# CLIQUES, CHROMATIC NUMBER, AND INDEPENDENT SETS IN THE SEMI-RANDOM PROCESS

DAVID GAMARNIK, MIHYUN KANG, AND PAWEŁ PRAŁAT

Abstract. The semi-random graph process is a single player game in which the player is initially presented an empty graph on $n$ vertices. In each round, a vertex $u$ is presented to the player independently and uniformly at random. The player then adaptively selects a vertex $v$, and adds the edge $uv$ to the graph. For a fixed monotone graph property, the objective of the player is to force the graph to satisfy this property with high probability in as few rounds as possible. In this paper, we investigate the following three properties: containing a complete graph of order $k$, having the chromatic number at least $k$, and not having an independent set of size at least $k$.

## 1. Introduction and Main Results

1.1. **Definitions.** In this paper, we consider the **semi-random graph process** suggested by Peleg Michaeli, introduced formally in [5], and studied recently in [3, 4, 8, 12, 14, 15, 16, 23, 11, 24] that can be viewed as a "one player game". The process starts from $G_0$, the empty graph on the vertex set $[n] := \{1, \dots, n\}$ where $n \in \mathbb{N}$. In each **round** $t \in \mathbb{N}$, a vertex $u_t$ is chosen uniformly at random from $[n]$. Then, the player (who is aware of graph $G_t$ and vertex $u_t$) must select a vertex $v_t$ and add the edge $u_t v_t$ to $G_t$ to form $G_{t+1}$. The goal of the player is to build a (multi)graph satisfying a given property $\mathcal{P}$ as quickly as possible. It is convenient to refer to $u_t$ as a **square**, and $v_t$ as a **circle** so every edge in $G_t$ joins a square with a circle. We say that $v_t$ is paired to $u_t$ in step $t$. Moreover, we say that vertex $x \in [n]$ is **covered** by the square $u_t$ arriving at round $t$, provided $u_t = x$. The analogous definition extends to the circle $v_t$. Equivalently, we may view $G_t$ as a directed graph where each arc directs from $u_t$ to $v_t$, and thus we may use $(u_t, v_t)$ to denote the edge added in step $t$. For this paper, it is easier to consider squares and circles for counting arguments.

A **strategy** $\mathcal{S}$ is defined by specifying for each $n \geq 1$, a sequence of functions $(f_t)_{t=1}^{\infty}$, where for each $t \in \mathbb{N}$, $f_t(u_1, v_1, \dots, u_{t-1}, v_{t-1}, u_t)$ is a distribution over $[n]$ which depends on the vertex $u_t$, and the history of the process up until step $t-1$. Then, $v_t$ is chosen according to this distribution. If $f_t$ is an atomic distribution, then $v_t$ is determined by $u_1, v_1, \dots, u_{t-1}, v_{t-1}, u_t$. We then denote $(G_i^{\mathcal{S}}(n))_{i=0}^{t}$ as the sequence of random (multi)graphs obtained by following the strategy $\mathcal{S}$ for $t$ rounds; where we shorten $G_t^{\mathcal{S}}(n)$ to $G_t$ or $G_t(n)$ when clear.

1.2. **Notation.** Results presented in this paper are asymptotic by nature. We say that some property $\mathcal{P}$ holds **asymptotically almost surely** (or **a.a.s.**) if the probability that the semi-random process has this property (after possibly applying some given strategy) tends to 1 as $n$ goes to infinity. Given two functions $f = f(n)$ and $g = g(n)$, we will write $f(n) = \mathcal{O}(g(n))$ if there exists an absolute constant $c \in \mathbb{R}_+$ such that $|f(n)| \leq c|g(n)|$ for all $n$, $f(n) = \Omega(g(n))$ if $g(n) = \mathcal{O}(f(n))$, $f(n) = \Theta(g(n))$ if $f(n) = \mathcal{O}(g(n))$ and $f(n) = \Omega(g(n))$, and we write $f(n) = o(g(n))$ or $f(n) \ll g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. In addition, we write $f(n) \gg g(n)$ if $g(n) = o(f(n))$ and we write $f(n) \sim g(n)$ if $f(n) = (1 + o(1))g(n)$, that is, $\lim_{n\to\infty} f(n)/g(n) = 1$.

We will use $\log n$ to denote a natural logarithm of $n$. As mentioned earlier, for a given $n \in \mathbb{N} := \{1, 2, \dots\}$, we will use $[n]$ to denote the set consisting of the first $n$ natural numbers, that is, $[n] := \{1, 2, \dots, n\}$. Finally, as typical in the field of random graphs, for expressions that clearly have to be an integer, we round up or down but do not specify which: the choice of which does not affect the argument.

1.3. **Main Results—Complete Graphs.** In this paper, we investigate three monotone properties. The first one is the property of containing $K_k$, a complete graph of order $k$. In the very first paper on the semi-random process [5], it was proved that a.a.s. one may construct a complete graph of a constant order $k$ once there are vertices with at least $k$ squares on them. On the other hand, if no vertex receives at least $k$ squares, it is impossible to achieve it a.a.s. Specifically, the following result was proved.

**Observation 1.1** ([5]). *Fix an integer $k \geq 3$ and any function $\omega = \omega(n)$ that tends to infinity as $n \to \infty$. Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $K_k$ at time $t = \omega n^{(k-2)/(k-1)}$.*
*(b) There is no strategy that a.a.s. creates $K_k$ at time $t = n^{(k-2)/(k-1)}/\omega$.*

In fact, part (a) of the above observation was proved for a larger family of graphs that are $k-1$ ($\geq 2$) degenerate (see Section 4 for the definition of degeneracy). Moreover, it was conjectured that part (b) can be generalized to such large family of graphs. The conjecture was proved recently in [3]. As a result, creating graphs of a constant size is well-understood—essentially, creating a fixed graph with degeneracy $d$ is possible once the process lasts long enough so that there are vertices with at least $d-1$ squares.

On the other hand, constructing complete graphs of order $k \gg \log n$ is very simple and can be done in almost optimal way. It follows immediately from Chernoff's bound (see (3) and (4)), together with the union bound over all vertices, that if $t \gg n \log n$, then a.a.s. all vertices receive

$$\frac{t}{n}\left(1+\mathcal{O}\left(\sqrt{\frac{\log n}{t/n}}\right)\right) \sim \frac{t}{n}$$

squares. One may try to create a complete graph on the vertex set $[k]$ for $k = 2\ell+1$ ($\ell \in \mathbb{N}$) by connecting the $j$th square ($j \in [\ell]$) landing on vertex $i$ with vertex $(i-1+j) \pmod{2\ell+1}+1$. This simple algorithm yields a lower bound for the size of the complete graph. To get an upper bound, we simply observe that it is impossible to create $K_{k'}$ if the maximum degree is smaller than $(k'-1)/2$. After combining the two observations, we get the following.

**Observation 1.2.** *Suppose that $t = t(n) \geq \omega n \log n$, where $\omega = \omega(n)$ is any function that tends to infinity as $n \to \infty$. Let*

$$\begin{aligned} k &= k(n) := \frac{2t}{n}\left(1-\omega^{-1/3}\right) \sim \frac{2t}{n} \\ k' &= k'(n) := \frac{2t}{n}\left(1+\omega^{-1/3}\right) \sim \frac{2t}{n} \sim k. \end{aligned}$$

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $K_{\min\{k,n\}}$ at time $t$.*
*(b) There is no strategy that a.a.s. creates $K_{k'}$ at time $t$.*

In fact, much stronger property holds. Let $H$ be an $n$-vertex graph of maximum degree $\Delta = \Delta(n) \gg \log n$. In [4] it was proved that there exists a strategy to build $H$ in $(\Delta n/2)(1+o(1))$ rounds.

In light of Observations 1.1 and 1.2, it remains to investigate how large complete graphs one can build in $t$ rounds, provided that $t = t(n) = n^{1+o(1)}$ and $t = \mathcal{O}(n \log n)$. Recall that $f(n) = o(1)$ only means that $\lim_{n\to\infty} f(n) = 0$ and so it might be negative. In particular, we allow $t \ll n$ as long as $t = n^{1+o(1)}$. If $t = o(n \log n)$, then one may create a complete graph of size which is asymptotic to the maximum number of squares on a single vertex—see Lemma 3.1(b). More importantly, asymptotically, this is the best one can do. This is our first main result which we state below.

**Theorem 1.3.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t \ll n \log n$. Let $\beta = \beta(n) := n \log n/t$. (In particular, $\beta \to \infty$, as $n \to \infty$.) Define*

$$\ell \;=\; \ell(n) \;:=\; \frac{\log n}{\log \beta - 2 \log \log \beta} \;\sim\; \frac{\log n}{\log \beta},$$

*and*

$$\epsilon \;=\; \epsilon(n) \;:=\; \begin{cases} e^2(\log \beta)/\beta & \text{if} \quad \beta \le \log n/\log\log n, \\ 15(\log \ell)/\ell & \text{if} \quad \log n/\log\log n < \beta \le \log^2 n, \\ e/\ell & \text{if} \quad \beta > \log^2 n. \end{cases}$$

*(In particular, $\epsilon = o(1)$, regardless of $\beta$.) Finally, let*

$$k \;=\; k(n) \;:=\; \frac{\log n - 2\log\log n - t/n}{\log \beta} \;\sim\; \frac{\log n}{\log \beta}$$

$$k' \;=\; k'(n) \;:=\; \ell(1 + 4\epsilon^{1/4}) \;\sim\; \frac{\log n}{\log \beta} \;\sim\; k.$$

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $K_k$ at time $t$.*
*(b) There is no strategy that a.a.s. creates $K_{k'}$ at time $t$.*

Unfortunately, when $t = \Theta(n \log n)$, then our bounds do not asymptotically match but they are at most a multiplicative factor of $2+o(1)$ away from each other, as we will show later (see Figure 2). Suppose that $t = t(n) = \gamma n \log n$ for some $\gamma \in (0, \infty)$. We will derive an asymptotic lower bound of $k_1 = \ell = \xi\gamma \log n$, where constant $\xi = \xi(\gamma) \in (1, \infty)$ is defined to be the unique solution to the following equation

$$1 - \xi\gamma(\log \xi - 1) - \gamma = 0, \tag{1}$$

which is equivalent to

$$\xi(\log \xi - 1) = \frac{1-\gamma}{\gamma} = \frac{1}{\gamma} - 1 \in (-1, \infty) \tag{2}$$

or to

$$\xi\gamma = \frac{1-\gamma}{\log \xi - 1}.$$

(The left hand side of (2) is a bijection from $(0, \infty)$ to $(-1, \infty)$ which proves the uniqueness of $\xi$.) As we will see in Lemma 3.1(c) below, $\ell$ defined in Theorem 1.3 is asymptotic to the maximum number of squares on a single vertex.

It is easy to see that $\xi$ is a decreasing function of $\gamma$. If $\gamma \to 0$, then $\xi \sim (1/\gamma)/\log(1/\gamma)$, which is consistent with Theorem 1.3 (applied with $\beta = 1/\gamma \to \infty$). If $\gamma = 1$, then $\xi = e$. More importantly, if $\gamma = \gamma_\ell = (2 \log 2 - 1)^{-1} \approx 2.59$, then $\xi = 2$. Finally, if $\gamma \to \infty$, then $\xi \to 1$. The constant $\gamma_\ell$ will play a special role in the lower bound in the statement of our result. We will show another asymptotic lower bound of $k_2 \sim 2\gamma \log n$ which is stronger than the previous one, provided that $\gamma > \gamma_\ell$ (see Figure 1, right side).

We will also show two upper bounds. The first one, $k_2' \sim 2\ell \sim 2\xi\gamma \log n$ is trivial but is best possible when $\gamma \to \infty$ (recall that $\xi \to 1$ as $\gamma \to \infty$). The second one, $k_1' \sim (1 + 2\sqrt{2}(e/\xi)^{1/4})\ell$, is stronger provided that $\gamma < \gamma_u = (64e \log 64 - 1)^{-1} \approx 0.00139$ (see Figure 1, left side). Indeed, if $\gamma < \gamma_u$, then $\xi > 64e$ and, as a consequence, $1 + 2\sqrt{2}(e/\xi)^{1/4} < 2$. This bound is best possible when $\gamma \to 0$.

**Theorem 1.4.** *Suppose that $t = t(n) = \gamma n \log n$ for some $\gamma \in (0, \infty)$. Let $\xi = \xi(\gamma) \in (1, \infty)$ be defined as in (1). Define*

$$\ell \;=\; \ell(n) \;:=\; \frac{1-\gamma}{\log \xi - 1} \log n \;=\; \xi\gamma \log n.$$

*Let*

$$\begin{aligned}
k_1 &= k_1(n) := \frac{(1-\gamma)\log n - 2\log\log n}{\log \xi - 1} \sim \ell = \xi\gamma\log n,\\
k_2 &= k_2(n) := 2\gamma\log n - 4\sqrt{\gamma\log n\log\log n} \sim 2\gamma\log n,\\
k_1' &= k_1'(n) := \left(1+\frac{3}{\log^{1/2} n}\right)\left(1+2\sqrt{2}(e/\xi)^{1/4}\right)\xi\gamma\log n \sim \left(1+2\sqrt{2}(e/\xi)^{1/4}\right)\xi\gamma\log n,\\
k_2' &= k_2'(n) := 2\xi\gamma\log n + 1 \sim 2\xi\gamma\log n.
\end{aligned}$$

*Finally, let*

$$\begin{aligned}
k &= k(n) := \max\{k_1, k_2\} \sim \max\{\xi, 2\}\gamma\log n,\\
k' &= k'(n) := \min\{k_1', k_2'\} \sim \min\{1+2\sqrt{2}(e/\xi)^{1/4}, 2\}\xi\gamma\log n.
\end{aligned}$$

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $K_k$ at time $t$.*

*(b) There is no strategy that a.a.s. creates $K_{k'}$ at time $t$.*

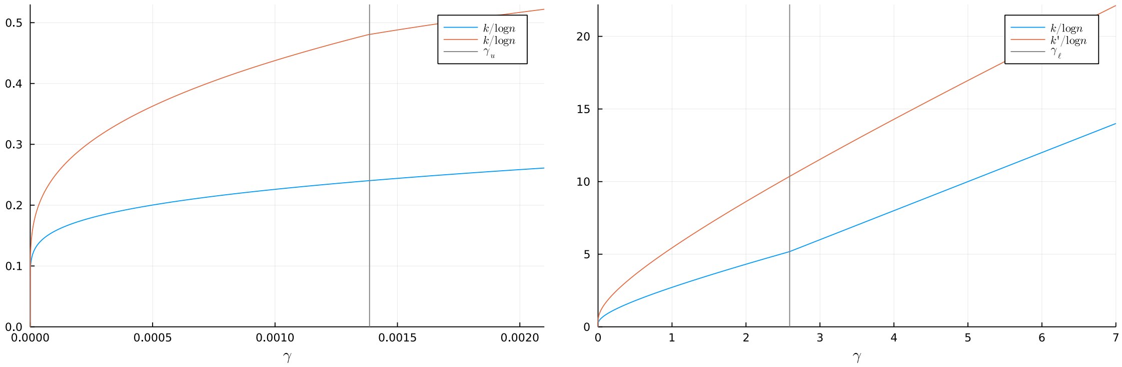

FIGURE 1. The upper ($k'$) and the lower ($k$) bound for the order of a largest complete graph: small (left figure) and large (right figure) values of $\gamma$.

1.4. **Main Results—Chromatic Number.** A **proper colouring** of a graph is a labeling of its vertices with colours such that no two vertices sharing the same edge have the same colour. The smallest number of colours in a proper colouring of a graph $G = (V, E)$ is called its **chromatic number**, and it is denoted by $\chi(G)$. Since this graph parameter is not well-defined for (multi)graphs with loops, we simply ignore them if they are present in $G_t$. Potential parallel edges do not cause any problems but, of course, can be ignored too.

The second monotone property we investigate in this paper is the property that $\chi(G_t) \geq k$ for some value of $k = k(n)$. Trivially, the player can achieve this property by constructing $K_k$ so earlier results immediately imply the corresponding lower bounds. We will prove matching upper bounds (up to a multiplicative factor of $2 + o(1)$) yielding the following three results. Hence, in all regimes, the chromatic number is of order of the clique number.

In the first regime, the ratio between the upper and the lower bound is $2 + o(1)$.

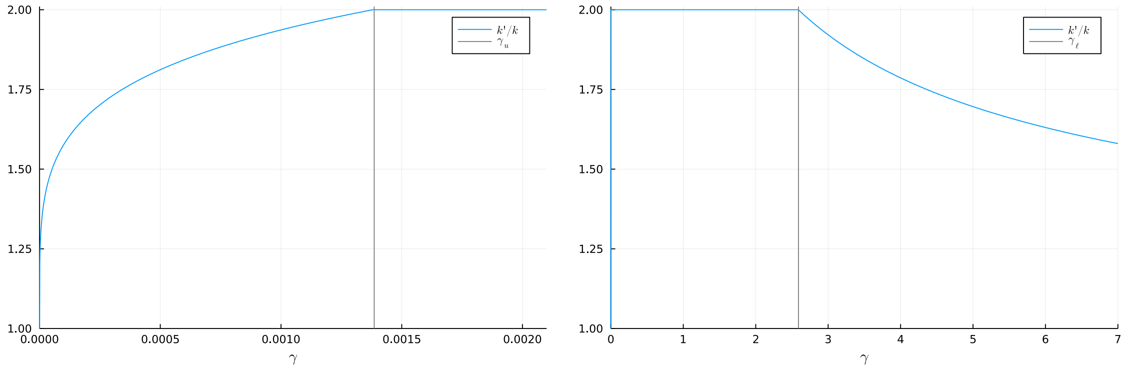

FIGURE 2. The ratio between the upper ($k'$) and the lower ($k$) bound for the order of a largest complete graph: small (left figure) and large (right figure) values of $\gamma$.

**Theorem 1.5.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t \ll n \log n$. Let $\beta = \beta(n) := n \log n/t$. (In particular, $\beta \to \infty$, as $n \to \infty$.) Define $\ell = \ell(n)$ and $k = k(n)$ as in Theorem 1.3.*

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq k$.*
*(b) There is no strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq 2\ell + 2 \sim 2k$.*

In the second regime, the ratio between the upper and the lower bound is at most $2 + o(1)$ (see Figure 3).

**Theorem 1.6.** *Suppose that $t = t(n) = \gamma n \log n$ for some $\gamma \in (0, \infty)$. Let $\xi = \xi(\gamma) \in (1, \infty)$ be defined as in (1). Define $\ell = \ell(n)$ and $k = k(n)$ as in Theorem 1.4.*

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq k$.*
*(b) There is no strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq 2\ell + 2 = \Theta(k)$.*

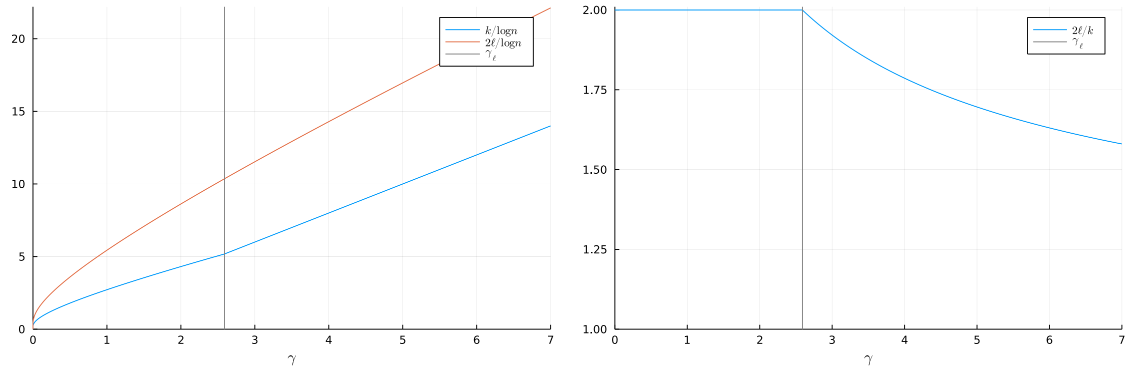

FIGURE 3. The upper ($2\ell + 2$) and the lower ($k$) bound for $\chi(G_t)$ (left figure) as well as the ratio between the two (right figure).

Note that if $\gamma \to \infty$ in the above result, then $\xi \to 1$ and so both bounds are asymptotically tight: $\chi(G_t) \sim 2\gamma \log n$.

**Theorem 1.7.** *Suppose that $t = t(n) \geq \omega n \log n$, where $\omega = \omega(n)$ is any function that tends to infinity as $n \to \infty$. Define $k = k(n)$ and $k' = k'(n)$ as in Observation 1.2.*

*Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq \min\{k, n\}$.*
*(b) There is no strategy that a.a.s. creates $G_t$ such that $\chi(G_t) \geq k' \sim k$.*

1.5. **Main Results—Independent Sets.** An **independent set** is a set of vertices in a graph, no two of which are adjacent. The **independence number** $\alpha(G)$ of a graph $G = (V, E)$ is the cardinality of a maximum independent set of vertices. As for the chromatic number, we simply ignore loops if they are present in $G_t$.

The last monotone property we investigate in this paper is the property that $\alpha(G_t) \le k$ for a given value of $k = k(n)$. We have a good understanding of the independence number of $G_t$ when the average degree tends to infinity together with $n$.

**Theorem 1.8.** *Suppose that $t = t(n)$ is such that $n \ll t \ll n^2$. Let $\lambda = \lambda(n) = t/n$. Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) \;\le\; \frac{n}{2\lambda}\left(1 + \mathcal{O}(\sqrt{\log \lambda/\lambda}) + \mathcal{O}(\lambda/n)\right) \;\sim\; \frac{n}{2\lambda}.$$

*(b) There is no strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) \;<\; \frac{n}{2\lambda + 1}.$$

Suppose now that the average degree is of the same order as the order of a graph, that is, $t = t(n) \sim cn^2$. In this case, we determine the independence number precisely unless $c = 1/(2q)$ for some $q \in \mathbb{N}$. If $c = 1/(2q)$ for some $q \in \mathbb{N}$, then the upper and the lower bounds may be off by one.

**Theorem 1.9.** *Suppose that $t = t(n) \sim cn^2$ for some constant $c \in (0, 1]$. Let $\lambda = \lambda(n) = t/n \sim cn$. Then, the following hold.*

*(a) There exists a strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) \;\le\; \left\lceil \frac{n}{2\lambda}\left(1 + \mathcal{O}(\sqrt{\log \lambda/\lambda})\right)\right\rceil.$$

*(b) There is no strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) \;<\; \frac{n}{2\lambda + 1}.$$

On the other extreme case, if $t = t(n) \ll n$, then the number of vertices that are *not* isolated is (deterministically) at most $2t = o(n)$ and so $\alpha(G_t) \sim n$. Understanding $\alpha(G_t)$ seems to be more challenging when $t \sim cn$ for some constant $c \in (0, \infty)$. It is easy to see that $\alpha(G_t) = \Theta(n)$ but determining the constants hidden in the $\Theta(\cdot)$ notation appears to be difficult. Indeed, in this regime we do not even know the behaviour of the independence number of the binomial random graph [2]. This random graph, much easier model to analyze, is a special case of the semi-random process and can be easily mimicked by it. We provide a few natural upper and lower bounds later on but more work needs to be done to have a better understanding of this graph parameter.

1.6. **Structure of the Paper.** We first introduce some basic concentration tools and present known results about the semi-random process (Section 2). Complete graphs are investigated in Section 3, chromatic number in Section 4, and independent sets in Section 5. We conclude the paper with summarizing open problems that are left to be investigated (Section 6).

## 2. Preliminaries

2.1. **Concentration Tools.** Let us first state a few specific instances of Chernoff's bound that we will find useful. Let $X \in \text{Bin}(n, p)$ be a random variable distributed according to a Binomial distribution with parameters $n$ and $p$. Then, a consequence of **Chernoff's bound** (see e.g., [18, Theorem 2.1]) is that for any $t \ge 0$ we have

$$\mathbb{P}(X \ge \mathbb{E}X + t) \;\;\le\;\; \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right) \tag{3}$$

$$\mathbb{P}(X \le \mathbb{E}X - t) \quad \le \quad \exp\left(-\frac{t^2}{2\mathbb{E}X}\right). \tag{4}$$

Moreover, let us mention that the bound holds in a more general setting as well, that is, for $X = \sum_{i=1}^{n} X_i$ where $(X_i)_{1\le i\le n}$ are independent variables and for every $i \in [n]$ we have $X_i \in \mathrm{Bernoulli}(p_i)$ with (possibly) different $p_i$-s (again, see e.g., [18] for more details). Finally, it is well-known that the Chernoff bound also applies to negatively correlated Bernoulli random variables [10].

2.2. **The Differential Equation Method.** For one of our bounds, we will use the differential equation method (see [6] for a gentle introduction) to establish dynamic concentration of our random variables. The origin of the differential equation method stems from work done at least as early as 1970 (see Kurtz [22]), and which was developed into a very general tool by Wormald [28, 29] in the 1990's. Indeed, Wormald proved a "black box" theorem, which gives dynamic concentration so long as some relatively simple conditions hold. Warnke [26] recently gave a short proof of a somewhat stronger black box theorem.

2.3. **Literature Review.** Since the semi-random process is still a relatively new model, let us highlight a few results on the model.

2.3.1. *Perfect Matchings.* In the very first paper [5], it was shown that the semi-random process is general enough to approximate (using suitable strategies) several well-studied random graph models, including an extensively studied $k$-out process (see, for example, Chapter 18 in [19]). In the $k$-out process, each vertex independently connects to $k$ randomly selected vertices which results in a random graph on $n$ vertices and $kn$ edges.

Since the 2-out process has a perfect matching *a.a.s.* [13], we immediately get that one can create a perfect matching in $(2+o(1))n$ rounds. By coupling the semi-random process with another random graph that is known to have a perfect matching *a.a.s.* [20], the upper bound can be improved to $(1 + 2/e + o(1))n < 1.73576n$. This bound was consecutively improved by investigating a fully adaptive algorithm [16]. The currently best upper bound is $1.20524n$. On the other hand, the lower bound observed in [5] ($(\ln(2) + o(1))n > 0.69314n$) was improved as well, and now we know that one needs at least $0.93261n$ rounds to create a perfect matching [16].

2.3.2. *Hamilton Cycles.* It is known that a.a.s. the famous 3-out process is Hamiltonian [7]. Since the semi-random process can be coupled with the $k$-out process [5] (for any $k \in \mathbb{N}$), we get that a.a.s. one may create a Hamilton cycle in $(3 + o(1))n$ rounds. A new upper bound was obtained in [14] in terms of an optimal solution to an optimization problem whose value is believed to be $2.61135n$ by numerical support.

The upper bound of $(3 + o(1))n$ obtained by simulating the 3-out process is *non-adaptive.* That is, the strategy does *not* depend on the history of the semi-random process. The above mentioned improvement proposed in [14] uses an adaptive strategy but in a weak sense. The strategy consists of 4 phases, each lasting a linear number of rounds, and the strategy is adjusted *only* at the end of each phase (for example, the player might identify vertices of low degree, and then focus on connecting circles to them during the next phase).

In [15], a fully adaptive strategy was proposed that pays attention to the graph $G_t$ and the position of $u_t$ for every single step $t$. As expected, such a strategy creates a Hamilton cycle substantially faster than the weakly adaptive or non-adaptive strategies, and it allows to improve the upper bound from $2.61135n$ to $2.01678n$. One more trick was observed recently which further improves an upper bound to $1.84887n$ [12]. After combining all the ideas together, the currently best upper bound is equal to $1.81696n$ [11].

Let us now move to the lower bounds. As observed in the initial paper introducing the semi-random process [5], if $G_t$ has a Hamiltonian cycle, then $G_t$ has minimum degree at least 2. Thus, a.a.s. it takes at least $(\ln 2 + \ln(1 + \ln 2) + o(1))n \ge 1.21973n$ rounds to achieve this property as this is exactly how many rounds are needed to get the property of having the minimum degree at least 2 [5]. In [14], the lower bound mentioned above was shown to not be tight. The lower bound was increased by $\varepsilon n = 10^{-8} n$ and so numerically negligible. Better bound was obtained in [15] and now we know that a.a.s. it takes at least $1.26575n$ rounds to create a Hamilton cycle.

2.3.3. *Spanning Subgraphs.* Let us now discuss what is known about the property of containing a given spanning graph $H$ as a subgraph. It was asked by Noga Alon whether for any bounded-degree $H$, one can construct a copy of $H$ *a.a.s.* in $O(n)$ rounds. This question was answered positively in a strong sense in [4], in which it was shown that any graph with maximum degree $\Delta$ can be constructed *a.a.s.* in $(3\Delta/2 + o(\Delta))n$ rounds. They also proved that if $\Delta = \omega(\log(n))$, then this upper bound improves to $(\Delta/2 + o(\Delta))n$ rounds. Note that both of these upper bounds are asymptotic in $\Delta$. When $\Delta$ is constant in $n$, such as in both the perfect matching and Hamiltonian cycle setting, determining the optimal dependence on $\Delta$ for the number of rounds needed to construct $H$ remains open. Moving to this direction, $k$-factors and $k$-connectivity was studied recently in [21].

2.3.4. *A Few Other Directions.* Finally, let us mention that sharp thresholds were studied recently in [23]. In fact, the results from there apply to a more general class of processes including the semi-random process. Moreover, some interesting variants of the semi-random process are already considered. In [8], a random spanning tree of $K_n$ is presented to the player who needs to keep one of the edges. In [17], squares presented by the process follow a random permutation. In [25], the process in which $k$ random squares are presented and the player needs to select one of them before creating an edge is considered. Hypergraphs are investigated in [3] and [24].

## 3. Complete Graphs

Let us start with a well-known observation, closely related to the coupon collector problem, that the number of squares on a given vertex is a random variable that is asymptotically distributed as the Poisson random variable.

**Lemma 3.1.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t = \mathcal{O}(n \log n)$. For any vertex $v \in [n]$, let $X_v = X_v(n)$ be the number of squares that land on $v$ until time $t$ and let*

$$\lambda \ := \ \mathbb{E}[X_v] = t/n.$$

*Then, the following properties hold.*

*(a) For any vertex $v \in [n]$ and for any $k = o(\sqrt{n})$,*

$$\mathbb{P}(X_v = k) \ \sim \ \frac{\lambda^k}{k!} \exp(-\lambda).$$

*(b) Suppose that $t = t(n) = n \log n / \beta$ for some $\beta = \beta(n) \to \infty$ as $n \to \infty$. As in Theorem 1.3, define*

$$\ell \ = \ \ell(n) \ := \ \frac{\log n}{\log \beta - 2 \log\log \beta} \ \sim \ \frac{\log n}{\log \beta} \ = \ \frac{\beta}{\log \beta} \cdot \lambda.$$

*Then, a.a.s.*

$$\max\Big( X_v : v \in [n] \Big) \ \le \ \ell \ \sim \ \frac{\beta}{\log \beta} \cdot \lambda.$$

*(c) Suppose that $t = t(n) = \gamma n \log n$ for some $\gamma \in (0, \infty)$. Let $\xi = \xi(\gamma) \in (1, \infty)$ be as in (1). As in Theorem 1.4, define*

$$\ell \ = \ \ell(n) \ := \ \frac{1-\gamma}{\log \xi - 1} \log n \ = \ \xi \gamma \log n \ = \ \xi \lambda.$$

*Then, a.a.s.*

$$\max\Big( X_v : v \in [n] \Big) \ \le \ \ell \ = \ \xi \lambda.$$

*Proof.* Note that for any $k = o(\sqrt{n})$,

$$\begin{aligned}
\mathbb{P}(X_v = k) &= \binom{t}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{t-k} \\
&= \frac{t^k}{k!}\left(1+\mathcal{O}(k/t)\right)^k\left(\frac{1}{n}\right)^k\exp\left(-\frac{1}{n}+\mathcal{O}(1/n^2)\right)^{t-k} \\
&= \frac{(t/n)^k}{k!}\left(1+\mathcal{O}(k^2/t)\right)\exp\left(-t/n+\mathcal{O}(k/n)+\mathcal{O}(t/n^2)\right) \\
&\sim \frac{\lambda^k}{k!}\exp(-\lambda).
\end{aligned}$$

The property (a) holds.

To show property (b), first note that, since $t = n^{1+o(1)}$, we get that $\beta = n^{o(1)}$ (but also, by definition, $\beta \to \infty$ as $n \to \infty$) and so $\ell \to \infty$ as $n \to \infty$. It follows from part (a) and the Stirling's formula ($\ell! \sim \sqrt{2\pi\ell}(\ell/e)^\ell$) that for any vertex $v \in [n]$,

$$\begin{aligned}
\mathbb{P}\Big(X_v = \ell\Big) &\sim \frac{\lambda^\ell}{\ell!}e^{-\lambda} \le \frac{\lambda^\ell}{\ell!} = o\left(\frac{(t/n)^\ell}{(\ell/e)^\ell}\right) = o\left(\left(\frac{e\log n}{\ell\beta}\right)^\ell\right) = o\left(\left((1+o(1))\frac{e\log\beta}{\beta}\right)^\ell\right) \\
&= o\left(\exp\left(-\ell\big(\log\beta - \log\log\beta - \mathcal{O}(1)\big)\right)\right) \\
&= o\left(\exp\left(-\frac{\log n}{\log\beta - 2\log\log\beta}\big(\log\beta - \log\log\beta - \mathcal{O}(1)\big)\right)\right) \\
&= o\left(\exp\left(-\log n\right)\right) = o(1/n).
\end{aligned}$$

Now, note that for any $\ell \le k < n^{1/3}$ we have

$$\frac{\mathbb{P}(X_v = k+1)}{\mathbb{P}(X_v = k)} \sim \frac{\lambda}{k+1} \le \frac{\lambda}{\ell} \sim \frac{\log n/\beta}{\log n/\log\beta} = \frac{\log\beta}{\beta} = o(1),$$

and so $\mathbb{P}(\ell \le X_v \le n^{1/3}) \sim \mathbb{P}(X_v = \ell)$. Finally, note that

$$\mathbb{P}(X_v \ge n^{1/3}) \le \binom{t}{n^{1/3}}\left(\frac{1}{n}\right)^{n^{1/3}} \le \left(\frac{et}{n^{4/3}}\right)^{n^{1/3}} = \exp\Big(-\Theta(n^{1/3}\log n)\Big) = o(1/n).$$

Combining all of these properties together we get that

$$\mathbb{P}(X_v \ge \ell) \le \mathbb{P}(\ell \le X_v \le n^{1/3}) + \mathbb{P}(X_v \ge n^{1/3}) = o(1/n), \tag{5}$$

and the property (b) holds by the union bound over all vertices $v$.

The same argument can be used to show property (c). This time, for any vertex $v \in [n]$,

$$\begin{aligned}
\mathbb{P}\Big(X_v = \ell\Big) &\sim \frac{\lambda^\ell}{\ell!}e^{-\lambda} = o\left(\left(\frac{\gamma\log n}{\ell/e}\right)^\ell\exp\Big(-\gamma\log n\Big)\right) \\
&= o\Big(\exp\Big(\ell\log(e/\xi) - \gamma\log n\Big)\Big) \\
&= o\Big(\exp\Big(-\ell(\log\xi - 1) - \gamma\log n\Big)\Big) \\
&= o\Big(\exp\Big(-(1-\gamma)\log n - \gamma\log n\Big)\Big) \\
&= o(1/n).
\end{aligned}$$

Since $\xi > 1$, for any $\ell \le k < n^{1/3}$ we have

$$\frac{\mathbb{P}(X_v = k+1)}{\mathbb{P}(X_v = k)} \sim \frac{\lambda}{k+1} \le \frac{\lambda}{\ell} = \frac{1}{\xi} < 1,$$

and so $\mathbb{P}(\ell \leq X_v \leq n^{1/3}) = \mathcal{O}(\mathbb{P}(X_v = \ell))$. The conclusion is the same, $\mathbb{P}(X_v \geq \ell) = o(1/n)$ (see (5)), and the property (c) holds by the union bound over all vertices $v$. □

We independently consider lower and upper bounds for the order of complete graphs one may build during the semi-random process.

### 3.1. Lower Bounds.

We present two algorithms that can be used to build complete graphs. Both can be used for any value of $t$ but the first algorithm will turn out to be better than the second one, provided that $t \leq \gamma_\ell n \log n$, where $\gamma_\ell = (2\log 2 - 1)^{-1} \approx 2.59$.

**Algorithm 3.2.** *The algorithm consists of $k-1$ phases. The first phase has only one round in which the player creates $K_2$, a single edge.*

*At the beginning of phase $i$, $i \geq 2$, a complete graph $K_i$ on the vertex set $v_1, v_2, \ldots, v_i$ is already constructed. Any other vertex $v$ that is covered by $s = s(v) \geq 1$ squares has the following property: for any $j \in [s]$, the $j$-th square is connected to a circle on vertex $v_j$. The player maintains this property by applying the following strategy. If a square lands on vertex $v \notin \{v_1, v_2, \ldots, v_i\}$ that is already covered by $s$ squares, then she connects $v$ to vertex $v_{s+1}$. (If $v = v_j$ for some $j \in [i]$, then she plays arbitrarily—that edge is ignored anyway.) The $i$-th phase ends when a square lands on a vertex with $i-1$ squares and a copy of a complete graph $K_{i+1}$ is created.*

*The algorithm ends at the end of phase $k-1$ when a copy of $K_k$ is created.*

The analysis of Algorithm 3.2 proves Theorem 1.3(a) and the first part of Theorem 1.4(a), namely, the lower bound of $k_1$.

*Proof of Theorem 1.3(a) and the first part of Theorem 1.4(a).* Let us first prove Theorem 1.3(a). Recall that

$$k = k(n) = \frac{\log n - 2\log\log n - \lambda}{\log \beta} \sim \frac{\log n}{\log \beta} = \frac{\beta}{\log \beta} \cdot \lambda,$$

where $\lambda = \lambda(n) = t/n \ll \log n$ is the expected number of squares on a given vertex and $\beta = \beta(n) = n\log n/t \to \infty$, as $n \to \infty$. Note that $\beta = n^{o(1)}$ (but also, by definition, $\beta \to \infty$ as $n \to \infty$) and so $k \to \infty$ as $n \to \infty$. On the other hand, since $\beta \to \infty$ as $n \to \infty$, we get that $k = o(\log n)$.

Suppose that the player uses Algorithm 3.2 to play the game against the semi-random process. We will show that at time $t$ a.a.s. there are at least $k$ vertices that are covered by at least $k$ squares. It is easy to see that the algorithm has to end before that round. This will imply the lower bound of $k$.

Let $Y = Y(t)$ be the number of vertices that are covered by $k$ squares at time $t$. Using Lemma 3.1 and the Stirling's formula ($k! \sim \sqrt{2\pi k}(k/e)^k$), we get that

$$\begin{aligned}
\mathbb{E}[Y] &\sim n \cdot \frac{\lambda^k}{k!}\exp(-\lambda) \\
&= \Theta(k^{-1/2}) \cdot n \cdot \left(\frac{e\lambda}{k}\right)^k \exp(-\lambda) \\
&= \Theta(k^{-1/2}) \cdot \exp\left(\log n - k\log\left(\frac{k}{e\lambda}\right) - \lambda\right). \qquad (6)
\end{aligned}$$

Since $k \leq \frac{\beta}{\log \beta} \cdot \lambda$, we get that $\frac{k}{e\lambda} \leq \frac{\beta}{e\log\beta} \leq \beta$ and so

$$\begin{aligned}
\mathbb{E}[Y] &= \Omega(k^{-1/2}) \cdot \exp\left(\log n - k\log\beta - \lambda\right) \\
&= \Omega(k^{-1/2}) \cdot \exp\left(\log n - (\log n - 2\log\log n - \lambda) - \lambda\right) \\
&= \Omega\left(\frac{\log^2 n}{k^{1/2}}\right) \gg k^{3/2} \gg k,
\end{aligned}$$

as $k = o(\log n)$. Finally, note that events "vertex $v$ is covered by $k$ squares" associated with different vertices are negatively correlated. As a result, it follows immediately from Chernoff's bound (4) and the comment right after it that a.a.s. $Y \ge k$. (Alternatively, one could us the second moment method to get the same conclusion.) As argued above, this finishes the proof of Theorem 1.3(a).

The same argument implies the lower bound of $k_1$ in Theorem 1.4(a). Recall that this time $\lambda = \lambda(n) = t/n = \gamma \log n$ and

$$k_1 = k_1(n) = \frac{(1-\gamma)\log n - 2\log\log n}{\log \xi - 1} \sim \frac{1-\gamma}{\log \xi - 1} \log n = \xi\gamma \log n.$$

Computations performed in (6) still apply. Since $k_1 \le \xi\gamma \log n$, we get that

$$\begin{aligned} \mathbb{E}[Y] &= \Omega(k_1^{-1/2}) \cdot \exp\left(\log n - k_1 \log(\xi/e) - \gamma \log n\right) \\ &= \Omega(k^{-1/2}) \cdot \exp\left(\log n - ((1-\gamma)\log n - 2\log\log n) - \gamma \log n\right) \\ &= \Omega\left(\frac{\log^2 n}{k^{1/2}}\right) = \Omega\left(k^{3/2}\right) \gg k, \end{aligned}$$

as $k = \Theta(\log n)$. As before, we conclude that a.a.s. $Y \ge k$ and we are done. □

To prove the second part of Theorem 1.4(a) (that is, the lower bound of $k_2$), we need to analyze the second algorithm which performs better when $t \ge \gamma_\ell n \log n$.

**Algorithm 3.3.** *Suppose that $k = 2\ell + 1$ for some $\ell \in \mathbb{N}$. Before the game starts, select arbitrarily $k$ vertices $v_0, v_1, \ldots, v_{2\ell}$ from the vertex set $[n]$. For each $i \in [2\ell] \cup \{0\}$ and for each $j \in [\ell]$, the player connects the $j$th square landing on vertex $v_i$ with vertex $v_{(i+j) \pmod{2\ell}}$. The algorithm ends when each vertex $v_i$ is covered by at least $\ell$ squares, and a copy of $K_k$ is created.*

Clearly, the player could partition the vertex set into $n/k$ sets and then use a slightly more sophisticated algorithm in which she tries to simultaneously create $n/k$ complete graphs, by applying Algorithm 3.3 independently to each part. Such algorithm would finish once at least one set induces a complete graph. We do not use and analyze it as it would give an asymptotically negligible improvement on the lower bound. However, it will be useful later on once we analyze the chromatic number.

*Proof of the second part of Theorem 1.4(a).* Recall that $t = t(n) = \gamma n \log n$ for some $\gamma \in (0, \infty)$ and

$$k_2 = k_2(n) = 2\gamma \log n - 4\sqrt{\gamma \log n \log\log n} \sim 2\gamma \log n.$$

Let $k$ be the smallest odd integer that is at least $k_2$ and assume $k = 2\ell + 1$ for some $\ell \in \mathbb{N}$. Clearly, $k = k_2 + \mathcal{O}(1)$.

Suppose that the player uses Algorithm 3.3 to play the game against the semi-random process. For any $i \in [2\ell] \cup \{0\}$, let $X_i$ be the number of squares on $v_i$ at time $t$. Note that $X_i \in \mathrm{Bin}(t, 1/n)$ with $\mathbb{E}[X_i] = \gamma \log n$. It follows from Chernoff's bound (4) that

$$\begin{aligned} \mathbb{P}(X_i < \ell) &= \mathbb{P}\Big(X_i \le \mathbb{E}[X_i] - 2\sqrt{\gamma \log n \log\log n} + \mathcal{O}(1)\Big) \\ &\le \exp\left(-\frac{(2\sqrt{\gamma \log n \log\log n} + \mathcal{O}(1))^2}{2\gamma \log n}\right) \\ &= \exp\Big(-2\log\log n + o(1)\Big) \\ &\sim (\log n)^{-2}. \end{aligned}$$

Hence, by the union bound over all vertices $v_i$, the algorithm does not finish before time $t$ with probability at most $(2\ell + 1) \cdot \mathcal{O}((\log n)^{-2}) = \mathcal{O}((\log n)^{-1}) = o(1)$. Hence the desired bound holds a.a.s., and the proof is finished. □

3.2. **Upper Bounds.** Suppose first that $t = t(n) = n \log n/\beta$ for some $\beta = \beta(n) \to \infty$ as $n \to \infty$. Lemma 3.1(b) shows that a.a.s. the number of squares on any vertex is at most $\ell \sim \log n/\log \beta$. It implies immediately that one cannot construct a complete graph of size larger than $2\ell + 1$. But the truth is that a.a.s. it is only possible to build a complete graph of size asymptotic to $\ell$. The main difficulty in showing this lies in the fact that the player can easily create vertices of large degree by placing a large number of circles on some vertices. So this is certainly not the bottleneck for this problem. The key observation used in the proof is that in order to create a large complete graph, many squares need to land on vertices that are already covered by circles, but this happens quite rarely.

Suppose that vertex $u$ is covered by $k$ squares, $u_{t_1}, u_{t_2}, \ldots, u_{t_k}$, for some $k \le \ell$. We will first estimate the number of squares that land "on top" of the associated circles $v_{t_1}, v_{t_2}, \ldots, v_{t_k}$. Formally, we will say that $(v_{t_i}, u_s)$ is a **rare pair** if $v_{t_i}$ and $u_s$ cover the same vertex and $s > t_i$. Note, in particular, that when distinct squares fall on the same circle, those are still counted as distinct rare pairs. The name is justified by the next lemma which shows that on average only $o(\ell)$ squares land on a given circle.

**Lemma 3.4.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t \ll n \log n$. Let $\beta = \beta(n) = n \log n/t \to \infty$, as $n \to \infty$. Let $\ell = \ell(n) \sim \log n/\log \beta$ and $\epsilon = \epsilon(n) = o(1)$ be defined as in Theorem 1.3. Then, the following property holds a.a.s.: for any vertex $u$,*

*(a) $u$ is covered by $k \le \ell$ squares, $u_{t_1}, u_{t_2}, \ldots, u_{t_k}$,*
*(b) the associated circles, $v_{t_1}, v_{t_2}, \ldots, v_{t_k}$, belong to less than $\ell^2\epsilon = o(\ell^2)$ rare pairs.*

*Proof.* Property (a) follows immediately from Lemma 3.1(b). Since we aim for a statement that holds a.a.s. we may assume that property (a) holds when proving property (b).

Note that if $(v_{t_i}, u_s)$ forms a rare pair, then the square $u_s$ needs to arrive after the circle $v_{t_i}$ is placed (that is, $s > t_i$). Hence, the probability that a given vertex $u$ fails to satisfy property (b) does not depend on the strategy of the player and can be upper bounded by

$$p := \binom{t}{\ell^2\epsilon}\left(\frac{\ell}{n}\right)^{\ell^2\epsilon} \le \left(\frac{et}{\ell^2\epsilon}\cdot\frac{\ell}{n}\right)^{\ell^2\epsilon} = \left(\frac{e\log n}{\ell\beta\epsilon}\right)^{\ell^2\epsilon} = \exp\left(-\ell^2\epsilon\log\left(\frac{\ell\beta\epsilon}{e\log n}\right)\right).$$

Suppose first that $\beta \le \log n/\log\log n$. Then, since $\ell \ge \log n/\log \beta$, we get that $\epsilon = e^2 \log\beta/\beta \ge e^2 \log n/(\beta\ell)$ and so

$$p \le \exp\left(-\ell^2\epsilon\right) \le \exp\left(-e^2\ell\log n/\beta\right).$$

Since $\beta \le \log n/\log\log n$ (so, in particular, $\log\beta \le \log\log n$),

$$p \le \exp\left(-e^2\ell\log\log n\right) \le \exp\left(-e^2\ell\log\beta\right) \le \exp\left(-e^2\log n\right) = o(1/n).$$

Suppose now that $\log n/\log\log n < \beta \le \log^2 n$. In particular, $(1 + o(1))\log\log n \le \log\beta \le 2\log\log n$. Then, since

$$\log\ell \ge \log\log n - \log\log\beta = (1 + o(1))\log\log n,$$

we get that

$$\epsilon = \frac{15\log\ell}{\ell} \ge \frac{2e^2\log\log n}{\ell} \ge \frac{e^2\log\beta}{\ell} = \frac{e^2\ell\log\beta}{\ell^2} \ge \frac{e^2\log n}{\ell^2}.$$

It follows that

$$p \le \exp\left(-e^2\log n\cdot\log\left(\frac{e\beta}{\ell}\right)\right) = \exp\left(-e^2\log n\cdot\log\left((e + o(1))\frac{\beta\log\beta}{\log n}\right)\right).$$

Since $\beta \ge \log n/\log\log n$ and $\log\beta \ge (1 + o(1))\log\log n$, we conclude that

$$p \le \exp\left(-e^2\log n\cdot\log\left((e + o(1))\right)\right) = \exp\left(-(e^2 + o(1))\log n\right) = o(1/n).$$

Finally, suppose that $\beta > \log^2 n$. This time $\epsilon = e/\ell$ and, since $\sqrt{\beta} > \log n$,

$$\begin{aligned} p &\leq \exp\Big(-e\ell \log\Big(\frac{\beta}{\log n}\Big)\Big) \leq \exp\Big(-e\ell\log\Big(\sqrt{\beta}\Big)\Big) = \exp\Big(-(e/2)\ell\log\beta\Big) \\ &= \exp\Big(-(e/2)\log n\Big) = o(1/n). \end{aligned}$$

In all scenarios, the conclusion follows by the union bound over all vertices. □

For a given graph $G = (V, E)$ and any set of vertices $S \subseteq V$, we will use $G[S]$ to denote the graph **induced** by set $S$, that is, $G[S] = (S, E')$ and edge $uv \in E$ is in $E'$ if and only if $u \in S$ and $v \in S$. The above lemma immediately implies the following useful corollary.

**Corollary 3.5.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t \ll n\log n$. Let $\ell = \ell(n)$ and $\epsilon = \epsilon(n)$ be defined as in Lemma 3.4. Then, a.a.s. for any set $S \subseteq [n]$, $G_t[S]$ has at most $|S|\ell^2\epsilon$ rare pairs.*

Our next observation shows that one can remove only a few edges from $G_t[S]$ in order to destroy all rare pairs.

**Lemma 3.6.** *Suppose that $t = t(n)$ is such that $t = n^{1+o(1)}$ and $t \ll n\log n$. Let $\ell = \ell(n)$ and $\epsilon = \epsilon(n)$ be defined as in Theorem 1.3. Then, a.a.s. for any set $S \subseteq [n]$, one can remove at most $2|S|\ell\sqrt{\epsilon}$ edges from $G_t[S]$ to remove all rare pairs.*

*Proof.* Suppose that some vertex $w$ yields $r$ rare pairs. Let

$$R_w = \{(v_{t_i}, u_{s_i}) : v_{t_i} = u_{s_i} = w \text{ and } s_i > t_i\}$$

be the set of rare pairs associated with vertex $w$, let $V_w = \{v_{t_i} : (v_{t_i}, u_{s_i}) \in R_w \text{ for some } u_{s_i}\}$ be the set of the associated circles, and let $U_w = \{u_{s_i} : (v_{t_i}, u_{s_i}) \in R_w \text{ for some } v_{t_i}\}$ be the set of the associated squares. We will first show that one can remove at most $2\sqrt{r}$ edges to destroy all rare pairs from $R_w$.

If $|V_w| \leq \sqrt{r}$, then one can remove all edges associated with circles from $V_w$ which clearly destroys all rare pairs from $R_w$. Similarly, if $|U_w| \leq \sqrt{r}$, then one can achieve the same by removing all edges associated with squares from $U_w$. We may then assume that $|V_w| > \sqrt{r}$ and that $|U_w| > \sqrt{r}$.

Let $\hat{s}$ be the largest integer from $[t]$ with the property that $\hat{U}_w = \{u_{s_i} \in U_w : s_i \geq \hat{s}\}$ has cardinality $\sqrt{r}$. In other words, $\hat{U}_w \subseteq U_w$ consists of the $\sqrt{r}$ "youngest" squares from $U_w$. Similarly, let $\hat{t}$ be the smallest integer from $[t]$ with the property that $\hat{V}_w = \{v_{t_i} \in V_w : t_i \leq \hat{t}\}$ has cardinality $\sqrt{r}$. In other words, $\hat{V}_w \subseteq V_w$ consists of the $\sqrt{r}$ "oldest" circles from $V_w$. Let us remove all edges associated with squares from $\hat{U}_w$, and let us remove all edges associated with circles from $\hat{V}_w$, (or both), for a total of at most $2\sqrt{r}$ edges. We claim that this procedure removes all rare pairs from $R_w$.

For a contradiction, suppose that a rare pair $(v_{t_i}, u_{s_i})$ is not removed. In particular, $t_i > \hat{t}$ and $s_i < \hat{s}$. On the other hand, by the definition of being a rare pair, $t_i < s_i$. We conclude that $\hat{t} < \hat{s}$. But it means that each circle from $\hat{V}_w$ forms a rare pair with any square from $\hat{U}_w$ for a total of $\sqrt{r} \cdot \sqrt{r} = r$ rare pairs. With the additional rare pair $(v_{t_i}, u_{s_i})$, there are at least $r + 1$ rare pairs which gives us the desired contradiction, and the claim is proved: one can remove at most $2\sqrt{r}$ edges to destroy all rare pairs from $R_w$.

Consider any set $S \subseteq [n]$. By Corollary 3.5, since we aim for a statement that holds a.a.s., we may assume that $G_t[S]$ yields at most $|S|\ell^2\epsilon$ rare pairs, that is, $\sum_{w\in S} r_w \leq |S|\ell^2\epsilon$, where $r_w$ is the number of rare pairs associated with vertex $w$. By the above observation, one may remove at most $\sum_{w\in S} 2\sqrt{r_w}$ edges from $G_t[S]$ to destroy all of them. Clearly, the optimization problem

$$\max \sum_{w\in S} 2\sqrt{r_w} \qquad \text{subject to} \qquad \sum_{w\in S} r_w \leq |S|\ell^2\epsilon \text{ and } r_w \geq 0 \text{ for all } w \in S$$

attains its maximum when all $r_w$'s are equal. We conclude that it is possible to remove at most

$$\sum_{w\in S} 2\sqrt{r_w} \le \sum_{w\in S} 2\sqrt{\ell^2\epsilon} = 2|S|\ell\sqrt{\epsilon}$$

edges from $G_t[S]$ to destroy all rare pairs, and the proof of the lemma is finished. □

Now, we are ready to finish the proof of Theorem 1.3.

*Proof of Theorem 1.3(b).* Let us apply *any* strategy to play the game. For a contradiction, suppose that there exists a $K_{k'}$ at time $t$, where $k' = \ell(1+4\epsilon^{1/4})$ as in the statement of the theorem. Recall that $\ell \sim \log n/\log\beta$ and $\epsilon = \epsilon(n) = o(1)$. It is also straightforward to check that $\ell\epsilon^{1/4} \to \infty$ as $n \to \infty$. Let $S \subseteq [n]$ be any set of cardinality $k'$ that induces a complete graph. Since we aim for a statement that holds a.a.s., we may apply Lemma 3.6. It follows that one can remove at most $2k'\ell\epsilon^{1/2} < 4\ell^2\epsilon^{1/2}$ edges from $G_t[S]$ in order to destroy all rare pairs. Note that after that operation, set $S$ satisfies the following properties:

(a) $S$ has cardinality $k' = \ell(1+4\epsilon^{1/4})$,

(b) the number of edges induced by $S$ (and so also the number of squares) is more than

$$\binom{k'}{2} - 4\ell^2\epsilon^{1/2} = \binom{\ell}{2} + 4\ell^2\epsilon^{1/4} + \binom{4\ell\epsilon^{1/4}}{2} - 4\ell^2\varepsilon^{1/2} > \binom{\ell}{2} + 4\ell^2\epsilon^{1/4},$$

(note that the first equality is due to a simple fact that $\binom{a+b}{2} = \binom{a}{2} + ab + \binom{b}{2}$ for $a, b \in \mathbb{N}$)

(c) $S$ induces no rare pair,

(d) there are at most $\ell$ squares on any vertex.

Let us now remove all edges induced by $S$ and put them back, one by one, following the order they appeared during the semi-random process. We will distinguish $k'$ phases. The first phase starts when the circle associated with the first edge lands on vertex $v_1 \in S$. Since there are no rare pairs (property (c)), no square will land on $v_1$ but other circles might end up there. The first phase continues as long as circles continue landing on $v_1$. The second phase starts when some circle lands on vertex $v_2 \neq v_1$. Note that, since all edges introduced during the first phase are edges of the graph and all the circles cover vertex $v_1$, all squares cover unique vertices at the beginning of the second phase. In particular, there is at most one square on vertex $v_2$ when the first circle is placed on $v_2$.

In general, phase $i$ starts when some circle is placed on vertex $v_i \notin \{v_1, v_2, \ldots, v_{i-1}\}$. Arguing as above, we conclude that at that point there are at most $i-1$ squares on $v_i$, and no more squares can land on it in the future. This is a useful upper bound for the number of squares, provided that $i \le \ell$. For larger values of $i$ (that is, $\ell < i \le k' = \ell(1+4\epsilon^{1/4})$), we may apply property (d). We conclude that the number of squares is at most $\binom{\ell}{2} + 4\epsilon^{1/4}\ell\cdot\ell$, which contradicts property (b). This finishes the proof of the theorem. □

Suppose now that $t = t(n) = \gamma n\log n$ for some $\gamma \in (0,\infty)$. The upper bound of $k_2'$ in Theorem 1.4(b) is trivial and the argument above can be easily adjusted to show the upper bound of $k_1'$. We carefully explain the adjustment needed below.

*Proof of Theorem 1.4(b).* First, let us note that the upper bound of $k_2' = 2\ell+1$ is indeed trivial. For a contradiction, suppose that some set $S$ of cardinality at least $2\ell+2$ induces a complete graph on $(2\ell+2)(2\ell+1)/2$ edges (and so it induces that many squares). Hence, by averaging argument, there is a vertex with at least $(2\ell+1)/2 > \ell$ squares which contradicts Lemma 3.1(c).

It remains to prove the upper bound of $k_1'$. As in Lemma 3.4, we first need to upper bound the number of rare pairs generated by one vertex. The probability that a given vertex $u$ generates at least $c\ell^2$ rare pairs is at most

$$\begin{aligned} p &:= \binom{t}{c\ell^2}\left(\frac{\ell}{n}\right)^{c\ell^2} \le \left(\frac{e\gamma n\log n}{c\ell^2}\cdot\frac{\ell}{n}\right)^{c\ell^2} = \left(\frac{e\gamma\log n}{c\ell}\right)^{c\ell^2} = \left(\frac{e}{c\xi}\right)^{c\ell^2} \\ &= \exp\left(-c(\xi\gamma\log n)^2\log\left(\frac{c\xi}{e}\right)\right) \\ &= \exp\left(-\Theta(\log^2 n)\log\left(1+\frac{1}{\sqrt{\log n}}\right)\right) \\ &= \exp\left(-\Theta(\log^{3/2} n)\right) = o(1/n), \end{aligned}$$

when $c = (e/\xi)(1+\log^{-1/2} n)$. (Note that $\log(1+x) = x + \mathcal{O}(x^2)$.) By the union bound over all vertices $u$ we get that a.a.s. no vertex generates at least $c\ell^2$ rare pairs. We conclude that a.a.s. for any set $S \subseteq [n]$, $G_t[S]$ induces at most $c\ell^2|S|$ rare pairs (the counterpart of Corollary 3.5). Lemma 3.6 still applies and we get that a.a.s. for any set $S$, one can remove at most $2\sqrt{c}\ell|S|$ edges from $G_t[S]$ to remove all rare pairs.

We finish the proof as before. For a contradiction, suppose that there exists a complete graph on $k' = \ell(1+bc^{1/4})$ vertices, where $b$ is a constant that will be properly tuned soon. We may assume that $1+bc^{1/4} \le 2$; otherwise, the trivial upper bound of $k_2' = 2\ell+1$ applies. A.a.s. for any set $S$ of cardinality $k'$, after destroying all rare pairs from $G_t[S]$, the number of edges left is at least

$$\begin{aligned} \binom{k'}{2} - 2\sqrt{c}\ell k' &= \binom{\ell}{2} + bc^{1/4}\ell^2 + \binom{bc^{1/4}\ell}{2} - 2\sqrt{c}\ell k' \\ &> \binom{\ell}{2} + bc^{1/4}\ell^2 + \frac{b^2c^{1/2}\ell^2}{2}\left(1-\mathcal{O}\left(\frac{1}{\ell}\right)\right) - 4\sqrt{c}\ell^2 \\ &= \binom{\ell}{2} + bc^{1/4}\ell^2 + \frac{c^{1/2}\ell^2}{2}\left(b^2 - 8 - \mathcal{O}\left(\frac{1}{\log n}\right)\right) \\ &> \binom{\ell}{2} + bc^{1/4}\ell^2, \end{aligned}$$

when, for example, $b = \sqrt{8}(1+\log^{-1/2} n)$. As before, we get the desired contradiction which implies the upper bound of

$$\begin{aligned} k' = \left(1+bc^{1/4}\right)\ell &= \left(1+2\sqrt{2}(e/\xi)^{1/4}(1+\log^{-1/2} n)^2\right)\ell \\ &\le \left(1+2\sqrt{2}(e/\xi)^{1/4}\right)\ell\left(1+3\log^{-1/2} n\right) = k_1'. \end{aligned}$$

This finishes the proof of the theorem. □

## 4. Chromatic Number

Parts (a) in Theorems 1.5, 1.6 and 1.7 follow immediately from our results for complete graphs, namely, parts (a) in Theorems 1.3 and 1.4, and Observation 1.2. Parts (b) will follow from upper bounds for the number of squares that land on vertices, Lemma 3.1.

Let us start with some useful basic facts about degeneracy of graphs. Recall that for a given $d \in \mathbb{N}$, a graph $H$ is **$d$-degenerate** if every sub-graph $H' \subseteq H$ has minimum degree $\delta(H') \le d$ (where the minimum degree of a graph is the minimum degree over all vertices). The **degeneracy** of $H$ is the smallest value of $d$ for which $H$ is $d$-degenerate.

The **$d$-core** of a graph $H$ is the maximal induced subgraph $H' \subseteq H$ with minimum degree $\delta(H') \ge d$. (Note that the $d$-core is well defined, though it may be empty. Indeed, if $S \subseteq V(H)$ and $T \subseteq V(H)$ induce sub-graphs with minimum degree at least $d$, then the same is true for $S \cup T$.)

If $H$ has degeneracy $d$, then it has a non-empty $d$-core. Indeed, by definition, $H$ is *not* $(d-1)$-degenerate and so it has a sub-graph $H'$ with $\delta(H') \ge d$. Moreover, it follows immediately from the definition that if $H$ has degeneracy $d$, then there exists a permutation of the vertices of $H$, $(v_1, v_2, \ldots, v_k)$, such that for each $\ell \in [k]$ vertex $v_\ell$ has degree at most $d$ in the sub-graph induced by the set $\{v_1, v_2, \ldots, v_\ell\}$. Indeed, one can define such permutation recursively. Let $v_k$ be any vertex in $H$ that is of degree at most $d$. Then, let $v_{k-1}$ be any vertex of degree at most $d$ in the graph $H'$ induced by the set $\{v_1, v_2, \ldots, v_{k-1}\}$, etc.

The above properties imply a useful reformulation of degeneracy: a graph $H$ is $d$-degenerate if and only if the edges of $H$ can be oriented to form a directed acyclic graph $D$ with maximum out-degree at most $d$. In other words, there exists a permutation of the vertices of $H$, $(v_1, v_2, \ldots, v_k)$, such that for every directed edge $(v_i, v_j) \in D$ we have $i > j$ and the out-degrees are at most $d$. As a consequence, we get another well-known but useful property: for any $d$-degenerate graph $H$ we have $\chi(H) \le d+1$. Indeed, one may colour vertices of $H$ greedily using the permutation $(v_k, v_{k-1}, \ldots, v_1)$.

With these properties, we maye easily prove Theorems 1.5 and 1.6.

*Proof of Theorems 1.5(b) and 1.6(b).* Fix an arbitrary strategy for the player, and consider graph $G_t$ generated at time $t$. Let $X_v = X_v(n)$ be the number of squares that land on $v$ until time $t$. By Lemma 3.1, we know that a.a.s.

$$\max\Big(X_v : v \in [n]\Big) \ \le \ \ell.$$

Since we aim for a statement that holds a.a.s., we may assume that this property is satisfied. Let $S \subseteq [n]$ be any subset of vertices. Since each edge connects a square with a circle, $G_t[S]$ induces at most $|S|\ell$ edges and so the average degree in $G_t[S]$ is at most $2\ell$. It follows that $\delta(G_t[S]) \le 2\ell$ for any $S \subseteq [n]$ and so $G_t$ is $2\ell$-degenerate. The above observation implies that $\chi(G_t) \le 2\ell + 1$ which finishes the proof of the theorem. □

## 5. Independent Sets

5.1. **Upper Bound.** We will first prove an upper bound that not only implies Theorem 1.8(a) and Theorem 1.9(a) but also provides a good upper bound when the average degree of $G_t$ is a constant, especially when that constant is large. Having said that, we do not tune our argument to get the best possible bound but rather aim for an easy argument that provides the upper bound that matches the lower bound when the average degree tends to infinity as $n \to \infty$.

**Lemma 5.1.** *Suppose that $t = t(n) = \Omega(n)$. Let $\lambda = \lambda(n) = t/n$, let $\ell = \ell(n) = \lambda - \sqrt{5\lambda \log \lambda}$, and let $k = k(n) = 2\lceil \ell \rceil + 1$. Finally, let*

- $u = u(n) = \lceil n/k \rceil = \left\lceil \frac{n}{2\lambda}(1 + \mathcal{O}(\sqrt{\log \lambda / \lambda})) \right\rceil$, *if* $\lambda \gg n^{2/5}$,
- $u = u(n) = \lceil n/k \rceil (1 + k^2 \sqrt{\log \lambda}/\lambda^{5/2}) = \frac{n}{2\lambda}(1 + \mathcal{O}(\sqrt{\log \lambda/\lambda}))$, *if* $1 \ll \lambda = \mathcal{O}(n^{2/5})$,
- $u = u(n) = n\left(\frac{1}{2\lceil \ell \rceil + 1} + \frac{2\lceil \ell \rceil}{\lambda^{5/2}}\right) + n^{3/4}$, *if* $\lambda = \Theta(1)$.

*Then, there exists a strategy that a.a.s. creates $G_t$ such that $\alpha(G_t) \le u$.*

*Proof.* Let us arbitrarily partition the set of vertices $[n]$ into $\lceil n/k \rceil$ parts, each of size at most $k = 2\lceil \ell \rceil + 1$. We will independently apply Algorithm 3.3 to each part. The algorithm succeeds on a given part and produces a complete graph of order $k$ if all vertices in that part receive at least $\ell$ squares at time $t$. For a given $i \in \{1, 2, \ldots, \lceil n/k \rceil\}$, let $X_i$ be the indicator random variable for the event that Algorithm 3.3 fails on part $i$, and let $X = \sum_{i=1}^{\lceil n/k \rceil} X_i$ be the number of parts that failed.

For a given vertex $v \in [n]$, let $Y_v$ be the random variable counting the number of squares on $v$ at time $t$. Clearly, $Y_v \in \mathrm{Bin}(t, 1/n)$ with $\mathbb{E}[Y_v] = t/n = \lambda$. It follows immediately from Chernoff's bound (4) that

$$\mathbb{P}(Y_v < \ell) = \mathbb{P}\Big(Y_v < \lambda - \sqrt{5\lambda \log \lambda}\Big) \le \exp\left(-\frac{5\lambda \log \lambda}{2\lambda}\right) = 1/\lambda^{5/2}.$$

Since the algorithm fails on a given part if at least one vertex in that part (out of at most $k$ vertices) receives less than $\ell$ squares at time $t$,

$$\mathbb{P}(X_i = 1) \le k/\lambda^{5/2} = \mathcal{O}(1/\lambda^{3/2}).$$

Hence, the expected number of parts that fail can be estimated as follows:

$$\mathbb{E}[X] \le \lceil n/k \rceil \cdot \frac{k}{\lambda^{5/2}} = \mathcal{O}(n/\lambda^{5/2}).$$

If $\lambda \gg n^{2/5}$, then $\mathbb{E}[X] \to 0$ as $n \to \infty$ and so, by Markov's inequality, we get that a.a.s. $X = 0$. Since each independent set can have at most one vertex from each part (as all of them are successful a.a.s.), we get that a.a.s. $\alpha(G_t) \le \lceil n/k \rceil$ and the desired bound holds.

Suppose then that $1 \ll \lambda = \mathcal{O}(n^{2/5})$. This time, by Markov's inequality we get that a.a.s. $X \le \mathbb{E}[X]\sqrt{\log \lambda} = \mathcal{O}(n\sqrt{\log \lambda}/\lambda^{5/2})$. As before, each independent set can have at most one vertex from each successful part and, trivially, at most $kX$ vertices from parts that failed. We get that a.a.s.

$$\begin{aligned} \alpha(G_t) &\le \left\lceil \frac{n}{k} \right\rceil + kX = \frac{n}{2\lambda}(1 + \mathcal{O}(\sqrt{\log \lambda/\lambda})) + \mathcal{O}(n\sqrt{\log \lambda}/\lambda^{3/2}) \\ &= \frac{n}{2\lambda}(1 + \mathcal{O}(\sqrt{\log \lambda/\lambda})). \end{aligned}$$

Finally, suppose that $\lambda = \Theta(1)$. Note first that $(X_i)$ is a sequence of negatively correlated random variables. Combining this with earlier observations, we conclude that $X$ can be stochastically upper bounded by $Y = \sum_{i=1}^{\lceil n/k \rceil} Y_i$, where $(Y_i)$ are negatively correlated Bernoulii random variables with parameter $k/\lambda^{5/2}$. It follows from Chernoff's bound (3) and the comment right after it that a.a.s. $X \le Y \le \frac{n}{\lambda^{5/2}} + n^{2/3}$. We use the same observation as before, namely, that each independent set can have at most one vertex from each of the $\lceil n/k \rceil - X$ successful parts and at most $kX$ vertices from the $X$ parts that failed. We get that a.a.s.

$$\begin{aligned} \alpha(G_t) &\le \left(\left\lceil \frac{n}{k} \right\rceil - X\right) + kX = \left\lceil \frac{n}{k} \right\rceil + (k-1)X \\ &\le n\left(\frac{1}{2\lceil \ell \rceil + 1} + \frac{2\lceil \ell \rceil}{\lambda^{5/2}}\right) + O(n^{2/3}). \end{aligned}$$

This finishes the proof of the lemma. □

Let us note that the upper bounds we just proved are asymptotically tight when $\lambda = \lambda(n) \gg 1$. On the other hand, the above bound is not sharp when $\lambda = \Theta(1)$. There are many ways one may improve it. For example, a more careful argument could estimate the size of a largest independent set of a given part that fails (right now, we simply use a trivial upper bound of $k$). Moreover, some parts that succeed receive more squares than needed for the argument (for example, perhaps each vertex in that part receives more than $\ell$ squares). Such additional squares could be used to create slightly larger cliques. Finally, when $\lambda$ is a small constant, a better strategy could be to create a large perfect matching using the adaptive algorithm analyzed in [16].

5.2. **Lower Bounds.** It is easy to prove by induction that for any graph $G = (V, E)$, $\alpha(G) \ge n/(\Delta + 1)$, where $\Delta$ is the maximum degree. Caro [9] and Wei [27] independently proved the following, more refined, version of this observation. (See also [1].)

**Observation 5.2** ([9, 27])**.** *For any graph $G = (V, E)$,*

$$\alpha(G) \ge \sum_{v\in V} \frac{1}{\deg(v)+1} \ge \frac{n}{d+1} \ge \frac{n}{\Delta+1},$$

*where $d = \frac{1}{n}\sum_{v\in V}\deg(v)$ is the average degree and $\Delta = \max_{v\in V}\deg(v)$ is the maximum degree.*

This observation immediately proves parts (b) of Theorems 1.8 and 1.9 since the average degree of $G_t$ is $d = 2|E(G_t)|/n = 2t/m = 2\lambda$. As mentioned above, this simple observation is asymptotically tight when $\lambda = \lambda(n) \gg 1$.

We propose two improvements when $\lambda = \Theta(1)$. Observation 5.2 still applies but the lower bound of $n/(d+1) = n/(2\lambda+1)$ that holds deterministically may be improved with slightly more effort. Indeed, the existence of an independent set of size

$$L(G_t) := \sum_{v\in V} \frac{1}{\deg(v)+1}$$

is still deterministically guaranteed. Understanding the graph parameter $\alpha(G_t)$ is challenging but $L(G_t)$ is relatively easy to deal with. In order to minimize $L(G_t)$, the player should keep the degree distribution as "flat" as possible; in particular, note that $L(G_t)$ is minimized when all vertices have degrees $\lfloor d \rfloor$ or $\lceil d \rceil$. However, she cannot achieve such distribution since squares arrive uniformly at random and so a.a.s. some vertices will receive more than $\lceil d \rceil$ squares.

To get a weaker lower bound we may consider an "off-line" version of the semi-random process, that is, let the player wait till time $t$ before placing all of her circles at once. Clearly, the original process (the "on-line" version) is at least as challenging to the player as its off-line counterpart, so the obtained lower bound also applies there and the lower bound for $\alpha(G_t)$ holds.

Let $Y_k$ be the number of vertices that received $k$ squares at time $t$. It is easy to see (see Lemma 3.1) that a.a.s. for any $k \in \mathbb{N}\cup\{0\}$,

$$Y_k \ \sim\ n\frac{\lambda^k}{k!}\exp(-\lambda).$$

The player will place her $t = \lambda n$ circles greedily on vertices with minimum degree. Let $M \in \mathbb{N}\cup\{0\}$ be the largest integer $m$ such that

$$f(m) := \sum_{k=0}^{m-1}(m-k)\frac{\lambda^k}{k!}\exp(-\lambda) \le \lambda.$$

For each $k \in \{0, 1, \dots, M-1\}$, the player may put $M-k$ circles on each vertex with $k$ squares to make them of degree $M$. The total number of circles used so far is $(1+o(1))nf(M)$ and the fraction of vertices of degree $M$ at this point is asymptotic to

$$g(M) := \sum_{k=0}^{M}\frac{\lambda^k}{k!}\exp(-\lambda).$$

The remaining $(1+o(1))n(\lambda - f(M))$ circles are places on such vertices. Once this is done there are $(1+o(1))nh(k)$ vertices of degree $k$, where

$$h(k) = \begin{cases} g(M) - (\lambda - f(M)) = g(M) - \lambda + f(M) & \text{if} \qquad k = M \\ \frac{\lambda^{M+1}}{(M+1)!}\exp(-\lambda) + \lambda - f(M) & \text{if} \qquad k = M+1 \\ \frac{\lambda^k}{k!}\exp(-\lambda) & \text{if} \qquad k \ge M+2. \end{cases}$$

These observations imply the following lower bound.

**Lemma 5.3.** *Suppose that $t = t(n) \sim \lambda n$ for some $\lambda \in (0, \infty)$. Let $\epsilon > 0$ be any (arbitrarily small) constant. Then, there is no strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) < (1-\epsilon)\, n \sum_{k \geq M} \frac{h(k)}{k+1}.$$

Finally, we analyze how small $L(G_t)$ can get for the original ("on-line") semi-random process. It is easy to see that in order to minimize $L(G_t)$ the player needs to apply a greedy strategy. In this strategy, in each round $s \leq t$ of the process, the player puts a circle on a vertex with minimum degree; if there is more than one such vertex to choose from, the decision which one to select is made arbitrarily.

Note that in each round $s \leq t$, the minimum degree in $G_s$ is at most the average degree, that is, at most $2s/n \leq 2t/n = 2\lambda$ so the player will never put a circle on a vertex of degree more than $2\lambda$. In the greedy strategy, we distinguish phases by labelling them with integers $q \in \{0, 1, \ldots, r\}$, where $r = \lfloor 2\lambda \rfloor$. During the $q$th phase, the minimum degree in $G_s$ is equal to $q$. In order to analyze the evolution of the semi-random process, we will track the following sequence of $r+1$ variables: for $0 \leq i \leq r$, let $Y_i = Y_i(s)$ denote the number of vertices in $G_s$ of degree $i$.

Phase 0 starts at the beginning of the process. Since $G_0$ is empty, $Y_0(0) = n$ and $Y_i(0) = 0$ for $1 \leq i \leq r$. There are initially many isolated vertices but they quickly disappear. Phase 0 ends at time $s$ which is the smallest value of $s$ for which $Y_0(s) = 0$. The DEs method (see Subsection 2.2) will be used to show that a.a.s. Phase 0 ends at time $s_0 \sim x_0 n$, where $x_0$ is an explicit constant which will be obtained by investigating the associated system of DEs. Moreover, the number of vertices of degree $i$ ($1 \leq i \leq r$) at the end of this phase is well concentrated around some values that are also determined based on the solution to the same system of DEs: a.a.s. $Y_i(s_0) \sim y_i(x_0)n$. With that knowledge, we move on to Phase 1 in which we prioritize vertices of degree 1.

Consider any Phase $q$, where $q \in \{0, 1, \ldots, r\}$. This phase starts at time $s_{q-1}$, exactly when the previous phase ends (or at time $s_{-1} := 0$ if $q = 0$). At that point, the minimum degree of $G_{s_{q-1}}$ is $q$, so $Y_i(s) = 0$ for any $s \geq s_{q-1}$ and $i < q$. Hence, we only need to track the behaviour of the remaining $r + 1 - q$ variables. Let us denote $H(s) = (Y_q(s), Y_{q+1}(s), \ldots, Y_r(s))$. Let $\delta_A$ be the Kronecker delta for the event $A$, that is, $\delta_A = 1$ if $A$ holds and $\delta_A = 0$ otherwise. Then, for any $i$ such that $q \leq i \leq r$,

$$\mathbb{E}\Big(Y_i(s+1) - Y_i(s) \mid H(s)\Big) = -\delta_{i=q} + \delta_{i=q+1} - \frac{Y_i(s)}{n} + \delta_{i\geq q+1} \cdot \frac{Y_{i-1}(s)}{n}. \tag{7}$$

Indeed, since the circle is put on a vertex of degree $q$, we always lose one vertex of degree $q$ (term $-\delta_{i=q}$) that becomes of degree $q+1$ (term $\delta_{i=q+1}$). We might lose a vertex of degree $i$ when the square lands on a vertex of degree $i$ (term $Y_i(s)/n$). We might also gain one of them when the square lands on a vertex of degree $i-1$ (term, $Y_{i-1}(s)/n$); note that this is impossible if $i = q$ (term $\delta_{i\geq q+1}$). This suggests the following system of DEs: for any $i$ such that $q \leq i \leq r$,

$$y_i'(x) = -\delta_{i=q} + \delta_{i=q+1} - y_i(x) + \delta_{i\geq q+1} \cdot y_{i-1}(x). \tag{8}$$

It is easy to check that the assumptions of the DEs method are satisfied (we omit details since we did not formally introduce the tool). The conclusion is that a.a.s. during the entire Phase $q$, for any $q \leq i \leq r$), $|Y_i(s) - y_i(s/n)n| = o(n)$. In particular, Phase $q$ ends at time $s_q \sim x_q n$, where $x_q > x_{q-1}$ is the solution of the equation $y_q(x) = 0$. Using the final values $y_i(x_q)$ in Phase $q$ as initial values for Phase $q+1$ we can repeat the argument inductively moving from phase to phase.

We stop the analysis at the end of Phase $r$ when a graph of minimum degree equal to $r+1$ is reached. As discussed earlier, it happens at time $s > t$ and so we may "rewind" the process back to round $t$ to check the degree distribution of $G_t$. Based on that, we may compute $L(G_t)$ which gives us the desired lower bound for $\alpha(G_t)$. Suppose that round $t$ occurs during phase $q \leq r$. A.a.s.

there are $(1+o(1))w(k)n$ vertices of degree $k \geq q$ in $G_t$, where

$$w(k) = \begin{cases} y_k(t/n) & \text{if } q \leq k \leq r \\ \frac{\lambda^k}{k!}\exp(-\lambda) & \text{if } k \geq r+1. \end{cases}$$

These observations imply the following lower bound.

**Lemma 5.4.** *Suppose that $t = t(n) \sim \lambda n$ for some $\lambda \in (0,\infty)$. Let $\epsilon > 0$ be any (arbitrarily small) constant. Then, there is no strategy that a.a.s. creates $G_t$ such that*

$$\alpha(G_t) < (1-\epsilon)\, n \sum_{k \geq M} \frac{w(k)}{k+1}.$$

## 6. Conclusions

In this paper, we investigated three monotone properties. Our bounds are off by at most a multiplicative factor of $2+o(1)$. It would be interesting to close the gap between the upper and the lower bounds (or, at least, narrow them down).

- The property of containing a complete graph of order $k$ is well understood unless $t = t(n) = \Theta(n \log n)$.
- The property of creating a graph with the chromatic number at least $k$, is well understood when $t \gg n \log n$. More work is needed when $t = \mathcal{O}(n \log n)$.
- The property of not having an independent set of size at least $k$ remains to be investigated when $t = \Theta(n)$. In other regimes, the asymptotic behaviour is determined.

## 7. Acknowledgements

This work was done while the authors were visiting the Simons Institute for the Theory of Computing. The second author is supported in part by the Austrian Science Fund (FWF) [10.55776/I6502]. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## References

[1] Noga Alon and Joel H Spencer. *The probabilistic method.* John Wiley & Sons, 2016.

[2] Mohsen Bayati, David Gamarnik, and Prasad Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 105–114, 2010.

[3] Natalie Behague, Trent Marbach, Paweł Prałat, and Andrzej Ruciński. Subgraph games in the semi-random graph process and its generalization to hypergraphs. *arXiv preprint arXiv:2105.07034*, 2022.

[4] Omri Ben-Eliezer, Lior Gishboliner, Dan Hefetz, and Michael Krivelevich. Very fast construction of bounded-degree spanning graphs via the semi-random graph process. *Proceedings of the 31st Symposium on Discrete Algorithms (SODA'20)*, pages 728–737, 2020.

[5] Omri Ben-Eliezer, Dan Hefetz, Gal Kronenberg, Olaf Parczyk, Clara Shikhelman, and Miloš Stojaković. Semi-random graph process. *Random Structures & Algorithms*, 56(3):648–675, 2020.

[6] Patrick Bennett and Andrzej Dudek. A gentle introduction to the differential equation method and dynamic concentration. *Discrete Mathematics*, 345(12):113071, 2022.

[7] Tom Bohman and Alan Frieze. Hamilton cycles in 3-out. *Random Structures & Algorithms*, 35(4):393–417, 2009.

[8] Sofiya Burova and Lyuben Lichev. The semi-random tree process. *arXiv preprint arXiv:2204.07376*, 2022.

[9] Yair Caro. New results on the independence number. Technical report, Technical Report, Tel-Aviv University, 1979.

[10] Devdatt Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *Random Structures & Algorithms*, 13(5):99–124, 1998.

[11] Alan Frieze, Pu Gao, Calum MacRury, Paweł Prałat, and Gregory Sorkin. Building hamiltonian cycles in the semi-random graph process in less than $2n$ rounds. *arXiv preprint arXiv:2311.05533*, 2023.

[12] Alan Frieze and Gregory B Sorkin. Hamilton cycles in a semi-random graph model. *arXiv preprint arXiv:2208.00255*, 2022.

[13] Alan M Frieze. Maximum matchings in a class of random graphs. *Journal of Combinatorial Theory, Series B*, 40(2):196–212, 1986.

[14] Pu Gao, Bogumił Kamiński, Calum MacRury, and Paweł Prałat. Hamilton cycles in the semi-random graph process. *European Journal of Combinatorics*, 99:103423, 2022.

[15] Pu Gao, Calum MacRury, and Pawel Pralat. A fully adaptive strategy for hamiltonian cycles in the semi-random graph process. In Amit Chakrabarti and Chaitanya Swamy, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2022, September 19-21, 2022, University of Illinois, Urbana-Champaign, USA (Virtual Conference)*, volume 245 of *LIPIcs*, pages 29:1–29:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[16] Pu Gao, Calum MacRury, and Paweł Prałat. Perfect matchings in the semirandom graph process. *SIAM Journal on Discrete Mathematics*, 36(2):1274–1290, 2022.

[17] Shoni Gilboa and Dan Hefetz. Semi-random process without replacement. In *Extended Abstracts EuroComb 2021*, pages 129–135. Springer, 2021.

[18] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[19] Michal Karoński and Alan Frieze. *Introduction to Random Graphs.* Cambridge University Press, 2016.

[20] Michal Karoński, Ed Overman, and Boris Pittel. On a perfect matching in a random digraph with average out-degree below two. *Journal of Combinatorial Theory, Series B*, 143, 03 2020.

[21] Hidde Koerts. k-connectedness and k-factors in the semi-random graph process. Master's thesis, University of Waterloo, 2022.

[22] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probability*, 7:49–58, 1970.

[23] Calum MacRury and Erlang Surya. Sharp thresholds in adaptive random graph processes. *Random Structures & Algorithms*, 64(3):741–767, 2024.

[24] Michael Molloy, Pawel Pralat, and Gregory B Sorkin. Matchings and loose cycles in the semirandom hypergraph model. *arXiv preprint arXiv:2401.00559*, 2023.

[25] Paweł Prałat and Harjas Singh. Power of $k$ choices in the semi-random graph process. *Electronic Journal of Combinatorics*, 31(1):#P1.11, 2024.

[26] Lutz Warnke. On Wormald's differential equation method. *Combin. Probab. Comput.*, in press.

[27] VK Wei. A lower bound on the stability number of a simple graph. Technical report, Bell Laboratories Technical Memorandum New Jersey, 1981.

[28] Nicholas C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.*, 5(4):1217–1235, 1995.

[29] Nicholas C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *In Lectures on Approximation and Randomized Algorithms, PWN, Warsaw*, pages 73–155, 1999.

MIT Sloan School of Management, MIT, Cambridge, USA
*Email address*: gamarnik@mit.edu

Institute of Discrete Mathematics, Graz University of Technology, Graz, Austria
*Email address*: kang@math.tugraz.at

Department of Mathematics, Toronto Metropolitan University, Toronto, Canada
*Email address*: pralat@torontomu.ca