered, among which at least $\frac{n}{3}$ vertices have been offered at least twice, then **Builder** can already construct a triangle-factor (in the offline setting). Indeed, **Builder** simply partitions the vertices into triples $\{u, v, w\}$, where u was offered at least twice and v was offered at least once, and chooses the edges $\{u, v\}$ and $\{u, w\}$ at rounds when u was offered, and the edge $\{v, w\}$ at a round when v was offered. It is easy to check that $O(n)$ rounds suffice w.h.p. to have at least $\frac{2n}{3}$ different vertices offered, and moreover to have at least $\frac{n}{3}$ vertices offered at least twice. Thus, in the offline setting **Builder** can construct a triangle-factor in $O(n)$ rounds. This $O(n)$ bound can in fact be generalized in a strong sense to *any* bounded-degree target graph H . Proposition 4.1 in [3] – which presents necessary and sufficient general winning conditions for **Builder** in the offline setting – implies that **Builder** wins the game as soon as the list of offered vertices allows the construction of a suitable *orientation* of H . In Section 3.1 we show how this can be used to get the following general offline result.

Proposition 1.2. *Let $\Delta, n > 0$ be integers, and let H be an n -vertex graph of maximum degree Δ . In the offline version of the semi-random process on n vertices, **Builder** has a strategy allowing him to construct a copy of H in $(\Delta/2 + o(\Delta))n$ rounds w.h.p.*

The $o(\Delta)$ term here is a function of Δ satisfying $o(\Delta)/\Delta \rightarrow 0$ as $\Delta \rightarrow \infty$. Proposition 1.2 substantially extends Theorem 1.9 in [3], which showed a similar result for the property of having minimum degree at least k (with an explicit dependence on k). Proposition 1.2 is clearly optimal up to the $o(\Delta)$ term, since Δ -regular graphs on n vertices have exactly $\Delta n/2$ edges (and hence trivially require at least this number of rounds). More interestingly, it turns out that the $o(\Delta)$ -term is unavoidable; it follows from [3, Theorem 1.9] that at least $(1/2 + \varepsilon_\Delta)\Delta n$ rounds are required for **Builder** to construct a graph of minimum degree at least Δ , where $\varepsilon_\Delta > 0$ (and $\varepsilon_\Delta \rightarrow 0$ as $\Delta \rightarrow \infty$).

Main result: online strategy for constructing bounded-degree spanning graphs We now return to the more challenging *online* setting, where **Builder** is offered vertices one-by-one and must (irrevocably) decide which edge to add immediately after being offered a vertex. Our main result in this paper, Theorem 1.3, asserts that **Builder** can construct any given bounded-degree spanning graph in $O(n)$ rounds w.h.p.

Theorem 1.3. *Let $\Delta, n > 0$ be integers and let H be an n -vertex graph of maximum degree Δ . In the online version of the semi-random graph process on n vertices, **Builder** has a strategy guaranteeing that w.h.p., after*

$$\begin{cases} (\Delta/2 + o(\Delta))n, & \text{if } \Delta = \omega(\log n) \\ (3\Delta/2 + o(\Delta))n, & \text{otherwise} \end{cases}$$

rounds of the process, the constructed graph will contain a copy of H .

As before, the $o(\Delta)$ term here is a function of Δ satisfying $o(\Delta)/\Delta \rightarrow 0$ as $\Delta \rightarrow \infty$. Note that Δ is allowed to depend on n arbitrarily. Theorem 1.3 answers Question 1.1 in a strong sense: not only can any bounded-degree graph be constructed w.h.p. in a linear number of rounds, but in fact, the dependence on the maximum degree is very modest. This result clearly illustrates the power of semi-random algorithms compared to their truly random counterparts; see the discussion below on the appearance of various spanning structures in the random graph process.

The notion of *competitive ratio* [6] refers to the performance of an online algorithm compared to the best offline algorithm for the same problem. In view of the trivial $\Delta n/2$ lower bound and Proposition 1.2, our algorithm is $(3 + o_\Delta(1))$ -competitive for general H , where the $o_\Delta(1)$ term tends to zero as Δ tends to infinity.

²For a graph F and an integer n divisible by $|V(F)|$, the n -vertex F -factor is the graph which consists of $n/|V(F)|$ vertex-disjoint copies of F .

As an open question, it will be very interesting to determine the optimal competitive ratio of an online algorithm for this problem.

Problem 1.4. *In the online version of the semi-random process on n vertices, is it true that for every n -vertex graph H of maximum degree Δ , Builder has a strategy to construct a copy of H w.h.p. in $(\Delta/2 + o(\Delta))n$ rounds?*

We end this discussion with two additional problems, which ask for tight bounds on the number of rounds required to construct some specific graphs H of particular interest.

Problem 1.5 (Tight bounds for constructing a Hamilton cycle). *Is there a number α_{Ham} such that in the online version of the semi-random process on n vertices, Builder can w.h.p. construct a Hamilton cycle in $(\alpha_{\text{Ham}} + o(1))n$ rounds, but w.h.p. cannot accomplish this in $(\alpha_{\text{Ham}} - o(1))n$ rounds? If so, what is the value of α_{Ham} ?*

Problem 1.6 (Tight bounds for constructing a K_r -factor). *For each $r \geq 2$, is there a number α_r such that in the online version of the semi-random process on n vertices, for n which is divisible by r , Builder can w.h.p. construct a K_r -factor in $(\alpha_r + o(1))n$ rounds, but w.h.p. cannot accomplish this in $(\alpha_r - o(1))n$ rounds? If so, what is the value of α_r ?*

We note that Problem 1.6 is open even for $r = 2$, in which case Builder's goal is to construct a perfect matching. Both the perfect matching problem and the Hamilton cycle problem were already considered in [3], where some bounds for these problems were obtained.

Constructing a bounded-degree spanning forest Theorem 1.3 establishes that the (typical) number of rounds needed to construct a general spanning graph H of maximum degree Δ is $O(\Delta n)$. This is clearly tight for graphs whose average degree is $\Theta(\Delta)$. It is now natural to ask if we can break the $\Theta(\Delta n)$ barrier for graphs with a much smaller average degree, such as trees. The next result answers this question positively.

Theorem 1.7. *Let $\Delta, n > 0$ be integers and let T be an n -vertex forest of maximum degree Δ . In the online version of the semi-random process on n vertices, Builder has a strategy guaranteeing that w.h.p., after $O(n \log \Delta)$ rounds of the process, the constructed graph will contain a copy of T .*

The next proposition shows that the dependence on n and Δ in Theorem 1.7 is tight even for the offline version of the semi-random process.

Proposition 1.8. *For every $\Delta \geq 1$ and for every $n \geq n_0(\Delta)$, there exists a forest T with n vertices and maximum degree Δ satisfying the following. In the offline version of the semi-random graph process, w.h.p. Builder needs $\Omega(n \log \Delta)$ rounds in order to construct a copy of T .*

We conclude this section by proposing the following problem:

Problem 1.9. *How many rounds are required to construct n -vertex d -degenerate graphs of maximum degree Δ ?*

An answer to Problem 1.9 would generalize Theorem 1.7, as forests are exactly the 1-degenerate graphs.

Non-adaptive strategies It is fairly natural to inquire whether imposing the restriction of *non-adaptivity* handicaps Builder, and if so, to which extent exactly. Here, by non-adaptivity we mean that, in a sense, Builder's choices are decided upon beforehand, and do not depend on the situation at any given round of the process. The precise definition that we use is as follows. A *non-adaptive strategy* consists of a family $\mathcal{L}$ of adjacency lists $\mathcal{L} = \{L^w : w \in [n]\}$, where for each $w \in [n]$, the list $L^w = (L^w(i) : i = 1, \dots, n-1)$ is a permutation of $[n] \setminus \{w\}$. The lists $\mathcal{L} = \{L^w : w \in [n]\}$ are specified *in advance* (i.e., before the sequence of random vertices starts being exposed). Playing according to such a strategy means that during the vertex exposure process $w_1, w_2, \dots$, if in a given round vertex w appears for the i th time, $i \geq 1$, then Builder is obliged

to connect w to the i th vertex $L^w(i)$ on its list. (To avoid ambiguities, let us assume that if w has already been connected to $L^w(i)$, then **Builder** simply skips his move. This assumption will not change much in our analysis.)

It turns out that the non-adaptivity assumption indeed hampers **Builder** — it takes him typically $\Omega(n\sqrt{\log n})$ rounds to get rid of isolated vertices, as stated in Theorem 1.10 below. This is in rather sharp contrast with the situation for general (i.e. adaptive) strategies. Indeed, it is easy to see that **Builder** can construct a connected graph in $n - 1$ rounds (with probability 1); constructing a graph with no isolated vertices can be done w.h.p. even faster, in $(\ln 2 + o(1))n$ rounds (see [3]); and finally, Theorem 1.3 shows that in fact every bounded-degree graph can be constructed in $O(n)$ rounds. The three theorems below are for the online version of the semi-random graph process.

Theorem 1.10. *In the semi-random process on n vertices, any non-adaptive strategy requires w.h.p. $\Omega(n\sqrt{\log n})$ rounds to construct a graph in which none of the n vertices is isolated.*

It turns out that the lower bound of Theorem 1.10 is tight in a strong sense: by executing an appropriate non-adaptive strategy for $O(n\sqrt{\log n})$ rounds, **Builder** can construct w.h.p. several important spanning structures, such as a Hamilton cycle and a K_r -factor (for fixed r). This is expressed in the following two theorems.

Theorem 1.11. *In the semi-random process on n vertices, there is a non-adaptive strategy allowing **Builder** to construct a Hamilton cycle in $8n\sqrt{\log n}$ rounds w.h.p.*

Theorem 1.12. *For every $r \geq 2$ there is $C = C(r)$ such that for every n which is divisible by r , there is a non-adaptive strategy allowing **Builder** to construct a K_r -factor in $Cn\sqrt{\log n}$ rounds w.h.p.*

We remark that while the above definition of non-adaptivity is deterministic in nature (in the sense that the lists $\{L^w : w \in [n]\}$ are predetermined), the proof of Theorem 1.10 can be easily adapted (with the same asymptotic lower bound) to apply also to “random non-adaptive strategies”, i.e. strategies in which for every $w \in [n]$ and $i \geq 1$, the vertex which **Builder** connects to w at the i th time that w is sampled is drawn from some predetermined probability distribution on $[n] \setminus \{w\}$.

The situation in the (“purely”)-random graph process A common theme in our results is that introducing “intelligent choices” into a random setting can allow for a dramatic improvement. (Perhaps the first appearance of this theme is in the aforementioned work of Azar et al. [2].) To illustrate this phenomenon in the setting of random graph processes, let us compare our Theorem 1.3 with the situation in which all edge-choices are made completely randomly. Recall that the *random graph process* $\tilde{G} = (G_m)_{m=0}^N$ is defined by choosing a random permutation $e_1, \dots, e_N$ of all $N = \binom{n}{2}$ edges of K_n , and letting G_m be the graph whose edge-set is $\{e_1, \dots, e_m\}$. We refer the reader to [9, 11] for an overview of this classical object, as well as the standard random graph models $G(n, m)$ and $G(n, p)$, which are mentioned below. Note that for each $0 \leq m \leq \binom{n}{2}$, the graph G_m is distributed as the Erdős-Rényi graph $G(n, m)$, i.e. as a random graph chosen uniformly among all graphs with m edges and n (labeled) vertices.

The appearance of bounded-degree spanning graphs in the random graph process (or equivalently, in $G(n, m)$) has been thoroughly investigated³. For starters, a standard first moment argument (see e.g. [11]) shows that a copy of $K_{\Delta+1}$ only appears in the random graph process after roughly $n^{2-2/\Delta}$ rounds. Thus, obtaining even a single $K_{\Delta+1}$ -copy (let alone a $K_{\Delta+1}$ -factor) requires much more than a linear number of rounds. Determining the typical time of the appearance of a $K_{\Delta+1}$ -factor turned out to be a difficult problem. Following a long line of research, this problem was settled by (a special case of) a celebrated result by Johansson, Kahn and Vu [12], which states that a $K_{\Delta+1}$ -factor appears in the random graph process at

³We note that the results surveyed here were actually proved for the binomial random graph $G(n, p)$, which is the graph obtained by selecting each of the $\binom{n}{2}$ edges of K_n with probability p and independently. It is well known — see e.g. [11, Section 1.4] or [9, Section 1.1] — that $G(n, m)$ is closely related to $G(n, p)$ when $p = m/\binom{n}{2}$, allowing the transfer of results between the two models.

around $m = n^{2-2/(\Delta+1)}(\log n)^{1/\binom{\Delta+1}{2}}$. A more general discussion on the appearance of bounded-degree graphs H other than clique-factors (and on the closely-related notion of *universality*) can be found in the recent work of Ferber, Kronenberg, and Luh [7]; see in particular Conjecture 1.5 there.

Similar superlinear lower bounds (on the number of edges required in order to typically contain a $K_{\Delta+1}$ -factor, or even a single $K_{\Delta+1}$) are known or can be shown for various other random graph models, such as the random regular graph, or the model G_{k-out} (where one connects each vertex to exactly k other randomly chosen vertices, discarding repetitions). Thus, while it was shown in [3] that the semi-random graph process can *simulate* the random models $G(n, p)$, $G(n, m)$, and G_{k-out} , the above discussion indicates that these cannot help in solving Question 1.1, and we must utilize the power of the intelligent player, **Builder**, in a more imaginative way.

A similar comparison can be made between Theorem 1.7 and the emergence of (bounded-degree) spanning trees in the random graph process. It is well known that in the random graph process, w.h.p. the last isolated vertex disappears only at around $m = \frac{1}{2}n \log n$. Thus, a superlinear number of rounds is required in order to contain a spanning tree w.h.p. Again, we observe here that intelligent choices speed up the time required to reach the given goal: it takes $\Theta(n \log n)$ rounds for the random graph process to contain even a single spanning tree, whilst in our semi-random graph process, the number of rounds required to contain a prescribed bounded-degree spanning tree is only $O(n)$.

It is worth mentioning that a recent breakthrough of Montgomery [18], which confirms a conjecture of Kahn [13], shows that for a fixed Δ , w.h.p. all spanning trees with maximum degree Δ appear in G_m after $m = Cn \log n$ rounds (where $C = C(\Delta)$ is a large enough constant). We refer the reader to [18] for further references to many other related works on this subject.

1.1 Related Work : Semi-Random Processes

Perhaps the most studied semi-random graph process is now known as the *Achlioptas process*, and was proposed by Dimitris Achlioptas in 2000. Similarly to our semi-random process, the Achlioptas process is a one-player game in which the player, **Builder**, gradually constructs a graph. The process runs in rounds, where in each round, two uniformly random edges are picked from the set of all $\binom{n}{2}$ edges of the n -vertex complete graph, or alternatively (depending on the version of the process) from all edges untaken at this point. These two edges are offered to **Builder**, who then must choose exactly one of them and add it to the graph.

While in our random graph process **Builder**'s goal is always to make his graph *satisfy* some given graph property, in the context of the Achlioptas process the goal is often to *avoid* satisfying a given property for as long as possible. In fact, Achlioptas's original question was whether **Builder** can delay the appearance of a giant component beyond its typical time of appearance in the ("purely"-)random graph process. This question was answered positively by Bohman and Frieze [4], see also [1, 5, 14, 23, 24]. Similar problems have also been studied for other properties or objectives; for example, the problem of avoiding a fixed subgraph [15, 20], or the problem of speeding up the appearance of a Hamilton cycle [16]. Achlioptas-like processes involving two choices were also investigated in other contexts, see e.g. [19] for a geometric perspective.

1.2 Paper Organization and Notation

In Section 2 we state several auxiliary results. Section 3 contains the proofs of Theorem 1.3 and Proposition 1.2, as well as the description and analysis of Strategy 3.7, which is the key tool used in our proofs. Theorem 1.7 and Proposition 1.8 are proved in Section 4. Finally, the proofs of Theorems 1.10, 1.11 and 1.12 appear in Section 5. Since the statements of Theorems 1.3 and 1.7 are asymptotic (in both n and Δ), we always assume, where needed, that n and Δ are sufficiently large. All logarithms are base e . We omit floor and ceiling signs whenever these are not crucial.

2 Preliminaries

We start by stating three known concentration inequalities that will be used in this paper. The first is a standard Chernoff-type bound (see, e.g., [11]), the second is a simplified version of Azuma's inequality (see, e.g., [11, Theorem 2.27]), and the third is a simplified version of Talagrand's inequality (see, e.g., [17]).

Lemma 2.1. *Let X be a binomial random variable. Then, for every $\lambda \geq 0$, it holds that*

$$\mathbb{P}[X \leq \mathbb{E}[X] - \lambda] \leq e^{-\frac{\lambda^2}{2\mathbb{E}[X]}}$$

and that

$$\mathbb{P}[X \geq \mathbb{E}[X] + \lambda] \leq e^{-\frac{\lambda^2}{2(\mathbb{E}[X] + \lambda/3)}}.$$

Lemma 2.2. [11, Theorem 2.27] *Let X be a non-negative random variable, not identically 0, which is determined by T independent trials $w_1, \dots, w_T$. Suppose that $c \in \mathbb{R}$ is such that changing the outcome of any one of the trials can change the value of X by at most c . Then, for every $\lambda \geq 0$, it holds that*

$$\mathbb{P}[X \leq \mathbb{E}[X] - \lambda] \leq e^{-\frac{\lambda^2}{2c^2T}}$$

and that

$$\mathbb{P}[X \geq \mathbb{E}[X] + \lambda] \leq e^{-\frac{\lambda^2}{2c^2T}}.$$

Lemma 2.3. [17, Pages 80-81] *Let X be a non-negative random variable, not identically 0, which is determined by T independent trials $w_1, \dots, w_T$. Suppose that $c, g > 0$ are such that*

1. *Changing the outcome of any one of the trials can change the value of X by at most c .*
2. *For every s , if $X \geq s$ then there is a set of at most $g \cdot s$ trials whose outcomes certify⁴ that $X \geq s$.*

Then, for every $0 \leq \lambda \leq \mathbb{E}[X]$, it holds that

$$\mathbb{P}[|X - \mathbb{E}[X]| > \lambda + 60c\sqrt{g\mathbb{E}[X]}] \leq 4e^{-\frac{\lambda^2}{8c^2g\mathbb{E}[X]}}.$$

We will also need the following lemma regarding “balanced” orientations of graphs.

Lemma 2.4. *Let H be an n -vertex graph of maximum degree Δ . Then there exists an orientation D of the edges of H which satisfies the following two conditions:*

- (a) $d_D^+(u) \leq \lfloor \Delta/2 \rfloor + 1$ for every $u \in V(H)$;
- (b) *there exists a set $A \subseteq V(H)$ of size $|A| \geq \frac{n}{\Delta^2+1}$ such that $d_D^+(u) = 0$ for every $u \in A$.*

Proof. Let H^2 denote the square of H , that is, the graph obtained from H by adding an edge between any two vertices at distance 2 in H . Let $A \subseteq V(H)$ be a maximum independent set in H^2 ; clearly $|A| \geq \frac{n}{\Delta^2+1}$. Let $H_0 = H \setminus A$. If $d_{H_0}(u)$ is even for every $u \in V(H_0)$, then let $H_1 = H_0$. Otherwise, let H_1 denote the graph obtained from H_0 by adding a new vertex x and connecting it by an edge to every vertex of odd degree in H_0 . Orient the edges of each connected component of H_1 along some Eulerian cycle of that component. Orient every edge of $E_H(V(H) \setminus A, A)$ from $V(H) \setminus A$ to A . Delete x and denote the resulting oriented graph by D . Observe that A is independent in H and thus D is an orientation of all the edges of H . It is evident that D satisfies (b). Moreover, D satisfies (a) as A is independent in H^2 and thus $d_H(u, A) \leq 1$ for every $u \in V(H) \setminus A$. $\square$

⁴To be precise, this means that if $(w_1, \dots, w_T)$ is such that $X \geq s$, then there is $k \leq g \cdot s$ and $1 \leq i_1 < \dots < i_k \leq T$ such that changing the outcome of any trials other than $w_{i_1}, \dots, w_{i_k}$ does not change the fact that $X \geq s$.

3 Constructing Spanning Graphs of Given Maximum Degree

In this section we prove Theorem 1.3 and Proposition 1.2. The tools we develop here will also be used in the proof of Theorem 1.7 in Section 4. We start by introducing some definitions and sketching a rough outline of the proof scheme that we will use to prove Theorem 1.3. From this point onward, we fix an n -vertex graph H with maximum degree Δ . We assume that the ground-set of vertices for the semi-random process is $[n]$. The following definition will play an important role in our arguments.

Definition 3.1. *Let G be a graph on the vertex set $[n]$, and let $\varphi : V(H) \rightarrow [n]$ be a bijection. We say that a subset $A \subseteq V(H)$ is (G, φ) -good if φ maps all edges of H contained in A to edges of G ; that is, A is (G, φ) -good if $\{\varphi(x), \varphi(y)\} \in E(G)$ for every $x, y \in A$ such that $\{x, y\} \in E(H)$.*

The graph G under consideration will always be Builder's graph at some given moment during the process. It will usually be clear which moment we are considering, and so we will omit G from the notation, simply writing " φ -good". Note that, if at some point during the process, there is a bijection $\varphi : V(H) \rightarrow [n]$ such that $V(H)$ is φ -good, then Builder has succeeded in constructing a copy of H .

Our strategy consists of two stages: in the first stage, Builder fixes an arbitrary bijection $\varphi : V(H) \rightarrow [n]$ and plays so as to construct a φ -good set which is as large as possible. We show (see Lemma 3.5 below) that Builder can w.h.p. guarantee the existence of a φ -good set which covers "almost all" of the vertices of H , within $(\Delta/2 + o(\Delta))n$ rounds. This part of the argument is fairly straightforward, and is given in Section 3.1. In the second stage, which is far more involved and constitutes the heart of the proof, Builder tries to iteratively extend this φ -good set by updating the embedding φ . We will show that by using a suitable "role-switching" strategy (i.e., Strategy 3.7), Builder can ensure that w.h.p. $V(H)$ will be φ -good after $(\Delta + o(\Delta))n$ additional rounds. This is done in Lemma 3.14.

3.1 The Initial Embedding, Proposition 1.2, and the Easy Case of Theorem 1.3

In this section we describe and analyze the first stage of Builder's strategy. Along the way we prove Proposition 1.2, as well as the easy part of Theorem 1.3, which corresponds to the regime $\Delta = \omega(\log n)$. In what follows, we will need the following lemma.

Lemma 3.2. *Let n, d be positive integers, let $\alpha \in (0, 0.1)$, and suppose that $d \gg \log(1/\alpha)$. In the course of $(d + o(d))n$ rounds of the semi-random process on n vertices, the number of vertices $i \in [n]$ which were offered at most d times is w.h.p. less than αn .*

Proof. Set $\ell := (d + \sqrt{6d \log(1/\alpha)})n = (d + o(d))n$ (here we use our assumption that $d \gg \log(1/\alpha)$), and let $w_1, \dots, w_\ell$ be the first ℓ random vertices offered to Builder. Let U be the set of all vertices $i \in [n]$ such that $|\{1 \leq j \leq \ell : w_j = i\}| \leq d$. Evidently, $|\{1 \leq j \leq \ell : w_j \in U\}| \leq d \cdot |U|$. So in order to prove the lemma, it is enough to show that w.h.p., every $U \subseteq [n]$ of size αn satisfies the inequality $X_U := |\{1 \leq j \leq \ell : w_j \in U\}| > d \cdot |U| = d\alpha n$.

Fix any $U \subseteq [n]$ of size $|U| = \alpha n$. Note that $X = X_U$ has the distribution $\text{Bin}(\ell, \frac{|U|}{n})$, and in particular $\mathbb{E}[X] = \ell|U|/n = \alpha\ell$. By Lemma 2.1 with $\lambda = \alpha\ell - d\alpha n$, we have

$$\begin{aligned} \mathbb{P}[X < d\alpha n] &= \mathbb{P}[X < \mathbb{E}[X] - \lambda] \leq \exp\left(-\frac{\alpha^2 n^2 \cdot (\ell/n - d)^2}{2\alpha\ell}\right) \\ &\leq \exp\left(-\frac{\alpha n \cdot (\ell/n - d)^2}{2d + o(d)}\right) = e^{-(3+o_d(1))\alpha \log(1/\alpha)n} \leq e^{-2.5\alpha \log(1/\alpha)n}, \end{aligned}$$

where the last inequality holds if d is large enough. Now our assertion follows by taking the union bound over all

$$\binom{n}{\alpha n} \leq (e/\alpha)^{\alpha n} = e^{(\log(1/\alpha)+1) \cdot \alpha n} \leq e^{2\alpha \log(1/\alpha)n}$$

sets $U \subseteq [n]$ of size $|U| = \alpha n$. $\square$

Proof of Proposition 1.2. Since the statement of the proposition is asymptotic in Δ , we may (and will) assume that Δ is large enough, where needed. Moreover, in light of Theorem 1.3, we may assume that $\Delta = O(\log n)$ (because otherwise, Builder can w.h.p. construct a copy of H in $(\Delta/2 + o(\Delta))n$ rounds even in the online version of the semi-random process). Let D be an orientation of the edges of H as in Lemma 2.4. Let $w_1, w_2, \dots$ denote the sequence of random vertices Builder is offered. Let $m(D)$ denote the smallest integer m for which there exists a bijection $\varphi : V(H) \rightarrow [n]$ such that $\varphi(u)$ appears in $(w_1, w_2, \dots, w_m)$ at least $d_D^+(u)$ times for every $u \in V(H)$. It follows by Proposition 4.1 in [3] that the number of rounds needed for Builder to construct a copy of H in the offline version of the semi-random process is at most $m(D)$. Hence, in order to complete the proof of this proposition, it suffices to show that w.h.p., in the course of $m := (\Delta/2 + o(\Delta))n$ rounds, at least $n - n/(\Delta^2 + 1)$ vertices are offered at least $\lfloor \Delta/2 \rfloor + 1$ times each. But this is just the statement of Lemma 3.2 with parameters $d = \lfloor \Delta/2 \rfloor$ and $\alpha = 1/(\Delta^2 + 1)$. $\square$

In subsequent proofs, Builder will employ the following simple (non-adaptive) strategy.

Strategy 3.3. Let H be an n -vertex graph of maximum degree Δ . Fix an orientation D of the edges of H which satisfies property (a) of Lemma 2.4. Let $\varphi : V(H) \rightarrow [n]$ be an arbitrary bijection. At any given round of the process, having been offered a random vertex w , Builder chooses an arbitrary vertex $u \in \varphi(N_D^+(\varphi^{-1}(w)))$ which is not adjacent to w in his current graph, and claims the edge $\{u, w\}$; if no such vertex u exists, then Builder claims an arbitrary edge incident with w .

In the following proposition we prove the easy part of Theorem 1.3.

Proposition 3.4. For every $\varepsilon \in (0, 1)$ there exists an integer C for which the following holds. Let n and $\Delta = \Delta(n) \geq C \log n$ be positive integers and let H be an n -vertex graph of maximum degree at most Δ . Then, in the online version of the semi-random process on n vertices, Builder has a strategy guaranteeing that w.h.p., after $(1 + \varepsilon) \frac{\Delta n}{2}$ rounds of the process, his graph will contain a copy of H .

Proof. Builder executes Strategy 3.3 for $\ell := (1 + \varepsilon) \frac{\Delta n}{2}$ rounds. In the notation of Strategy 3.3, it is evident that if for every $x \in V(H)$, the vertex $\varphi(x)$ is offered at least $d_D^+(x)$ times, then Builder is successful in building a copy of H . Therefore, in order to complete the proof of the proposition, it suffices to show that w.h.p., for every $1 \leq i \leq n$, the vertex i is offered at least $\lfloor \Delta/2 \rfloor + 1$ times in the course of these ℓ rounds of the process.

Fix any $1 \leq i \leq n$, and let Z_i be the random variable counting the number of times i is offered during the ℓ rounds. Then $Z_i \sim \text{Bin}(\ell, 1/n)$, implying that $\mathbb{E}[Z_i] = \ell/n = (1 + \varepsilon)\Delta/2$. Applying Lemma 2.1 with $\lambda = \varepsilon\Delta/2$, we obtain

$$\mathbb{P}[Z \leq \Delta/2] = \mathbb{P}[Z \leq \mathbb{E}[Z] - \varepsilon\Delta/2] \leq \exp\left(-\frac{(\varepsilon\Delta/2)^2}{(1 + \varepsilon)\Delta}\right) \leq e^{-\Delta\varepsilon^2/8} \leq n^{-C\varepsilon^2/8} \leq 1/n^2$$

where the penultimate inequality holds since $\Delta \geq C \log n$, and the last inequality holds for $C \geq 16\varepsilon^{-2}$. A union bound then implies that with probability at least $1 - \frac{1}{n} = 1 - o(1)$, every $1 \leq i \leq n$ was offered at least $\lfloor \Delta/2 \rfloor + 1$ times, as required. $\square$

We now return to the main case of Theorem 1.3, assuming henceforth that $\Delta = O(\log n)$. In the following lemma, we analyze the aforementioned first stage of Builder's strategy.

Lemma 3.5. Let n, Δ be positive integers, let $\alpha \in (0, 0.1)$, and suppose that $\Delta \gg \log(1/\alpha)$. Let H be an n -vertex graph of maximum degree Δ and let $\varphi : V(H) \rightarrow [n]$ be a bijection. Then, Builder has a strategy guaranteeing that after $(\Delta/2 + o(\Delta))n$ rounds of the process, w.h.p. there will be a φ -good set $A \subseteq V(H)$ of size at least $(1 - \alpha)n$.

Proof. Builder executes Strategy 3.3 for $\ell = (\Delta/2 + o(\Delta))n$ rounds. Let A' be the set of all vertices $x \in V(H)$ such that $\varphi(x)$ was offered at least $\lfloor \Delta/2 \rfloor + 1$ times in the course of the ℓ rounds. Apply Lemma 3.2 with $d = \lfloor \Delta/2 \rfloor$ and with $\frac{\alpha}{2\Delta}$ in place of α , to conclude that w.h.p. we have $|A'| \geq (1 - \frac{\alpha}{2\Delta})n$. Here we use the assumption that $\Delta \gg \log(1/\alpha)$, which is necessary in order to apply Lemma 3.2 with the above parameters. From now on we assume that $|A'| \geq (1 - \frac{\alpha}{2\Delta})n$ (which happens with high probability).

Now, let $B = \{u \in V(H) : N_D^-(u, V(H) \setminus A') \neq \emptyset\}$, where D is the orientation from Strategy 3.3. It readily follows from the description of Builder's strategy (namely, Strategy 3.3) that if $x \in A' \setminus B$, then after ℓ rounds of the process, $\varphi(x)$ is adjacent in Builder's graph to every vertex of $\varphi(N_H(x))$. Thus, $A := A' \setminus B$ is φ -good. Since the maximum degree of H is Δ , it follows that $|B| \leq \Delta \cdot |V(H) \setminus A'| \leq \frac{\alpha n}{2}$. We conclude that w.h.p. $|A| \geq (1 - \alpha)n$, completing the proof. $\square$

3.2 Improving the Embedding

In this section we introduce and analyze Builder's strategy for the second stage (see Strategy 3.7 below). Our starting point is the state of Builder's graph immediately after applying the strategy given by Lemma 3.5. (The value of the parameter α , with which this lemma is applied, will be chosen later; the bijection φ can be chosen arbitrarily.) Our goal is to iteratively update φ , so as to maintain a φ -good set which gradually increases in size until it equals $V(H)$.

Before delving into the details, let us illustrate the idea behind Strategy 3.7 by considering the following “toy” example: suppose that at some point during the process, Builder has already managed to obtain a bijection $\varphi : V(H) \rightarrow [n]$ for which there is a φ -good set A of size $n - 1$. Let b denote the unique element of $V(H) \setminus A$. The fact that A is φ -good means that $\varphi(A)$ spans a copy of $H[A]$ in Builder's graph. So in order to make φ an embedding of H into Builder's graph, it remains to connect $\varphi(b)$ to all of the vertices in $\varphi(N_H(b))$. A naive way of doing this would be for Builder to wait until $\varphi(b)$ will have been offered $d_H(b)$ times, and at each such time, to connect $\varphi(b)$ to a new vertex in $\varphi(N_H(b))$. This, however, will not work, since the probability that $\varphi(b)$ is offered (even once) in the course of $O(n)$ rounds does not tend to 1. So instead, Builder will try to find another vertex in $[n]$ to “play” the role $\varphi(b)$, and to have $\varphi(b)$ play the role which was previously played by that other vertex. To this end, Builder fixes (a large number of) vertices $a_1, \dots, a_m \in A$ which are not adjacent to b . (We note that in order to make the strategy work, we need the additional assumption that the neighbourhoods of $a_1, \dots, a_m$ are pairwise-disjoint, but the reader may ignore this issue at the moment.) Now Builder acts as follows: each time a vertex of $\{\varphi(a_1), \dots, \varphi(a_m)\}$ is offered, Builder connects it to some new vertex of $\varphi(N_H(b))$; and each time a vertex of $\bigcup_{i=1}^m \varphi(N_H(a_i))$ is offered, Builder connects it to $\varphi(b)$. Now, if at some point there is an index $1 \leq i \leq m$ such that $\varphi(a_i)$ has already been offered at least $\Delta(H)$ times and every vertex in $\varphi(N_H(a_i))$ has already been offered at least once, then at this point $\varphi(a_i)$ is adjacent in Builder's graph to every vertex of $\varphi(N_H(b))$, and $\varphi(b)$ is adjacent in Builder's graph to every vertex of $\varphi(N_H(a_i))$. Hence, Builder can now safely “switch” the roles of $\varphi(a_i)$ and $\varphi(b)$. Formally, Builder defines a new bijection $\varphi' : V(H) \rightarrow [n]$ by setting $\varphi'(b) = \varphi(a_i)$, $\varphi'(a_i) = \varphi(b)$, and $\varphi'(x) = \varphi(x)$ for every $x \in V(H) \setminus \{b, a_i\}$. Then φ' is an embedding of H into Builder's graph. A key point of this method is that, since there are many “candidates” for the role of b (i.e. the vertices $a_1, \dots, a_m$), it is very likely that one of them will indeed be chosen to be switched with b .

We now give a precise definition of the setting in which we will apply our “role-switching” strategy.

Setting 3.6. *We are given a graph⁵ G on the vertex-set $[n]$, a bijection $\varphi : V(H) \rightarrow [n]$, and a φ -good set $A \subseteq V(H)$. We set $B = V(H) \setminus A$ and write $B = \{b_1, \dots, b_r\}$. We are also given an integer $m > 0$ and distinct vertices $a_{i,k} \in A$, where $1 \leq i \leq r$ and $1 \leq k \leq m$. Finally, we are given an integer⁶ d such that $d_H(x) \leq d$ for every $x \in \bigcup_{i=1}^r \{b_i, a_{i,1}, \dots, a_{i,m}\}$. We assume that the following two properties are satisfied.*

⁵We think of G as Builder's graph immediately after employing the strategy whose existence is guaranteed by Lemma 3.5 for the number of rounds specified in that lemma.

⁶The reason for allowing flexibility in the choice of d (as opposed to simply letting d be the maximum degree of H), is that in one application (namely, Theorem 1.7), we will be able to make sure that the degrees of the vertices $b_i, a_{i,1}, \dots, a_{i,m}$ ($1 \leq i \leq r$) are much smaller than $\Delta(H)$, which will be crucial for obtaining the desired bound.

1. There are no edges in H between $\{a_{i,k} : (i,k) \in [r] \times [m]\}$ and B .
2. The sets $\{a_{i,k}\} \cup N_H(a_{i,k})$ are pairwise-disjoint, where (i,k) run over all pairs in $[r] \times [m]$.

Throughout the second stage of his strategy, Builder maintains and updates sets $A_t \subseteq V(H)$ and bijections $\varphi_t : V(H) \rightarrow [n]$. Initially, $A_0 = A$ and $\varphi_0 = \varphi$. For every positive integer t , the pair (A_t, φ_t) will be defined immediately after round t of the second stage. We also set $B_t = V(H) \setminus A_t$ (so in particular, $B_0 = B$). Finally, we let G_t denote Builder's graph after exactly t rounds of the second stage (so in particular, $G_0 = G$). We are now ready to describe Builder's strategy for round t of the second stage (for any integer $t \geq 1$).

Strategy 3.7. Let $w_t \in [n]$ be the random vertex Builder is offered at round t of the second stage.

1. If some pair $(i,k) \in [r] \times [m]$ is such that $w_t \in \varphi(\{a_{i,k}\} \cup N_H(a_{i,k}))$ and $b_i \in B_{t-1}$, then do:
 - (a) If $w_t \in \varphi(N_H(a_{i,k}))$, then claim the edge $\{w_t, \varphi(b_i)\}$.
 - (b) If $w_t = \varphi(a_{i,k})$, then choose an arbitrary vertex $u \in \varphi_{t-1}(N_H(b_i) \cap A_{t-1})$ which is not adjacent to $\varphi(a_{i,k})$ in G_{t-1} , and claim the edge $\{\varphi(a_{i,k}), u\}$.
 - (c) Check whether $\varphi(b_i)$ is adjacent in G_t to every vertex of $\varphi(N_H(a_{i,k}))$ and, moreover, $\varphi(a_{i,k})$ is adjacent in G_t to every vertex of $\varphi_{t-1}(N_H(b_i) \cap A_{t-1})$. If so, then set $A_t = A_{t-1} \cup \{b_i\}$ (and hence $B_t = V(H) \setminus A_t = B_{t-1} \setminus \{b_i\}$), and

$$\varphi_t(x) = \begin{cases} \varphi_{t-1}(b_i) & x = a_{i,k}, \\ \varphi_{t-1}(a_{i,k}) & x = b_i, \\ \varphi_{t-1}(x) & x \in V(H) \setminus \{a_{i,k}, b_i\}. \end{cases}$$

- (d) Otherwise (i.e., if the condition in Item 1(c) does not hold), set $A_t = A_{t-1}$ and $\varphi_t = \varphi_{t-1}$.
2. Else (i.e., if there is no pair $(i,k) \in [r] \times [m]$ which satisfies the condition in Item 1), claim an arbitrary edge which is incident with w_t ; this edge will not be considered as part of Builder's graph in our analysis. Set $A_t = A_{t-1}$ and $\varphi_t = \varphi_{t-1}$.

The operation of defining A_t and φ_t as done in Item 1(c), is referred to as *switching* $a_{i,k}$ and b_i . This name stems from the fact that we swap the vertices which play the roles of b_i and $a_{i,k}$ in our current partial embedding φ_t of H into G_t . Evidently, switching $a_{i,k}$ and b_i does not change the role of any of the other $n - 2$ vertices. Switching $a_{i,k}$ and b_i is only done if, roughly speaking, the vertex currently playing the role of b_i can play the role of $a_{i,k}$, and the vertex currently playing the role of $a_{i,k}$ can play the role of b_i ; the exact condition for switching is stated in Item 1(c) above. Note that the only pairs of vertices which can be switched are of the form $(a_{i,k}, b_i)$ for some $1 \leq i \leq r$ and $1 \leq k \leq m$. In the following lemma we collect several simple facts regarding Strategy 3.7.

Lemma 3.8. Consider the execution of Strategy 3.7 for ℓ consecutive rounds, where ℓ is an arbitrary positive integer. Then the following statements hold.

1. If $a_{i,k}$ and b_i were switched in round t , then none of the vertices $b_i, a_{i,1}, \dots, a_{i,m}$ was switched at any other round.
2. If $x \in V(H)$ was not switched at any round, then $\varphi_s(x) = \varphi(x)$ for every $0 \leq s \leq \ell$. If x was switched in round t , then $\varphi_s(x) = \varphi(x)$ for every $0 \leq s \leq t - 1$ and $\varphi_s(x) = \varphi_t(x)$ for every $t \leq s \leq \ell$.
3. For every $x \in V(H) \setminus (\bigcup_{i=1}^r \{b_i, a_{i,1}, \dots, a_{i,m}\})$, we have $\varphi_s(x) = \varphi(x)$ for every $0 \leq s \leq \ell$.
4. Fix $1 \leq i \leq r$, and define $N_t := \varphi_{t-1}(N_H(b_i) \cap A_{t-1})$ for every $1 \leq t \leq \ell$. Then, $N_t \subseteq N_{t'}$ holds for every $1 \leq t < t' \leq \ell$.

Proof. We start with Item 1. Switching $a_{i,k}$ and b_i in round t forces $b_i \in A_t$. It then follows by the description of Strategy 3.7 that $b_i \in A_s$ for every $t < s \leq \ell$, making the condition in Item 1 of Strategy 3.7 false for b_i in each of the rounds $t+1, \dots, \ell$. Hence, if $a_{i,k}$ and b_i were switched in round t , then none of the vertices $b_i, a_{i,1}, \dots, a_{i,m}$ could have been switched in any subsequent round. Moreover, none of these vertices could have been switched in any round prior to round t as this would have made b_i ineligible for switching in round t . This proves Item 1. Item 2 can be easily proved by induction, using Item 1 and the definition of the functions $(\varphi_s : 0 \leq s \leq \ell)$ in Strategy 3.7. Item 3 follows from Item 2 and the fact that only vertices in $\bigcup_{i=1}^r \{b_i, a_{i,1}, \dots, a_{i,m}\}$ can be switched.

Let us prove Item 4. Fix $1 \leq i \leq r$ and $1 \leq t < t' \leq \ell$, and let $v \in N_t$ be an arbitrary vertex; we will prove that $v \in N_{t'}$. Set $x = \varphi_{t-1}^{-1}(v)$, and note that $x \in N_H(b_i) \cap A_{t-1} \subseteq N_H(b_i) \cap A_{t'-1}$. We will show that $\varphi_{t'-1}(x) = \varphi_{t-1}(x) = v$, which would imply that $v \in \varphi_{t'-1}(N_H(b_i) \cap A_{t'-1}) = N_{t'}$, as required. Assume first that $x \in A$. It then follows by Item 1 of Setting 3.6 that $x \notin \bigcup_{j=1}^r \{b_j, a_{j,1}, \dots, a_{j,m}\}$. By Item 3 of Lemma 3.8 we then have $\varphi_s(x) = \varphi(x)$ for every $0 \leq s \leq \ell$; in particular, $\varphi_{t'-1}(x) = \varphi_{t-1}(x) = v$, as claimed. Suppose now that $x \in B$, that is, $x = b_j$ for some $1 \leq j \neq i \leq r$. Since $b_j = x \in A_{t-1}$, the vertex b_j must have been switched with some $a_{j,k}$ prior to round t . Now, Item 2 of Lemma 3.8 implies that $\varphi_{t'-1}(x) = \varphi_{t-1}(x) = v$ in this case as well. $\square$

The following lemma can be thought of as a proof of the “correctness” of Strategy 3.7.

Lemma 3.9. *For every non-negative integer t , the set A_t is (G_t, φ_t) -good.*

Proof. The proof is by induction on t . The base case $t = 0$ is immediate from our assumption that A is φ -good (see Setting 3.6), and the fact that $A_0 = A$, $\varphi_0 = \varphi$ and $G_0 = G$. For the induction step, fix some $t \geq 1$ and suppose that the assertion of the lemma holds for $t-1$. Consider the execution of Strategy 3.7 in round t . If either the condition in Item 1 or the condition in Item 1(c) does not hold, then there is nothing to prove, since in that case $A_t = A_{t-1}$ and $\varphi_t = \varphi_{t-1}$. Suppose then that both of these conditions hold, and let $(i, k) \in [r] \times [m]$ be the pair satisfying the condition in Item 1 of Strategy 3.7. That is, we assume that $a_{i,k}$ and b_i were switched in round t .

We need to show that for every $x, y \in A_t$, if $\{x, y\} \in E(H)$, then $\{\varphi_t(x), \varphi_t(y)\} \in E(G_t)$. Hence, let $x, y \in A_t$ be such that $\{x, y\} \in E(H)$. Note that $\{x, y\} \neq \{a_{i,k}, b_i\}$, as $\{a_{i,k}, b_i\} \notin E(H)$ by Item 1 in Setting 3.6. If $x, y \notin \{a_{i,k}, b_i\}$ then we have $x, y \in A_{t-1}$, $\varphi_t(x) = \varphi_{t-1}(x)$ and $\varphi_t(y) = \varphi_{t-1}(y)$; so our assertion that $\{\varphi_t(x), \varphi_t(y)\} \in E(G_t)$ follows from the induction hypothesis for $t-1$. Therefore, without loss of generality, we may assume that $x \in \{a_{i,k}, b_i\}$ and $y \notin \{a_{i,k}, b_i\}$. This assumption implies that $\varphi_t(y) = \varphi_{t-1}(y)$ and that $y \in A_{t-1}$.

Suppose first that $x = a_{i,k}$. Since $\{x, y\} \in E(H)$, it follows that $y \in N_H(x) = N_H(a_{i,k})$. Now Items 1 and 2 of Setting 3.6 imply that $y \notin \bigcup_{i=1}^r \{b_i, a_{i,1}, \dots, a_{i,m}\}$, which in turn implies that $\varphi_t(y) = \varphi(y) \in \varphi(N_H(a_{i,k}))$, see Item 3 of Lemma 3.8. As for x , it follows from the definition of φ_t in Item 1(c) of Strategy 3.7 that $\varphi_t(x) = \varphi_t(a_{i,k}) = \varphi_{t-1}(b_i) = \varphi(b_i)$, where the last equality holds by Item 2 of Lemma 3.8. Since $a_{i,k}$ and b_i were switched in round t , it follows by Item 1(c) of Strategy 3.7 that $\varphi_t(x) = \varphi(b_i)$ is adjacent in G_t to all vertices of $\varphi(N_H(a_{i,k}))$. In particular, $\{\varphi_t(x), \varphi_t(y)\} \in E(G_t)$ as required.

Suppose now that $x = b_i$. Since $\{x, y\} \in E(H)$, it follows that $y \in N_H(b_i)$. Therefore, $\varphi_t(y) = \varphi_{t-1}(y) \in \varphi_{t-1}(N_H(b_i) \cap A_{t-1})$. Observe that $\varphi_t(x) = \varphi_t(b_i) = \varphi_{t-1}(a_{i,k}) = \varphi(a_{i,k})$, where the last equality holds by Item 2 of Lemma 3.8. Since $a_{i,k}$ and b_i were switched in round t , it follows by Item 1(c) of Strategy 3.7 that $\varphi(a_{i,k})$ is adjacent in G_t to all vertices of $\varphi_{t-1}(N_H(b_i) \cap A_{t-1})$. In particular, $\{\varphi_t(x), \varphi_t(y)\} \in E(G_t)$. This concludes the proof of the lemma. $\square$

In the following three lemmas, we consider the execution of Strategy 3.7 for ℓ rounds for some positive integer ℓ . For every $1 \leq i \leq r$ and $1 \leq k \leq m$, we denote by $\mathcal{A}_{i,k}$ the event: “ $\varphi(a_{i,k})$ was offered at least d times **after** each of the vertices in $\varphi(N_H(a_{i,k}))$ had already been offered”. In other words, $\mathcal{A}_{i,k}$ is the event that there are indices $1 \leq t_1 < \dots < t_q < s_1 < \dots < s_d \leq \ell$, where $q = |N_H(a_{i,k})|$, such that each element of $\varphi(N_H(a_{i,k}))$ was offered in one of the rounds $t_1, \dots, t_q$, and $\varphi(a_{i,k})$ was offered in each of the rounds $s_1, \dots, s_d$.

Lemma 3.10. *Let $1 \leq i \leq r$. If there exists some $1 \leq k \leq m$ for which $\mathcal{A}_{i,k}$ occurred, then $b_i \in A_\ell$.*

Proof. Suppose for a contradiction that $\mathcal{A}_{i,k}$ occurred for some $1 \leq i \leq r$ and $1 \leq k \leq m$, but $b_i \notin A_\ell$, i.e., $b_i \in B_\ell$. This means that b_i was not switched at any of the ℓ rounds for which we execute Strategy 3.7. Set $q = |N_H(a_{i,k})|$ and let $1 \leq t_1 < \dots < t_q < s_1 < \dots < s_d \leq \ell$ be the round numbers appearing in the definition of $\mathcal{A}_{i,k}$. Item 1(a) of Strategy 3.7 dictates that whenever a vertex from $\varphi(N_H(a_{i,k}))$ is sampled, Builder connects it to $\varphi(b_i)$. This implies that, for every $t_q \leq t \leq \ell$, every vertex of $\varphi(N_H(a_{i,k}))$ is adjacent in G_t to $\varphi(b_i)$.

As in Lemma 3.8, we let $N_t = \varphi_{t-1}(N_H(b_i) \cap A_{t-1})$ for each $1 \leq t \leq \ell$. Suppose first that there exists some $1 \leq j \leq d$ such that $\varphi(a_{i,k})$ is adjacent in G_{s_j} to every vertex of N_{s_j} . Then by Item 1(c) of Strategy 3.7, Builder would have switched $a_{i,k}$ and b_i in round s_j , which would contradict our assumption that b_i was never switched. Hence, for every $1 \leq j \leq d$ there exists a vertex of N_{s_j} which is not adjacent in G_{s_j} to $\varphi(a_{i,k})$. It follows by Item 1(b) of Strategy 3.7 that, in round s_j , Builder claims an edge $\{\varphi(a_{i,k}), u_j\}$ for some $u_j \in N_{s_j}$ which is not adjacent to $\varphi(a_{i,k})$ in G_{s_j-1} . Note that $u_1, \dots, u_d$ are distinct. It follows by Item 4 of Lemma 3.8 that $u_1, \dots, u_d \in N_{s_d}$. On the other hand, $|N_{s_d}| \leq |N_H(b_i)| \leq d$, implying that $N_{s_d} = \{u_1, \dots, u_d\}$. But this means that in the graph G_{s_d} , the vertex $\varphi(a_{i,k})$ is adjacent to every vertex of N_{s_d} , contrary to the above. $\square$

The following technical lemma provides lower bounds on the probability of the events $\mathcal{A}_{i,k}$.

Lemma 3.11. *Fix any $1 \leq i \leq r$ and $1 \leq k \leq m$.*

(a) *If $\ell \geq 2d$, then*

$$\mathbb{P}[\mathcal{A}_{i,k}] \geq \left(\frac{\ell^2}{12dn^2} \right)^d \cdot e^{-\frac{(d+1)(\ell-2d)}{n-d-1}}.$$

(b) *If $\ell \geq (\log(2d) + d + 3\sqrt{d})n$, then $\mathbb{P}[\mathcal{A}_{i,k}] \geq \frac{1}{4}$.*

Proof. We start with Item (a). Since $|N_H(a_{i,k})| \leq d$, it follows that

$$\mathbb{P}[\mathcal{A}_{i,k}] \geq \binom{\ell}{2d} \cdot d! \cdot \left(\frac{1}{n} \right)^{2d} \cdot \left(1 - \frac{d+1}{n} \right)^{\ell-2d} \geq \left(\frac{\ell}{2dn} \right)^{2d} \cdot \left(\frac{d}{e} \right)^d \cdot e^{-\frac{(d+1)(\ell-2d)}{n-d-1}} \geq \left(\frac{\ell^2}{12dn^2} \right)^d \cdot e^{-\frac{(d+1)(\ell-2d)}{n-d-1}},$$

where in the second inequality we used the estimates $1 - x \geq e^{-\frac{x}{1-x}}$ (which holds for every $0 < x < 1$) and $d! \geq \left(\frac{d}{e} \right)^d$ (which holds for every $d \geq 1$).

Next, we prove (b). Let $\mathcal{E}_k$ be the event that every vertex of $\varphi(N_H(a_{i,k}))$ was offered in the course of the first $\log(2d)n$ rounds, and let $\mathcal{F}_k$ be the event that $\varphi(a_{i,k})$ was offered at least d times in the course of the last $(d + 3\sqrt{d})n$ rounds. Since $\ell \geq (\log(2d) + d + 3\sqrt{d})n$ by assumption, the events $\mathcal{E}_k$ and $\mathcal{F}_k$ are independent. Note that $\mathcal{E}_k \cap \mathcal{F}_k \subseteq \mathcal{A}_{i,k}$; that is, if both $\mathcal{E}_k$ and $\mathcal{F}_k$ occur, then so does $\mathcal{A}_{i,k}$. Therefore, $\mathbb{P}[\mathcal{A}_{i,k}] \geq \mathbb{P}[\mathcal{E}_k \cap \mathcal{F}_k] = \mathbb{P}[\mathcal{E}_k] \cdot \mathbb{P}[\mathcal{F}_k]$. The probability that $\mathcal{E}_k$ did not occur is at most

$$d \cdot \left(1 - \frac{1}{n} \right)^{\log(2d)n} \leq d \cdot e^{-\log(2d)} = \frac{1}{2},$$

and the probability that $\mathcal{F}_k$ did not occur equals the probability that $\text{Bin}\left(\left(d + 3\sqrt{d}\right)n, 1/n\right)$ is smaller than d , which is at most

$$\mathbb{P}\left[\text{Bin}\left(\left(d + 3\sqrt{d}\right)n, 1/n\right) < d\right] \leq e^{-\frac{9d}{2(d+3\sqrt{d})}} \leq e^{-\frac{9d}{8d}} \leq \frac{1}{2}.$$

Here in the first inequality we used Lemma 2.1 with $\lambda = 3\sqrt{d}$. We thus conclude that $\mathbb{P}[\mathcal{A}_{i,k}] \geq 1/2 \cdot 1/2 = 1/4$ as claimed. $\square$

The following lemma forms the main result of Section 3.2, and plays a key role in the proofs of Theorems 1.3 and 1.7. Roughly speaking, this lemma states that if $\varphi : V(H) \rightarrow [n]$ is a bijection admitting a φ -good set that misses only a small fraction of $V(H)$, then by following Strategy 3.7 for a suitable (and not too large) number of rounds, Builder can obtain a bijection $\varphi' : V(H) \rightarrow [n]$ which admits a φ' -good set that misses significantly fewer vertices. The actual statement is somewhat technical. The proof of Lemma 3.12 utilizes Lemmas 3.9, 3.10 and 3.11, as well as some of the concentration inequalities from Section 2.

Lemma 3.12. *Let H be an n -vertex graph, let $\varphi : V(H) \rightarrow [n]$ be a bijection, let $V(H) = A \cup B$ be a partition, and consider a moment in the semi-random process at which A is φ -good with respect to Builder's graph. Write $B = \{b_1, \dots, b_r\}$. Let m be a positive integer, and let $\{a_{i,k} \in A : 1 \leq i \leq r \text{ and } 1 \leq k \leq m\}$ be vertices which satisfy Items 1 and 2 of Setting 3.6. Let $d \geq 100$ be such that $d_H(x) \leq d$ for every $x \in \bigcup_{i=1}^r \{b_i, a_{i,1}, \dots, a_{i,m}\}$. Let $\ell_1 = (\log(2d) + d + 3\sqrt{d})n$, $\ell_2 = \lceil n \cdot m^{-1/4d} \rceil$, $q_1 = \frac{1}{4}$, and $q_2 = d^{-2d}m^{-1/2}$. Fix any $j \in \{1, 2\}$, and suppose that $mq_j \geq 10^6d$. Let*

$$p = \begin{cases} o(1), & \frac{mq_j}{64d} \geq \log n, \\ e^{-\frac{\sqrt{n}}{O(d)}}, & \frac{mq_j}{64d} < \log n. \end{cases}$$

Suppose that Builder executes Strategy 3.7 for ℓ_j additional rounds. Then, with probability at least $1 - p$, after ℓ_j rounds of the process Builder's graph G will satisfy the following: there will be a bijection $\varphi' : V(H) \rightarrow V(G) = [n]$ and a partition $V(H) = A' \cup B'$ such that A' is (G, φ') -good, $A \subseteq A'$, and

$$|B'| \leq \begin{cases} 0, & \frac{mq_j}{64d} \geq \log n, \\ 5n \cdot e^{-\frac{mq_j}{256d}}, & \frac{mq_j}{64d} < \log n. \end{cases} \quad (1)$$

Proof. We will show that the assertion of the lemma holds with $\varphi' = \varphi_{\ell_j}$, $A' = A_{\ell_j}$, and $B' = B_{\ell_j}$. The fact that A_{ℓ_j} is φ_{ℓ_j} -good follows immediately from Lemma 3.9. It thus remains to prove that (1) holds for $B' = B_{\ell_j}$ with probability at least $1 - p$.

Fix arbitrary indices $1 \leq i \leq r$ and $1 \leq k \leq m$. For $j \in \{1, 2\}$, let us denote by $\mathbb{P}_j(\mathcal{A}_{i,k})$ the probability that $\mathcal{A}_{i,k}$ occurred in the course of the first ℓ_j rounds of the process. We claim that $\mathbb{P}_j[\mathcal{A}_{i,k}] \geq q_j$. The fact that $\mathbb{P}_1[\mathcal{A}_{i,k}] \geq \frac{1}{4} = q_1$ follows immediately from Item (b) of Lemma 3.11 and our choice of ℓ_1 . As for $j = 2$, recall that $\ell_2 = \lceil n \cdot m^{-1/4d} \rceil$, which implies that $\ell_2 \geq 2d$ holds for n which is sufficiently large with respect to d (since we trivially have $m \leq n$, and since $d \geq 100$, something like $n \geq (2d)^{1.01}$ would suffice). Therefore, Item (a) of Lemma 3.11 yields

$$\mathbb{P}_2[\mathcal{A}_{i,k}] \geq \left(\frac{\ell_2^2}{12dn^2} \right)^d \cdot e^{-\frac{(d+1)(\ell_2-2d)}{n-d-1}} \geq \left(\frac{1}{12d} \right)^d \cdot m^{-1/2} \cdot e^{-d-1} \geq d^{-2d}m^{-1/2} = q_2, \quad \text{where the last inequality holds for sufficiently large } d \text{ (our assumption that } d \geq 100 \text{ suffices).}$$

where the last inequality holds for sufficiently large d (our assumption that $d \geq 100$ suffices).

We have thus proved our assertion that $\mathbb{P}_j[\mathcal{A}_{i,k}] \geq q_j$ holds for every $j \in \{1, 2\}$. For the remainder of the proof, we fix an arbitrary $j \in \{1, 2\}$ and suppose (as in the statement of the lemma) that $mq_j \geq 10^6d$. For convenience, we put $\ell := \ell_j$ and $q := q_j$. For every $1 \leq i \leq r$, let X_i be the random variable counting the number of indices $1 \leq k \leq m$ for which $\mathcal{A}_{i,k}$ occurred in the course of the first ℓ rounds. It follows by linearity of expectation that $\mathbb{E}[X_i] \geq mq$.

Now, consider the sequence $(w_1, \dots, w_\ell)$ of random vertices offered to Builder, and observe that changing any one coordinate in this sequence can change the value of X_i by at most 1 (here we use Item 2 in Setting 3.6). Furthermore, for every s , if $X_i \geq s$, then there is a set of at most $2ds$ coordinates in the sequence $(w_1, \dots, w_\ell)$ which certify that $X_i \geq s$ (indeed, each event $\mathcal{A}_{i,k}$ that occurred is certified by a set of at most $2d$ coordinates). It thus follows by Lemma 2.3 with $c = 1$, $g = 2d$, and $\lambda = \frac{\mathbb{E}[X_i]}{2}$, that

$$\mathbb{P} \left[X_i < \frac{\mathbb{E}[X_i]}{2} - 60\sqrt{2d\mathbb{E}[X_i]} \right] \leq 4e^{-\frac{(\mathbb{E}[X_i]/2)^2}{16d\mathbb{E}[X_i]}} = 4e^{-\frac{\mathbb{E}[X_i]}{64d}} \leq 4e^{-\frac{mq}{64d}}.$$

Therefore, with probability at least $1 - 4e^{-\frac{mq}{64d}}$, it holds that

$$X_i \geq \frac{\mathbb{E}[X_i]}{2} - 60\sqrt{2d\mathbb{E}[X_i]} = \sqrt{\mathbb{E}[X_i]} \cdot \left(\frac{\sqrt{\mathbb{E}[X_i]}}{2} - 60\sqrt{2d} \right) \geq \sqrt{\mathbb{E}[X_i]} \cdot \left(\frac{\sqrt{mq}}{2} - 60\sqrt{2d} \right) > 0,$$

where the last inequality follows from our assumption that $mq \geq 10^6d$.

Now let $\mathcal{I}$ be the set of all $1 \leq i \leq r$ such that $X_i = 0$. It follows by Lemma 3.10 that if some $1 \leq i \leq r$ satisfies $X_i > 0$, then $b_i \notin B_\ell$. Hence, we have $B_\ell \subseteq \{b_i : i \in \mathcal{I}\}$. So to complete the proof it is enough to show that the bounds in (1) hold for the set $\mathcal{I}$. We have seen that $\mathbb{P}[i \in \mathcal{I}] \leq 4e^{-\frac{mq}{64d}}$ holds for every $1 \leq i \leq r$. Hence,

$$\mathbb{E}[|\mathcal{I}|] \leq r \cdot 4e^{-\frac{mq}{64d}}.$$

Suppose first that $\frac{mq}{64d} \geq \log n$, and note that we have $r \leq \frac{n}{m} \leq \frac{n}{\log n}$. It follows that

$$\mathbb{E}[|\mathcal{I}|] \leq r \cdot 4e^{-\frac{mq}{64d}} \leq \frac{n}{\log n} \cdot \frac{4}{n} = o(1).$$

So by Markov's inequality, we have $|\mathcal{I}| = 0$ w.h.p., as required.

Suppose now that $\frac{mq}{64d} < \log n$. Observe that changing any one coordinate in the sequence $(w_1, \dots, w_\ell)$ of random vertices, can change the value of $|\mathcal{I}|$ by at most 1. Hence, it follows by Lemma 2.2 with $c = 1$ and $\lambda = n \cdot e^{-\frac{mq}{256d}} \geq n^{3/4}$, that

$$\mathbb{P}[|\mathcal{I}| \geq \mathbb{E}[|\mathcal{I}|] + \lambda] \leq e^{-\frac{\lambda^2}{2\ell}} \leq e^{-\frac{n^{3/2}}{2\ell}} \leq e^{-\frac{\sqrt{n}}{O(d)}},$$

where the last inequality holds since $\ell_1, \ell_2 = O(dn)$. We conclude that with probability at least $1 - e^{-\frac{\sqrt{n}}{O(d)}}$, we have $|\mathcal{I}| \leq \mathbb{E}[|\mathcal{I}|] + \lambda \leq (4r + n) \cdot e^{-\frac{mq}{256d}} \leq 5n \cdot e^{-\frac{mq}{256d}}$, as required. $\square$

3.3 Putting it All Together

In this section we iterate Lemma 3.12 to prove Lemma 3.14, from which Theorem 1.3 then easily follows. Lemma 3.14 roughly states that given a bijection $\varphi : V(H) \rightarrow [n]$ which admits a φ -good set covering almost all of $V(H)$, Builder can use φ as a basis for constructing a copy of H in his graph, and he can achieve this objective w.h.p. fairly quickly. In the proof we will need the following simple claim, which asserts that we can satisfy the conditions listed in Setting 3.6 with a relatively large choice of m .

Claim 3.13. *Let H be an n -vertex D -degenerate graph of maximum degree Δ . Let $A \cup B$ be a partition of $V(H)$ and suppose that $|A| \geq 4\Delta|B|$. Set $r := |B|$ and*

$$m := \left\lceil \frac{|A|}{8\Delta^2|B|} \right\rceil.$$

Then, there exist vertices $(a_{i,k} : 1 \leq i \leq r \text{ and } 1 \leq k \leq m)$ such that $a_{i,k} \in A$ and $d_H(a_{i,k}) \leq 2D$ for every $(i, k) \in [r] \times [m]$, and such that Conditions 1 and 2 of Setting 3.6 are satisfied.

Proof. Let $A' := \{a \in A : |N_H(a) \cap A| \leq 2D \text{ and } N_H(a) \cap B = \emptyset\}$. Note that every $a \in A'$ satisfies $d_H(a) = |N_H(a) \cap A| \leq 2D$. Since H is D -degenerate, there are less than $|A|/2$ vertices $a \in A$ which satisfy $|N_H(a) \cap A| > 2D$. As $\Delta(H) = \Delta$, there are at most $\Delta|B| \leq |A|/4$ vertices $a \in A$ which have a neighbour in B . Altogether, we get $|A'| \geq |A|/4$. Using again the assumption $\Delta(H) = \Delta$, we infer that there exists an integer

$$M \geq \frac{|A'|}{\Delta^2 + 1} \geq \frac{|A|}{8\Delta^2}$$

and vertices $a_1, \dots, a_M \in A'$ such that $\text{dist}_H(a_i, a_j) \geq 3$ for every $1 \leq i < j \leq M$. Recalling our choice of m , we index (a subset of) the vertices $a_1, \dots, a_M$ by pairs $(i, k) \in [r] \times [m]$. For every $(i, k) \in [r] \times [m]$, let $a_{i,k}$ be the vertex in $\{a_1, \dots, a_M\}$ which is indexed by the pair (i, k) . Now Condition 1 of Setting 3.6 is satisfied due to our choice of A' , and Condition 2 of Setting 3.6 is satisfied because $\text{dist}_H(a_{i,k}, a_{i',k'}) \geq 3$ for every choice of distinct pairs $(i, k), (i', k') \in [r] \times [m]$. $\square$

Lemma 3.14. *Let H be an n -vertex D -degenerate graph of maximum degree Δ . Set $d = \min\{2D, \Delta\}$. Suppose that at some point in the semi-random process, there is a bijection $\varphi : V(H) \rightarrow [n]$ and a partition $V(H) = A \cup B$ such that A is φ -good with respect to Builder's graph; such that $|B| \leq 10^{-8}\Delta^{-5}n$; and such that $d_H(b) \leq d$ for every $b \in B$. Then Builder has a strategy guaranteeing that w.h.p., after $(d + o(d))n$ additional rounds his graph will contain a copy of H .*

Proof. We may and will assume that d is large enough, say $d \geq 100$. Builder's strategy consists of two phases, which correspond to the two cases (i.e. $j = 1$ and $j = 2$) in Lemma 3.12.

Phase 1: Let

$$m_0 := \left\lceil \frac{n}{8\Delta^2|B|} \right\rceil.$$

Find vertices $(a_{i,k} : 1 \leq i \leq |B| \text{ and } 1 \leq k \leq m_0)$ such that $a_{i,k} \in A$ and $d_H(a_{i,k}) \leq d$ for every $(i,k) \in [r] \times [m_0]$, and such that Conditions 1 and 2 of Setting 3.6 are satisfied (with $m = m_0$). Apply the strategy whose existence is guaranteed by Lemma 3.12 with parameter $j = 1$ and with $m = m_0$. Lemma 3.12 (with $j = 1$) ensures that after $\ell_1 = (\log(2d) + d + 3\sqrt{d})n = (d + o(d))n$ rounds, there will be a bijection $\varphi_0 : V(H) \rightarrow [n]$ and a partition $V(H) = A_0 \cup B_0$ such that A_0 is φ_0 -good with respect to Builder's graph, $A \subseteq A_0$ (and hence $B_0 \subseteq B$), and

$$|B_0| \leq \begin{cases} 0, & \frac{m_0}{256d} \geq \log n, \\ 5n \cdot e^{-\frac{m_0}{1024d}}, & \frac{m_0}{256d} < \log n. \end{cases} \quad (2)$$

If $B_0 = \emptyset$ then $A_0 = V(H)$ is φ_0 -good, implying that Builder has successfully embedded H into his graph, and so Builder is done. Otherwise, proceed to Phase 2.

Phase 2: Define a sequence of bijections $\varphi_1, \varphi_2, \dots$ from $V(H)$ to $[n]$, and a sequence of partitions $A_1 \cup B_1, A_2 \cup B_2, \dots$ of $V(H)$, by performing the following steps for every integer $t \geq 1$ for which $B_{t-1} \neq \emptyset$.

- (a) Find vertices $(a_{i,k} : 1 \leq i \leq |B_{t-1}| \text{ and } 1 \leq k \leq m_t := \lceil \frac{n}{8\Delta^2|B_{t-1}|} \rceil)$ such that $a_{i,k} \in A_{t-1}$ and $d_H(a_{i,k}) \leq d$ for every $(i,k) \in [r] \times [m_t]$, and such that Conditions 1 and 2 of Setting 3.6 are satisfied for the partition $A_{t-1} \cup B_{t-1}$ with $m = m_t$.
- (b) Invoke the strategy whose existence is guaranteed by Lemma 3.12 with $j = 2$, with $m = m_t$, and with input φ_{t-1} and $A_{t-1} \cup B_{t-1}$. Using this strategy, Builder obtains a bijection $\varphi_t : V(H) \rightarrow [n]$ and a partition $A_t \cup B_t$ of $V(H)$ such that A_t is φ_t -good with respect to Builder's graph, $A_{t-1} \subseteq A_t$ (and hence $B_t \subseteq B_{t-1}$), and

$$|B_t| \leq \begin{cases} 0, & m'_t \geq \log n, \\ 5n \cdot e^{-m'_t/4}, & m'_t < \log n, \end{cases} \quad (3)$$

where

$$m'_t := \frac{m_t \cdot q_2}{64d} = \frac{\sqrt{m_t}}{64d^{2d+1}},$$

and q_2 is as in Lemma 3.12. (Here and later on we slightly abuse notation by hiding the fact that $q_2 = d^{-2d}/\sqrt{m_t}$ depends on t .)

Having described Builder's strategy, we now turn to prove that w.h.p. Builder can follow it. First, note that by Claim 3.13, there exist vertices $(a_{i,k} : 1 \leq i \leq |B| \text{ and } 1 \leq k \leq m_0)$ such that $a_{i,k} \in A$ and $d_H(a_{i,k}) \leq d$ for every $(i,k) \in [r] \times [m_0]$, and such that Conditions 1 and 2 of Setting 3.6 hold. Moreover, the conditions required for the application of Lemma 3.12 with $j = 1$ are satisfied as $m_0 q_1 = \frac{m_0}{4} \geq \frac{n}{32\Delta^2|B|} > 10^6 \Delta \geq 10^6 d$, where the second inequality follows from our assumption that $|B| \leq 10^{-8}\Delta^{-5}n$. This shows that Builder can follow Phase 1 of his strategy. It remains to show that w.h.p. Builder can follow Phase 2 of his strategy.

Similarly to Phase 1, the existence of the desired vertices $(a_{i,k} \in A_{t-1} : 1 \leq i \leq |B_{t-1}| \text{ and } 1 \leq k \leq m_t)$ for every given integer $t \geq 1$ follows from Claim 3.13 with input $A_{t-1} \cup B_{t-1}$. It remains to prove that the conditions of Lemma 3.12 are met whenever **Builder** wishes to apply it (with $j = 2$). The fact that A_{t-1} is φ_{t-1} -good for every positive integer t is guaranteed by the previous applications of Lemma 3.12. The fact that $d_H(b) \leq d$ for every $b \in B_{t-1}$ follows from our assumption that the same holds for every $b \in B$, and the fact that $B_0 \subseteq B$ and $B_i \subseteq B_{i-1}$ for every $i \geq 1$. We now show that $\sqrt{m_t} \cdot d^{-2d} = m_t \cdot q_2 \geq 10^6 d$ holds for every integer $t \geq 1$ for which $B_{t-1} \neq \emptyset$. To this end, first note that

$$|B_0| \leq 5n \cdot e^{-\frac{m_0}{1024d}} \leq 5n \cdot \exp\left(-\frac{n}{2^{13}\Delta^3|B|}\right) e \leq n \cdot e^{-\Delta^2}, \quad (4)$$

where the first inequality follows from (2), and the third from our assumption that $|B| \leq 10^{-8}\Delta^{-5}n$. Since $B_i \subseteq B_{i-1}$ for every $i \geq 1$, we have $|B_i| \leq n \cdot e^{-\Delta^2}$ for each $i \geq 0$. Now we obtain

$$m_t \cdot q_2 = \sqrt{m_t} \cdot d^{-2d} \geq \frac{\sqrt{n}}{3\Delta^{2\Delta+1}\sqrt{|B_{t-1}|}} \geq \frac{e^{\Delta^2/2}}{3\Delta^{2\Delta+1}} \geq 10^6 \Delta \geq 10^6 d, \quad (5)$$

where the penultimate inequality holds for sufficiently large Δ (say, $\Delta \geq 100$). This shows that we can indeed apply Lemma 3.12 with $j = 2$ and with input $A_{t-1} \cup B_{t-1}$ for every integer $t \geq 1$ for which $B_{t-1} \neq \emptyset$. We conclude that **Builder** can follow Phase 2 of his strategy.

Before moving on to prove the correctness of **Builder's** strategy, we first prove the following claim.

Claim 3.15. *Let $\zeta = e^{-\Delta^2}$ and, for every non-negative integer t , let $\beta_t = |B_t|/n$. Suppose either that $t = 0$ or that $t \geq 1$ and $m'_t < \log n$. Then $\beta_t \leq \zeta^{t+1}$.*

Proof. Our proof proceeds by induction on t . The base case $t = 0$ follows from (4). Let then $t \geq 1$ and suppose that $m'_t < \log n$. Observe that the sequence m'_s is monotone non-decreasing in s (this follows from the fact that $B_i \subseteq B_{i-1}$ for every $i \geq 1$). So either $t - 1 = 0$, or $m'_{t-1} < \log n$. In either case we can apply the induction hypothesis to get $\beta_{t-1} \leq \zeta^t$. Note that

$$m'_t = \frac{\sqrt{m_t}}{64d^{2d+1}} \geq \frac{\sqrt{n}}{200\Delta^{2\Delta+2}\sqrt{|B_{t-1}|}}. \quad (6)$$

Now we get

$$\begin{aligned} \beta_t &\leq 5 \cdot e^{-m'_t/4} \leq 5 \cdot \exp\left(-\frac{1}{800\Delta^{2\Delta+2}\sqrt{\beta_{t-1}}}\right) \leq 5 \cdot \exp\left(-\frac{1}{800\Delta^{2\Delta+2}\zeta^{t/2}}\right) \\ &= 5 \cdot \exp\left(-\frac{e^{\Delta^2 t/2}}{800\Delta^{2\Delta+2}}\right) \leq e^{-\Delta^2(t+1)} = \zeta^{t+1}, \end{aligned}$$

where the first inequality holds by (3), the second inequality holds by (6), the third inequality holds by the induction hypothesis for $t - 1$, and the last inequality holds for every $t \geq 1$, provided that Δ is larger than some suitable absolute constant (again, $\Delta \geq 100$ suffices). This proves the claim. $\square$

Returning to the proof of the lemma, we now prove the correctness of **Builder's** strategy. For the time being, we will assume that all applications of Lemma 3.12 throughout **Builder's** strategy are successful; later we will show that w.h.p. this is indeed the case. It follows from (6) and from Claim 3.15 that $m'_t \geq e^{\Delta^2 t/2} \cdot \frac{1}{200\Delta^{2\Delta+2}}$ holds for every $t \geq 1$. Hence, $m'_t \geq \log n$ must hold for some $t \leq \log \log n$ (and in fact much earlier, but we will not need this). Now, if $m'_t \geq \log n$ then by (3) we have $B_t = \emptyset$, which in turn implies that **Builder** has successfully embedded H into the graph he is constructing.

Next, we estimate the probability that **Builder's** strategy fails. Recall that Lemma 3.12 is only applied once with parameter $j = 1$, and that this application is w.h.p. successful. Let us now consider the applications

of Lemma 3.12 with $j = 2$ (in Phase 2). As previously noted, there is at most one such application with $m'_t \geq \log n$, and at most $\log \log n$ such applications with $m'_t < \log n$. The failure probability of the former application is $o(1)$, and the failure probability of each of the latter applications is at most $e^{-\frac{\sqrt{n}}{O(d)}}$. We thus conclude that w.h.p. all of the above applications of Lemma 3.12 are successful, as required. This completes the proof of correctness of Builder's strategy.

It remains to estimate the overall number of rounds required for implementing Builder's strategy. Recall that the sole application of Lemma 3.12 with $j = 1$ requires $(d + o(d))n$ rounds. It thus remains to bound from above the number of rounds required for Phase 2 of Builder's strategy. To this end, let t^* denote the smallest integer t satisfying $m'_t \geq \log n$, and note that $t^* \leq \log \log n$. Then in Phase 2, Lemma 3.12 was invoked at most t^* times. Moreover, for each $1 \leq t \leq t^*$, invoking Lemma 3.12 with input $A_{t-1} \cup B_{t-1}$ (and with $j = 2$) required at most

$$\lceil n \cdot m_t^{-\frac{1}{4d}} \rceil \leq n \cdot m_t^{-\frac{1}{4d}} + 1 \leq n \cdot \left(\frac{n}{8\Delta^2 |B_{t-1}|} \right)^{-\frac{1}{4d}} + 1 \leq O(n) \cdot \left(\frac{1}{\beta_{t-1}} \right)^{-\frac{1}{4d}} + 1 \leq O(n) \cdot e^{-\Delta t/4} + 1$$

rounds, where in the last inequality we used Claim 3.15. Therefore, the overall number of rounds required for the (at most) t^* applications of Lemma 3.12 in Phase 2 is no more than

$$\sum_{t=1}^{t^*} \left(O(n) \cdot e^{-\Delta t/4} + 1 \right) \leq O(n) \cdot \sum_{t=1}^{\infty} e^{-\Delta t/4} + \log \log n = O(n).$$

We conclude that the overall number of rounds required for implementing Builder's strategy is at most $(d + o(d))n$, thus completing the proof of Lemma 3.14. $\square$

Equipped with Lemma 3.14, we can finally prove Theorem 1.3.

Proof of Theorem 1.3. Let n , Δ , and H be as in the statement of the theorem. Due to Proposition 3.4, we only need to handle the case $\Delta = O(\log n)$. Set $\alpha = 10^{-8}\Delta^{-5}$. Builder's strategy for embedding H is as follows. Fix an arbitrary bijection $\varphi : V(H) \rightarrow [n]$. In the first $(\Delta/2 + o(\Delta))n$ rounds, Builder invokes the strategy whose existence is guaranteed by Lemma 3.5, and thus obtains w.h.p. a φ -good set $A \subseteq V(H)$ of size at least $(1 - \alpha)n$. Setting $B := V(H) \setminus A$, observe that the requirements of Lemma 3.14 are satisfied for $D := \Delta$. Builder now applies the strategy given by Lemma 3.14 for an additional $(\Delta + o(\Delta))n$ rounds, and by doing so successfully constructs a copy of H in his graph w.h.p. The overall number of rounds is then $(3\Delta/2 + o(\Delta))n$, as required. $\square$

4 Constructing Spanning Forests

In this section we prove Theorem 1.7 and Proposition 1.8. We start with the following simple lemma, whose proof demonstrates a strategy for greedily embedding an almost-spanning forest.

Lemma 4.1. *Let n be a positive integer and let $\alpha \in (0, 1)$ be such that $n \gg \alpha^{-2} \log(1/\alpha)$. Let T' be a forest on $(1 - \alpha)n$ vertices and let $\ell = \log(2/\alpha) \cdot n$. Then, in the semi-random process on n vertices, Builder has a strategy which w.h.p. allows him to construct a copy of T' within ℓ rounds.*

Proof. Assume without loss of generality that T' is a tree (otherwise simply replace T' with a tree containing it). Let $t = (1 - \alpha)n$ and let $v_1, \dots, v_t$ be an ordering of the vertices of T' such that $T'[\{v_1, \dots, v_i\}]$ is a tree for every $1 \leq i \leq t$. Throughout the process, Builder maintains a partial function φ which is initially empty. For every positive integer i , let w_i denote the vertex Builder is offered in the i th round. In the first round, Builder connects w_1 to an arbitrary vertex u ; he then sets $\varphi(v_1) = w_1$ and $\varphi(v_2) = u$. For every $i \geq 2$, Builder plays the i th round as follows. Let r denote the largest integer for which $\varphi(v_r)$ has already been defined. If $w_i \notin \{\varphi(v_1), \dots, \varphi(v_r)\}$, then Builder connects w_i to $\varphi(v_j)$, where $j \leq r$ is the unique integer for which

$\{v_{r+1}, v_j\} \in E(T')$; he then sets $\varphi(v_{r+1}) = w_i$. Otherwise, **Builder** claims an arbitrary edge incident with w_i , which he does not consider to be part of the tree he is building (alternatively, **Builder** skips this round).

It is evident that, by following the proposed strategy, **Builder**'s graph contains a copy of T' as soon as t different vertices are offered. Hence, it suffices to prove that w.h.p. at least t different vertices are offered during the first ℓ rounds.⁷ For every $1 \leq j \leq n$, let I_j be the indicator random variable for the event: "vertex j was not offered during the first ℓ rounds of the process". Let $X = \sum_{j=1}^n I_j$; then

$$\mathbb{E}(X) = \sum_{j=1}^n \mathbb{E}(I_j) = n(1 - 1/n)^\ell \leq n \cdot e^{-\ell/n} = \alpha n/2.$$

Observe that changing any one coordinate in the sequence $(w_1, \dots, w_\ell)$ of random vertices, can change the value of X by at most 1. Hence, it follows by Lemma 2.2 with $c = 1$ and $\lambda = \alpha n/2$, that

$$\mathbb{P}[X \geq \alpha n] \leq \mathbb{P}[X \geq \mathbb{E}(X) + \alpha n/2] \leq e^{-\frac{(\alpha n/2)^2}{2\ell}} = e^{-\frac{\alpha^2 n^2}{8 \log(2/\alpha)n}} = o(1),$$

where the last equality holds by our assumption that $n \gg \alpha^{-2} \log(1/\alpha)$. $\square$

We are now in a position to prove Theorem 1.7.

Proof of Theorem 1.7. Let n , Δ , and T be as in the statement of the theorem. Assume first that $\Delta \geq n^{1/11}$. In this case **Builder** employs the strategy presented in the proof of Lemma 4.1. It is easy to see that as soon as each of the n vertices has been offered, **Builder**'s graph contains a copy of T . It is well-known (and easy to prove) that this will happen w.h.p. in $(1 + o(1))n \log n = O(n \log \Delta)$ rounds. For the remainder of the proof we thus assume that $\Delta < n^{1/11}$.

Set $\alpha = 10^{-8} \Delta^{-5}$, and note that $n \gg \alpha^{-2} \log(1/\alpha)$, as required by Lemma 4.1. Since T is a forest, there are at least $n/2$ vertices of T whose degree is at most 2. Fix a set B of αn such vertices. Set $A := V(H) \setminus B$ and $T' := T[A]$. At the first stage of his strategy, **Builder** invokes Lemma 4.1, which enables him to construct a copy of T' w.h.p. in $\log(2/\alpha)n = O(n \log \Delta)$ rounds.

Let $\varphi : V(T) \rightarrow [n]$ be a bijection such that $\varphi|_A$ is an embedding of T' into **Builder**'s graph. Note that the conditions of Lemma 3.14 are satisfied with $D = 1$, since forests are 1-degenerate, and since we made sure that all vertices in B have degree at most 2. In the second stage of his strategy, **Builder** applies (the strategy given by) Lemma 3.14 in order to construct a copy of T in his graph w.h.p. This requires $O(n)$ additional rounds. The total number of rounds is thus $O(n \log \Delta)$, as required. $\square$

4.1 A Lower Bound : Proof of Proposition 1.8

Proof of Proposition 1.8. Let Δ and $n \geq n_0(\Delta)$ be as in the statement of the proposition. Since **Builder** clearly needs at least $n - 1$ rounds in order to build a tree on n vertices, we can assume that Δ is a sufficiently large constant. We prove the proposition for the n -vertex forest T consisting of $\lfloor \frac{n}{\Delta+1} \rfloor$ pairwise-disjoint $(\Delta+1)$ -vertex stars, together with some (at most Δ) isolated vertices (if needed). Let us denote the center of the i th star by u_i , and its leaves by $x_{i,1}, \dots, x_{i,\Delta}$ ($1 \leq i \leq \lfloor \frac{n}{\Delta+1} \rfloor$).

Let $w_1, w_2, \dots$ denote the sequence of random vertices offered to **Builder, and let $m = 0.1n \log \Delta$. Suppose that **Builder** did manage to build a copy of T within m rounds, and let $\varphi : V \rightarrow [n]$ be a bijection such that $\{\varphi(u), \varphi(v)\}$ is an edge in **Builder**'s graph for every $\{u, v\} \in E(T)$. It is then evident that, for every $1 \leq i \leq \lfloor \frac{n}{\Delta+1} \rfloor$, either $\varphi(u_i)$ appears at least $\sqrt{\Delta}$ times in $(w_1, w_2, \dots, w_m)$, or at least $\Delta - \sqrt{\Delta}$ of the elements of the set $\{\varphi(x_{i,j}) : 1 \leq j \leq \Delta\}$ appear at least once in $(w_1, w_2, \dots, w_m)$. A straightforward calculation then shows that either at least $\frac{1}{\sqrt{\Delta}} \cdot \lfloor \frac{n}{\Delta+1} \rfloor \geq \frac{n}{2\Delta^{3/2}}$ of the vertices in $\{\varphi(u_i) : 1 \leq i \leq \lfloor \frac{n}{\Delta+1} \rfloor\}$ were**

⁷This can be restated as saying that in the coupon collector's problem, $\ell = \log(2/\alpha)n$ rounds suffice in order to collect $t = (1 - \alpha)n$ different coupons w.h.p.

offered at least $\sqrt{\Delta}$ times each, or all but at most

$$\Delta + \Delta \cdot \frac{1}{\sqrt{\Delta}} \cdot \left\lfloor \frac{n}{\Delta+1} \right\rfloor + \left\lfloor \frac{n}{\Delta+1} \right\rfloor \cdot (\sqrt{\Delta} + 1) \leq \frac{3n}{\sqrt{\Delta}}$$

of the n vertices were offered at least once. In the inequality above we use the assumption that $n \geq n_0(\Delta)$ for some suitable $n_0(\Delta)$. So in order to prove that w.h.p. **Builder** needs more than m rounds to build T , it suffices to show that w.h.p. there are more than $3n/\sqrt{\Delta}$ vertices $1 \leq i \leq n$ that do not appear in $(w_1, w_2, \dots, w_m)$, and less than $\frac{n}{2\Delta^{3/2}}$ vertices $1 \leq i \leq n$ which appear in $(w_1, w_2, \dots, w_m)$ at least $\sqrt{\Delta}$ times.

Let X be the random variable which counts the number of vertices $1 \leq i \leq n$ that do not appear in $(w_1, w_2, \dots, w_m)$. Our goal is to show that w.h.p. $X > 3n/\sqrt{\Delta}$. We have

$$\mathbb{E}(X) = n(1 - 1/n)^m \geq n \cdot e^{-m/(n-1)} = n \cdot e^{-0.1 \log \Delta \cdot \frac{n}{n-1}} = n \cdot \Delta^{-0.1 - o(1)} > 4n/\sqrt{\Delta},$$

where the first inequality follows from the fact that $(1 - 1/n)^{n-1} \geq 1/e$ for each $n \geq 1$, and the last inequality holds for sufficiently large n and Δ . We will use Lemma 2.2 to prove that w.h.p. X is not much smaller than its expected value. Observe that changing any single coordinate in the sequence of random vertices $(w_1, w_2, \dots, w_m)$ can change the value of X by at most 1. Therefore, applying Lemma 2.2 with parameters $c = 1$ and $\lambda = n/\sqrt{\Delta}$ yields

$$\mathbb{P}[X \leq 3n/\sqrt{\Delta}] \leq \mathbb{P}[X \leq \mathbb{E}(X) - n/\sqrt{\Delta}] \leq e^{-\frac{(n/\sqrt{\Delta})^2}{2m}} \leq e^{-\frac{(n/\sqrt{\Delta})^2}{n \log \Delta}} = e^{-\frac{n}{\Delta \log \Delta}} = o(1),$$

where the last equality holds since n is sufficiently large with respect to Δ .

Let Z be the random variable which counts the number of vertices $1 \leq i \leq n$ that appear at least $\sqrt{\Delta}$ times in $(w_1, w_2, \dots, w_m)$. Our goal is to prove that w.h.p. $Z < \frac{n}{2\Delta^{3/2}}$. For every $1 \leq i \leq n$, let Z_i be the random variable counting the number of times i appears in $(w_1, w_2, \dots, w_m)$. Then $Z_i \sim \text{Bin}(m, 1/n)$, implying that $\mathbb{E}[Z_i] = m/n \leq 0.1 \log \Delta$. Applying Lemma 2.1 with parameter $\lambda = 2 \log \Delta \leq \sqrt{\Delta} - 0.1 \log \Delta$, we obtain

$$\mathbb{P}[Z_i \geq \sqrt{\Delta}] \leq \mathbb{P}[Z_i \geq \mathbb{E}[Z_i] + \lambda] \leq \exp\left(-\frac{(2 \log \Delta)^2}{2(\mathbb{E}[Z_i] + \frac{2}{3} \log \Delta)}\right) \leq e^{-2 \log \Delta} \leq \frac{1}{4\Delta^{3/2}}, \quad (7)$$

where the last inequality holds for sufficiently large Δ . For every $1 \leq i \leq n$, let I_i be the indicator random variable for the event $Z_i \geq \sqrt{\Delta}$; note that $Z = \sum_{i=1}^n I_i$. It follows by (7) and by the linearity of expectation that $\mathbb{E}(Z) \leq \frac{n}{4\Delta^{3/2}}$. Since changing any single coordinate in the sequence of random vertices $(w_1, w_2, \dots, w_m)$ can change the value of Z by at most 1, applying Lemma 2.2 with parameters $c = 1$ and $\lambda = \frac{n}{4\Delta^{3/2}}$ yields

$$\mathbb{P}\left[Z \geq \frac{n}{2\Delta^{3/2}}\right] \leq \mathbb{P}\left[Z \geq \mathbb{E}(Z) + \frac{n}{4\Delta^{3/2}}\right] \leq e^{-\frac{n^2}{32\Delta^3 m}} \leq e^{-\frac{n^2}{32\Delta^3 n \log \Delta}} = o(1),$$

where the equality holds since n is sufficiently large with respect to Δ . □

5 Non-Adaptive Strategies

In this section we prove Theorems 1.10, 1.11 and 1.12.

Proof of Theorem 1.10. Let $\mathcal{L} = \{L^w : w \in [n]\}$ be a family of lists as in the definition of a non-adaptive strategy. Recall that for each $w \in [n]$, the list L^w is a permutation of $[n] \setminus \{w\}$. Our goal is to show that the strategy corresponding to $\mathcal{L}$ requires w.h.p. at least $\Omega(n\sqrt{\log n})$ rounds to make all n vertices non-isolated.

Set $t = n\sqrt{\log n}/4$, and let $w_1, \dots, w_t$ be the first t random vertices **Builder** is offered. For every $v \in [n]$, let t_v denote the number of appearances of v in the sequence $(w_1, \dots, w_t)$. Let $U = \{v \in [n] : t_v > \sqrt{\log n}/2\}$ and let $W = \bigcup_{v \in U} \{L^v(i) : 1 \leq i \leq t_v\}$. Our main observation (which follows immediately from the definitions of U and W) is that a vertex $u \in [n]$ will be left isolated after t rounds, if all of the following conditions hold:

- (1) u does not appear in $(w_1, \dots, w_t)$;
- (2) none of the vertices v , for which u is included among the first $\sqrt{\log n}/2$ elements of L^v , appear in $(w_1, \dots, w_t)$;
- (3) $u \notin W$.

So in order to complete the proof of the theorem, it remains to prove that w.h.p. there exists a vertex $u \in [n]$ which satisfies Conditions (1), (2), and (3) above. To this end, we will use a two-round exposure argument. Let Z denote the set of vertices which do not appear in $(w_1, \dots, w_t)$; clearly $Z \cap U = \emptyset$. In the following claim we collect some simple facts regarding the sets U, Z and the integers $(t_v : v \in [n])$.

Claim 5.1. *The following hold w.h.p.*

- (a) $|U| \leq e^{-\Omega(\sqrt{\log n})}n$.
- (b) $t_v < \log n$ for every $v \in [n]$.
- (c) $|Z| \geq (1 - o(1))e^{-\sqrt{\log n}/4}n$.

Proof. We start with Item (a). Recall that for a given vertex $v \in [n]$, we have $t_v \sim \text{Bin}(t, \frac{1}{n})$; hence, $\mathbb{E}[t_v] = \sqrt{\log n}/4$. Now, by Lemma 2.1 with $\lambda = \sqrt{\log n}/4$, we have

$$\mathbb{P}[v \in U] = \mathbb{P}[t_v > \sqrt{\log n}/2] = \mathbb{P}[t_v > \mathbb{E}[t_v] + \sqrt{\log n}/4] \leq e^{-\frac{\log n}{O(\sqrt{\log n})}} = e^{-\Omega(\sqrt{\log n})}.$$

It thus follows by Markov's inequality that w.h.p. $|U| \leq e^{-\Omega(\sqrt{\log n})}n$.

We now prove Item (b). Observe that for every $v \in [n]$ we have

$$\mathbb{P}[t_v \geq \log n] \leq \binom{t}{\log n} \left(\frac{1}{n}\right)^{\log n} \leq \left(\frac{et}{n \log n}\right)^{\log n} \leq \left(\frac{1}{\sqrt{\log n}}\right)^{\log n} = o(1/n).$$

A union bound over $[n]$ then shows that w.h.p. $t_v < \log n$ for every $v \in [n]$.

Finally, we prove Item (c). For each $v \in [n]$, the probability that $v \in Z$ is $(1 - 1/n)^t = (1 - o(1))e^{-\sqrt{\log n}/4}$. Therefore, $\mathbb{E}[|Z|] = (1 - o(1))e^{-\sqrt{\log n}/4}n$. To show that $|Z|$ is concentrated around its expected value, observe that changing any single coordinate in the sequence $(w_1, \dots, w_t)$ of random vertices, can change the value of $|Z|$ by at most 1. Hence, by Lemma 2.2 with $c = 1$ and (say) $\lambda = n^{2/3}$, we have

$$\mathbb{P}\left[|Z| \leq \mathbb{E}[|Z|] - n^{2/3}\right] \leq e^{-\frac{n^{4/3}}{2t}} = e^{-\frac{n^{4/3}}{O(n\sqrt{\log n})}} = o(1).$$

We conclude that w.h.p. $|Z| \geq (1 - o(1))e^{-\sqrt{\log n}/4}n$. □

From now on we condition on the events stated in Items (a)-(c) of Claim 5.1 (which hold w.h.p. by that claim). Items (a) and (b) imply that $|W| \leq |U| \log n \leq \log n \cdot e^{-\Omega(\sqrt{\log n})}n = o(n)$. Observe that conditioning on their sizes, U, Z are uniformly distributed among all pairs of disjoint subsets of $[n]$ of the corresponding sizes. From this point on we condition on U , which in turn determines W .

For each $u \in [n]$, let A_u be the set of all vertices $v \in [n] \setminus \{u\}$ such that u is included among the first $\sqrt{\log n}/2$ elements of L^v . Let V_0 be the set of all $u \in [n]$ satisfying $|A_u| \leq \sqrt{\log n}$. Since the union (as a multiset) of the first $\sqrt{\log n}/2$ elements in all lists has size $n\sqrt{\log n}/2$ altogether, we deduce that $|V_0| \geq n/2$; hence $|V_0 \setminus W| \geq (1/2 - o(1))n$.

Now expose Z , conditioning on its size. Observe that if $u \in [n] \setminus W$ is such that $A_u \cup \{u\} \subseteq Z$, then u satisfies Conditions (1), (2) and (3). So from now on our goal is to show that w.h.p. there exists a vertex $u \in [n] \setminus W$ for which $A_u \cup \{u\} \subseteq Z$. Since each $u \in [n]$ belongs to at most $\sqrt{\log n}/2 + 1$ of the sets $\{A_v \cup \{v\} : v \in [n]\}$, and since $|A_u \cup \{u\}| \leq \sqrt{\log n} + 1$ for each $u \in V_0$, one can find a collection $B_1, \dots, B_{s'}$ of

$$s' \geq \left\lfloor \frac{|V_0 \setminus W|}{(\sqrt{\log n}/2 + 1)(\sqrt{\log n} + 1) + 1} \right\rfloor = \Omega\left(\frac{n}{\log n}\right)$$

pairwise-disjoint sets among the sets $\{A_u \cup \{u\} : u \in V_0 \setminus W\}$. Since $B_1, \dots, B_{s'}$ are pairwise-disjoint, at least $s := s' - |U| \geq \Omega(n/\log n) - e^{-\Omega(\sqrt{\log n})}n = \Omega(n/\log n)$ of these sets are contained in $[n] \setminus U$. So suppose, without loss of generality, that $B_1, \dots, B_s \subseteq [n] \setminus U$, and let us show that w.h.p., there is $1 \leq i \leq s$ such that $B_i \subseteq Z$. To this end, we will couple Z with a binomial random set of slightly smaller size. Recall that we are conditioning on the size of Z and on the event $|Z| \geq (1 - o(1))e^{-\sqrt{\log n}/4}n$, which occurs w.h.p. by Claim 5.1. We denote $z = |Z|$, recalling that (under this conditioning), Z is distributed uniformly among all subsets of $[n] \setminus U$ of size z . We generate Z by performing the following experiment: set $p = \frac{z}{2n}$, and let R be a random subset of $[n] \setminus U$, obtained by independently including each element of $[n] \setminus U$ with probability p . If $|R| \leq z$, then we uniformly choose a set $Z' \subseteq [n] \setminus U$ of size z which contains R . It is easy to see that, conditioned on $|R| \leq z$, the set Z' is distributed uniformly among all subsets of $[n] \setminus U$ of size z . Hence (conditioned on $|R| \leq z$), Z' has the same distribution as Z (conditioned on $|Z| = z$). Note that $|R|$ is stochastically dominated by $\text{Bin}(n, \frac{z}{2n})$, so by Lemma 2.1 with $\lambda = \frac{z}{2}$ we have

$$\mathbb{P}[|R| > z] \leq \mathbb{P}\left[|R| \geq \mathbb{E}[|R|] + \frac{z}{2}\right] \leq e^{-\frac{(z/2)^2}{2(\mathbb{E}[|R|] + z/6)}} \leq e^{-\Omega(z)} = o(1).$$

Since the sets $B_1, \dots, B_s$ are pairwise-disjoint and of size at most $\sqrt{\log n} + 1$ each, the probability that R contains none of these sets is at most

$$\begin{aligned} \left(1 - p^{\sqrt{\log n} + 1}\right)^s &\leq \exp\left(-p^{\sqrt{\log n} + 1} \cdot s\right) = \exp\left(-\left(\frac{z}{2n}\right)^{\sqrt{\log n} + 1} \cdot s\right) \\ &\leq \exp\left(-e^{-\log n/2} \cdot s\right) = e^{-s/\sqrt{n}} = e^{-\Omega(\sqrt{n}/\log n)} = o(1), \end{aligned}$$

where in the second inequality we used the assumption that $z \geq (1 - o(1))e^{-\sqrt{\log n}/4}n$. We conclude that w.h.p. there will be some $1 \leq i \leq s$ such that $B_i \subseteq Z$, as required. $\square$

Proof of Theorem 1.11. Setting $k = n/\sqrt{\log n}$, partition $[n]$ into sets $V_1, \dots, V_k$, each of size either $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. For each $1 \leq i \leq k$ and $v \in V_i$, Builder sets the adjacency list L^v so that first appear all the vertices of $V_i \setminus \{v\}$, then all the vertices of V_{i+1} (where $i+1$ is taken modulo k), and finally all other vertices. In each of the three “segments”, the inner order among the vertices is arbitrary.

Now, set $t = 8n\sqrt{\log n}$, and let $(w_1, \dots, w_t)$ be the first t random vertices Builder is offered. Let W be the set of vertices appearing at most $\lceil n/k \rceil$ times in $(w_1, \dots, w_t)$. Let $\mathcal{A}$ be the event that $|V_i \cap W| \leq |V_i|/2 - 1$ for each $1 \leq i \leq k$. We will show that $\mathcal{A}$ happens w.h.p., and that if $\mathcal{A}$ happens then after t rounds, Builder’s graph contains a Hamilton cycle.

We start by estimating $\mathbb{P}[\mathcal{A}]$. Fixing $1 \leq i \leq k$, note that if $|V_i \cap W| \geq |V_i|/2$ then there is a set $U \subseteq V_i$ of size $|U| = |V_i|/2 = (1 + o(1))\sqrt{\log n}/2$, such that $X := |\{1 \leq j \leq t : w_j \in U\}| \leq |U| \cdot \lceil n/k \rceil \leq \log n$. Note that X has the distribution $\text{Bin}(t, |U|/n)$. By Lemma 2.1 with $\lambda = (3 + o(1))\log n$, and by our choice of t , we get

$$\mathbb{P}[X \leq \log n] \leq \mathbb{P}\left[\text{Bin}\left(8n\sqrt{\log n}, \frac{(1 + o(1))\sqrt{\log n}}{2n}\right) \leq \log n\right] \leq \exp\left(-\frac{(1 + o(1))9\log^2 n}{8\log n}\right) = o(n^{-1.1}).$$

By taking the union bound over all $k = o(n)$ indices $1 \leq i \leq k$, and over all at most $2^{|V_i|} = 2^{(1+o(1))\sqrt{\log n}} = n^{o(1)}$ choices of $U \subseteq V_i$, we obtain $\mathbb{P}[\mathcal{A}] = 1 - o(1)$.

Suppose now that $\mathcal{A}$ happened. Then for each $1 \leq i \leq k$, each of the at least $|V_i|/2 + 1$ vertices $v \in V_i \setminus W$ has been connected to all vertices in V_i , and to at least 2 vertices in V_{i+1} (indeed, this is due to our choice of the lists L_v , and the definition of $\mathcal{A}$). It follows that the minimum degree inside V_i is at least $|V_i|/2 + 1$ (for each $1 \leq i \leq k$), and that we can choose distinct vertices $x_i, y_i \in V_i$ such that y_i is connected to x_{i+1} for each $1 \leq i \leq k$ (with indices taken modulo k).

Recall that a graph is called *Hamilton-connected* if for each pair of distinct vertices u, v , there is a Hamilton path whose endpoints are u and v . It follows from a classical result of Ore [21] that any m -vertex graph with minimum degree at least $\frac{m+1}{2}$ is Hamilton-connected. Let G denote Builder’s graph immediately after t

rounds of the process. By the above result of Ore, $G[V_i]$ is Hamilton-connected for each $1 \leq i \leq k$. So fix, for each $1 \leq i \leq k$, a Hamilton path P_i in $G[V_i]$, whose endpoints are x_i and y_i . Now it is easy to see that $P_1, \{y_1, x_2\}, P_2, \{y_2, x_3\}, \dots, \{y_{k-1}, x_k\}, P_k, \{y_k, x_1\}$ is a Hamilton cycle in G , as required. $\square$

Proof of Theorem 1.12. The proof is somewhat similar to the proof of Theorem 1.11, and so we only give a rough sketch. Partition $[n]$ into $k = n/\sqrt{\log n}$ parts $V_1, \dots, V_k$ whose sizes are all divisible by r and are as close to each other as possible. Let $s_{\min} = \min\{|V_i| : 1 \leq i \leq k\}$ and let $s_{\max} = \max\{|V_i| : 1 \leq i \leq k\}$; observe that $s_{\min}, s_{\max} = (1 \pm o(1))\sqrt{\log n}$. Then, for every $1 \leq i \leq k$ and every $v \in V_i$, Builder sets the adjacency list L^v so that first appear all the vertices of $V_i \setminus \{v\}$ (in an arbitrary order), and then all other vertices (in an arbitrary order).

Now, set $t = Cn\sqrt{\log n}$ (where $C = C(r)$ will be chosen later), and let $(w_1, \dots, w_t)$ be the first t random vertices offered to Builder. Let W be the set of vertices appearing at most $s_{\max} - 2$ times in $(w_1, \dots, w_t)$. Observe that if $|W \cap V_i| < s_{\min}/r$, then the resulting induced subgraph of Builder $G[V_i]$ has minimum degree at least $(1 - 1/r)|V_i|$ and thus admits a K_r -factor by the Hajnal-Szemerédi Theorem [10]. If this happens for every $1 \leq i \leq k$, then the union over $1 \leq i \leq k$ of these K_r -factors obviously forms a K_r -factor of G . It remains to prove that w.h.p. $|W \cap V_i| < s_{\min}/r$ holds for every $1 \leq i \leq k$. Fix some $1 \leq i \leq k$. Then

$$\begin{aligned} \mathbb{P}[|W \cap V_i| \geq s_{\min}/r] &\leq \binom{s_{\max}}{s_{\min}/r} \cdot \mathbb{P}[\text{Bin}(t, s_{\min}/(rn)) \leq (s_{\max} - 2) \cdot s_{\min}/r] \\ &\leq (3r)^{s_{\min}/r} \cdot \exp\left\{-\frac{rn}{2ts_{\min}} \cdot \frac{C^2 \cdot s_{\min}^4}{4r^2}\right\} \\ &= (3r)^{s_{\min}/r} \cdot \exp\left\{-\frac{C \cdot s_{\min}^3}{8r\sqrt{\log n}}\right\} \\ &\leq (3r)^{\sqrt{\log n}} \cdot \exp\left\{-\frac{C \cdot (1 - o(1)) \log n}{8r}\right\} \\ &\leq \exp\left\{-\frac{C \log n}{9r}\right\} = o(1/k), \end{aligned}$$

where the second inequality holds by Lemma 2.1 with $\lambda = \frac{C}{2} \cdot \frac{s_{\min}^2}{r}$, and the equality holds if, say, $C = 9r$. We also assumed throughout that n is large enough with respect to r . A union bound over all $1 \leq i \leq k$ then shows that w.h.p. $|W \cap V_i| < s_{\min}/r$ holds for every $1 \leq i \leq k$, as required. $\square$

References

- [1] Dimitris Achlioptas, Raissa M. D’Souza, and Joel Spencer. Explosive percolation in random networks. *Science*, 323(5920):1453–1455, 2009.
- [2] Yossi Azar, Andrei Z. Broder, Anna R. Karlin, and Eli Upfal. Balanced allocations. *SIAM Journal on Computing*, 29(1):180–200, 1999.
- [3] Omri Ben-Eliezer, Dan Hefetz, Gal Kronenberg, Olaf Parczyk, Clara Shikhelman, and Miloš Stojaković. Semi-random graph process. *Random Structures and Algorithms*, 56(3):648–675, 2020.
- [4] Tom Bohman and Alan M. Frieze. Avoiding a giant component. *Random Structures and Algorithms*, 19(1):75–85, 2001.
- [5] Tom Bohman and David Kravitz. Creating a giant component. *Combinatorics, Probability and Computing*, 15:489–511, 2006.
- [6] Allan Borodin and Ran El-Yaniv. *Online Computation and Competitive Analysis*. Cambridge University Press, New York, NY, USA, 1998.

- [7] Asaf Ferber, Gal Kronenberg, and Kyle Luh. Optimal threshold for a random graph to be 2-universal. *Transactions of the American Mathematical Society*, 372:4239–4262, 2019.
- [8] Michael J. Freedman, Kobbi Nissim, and Benny Pinkas. Efficient private matching and set intersection. In *Advances in Cryptology - EUROCRYPT 2004*, pages 1–19, 2004.
- [9] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2016.
- [10] András Hajnal and Endre Szemerédi. Proof of a conjecture of P. Erdős. In *Combinatorial theory and its applications II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970.
- [11] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs*. Wiley, 2000.
- [12] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures and Algorithms*, 33(1):1–28, 2008.
- [13] Jeff Kahn, Eyal Lubetzky, and Nicholas Wormald. The threshold for combs in random graphs. *Random Structures and Algorithms*, 48(4):794–802, 2016.
- [14] Mihyun Kang, Will Perkins, and Joel Spencer. The Bohman-Frieze process near criticality. *Random Structures and Algorithms*, 43(2):221–250, 2013.
- [15] Michael Krivelevich, Po-Shen Loh, and Benny Sudakov. Avoiding small subgraphs in Achlioptas processes. *Random Structures and Algorithms*, 34:165–195, 2009.
- [16] Michael Krivelevich, Eyal Lubetzky, and Benny Sudakov. Hamiltonicity thresholds in Achlioptas processes. *Random Structures and Algorithms*, 37:1–24, 2010.
- [17] Michael Molloy and Bruce Reed. *Graph Colouring and the Probabilistic Method*. Springer, 2002.
- [18] Richard Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.
- [19] Tobias Müller and Reto Spöhel. A geometric Achlioptas process. *The Annals of Applied Probability*, 25(6):3295–3337, 2015.
- [20] Torsten Mütze, Reto Spöhel, and Henning Thomas. Small subgraphs in random graphs and the power of multiple choices. *J. Combin. Theory Ser. B*, 101(4):237–268, 2011.
- [21] Oystein Ore. Hamilton connected graphs. *Journal de Mathématiques Pures et Appliquées*, 42:21–27, 1963.
- [22] Rasmus Pagh and Flemming Friche Rodler. Cuckoo hashing. *Journal of Algorithms*, 51(2):122–144, 2004.
- [23] Oliver Riordan and Lutz Warnke. Achlioptas process phase transitions are continuous. *The Annals of Applied Probability*, 22(4):1450–1464, 2012.
- [24] Joel Spencer and Nicholas Wormald. Birth control for giants. *Combinatorica*, 27:587–628, 2007.
- [25] Sasu Tarkoma, Christian Esteve Rothenberg, and Eemil Lagerspetz. Theory and practice of Bloom filters for distributed systems. *IEEE Communications Surveys and Tutorials*, 14(1):131–155, 2012.