# Sharp Thresholds in Adaptive Random Graph Processes

Calum MacRury [*]        Erlang Surya [†]

**Abstract**

The $\mathcal{D}$-process is a single player game in which the player is initially presented the empty graph on $n$ vertices. In each step, a subset of edges $X$ is independently sampled according to a distribution $\mathcal{D}$. The player then selects one edge $e$ from $X$, and adds $e$ to its current graph. For a fixed monotone increasing graph property $\mathcal{P}$, the objective of the player is to force the graph to satisfy $\mathcal{P}$ in as few steps as possible. The $\mathcal{D}$-process generalizes both the Achlioptas process and the semi-random graph process.

We prove a sufficient condition for the existence of a sharp threshold for $\mathcal{P}$ in the $\mathcal{D}$-process. Using this condition, in the semi-random process we prove the existence of a sharp threshold when $\mathcal{P}$ corresponds to being Hamiltonian or to containing a perfect matching. This resolves two of the open questions proposed by Ben-Eliezer et al. (RSA, 2020).

Keywords: Random graph processes, sharp thresholds, online decision-making

# 1 Introduction

Let $n \in \mathbb{N}$, and $K_n$ be the complete graph on vertex set $[n] := \{1, \ldots, n\}$. Suppose that $\mathcal{D}$ is a fixed distribution on (non-empty) subsets of edges of $K_n$. The $\mathcal{D}$-**adaptive random graph process** (shortly, $\mathcal{D}$-process) is a single player game in which the player is initially presented a graph $G_0$ on vertex set $[n]$, which unless specified otherwise, will be the empty graph. In each **step** (or round) $t \in \mathbb{N}$, a subset of edges $X_t$ is sampled from $\mathcal{D}$. The player (who is aware of graph $G_t$ and the subset $X_t$) must then select an edge $Y_t$ from $X_t$ and add it to $G_{t-1}$ to form $G_t$. In this paper, the goal of the player is to devise a strategy which builds a (multi)graph satisfying a given monotone increasing property $\mathcal{P}$ in as few rounds as possible. Some examples of $\mathcal{D}$-processes are the Erdős–Rényi random graph process [11] (where multi-edges are allowed), the Achlioptas process [6], the semi-random graph process [5] (see Section 1.2), and the semi-random tree process [7].

Formally, a **strategy** (i.e., algorithm) $\mathcal{S} = \mathcal{S}_n$ is defined by specifying a sequence of functions $(s_t)_{t=1}^{\infty}$, where for each $t \in \mathbb{N}$, $s_t(G_{t-1}, X_t)$ is a distribution on $X_t$ which depends on the graph at step $t-1$ (and the edges of $X_t$). Then, an edge $Y_t \in X_t$ is chosen according to this distribution. If $s_t$ is an atomic distribution, then $Y_t$ is determined by $G_{t-1}$ and $X_t$. Note that if for each $t \geq 1$, $s_t$ is atomic, then we say that the strategy $\mathcal{S}$ is **deterministic**. In this case, we may assume that each $s_t$ is a function which depends only on $X_1, \ldots, X_t$.

We 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 there is no ambiguity. Moreover, we define the stopping time $T_{\mathcal{S}} = T_{\mathcal{S}}(n)$ to be the minimum $t \geq 1$ such that $G_t^{\mathcal{S}}(n)$ satisfies $\mathcal{P}$, where $T_{\mathcal{S}} := \infty$ if no such $t$ exists. All our asymptotics are with respect to $n \to \infty$, and **with high probability** (w.h.p.) means with probability tending to 1 as $n \to \infty$.

**Definition 1** (Sharp Threshold). Given an edge monotonic property $\mathcal{P}$, we say that there exists a **sharp threshold** for $\mathcal{P}$ in the $\mathcal{D}$-process (or $\mathcal{P}$ admits a sharp threshold), provided there exists a function $m^* = m_{\mathcal{P},\mathcal{D}}^*(n)$ such that for every $\epsilon > 0$:

1. There *exists* a strategy $\mathcal{S}_n'$ such that $\mathbb{P}(T_{\mathcal{S}_n'} \leq (1 + \epsilon)m^*) = 1 - o(1)$.

2. *Every* strategy $\mathcal{S}_n$ satisfies $\mathbb{P}(T_{\mathcal{S}_n} \leq (1 - \epsilon)m^*) = o(1)$.

When $m^*$ satisfies these conditions, we say that it is a sharp threshold of $\mathcal{P}$ in the $\mathcal{D}$-process.

In general, it is difficult to show that a property has a sharp threshold for a $\mathcal{D}$-process due to the fact that one needs to prove a "negative" result that *any* algorithm which executes for too few rounds has win probability at most $o(1)$. This contrasts with many sharp threshold results for (non-adaptive) random graph processes, where the negative result follows easily via the first-moment method. In this paper, we develop a framework which allows us to show that if a property satisfies a "positive" result involving the existence of a certain algorithm, then it must admit a sharp threshold. We formalize this framework in a condition we refer to as *edge-replaceability* (see Definition 2).

There have been a few results which establish the existence of sharp thresholds for adaptive random graph processes. In [3, 4], Ben-Eliezer et al. showed that the property of containing an arbitrary spanning graph with $(1 + o(1))\Delta n/2$ edges has $\Delta n/2$ as a sharp threshold, provided its maximum degree $\Delta$ satisfies $\Delta = \omega(\log n)$. For certain types of Achlioptas processes, Krivelevich et al. [16] [1] showed that the property of being Hamiltonian admits a sharp threshold. Both of

---

[1]The model in [16] samples $k$ edges uniformly at random from the set of currently missing edges instead of the set of all edges (i.e. $\mathcal{D}$ slightly changes over time). However, this distinction does not change the existence of sharp thresholds in the regime of interest.

these papers follow the same high-level approach:

1. A naive lower bound $L$ is obtained from a standard analysis, such that any strategy needs at least approximately $L$ steps to succeed w.h.p..

2. An explicit strategy $\mathcal{S}$ is devised and shown to satisfy the desired property in approximately $L$ steps w.h.p. This second step establishes the existence of a sharp threshold, and essentially all the work is done here.

As one might expect, there are certain limitations to such an explicit approach. First, it is not always the case that the naive lower bound is the right answer. For instance, in the semi-random graph process, Gao et al. [13, 15, 14] established that the naive lower bound can be improved substantially when the property corresponds to containing a Hamiltonian cycle [13, 14], or to containing a perfect matching [15]. Second, even if a sharp threshold does exist, it is not clear that it can be identified by a strategy with an explicit description. For example, in the semi-random graph process, Gao et al. introduce algorithms for constructing Hamiltonian cycles [13, 14], and perfect matchings [15]. While each algorithm satisfies the relevant property in number of steps close to the best known lower bound, the authors indicate that they do *not* believe their algorithms are optimal. This greatly limits their usefulness in terms of proving the existence of sharp thresholds. In this paper, we circumvent these limitations by developing a general machinery which allows us to establish the existence of sharp thresholds in the $\mathcal{D}$-process *implicitly*. That is, without explicitly identifying lower bounds or finding (asymptotically) optimal strategies.

While we are unaware of any work applying this implicit approach to any "truly adaptive" random graph process, it has been used when the player has no real control (i.e., $\mathcal{D}$ is supported on singletons). In his seminal paper, Friedgut [12] proved the existence of sharp thresholds for "global" properties in the Erdős–Rényi random graph process in an implicit way (the model he considered does not allow multi-edges, but this distinction is irrelevant in many regimes of interest). To do so, he identifies the Erdős–Rényi random graph with the product space measure on $\{0,1\}^{\binom{n}{2}}$, and applies Fourier analysis to the Boolean function indicating whether or not the random graph satisfies the given property. It is not clear how such techniques can be generalized to the $\mathcal{D}$-process, as in general, the $\mathcal{D}$-process depends on the decisions of the player, and so it cannot be obviously modelled by a product space measure.

## 1.1 Main result

Given an arbitrary distribution $\mathcal{D}$, Theorem 1 provides a sufficient condition for when a monotone increasing property $\mathcal{P}$ admits a sharp threshold in the $\mathcal{D}$-process. For any $\theta \in (0,1)$, define $m_{\mathcal{P}}(\theta, n)$ to be the minimum $t \geq 1$, such that there exists a strategy $\mathcal{S}'_n$ which satisfies $\mathbb{P}[T_{\mathcal{S}'_n} \leq m_{\mathcal{P}}(\theta, n)] \geq \theta$, and for every strategy $\mathcal{S}_n$, $\mathbb{P}[T_{\mathcal{S}_n} \leq m_{\mathcal{P}}(\theta, n) - 1] < \theta$. For convenience we let $m^* := m^*(n) := m_{\mathcal{P}}(1/2, n)$, and define the sufficient condition used in Theorem 1:

**Definition 2** (Edge-Replaceable)**.** We say that $\mathcal{P}$ is $\omega$-**edge-replaceable** (or just **edge-replaceable**) if there exists $\omega := \omega(n) \to \infty$ such that the following guarantee holds: For any $G \in \mathcal{P}$, and $e \in E(G)$, if we begin the $\mathcal{D}$-process with graph $G_0 = G - e$, then there exists a strategy for the player which constructs some $G' \in \mathcal{P}$ in $\sqrt{m^*}/\omega$ steps with probability at least $1 - o(1/\sqrt{m^*})$. We refer to this strategy as an **edge-replacement procedure** of $\mathcal{P}$.

We remark that for the properties we consider, $G'$ will typically be distinct from $G$.

**Theorem 1** (Sharp threshold). *If $\mathcal{P}$ is $\omega$-edge-replaceable, then for any constants $0 < \theta_1 < \theta_2 < 1$, we have that*

$$m_{\mathcal{P}}(\theta_2, n) - m_{\mathcal{P}}(\theta_1, n) = O_{\theta_1, \theta_2}\left(\frac{m^*}{\omega}\right), \tag{1}$$

*where the implicit constant in the $O$ term depends on $\theta_1, \theta_2$. Thus, $m^*$ is a sharp threshold of $\mathcal{P}$.*

We prove Theorem 1 by fixing an arbitrary strategy which succeeds with probability at least $\theta \in [\theta_1, \theta_2]$ in $m_{\mathcal{P}}(\theta, n)$ steps, and identifying a strategy modification which has the *potential* to increase the strategy's winning probability by $\Omega(1/\sqrt{m^*})$. The proof relies on a martingale concentration inequality, whose full statement we defer to Section 3. After performing this strategy modification, the final graph $G_0$ we are left with may be lacking an edge $e$ necessary to satisfy $\mathcal{P}$, however we can apply an augmentation via the edge-replacement procedure of $\mathcal{P}$ to $G_0$ to recover a graph $G'$ which *does* satisfy $\mathcal{P}$ in $\sqrt{m^*}/\omega$ steps with probability at least $1 - o(1/\sqrt{m^*})$. Thus, we boost the win probability of $\mathcal{S}$ by $\Theta(1/\sqrt{m^*})$ in $\sqrt{m^*}/\omega$ steps. By applying this procedure $\sqrt{m^*}$ times, we increase the original strategy's win probability by $\Theta(1)$. Since this only requires an extra $\Theta(\sqrt{m^*}) \cdot \sqrt{m^*}/\omega = o(m^*)$ steps in total, we are able to establish the existence of a sharp threshold.

## 1.2 Application: The Semi-Random Graph Process

The **semi-random graph process** was suggested by Peleg Michaeli, introduced formally in [5], and studied in [3, 13, 4, 15, 2, 14, 7]. The process is a one player game in which the player begins with the empty graph on $[n]$. In each step $t \geq 1$, the player is given a vertex $u_t$ drawn independently and uniformly at random (u.a.r.) from $[n]$, often referred to as a **square**. They then adaptively pick a vertex $v_t$ (called a **circle**), and add the edge $(u_t, v_t)$ to their current graph. Observe that if $\mathcal{D}$ is the uniform distribution over all spanning stars on $K_n$, then the $\mathcal{D}$-process encodes the semi-random graph process.

To warm-up, we first consider the property $\mathcal{P}_k$ of attaining minimum degree $k \geq 1$ in the semi-random graph process. In [5], Ben-Eliezer et al. identified an explicit constant $h_k$ such that $h_k n$ is a sharp threshold for $\mathcal{P}_k$. Clearly, $\mathcal{P}_k$ is $\omega$-edge-replaceable with $\omega = \sqrt{n}/2$. Thus, applying Theorem 1 to $\mathcal{P}_k$ yields an alternative proof of the existence of a sharp threshold.

Moving to our main applications, let $\mathcal{M}$ be the property of containing a perfect matching (by perfect matching on an odd number of vertices, we mean a matching which saturates all but one vertex), and $\mathcal{H}$ be the property of containing a Hamiltonian cycle. As an application of Theorem 1 and the tools we develop in Section 4, we prove the following sharp threshold result. This result answers two of the open problems proposed by Ben-Eliezer et al. in [4] (the journal version of [3]).

**Theorem 2.** *Let $\mathcal{P} \in \{\mathcal{M}, \mathcal{H}\}$. In this case, if $m^*(n) := m_{\mathcal{P}}^*(n) := m_{\mathcal{P}}(1/2, n)$, then*

1. *'Existence of a threshold': In the semi-random graph process, $m^*$ is a sharp threshold for $\mathcal{P}$*

2. *'Linear growth': There exists some constant $C_{\mathcal{P}} > 0$, such that $m^* = (C_{\mathcal{P}} + o(1))n$.*

There are a few notable complications in proving Theorem 2. First, it turns out that the condition in Theorem 1 does *not* hold for either $\mathcal{M}$ or $\mathcal{H}$. However, for each of $\mathcal{M}$ and $\mathcal{H}$, we can define an *approximate* property that does satisfy the required conditions, and thus admits a sharp threshold. Since each approximate property is closely related to $\mathcal{M}$ and $\mathcal{H}$, we are able to argue that $\mathcal{M}$ and $\mathcal{H}$ also have sharp thresholds. Relating each approximate property with its "full" property relies on the "clean-up" algorithms of Gao et al. [15, 14]. When $\mathcal{P}$ is $\mathcal{M}$, this clean-up algorithm allows one to extend a large matching to a perfect matching in a sublinear number of steps. When $\mathcal{P}$ is $\mathcal{H}$, the clean-up algorithm has a similar guarantee.

We emphasize that it is easy to establish $m^* = \Theta(n)$ for both properties $\mathcal{M}$ and $\mathcal{H}$. Hence the first part of Theorem 2 implies the existence of a function $C_{\mathcal{P}}(n) = \Theta(1)$ such that $m^* = (C_{\mathcal{P}}(n) + o(1))n$. The second part of the theorem shows that $\lim_{n\to\infty} C_{\mathcal{P}}(n)$ exists. Showing the existence of such limit is non-trivial. Indeed, showing the limit of such function is the central question in many topics [1, 10]. We achieved this by considering the optimal strategy $\mathcal{S}_n$ that minimizes $\mathbb{E} T_{\mathcal{S}_n} =: I_n$, then showing that $I_n$ satisfies a certain set of inequalities (see Lemma 22). We then use a purely analytic argument to show the existence of the limit $\lim_{n\to\infty} I_n/n$ (Lemma 19), which quickly leads to the desired result. This is reminiscent of the approach seen in [1], where a more standard subadditive inequality is combined with the brilliant use of interpolation to prove the existence of scaling limit for the size of independent sets in sparse random graphs. In our setting, it is much easier to show that the required inequalities hold by standard "strategy stealing" arguments.

We conclude with the observation that a sharp threshold does not always exist for the semi-random process, which follows as a byproduct of the result in [2].

**Theorem 3.** *Let $L$ be a fixed (finite) list of fixed graphs, none of which are forests. Then in the semi-random graph process the property of containing a graph in $L$ does not admit a sharp threshold.*

The rest of the paper is organized as follows. In Section 2 we prove Theorem 1. In Section 3 we establish the martingale concentration inequality needed to prove Theorem 1. In Section 4 we prove Theorems 2 and 3. Section 5 list some final remarks.

## 2 Proving Theorem 1

Suppose that $\mathcal{P}$ is an edge-replaceable property with respect to $\omega \to \infty$ (see Definition 2) in some arbitrary $\mathcal{D}$-process. Moreover, take $0 < \theta_1 < \theta_2 < 1$. We wish to show that if we are given a strategy which wins after $m(\theta_1, n)$ steps with probability at least $\theta_1$, then we can augment the strategy to boost its win probability to $\theta_2$ in $O(m^*/\omega)$ additional steps. If we can prove this, then it will imply that $m(\theta_2, n) - m(\theta_1, n) = O(m^*/\omega)$. Now, suppose that we have boosted to a win probability of $\theta \in [\theta_1, \theta_2]$, and $\theta^* := \min(\theta + \theta(1-\theta)^3/32, \theta_2)$ is the next *target* probability we wish to boost to. We claim that this increase is attainable in an appropriate number of steps. That is, for each $\theta \in [\theta_1, \theta_2]$,

$$m(\theta^*, n) - m(\theta, n) \leq m^*/\omega. \tag{2}$$

By beginning with $\theta = \theta_1$, and iterating (2) a constant number of times, Theorem 1 follows (see the proof of Theorem 1 below for the details).

### 2.0.1 Reducing (2) to Small Boosts

Instead of trying to directly describe a strategy which implies (2), we first prove that we can boost the winning probability by $\Theta(1/\sqrt{m^*})$ in $O(\sqrt{m^*}/\omega)$ extra steps. More precisely, if a strategy wins with probability $\theta$ after $m(\theta, n)$ steps, then we can augment the strategy such that its winning probability is $\theta + \frac{\theta(1-\theta)^3}{4\sqrt{m^*}}$ after $\sqrt{m^*}/\omega$ additional steps. This is the content of Lemma 4:

**Lemma 4** (Small Boost). *Given constants $0 < \theta_1 < \theta_2 < 1$, for any sufficiently large $n \geq 1$ (depending only on $\theta_1, \theta_2$) and any $\theta \in [\theta_1, \theta_2]$, we have that*

$$m\left(\theta + \frac{\theta(1-\theta)^3}{4\sqrt{m^*}}, n\right) - m(\theta, n) \leq \sqrt{m^*}/\omega.$$

Let us assume that Lemma 4 holds for now. We can then prove (2) by iteratively applying Lemma 4 $\sqrt{m^*}$ times to increase the win probability from $\theta$ to $\theta^*$ in $\sqrt{m^*} \cdot \sqrt{m^*}/\omega = m^*/\omega$ additional steps. We include the details below, and complete the proof of Theorem 1.

*Proof of Theorem 1.* Let us take $n \geq 1$ sufficiently large (as in Lemma 4) and $\theta \in [\theta_1, \theta_2]$. Recall that $\theta^* := \min(\theta + \theta(1-\theta)^3/32, \theta_2)$, and we first must show that (2) holds. I.e.,

$$m\left(\min(\theta + \theta(1-\theta)^3/32, \theta_2), n\right) - m(\theta, n) \leq m^*/\omega. \tag{3}$$

In order to prove this, we iterate Lemma 4 $\sqrt{m^*}$ times. Formally, we define $\gamma_0 := \theta$, and $\gamma_{i+1} := \gamma_i + \gamma_i(1-\gamma_i)^3/(4\sqrt{m^*})$. Observe then that by Lemma 4, for each $i \geq 0$, with $\gamma_i \leq \theta_2$,

$$m(\gamma_{i+1}, n) - m(\gamma_i, n) \leq \sqrt{m^*}/\omega.$$

In particular, if $\gamma_i \leq \min(\theta + \theta(1-\theta)^3, \theta_2) \leq (1+\theta)/2$, then

$$\gamma_{i+1} - \gamma_i \geq \frac{\theta(1 - (1+\theta)/2)^3}{4\sqrt{m^*}} = \frac{\theta(1-\theta)^3}{32\sqrt{m^*}}.$$

Therefore

$$\gamma_{\sqrt{m^*}} \geq \min\left(\theta_2, \theta + \frac{\theta(1-\theta)^3}{32}\right),$$

and so (3) holds. By iterating (3) a constant number of times in a similar manner, Theorem 1 follows. $\qquad\square$

## 2.1 Proving Lemma 4

In this section, we explain the main tools used in the proof of Lemma 4. Fix $\theta \in [\theta_1, \theta_2]$, and set $N := m(\theta, n)$ for convenience. Let us suppose that $\mathcal{S}$ is a strategy which satisfies $\mathcal{P}$ with probability at least $\theta$ after $N$ steps. First notice that we can assume that $\mathcal{S}$ is *deterministic* without loss of generality. This is because the optimal strategy for winning in at most $N$ steps deterministically chooses $Y_i \in X_i$ so as to maximize its win probability (conditional on the current history $X_1, \ldots, X_i$). Since $\mathcal{S}$ is deterministic, there exists an **indicator function** $f$ of $\mathcal{S}$, where $f(X) := 1$ if strategy $\mathcal{S}$ wins when presented the edge subsets of $X := (X_1, \ldots, X_N)$ in order. To prove Lemma 4, we augment $\mathcal{S}$ to get another strategy $\mathcal{S}'$ which wins with probability at least $\theta^* := \theta + \frac{\theta(1-\theta)^3}{4\sqrt{m^*}}$ after $m(\theta, n) + \sqrt{m^*}/\omega$ steps.

We now give an informal overview of the three main parts to the proof of Lemma 4. The full details appear in the appropriate sections.

1. 'Reducing to the free-move $\mathcal{D}$-process': We introduce a new game which gives slightly more power to the player called the **free-move $\mathcal{D}$-process**. The free-move $\mathcal{D}$-process is played in the same way as the $\mathcal{D}$-process, except that the player has one opportunity to pick the subset they desire instead of the subset they received (and then select an edge from this subset). Since $\mathcal{P}$ is edge-replaceable, the win probability of any free-move strategy can be matched by a (regular) strategy, provided the regular strategy is given an additional $\sqrt{m^*}/\omega$ steps (see Lemma 5). Thus it suffices to define a free-move strategy $\mathcal{F}$ which wins with probability at least $\theta^*$ after $m(\theta, n)$ steps.

2. 'Defining `PotentialBoost`': The free-move strategy `PotentialBoost` analyzes the Doob-martingale $M = (M_j)_{j=0}^N$ of $f(X)$ with respect to $(X_j)_{j=1}^N$. Informally, $M_j$ measures the probability that $\mathcal{S}$ will win, given the first $j$ arriving edge subsets $X_1, \ldots, X_j$. Based on this interpretation, `PotentialBoost` follows the strategy of $\mathcal{S}$ up until the first time $\tau \geq 1$ that there is *potential* to increase its win probability. In particular, there is an edge subset $W_\tau$ such that replacing the edge set $X_\tau$ with $W_\tau$ increases the probability that $\mathcal{S}$ will win by at least $c$. At this point, it invokes its free-move to swap $X_\tau$ with $W_\tau$, and then follows the strategy $\mathcal{S}$ as if $X_1, \ldots, W_\tau$ where the first $\tau$ subsets to arrive. Conditional on $\tau \leq N$, this guarantees that `PotentialBoost` has a win probability at least $c$ greater than $\mathcal{S}$.

3. 'Bounding the win probability of `PotentialBoost`': In order to prove that `PotentialBoost` attains a win probability significantly better than $\mathcal{S}$, we must prove that $\mathbb{P}[\tau \leq N] = \Omega(1)$. We do so by proving a martingale concentration result (Theorem 12), and then applying it in a non-standard way. Observe that the function $f$ is $\{0,1\}$-valued, and so since $0 < \theta < 1$, $f(X)$ *cannot* be concentrated about $\mathbb{E}[f(X)]$. On the other hand, we argue that if $\mathbb{P}[\tau \leq N] = o(1)$, then Theorem 12 would force $f(X)$ to be concentrated. Thus, we can conclude that $\mathbb{P}[\tau \leq N] = \Omega(1)$.

### 2.1.1 The Free-Move $\mathcal{D}$-Process

The **free-move $\mathcal{D}$-process** is defined in the same way as the $\mathcal{D}$-process, except that the player can adaptively choose a time $\tau \geq 1$, such that if $X_1, \ldots, X_\tau$ were the previously presented subsets of edges, then they can choose an arbitrary subset $W_\tau$ from $\mathrm{Supp}(\mathcal{D})$ (the support of $\mathcal{D}$). They then get to add an edge $Y_\tau \in W_\tau$ to $G_{\tau-1}$, opposed to an edge from $X_\tau$ (as in the standard game).

Clearly, any strategy for the standard $\mathcal{D}$-process is a strategy for the free-move $\mathcal{D}$-process. Thus, satisfying an edge-monotone property $\mathcal{P}$ in the latter game is no harder than in the former game. However, if $\mathcal{P}$ is edge-replaceable then the advantage gained by the player is not very significant, and so this new game is a good approximation of the original game. We extend all the definitions from the standard $\mathcal{D}$-process to formalize this intuition. Specifically, if $\mathcal{F}$ is a strategy for the free-move $\mathcal{D}$-process, then $G_t^{\mathcal{F}}$ is the graph constructed by following $\mathcal{F}$ in the first $t$ steps. Moreover, $T_{\mathcal{F}}$ is defined to be the first $t \geq 1$ such that $G_t^{\mathcal{F}} \in \mathcal{P}$ (where $T_{\mathcal{F}} := \infty$ if no such $t$ exists.)

**Lemma 5.** *Let $\mathcal{F}$ be a strategy for the free-move $\mathcal{D}$-process process for satisfying a property $\mathcal{P}$ which is $\omega$-edge-replaceable. In this case, there exists a strategy $\mathcal{F}'$ for the (standard) $\mathcal{D}$-process, such that for each $k \geq 1$, $\mathbb{P}[G_{k+\sqrt{m^*}/\omega}^{\mathcal{F}'} \in \mathcal{P}] \geq (1 - o(1/\sqrt{m^*})) \cdot \mathbb{P}[G_k^{\mathcal{F}} \in \mathcal{P}].$*

*Proof of Lemma 5.* Let us assume that $\mathcal{P}$ is $\omega$-edge-replaceable, and $\mathcal{F}$ is a strategy for the free-move $\mathcal{D}$-process process. In order to prove the lemma, it suffices to show that there exists a strategy $\mathcal{F}'$ for the (standard) $\mathcal{D}$-process, such that if both strategies are presented the same (random) edge subsets $(X_t)_{t=1}^\infty$, then with probability $1 - o(1/\sqrt{m^*})$ we have that

$$T_{\mathcal{F}'} \leq T_{\mathcal{F}} + \sqrt{m^*}/\omega. \tag{4}$$

We begin by defining $\mathcal{F}'$ to follow the same decisions of $\mathcal{F}$ up until time $T_{\mathcal{F}}$, where if $\mathcal{F}$ invokes a free-move at some time $1 \leq \tau \leq T_{\mathcal{F}}$, then we define $\mathcal{F}'$ to choose an edge of $X_\tau$ arbitrarily. If $\mathcal{F}$ does not invoke a free-move, then $\tau := \infty$, and the strategies execute identically.

Let $G_{T_{\mathcal{F}}}$ and $G'_{T_{\mathcal{F}}}$ be the graphs constructed by $\mathcal{F}$ and $\mathcal{F}'$ after $T_{\mathcal{F}}$ steps, respectively. At this point, $G_{T_{\mathcal{F}}} \in \mathcal{P}$ (by definition of $T_{\mathcal{F}}$), yet $G'_{T_{\mathcal{F}}}$ may not satisfy $\mathcal{P}$. Specifically, if $\tau < \infty$, then $G'_{T_{\mathcal{F}}}$ will be missing the edge $e$ that $\mathcal{F}$ added at step $\tau$. Note that $G'_{T_{\mathcal{F}}} + e \in \mathcal{P}$, so after $T_{\mathcal{F}}$ steps we define $\mathcal{F}'$ to run the edge-replacement procedure of $\mathcal{P}$ to ensure that after another $\sqrt{m^*}/\omega$ steps,

7

it will be left with a graph which satisfies $\mathcal{P}$ with probability $1 - o(1/\sqrt{m^*})$. This completes the proof of (4), and so the lemma is proven.

$\square$

### 2.1.2 Defining `PotentialBoost`

Recall that $\mathcal{S}$ is a deterministic strategy which wins with probability at least $\theta$ after $N := N(\theta) = m(\theta, n)$ steps, and $f$ is its indicator function. Observe that $\mu := \mathbb{E}[f(X)] = \mathbb{P}[f(X) = 1] \geq \theta$ for $X = (X_1, \ldots, X_N)$, where each $X_i$ is drawn independently from $\mathcal{D}$. Setting $C(\theta) := 1 + \log_2\left(\frac{1}{1-\theta}\right)$, we define

$$c := \frac{\mu(1-\mu)}{\sqrt{2C(\theta)m^*}}. \tag{5}$$

The dependence of $c$ on $\theta$ and $\mu$ is for technical reasons which will only become relevant in Section 2.1.3. For now, it suffices to think of $c$ as $\Theta(1/\sqrt{m^*})$. Our goal is to identify instantiations of $X$ in which by using the *free-move* of `PotentialBoost`, we can boost the win probability of $\mathcal{S}$ by $c$.

We first consider the Doob-martingale $M = (M_j)_{j=0}^N$ of $f(X)$ with respect to $(X_j)_{j=1}^N$. That is, $M_0 := \mathbb{E}[f(X)]$ and $M_j := \mathbb{E}[f(X) \mid X_1, \ldots X_j]$ for $j \in [N]$, where $M_N = f(X)$. Moreover, for each $1 \leq j \leq N$, define the function $f_j$, where for each $(r_1, \ldots, r_j) \in \text{Supp}(\mathcal{D})^j$,

$$f_j(r_1, \ldots, r_j) := \mathbb{E}[f(X) \mid (X_i)_{i=1}^j = (r_i)_{i=1}^j]. \tag{6}$$

Equivalently, $f_j(r_1, \ldots, r_j)$ is the probability that $\mathcal{S}$ wins after $N$ steps, conditional on $X_1 = r_1, \ldots, X_j = r_j$. Observe that $M_j = f_j(X_1, \ldots, X_j)$ by construction. We say that $(r_1, \ldots, r_j) \in \text{Supp}(\mathcal{D})^j$ has **potential**, provided there exists $w_j \in \text{Supp}(\mathcal{D})$ such that

$$f_j(r_1, \ldots, r_j) + c < f_j(r_1, \ldots, w_j). \tag{7}$$

In this case, we refer to $w_j$ as a **witness** for $(r_1, \ldots, r_j)$. Note that there may be multiple witnesses for $(r_1, \ldots, r_j)$. Intuitively, if $(r_1, \ldots, r_j)$ has potential, then $\mathcal{S}$ has a better win probability when $X_1 = r_1, \ldots, X_j = w_j$, opposed to when $X_1 = r_1, \ldots, X_j = r_j$, While we cannot ensure that $X_j = w_j$ in the standard $\mathcal{D}$-process, we *can* in the free-move $\mathcal{D}$-process.

Algorithm `PotentialBoost` runs for $N$ steps, and yet has a slightly higher win probability than $\mathcal{S}$. We assume that the algorithm is presented the subsets $X_1, \ldots, X_N$ in order. We choose the edges in the same way as $\mathcal{S}$ up until the first step $1 \leq t \leq N$ such that $(X_1, \ldots, X_t)$ has potential. Let us define $1 \leq \tau \leq N$ to be this step, where $\tau := \infty$ if no such step exists. Assuming $\tau \leq N$, we identify an arbitrary witness $W_\tau$ of $(X_1, \ldots, X_\tau)$. At this point, we invoke our free-move, and replace $X_\tau$ with $W_\tau$. For step $\tau$ and each subsequent step, we choose the edges by following the strategy of $\mathcal{S}$ with $X_\tau$ replaced by $W_\tau$. Below is a formal description of the algorithm:

**Algorithm `PotentialBoost` Free-move Strategy**

---

**Input:** $\widetilde{G}_0 = ([n], \emptyset)$.
**Output:** $\widetilde{G}_N$

1: **for** $t = 1, \ldots, \min\{\tau - 1, N\}$ **do** ▷ follow decisions of $\mathcal{S}$
2:      Define $Y_t$ to be the edge chosen by $\mathcal{S}$ when given $(X_1, \ldots, X_t)$.
3:      $\widetilde{G}_t := \widetilde{G}_{t-1} \cup Y_t$.
4: **end for**
5: **if** $\tau \leq N$ **then** let $W_\tau$ be an arbitrary witness of $(X_1, \ldots, X_\tau)$.    ▷ $(X_1, \ldots, X_\tau)$ has potential
6:      Define $Y_\tau$ be the edge chosen by $\mathcal{S}$ when given $(X_1, \ldots, W_\tau)$.
7:      $\widetilde{G}_\tau := \widetilde{G}_{\tau-1} \cup Y_\tau$.                             ▷ execute a free-move
8:      **for** $t = \tau + 1, \ldots, N$ **do**                 ▷ follow $\mathcal{S}$ with $X_\tau$ replaced with $W_\tau$
9:          Define $Y_t$ be the edge chosen by $\mathcal{S}$ when given $(X_1, \ldots, W_\tau, \ldots, X_t)$.
10:          $\widetilde{G}_t := \widetilde{G}_{t-1} \cup Y_t$.
11:      **end for**
12: **end if**
13: **return** $\widetilde{G}_N$.

---

Let $G_N = G_N^{\mathcal{S}}$ be the graph formed by $\mathcal{S}$ when passed edge subsets $X_1, \ldots, X_N$ and let $\widetilde{G}_N$ be the corresponding output from the `PotentialBoost` algorithm. We compare $\widetilde{G}_N$ to $G_N$:

**Lemma 6.** *The graph $\widetilde{G}_N$ satisfies the following:*

1. *If $\mathbb{P}[\tau > N] > 0$, then $\mathbb{P}[\widetilde{G}_N \in \mathcal{P} \mid \tau > N] = \mathbb{P}[G_N \in \mathcal{P} \mid \tau > N]$.*

2. *If $\mathbb{P}[\tau \leq N] > 0$, then $\mathbb{P}[\widetilde{G}_N \in \mathcal{P} \mid \tau \leq N] > \mathbb{P}[G_N \in \mathcal{P} \mid \tau \leq N] + c$*

3. *$\mathbb{P}[\widetilde{G}_N \in \mathcal{P}] \geq \mathbb{P}[G_N \in \mathcal{P}] + c \cdot \mathbb{P}[\tau \leq N]$.*

*Proof.* We prove the properties of Lemma 6 in order. First observe that $\tau \leq N$ if and only if `PotentialBoost` makes a free-move at some step. Moreover, if `PotentialBoost` does *not* make a free move, then the algorithm simply executes $\mathcal{S}$ as the subsets $X_1, \ldots, X_N$ arrive. Thus, $\widetilde{G}_N$ and $G_N$ are the same graph, and so in particular,

$$\mathbb{P}[\widetilde{G}_N \in \mathcal{P} \mid \tau > N] = \mathbb{P}[G_N \in \mathcal{P} \mid \tau > N].$$

Let us now consider the case $\tau \leq N$. It will be convenient to define $R$ to be those $(r_1, \ldots, r_k) \in \cup_{i=1}^N \mathrm{Supp}(\mathcal{D})^i$, such that $(r_1, \ldots, r_k)$ has potential, yet no proper prefix of $(r_1, \ldots, r_k)$ has potential. Observe that conditional on $\tau \leq N$, $(X_1, \ldots, X_\tau)$ is supported on $R$. Now, fix $(r_1, \ldots, r_k) \in R$, and condition on $(X_1, \ldots, X_k) = (r_1, \ldots, r_k)$. Observe then that $\widetilde{G}_N$ is distributed as $G_N$ conditional on $(X_1, \ldots, X_k) = (r_1, \ldots, w_k)$. Thus, for each $(r_1, \ldots, r_k) \in R$,

$$\begin{aligned}
\mathbb{P}[\widetilde{G}_N \in \mathcal{P} \mid (X_i)_{i=1}^k = (r_i)_{i=1}^k] &= \mathbb{P}[G_N \in \mathcal{P} \mid (X_i)_{i=1}^{k-1} = (r_i)_{i=1}^{k-1}, X_k = w_k] \\
&= \mathbb{E}[f(X) \mid (X_i)_{i=1}^{k-1} = (r_i)_{i=1}^{k-1}, X_k = w_k] \\
&= f_k(r_1, \ldots, w_k) > f_k(r_1, \ldots, r_k) + c,
\end{aligned}$$

where the second equality uses the definition of $f$, and the final inequality holds since $(r_1, \ldots, r_k)$ has potential. By averaging over all the elements of $R$, property (2) follows. Property (3) is implied by (1) and (2):

$$\mathbb{P}[\widetilde{G}_N \in \mathcal{P}] \geq \mathbb{P}[G_N \in \mathcal{P}, \tau > N] + \mathbb{P}[G_N \in \mathcal{P}, \tau \leq N] + c \cdot \mathbb{P}[\tau \leq N] = \mathbb{P}[G_N \in \mathcal{P}] + c \cdot \mathbb{P}[\tau \leq N].$$

$\square$

### 2.1.3 Bounding the Win Probability of `PotentialBoost`

Observe that property (3) of Lemma 6 ensures `PotentialBoost` has a win probability at least as large as $\mathcal{S}$. Moreover, by definition, $c = \Theta(1/\sqrt{m^*})$. Thus, if we can show that the stopping time $\tau$ of `PotentialBoost` satisfies $\mathbb{P}[\tau \leq N] = \Omega(1)$, then this will prove that `PotentialBoost` boosts the win probability of $\mathcal{S}$ by $\Omega(1/\sqrt{m^*})$, as (roughly) claimed by Lemma 4. Before proceeding with this lower bound, we state the following upper bound on $N(\theta)$, which relies on a standard multi-round exposure argument to boost the win probability from $1/2$ to $\theta$.

**Proposition 7.** *If $C(\theta) = 1 + \log_2\left(\frac{1}{1-\theta}\right)$, then $N(\theta) \leq C(\theta)m^*$.*

*Proof.* Suppose $\mathcal{S}'$ is a strategy that succeed with probability at least $1/2$ in $m$ steps. For any integer $k$, consider a strategy that runs for $km$ steps where for any $i = 0, \ldots, k-1$, in steps $\{im+1, im+2, \ldots, (i+1)m\}$ we run the strategy $\mathcal{S}'$ as if the graph is empty. Then the probability of failure after $kN(1/2)$ steps is at most $(1/2)^k$. By letting $k = \lceil \log_2(1/(1-\theta)) \rceil$ and noting that $1 - (1/2)^k \geq \theta$, we get that

$$N(\theta) \leq km \leq \left(1 + \log_2\left(\frac{1}{1-\theta}\right)\right)m$$

as desired. $\qquad\square$

**Lemma 8.** *If $\mu := \mathbb{P}[G_N \in \mathcal{P}]$, $\mathbb{P}[\tau \leq N] \geq \frac{1-\mu}{2}$.*

To establish Lemma 8, we invoke a concentration inequality for the Doob martingale of $f(X)$ with respect to $(X_j)_{j=1}^N$ (see Corollary 10). We state and prove the full theorem in Section 3, and for now just indicate how we apply a special case of this theorem for our specific needs. The rough idea is as follows. If $\mathbb{P}[\tau \leq N]$ were $o(1)$, then our concentration inequality would imply that $f(X)$ must be concentrated about its expectation. But $f(X) \in \{0,1\}$, and $\mathbb{E}[f(X)] = \mu \geq \theta$, so since we may assume that $\mu$ is bounded away from 1, this is not possible. Thus, $\mathbb{P}[\tau \leq N]$ must be $\Omega(1)$.

To formalize this intuition, let us say that $r = (r_1, \ldots, r_N) \in \mathrm{Supp}(\mathcal{D})^N$ is **stable** if no prefix of $r$ has potential. That is, for each $1 \leq j \leq N$ and $w_j \in \mathrm{Supp}(\mathcal{D})$,

$$f_j(r_1, \ldots, w_j) - f_j(r_1, \ldots, r_j) \leq c. \tag{8}$$

Define $\Gamma \subseteq \mathrm{Supp}(\mathcal{D})^N$ to be the stable elements of $\mathrm{Supp}(\mathcal{D})^N$. We relate $\Gamma$ to the stopping time $\tau$ of `PotentialBoost` in the following way:

**Proposition 9.** *$X = (X_1, \ldots, X_N) \in \Gamma$ if and only if $\tau > N$. In particular, $\mathbb{P}[X \notin \Gamma] = \mathbb{P}[\tau \leq N]$.*

We then invoke the following one-sided concentration inequality to lower bound $\mathbb{P}[X \notin \Gamma]$:

**Corollary 10** (of Theorem 12). *For each $t \geq 0$, $\mathbb{P}[f(X) \leq \mathbb{E}f(X) - t] \leq \exp\left(\frac{-2t^2}{Nc^2}\right) + \mathbb{P}[X \notin \Gamma]$.*

*Proof of Lemma 8.* By setting $t = \mu/2$ where $\mu = \mathbb{P}[f(X) = 1] = \mathbb{E}[f(X)]$, Corollary 10 implies that

$$1 - \mu = \mathbb{P}[f(X) = 0] = \mathbb{P}[f(X) \leq \mu/2] \leq \exp\left(\frac{-\mu^2}{2Nc^2}\right) + \mathbb{P}[X \notin \Gamma].$$

Thus, since $c := \frac{\mu(1-\mu)}{\sqrt{2C(\theta)m^*}}$, and $N \leq C(\theta)m^*$ by Proposition 7, we get that $\mu^2/(2Nc^2) \geq (1-\mu)^{-2}$. Now, $\mathbb{P}[X \notin \Gamma] = \mathbb{P}[\tau \leq N]$, by Proposition 9, so it follows that

$$\mathbb{P}[\tau \leq N] \geq 1 - \mu - \exp\left(\frac{-1}{(1-\mu)^2}\right) \geq \frac{1-\mu}{2},$$

where the last step uses the elementary inequality $\exp(-1/z^2) \leq z/2$ for $z \in (0,1)$. $\qquad\square$

### 2.1.4 Putting it All Together

*Proof of Lemma 4.* Let us set $N' := N + \sqrt{m^*}/\omega$ for convenience. Observe that by Lemma 5, we are guaranteed a strategy for the standard $\mathcal{D}$-process which constructs $G_{N'}$ such that

$$\mathbb{P}[G_{N'} \in \mathcal{P}] \geq \left(1 - o(1/\sqrt{m^*})\right) \cdot \mathbb{P}[\widetilde{G}_N \in \mathcal{P}].$$

Now, after applying Lemmas 6 and 8, we get that $\mathbb{P}[\widetilde{G}_N \in \mathcal{P}] \geq \mu + \frac{\mu(1-\mu)^2}{\sqrt{8C(\theta)m^*}}$, for $\mu = \mathbb{P}[G_N \in \mathcal{P}]$. On the other hand, $\mathbb{P}[G_N \in \mathcal{P}] \geq \theta$, and $z \to z + \frac{z(1-z)^2}{\sqrt{8C(\theta)m^*}}$ is increasing as a function of $z$ (it is routine to check that it has a positive derivative), so we get that

$$\mathbb{P}[\widetilde{G}_N \in \mathcal{P}] \geq \theta + \frac{\theta(1-\theta)^2}{\sqrt{8C(\theta)m^*}}.$$

However, $C(\theta) := 1 + \log_2\left(\frac{1}{1-\theta}\right)$, so $C(\theta) \leq 1/(1-\theta)^2$ by the elementary inequality $1 + \log_2(z) \leq z^2$ for $z \geq 1$. Thus, $\mathbb{P}[\widetilde{G}_N \in \mathcal{P}] \geq \frac{\theta(1-\theta)^3}{\sqrt{8m^*}}$, and so

$$\mathbb{P}[G_{N'} \in \mathcal{P}] \geq \left(1 - o(1/\sqrt{m^*})\right)\left(\theta + \frac{\theta(1-\theta)^3}{\sqrt{8m^*}}\right) \geq \theta + \frac{\theta(1-\theta)^3}{4\sqrt{m^*}},$$

where the last inequality holds for sufficiently large $n$ (dependent on $\theta_1$ and $\theta_2$). $\qquad\square$

## 3 On Approximately Balanced Martingales

Let $S_0, \ldots, S_k$ be finite sets, and suppose that $X = (X_j)_{j=0}^k$ is a random variable in $S := S_0 \times \cdots \times S_k$, where $S_j := \mathrm{Supp}(X_j)$. Moreover, assume that $M = (M_j)_{j=0}^k$ is a martingale with respect to $(X_j)_{j=0}^k$. Thus, there exists a function $m_j : S_0 \times \cdots \times S_j \to \mathbb{R}$, such that $M_j = m_j(X_0, \ldots, X_j)$. Given a constant $c_j \geq 0$, we say that $M_j$ is **balanced** (with respect to $c_j$), provided for all $(s_0, \ldots, s_j) \in S_0 \times \cdots \times S_j$ and $s'_j \in S_j$,

$$m_j(s_0, \ldots, s'_j) - m_j(s_0, \ldots, s_j) \leq c_j. \tag{9}$$

From the definition of martingale, we get the following:

**Proposition 11.** *If $M_j$ is balanced, then $|M_j - M_{j-1}| \leq c_j$.*

If we are given constants $c = (c_j)_{j=1}^k$, such that each $M_j$ is balanced with respect to $c_j$, then we say that $M$ is **balanced** (with respect to $c$). Observe that if $M$ is balanced, then $|M_j - M_{j-1}| \leq c_j$ for all $j \in [k]$ (i.e., $M$ is $c$-**Lipschitz**). As a result, one can apply the Azuma-Hoeffding inequality to argue that $M_k$ is concentrated about $M_0$. On the other hand, if $M$ is $c$-Lipschitz, then $M$ is $(2c_j)_{j=0}^k$ balanced. Thus, the balanced property is also *necessary* to apply the Azuma-Hoeffding inequality.

This raises the question of what can be done if $M$ is not balanced. We provide a lower tail concentration inequality which depends on the probability *each* $M_j$ satisfies (9) on the randomly chosen point $(X_0, \ldots, X_{j-1}, X_j)$, for all $s'_j \in S_j$. More formally, we say that $(s_0, \ldots, s_k) \in S$ is **stable** with respect to $M$ and $c$, provided for all $1 \leq j \leq k$ and $s'_j \in S_j$,

$$m_j(s_0, \ldots, s'_j) - m_j(s_0, \ldots, s_j) \leq c_j. \tag{10}$$

Define $\Gamma_M \subseteq S$ to be the stable elements of $S$. We measure the balance of $M$ based on the value of $\mathbb{P}[X \in \Gamma_M]$, where $\mathbb{P}[X \in \Gamma_M] = 1$ indicates that $M$ perfectly satisfies the balanced definition.

**Theorem 12.** *Suppose $M = (M_j)_{j=0}^k$ is martingale with respect to a sequence of discrete random variables $X = (X_j)_{j=0}^k$ in $S = S_0 \times \cdots \times S_k$, where $S_j := \mathrm{Supp}(X_j)$. Given constants $c = (c_j)_{j=1}^k$, let $\Gamma_M \subseteq S$ be the stable elements of $S$ with respect to $M$ and $c$. In this case, for any $t \geq 0$,*

$$\mathbb{P}[M_k \leq M_0 - t] \leq \exp\left(\frac{-2t^2}{\sum_{j=1}^k c_j^2}\right) + \mathbb{P}[X \notin \Gamma_M].$$

**Remark 13.** We can derive an upper tail concentration inequality by negating the left-hand side of (10) to modify the definition of $\Gamma_M$. We also note that our approach can be seen as a refinement of the decision tree approach of [8], which was used to prove various concentration inequalities for martingales which are tolerant to "bad" events.

In order to prove Theorem 12, we couple $M = (M_j)_{j=0}^k$ with another martingale $M' = (M'_j)_{j=0}^k$ which is balanced and dominated by $M$ on $\Gamma_M$.

**Lemma 14.** *There exists a coupling of $M$, and another martingale $M' = (M'_j)_{j=0}^k$ with respect to $(X_j)_{j=0}^k$, such that the following conditions hold:*

$(Q_1)$ *'Initial values': $M'_0 = M_0$.*

$(Q_2)$ *'Balanced': $M'$ is balanced with respect to $c_1, \ldots, c_k$.*

$(Q_3)$ *'Domination': If $X \in \Gamma_M$, then $M'_j \leq M_j$ for all $j \in [k]$.*

*Proof.* In order to prove the lemma for $M = (M_j)_{j=0}^k$, we proceed inductively on the value $k$. Firstly, observe that if $k = 0$, then we may set $M'_0 := M_0$, and so the required properties hold trivially. Let us now take $k \geq 1$, and assume that the lemma holds for $k - 1$.

In order to simplify the notation below, let $\mathcal{F}_j = \sigma(X_0, \ldots, X_j)$ be the sigma-algebra generated by $X_0, \ldots, X_j$. It will be convenient to first assume that $X_0$ is constant, so that $\mathcal{F}_0$ is the trivial sigma-algebra (i.e., $\mathcal{F}_0 = \{\emptyset, \Omega\}$), and $M_0$ is constant. Note that then the function $m_j$ satisfies $M_j = m_j(X_1, \ldots, X_j)$ for $1 \leq j \leq k$.

As in the base case, we first set $M'_0 := M_0$ so that $(Q_1)$ is satisfied. In order to define $M'_1$, the high level idea is to modify $M_1$ in such a way that $M'_1$ is balanced, while maintaining the martingale property. This requires us to shift each element's value either up or down. In order to also satisfy the domination property, we must ensure that certain elements are only ever downshifted.

Let us say that $s_1 \in S_1$ is *small*, provided $m_1(s_1) < m_1(s'_1) - c_1$ for some $s'_1 \in S_1$. Let $A$ be the small elements of $S_1$, and $B := S_1 \setminus A$. Observe that if $A \neq \emptyset$, then $B \neq \emptyset$. Moreover, for each $b, b' \in B$, we have that

$$|m_1(b) - m_1(b')| \leq c_1. \tag{11}$$

We refer to $B$ as the *large* elements of $S_1$. Let us proceed with our construction under the assumption that $A \neq \emptyset$, so that $\mathbb{P}[X_1 \in A] > 0$ and $\mathbb{P}[X_1 \in B] > 0$. When $A = \emptyset$, the construction follows easily from the inductive assumption.

Observe that $\mathbb{E}[M_1 \mid X_1 \in A] \leq \max_{b \in B} m_1(b)$ since $m_1(a) \leq m_1(b)$ for all $a \in A, b \in B$. Thus, there exists $\gamma \geq 0$ such that

$$\mathbb{E}[M_1 \mid X_1 \in A] + \frac{\gamma}{\mathbb{P}[X_1 \in A]} + \frac{\gamma}{\mathbb{P}[X_1 \in B]} \in \left[\min_{b \in B} m_1(b), \max_{b \in B} m_1(b)\right],$$

and so

$$\mathbb{E}[M_1 \mid X_1 \in A] + \frac{\gamma}{\mathbb{P}[X_1 \in A]} \in \left[\min_{b \in B} m_1(b) - \frac{\gamma}{\mathbb{P}[X_1 \in B]}, \max_{b \in B} m_1(b) - \frac{\gamma}{\mathbb{P}[X_1 \in B]}\right]. \tag{12}$$

12

Setting $\gamma_A := \gamma/\mathbb{P}[X_1 \in A]$ and $\gamma_B := \gamma/\mathbb{P}[X_1 \in B]$ for convenience, we define

$$M_1' := (\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]} + (M_1 - \gamma_B) \cdot \mathbf{1}_{[X_1 \in B]}. \tag{13}$$

Thus, relative to $M_1$, $M_1'$ lowers the value of each $b \in B$ by $\gamma_B$, and assigns $\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A$ to every $a \in A$. Observe first that because of (11) and (12), we have that $|m_1'(s_1) - m_1'(s_1')| \le c_1$ for each $s_1, s_1' \in S_1$. In addition, observe that

$$\begin{aligned}
\mathbb{E}[M_1' \mid \mathcal{F}_0] &= \mathbb{E}[\,(\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]}] + \mathbb{E}[(M_1 - \gamma_B) \cdot \mathbf{1}_{[X_1 \in B]}] \\
&= \mathbb{E}[M_1 \cdot \mathbf{1}_{[X_1 \in A]}] + \gamma_A \mathbb{P}[X_1 \in A] + \mathbb{E}[M_1 \cdot \mathbf{1}_{[X_1 \in B]}] - \gamma_B \mathbb{P}[X_1 \in B] \\
&= \mathbb{E}[M_1 \cdot (\mathbf{1}_{[X_1 \in A]} + \mathbf{1}_{[X_1 \in B]})] + \gamma - \gamma \\
&= \mathbb{E}[M_1] = M_0,
\end{aligned}$$

where the last line follows from the martingale property of $M_1$. Thus, $\mathbb{E}[M_1' \mid \mathcal{F}_0] = M_0'$, and so $M_1'$ also satisfies the martingale property.

We now construct $(M_j')_{j=2}^k$ and verify the remaining properties. For each $j \in [k]$, let

$$Y_j := (\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]} + (M_j - \gamma_B) \cdot \mathbf{1}_{[X_1 \in B]}.$$

(Note that $Y_1 = M_1'$). We claim that $Y = (Y_j)_{j=1}^k$ is a martingale with respect to $(X_j)_{j=1}^k$. In order to see this, fix $2 \le j \le k$, and take the conditional expectation with respect to $\mathcal{F}_{j-1}$:

$$\begin{aligned}
\mathbb{E}[Y_j \mid \mathcal{F}_{j-1}] &= \mathbb{E}[(M_j - \gamma_B) \cdot \mathbf{1}_{[X_1 \in B]} \mid \mathcal{F}_{j-1}] + \mathbb{E}[\,(\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]} \mid \mathcal{F}_{j-1}] \\
&= \mathbb{E}[(M_j - \gamma_B) \mid \mathcal{F}_{j-1}] \cdot \mathbf{1}_{[X_1 \in B]} + (\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]} \\
&= (M_{j-1} - \gamma_B) \cdot \mathbf{1}_{[X_1 \in B]} + (\mathbb{E}[M_1 \mid X_1 \in A] + \gamma_A) \cdot \mathbf{1}_{[X_1 \in A]} =: Y_{j-1}.
\end{aligned}$$

The first equality follows since the random variables $\mathbf{1}_{[X_1 \in B]}, \mathbf{1}_{[X_1 \in A]}$ and $\mathbb{E}[M_1 + \gamma_A \mid X_1 \in A]$ are determined by $X_1, \ldots, X_{j-1}$ (and thus can be viewed as constants), and the second uses the martingale property of $M$.

Let $\Gamma_Y$ be the stable elements of $S_1 \times \cdots \times S_k$ with respect to $Y$ and $c$. By applying the inductive assumption to $Y$, we get a martingale which can be coupled with $Y$, and whose initial term is $Y_1$. Since $Y_1 = M_1'$, we can denote this martingale unambiguously by $(M_j')_{j=1}^k$. Observe that it has the following properties:

1. $(M_j')_{j=1}^k$ is balanced with respect to $c_2, \ldots, c_k$.

2. If $(X_1, \ldots, X_k) \in \Gamma_Y$, then $M_j' \le Y_j$ for $j = 1, \ldots, k$.

We claim that $M' = (M_j')_{j=0}^k$ is a martingale which satisfies properties $(Q_1)$, $(Q_2)$, and $(Q_3)$. We prove these statements in order.

We have already verified that $\mathbb{E}[M_1' \mid \mathcal{F}_0] = M_0'$. Moreover, $(M_j')_{j=1}^k$ satisfies the martingale property by the inductive assumption. Thus, $M' = (M_j')_{j=0}^k$ is a martingale with respect to $(X_j)_{j=0}^k$.

By construction, $M_0' = M_0$, and so $(Q_1)$ holds. Now, $M_2', \ldots M_k'$ are balanced by (1), and we have already verified that $M_1'$ is balanced. Thus, $M'$ satisfies $(Q_2)$. It remains to verify $(Q_3)$.

We shall first show that if $(X_0, X_1, \ldots, X_k) \in \Gamma_M$, then $(X_1, \ldots, X_k) \in \Gamma_Y$. Now, by the definition of $\Gamma_M$, we have that for each $1 \le j \le k$ and $s_j' \in \mathrm{Supp}(S_j)$,

$$m_j(X_1, \ldots, X_{j-1}, s_j') - m_j(X_1, \ldots, X_j) \le c_j. \tag{14}$$

13

It suffices to show that $y_j(X_1, \ldots, X_{j-1}, s'_j) - y_j(X_1, \ldots, X_j) \leq c_j$, where $y_j$ is the function which satisfies $y_j(X_1, \ldots, X_j) = Y_j$. Observe that if $X_1 \in A$, then $y_j(X_1, \ldots, X_{j-1}, s'_j) = y_j(X_1, \ldots, X_j)$. Otherwise, if $X_1 \in B$, then

$$y_j(X_1, \ldots, X_{j-1}, s'_j) - y_j(X_1, \ldots, X_j) = m_j(X_1, \ldots, X_{j-1}, s'_j) - m_j(X_1, \ldots, X_j) \leq c_j,$$

where the inequality follows from (14). Thus, $(X_1, \ldots, X_k) \in \Gamma_Y$, and so by the inductive assumption, $M'_j \leq Y_j$ for $j = 1, \ldots, k$. On the other hand, since $(X_1, \ldots, X_k) \in \Gamma_M$, $m_1(X_1) \geq m_1(b') - c_1$ for all $b' \in S_1$. Thus, $X_1$ is large (i.e., $X_1 \in B$). It follows that $Y_j = M_j - \gamma_B \leq M_j$, and so $M'_j \leq M_j$ for $j = 1, \ldots, k$, which proves that $(Q_3)$ holds.

To complete the inductive step, we must handle the case when $X_0$ is not necessarily constant. We can handle this by applying the above martingale construction to $(m_j(s_0, X_1 \ldots, X_j))_{j=0}^k$ for each $s_0 \in \mathrm{Supp}(X_0)$. The proof is thus complete. $\square$

*Proof of Theorem 12.* Fix $t \geq 0$, and let $M' = (M'_j)_{j=0}^k$ be the martingale with respect to $(X_j)_{j=1}^k$ guaranteed by Lemma 14. Now, $M'$ is balanced, and so $|M'_j - M'_{j-1}| \leq c_j$ for each $j \in [k]$ by Proposition 11. Thus, we can apply can apply the (one-sided) Azuma-Hoeffding inequality to ensure that

$$\mathbb{P}[M'_k \leq M_0 - t] \leq \exp\left(\frac{-2t^2}{\sum_{j=1}^k c_j^2}\right),$$

where we have used that $M'_0 = M_0$. Returning to $M$, observe that

$$\mathbb{P}[M_k \leq M_0 - t] \leq \mathbb{P}[M_k \leq M_0 - t \text{ and } X \in \Gamma] + \mathbb{P}[X \notin \Gamma].$$

Moreover, if $X \in \Gamma$, then $M'_k \leq M_k$. Thus, $\mathbb{P}[M_k \leq M_0 - t \text{ and } X \in \Gamma] \leq \mathbb{P}[M'_k \leq M_0 - t]$, and so the theorem follows after combining the above equations. $\square$

# 4 Proving Theorem 2

As an application of Theorem 1, we prove that the properties $\mathcal{M}$ and $\mathcal{H}$ admit sharp thresholds in the semi-random graph process. In order to prove this, we first establish the existence of sharp thresholds for the *approximate* properties $\mathcal{M}'$ and $\mathcal{H}'$, and then we transfer these thresholds to $\mathcal{M}$ and $\mathcal{H}$, respectively. Note that many of the definitions (respectively, lemmas) we introduce (respectively, prove) in this section apply to the $\mathcal{D}$-process. While our main application is in proving Theorem 2 – which *is* specific to the semi-random graph process – we develop our techniques for the $\mathcal{D}$-process whenever possible.

Suppose we are given a property $\mathcal{P}$ in the $\mathcal{D}$-process, and $m := m_{\mathcal{P}}(1/2, n)$. We say that $\mathcal{P}'$ is an **approximate property** of $\mathcal{P}$ if $\mathcal{P} \subseteq \mathcal{P}'$, and for any $G_0 \in \mathcal{P}'$, we can play the $\mathcal{D}$-process starting with $G_0$ and obtain a graph in $\mathcal{P}$ in $o(m)$ steps w.h.p.

**Lemma 15.** *Let $\mathcal{P}$ be a property and let $\mathcal{P}'$ be an approximate property of $\mathcal{P}$ in the $\mathcal{D}$-process. If $m^*$ is a sharp threshold for $\mathcal{P}'$, then $m^*$ is also a sharp threshold for $\mathcal{P}$ (in the $\mathcal{D}$-process).*

*Proof of Lemma 15.* Recall that $m := m_{\mathcal{P}}(1/2, n)$ is the minimum number of steps needed to ensure that $\mathcal{P}$ is satisfied with probability at least $1/2$ in the $\mathcal{D}$-process. For any strategy $\mathcal{S}$, since $\mathcal{P} \subseteq \mathcal{P}'$, we have that $T_{\mathcal{S}, \mathcal{P}'} \leq T_{\mathcal{S}, \mathcal{P}}$. Therefore, for any constant $\epsilon > 0$,

$$\mathbb{P}[T_{\mathcal{S}, \mathcal{P}} \leq (1 - \epsilon)m^*] \leq \mathbb{P}[T_{\mathcal{S}, \mathcal{P}'} \leq (1 - \epsilon)m^*] = o(1), \tag{15}$$

and so $m^*$ satisfies the second sharp threshold property for $\mathcal{P}$.

Now, by assumption there exists a strategy $\mathcal{S}'$ for $\mathcal{P}'$ such that $\mathbb{P}[T_{\mathcal{S}',\mathcal{P}'} \leq (1+\epsilon/2)m^*] = 1-o(1)$. Consider the strategy $\mathcal{S}$ for $\mathcal{P}$ that follows $\mathcal{S}'$ until we obtain a graph in $\mathcal{P}'$, and then obtains a graph in $\mathcal{P}$ in $o(m)$ additional steps w.h.p. (this possible from the definition of approximate property). It follows that

$$\mathbb{P}[T_{\mathcal{S},\mathcal{P}} \leq (1+\epsilon/2)m^* + o(m)] \geq (1-o(1))\mathbb{P}[T_{\mathcal{S}',\mathcal{P}'} \leq (1+\epsilon/2)m^*] = 1-o(1). \tag{16}$$

Observe that (16) *almost* establishes $m^*$ satisfies the first sharp threshold property for $\mathcal{P}$, however we must control the $o(m)$ term (note that this term does *not* depend on $\epsilon$). It suffices to show that $m = (1+o(1))m^*$. In order to see this, observe that due to the definition of $m = m_{\mathcal{P}}(1/2, n)$, (15) and (16) imply that for each $\epsilon > 0$,

$$(1-\epsilon)m^* \leq (1-o(1))m \leq (1+\epsilon/2)m^*.$$

Since this holds for each $\epsilon > 0$, $m = (1+o(1))m^*$ as required, and so the proof is complete. $\qquad\square$

Let $\mathcal{M}'$ be the property of having a matching that contains at least $n-n^{0.99}$ vertices, and let $\mathcal{H}'$ be the property of having a path of length at least $n - n^{0.99}$. When restricted to the semi-random graph process, the following "clean-up" algorithm results of Gao et. al. [15, 14] show that $\mathcal{M}'$ and $\mathcal{H}'$ are approximate properties of $\mathcal{M}$ and $\mathcal{H}$, respectively.

**Lemma 16** ([15], and [Lemma 2.5, [14]])**.** *Suppose $G_0$ is a graph with a matching (respectively, a path) that saturates $n - o(n)$ vertices. If we start the semi-random graph process with $G_0$, then there exists a strategy that constructs $G' \in \mathcal{M}$ (respectively, $G' \in \mathcal{H}$) in $o(n)$ steps w.h.p.*

**Remark 17.** We state the quantitative versions of the clean-up algorithms in Appendix A, as these will be useful in the second part of the proof of Theorem 2.

It is not difficult to show that in the semi-random graph process, $\mathcal{M}'$ and $\mathcal{H}'$ are edge-replaceable. Thus, Theorem 1 and Lemma 15 together imply that $\mathcal{M}$ and $\mathcal{H}$ admit sharp thresholds (thus proving the first part of Theorem 2). To prove the second part of Theorem 2, it remains to show that there is a sharp threshold of the form $C_{\mathcal{P}}n$ for both properties. We prove this via an analytic argument in Sections 4.1, 4.2, and 4.3.

In Section 4.4, we show that non-trivial local properties do *not* admit sharp thresholds in the semi-random graph process. Since Theorem 2 confirms that two of the most extensively studied global properties admit sharp thresholds, our results suggest that the dichotomy between thresholds for local and global properties that Friedgut [12] observed for the Erdős–Rényi random graph also applies to the semi-random graph process.

## 4.1 Linear Function as a Sharp Threshold

In this subsection, all of our results apply in the full generality of the $\mathcal{D}$-process. For a given property $\mathcal{P}$, let $I_n(\mathcal{P}) := I_n = \min_{\mathcal{S}_n} \mathbb{E}T_{\mathcal{S}_n}$ where $\mathcal{S}_n$ is taken over all possible strategies. We focus on properties $\mathcal{P}$ with $I_n = \Theta(n)$, which we refer to as **linear** (in $n$). The restriction to the linear regime is a typical feature of results which guarantee the existence of certain limits (see, for example, the interpolation method in [1]). Recall that Theorem 1 ensures that if $\mathcal{P}$ is edge-replaceable, then $\mathcal{P}$ has a sharp threshold. By imposing additional analytic conditions on $(\frac{I_n}{n})_{n\geq 1}$, we can prove the existence of some constant $C > 0$ such that $Cn$ is a sharp threshold for $\mathcal{P}$.

**Theorem 18.** *Let $\alpha > 0$ and $\delta \in (0,1)$ be constants. Suppose that $\mathcal{P}$ is linear, $n^{\alpha}$-edge-replaceable property satisfying the following conditions for all $n$ sufficiently large:*

1. $|I_n - I_{n+1}| < n^\delta$.

2. For all $i \in [n]$ such that $\min(i, n-i) \geq n^\delta$,

$$\frac{I_n}{n} \leq \max\left(\frac{I_i}{i}, \frac{I_{n-i}}{n-i}\right) + O(n^{\delta-1}).$$

Then the limit

$$\lim_{n\to\infty} \frac{I_n}{n} =: C$$

exists. Moreover, $Cn$ is a sharp threshold for $\mathcal{P}$.

In order to prove Theorem 18, we first show that the limit $\lim_{n\to\infty} \frac{I_n}{n}$ exists. We prove a lemma which applies to an arbitrary sequence $(a_n)_n$ of reals which is bounded and which satisfies the analogous conditions 1. and 2. from Theorem 18.

**Lemma 19.** *Let $\delta \in (0,1)$ be a constant. Suppose that $(a_n)_n$ is a bounded sequence of real numbers which satisfies the following conditions for all $n$ sufficiently large:*

$$|na_n - (n+1)a_{n+1}| \leq n^\delta, \tag{17}$$

*and*

$$a_n \leq \max(a_i, a_{n-i}) + O(n^{-\delta}) \tag{18}$$

*whenever $\min(i, n-i) \geq n^{1-\delta}$. Then $\lim_{n\to\infty} a_n$ exists.*

We remark the similarity between Lemma 19 and Fekete's lemma [9], the latter of which states that if for all $i < n$,

$$a_n \leq \frac{i}{n}a_i + \frac{n-i}{n}a_{n-i},$$

then the limit $\lim_{n\to\infty} a_n$ exists. Indeed the proof of both results are quite similar, and so we defer the proof of Lemma 19 to Subsection 4.3.

Before we prove Theorem 18, we require one more technical lemma. Roughly speaking, we shall prove that if $\mathcal{P}$ is edge-replaceable, then there exists a strategy which does nearly as well as $I_n$ with polynomially small failure probability.

**Lemma 20.** *Suppose $\mathcal{P}$ is a linear $n^\alpha$-edge-replaceable property for some fixed $\alpha > 0$. Then there exists a strategy $\mathcal{S}_n^c$ such that*

$$\mathbb{P}(T_{\mathcal{S}_n^c} > I_n + n^{1-\alpha/4}) = O(n^{-\alpha/4}).$$

**Remark 21.** Lemma 20 can be proven by a very careful refinement of Theorem 1 from Section 2. In particular, one would have to allow $\theta_1$ and $\theta_2$ to depend on $n$ and approach 0 and 1, respectively, sufficiently fast as $n \to \infty$. We instead opt for a self-contained proof that is simpler. Note that the argument is similar to the proof of Lemma 5.

*Proof of Lemma 20.* Observe first that the $n^\alpha$-edge-replacement procedure guaranteed by the lemma succeeds with probability $1 - o(1/\sqrt{n})$, and runs for $O(n^{1/2-\alpha})$ steps. By executing this procedure multiple times independently until the first successful run occurs, we get an edge-replacement procedure which takes $O(n^{1/2-\alpha})$ steps in expectation, yet which succeeds with probability 1. We assume we are working with such a procedure in what follows. Let us refer to a graph $G'$ on $[n]$ as **edge-completable**, provided $G' + e \in \mathcal{P}$ for some edge $e$.

16

Let $\mathcal{S}^c = \mathcal{S}_n^c$ be a strategy on $[n]$ that minimizes expected number of steps needed to satisfy $\mathcal{P}$ in the $\mathcal{D}$-process. That is, if $T := T_{\mathcal{S}^c}$, then $\mathbb{E}T = I_n$. Observe that for each $t \geq 1$, $\mathcal{S}^c$ selects $Y_t \in X_t$ which minimizes the expected number of steps to satisfy $\mathcal{P}$, conditional on $X_t$, and the current graph $G_{t-1}$ Thus, we can assume that $\mathcal{S}^c$ is deterministic without loss of generality.

Let $(X_i)_{i=1}^{\infty}$ be the sequence of random subsets the player receives. For convenience, define $\mathcal{H}_j := \sigma(X_1, \ldots, X_j)$ for each $j \geq 1$, and let $\mathcal{H}_0$ be the trivial sigma-algebra. Consider the Doob martingale $Z_j := \mathbb{E}[T \mid \mathcal{H}_j]$ for each $j \geq 0$. We shall first prove that that $|Z_j - Z_{j-1}| = O(n^{1/2-\alpha})$ for each $j \geq 1$. In order to show this, it suffices to argue that for *any* $s_j, s_j' \in \mathrm{Supp}(\mathcal{D})$,

$$\mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j] - \mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j'] = O(n^{1/2-\alpha}). \tag{19}$$

Observe first that since $\mathcal{S}^c$ is deterministic, the event $\{T < j\}$ is $\mathcal{H}_{j-1}$-measureable. Moreover, if $T < j$, then $\mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j] = \mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j']$, so clearly (19) holds when $T < j$.

It remains to prove (19) when $T \geq j$. Our approach is to consider a "stolen" strategy $\mathcal{S}' = \mathcal{S}_n'$ for $\mathcal{P}$ defined on the same sequence $(X_i)_{i=1}^{\infty}$, and which constructs graphs $(G_i')_{i=0}^{\infty}$. We design the strategy so that its "winning time" $T' := T_{\mathcal{S}'}$ is easily compared to $\mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j']$, which will allow us to prove (19).

In each step $i \geq 1$, $\mathcal{S}'$ selects $Y_i' \in X_i$ as follows: If $i \leq j-1$, $\mathcal{S}'$ chooses $Y_i' \in X_i$ in the same way as $\mathcal{S}^c$. If $i = j$, $\mathcal{S}'$ passes on the round (does not select an edge). Else if $i > j$, then there are two cases to consider: While $G_{i-1}'$ is *not* edge-completable, $\mathcal{S}'$ "pretends" that $X_j = s_j'$ when copying the decisions of $\mathcal{S}^c$. That is, $Y_i' \in X_i$ is chosen in the same way as $\mathcal{S}^c$ when given $(X_1, \ldots, s_j', \ldots X_i)$. Once $G_{i-1}'$ becomes edge-completable, $\mathcal{S}'$ executes the edge-replacement procedure and returns the resulting graph.

First observe that since $\mathcal{S}'$ always passes on step $j$,

$$\mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j] = \mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j']. \tag{20}$$

Moreover, we claim that

$$\mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]} = \mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]} + O(n^{1/2-\alpha}). \tag{21}$$

In order to prove (21), we first observe that if $X_j = s_j'$ and $T \geq j$, then $G_T'$ is edge-completable. To see this, note that if $X_j = s_j'$ and $T \geq j$, then we know that $\mathcal{S}^c$ and $\mathcal{S}'$ made decisions in the same way up until step $T$, except for when $\mathcal{S}'$ passed on step $j$. Thus, $G_T'$ has exactly one fewer edge than $G_T$, the graph constructed by $\mathcal{S}^c$ after $T$ steps. Since $G_T \in \mathcal{P}$, $G_T'$ must be edge-completable.

On the other hand, if $G_T'$ is edge-completable, then the definition of $\mathcal{S}'$ implies that $T' \leq T + R$, where $R$ is the number of steps used by the edge-replacement procedure when executed on $G_T'$. Combined with the previous paragraph, we get the following inequality: if $X_j = s_j'$ and $T \geq j$, then $T' \leq T + R$. Since $\{T \geq j\}$ is $\mathcal{H}_{j-1}$-measurable (here $\mathcal{H}_{j-1} = \sigma(X_1, \ldots, X_{j-1})$), we can apply conditional expectations to this inequality to conclude that

$$\mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]} \leq \mathbb{E}[T + R \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]}.$$

But since $T \geq j$, the edge-replacement procedure begins after step $j$, and so it uses edge subsets which are distinct from $(X_i)_{i \leq j}$. Thus, $\mathbb{E}[R \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]} = O(n^{1/2-\alpha})$, and so (21) holds.

To complete the proof of (19) when $T \geq j$, observe that the optimality of $\mathcal{S}^c$ implies that $\mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j] \leq \mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j]$. Thus, by applying (20) followed by (21),

$$\mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j] \cdot \mathbf{1}_{[T \geq j]} \leq \mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j] \cdot \mathbf{1}_{[T \geq j]}$$
$$= \mathbb{E}[T' \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]}$$
$$= \mathbb{E}[T \mid (X_i)_{i<j}, X_j = s_j'] \cdot \mathbf{1}_{[T \geq j]} + O(n^{1/2-\alpha}),$$

which, after rearrangement, is precisely (19) when $T \geq j$.

As we've now verified (19) (for both cases of $T \leq j$ and $T > j$), we know that $|Z_j - Z_{j-1}| = O(n^{1/2-\alpha})$ for each $j \geq 1$. We next apply the Azuma-Hoeffding inequality to $(Z_i)_{i=0}^{\infty}$ to get that for any $\gamma > 0, \beta > 0$,

$$\mathbb{P}(|Z_{\beta n} - \mathbb{E}T| \geq \gamma \mathbb{E}T) \leq \exp\left(-\Theta\left(\frac{\gamma^2 (\mathbb{E}T)^2}{\beta n (n^{1/2-\alpha})^2}\right)\right) \leq \exp\left(-\Theta\left(\frac{\gamma^2 n^{2\alpha}}{\beta}\right)\right).$$

Observe now that since $T$ is a stopping time with respect to $(\mathcal{H}_i)_{i\geq 0}$, the random variable $T \cdot \mathbf{1}_{[T \leq \beta n]}$ is $\mathcal{H}_{\beta n}$-measureable. Thus, if $T \leq \beta n$, then $Z_{\beta n} = T$. By applying Markov's inequality to $T$,

$$\mathbb{P}(T \neq Z_{\beta n}) \leq \mathbb{P}(T > \beta n) \leq \frac{\mathbb{E}T}{\beta n} = O(1/\beta).$$

Combining the two previous equations,

$$\mathbb{P}(|T - \mathbb{E}T| \geq \gamma \mathbb{E}T) \leq \mathbb{P}(T \neq Z_{\beta n}) + \mathbb{P}(|Z_{\beta n} - \mathbb{E}T| \geq \gamma \mathbb{E}T) \leq O(1/\beta) + \exp\left(-\Theta\left(\frac{\gamma^2 n^{2\alpha}}{\beta}\right)\right)$$

Now, let $\gamma = n^{-\alpha/4}, \beta = n^{\alpha/4}$. Since $\mathbb{E}T = I_n$, we get that $\mathbb{P}(T > I_n + n^{1-\alpha/4}) = O(n^{-\alpha/4})$, which proves the theorem. $\qquad\square$

*Proof of Theorem 18.* Let $a_n = I_n/n$ for all $n \geq 1$. Then by Lemma 19, and the assumptions of the theorem, the limit $\lim_{n\to\infty} a_n = C$ exists.

We will now show that $Cn$ is a sharp threshold for $\mathcal{P}$. Fix an arbitrary constant $\epsilon > 0$. Since $C = \lim_{n\to\infty} I_n/n$, we know that for $n$ sufficiently large, $Cn/I_n < (1 + \epsilon)/(1 + \epsilon/2)$. Thus, by Lemma 20, the strategy $\mathcal{S}^c = \mathcal{S}_n^c$ corresponding to $I_n$ satisfies

$$\mathbb{P}(T_{\mathcal{S}_n^c} > (1 + \epsilon)Cn) \leq \mathbb{P}(T_{\mathcal{S}_n^c} > (1 + \epsilon/2)I_n) = o(1),$$

and so the first part of the sharp threshold definition is verified.

We now verify the second part of the sharp threshold definition. Let $\gamma > 0$ be a constant (dependent on $\epsilon$) chosen sufficiently small such that

$$(1 - \epsilon)\frac{1 + \gamma}{1 - \gamma} < 1. \tag{22}$$

By definition, there exists a strategy which wins in $m(1 - \gamma, n)$ steps with probability $1 - \gamma$. In the case of failure, we can execute $\mathcal{S}^c$ for an additional $I_n$ steps in expectation. By optimality of $\mathcal{S}^c$, this implies that

$$I_n \leq (1 - \gamma)m(1 - \gamma, n) + \gamma \cdot (m(1 - \gamma, n) + I_n),$$

and so $I_n \leq \frac{1}{1-\gamma}m(1 - \gamma, n)$ after rearrangement. On the other hand, for $n$ sufficiently large we have that $Cn/I_n \leq (1 + \gamma)$. Therefore, by applying both inequalities,

$$(1 - \epsilon)Cn \leq (1 - \epsilon)(1 + \gamma)I_n \leq (1 - \epsilon)\frac{1 + \gamma}{1 - \gamma}m(1 - \gamma, n). \tag{23}$$

On the other hand, since $\mathcal{P}$ is edge-replaceable, we can apply Theorem 1 to argue that $m(1 - \gamma, n)$ is a sharp threshold for $\mathcal{P}$. Therefore, by applying (22) and (23), we have that for any strategy $\mathcal{S}_n$,

$$\mathbb{P}(T_{\mathcal{S}_n} < (1 - \epsilon)Cn) \leq \mathbb{P}\left(T_{\mathcal{S}_n} < (1 - \epsilon)\frac{1 + \gamma}{1 - \gamma}m(1 - \gamma, n)\right) = o(1).$$

This establishes the second part of the sharp threshold definition, and so the proof is complete. $\quad\square$

## 4.2 Proving Theorem 2

In this section, we restrict our attention to the semi-random graph process and complete the proof of Theorem 2. (The definitions we use specific to this process are introduced in Subsection 1.2). Recall that $\mathcal{M}'$ and $\mathcal{H}'$ are the approximate properties of $\mathcal{M}$ and $\mathcal{H}$. Our approach is to verify that $\mathcal{M}'$ and $\mathcal{H}'$ satisfy the conditions of Theorem 18. This will complete the proof of Theorem 2 due to Lemma 15.

While verifying condition 1. of Theorem 18 for a property $\mathcal{P}$ is straightforward, verifying condition 2. is more involved. Because of this, we first define a condition which is more easily verified.

**Definition 3** (Splittable). Let $\omega = \omega(n)$. We refer to $\mathcal{P}$ as $\omega$-**splittable** (or **splittable** if clear), provided the following guarantee holds: Suppose $i \in [n]$ and $G_0$ is an arbitrary graph on $[n]$. If the induced graphs $G_0[1, \ldots, i]$ and $G_0[i+1, \ldots, n]$ each satisfy $\mathcal{P}$, then there is a strategy with initial graph $G_0$ which satisfies $\mathcal{P}$ after $\omega$ steps in expectation.

We now state a technical lemma which allows us to establish condition 2. from Theorem 18.

**Lemma 22.** *Suppose $\mathcal{P}$ is a linear property which is $n^\alpha$-edge-replaceable and $n^\beta$-splittable for $\alpha, \beta \in (0, 1)$. Then there exists $\delta \in (0, 1)$ such that*

$$\frac{I_n}{n} \le \max\left(\frac{I_i}{i}, \frac{I_{n-i}}{n-i}\right) + O(n^{\delta - 1}) \tag{24}$$

*for all $i \in [n]$ such that $\min(i, n - i) \ge n^\delta$.*

*Proof.* In order to prove the lemma, we describe a strategy $\mathcal{S}$ to be played on $[n]$ whose expected number of steps is upper bounded by (24).

First, partition $[n]$ into $A = \{1, \ldots, i\}$ and $B = \{i+1, \ldots, n\}$. Let $\mathcal{S}_A^c$ (respectively, $\mathcal{S}_B^c$) be the strategy on vertex set $A$ (respectively, $B$) guaranteed from Lemma 20 due to the $n^\alpha$-edge replacable assumption. Define

$$N := \max\left(\frac{n}{i} I_i, \frac{n}{n-i} I_{n-i}\right) + n^{1-x}$$

where $\delta \in (0, 1)$ is a constant to be specified later, $\delta_1 := 1 - \alpha/4$, and $x := \delta(1 - \delta_1)/2$. During the first $N$ steps, we define $\mathcal{S}$ to essentially play two games at once: Each time we are given a square in $A$, we choose a circle of $A$ via strategy $\mathcal{S}_A^c$, and similarly if we are given a square in $B$, we choose a circle of $B$ via strategy $\mathcal{S}_B^c$.

For $i \ge n^\delta$ the number of steps where we play on $A$ is $\mathrm{Bin}(N, i/n)$, so

$$\mathbb{P}(\mathrm{Bin}(N, i/n) \le I_i + i^{\delta_1}) \le \exp\left(-\Theta\left(\frac{(n^{-x}i - i^{\delta_1})^2}{Ni/n}\right)\right)$$

$$\le \exp\left(-\Theta\left(\frac{(n^{-x}i)^2}{i}\right)\right) = \exp(-\Theta(n^{\delta - 2x})) = O(1/n).$$

where the first inequality follows from Chernoff bound, the second inequality follows from $i^{\delta_1} \ll n^{-x}i$, the first equality follows from $i \ge n^\delta$ and the last equality follows from $\delta > 2x$. Therefore, by Lemma 20 the probability that we did not finish the game on $A$ is at most $O(1/n + i^{\delta_1 - 1})$. Similarly when $n - i \ge n^\delta$, the probability that we did not finish the game on $B$ is at most $O(1/n + (n-i)^{\delta_1 - 1})$. If we finish the game on $A$ and $B$, then since $\mathcal{P}$ is $n^\beta$-splittable, we can construct a graph in $\mathcal{P}$ on the vertex set $[n]$ in $n^\beta$ expected steps. Otherwise, using the linearity assumption, we just play the game as if the graph is empty and finish in an additional $I_n = O(n)$ expected steps.

19

Therefore, for an appropriate choice of $\delta \in (0,1)$ sufficiently close to 1, if $\min(i, n-i) \geq n^\delta$ then the total expected number of steps for our strategy is at most

$$N + n^\delta + I_n \cdot O\left(\frac{1}{n} + i^{\delta_1 - 1} + (n-i)^{\delta_1 - 1}\right) \leq \max\left(\frac{nI_i}{i}, \frac{nI_{n-i}}{n-i}\right) + O(n^\delta),$$

which establishes (24). □

We are now ready to prove Theorem 2. As mentioned, the proof relies on the quantitative clean-up algorithms of Gao et al. (see Appendix A).

*Proof of Theorem 2.* We will show that the property $\mathcal{M}'$ satisfies the condition of Theorem 18. First, it is a linear property since it is known that $I_n/n \in [1/2, 2]$ (as first observed in [5]).

Next, we verify that it is edge-replaceable. For any matching that saturates less than $n - n^{0.99}$ vertices, with probability $1 - o(1/\sqrt{n})$, one of the given squares will land on an unsaturated vertex in at most $n^{0.02}$ steps, and we can then form an edge between two unsaturated vertices to form a larger matching. Therefore, $\mathcal{M}'$ is $n^{0.48}$-edge-replaceable and we can take $\alpha = 0.48$.

Condition 1. of Theorem 18 is routine to check, but we include the argument here for the sake of completeness. To show $I_{n+1} < I_n + n^\delta$ for any $\delta > 0.01$ and large enough $n$, we simply use the strategy that obtains a matching that saturates $n - n^{0.99}$ vertices on the first $n$ vertices in $I_n + O(1)$ expected steps (there are $O(1)$ expected steps where the given square is vertex $n+1$). To obtain a matching that saturates $n + 1 - (n+1)^{0.99}$ vertices, we simply wait until we are given a square on a vertex that is not saturated by a matching on $n - n^{0.99}$ vertices, which happens in $O(n^{0.01})$ expected steps. Similarly we can show $I_n < I_{n+1} + n^\delta$ by analyzing the optimal strategy that obtains $I_{n+1}$, while ignoring steps that involved vertex $n+1$.

Next, we check condition 2. of Theorem 18 by verifying $\mathcal{M}'$ is $n^\beta$-splittable for some $\beta \in (0,1)$. To prove this, we apply Lemma 27 of Appendix A (the quantitative version of the perfect matching clean-up algorithm). Observe that given any $i \in [n]$ from Definition 3, we begin with an initial matching of which contains at least $n - 2n^{0.99}$ vertices. By applying Lemma 27 with $\epsilon = 1/n^{0.01}$, we can recover a perfect matching in $O(n^{1-0.005})$ steps in expectation. Thus, any $\beta > 1 - 0.005$ suffices. By Lemma 22, we are guaranteed some $\delta \in (0,1)$ for which condition 1. of Theorem 18 is satisfied.

Therefore $\lim_{n \to \infty} I_n/n =: C_\mathcal{M}$ exists and $C_\mathcal{M} n$ is a sharp threshold for the property $\mathcal{M}'$. Since $\mathcal{M}'$ is an approximate property of $\mathcal{M}$, it follows from Lemma 15 that $C_\mathcal{M} n$ is a sharp threshold for $\mathcal{M}$.

The proof that $\mathcal{H}'$ satisfies the condition of Theorem 18 is similar, and we will only sketch the argument. First, it is a linear property since $I_n/n \in [1,3]$ (as first observed in [5]). We will show that if $G_0$ contains a vertex-disjoint union of 2 paths $P_1, P_2$ with total length $\ell - 1 := n - n^{0.99} - 1$, then we can obtain a path $P$ with length $\ell$ in $O(n^{2/5})$ steps. Without loss of generality, suppose that $P_1$ is the longer path. It is routine to check that in $n^{2/5}$ steps, with probability $1 - o(1/\sqrt{n})$, we will receive 2 squares with distance at most $n^{1/4}$ in $P_1$. Therefore, we can consider the strategy that matches all given squares in $P_1$ with one endpoint of $P_2$ if all previous squares are of distance at least $n^{1/4}$ in $P_1$, and then match the first square in $P_1$ that does not satisfy that property with the other endpoint of $P_2$. This strategy will construct a path of length at least $\ell - 1 - n^{1/4}$ in $O(n^{2/5})$ steps with probability at least $1 - o(1/\sqrt{n})$. To extend this to a path of length $\ell$, we simply attached any given unsaturated square with an endpoint of our path. We need to do so $n^{1/4} + 1$ times and the expected number of round between receiving unsaturated squares is at most $O(n^{0.01})$. It routinely follows that in, say, $n^{0.27}$ steps we can obtain a path of length $\ell$ with probability at least $1 - o(1/\sqrt{n})$. Therefore, $\mathcal{H}'$ is $n^{0.1}$-edge-replaceable (with room to spare).

By an argument similar to the one seen for $\mathcal{M}'$, it can be seen that condition 1. of Theorem 18 holds. Similarly, using Lemma 26 of Appendix A (quantitative version of Hamiltonian cycle clean-up algorithm), $\mathcal{H}'$ is splittable, and thus satisfies condition 2. Therefore, by Theorem 18 and Lemma 15, there exists a constant $C_{\mathcal{H}} > 0$ such that $C_{\mathcal{H}} n$ is a sharp threshold for $\mathcal{H}$. $\qquad\square$

## 4.3 Proving Lemma 19

*Proof.* Let $\liminf_{n\to\infty} a_n = L$. For convenience let $a_n = a_{\lfloor n \rfloor}$ for non-integer $n$. We note that the inequalities are still true when $i, n$ are not integers (possibly changing $\delta$ and the implicit constants if necessary). Given any $\epsilon > 0$, we will pick some large $N_0$ and $k$ to be specified later. Then for any $N \geq k N_0$, we get that there exists an integer $r$ and $i \in \{0, \ldots, k-1\}$ such that

$$(k+i)N_0 \leq \frac{N}{2^r} \leq (k+i+1)N_0$$

We will then show that

$$a_{N_0} \leq L + \epsilon \tag{25}$$

$$a_N \leq a_{N/2^r} + O(N_0^{-\delta}) \tag{26}$$

$$a_{N/2^r} \leq a_{(k+i+1)N_0}(1 + O(1/k)) + O\left(\frac{N_0^{\delta}}{k^{1-\delta}}\right) \tag{27}$$

$$a_{(k+i+1)N_0} \leq a_{N_0} + O((\log k)N_0^{-\delta}) \tag{28}$$

By picking $N_0$ large enough which satisfies (17), (18), (25) (which is possible by definition of $L$) and $k = N_0^{(1+\delta)/(1-\delta)}$, combining all four inequalities gives us

$$a_N \leq L + 2\epsilon$$

for all $N \geq k N_0$.

To prove (26), we simply use (18) iteratively, dividing the current index by 2 each time to get that

$$a_{N/2^\ell} \leq a_{N/2^{\ell+1}} + O((N/2^\ell)^{-\delta})$$

for all $\ell = 0, \ldots, r-1$. Hence

$$a_N \leq a_{N/2^r} + O\left(\sum_{0 \leq \ell \leq r-1} (N/2^\ell)^{-\delta}\right) = a_{N/2^r} + O\left(\sum_{i \geq 0} (N_0 2^i)^{-\delta}\right) = a_{N/2^r} + O(N_0^{-\delta}).$$

To prove (27), we will use (17) iteratively. We have

$$a_{(k+i+1)N_0-1} \leq \frac{(k+i+1)N_0}{(k+i+1)N_0 - 1} a_{(k+i+1)N_0} + ((k+i+1)N_0 - 1)^{\delta-1}.$$

By iterating this $t$ times where $t = (k+i+1)N_0 - N/2^r \leq N_0$, we have

$$a_{(k+i+1)N_0-t} \leq \frac{(k+i+1)N_0}{(k+i+1)N_0 - t} a_{(k+i+1)N_0} + \frac{\sum_{1 \leq j \leq t}((k+i+1)N_0 - j)^{\delta}}{(k+i+1)N_0 - t}$$

$$\leq a_{(k+i+1)N_0}(1 + O(1/k)) + O\left(\frac{N_0^{\delta}}{k^{1-\delta}}\right)$$

21

We will now prove (28). For any natural number $j$, we will invoke the following inequalities derived from (18) : if $j$ is even, then

$$a_{jN_0} \le a_{jN_0/2} + O(N_0^{-\delta}),$$

and if $j$ is odd, then

$$a_{jN_0} \le \max(a_{(j-1)N_0}, a_{N_0}) + O(N_0^{-\delta}).$$

We will start with $j_0 = k + i + 1$. Given $j_i$, if $j_i$ is odd let $j_{i+1} = j_i - 1$, and if $j_i$ is even let $j_{i+1} = j_i/2$. Clearly $j_{\lceil \log(2k) \rceil} = 0$. We get that

$$a_{j_\ell N_0} \le \max(a_{j_{\ell+1}N_0}, a_{N_0}) + \ell O(N_0^{-\delta})$$

Therefore

$$a_{j_0 N_0} \le a_{N_0} + \log(2k) O(N_0^{-\delta})$$

as desired. $\qquad\square$

## 4.4 Sharp Thresholds Do Not Exist for Local Properties

In this section we make some brief observations regarding thresholds for local properties. Given a list of fixed graphs, none of which are forests, we prove that a sharp threshold does *not* exist for the property of containing at least one of these fixed graphs. The results in this section follow from the below result of Behague et al. [2], and we include the proofs for completeness.

**Theorem 23** (Theorem 1.2 of [2])**.** *Let $H$ be a fixed subgraph of degeneracy $d \ge 2$. Then for any strategy $\mathcal{S}_n$, if $T_{\mathcal{S}_n}$ is the number of rounds needed for $\mathcal{S}_n$ to build a copy of $H$, then*

$$\mathbb{P}(T_{\mathcal{S}_n} \le n^{(d-1)/d}/\omega) = o(1)$$

*for any $\omega \to \infty$.*

We now prove the following theorem, which implies Theorem 3.

**Theorem 24.** *Let $L$ be a fixed (finite) list of fixed graphs, none of which are forests. Suppose $d$ is the minimum degeneracy of graphs in $L$. Let $\mathcal{P}$ be the property of containing a graph in $L$. Then*

1. *There exists constant $\alpha$ such that for any strategy $\mathcal{S}$,*

$$\mathbb{P}(G^{\mathcal{S}}_{\alpha n^{(d-1)/d}} \in \mathcal{P}) \le 1/2$$

2. *For any constant $\beta$, there exists a constant $\delta = \delta(\beta) > 0$ and a strategy $\mathcal{S}'$ such that*

$$\mathbb{P}(G^{\mathcal{S}'}_{\beta n^{(d-1)/d}} \in \mathcal{P}) \ge \delta$$

*Proof.* To prove 1., observe that by Theorem 23 there must exist some $\alpha > 0$ such that for any graph $H \in L$ and any strategy $\mathcal{S}$, we have that

$$\mathbb{P}(H \in G^{\mathcal{S}}_{\alpha n^{(d-1)/d}}) \le \frac{1}{2|L|}.$$

The proof now immediately follows from an application of union bound over all $H \in L$.

Proving 2. is slightly more involved. Fix a graph $H$ of $L$ of degeneracy $d$. We first describe the strategy $\mathcal{S}'$ from [2]. Since $H$ is $d$-degenerate, we may consider an ordering of the vertices of $H$ (say, $(v_1, \ldots, v_k)$) such that $v_i$ has at most $d$ neighbours in $\{v_1, \ldots, v_{i-1}\}$. We divide the game into $k$ phases, where in phase $i$ we build the induced graph $H[v_1, \ldots, v_i]$. Suppose $v_i$ is adjacent to $v_{i_1}, \ldots, v_{i_\ell}$ where $\ell \leq d$, and we have a copy of $H[v_1, \ldots, v_{i-1}]$. To complete phase $i$, if a vertex $v \neq v_1, \ldots, v_{i-1}$ is given as a square for the $j^{th}$ time, we match it with $v_{i_j}$. Therefore, if one of such $v$ is given as a square at least $d$ times, then we have successfully built a copy of $H[v_1, \ldots, v_i]$. By a standard analysis, this succeeds in $\frac{\beta}{k} n^{(d-1)/d}$ steps with probability $\Omega(1)$.

Therefore, the probability that $\mathcal{S}'$ constructs $H$ in $\beta n^{(d-1)/d}$ steps is $\Omega(1)$, and so the proof is complete. $\qquad\square$

*Proof of Theorem 3.* For the sake of contradiction, suppose that a sharp threshold $m^* = m^*(n)$ exists. Let $\mathcal{P}$ be the property of containing a graph in $L$. By the definition of sharp threshold, we have that for any strategy $\mathcal{S}_n$,

$$\mathbb{P}(G^{\mathcal{S}}_{m^*/2} \in \mathcal{P}) = o(1).$$

Thus, taking $\alpha$ and $\beta := \alpha/4$ as in Theorem 24, the second part of Theorem 24 implies that $m^*/2 \leq \alpha n^{(d-1)/d}/4$. From the definition of sharp threshold, we also know that there exists a strategy $\mathcal{S}'_n$ such that

$$\mathbb{P}(G^{\mathcal{S}'_n}_{2m^*} \in \mathcal{P}) = 1 - o(1).$$

But $2m^* \leq \alpha n^{(d-1)/d}$, and so by following the strategy of $\mathcal{S}'_n$,

$$1 - o(1) = \mathbb{P}(G^{\mathcal{S}'_n}_{2m^*} \in \mathcal{P}) \leq \mathbb{P}(G^{\mathcal{S}'_n}_{\alpha n^{(d-1)/d}} \in \mathcal{P}) \leq 1/2,$$

where the final inequality follows from the definition of $\alpha$ in first part of Theorem 24. This is a contradiction for sufficiently large $n$, and so the proof is complete. $\qquad\square$

**Remark 25.** The case $d = 1$ i.e. $H$ is a forest, is trivial. If $H$ is a fixed graph of degeneracy $d = 1$, then for any strategy $\mathcal{S}_n$

$$\mathbb{P}(T_{\mathcal{S}_n} < |V(H)| - 1) = o(1),$$

and there exists a strategy $\mathcal{S}_n$ such that

$$\mathbb{P}(T_{\mathcal{S}_n} = |V(H)| - 1) = 1 - o(1).$$

# 5 Conclusion

Our result allows us to prove the existence of sharp thresholds for edge-replaceable properties in adaptive random graph processes. As we have seen in this paper, being edge-replaceable is a strong enough restriction that natural properties such as $\mathcal{M}$ and $\mathcal{H}$ do not satisfy it in the semi-random graph process. We resolved this issue by proving that the properties of interest can be approximated by weaker properties which *are* edge-replaceable.

It would be of great interest if one can develop more powerful tools to establish sharp thresholds in adaptive random graph processes (or adaptive random processes in general) when edge-replaceable properties do not hold even in the approximate sense. A starting point would be to fully resolve the following problem proposed by Ben-Eliezer et al. [4]:

**Question 1.** *For all $r \geq 2$, does the property of having a $K_r$-factor admit a sharp threshold in the semi-random graph process?*

We resolved this for $r = 2$, and it would be interesting to solve this for all constant $r$ or $r = r(n)$ which grows slowly with $n$. More generally, one could ask for a sharp threshold result for the property of containing a certain spanning graph with large minimum degree. This may require new techniques, as it seems like such properties are generally not edge-replaceable, even in the approximate sense.

For each property $\mathcal{P} \in \{\mathcal{M}, \mathcal{H}\}$, we have shown that there exists a constant $C_{\mathcal{P}}$ such that $C_{\mathcal{P}}n$ is a sharp threshold in the semi-random graph process. That being said, currently only upper and lower bounds are known for $C_{\mathcal{P}}$ (as implied in [15, 14]).

**Question 2.** *What is the exact value of $C_{\mathcal{P}}$ in Theorem 2?*

This question currently appears out of reach, as it seems to necessitate designing an asymptotically optimal strategy for $\mathcal{P}$. Our sharp threshold results indicate that in order to identify $C_{\mathcal{P}}$, it suffices to find an optimal strategy which satisfies $\mathcal{P}$ with (small) constant probability. We hope that this reduction may prove useful in later works.

### Acknowledgement.

# References

[1] 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.

[2] Natalie C Behague, Trent G Marbach, Pawel Pralat, and Andrzej Rucinski. Subgraph games in the semi-random graph process and its generalization to hypergraphs. *arXiv preprint arXiv:2105.07034*, 2021.

[3] Omri Ben-Eliezer, Lior Gishboliner, Dan Hefetz, and Michael Krivelevich. Very fast construction of bounded-degree spanning graphs via the semi-random graph process. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 718–737. SIAM, 2020.

[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. *Random Structures & Algorithms*, 57(4):892–919, 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] Tom Bohman and Alan Frieze. Avoiding a giant component. *Random Structures & Algorithms*, 19(1):75–85, 2001.

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

[8] Fan Chung and Linyuan Lu. Concentration inequalities and martingale inequalities: a survey. *Internet mathematics*, 3(1):79–127, 2006.

[9] Nicolaas Govert de Bruijn and Paul Erdös. Some linear and some quadratic recursion formulas. ii. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen: Series A: Mathematical Sciences*, 14:152–163, 1952.

[10] Jian Ding, Allan Sly, and Nike Sun. Proof of the satisfiability conjecture for large k. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 59–68, 2015.

[11] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[12] Ehud Friedgut and Jean Bourgain. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999.

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

[14] Pu Gao, Calum MacRury, and Paweł Prałat. 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)*, volume 245 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 29:1–29:22, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

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

[16] Michael Krivelevich, Eyal Lubetzky, and Benny Sudakov. Hamiltonicity thresholds in achlioptas processes. *Random Structures & Algorithms*, 37(1):1–24, 2010.

[17] Calum MacRury. *Online Decision-Making in a Randomized Environment: Stochastic Matching and Adaptive Random Graph Processes*. PhD thesis, University of Toronto, 2023.

# A    Clean-up Algorithms

In this section, we state the explicit guarantees of the clean-up algorithms introduced by Gao et al. [14, 15]. Note that these algorithms were originally proven to hold w.h.p. in a fixed number of steps (i.e., a Monte Carlo algorithm). By executing independent runs of such an algorithm until the first success occurs, we can ensure that the clean-up succeeds with probability 1 while getting the same asymptotic upper bound on the *expected* number of steps (i.e., we convert to a Las Vegas algorithm). We make use of these expectation bounds in the proof of Theorem 2.

## A.1    Hamiltonian Cycles

**Lemma 26** (Lemma 2.5, [14])**.** *Let $0 < \epsilon = \epsilon(n) < 1/1000$, and suppose that $P$ is a path on $(1 - \epsilon)n$ vertices of $[n]$. Then, given $P$ initially, there exists a strategy for the semi-random graph process which builds a Hamiltonian cycle from $P$ in $O(\sqrt{\epsilon}n + n^{3/4}\log^2 n)$ steps w.h.p. Note that the constants hidden in the $O(\cdot)$ notation does not depend on $\epsilon$.*

## A.2 Perfect Matchings

Gao et al. [15] provide a clean-up algorithm with the following guarantee. For $\epsilon = 10^{-14}$, if the algorithm is presented a matching $M$ on at least $(1 - \epsilon)n$ vertices of $[n]$, then $M$ can be extended to a perfect matching in at most $100\sqrt{\epsilon}n$ steps w.h.p. The analysis of [15] holds when $\epsilon = \epsilon(n)$ satisfies $\epsilon(n) \to 0$ as $n \to \infty$, however an additional term must be added to $100\sqrt{\epsilon}n$ if $\epsilon$ tends to 0 sufficiently fast. This is explicitly proven in the first author's thesis [17], and we restate the lemma below for convenience.

**Lemma 27** (Lemma 5.1.8, [17]). *Let $0 < \epsilon = \epsilon(n) < 1$, and suppose that $M$ is a matching on $(1 - \epsilon)n$ vertices of $[n]$. Then, given $M$ initially, there exists a strategy for the semi-random graph process which builds a perfect matching from $M$ in $O(\sqrt{\epsilon}n + n^{3/4}\log^2 n)$ steps w.h.p. Note that the constants hidden in the $O(\cdot)$ notation does not depend on $\epsilon$.*