Offline thresholds for Ramsey-type games on random graphs

Michael Krivelevich* Reto Spöhel† Angelika Steger

School of Mathematical Sciences
Tel Aviv University
69978 Tel Aviv, Israel
krivelev@post.tau.ac.il

Institute of Theoretical Computer Science
ETH Zürich
8092 Zürich, Switzerland
{rspoehel|steger}@inf.ethz.ch

ABSTRACT. In this paper, we compare the offline versions of three Ramsey-type one-player games that have been studied in an online setting in previous work: the online Ramsey game, the balanced online Ramsey game, and the Achlioptas game. The goal in all games is to color the edges of the random graph $G_{n,m}$ according to certain rules without creating a monochromatic copy of some fixed forbidden graph H. While in general the three online games have different thresholds, we prove that for most graphs H, the offline threshold for all three problems is $m_0(n) = n^{2-1/m_2(H)}$, where $m_2(H) := \max_{H' \subseteq H}(e_{H'} - 1)/(v_{H'} - 2)$.

1. Introduction

The motivation for this work comes from three Ramsey-type one-player games that have been studied in an online setting in previous work: the online Ramsey game [2, 9, 10], the balanced online Ramsey game [7, 12], and the Achlioptas game [5, 11]. In all three games, the edges of the complete graph K_n appear in a random order, either one by one or in batches of some fixed size. In each step of the game, the player has to color the new edges immediately and irrevocably, according to certain rules and in particular *without creating a monochromatic copy of some fixed forbidden graph H*. The question we are interested in is how long the player can 'survive' in a given online game, i.e., how many random edges she can color without creating a monochromatic copy of H. The 'typical' number of edges the player is able to color in a given game (using an appropriate coloring strategy) is called the *threshold* of the game (cf. below for a precise definition), and the main goal when investigating these games is to determine their thresholds asymptotically as a function of n.

We use the word *online* to emphasize the fact that in each step, the player has to decide how to color the new edges *before* seeing the random edges that appear later in the game. In this paper, we investigate what happens if the player is allowed to 'see into the future' or, more precisely, is given a large number of random edges *all at once* and is asked to color them subject to the same rules as before. Throughout, we refer to this as the *offline version* of a given game (or as the offline problem corresponding to a given game). Note that when studying these offline versions we are investigating colorability properties of *static* random objects.

Before describing the three games in detail and giving the technical definitions needed to state our results precisely, we summarize our findings as follows: While in general the three online games have different thresholds, for most graphs H, the offline versions of all three games have the same threshold, namely $m_0(n) = n^{2-1/m_2(H)}$, where

$$m_2(H) := \max_{H' \subseteq H} \frac{e_{H'} - 1}{v_{H'} - 2}$$

(cf. Theorems 2, 4 and 5 below).

*Research supported in part by a USA-Israel BSF grant, by a grant from the Israel Science Foundation, and by a Pazy Memorial Award.

†Research supported by Swiss National Science Foundation, grant 200020-119918.

1.1. **Online Ramsey game.** Consider the following one-player game. Starting with the empty graph on n vertices, in every step a new edge is drawn uniformly at random from all non-edges and inserted into the current graph. This edge has to be colored immediately with one of r available colors, where $r \geq 2$ is a fixed integer. The player's goal is to avoid creating a monochromatic copy of some fixed graph H for as long as possible. We will refer to this game in the following as the *online H-avoidance game (with r colors)*. As usual, we use the phrase *asympotically almost surely (a.a.s.)* to indicate that some statement holds with probability $1 - o(1)$ as $n \to \infty$. We say that $N_0 = N_0(H, r, n)$ is a *threshold* for the game if, on the one hand, there exists a strategy such that for any function $N \ll N_0$, the player a.a.s. can play for N steps following this strategy, and if, on the other hand, for any function $N \gg N_0$ the player a.a.s. loses the game after at most N steps, regardless of her strategy. This game was introduced in [2] for the case $H = K_3$ and $r = 2$ colors. In [9, 10], the following theorem was shown.

Theorem 1 ([9, 10])**.** *Let H be a non-forest that has a subgraph $H_- \subset H$ with $e_H - 1$ edges satisfying*

$$m_2(H_-) \leq \overline{m}_2^2(H) \ ,$$

where

$$\overline{m}_2^2(H) := \max_{H' \subseteq H} \frac{e_{H'}}{v_{H'} - 2 + 1/m(H)}$$

and

$$m(H) := \max_{H' \subseteq H} \frac{e_{H'}}{v_{H'}} \ .$$

Then the threshold for the online H-avoidance game with two colors is

$$N_0(H, 2, n) = n^{2 - 1/\overline{m}_2^2(H)} \ .$$

This result applies in particular to cliques and cycles of arbitrary size. However, in general $n^{2-1/\overline{m}_2^2(H)}$ is only a lower bound for the threshold of the online H-avoidance game with two colors – if, e.g., H is the graph consisting of two triangles overlapping in exactly one vertex, the threshold is strictly higher (see [9]).

It seems plausible that e.g. for H a clique or cycle, the threshold for the game with r colors is

$$N_0(H, r, n) = n^{2 - 1/\overline{m}_2^r(H)} \ ,$$

where $\overline{m}_2^r(H)$ is inductively defined as

$$\overline{m}_2^r(H) := \max_{H' \subseteq H} \frac{e_{H'}}{v_{H'} - 2 + 1/\overline{m}_2^{r-1}(H)} \ .$$

This would be in line with known results for the corresponding vertex-coloring problem [8]. (In fact, it was shown in [9] that for any $r \geq 2$ and any non-forest H, $n^{2-1/\overline{m}_2^r(H)}$ is indeed a lower bound for the threshold of the game with r colors. However, the upper bound proof presented in [10] does not extend to the game with more than two colors.)

Note that, by symmetry, after N steps the board of the online Ramsey game is distributed uniformly over all graphs on n vertices with exactly N edges. Thus the corresponding offline problem is the following: Given a random graph $G_{n,m}$ (a graph drawn uniformly at random from all graphs on n vertices with m edges), is there an edge-coloring avoiding monochromatic copies of H? This question is answered by a classical result of Rödl and Ruciński. For any pair of graphs G and H, let $G \to (H)^e_r$ denote the property that every r-coloring of the edges of G contains a monochromatic copy of H. In this notation, the result of Rödl and Ruciński reads as follows.

Theorem 2 ([6, 13, 14])**.** *Let $r \geq 2$, and let H be a non-forest. Then there exist positive constants $c = c(H,r)$ and $C = C(H,r)$ such that*

$$\lim_{n\to\infty} \Pr[G_{n,m} \to (H)^e_r] = \begin{cases} 0, & m \leq cn^{2-1/m_2(H)} \\ 1, & m \geq Cn^{2-1/m_2(H)} \end{cases} . \tag{1}$$

In this paper, we prove results of a similar flavor for the offline problems corresponding to two variants of the online game that we will introduce now. In the following, we will refer to the setting of Theorem 2 as the classical case.

1.2. **Balanced online Ramsey game.** This balanced variant of the online Ramsey game is similar to the unbalanced game except that in each step, a set of r new edges is drawn uniformly at random (from all non-edges as before) and presented to the player. The player has to color these edges immediately subject to the restriction that each of the r available colors is used for exactly one of these edges. Intuitively, this makes the player's task more difficult than in the unbalanced game. Thus we expect that the thresholds for this balanced game do not exceed those of the unbalanced game, and indeed this turns out to be the case. We will refer to this game in the following as the *balanced online H-avoidance game (with r colors)*.

Extending results from [7], the following theorem was shown in [12].

Theorem 3 ([12])**.** *Let $r \geq 2$, and let H be a non-forest that has a subgraph $H_- \subset H$ with $e_H - 1$ edges satisfying*

$$m_2(H_-) \leq \overline{m}^r_{2b}(H) \ , \tag{2}$$

where

$$\overline{m}^r_{2b}(H) := \max_{H' \subseteq H} \frac{r(e_{H'}-1)+1}{r(v_{H'}-2)+2} \ .$$

Then the threshold for the balanced online H-avoidance game with r colors is

$$N_0(H,r,n) = n^{2-1/\overline{m}^r_{2b}(H)} \ .$$

This result applies to cycles of arbitrary size and to arbitrary integers $r \geq 2$. For cliques K_ℓ however, the result applies only if r is large enough (roughly, if $r \geq \ell$). Again there are examples of graphs for which the threshold is strictly higher than $n^{2-1/\overline{m}^2_{2b}(H)}$.

In order to state our results, we need to introduce some notation. An r-matched graph $G^r = (V, \mathcal{K})$ consists of a finite set V of vertices and a family $\mathcal{K}$ of pairwise disjoint sets of edges of cardinality r each. We refer to these as r-sets. A *valid* coloring of some r-matched graph is an r-edge-coloring with the property that each of the r colors appears exactly once in every r-set. Note that, by symmetry, after N steps the board of the balanced online Ramsey game is distributed uniformly over all r-matched graphs on n vertices with rN edges.

By $G^r_{n,m}$ we denote an r-matched graph drawn uniformly at random from all r-matched graphs on n vertices with m edges (we assume that m is divisible by r). The offline problem corresponding to the balanced online Ramsey game is the following: Given a random r-matched graph $G^r_{n,m}$, is there a valid edge-coloring avoiding monochromatic copies of H?

Note that finding such a coloring is harder than just finding a coloring of $G^r_{n,m}$ in which each color is used equally often. (In fact, it is fairly straightforward to see that the 0-statement of Theorem 2 remains true if only the latter requirement is made: A standard first moment calculation shows that if $m \leq cn^{2-(v_H-2)/(e_H-1)}$ for a small enough constant $c > 0$, a.a.s. all but a tiny fraction of the edges are not contained in any copy of H and can be colored arbitrarily. Combined with the observation that in order to avoid monochromatic copies of H it suffices to avoid monochromatic copies of H' for a subgraph $H' \subseteq H$ attaining $m_2(H) = (e_{H'}-1)/(v_{H'}-2)$, the claim follows.)

For any r-matched graph G^r and any graph H, let $G^r \overset{\text{valid}}{\to} (H)^e_r$ denote the property that every valid r-coloring of the edges of G^r contains a monochromatic copy of H. We say that a graph H on at least three vertices is *2-balanced* if $m_2(H) = (e_H - 1)/(v_H - 2)$, and is *strictly 2-balanced* if in addition $(e_{H'} - 1)/(v_{H'} - 2) < (e_H - 1)/(v_H - 2)$ for all proper subgraphs $H' \subseteq H$ on at least three vertices. Note that every graph H on at least three vertices has a strictly 2-balanced subgraph $H' \subseteq H$ such that $m_2(H) = (e_{H'} - 1)/(v_{H'} - 2)$.

Theorem 4 (Main Result 1). *Let $r \geq 2$, and let H be a non-forest that has a strictly 2-balanced subgraph $H' \neq K_3$ such that $m_2(H) = (e_{H'} - 1)/(v_{H'} - 2)$. Then there exist positive constants $c = c(H, r)$ and $C = C(H, r)$ such that*

$$\lim_{n\to\infty} \Pr[G^r_{n,m} \overset{\text{valid}}{\to} (H)^e_r] = \begin{cases} 0, & m \leq cn^{2-1/m_2(H)} \\ 1, & m \geq Cn^{2-1/m_2(H)} \end{cases} . \tag{3}$$

Note that the 1-statement is an immediate consequence of Theorem 2. Our proof of the 0-statement is based on the following observation. Assume that H is strictly 2-balanced, and consider the hypergraph $\mathcal{H}'$ which has the vertex set $\mathcal{K} = \mathcal{K}(G^r_{n,m})$, i.e., the family of all r-sets of the random r-matched graph, and as its edge set the copies of H such that a hyperedge is incident to a vertex $K \in \mathcal{K}$ if and only if the corresponding copy of H intersects K. This hypergraph has similar properties as the one considered in [13] for the classical case, which is defined analogously on the vertex set $E(G_{n,m})$ instead of $\mathcal{K}(G^r_{n,m})$. Specifically, it turns out that – analogously to [13] – a.a.s. all components of $\mathcal{H}'$ are unicylic and have at most logarithmic size as long as m is below the threshold. This insight allows us to extend the classical proof to our scenario.

The case where H' is a triangle is excluded in our result. We believe that Theorem 4 also holds in this case, but it seems a proof would have to proceed by somewhat different methods. The difficulties involved are essentially inherited from the classical case – for triangles the 0-statement of Theorem 2 was also proved separately in [6].

1.3. **Achlioptas game.** This last game is very similar to the balanced Ramsey game. The difference is that in each step, instead of coloring all r edges that are presented, the player simply has to select one edge and is allowed to discard the remaining $r - 1$ edges (each edge appears only once, so these edges will not be presented again later on). We will refer to this in the following as the *Achlioptas game*. Note that this can be viewed as a balanced Ramsey game in which the player needs to worry only about, say, red copies of H. Obviously this makes the player's task easier than in the balanced Ramsey game. Thus the thresholds for the balanced Ramsey game are lower bounds for the thresholds of the Achlioptas game, both in the online and in the offline version.

Our definition of the Achlioptas process differs slightly from other definitions found in the literature (e.g. in [5] or [15]), where in each step edges are sampled from all edges that have not been *selected* before (or even from all $\binom{n}{2}$ edges, allowing for multigraphs). In our setup, the requirement that discarded edges are completely removed from the game instead of placed back in the pool of available edges ensures that the edge sampling in a given step is not influenced by the player's earlier choices. This is needed in order to obtain an offline version of the problem with a well-defined distribution ($G^r_{n,m}$, in our case). As argued in [11], the threshold of the online game does not depend on the precise definition used.

The Achlioptas game was first investigated in [5], and solved completely in [11], where a general threshold formula $N_0(H, r, n)$ valid for all graphs H and every integer $r \geq 2$ was determined. This formula is rather complicated and contains both a minimization over all possible *orders* in which the edges of H may appear, and a maximization over entire *sequences* of appropriate subgraphs of H. As can be seen from that formula, in all cases where Theorem 3 applies, the Achlioptas threshold coincides with the one for the balanced Ramsey game. It is open whether in fact the thresholds of the two games are equal for all non-forests H (it is not hard to see that they differ if H is e.g. a star).

To state our results for the corresponding offline problem, we introduce some more notation. We say that A is an Achlioptas subgraph of some r-matched graph G^r (denoted $A \sqsubset G^r$) if A contains exactly one edge from every r-set of G^r. For any r-matched graph G^r and any graph H, let $G^r \overset{\text{Achlioptas}}{\to} (H)^e_r$ denote the property that every Achlioptas subgraph of $A \sqsubset G^r$ contains a copy of H.

Theorem 5 (Main Result 2)**.** *Let $r \geq 2$, and let H be a non-forest that has a strictly 2-balanced subgraph $H' \neq K_3$ such that $m_2(H) = (e_{H'} - 1)/(v_{H'} - 2)$. Then there exist positive constants $c = c(H, r)$ and $C = C(H, r)$ such that*

$$\lim_{n\to\infty} \Pr[G^r_{n,m} \overset{\text{Achlioptas}}{\to} (H)^e_r] = \begin{cases} 0, & m \leq cn^{2-1/m_2(H)} \\ 1, & m \geq Cn^{2-1/m_2(H)} \end{cases} . \tag{4}$$

Note that due to the assumptions on H, the 0-statement follows immediately from the 0-statement of Theorem 4. We will prove the 1-statement for any graph H with at least one edge. Our proof is inspired by [14] and proceeds by induction on $e(H)$. To make the inductive approach work, we prove the following strengthening of the desired result: For m as in the theorem, a.a.s. there is not only one, but $\Theta(n^{v_H}(m/n^2)^{e_H})$ many copies of H in every Achlioptas subgraph $A \sqsubset G^r_{n,m}$.

1.4. **Organization of this paper.** In Section 2 we prove the 0-statement of Theorem 4 (which, as discussed, also implies the 0-statement of Theorem 5). In Section 3 we prove the 1-statement of Theorem 5.

2. Lower bound for the balanced Ramsey problem

2.1. **Preliminaries.** Recall that an r-matched graph $G^r = (V, \mathcal{K})$ consists of a finite set V of vertices and a family $\mathcal{K}$ of pairwise disjoint sets of edges of cardinality r each, referred to as the r-sets of G^r. We use the notations $V(G^r)$ and $\mathcal{K}(G^r)$, and write $E(G^r) = \bigcup_{K\in\mathcal{K}(G^r)} K$ to refer to the edge set of the underlying unmatched graph. For each edge $e \in E(G^r)$, we let $K(e) \in \mathcal{K}$ denote the unique r-set containing e. For a subset $E' \subseteq E(G^r)$, we let

$$\mathcal{K}(E') := \bigcup_{e\in E'} \{K(e)\} \;\subseteq \mathcal{K} \qquad \text{and} \qquad V(E') := \bigcup_{e\in E'} e \;\subseteq V \;.$$

For any $\mathcal{K}' \subseteq \mathcal{K}$, we let

$$E(\mathcal{K}') := \bigcup_{K\in\mathcal{K}'} K \;\subseteq E(G^r) \qquad \text{and} \qquad V(\mathcal{K}') := V(E(\mathcal{K}')) = \bigcup_{K\in\mathcal{K}'} \bigcup_{e\in K} e \;\subseteq V \;.$$

Recall that by $G^r_{n,m}$ we denote an r-matched graph drawn uniformly at random from all r-matched graphs on n vertices with m edges (we assume that m is divisible by r). Note that, by symmetry, such a graph can be obtained by first generating a normal random graph $G_{n,m}$ and then choosing a random partition of its edge set into sets of size r uniformly at random. We will use the following elementary lemma.

Lemma 6. *For every integer $r \geq 1$, there exists a constant $C_r > 0$ such that, for n large enough, the expected number of copies of any r-matched graph $F^r = (V, \mathcal{K})$ with $|\mathcal{K}|r \leq 0.99\binom{n}{2}$ edges in $G^r_{n,m}$ is at most*

$$n^{|V|-|\mathcal{K}|\cdot 2r}(C_r m)^{|\mathcal{K}|} \;.$$

Proof. There are at most $n^{|V|}$ copies of F^r in K_n. For each of these, the probability that it is present in $G^r_{n,m}$ is the probability that all $|\mathcal{K}|r$ edges of $E(F^r)$ are present, multiplied with the probability that these edges are matched up correctly. It follows that the probability that a fixed copy of F^r in K_n is present in $G^r_{n,m}$ is exactly

$$\begin{aligned}
&\frac{\binom{\binom{n}{2}-|\mathcal{K}|r}{m-|\mathcal{K}|r}}{\binom{\binom{n}{2}}{m}} \cdot \frac{1}{\binom{m-1}{r-1}} \cdot \frac{1}{\binom{m-r-1}{r-1}} \cdot \ldots \cdot \frac{1}{\binom{m-(|\mathcal{K}|-1)r-1}{r-1}} \\
=&\frac{(\binom{n}{2}-|\mathcal{K}|r)! \cdot (r-1)!^{|\mathcal{K}|} \cdot m(m-r)\ldots(m-(|\mathcal{K}|-1)r)}{\binom{n}{2}!} \\
\leq&\left(\binom{n}{2}-|\mathcal{K}|r\right)^{-|\mathcal{K}|r} \cdot (r-1)!^{|\mathcal{K}|} \cdot m^{|\mathcal{K}|} \\
\leq&(n^{-2r} \cdot 201^r(r-1)! \cdot m)^{|\mathcal{K}|}
\end{aligned}$$

if n is large enough. Thus the statement of Lemma 6 follows for $C_r := 201^r(r-1)!$. ☐

2.2. **Overview of proof.** We present our proof algorithmically. We first give a deterministic algorithm that for any r-matched graph G^r either finds a valid coloring avoiding monochromatic copies of H or terminates with an error. For this part of the argument, H can be any graph with at least two edges.

In order to prove the 0-statement of Theorem 4, we then apply this algorithm for a strictly 2-balanced graph $H' \subseteq H$ with $m_2(H) = (e_{H'}-1)/(v_{H'}-2)$ and prove that a.a.s. it finds a valid coloring of $G^r_{n,m}$ avoiding monochromatic copies of H' (and thus also avoiding monochromatic copies of H) if $m \leq cn^{2-1/m_2(H)}$ for a suitable constant $c = c(H', r)$.

To be precise, we will prove this statement only for the case where H' is different from K_4 and C_4 at first (recall that the case $H' = K_3$ was excluded in the statement of the theorem). In order to deal with K_4 and C_4, the base algorithm needs an extra twist; we will sketch the necessary modifications at the end of this section.

2.3. **The algorithm.** In this section, we let H denote an arbitrary fixed graph with at least two edges. For any r-matched graph $G^r = (V, \mathcal{K})$, we define a family of copies of H in G^r as follows:

$$\mathcal{L}_{G^r} := \{L \subseteq E(G^r) : L \cong E(H) \;\wedge\; |\mathcal{K}(L)| = e_H\} \tag{5}$$

To simplify notation, these copies of H are viewed as edge-sets throughout. Note that the second condition excludes copies of H that contain several edges from the same r-set. Clearly, such copies will not be monochromatic in any valid coloring of G^r. Furthermore, for any subfamily $\mathcal{L} \subseteq \mathcal{L}_{G^r}$ we define a subsubfamily $\mathcal{L}^*_{G^r}(\mathcal{L}) \subseteq \mathcal{L}$ of '$\mathcal{L}$-critical' copies of H as follows:

$$\begin{aligned}
\mathcal{L}^*_{G^r}(\mathcal{L}) := \Big\{ L \in \mathcal{L} \text{ s.t. } \forall e \in L:\; &\big(\exists L' \in \mathcal{L} \text{ s.t. } L \cap L' = \{e\}\big) \\
&\vee\; \big(\forall f \in K(e) \setminus \{e\}\; \exists L_f \in \mathcal{L} \text{ s.t. } f \in L_f\big) \Big\} \;\subseteq\; \mathcal{L}\;. \qquad (6)
\end{aligned}$$

In words, a copy L is $\mathcal{L}$-critical if it is in $\mathcal{L}$ and if each of its edges (a) is the intersection of L with another copy from $\mathcal{L}$, or (b) satisfies that all $r-1$ other edges in its r-set are also contained in copies from $\mathcal{L}$.

Consider now the algorithm Bal-Edge-Col given in Figure 1. The algorithm is started with $G' = (V, \mathcal{K}')$ being a copy of G^r, and proceeds by removing and inserting r-sets into G'.

In the first while-loop the algorithm tries to successively remove r-sets from G' in such a way that when they are reinserted in the reverse order during the second loop, a valid coloring of G^r avoiding

1: **procedure** BAL-EDGE-COL($G^r = (V, \mathcal{K})$)
2: $\quad s \leftarrow$ EMPTY-STACK()
3: $\quad \mathcal{K}' \leftarrow \mathcal{K}$
4: $\quad \mathcal{L} \leftarrow \mathcal{L}_{G^r}$
5: $\quad$ **while** $G' = (V, \mathcal{K}')$ is not empty **do**
6: $\quad\quad$ **if** $\exists\ K \in \mathcal{K}'$ s.t. $\forall L \in \mathcal{L} : K \cap L = \emptyset$ **then**
7: $\quad\quad\quad s.$PUSH(K)
8: $\quad\quad\quad \mathcal{K}'.$REMOVE($K$)
9: $\quad\quad$ **else**
10: $\quad\quad\quad$ **if** $\exists L \in \mathcal{L} \setminus \mathcal{L}^*_{G'}(\mathcal{L})$ **then**
11: $\quad\quad\quad\quad s.$PUSH(L)
12: $\quad\quad\quad\quad \mathcal{L}.$REMOVE($L$)
13: $\quad\quad\quad$ **else**
14: $\quad\quad\quad\quad$ **error** "stuck"
15: $\quad\quad\quad$ **end if**
16: $\quad\quad$ **end if**
17: $\quad$ **end while**
18: $\quad$ **while** $s \neq \emptyset$ **do**
19: $\quad\quad$ **if** $s.$TOP() is an r-set **then**
20: $\quad\quad\quad K \leftarrow s.$POP()
21: $\quad\quad\quad \mathcal{K}'.$ADD($K$)
22: $\quad\quad\quad$ COLOR-ARBITRARILY(K)
23: $\quad\quad$ **else**
24: $\quad\quad\quad L \leftarrow s.$POP()
25: $\quad\quad\quad \mathcal{L}.$ADD($L$)
26: $\quad\quad\quad$ **if** L is monochromatic **then**
27: $\quad\quad\quad\quad e' \leftarrow$ any $e \in L$ s.t. $\big(\nexists L' \in \mathcal{L} : L \cap L' = \{e\}\big)$
$\quad\quad\quad\quad\quad$ and $\big(\exists f \in K(e) \setminus \{e\}$ s.t. $\nexists L_f \in \mathcal{L} : f \in L_f\big)$
28: $\quad\quad\quad\quad f' \leftarrow$ any $f \in K(e') \setminus \{e'\}$ s.t. $\nexists L_f \in \mathcal{L} : f \in L_f$
29: $\quad\quad\quad\quad$ SWAP-COLORS(e', f')
30: $\quad\quad\quad$ **end if**
31: $\quad\quad$ **end if**
32: $\quad$ **end while**
33: **end procedure**

FIGURE 1. The implementation of algorithm BAL-EDGE-COL.

monochromatic copies of H can be obtained by a simple local recoloring strategy. These r-sets are stored in the stack s, cf. lines 7–8 and 20–21.

The algorithm uses a local variable $\mathcal{L}$ to keep track of the copies of H in G' it 'still needs to worry about' (cf. the next paragraph). Throughout we have $\mathcal{L} \subseteq \mathcal{L}_{G'}$, and at the beginning we have $\mathcal{L} = \mathcal{L}_{G'} = \mathcal{L}_{G^r}$. These copies are also handled by the stack s, cf. lines 11–12 and 24–25. (Thus s contains both r-sets and copies of H.)

In lines 6–8, the r-sets that do not intersect any of the copies of H in $\mathcal{L}$ are removed from the graph G'. If no such r-sets exist, the algorithms checks in lines 10–12 whether there are non-$\mathcal{L}$-critical copies of H in $\mathcal{L}$. Such copies of H are deleted from $\mathcal{L}$ (and therefore ignored in all subsequent checks of the condition in line 6). This may allow more r-sets to be removed from G', and may also cause further copies of H in $\mathcal{L}$ to become non-$\mathcal{L}$-critical.

As we shall show next, if all r-sets can be removed from G^r during the first while-loop, the second while-loop creates a valid coloring of G^r avoiding monochromatic copies of H when reinserting the r-sets in the reverse order. In order to do so, it uses the two procedures Color-Arbitrarily and Swap-Colors. The procedure Color-Arbitrarily(K) assigns each of the r available colors to exactly one edge of the r-set K, and Swap-Colors(e, f) simply interchanges the colors of two edges e and f.

Lemma 7. *Let H be any fixed graph with at least two edges. On any r-matched graph G^r, algorithm* Bal-Edge-Col *either terminates with an error in line* 14 *or finds a valid edge-coloring of G^r avoiding monochromatic copies of H.*

Proof. Assume that the first while-loop of Bal-Edge-Col terminated without error. We show that the second loop creates a valid edge-coloring of G^r avoiding monochromatic copies of H.

First we show that the edges e' and f' in lines 27 and 28 always exist. Since the second loop inserts r-sets into G' in the reverse order in which they were deleted during the first loop, when we select e' and f' in lines 27 and 28, G' and $\mathcal{L}$ are exactly as at the time when L was pushed on the stack in line 11. Thus L satisfies the condition of the if-clause in line 10, i.e., L is not $\mathcal{L}$-critical. Hence there exist edges e' and f' as specified, cf. (6).

Moreover, note that since the algorithm only calls Swap-Colors for edges that are in the same r-set, a valid coloring of G' is maintained throughout.

We conclude the proof by showing that the algorithm maintains the following invariant during the second loop: after each execution of line 29, no copy $L \in \mathcal{L}$ is monochromatic. Since at the very end we have $\mathcal{L} = \mathcal{L}_{G^r}$, the lemma then follows. (Recall that copies of H that are not in $\mathcal{L}_{G^r}$ will not be monochromatic in any valid coloring of G^r.)

We prove this by induction. Clearly, the statement is true at the beginning of the second loop since then $\mathcal{L}$ is empty. Consider thus the situation immediately after some copy L is popped from the stack and inserted back into $\mathcal{L} \subseteq \mathcal{L}_{G'}$ in lines 24 and 25. If L is monochromatic at this moment, exactly one of its edges is recolored in line 29, so clearly L is not monochromatic after that. It remains to prove that the execution of line 29 does not cause other copies of H in $\mathcal{L}$ to become monochromatic. This is an easy consequence of the following two observations: By our choice of f', there are no copies of H in $\mathcal{L}$ that contain f'. By our choice of e', copies of H in $\mathcal{L}$ that contain e' overlap in more than one edge with L and will therefore not become monochromatic when e' is recolored. This proves that no monochromatic copy in $\mathcal{L}$ is created, concluding the proof of Lemma 7. □

The algorithm Bal-Edge-Col gets stuck if and only if the conditions in lines 6 and 10 both fail. Then all copies in $\mathcal{L}$ are $\mathcal{L}$-critical, and each of the remaining r-sets $K \in \mathcal{K}'$ intersects with such a critical copy. In the following we denote by G' the r-matched graph $(V, \mathcal{K}')$ at the moment Bal-Edge-Col gets stuck, and – with slight abuse of notation – by $\mathcal{L}^*_{G'}$ the family $\mathcal{L} = \mathcal{L}^*_{G'}(\mathcal{L}) \subseteq \mathcal{L}_{G'}$ at the moment the algorithm gets stuck.

2.4. **Analysis.** In order to prove the 0-statement of Theorem 4, we consider a subgraph $H' \subseteq H$ that is strictly 2-balanced and satisfies $m_2(H') = m_2(H)$. If we show that Bal-Edge-Col a.a.s. finds a valid coloring of $G^r_{n,m}$ with m as claimed avoiding monochromatic copies of H', the claim immediately follows. As already mentioned, we will deal with the special cases $H' = K_4$ and $H' = C_4$ later.

Lemma 8. *Let H be a strictly 2-balanced non-forest different from K_3, K_4, C_4. Then there exists a constant $c = c(H, r) > 0$ such that for $m \leq cn^{2-1/m_2(H)}$, a.a.s. algorithm* Bal-Edge-Col *terminates on $G^r_{n,m}$ without error.*

For the rest of this section, we consider H fixed as in the lemma. Our approach for the proof of Lemma 8 is as follows. We describe a procedure Grow that takes as input the r-matched graph G' and the family $\mathcal{L}^*_{G'} \subseteq \mathcal{L}_{G'}$ left when Bal-Edge-Col gets stuck, and constructs as output an r-matched subgraph $F^r \subseteq G'$ that is either too dense or too large to appear in $G^r_{n,m}$ with m as claimed. Together with a bound on the number of non-isomorphic graphs Grow can output, this will imply that Bal-Edge-Col succeeds a.a.s. In the next few pages we describe the procedure Grow and prove a series of preparatory claims. The proof of Lemma 8 is then carried out at the end of this section on page 15.

The procedure Grow starts with any $F_1 = L_1 \in \mathcal{L}^*_{G'}$ and in every step adds a new copy $L_i \in \mathcal{L}^*_{G'}$ of H to the graph F_{i-1} already found. Throughout, $F_i = \bigcup_{j=1}^{i} L_j$ is viewed simply as a subset of $E(G')$. By $F_i^r := (V(\mathcal{K}(F_i)), \mathcal{K}(F_i))$ we denote the r-matched graph spanned by the complete r-sets that intersect F_i.

In every step, Grow determines an edge $f_i \in E(\mathcal{K}(L_{i-1}))$ that satisfies a set of properties we will specify later. The main purpose of these properties is to ensure the existence of a copy $L_i \in \mathcal{L}^*_{G'}$ that contains f_i and is not contained completely in $E(F^r_{i-1})$. Such an L_i is then added to F_{i-1}. As a consequence, L_{i-1} and L_i either intersect directly ($f_i \in L_{i-1}$, option A below) or are 'linked' via a common r-set ($f_i \in E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}$, option B below).

If L_i and $E(F^r_{i-1})$ intersect only in f_i, then F^r_{i-1} is extended by exactly $e_H - 1$ many r-sets since by definition of $\mathcal{L}_{G'}$ each 'new' edge in $L_i \setminus \{f_i\}$ is in a different r-set. In general however, L_i and $E(F^r_{i-1})$ might intersect in more edges, and thus F^r_{i-1} might be extended by fewer r-sets. For $i \geq 2$ we denote by x_i the amount by which the actual number of new r-sets differs from the upper bound $e_H - 1$, i.e., we set

$$\begin{aligned} x_i :&= e_H - 1 - \big(|\mathcal{K}(F_i)| - |\mathcal{K}(F_{i-1})|\big) \\ &= e_H - 1 - |L_i \setminus E(F^r_{i-1})| \\ &= |L_i \cap E(F^r_{i-1})| - 1 \ \geq 0 \ . \end{aligned} \tag{7}$$

Similarly, F^r_{i-1} is extended by at most $v_H - 2$ 'inner' vertices from $V(L_i)$, and in addition to that, by at most $|L_i \setminus E(F^r_{i-1})| \cdot 2(r-1)$ 'outer' vertices incident to edges from $E(\mathcal{K}(L_i)) \setminus L_i$. Note that the latter bound depends on $|L_i \setminus E(F^r_{i-1})|$ and thus on x_i. For $i \geq 2$ we denote by y_i the amount by which the actual number of new vertices differs from this upper bound of inner and outer vertices, i.e., we set

$$\begin{aligned} y_i :&= v_H - 2 + |L_i \setminus E(F^r_{i-1})| \cdot 2(r-1) - \big(|V(F^r_i)| - |V(F^r_{i-1})|\big) \\ &= v_H - 2 - |V(L_i) \setminus V(F^r_{i-1})| \\ &\quad + |L_i \setminus E(F^r_{i-1})| \cdot 2(r-1) - |V(F^r_i) \setminus (V(F^r_{i-1}) \cup V(L_i))| \\ &\overset{(7)}{=} |V(L_i) \cap V(F^r_{i-1})| - 2 \\ &\quad + (e_H - 1 - x_i) \cdot 2(r-1) - |V(E(\mathcal{K}(L_i)) \setminus L_i) \setminus V(E(F^r_{i-1}) \cup L_i)| \ \geq 0 \ . \end{aligned} \tag{8}$$

In the last expression, the term $|V(L_i) \cap V(F^r_{i-1})| - 2$ accounts for inner vertices analogously to (7), and the remaining terms account for outer vertices that are lost due to outer edges $E(\mathcal{K}(L_i)) \setminus L_i$ intersecting among themselves or with $E(F^r_{i-1}) \cup L_i$. For $i = 1$ we define

$$x_1 := 0 \qquad \text{and} \qquad y_1 := v_H + e_H \cdot 2(r-1) - |V(F^r_1)| \ \geq 0 \ . \tag{9}$$

We will relate y_i to x_i by considering the graph

$$J_i := (V(L_i) \cap V(F^r_{i-1}), L_i \cap E(F^r_{i-1})) \ . \tag{10}$$

We have

$$x_i \overset{(7)}{=} e(J_i) - 1 \tag{11}$$

and

$$y_i \overset{(8)}{\geq} v(J_i) - 2 \ , \tag{12}$$

which implies in particular that x_i can only be positive if y_i is positive. In order to make use of the assumption that H is strictly 2-balanced, we will show that J_i is a *proper* subgraph of $(V(L_i), L_i) \cong H$. (Note that this is a reformulation of the requirement mentioned above that L_i is not contained completely in $E(F_{i-1}^r)$.)

Procedure GROW stops and returns the current r-matched graph F_i^r as soon as $\sum_{1 \leq j \leq i} y_j \geq 3$ or $i \geq \log n$. We will show that if the procedure stops because $\sum_{1 \leq j \leq i} y_j \geq 3$, the r-matched graph F_i^r is so dense that a.a.s. it will not appear in $G_{n,m}^r$. If on the other hand GROW stops because i reaches $\log n$, then at this point F_i^r is so large that a.a.s. it will not appear in $G_{n,m}^r$.

To complete the description of procedure GROW, we need to specify the precise properties we require f_i and L_i to satisfy. For f_i, we require that either

A. $f_i \in L_{i-1}$, and
- (i) there is a copy $L \in \mathcal{L}_{G'}^*$ with $L \cap L_{i-1} = \{f_i\}$, and
- (ii) at most one vertex of f_i is incident to an edge from $E(F_{i-1}^r) \setminus L_{i-1}$, and
- (iii) if $y_{i-1} = 0$, no vertex of f_i is incident to an edge from $E(F_{i-1}^r) \setminus L_{i-1}$,

or

B. $f_i \in E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}$, and
- (i) there is a copy $L \in \mathcal{L}_{G'}^*$ containing f_i, and
- (ii) at most one vertex of f_i is incident to an edge from $E(F_{i-1}^r) \setminus \{f_i\}$, and
- (iii) if $y_{i-1} = 0$, no vertex of f_i is incident to an edge from $E(F_{i-1}^r) \setminus \{f_i\}$.

Once such an f_i is picked, any L as in property $A.(i)$, respectively $B.(i)$, is a valid choice for L_i.

We now show that it is always possible to find an edge f_i as specified, and that this in turn ensures that F_i^r gets strictly larger in every step. Observe that whenever GROW needs to find an edge f_i, we have $y_{i-1} \leq 2$ since otherwise GROW terminates after step $i-1$.

Claim 9. *Let G' and $\mathcal{L}_{G'}^*$ be the r-matched graph and the family $\mathcal{L}$ at the moment* BAL-EDGE-COL *got stuck, and assume that* GROW *has successfully constructed an r-matched graph $F_{i-1}^r \subseteq G'$ in the first $i-1$ steps.*

If $y_{i-1} \leq 2$, then there exists an edge f_i satisfying properties $A.(i), (ii), (iii)$ or $B.(i), (ii), (iii)$ in step i of GROW*. As a consequence, the copy L_i added to F_{i-1} is not contained in $E(F_{i-1}^r)$ completely (i.e., J_i as defined in* (10) *is isomorphic to a* proper *subgraph of H).*

Proof. We will use that H as in Lemma 8 has at least five vertices, has minimum degree 2, and is 'spacious', i.e., for every edge e of H there exists an edge of H that is disjoint from e.

By definition of GROW, the copy L_{i-1} that was picked in step $i-1$ is in $\mathcal{L}_{G'}^*$. This implies that for every edge $e \in L_{i-1}$ there is an edge $f_e \in K(e)$ for which there exists $L \in \mathcal{L}_{G'}^*$ as specified in $A.(i)$ (then $f_e = e$) or $B.(i)$ (then $f_e \in K(e) \setminus \{e\}$), cf. (6) and the last paragraph of Section 2.3. Thus we have a set of 'candidate edges' $C_{i-1} := \{f_e \mid e \in L_{i-1}\}$ for which a copy $L \in \mathcal{L}_{G'}^*$ as required exists. Note that the two cases A and B give rise to a natural distinction between 'inner' candidate edges in $C_{i-1} \cap L_{i-1}$ and 'outer' candidate edges in $C_{i-1} \setminus L_{i-1}$.

We first deal with the easier case $y_{i-1} = 0$. Then we also have $x_{i-1} = 0$ by (11) and (12). Thus $\mathcal{K}(L_{i-1})$ and F^r_{i-2} intersect only in the two vertices of f_{i-1}, and none of the candidate edges is involved in any other undesired intersections. Since H is spacious, there exists an edge $e' \in L_{i-1}$ that is disjoint from f_{i-1}, and the corresponding candidate edge $f_{e'} \in C_{i-1}$ is an edge satisfying $A.(iii)$ (if $f_{e'} = e'$ is an inner candidate edge) or $B.(iii)$ (if $f_{e'} \in K(e') \setminus \{e'\}$ is an outer candidate edge). In either case, $f_{e'}$ can be chosen as f_i.

In the remaining case $1 \leq y_{i-1} \leq 2$, condition $A.(ii)$ or $B.(ii)$ might not hold for some of the edges in C_{i-1} due to undesired intersections. Our argument will be that there are not enough intersections to rule out all candidate edges. In order to show this, we decompose y_{i-1} as defined in (8) into smaller terms corresponding to different types of intersections.

We distinguish three types of contributions to y_{i-1}. The first one corresponds to intersections of L_{i-1} with $E(F^r_{i-2}) \setminus \{f_{i-1}\}$. These intersections contribute an amount of $|V(L_{i-1}) \cap V(F^r_{i-2})| - 2 = v(J_{i-1}) - 2$ to y_{i-1} (cf. the remarks after (8) and the definition of J_i in (10)). We let

$$V_1 := V(L_{i-1}) \cap V(F^r_{i-2}) = V(J_{i-1}) \ .$$

The second contribution comes from (outer) edges in $E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}$ that are not vertex-disjoint from each other. Their contribution to y_{i-1} is at least the cardinality of

$$V_2 := \{v : \exists e_1 \neq e_2 \in E(\mathcal{K}(L_{i-1})) \setminus L_{i-1} : v \in e_1 \cap e_2\}$$

(this is a lower bound because we ignore multiplicities). The last contribution to y_{i-1} comes from outer edges that are not vertex-disjoint from $E(F^r_{i-2}) \cup L_{i-1}$. Their contribution to y_{i-1} is the cardinality of

$$V_3 := V(E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}) \cap V(E(F^r_{i-2}) \cup L_{i-1}) \ .$$

Note that the sets V_1, V_2, V_3 are not necessarily disjoint from each other, as several intersections might happen at the same vertex. We have $y_{i-1} \geq (|V_1| - 2) + |V_2| + |V_3|$ and in particular

$$|V_1 \cup V_3| \leq y_{i-1} + 2 \tag{13}$$

and

$$|V_2 \cup V_3| \leq y_{i-1} \ . \tag{14}$$

A moment's thought reveals that $e \in L_{i-1}$ satisfies property $A.(iii)$ if it is disjoint from $V_1 \cup V_3$, and that it satisfies $A.(ii)$ if it shares at most one vertex with $V_1 \cup V_3$. Similarly, an edge $e \in E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}$ satisfies property $B.(ii)$ or $B.(iii)$ if it shares at most one, resp. no vertex with $V_2 \cup V_3$.

It follows with $y_{i-1} \leq 2$ from (14) that there is at most one outer candidate edge $e \in C_{i-1} \setminus L_{i-1}$ that cannot be picked as f_i, namely the edge $V_2 \cup V_3$ (if it is indeed a candidate edge). Thus if $|C_{i-1} \setminus L_{i-1}| \geq 2$, there always is a valid choice for f_i. On the other hand, if $|C_{i-1} \setminus L_{i-1}| \leq 1$, the assumption that H has at least 5 vertices implies with (13) that there is a vertex $v \in V(L_{i-1}) \setminus (V_1 \cup V_3)$. Since H has minimum degree at least 2, there are at least two edges from L_{i-1} incident to v. Since $|C_{i-1} \setminus L_{i-1}| \leq 1$, at least one of them is an (inner) candidate edge and can be chosen as f_i.

This proves the first statement of Claim 9. It remains to show that L_i is not contained completely in $E(F^r_{i-1})$. We distinguish two cases that correspond to options A and B.

If $f_i \in L_{i-1}$, by property $A.(ii)$ there is a vertex $v \in f_i$ that is not incident to $E(F^r_{i-1}) \setminus L_{i-1}$. As H has minimum degree at least 2, there is at least one other edge $f' \in L_i$ incident to v besides f_i. Moreover, since $L_i \cap L_{i-1} = \{f_i\}$ due to property $A.(i)$, we have $f' \notin L_{i-1}$. It follows that f' is not in $E(F^r_{i-1})$, as the opposite would contradict our choice of v.

Similarly, if $f_i \in E(\mathcal{K}(L_{i-1})) \setminus L_{i-1}$, by property $B.(ii)$ there is a vertex $v \in f_i$ that is not incident to $E(F^r_{i-1}) \setminus \{f_i\}$. As H has minimum degree at least 2, there is at least one other edge $f' \in L_i$ incident to v besides f_i. Again f' is not in $E(F^r_{i-1})$ by our choice of v.

Thus in both cases L_i is not contained in $E(F^r_{i-1})$ completely. This concludes the proof of Claim 9. □

Our next goal is to prove that the output of Grow is either too large or too dense to appear in $G^r_{n,m}$ with m as claimed. Eventually, we shall apply Lemma 6 to the r-matched graph $F^r = (V, \mathcal{K})$ returned by Grow and use that the expected number of copies of F^r in $G^r_{n,m}$ with $m = cn^{2-1/m_2(H)}$ is bounded by

$$n^{|V|-|\mathcal{K}|\cdot 2r}(C_r \cdot cn^{2-1/m_2(H)})^{|\mathcal{K}|} = (c \cdot C_r)^{|\mathcal{K}|} n^{\lambda_H(F^r)} \ , \tag{15}$$

where λ_H is the function that assigns to every r-matched graph $F^r = (V, \mathcal{K})$ the value

$$\lambda_H(F^r) := |V| - |\mathcal{K}| \cdot \Big(2(r-1) + \frac{1}{m_2(H)}\Big) \ . \tag{16}$$

In order to show that (15) tends to zero fast enough, we take a closer look at the exponent $\lambda_H(F^r_i)$.

Claim 10. *After every step of* Grow *we have*

$$\lambda_H(F^r_i) = \frac{\sum_{1\le j\le i} x_j - 1}{m_2(H)} - \Big(\sum_{1\le j\le i} y_j - 2\Big) \ . \tag{17}$$

Proof. We have

$$\begin{aligned}
\lambda_H(F^r_i) \quad &= \quad \sum_{2\le j\le i}\big(|V(F^r_j)| - |V(F^r_{j-1})|\big) + |V(F^r_1)| \\
&\quad - \Big(\sum_{2\le j\le i}\big(|\mathcal{K}(F_j)| - |\mathcal{K}(F_{j-1})|\big) + |\mathcal{K}(F_1)|\Big)\cdot\Big(2(r-1) + \frac{1}{m_2(H)}\Big) \\
&\overset{(7),(8),(9)}{=} \sum_{2\le j\le i}\big(v_H - 2 + |L_j \setminus E(F^r_{j-1})|\cdot 2(r-1) - y_j\big) + \big(v_H + e_H\cdot 2(r-1) - y_1\big) \\
&\quad - \Big(\sum_{2\le j\le i}|L_j \setminus E(F^r_{j-1})| + e_H\Big)\cdot 2(r-1) \\
&\quad - \Big(\sum_{2\le j\le i}(e_H - 1 - x_j) + e_H - x_1\Big)\cdot\frac{1}{m_2(H)} \\
&= \quad i\cdot(v_H - 2) + 2 - \sum_{1\le j\le i} y_j - \frac{i\cdot(e_H-1) + 1 - \sum_{1\le j\le i} x_j}{m_2(H)} \\
&= \quad \frac{\sum_{1\le j\le i} x_j - 1}{m_2(H)} - \Big(\sum_{1\le j\le i} y_j - 2\Big) \ ,
\end{aligned}$$

where the terms containing i cancel out because H is (strictly) 2-balanced, i.e., $m_2(H) = (e_H - 1)/(v_H - 2)$. □

Using Claim 10 we now prove the desired upper bounds on the exponent $\lambda_H(F^r_i)$.

Claim 11. *There exist constants $\lambda_0 = \lambda_0(H) > 0$ and $\gamma = \gamma(H) > 0$ such that the following holds:*

- *The output of* Grow *satisfies $\lambda_H(F^r_i) \le \lambda_0$.*
- *If* Grow *terminates because $\sum_{1\le j\le i} y_j \ge 3$, its output satisfies $\lambda_H(F^r_i) \le -\gamma$.*

Proof. Let
$$\lambda_0 = \lambda_0(H) := 2 - 1/m_2(H) > 0$$
and
$$\gamma = \gamma(H) := \min_{\substack{H' \subseteq H \\ v(H') \geq 3}} (v(H') - 2) - \frac{e(H') - 1}{m_2(H)} > 0 \ , \tag{18}$$
where γ is positive due to our assumption that H is strictly 2-balanced. Note that considering H' with 3 vertices and one edge in (18) yields $\gamma \leq 1$.

We first prove that $\lambda_H(F_i^r)$ is non-increasing. Let
$$\Delta_i := \lambda_H(F_i^r) - \lambda_H(F_{i-1}^r) \stackrel{(17)}{=} \frac{x_i}{m_2(H)} - y_i$$
denote the change of λ_H in step i, and consider J_i as defined in (10). By (11) and (12) and using that J_i is isomorphic to a *proper* subgraph of H (cf. Claim 9), we have
$$\Delta_i \stackrel{(11),(12)}{\leq} \frac{e(J_i) - 1}{m_2(H)} - (v(J_i) - 2) \stackrel{(18)}{\leq} -\gamma \tag{19}$$
if $v(J_i) \geq 3$. Otherwise it follows from (11) and (12) that $x_i = e(J_i) - 1 = 0$ and $y_i \geq v(J_i) - 2 = 0$ (observe that J_i always contains the edge f_i), which implies $\Delta_i \leq 0$. Thus $\lambda(F_i^r)$ is non-increasing throughout, and the graph returned by Grow satisfies $\lambda(F_i^r) \leq \lambda(F_1^r) \leq \lambda_0$, where the last step follows from (17) using that $x_1 = 0$ and $y_1 \geq 0$. This proves the first part of Claim 11.

Consider now the graph J_i' obtained from J_i by removing the edge f_i and the vertices of f_i that have degree one in J_i. Clearly, we have
$$x_i \stackrel{(11)}{=} e(J_i) - 1 = e(J_i') \ , \tag{20}$$
and we now show that moreover we have
$$\sum_{1 \leq j \leq i} y_j \geq v(J_i') \ . \tag{21}$$
This is true because of the following: If $y_{i-1} = 0$, the edge f_i satisfies property $A.(iii)$ or $B.(iii)$, and all edges $f' \in L_i$ incident to f_i are not in $E(F_{i-1}^r)$ (cf. the arguments at the end of the proof of Claim 9). In other words, f_i is isolated in J_i, and we obtain from (12) that $v(J_i') = v(J_i) - 2 \leq y_i$. On the other hand, if $y_{i-1} \geq 1$ the edge f_i satisfies only property $A.(ii)$ or $B.(ii)$ and is not necessarily isolated in J_i. However, at most one of the two vertices of f_i is incident to an edge from $E(F_{i-1}^r)$ (cf. again the arguments at the end of the proof of Claim 9), and with (12) we obtain $v(J_i') \leq v(J_i) - 1 \leq 1 + y_i \leq y_{i-1} + y_i$, which proves (21).

Assume now that Grow terminates after step i because $\sum_{1 \leq j \leq i} y_j \geq 3$. Clearly, if $\sum_{1 \leq j \leq i} x_j = 0$, the output of Grow satisfies $\lambda_H(F_i^r) \leq 2 - \sum_{1 \leq j \leq i} y_j \leq -1 \leq -\gamma$. Otherwise, assume first that i is the only step with $x_i = e(J_i') \geq 1$. If $x_i \geq 2$, we have $v(J_i') \geq 3$, and thus
$$\lambda_H(F_i^r) \stackrel{(17)}{=} \frac{x_i - 1}{m_2(H)} - (\sum_{1 \leq j \leq i} y_j - 2) \stackrel{(20),(21)}{\leq} \frac{e(J_i') - 1}{m_2(H)} - (v(J_i') - 2) \stackrel{(18)}{\leq} -\gamma \ .$$
If $x_i = 1$, we obtain with (17) and $\sum_{1 \leq j \leq i} y_j \geq 3$ that $\lambda_H(F_i^r) \leq -1 \leq -\gamma$.

It remains to consider the case where $x_j \geq 1$ for some $j < i$. Since Grow did not terminate in step j, we have $\sum_{k=1}^{j} y_k \leq 2$, which implies with (20) and (21) that $x_j = 1$, $\sum_{k=1}^{j} y_k = 2$. Moreover, for all $j < k < i$ we have $x_k = y_k = 0$ since otherwise Grow would have terminated in step k. It follows that $\lambda_H(F_j^r) = 0$ and $\lambda_H(F_i^r) = \Delta_i$. Since, similarly to above, we have $\Delta_i \leq -\gamma$ by (19) if $v(J_i) \geq 3$ and $\Delta_i = -y_i \leq -1 \leq -\gamma$ if $v(J_i) \leq 2$ (since then $x_i = 0$), this concludes the proof of the second part of Claim 11. $\square$

In order to show that a.a.s. $G^r_{n,m}$ contains none of the r-matched graphs that can be generated by Grow, we prove an upper bound on the number of such graphs, making crucial use of the fact that only constantly many steps with $x_i > 0$ or $y_i > 0$ may occur before Grow terminates. For $i \geq 1$, let $\mathcal{F}^r(H, i)$ denote a family of representatives for the isomorphism classes of all r-matched graphs F^r_i that can be the output of Grow when it terminates after exactly i steps on some input G' and $\mathcal{L}^*_{G'}$ as in Claim 9. Note that, crucially, we do *not* consider a fixed input G' and $\mathcal{L}^*_{G'}$ here, but consider all possible outputs of Grow on all possible inputs G' and $\mathcal{L}^*_{G'}$. Moreover, let $\mathcal{F}^r(H, \leq i) := \bigcup_{j=1}^{i} \mathcal{F}^r(H, j)$.

Claim 12. *There exists a constant $C = C(H, r)$ such that for all $i \geq 1$ we have*

$$|\mathcal{F}^r(H, i)| \leq (i + 1)^C (re_H)^i \ . \tag{22}$$

Proof. In the following, we say that step i is *non-degenerate* if $y_i = 0$ (which implies in particular that $x_i = 0$, as argued above), and *degenerate* otherwise. For $0 \leq d \leq \min\{i, 3\}$, let $\mathcal{F}^r(H, i, d)$ denote a family of representatives for the isomorphism classes of all r-matched graphs F^r_i that Grow can generate in exactly i steps if it performs exactly d degenerate steps along the way (recall that Grow terminates after at most 3 degenerate steps).

In a non-degenerate step, the isomorphism class of the r-matched graph F^r_i is uniquely defined by the structure of F^r_{i-1} and the edge $f_i \in E(\mathcal{K}(L_{i-1}))$ chosen by Grow: since we have $x_i = y_i = 0$, F^r_i is obtained by attaching a new copy L_i to f_i such that $V(L_i) \cap V(F^r_{i-1}) = f_i$, and embedding each of the $e_H - 1$ edges in $L_i \setminus \{f_i\}$ into an r-set such that the $(e_H - 1)(r - 1)$ edges in $E(\mathcal{K}(L_i \setminus \{f_i\})) \setminus L_i$ are completely vertex-disjoint from each other and from $E(F^r_{i-1}) \cup L_i$. Thus regardless of the input instance G' and regardless of what happened in all previous steps, there are at most re_H ways to extend F^r_{i-1} to F^r_i in a non-degenerate step. This implies in particular that $|\mathcal{F}^r(H, i, 0)| \leq (re_H)^i$ for all i.

In order to analyze what happens in degenerate steps, we use a more generic argument. Observe that the newly added r-sets $\mathcal{K}(L_i) \setminus \mathcal{K}(F_{i-1})$ span an r-matched graph on at most

$$K := v_H + (e_H - 1) \cdot 2(r - 1)$$

vertices. In particular, at most K vertices are added in every step. Together with $|V(F^r_1)| \leq K + 2(r - 1)$, we obtain $|V(F^r_i)| \leq Ki + 2(r - 1)$.

In the following, $\mathcal{G}^r_K$ denotes the set of all r-matched graphs on at most K vertices. F^r_i is uniquely defined if one specifies the r-matched graph $G^r \in \mathcal{G}^r_K$ spanned by the new r-sets $\mathcal{K}(L_i) \setminus \mathcal{K}(F_{i-1})$, the number s of vertices in which G^r intersects F^r_{i-1}, and two ordered lists of vertices from G^r and F^r_{i-1} respectively of length s, which specify the mapping of the intersection vertices from G^r into F^r_{i-1}. Thus, the number of ways to extend F^r_{i-1} to F^r_i in a degenerate step is bounded from above by

$$\sum_{G^r \in \mathcal{G}^r_K} \sum_{s=2}^{|V(G^r)|} |V(G^r)|^s |V(F^r_i)|^s \leq |\mathcal{G}^r_K| \cdot K \cdot K^K (Ki + 2(r - 1))^K \leq (i + 1)^{C_0} \ ,$$

for a large constant C_0 depending only on H and r. It follows that for $0 \leq d \leq \min\{i, 3\}$ we have

$$|\mathcal{F}^r(H, i, d)| \leq \binom{i}{d} \left((i + 1)^{C_0}\right)^d (re_H)^{i-d} \leq (i + 1)^{d(C_0+1)} (re_H)^i \ .$$

Here the binomial coefficient corresponds to the choice of the d degenerate steps. We obtain

$$\begin{aligned} |\mathcal{F}^r(H,i)| &\leq \sum_{d=0}^{\min\{i,3\}} |\mathcal{F}^r(H,i,d)| \\ &\leq 4(i+1)^{3(C_0+1)}(re_H)^i \\ &\leq (i+1)^C(re_H)^i \end{aligned}$$

for an even larger constant C depending only on H and r. This concludes the proof of Claim 12. □

We now have all the ingredients to prove that a.a.s. $G^r_{n,m}$ does not contain one of the graphs that can be generated by Grow. In the informal language used at the beginning of this section, we will show that the graphs in $\mathcal{F}^r(H, \leq \lceil \log n\rceil - 1)$ are too dense, and the graphs in $\mathcal{F}^r(H, \lceil \log n\rceil)$ too large to appear in $G^r_{n,m}$ with m as claimed.

Claim 13. *There exists a constant $c = c(H,r) > 0$ such that for $m \leq cn^{2-1/m_2(H)}$, a.a.s. the random r-matched graph $G^r_{n,m}$ does not contain any r-matched graph from $\mathcal{F}^r(H, \leq \lceil \log n\rceil)$.*

Proof. In addition to the bounds on $\lambda(F^r_i)$ and $|\mathcal{F}^r(H,i)|$ proved in Claims 11 and 12, we use that due to Claim 9, every step i extends F^r_{i-1} by at least one r-set and thus all r-matched graphs $(V,\mathcal{K}) \in \mathcal{F}^r(H,i)$ satisfy $|\mathcal{K}| \geq i$. It follows from Lemma 6 that for

$$c = c(H,r) := \frac{\mathrm{e}^{-\lambda_0-\gamma}}{C_r \cdot re_H} \ , \tag{23}$$

the expected number of copies of r-matched graphs from $\mathcal{F}^r(H, \leq \lceil \log n\rceil)$ in $G^r_{n,m}$ with $m \leq cn^{2-1/m_2(H)}$ is bounded by

$$\begin{aligned} &\sum_{F^r=(V,\mathcal{K})\in\mathcal{F}^r(H,\leq\lceil\log n\rceil)} n^{|V|-2r\cdot|\mathcal{K}|}(C_r m)^{|\mathcal{K}|} \\ &\overset{(15)}{\leq} \sum_{i=1}^{\lceil \log n\rceil} \sum_{F^r=(V,\mathcal{K})\in\mathcal{F}^r(H,i)} (c\cdot C_r)^{|\mathcal{K}|} n^{\lambda_H(F^r)} \\ &\overset{(|\mathcal{K}|\geq i,\ c\cdot C_r \leq 1)}{\leq} \sum_{i=1}^{\lceil \log n\rceil} \sum_{F^r=(V,\mathcal{K})\in\mathcal{F}^r(H,i)} (c\cdot C_r)^{i} n^{\lambda_H(F^r)} \\ &\overset{\text{(Cl. 11)}}{\leq} \sum_{i=1}^{\lceil \log n\rceil -1} |\mathcal{F}^r(H,i)|\cdot(c\cdot C_r)^i n^{-\gamma} \ + \ |\mathcal{F}^r(H,\lceil \log n\rceil)|\cdot(c\cdot C_r)^{\lceil\log n\rceil} n^{\lambda_0} \\ &\overset{(22),\ (23)}{\leq} \sum_{i=1}^{\lceil \log n\rceil -1} (i+1)^C \mathrm{e}^{(-\lambda_0-\gamma)i} n^{-\gamma} \ + \ (\lceil \log n\rceil+1)^C \mathrm{e}^{(-\lambda_0-\gamma)\lceil\log n\rceil} n^{\lambda_0} \\ &\leq (\log n + 2)^{C+1}\cdot n^{-\gamma} = o(1) \ , \end{aligned}$$

which implies Claim 13 by Markov's inequality. □

Proof of Lemma 8. Suppose that the call to Bal-Edge-Col(G^r) gets stuck for some r-matched graph G^r, and consider $G' \subseteq G^r$ and $\mathcal{L}^*_{G'} \subseteq \mathcal{L}_{G'}$ at this moment. Applying Grow to G' and $\mathcal{L}^*_{G'}$ yields a copy of an r-matched graph $F^r \in \mathcal{F}^r(H, \leq \lceil \log n\rceil)$ that is contained in $G' \subseteq G^r$. However, by Claim 13, $G^r = G^r_{n,m}$ with m as claimed contains a.a.s. no such graph. Thus Bal-Edge-Col succeeds a.a.s. in finding a valid coloring of $G^r_{n,m}$ avoiding monochromatic copies of H. This proves

Lemma 8 (and thus the 0-statement of Theorem 4 if none of the special cases discussed below occurs). □

2.5. **Dealing with the special cases.** We first consider the case $H = K_4$. Let us point out where the proof of Lemma 8 goes wrong in this case. Recall that the purpose of Claim 9 was to show that we can always find an edge $f_i \in \mathcal{K}(L_{i-1})$ and a copy L_i containing f_i that is not contained completely in $E(F^r_{i-1})$. The argument presented there does not guarantee this for $H = K_4$. However, a more detailed analysis shows that the only case in which no such f_i and L_i exist occurs if $L_{i-3} \cup L_{i-2} \cup L_{i-1} \subseteq E(F^r_{i-1})$ is isomorphic to K_6. This happens if the edges f_{i-2} and f_{i-1} were chosen according to option $A.(iii)$ and consequently $f_{i-2} \cap f_{i-1} = \emptyset$, $L_{i-3} \cap L_{i-2} = \{f_{i-2}\}$, $L_{i-2} \cap L_{i-1} = \{f_{i-1}\}$ and if, moreover, $L_{i-1} \cap L_{i-3} = \{f\}$ for an edge f disjoint from f_{i-2} and f_{i-1}. Note that this implies in particular that $y_{i-1} = 2$, $x_{i-1} = 1$ and, since Grow did not terminate in step $i-1$, $y_j = x_j = 0$ for all $1 \leq j \leq i-2$.

This is not just a shortcoming of our analysis of Bal-Edge-Col. It is not hard to see that, for m as claimed, copies of K_6 appear with positive probability in $G^r_{n,m}$ and that, moreover, the algorithm Bal-Edge-Col as stated in Figure 1 is indeed unable to deal with these, despite the fact that K_6 is obviously colorable without creating monochromatic copies of K_4.

We now describe how this issue can be overcome. Let K_6^{r*} denote the r-matched graph obtained by embedding every edge e of K_6 into a r-set $K(e)$ such that the underlying unmatched graph consists of a copy of K_6 and $15(r-1)$ isolated edges. Standard first moment calculations show that a.a.s. every copy of K_6 in $G^r_{n,m}$ is contained in a copy of K_6^{r*}, and that moreover all copies of K_6^{r*} are vertex-disjoint. Consider now the algorithm Bal-Edge-Col-K_4 that proceeds exactly as Bal-Edge-Col, except that if the first loop gets stuck in line 14, it checks whether the current graph G' is a collection of vertex-disjoint copies of K_6^{r*}. If so, it colors these copies 'by hand' and then starts the second loop as usual; if not, it terminates with an error. It follows with the same arguments as before that Bal-Edge-Col-K_4 finds a valid coloring of any r-matched graph G^r if it terminates correctly.

In order to analyse Bal-Edge-Col-K_4, we consider a modified algorithm Grow-K_4. At the beginning, Grow-K_4 makes sure that the copy $F_1 = L_1 \in \mathcal{L}^*_{G'}$ it starts with (that is picked arbitrarily in Grow) is not contained in a copy of K_6. It then proceeds exactly as Grow until it either terminates regularly or it encounters a K_6 as described above. In the latter case, it performs a single exceptional step and picks the next edge f_i not in L_{i-1} (as Grow would try and fail to do), but chooses any edge in F^r_{i-4} that is different from $f_1, \ldots, f_{i-4}$ instead. Since $x_j = y_j = 0$ for all $1 \leq j \leq i-2$, F^r_{i-4} is free from any undesired intersections, and either f_i itself or some other edge in $K(f_i)$ is contained in a copy of $H = K_4$ that adds at least one new edge to F^r_{i-1}. After this exceptional step, Grow-K_4 continues to operate exactly like Grow. Note that if it encounters a second (or for a second time the same) copy of K_6 in step i', it terminates due to $\sum_{1 \leq j \leq i'} y_j \geq 3$. The proof of Lemma 8 can now be completed as before (note that Claim 12 still holds for Grow-K_4 since the choice of f_i in the single exceptional step contributes a factor of at most $\widetilde{C} i^2$ to $|\mathcal{F}^r(K_4, i)|$ for some constant $\widetilde{C} = \widetilde{C}(r)$). This settles the case $H = K_4$.

Very similar remarks apply to the case $H = C_4$. The only case in which we cannot find $f_i \in \mathcal{K}(L_{i-1})$ and a copy L_i containing f_i that is not contained completely in $E(F^r_{i-1})$ occurs if $L_{i-4} \cup L_{i-3} \cup L_{i-2} \cup L_{i-1} \subseteq E(F^r_{i-1})$ is isomorphic to the three-dimensional cube D_3. Again this is not just an issue in our analysis since copies of D_3 appear with positive probability in $G^r_{n,m}$ with m as claimed and cannot be handled by Bal-Edge-Col. The analogous tweak as in the case $H = K_4$ yields algorithms Bal-Edge-Col-C_4 and Grow-C_4 for which the proof can be completed as before.

3. Upper bound for the offline Achlioptas problem

3.1. **Preliminaries.** We will use the Azuma-Hoeffding inequality. Here we present the formulation given in [4].

Theorem 14 ([1, 3])**.** *If $(X_k)_0^n$ is a martingale with $X_n = X$ and $X_0 = \mathbb{E}[X]$, and there exist constants $c_k > 0$ such that*

$$|X_k - X_{k-1}| \leq c_k$$

for each $k \leq n$, then, for every $t > 0$,

$$\Pr[X \geq \mathbb{E}[X] + t] \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right) \ ,$$
$$\Pr[X \leq \mathbb{E}[X] - t] \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right) \ . \tag{24}$$

3.2. **Proof.** As mentioned in the introduction, we proceed by induction on $e(H)$ and prove the following strengthening of the 1-statement of Theorem 5. Note that H may be disconnected and even contain isolated vertices. For convenience we define $m_2(H) := 1/2$ if $e(H) = 1$.

Theorem 15. *Let H be a fixed graph with at least one edge, and let r be a fixed integer. There exist positive constants $C = C(H,r)$ and $a = a(H,r)$ such that for $m \geq Cn^{2-1/m_2(H)}$ with $m \ll n^2$, a.a.s $G^r_{n,m}$ has the property that every Achlioptas subgraph $A \sqsubset G^r_{n,m}$ contains $an^{v_H}(m/n^2)^{e_H}$ many copies of H.*

Proof. Note that any $A \sqsubset G^r_{n,m}$ has exactly m/r edges. Thus for the base case $e(H) = 1$, the statement holds deterministically for any $C > 0$ (in fact, for any $m \geq 0$) and $a = 1/(2r)$. For the induction step, we start by fixing some constants. Fix an arbitrary subgraph $H_- \subset H$ with $e_H - 1$ edges and v_H vertices, and set

$$a' := \min\left\{a(H_-, r),\ 1\right\} \ , \qquad C' := \max\left\{C(H_-, r),\ (3/a')^{1/(e_H-1)},\ 1\right\} \ , \tag{25}$$

where $a(H_-, r)$ and $C(H_-, r)$ are the constants guaranteed inductively by the theorem. Furthermore, let

$$\gamma := \frac{(a')^2}{12 \cdot 2^{20v_H^2}} < 1 \ , \qquad b := \frac{\gamma^{2r}}{8r} < 1 \ . \tag{26}$$

We shall prove Theorem 15 for

$$C := (1 + b^{-1})\, C' \ , \qquad a := \frac{a'}{(1 + b^{-1})^{e_H}} \ . \tag{27}$$

We proceed by a two-round approach as follows. For given $m \geq Cn^{2-1/m_2(H)}$, set

$$m_1 := \frac{1}{1 + b^{-1}}\, m \ , \qquad m_2 := \frac{b^{-1}}{1 + b^{-1}}\, m \ . \tag{28}$$

Note that $m_1 + m_2 = m$ and

$$m_1 \geq C'n^{2-1/m_2(H)} \geq C'n^{2-1/m_2(H_-)} \ . \tag{29}$$

We assume w.l.o.g. that both m_1 and m_2 are integers divisible by r. In the first round we generate a random r-matched graph G^r_{n,m_1} and ask the adversary to pick an Achlioptas subgraph $A_1 \sqsubset G^r_{n,m_1}$. In the second round, we add another m_2/r random r-sets (uniformly at random from all edges in $E(K_n) \setminus E(G^r_{n,m_1})$) and ask the adversary to extend her choice of $A_1 \sqsubset G^r_{n,m_1}$ to an Achlioptas subgraph A of the resulting r-matched graph $G^r_{n,m}$.

For a fixed choice of $A_1 \sqsubset G^r_{n,m_1}$ and any $e \in E(K_n)$, let x_e denote the number of copies of H_- in A_1 that e completes to copies of H, and set

$$\Gamma(A_1) := \{e \in E(K_n) \mid x_e \geq a' n^{v_H-2}(m_1/n^2)^{e_H-1}\} \ . \tag{30}$$

We shall prove the following two claims.

Claim 16. *A.a.s. G^r_{n,m_1} has the property that for every Achlioptas subgraph $A_1 \sqsubset G^r_{n,m_1}$ we have*

$$|\Gamma(A_1)| \geq \gamma n^2 \ .$$

Claim 17. *If G^r_{n,m_1} is as in Claim 16, for any fixed choice of $A_1 \sqsubset G^r_{n,m_1}$ we have that with probability at least $1-\mathrm{e}^{-bm_2}$, at least bm_2 many r-sets of the second round are completely in $\Gamma(A_1)$.*

These two claims imply the induction step of Theorem 15 as follows. A.a.s. G^r_{n,m_1} is as in Claim 16. Since there are $r^{m_1/r}$ choices for $A_1 \sqsubset G^r_{n,m_1}$, Claim 17 guarantees that with probability at least

$$1 - r^{m_1/r} \cdot \mathrm{e}^{-bm_2} \overset{(28)}{=} 1 - (r^{1/r}\,\mathrm{e}^{-1})^{m_1} = 1 - o(1) \ ,$$

at least bm_2 many r-sets of the second round are completely in $\Gamma(A_1)$ for *every* choice of $A_1 \sqsubset G^r_{n,m_1}$. Moreover, since one edge from every r-set must be included in $A \sqsubset G^r_{n,m}$, by the definition of $\Gamma(A_1)$ (cf. (30)), a.a.s. at least

$$bm_2 \cdot a' n^{v_H-2}(m_1/n^2)^{e_H-1} \overset{(27),(28)}{=} a n^{v_H}(m/n^2)^{e_H}$$

copies of H are created in the second round. This concludes the proof of Theorem 15 (which, as discussed, implies the 1-statement of Theorem 5). □

It remains to prove Claim 16 and Claim 17.

Proof of Claim 16. We will use several times that, by definition of $m_2(H)$, for every subgraph $J \subseteq H$ with at least one edge we have

$$n^{v_J - 2 - \frac{e_J - 1}{m_2(H)}} \geq n^0 = 1 \ . \tag{31}$$

Consider a fixed choice of $A_1 \sqsubset G^r_{n,m_1}$, and recall that for every edge $e \in E(K_n)$, x_e denotes the number of copies of H_- in A_1 that e completes to copies of H. Let $k(H_-)$ denote the number of copies of H_- in A_1. Since every copy of H_- in A_1 contributes to at least one of the x_e, we have

$$\sum_{e \in E(K_n)} x_e \geq k(H_-) \ .$$

By definition of $\Gamma(A_1)$ (cf. (30)) it follows that

$$\begin{aligned}\sum_{e \in \Gamma(A_1)} x_e &\geq k(H_-) - \binom{n}{2} \cdot a' n^{v_H-2}(m_1/n^2)^{e_H-1} \\ &\geq k(H_-) - a'/2 \cdot n^{v_H}(m_1/n^2)^{e_H-1} \ .\end{aligned}$$

Due to (25) and (29), we have by induction that a.a.s. $k(H_-) \geq a' n^{v_H}(m_1/n^2)^{e_H-1}$ and, consequently,

$$\sum_{e \in \Gamma(A_1)} x_e \geq a'/2 \cdot n^{v_H}(m_1/n^2)^{e_H-1} \ . \tag{32}$$

Note that in fact the induction hypothesis guarantees that a.a.s. this bound holds for all choices of $A_1 \sqsubset G^r_{n,m_1}$ simultaneously.

Let $\mathcal{T}$ be the family of all pairwise nonisomorphic graphs T which are unions of two copies of H_-, say $H_-^1 \cup H_-^2$, such that for some edge $f \in \binom{V(T)}{2} \setminus E(T)$, both $H_-^1 \cup \{f\}$ and $H_-^2 \cup \{f\}$ are isomorphic to H. Let $k(\mathcal{T})$ denote the number of copies of graphs from $\mathcal{T}$ in A_1. We have

$$\sum_{e \in E(K_n)} \binom{x_e}{2} \leq 2^{17v_H^2} \cdot k(\mathcal{T}) \ , \tag{33}$$

where the constant $2^{17v_H^2}$ is due to the fact that a given copy of some $T \in \mathcal{T}$ contributes at most $\binom{v_T}{v_H-2}^2 \binom{e_T}{e_H-1}^2 \leq 2^{2v_T+2e_T} \leq 2^{4v_T^2} \leq 2^{16v_H^2}$ to at most $\binom{v_T}{2} \leq 2v_H^2 \leq 2^{v_H^2}$ terms of the sum.

For a fixed graph $T \in \mathcal{T}$, let $I = H_-^1 \cap H_-^2$ denote the intersection of the two copies of H_-, and let J denote the graph obtained by adding the edge f to I (if there are multiple choices for H_-^1, H_-^2, and f, pick one arbitrarily). Letting the random variable X_T denote the number of copies of T in G_{n,m_1} (the underlying unmatched graph of G^r_{n,m_1}), we have

$$\begin{aligned}\mathbb{E}[X_T] \leq& (1+o(1)) n^{2v_H - v_I} (2m_1/n^2)^{2(e_H-1)-e_I} \leq 2^{2e_H} n^{2v_H - v_J} (m_1/n^2)^{2e_H-1-e_J} \\ &\leq 2^{v_H^2} \left(n^{2v_H-2}(m_1/n^2)^{2e_H-2}\right) / \left(n^{v_J-2}(m_1/n^2)^{e_J-1}\right) \overset{(25),(29),(31)}{\leq} 2^{v_H^2} n^{2v_H-2}(m_1/n^2)^{2e_H-2} \ .\end{aligned} \tag{34}$$

If $\mathbb{E}[X_T]$ is a growing function of n, then by a standard application of Chebyshev's inequality, we have $X_T \leq 2\mathbb{E}[X_T]$ a.a.s. Otherwise, Markov's inequality implies that for any growing function $\omega(n)$ we have $X_T \leq \omega(n)$ a.a.s. As the right hand side of (34) is indeed a growing function of n (due to (29) and (31), it grows at least quadratically), we obtain that in either case X_T exceeds the r.h.s. of (34) at most by a factor of two a.a.s. Since A_1 is a subgraph of G_{n,m_1}, it follows in particular that a.a.s.

$$\begin{aligned} k(\mathcal{T}) &\leq |\mathcal{T}| \cdot 2 \cdot 2^{v_H^2} n^{2v_H-2}(m_1/n^2)^{2e_H-2} \\ &\leq 2^{3v_H^2} \cdot n^{2v_H-2}(m_1/n^2)^{2e_H-2} \ , \end{aligned} \tag{35}$$

where we bounded $|\mathcal{T}|$ by the number of graphs on at most $2v_H$ vertices, which in turn is bounded by $\sum_{i=1}^{2v_H} 2^{\binom{i}{2}} \leq 2^{\binom{2v_H}{2}+1} \leq 2^{2v_H^2-1}$. Again a.a.s. the bound (35) holds for all choices of $A_1 \sqsubset G^r_{n,m_1}$ simultaneously. From (33) and (35) we obtain

$$\sum_{e \in \Gamma(A_1)} \binom{x_e}{2} \leq \sum_{e \in E(K_n)} \binom{x_e}{2} \leq 2^{20v_H^2} \cdot n^{2v_H-2}(m_1/n^2)^{2e_H-2} \ . \tag{36}$$

Jensen's inequality now yields

$$\sum_{e \in \Gamma(A_1)} \binom{x_e}{2} \geq |\Gamma(A_1)| \binom{\frac{\sum_{e\in\Gamma(A_1)} x_e}{|\Gamma(A_1)|}}{2} \geq \frac{\left(\sum_{e\in\Gamma(A_1)} x_e\right)^2}{3|\Gamma(A_1)|} \ , \tag{37}$$

where in the last step we used that for all $e \in \Gamma(A_1)$ we have

$$x_e \overset{(30)}{\geq} a' n^{v_H-2}(m_1/n^2)^{e_H-1} \overset{(29)}{\geq} a'(C')^{e_H-1} n^{v_H-2-\frac{e_H-1}{m_2(H)}} \overset{(25),(31)}{\geq} 3 \ .$$

It follows that a.a.s.

$$\begin{aligned} |\Gamma(A_1)| &\overset{(37)}{\geq} \frac{\left(\sum_{e\in\Gamma(A_1)} x_e\right)^2}{3\sum_{e\in\Gamma(A_1)} \binom{x_e}{2}} \\ &\overset{(26),(32),(36)}{\geq} \gamma n^2 \end{aligned}$$

for every choice of $A_1 \sqsubset G^r_{n,m_1}$. This concludes the proof of Claim 16. $\square$

Proof of Claim 17. Conditioning on G^r_{n,m_1} being as in Claim 16 and considering a fixed choice of $A_1 \sqsubset G^r_{n,m_1}$, we let X denote the number of r-sets of the second round that are completely in $\Gamma(A_1)$. From Claim 16 we obtain

$$\mathbb{E}[X] = m_2/r \cdot \frac{\binom{|\Gamma(A_1)| \setminus E(G^r_{n,m_1})}{r}}{\binom{\binom{n}{2}-m_1}{r}} = m_2/r \cdot (1+o(1)) \left(\frac{2|\Gamma(A_1)|}{n^2}\right)^r \overset{\text{Cl. 16}}{\geq} m_2 \cdot \frac{\gamma^r}{r} \ , \tag{38}$$

where we used that $|\Gamma(A_1)| = \Theta(n^2)$ and $m_1 \ll n^2$. We now apply the Azuma-Hoeffding inequality (Theorem 14) with $t = \mathbb{E}[X]/2$ to the Doob Martingale that arises if we draw the m_2/r many r-sets of the second round sequentially, using that each of these r-sets changes X by at most $c_k = 1$, $1 \leq k \leq m_2/r$. Observing that by (26) and (38) we have $bm_2 \leq \mathbb{E}[X]/2$, we obtain

$$\Pr[X \leq bm_2] \leq \Pr[X \leq \mathbb{E}[X]/2] \overset{(24)}{\leq} \exp\left(-\frac{(\mathbb{E}[X]/2)^2}{2 \cdot m_2/r}\right) \overset{(26),(38)}{\leq} \exp(-bm_2) \ .$$

This concludes the proof of Claim 17. □

Acknowledgement. The authors thank the anonymous referees for many helpful comments.

References

[1] K. Azuma. Weighted sums of certain dependent random variables. *Tôhoku Math. J. (2)*, 19:357–367, 1967.

[2] E. Friedgut, Y. Kohayakawa, V. Rödl, A. Ruciński, and P. Tetali. Ramsey games against a one-armed bandit. *Combin. Probab. Comput.*, 12(5-6):515–545, 2003. Special issue on Ramsey theory.

[3] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[4] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[5] M. Krivelevich, P.-S. Loh, and B. Sudakov. Avoiding small subgraphs in Achlioptas processes. *Random Structures Algorithms*, 34(1):165–195, 2009.

[6] T. Łuczak, A. Ruciński, and B. Voigt. Ramsey properties of random graphs. *J. Combin. Theory Ser. B*, 56(1):55–68, 1992.

[7] M. Marciniszyn, D. Mitsche, and M. Stojaković. Online balanced graph avoidance games. *European J. Combin.*, 28(8):2248–2263, 2007.

[8] M. Marciniszyn and R. Spöhel. Online vertex colorings of random graphs without monochromatic subgraphs. In *Proceedings of the 18th annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2007)*, pages 486–493, 2007. Journal version accepted for publication in *Combinatorica*.

[9] M. Marciniszyn, R. Spöhel, and A. Steger. Online Ramsey games in random graphs. *Combin. Probab. Comput.*, 18(1–2):271–300, 2009.

[10] M. Marciniszyn, R. Spöhel, and A. Steger. Upper bounds for online Ramsey games in random graphs. *Combin. Probab. Comput.*, 18(1–2):259–270, 2009.

[11] T. Mütze, R. Spöhel, and H. Thomas. Small subgraphs in random graphs and the power of multiple choices. Submitted for publication.

[12] A. Prakash, R. Spöhel, and H. Thomas. Balanced online Ramsey games in random graphs. *Electron. J. Combin.*, 16(1):Research Paper 11, 22, 2009.

[13] V. Rödl and A. Ruciński. Lower bounds on probability thresholds for Ramsey properties. In *Combinatorics, Paul Erdős is eighty, Vol. 1*, Bolyai Soc. Math. Stud., pages 317–346. János Bolyai Math. Soc., Budapest, 1993.

[14] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[15] J. Spencer and N. Wormald. Birth control for giants. *Combinatorica*, 27(5):587–628, 2007.