Colour-biased Hamilton cycles in random graphs

Lior Gishboliner   Michael Krivelevich   Peleg Michaeli

Abstract

We prove that a random graph $G(n, p)$, with p above the Hamiltonicity threshold, is typically such that for any r-colouring of its edges there exists a Hamilton cycle with at least $(2/(r+1) - o(1))n$ edges of the same colour. This estimate is asymptotically optimal.

1   Introduction

Hamiltonicity is one of the most flourishing and well-studied areas of research in the theory of random graphs, boasting a wide array of results over hundreds of papers. In fact, the question of finding the threshold for containing a Hamilton path has already been posed by Erdős and Rényi in their seminal paper on random graphs [13]. Building on the breakthrough work of Pósa [36], which introduced a method now known as *Pósa's rotation–extension technique*, Komlós and Szemerédi [27] and independently Bollobás [8] proved the fundamental result that the threshold for the appearance of a Hamilton cycle in the binomial random graph $G(n, p)$ is $p = (\log n + \log \log n)/n$. For a historical overview and a list of papers on this topic we refer the reader to an annotated bibliography by Frieze [17].

A central theme in this area is that the appearance of Hamilton cycles is closely tied to the disappearance of vertices of degree at most 1. In fact, having minimum degree 2 is often thought of as the "bottleneck" for the appearance of Hamilton cycles. This perspective is made remarkably precise in the hitting time results of Ajtai, Komlós and Szemerédi [1] and of Bollobás [8] (see also the survey [28] for a shorter proof, and [3] for yet another quantitative aspect of this phenomenon).

With the threshold for Hamiltonicity known, it is natural to ask about the typical structure of the set of Hamilton cycles appearing in $G(n, p)$, for p which is just above the Hamiltonicity threshold. For such values of p, one might expect the number of Hamilton cycles in $G(n, p)$ to be small, and their structure sparse and fragile. It turns out, however, that this is quite far from the truth. In fact, the set of Hamilton cycles of $G(n, p)$ (for p as above) typically possesses a rich and robust structure. Several concrete manifestations of this phenomenon have been demonstrated in prior works. For example, it is known that the number of Hamilton cycles in $G(n, p)$ is — in some well-defined quantitative sense — concentrated around its mean [20]; that the set of Hamilton cycles in $G(n, p)$ typically possesses local resilience properties [29,33,34,38]; and that random edge-colourings of $G(n, p)$ typically admit Hamilton cycles coloured according to any prescribed pattern [4,15].

In this paper, we establish yet another natural "robustness property" of the set of Hamilton cycles in $G(n, p)$ (for any p above the Hamiltonicity threshold). The precise problem we will be studying is as follows. For a graph G and an integer $r \geq 2$, let $M(G, r)$ be the largest integer M such that in any r-colouring of the edges of G, there will be a Hamilton cycle with at least M edges of the same colour (if G is not Hamiltonian, we set $M(G, r) = 0$).

The problem of estimating $M(G, r)$ is somewhat similar to (though slightly different from) multicolour discrepancy problems. In the general setting of combinatorial discrepancy theory, one is given a hypergraph H and tries to r-colour its vertices in such a way that every hyperedge is coloured as evenly as possible, in the sense that the numbers of vertices of a given colour in every hyperedge e deviates from its "mean", $|e|/r$, by as little as possible. The discrepancy of H is then defined as the maximal deviation one is guaranteed to have in any colouring. In the special setting we consider here, the vertices of the hypergraph H are the edges of G, and the hyperedges of H are the Hamilton cycles in G. We note, however, that the problem of estimating $M(G, r)$ differs from its discrepancy variant in that $M(G, r)$ is only concerned with "one-sided deviations", namely with colours appearing significantly more (and not less) than what is expected. It is worth noting that discrepancy-type problems in graphs were studied for various "target subgraphs", such as cliques [14], spanning trees [5,12], Hamilton cycles [5] and clique factors [6].

It is natural to expect that if G contains only few Hamilton cycles, then one can r-colour the edges of G in such a way that every Hamilton cycle sees approximately the same number, i.e. roughly n/r, of edges of each colour. Our main result, Theorem 1.1, shows that the situation in $G(n, p)$ (for p above the Hamiltonicity threshold) is typically very different: one is always guaranteed to find a Hamilton cycle which contains significantly more than n/r edges of the same colour. As alluded to earlier, this is yet another indication of the rich structure of the set of Hamilton cycles in $G(n, p)$.

Before stating our main result, let us recall some standard terminology. For a positive integer n and a real $p \in [0, 1]$, denote by $G(n, p)$ the **binomial random graph**, namely, the probability space of all simple labelled graphs on n given vertices, where each pair of vertices is connected by an edge independently with probability p. We say that an event $\mathcal{A}$ in our probability space occurs **with high probability** (or **whp**) if $\mathbb{P}(\mathcal{A}) \to 1$ as n goes to infinity.

Theorem 1.1. *Let $r \geq 2$ be an integer and let $p \geq (\log n + \log \log n + \omega(1))/n$. Then $G \sim G(n, p)$ is **whp** such that in any r-colouring of its edges there exists a Hamilton cycle with at least $(2/(r + 1) - o(1))n$ edges of the same colour.*

Using similar tools to those used in the proof of Theorem 1.1, we sketch a proof for the following analogous result for perfect matchings.

Theorem 1.2. *Let $r \geq 2$ be an integer and let $p \geq (\log n + \omega(1))/n$. Then, assuming n is even, $G \sim G(n, p)$ is **whp** such that in any r-colouring of its edges there exists a perfect matching with at least $(1/(r + 1) - o(1))n$ edges of the same colour.*

The fraction $1/(r+1)$ in Theorem 1.2, and hence also the fraction $2/(r+1)$ in Theorem 1.1, is tight. In fact, in *every* n-vertex graph G there exists an r-colouring in which in every matching, the maximum number of edges of the same colour is at most $n/(r+1)$. Such a colouring, which to the best of our knowledge first appeared in [9], can be described as follows. Partition $V(G)$ into sets $V_1, \ldots, V_r$ such that $|V_i| = n/(r+1)$ for $i = 1, \ldots, r-1$ and $|V_r| = 2n/(r+1)$. For $i = 1, \ldots, r$ (in increasing order), colour by i every edge touching V_i that has not already been coloured. Namely, for each $1 \leq i \leq r$, all edges contained in $V_i \cup \cdots \cup V_r$ and touching V_i are coloured with colour i (see Fig. 1). It is easy to see that any monochromatic matching in this colouring is of size at most $n/(r+1)$. Moreover, observe that any Hamilton cycle contains at most $2n/(r+1)$ edges of a given colour, as otherwise it would also contain a matching of size larger than $n/(r+1)$, hence also $M(G, r) \leq 2n/(r+1)$.
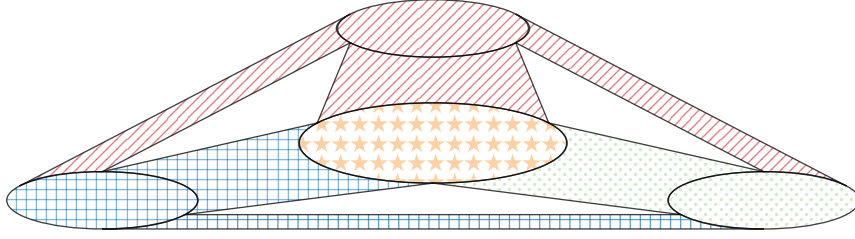
Figure 1: A 4-coloured complete graph on n vertices. Each of the small bulbs represents a set of $n/5$ vertices, and the large bulb in the middle represents a set of $2n/5$ vertices. Any monochromatic matching in this graph is of size at most $n/5$, hence also every Hamilton cycle contains at most $2n/5$ edges of the same colour.

The above construction and its analysis suggest a connection between the problem of estimating $M(G, r)$ and the problem of finding monochromatic matchings in r-colourings of (the edges of) G. Indeed, our proof of Theorem 1.1 relies on a new Ramsey-type result for matchings, which may be of independent interest.

A classical theorem of Cockayne and Lorimer [9] states that for integers $k_1, \ldots, k_r \geq 1$ and $n \geq \sum_{i=1}^{r} (k_i - 1) + \max\{k_1, \ldots, k_r\} + 1$, every r-colouring of the edges of the complete graph K_n contains a monochromatic matching of size k_i in colour i for some $1 \leq i \leq r$. The following theorem extends this result to *almost complete* host graphs.

Theorem 1.3. *Let $r \geq 2$, let $k_1, \ldots, k_r \geq 1$, let $0 \leq \delta \leq \frac{1}{2(r+1)}$, let G be a graph with n vertices and at least $(1 - \delta)\binom{n}{2}$ edges, and suppose that $(1 - (r+1)\delta)n \geq \sum_{i=1}^{r} (k_i - 1) + k + 1$, where $k := \max\{k_1, \ldots, k_r\}$. Then, for every r-colouring of the edges of G, there is $1 \leq i \leq r$ such that G contains a matching of size k_i, all of whose edges are coloured with colour i.*

For $k_1 = \cdots = k_r = k$, the condition in Theorem 1.3 becomes $(1 - (r+1)\delta)n \geq (r+1)(k-1) + 2$, which is satisfied if $k \leq (\frac{1}{r+1} - \delta)n$. Hence, for this case we have the following corollary.

Corollary 1.4. *Let $r \geq 2$, let $0 \leq \delta \leq \frac{1}{2(r+1)}$, and let G be a graph on n vertices and at least $(1 - \delta)\binom{n}{2}$ edges. Then, in every r-colouring of the edges of G there is a monochromatic matching of size at least $\lfloor (\frac{1}{r+1} - \delta)n \rfloor$.*

Our proof of Theorem 1.3 is inspired by a new proof of the Cockayne–Lorimer theorem, given in [39].

As a second step towards proving Theorem 1.1, we will combine Corollary 1.4 with a multicolour version of the sparse regularity lemma (stated here as Theorem 3.2) to prove that in any r-colouring of the edges of an n-vertex *pseudorandom* graph G, there must be a path of length $(2/(r + 1) - o(1))n$ in which all but a fixed number of edges are of the same colour. We will postpone the precise definition of pseudorandomness to Section 3, and for now only note that as a bi-product, we get the following aesthetically pleasing result.

Theorem 1.5. *Let $r \geq 2$ be an integer and let $\varepsilon > 0$. Then there exist $C = C(r, \varepsilon)$ and $K = K(r, \varepsilon)$ such that if $p \geq C/n$, the random graph $G \sim G(n, p)$ is **whp** such that in any r-colouring of its edges there exists a path of length at least $(2/(r + 1) - \varepsilon)n$ in which all but at most K of the edges are of the same colour.*

It is interesting to note that Theorems 1.1 and 1.5 are nontrivial (and new) even for the extreme case $p = 1$, i.e., where the coloured graph is the complete graph. An immediate corollary of Theorem 1.5 is that for large enough $C = C(r, \varepsilon)$, **whp** there is a monochromatic matching of size $(1/(r + 1) - \varepsilon)n$ in any r-colouring of the edges of $G(n, C/n)$. In this sense, Theorem 1.5 is again optimal, as explained before.

A closely related and in fact relatively well studied problem is that of finding long *monochromatic* paths in edge-colourings of graphs (which corresponds to requiring $K = 0$ in Theorem 1.5). For two colours, this problem was resolved by Gerencsér and Gyárfás in [19] for complete graphs and by Letzter in [30] for random graphs. For $r \geq 3$ colours, it is conjectured that every edge-colouring of K_n contains a monochromatic path of length $(1/(r-1)-o(1))n$, and that the same holds **whp** for $G(n, p)$ with $np \to \infty$ (see, e.g., [11]). It is known that if true, this would be best possible (even in the complete graph; see, e.g., [25] and the references therein). This conjecture was resolved for $r = 3$ by Gyárfás, Ruszinkó, Sárközy and Szemerédi in [21, 22] (for the complete graph) and by Dudek and Prałat in [11] (for random graphs), and it remains open for all $r \geq 4$. Accidentally, for $r = 2, 3$ the two problems — that of finding a large monochromatic path and that of finding a large path in which all but a constant number of the edges are of the same colour — have the same answer (both in random and in complete graphs; this follows from Theorem 1.5 and the aforementioned results of [11, 19, 21, 22, 30]). For $r \geq 4$, however, these two problems diverge; allowing a fixed number of edges to be coloured differently significantly increases the length of a path one can find, from at most $(1/(r-1)+o(1))n$ for monochromatic paths to $(2/(r + 1) - o(1))n$ for *almost monochromatic* ones.

A common technique for finding long monochromatic paths, pioneered by Figaj and Łuczak in [16] (following an idea by Łuczak [32]), consists of applying the (sparse) regularity lemma and finding large monochromatic *connected matchings* in the reduced graph of a regular partition. In contrast, in order to find an almost monochromatic path, it is sufficient to find a monochromatic (not necessarily connected) matching in the reduced graph. One can expect — and we show that this is indeed the case — that in almost complete graphs (such as the reduced graphs we consider here), one can find substantially larger monochromatic matchings when dropping the requirement that they be connected. As mentioned above, this is a key step in the proof of Theorem 1.5.

Let us now say a few words about the remaining ingredients which go into the proof of Theorem 1.1. With Theorem 1.5 at hand, the proof of Theorem 1.1 proceeds as follows. Theorem 1.5 gives us a path P of length $(2/(r + 1) - o(1))n$ in which all but a fixed number of edges are of the same colour. Our goal is therefore to extend this path into a Hamilton cycle, or, equivalently, to find a Hamilton path in the remaining set of vertices between neighbours of the endpoints of P. We achieve this by carefully splitting the remaining vertices into two equal sets, each containing many neighbours of the corresponding endpoint of P, so that the minimum degree of the graph spanned by each of these sets is at least 2. In fact, to do so we need to "prepare" our graph, putting aside small degree vertices with their neighbours, and finding P outside this set. We thus want to find a suitable path not in our random graph but rather in some large induced subgraph thereof; hence we need a generalisation of Theorem 1.5 to pseudorandom graphs, Theorem 3.1. We continue by showing that in each of the two above-mentioned sets there are many Hamilton paths which start at a given point (a neighbour of the corresponding endpoint of P), or, more precisely, Hamilton paths with many distinct ends. The argument relies on the so-called rotation-extension technique, which was invented by Pósa in [36] and has since been applied in numerous papers about Hamiltonicity of random graphs. We conclude our proof by using expansion properties of our graph to connect the ends of two

such Hamilton paths, thus extending P to a Hamilton cycle.

Organisation We begin by proving Theorem 1.3 in Section 2. In Section 3 we prove Theorem 3.1, a generalisation of Theorem 1.5 to pseudorandom graphs. At the end of the section we show how to connect the monochromatic linear forest we obtain to a long path, almost all of whose edges are of the same colour. The goal of Section 4 is to introduce fairly general machinery to prepare random graphs in such a way that a path found in some (large) part of the graph can always be extended to a Hamilton cycle.

Notation and terminology Let $G = (V, E)$ be a graph. For two vertex sets $U, W \subseteq V$ we denote by $E_G(U)$ the set of edges of G spanned by U and by $E_G(U, W)$ the set of edges having one endpoint in U and the other in W. The degree of a vertex $v \in V$ is denoted by $d_G(v)$, and we write $d_G(v, U) = |E_G(\{v\}, U)|$. We let $\delta(G)$ and $\Delta(G)$ denote the minimum and maximum degrees of G. When the graph G is clear from the context, we may omit the subscript G in the notations above.

If f, g are functions of n we use the notation $f \sim g$ to denote asymptotic equality, namely, $f \sim g$ if $f = (1 + o(1))g$, and we write $f \ll g$ if $f = o(g)$. For the sake of simplicity and clarity of presentation, we often make no particular effort to optimise the constants obtained in our proofs, and omit floor and ceiling signs when they are not crucial.

2 Large monochromatic matchings in almost complete graphs

The goal of this section is to prove Theorem 1.3. The primary tool used in the proof is the well-known Tutte–Berge formula (see, e.g., [31]), which we state as follows. For a graph G, let $\nu(G)$ denote the maximum size of a matching in G, and $\mathrm{odd}(G)$ denote the number of connected components of G whose size is odd.

Theorem 2.1 (Tutte–Berge formula). *Every graph G satisfies*

$$\nu(G) = \frac{1}{2} \cdot |V(G)| - \frac{1}{2} \cdot \max_{U \subseteq V(G)} (\mathrm{odd}(G - U) - |U|).$$

We will also need the following simple lemma.

Lemma 2.2. *Let G be a graph with n vertices and t connected components. Then $|E(G)| \leq \binom{n-t+1}{2}$.*

Proof. Let G be as in the lemma, and let $C_1, \ldots, C_t$ be the connected components of G. Evidently, $|E(G)| \leq \sum_{i=1}^{t} \binom{|C_i|}{2}$. Thus, in order to prove the lemma, it suffices to show that the function $g(x_1, \ldots, x_t) = \sum_{i=1}^{t} \binom{x_i}{2}$ with domain $\{(x_1, \ldots, x_t) \in [n]^t \ : \ x_1 + \cdots + x_n = t\}$ attains its maximum when $x_1 = n - t + 1, x_2 = \cdots = x_t = 1$, where it equals $\binom{n-t+1}{2}$. So let $(x_1, \ldots, x_t) \in [n]^t$ be a maximum point of g. It is enough to show that there is (at most) one $1 \leq i \leq t$ such that $x_i \geq 2$. So suppose by contradiction that $x_i, x_j \geq 2$ for some distinct $i, j \in [t]$. Without loss of generality, assume that $x_j \geq x_i$. Now, setting $y_i := x_i - 1, y_j := x_j + 1$, and $y_k := x_k$ for $k \in [t] \setminus \{i, j\}$, observe that $g(y_1, \ldots, y_k) = g(x_1, \ldots, x_k) + x_j - (x_i - 1) \geq g(x_1, \ldots, x_k) + 1$, in contradiction to the choice of $x_1, \ldots, x_k$. □

Proof of Theorem 1.3

Let G be a graph with n vertices and at least $(1 - \delta)\binom{n}{2}$ edges, and suppose that

$$\left(1 - (r+1)\delta\right)n \geq \sum_{i=1}^{r} (k_i - 1) + k + 1, \tag{1}$$

where $k := \max\{k_1, \ldots, k_r\}$. We may assume that $k \geq \left(\frac{1}{r+1} - \delta\right)n - 1$, because otherwise (1) also holds with k_1 replaced by $k_1 + 1$ (which increases k by at most 1), meaning that we may instead prove the theorem for $k_1 + 1, k_2, \ldots, k_r$ (which evidently implies the statement for $k_1, \ldots, k_r$).

Fix any r-colouring of the edges of G. For each $i \in [r]$, let G_i be the graph on $V(G)$ whose edges are the edges of G which are coloured with colour i. Our goal is to show that there is $1 \leq i \leq r$ such that $\nu(G_i) \geq k_i$. So suppose, for the sake of contradiction, that $\nu(G_i) \leq k_i - 1$ for every $i \in [r]$. By Theorem 2.1, for each $i \in [r]$ there must be $U_i \subseteq V(G_i) = V(G)$ such that

$$\frac{n}{2} - \frac{1}{2} \cdot (\text{odd}(G_i - U_i) - |U_i|) = \nu(G_i) \leq k_i - 1,$$

or, equivalently, $\text{odd}(G_i - U_i) \geq n - 2(k_i - 1) + |U_i|$. In particular, $G_i - U_i$ has at least $n - 2(k_i - 1) + |U_i|$ connected components. This means that $n - |U_i| = |V(G_i - U_i)| \geq n - 2(k_i - 1) + |U_i|$, and hence $|U_i| \leq k_i - 1$. By Lemma 2.2, the following holds for every $i \in [r]$:

$$|E(G_i - U_i)| \leq \binom{(n - |U_i|) - (n - 2(k_i - 1) + |U_i|) + 1}{2} = \binom{2(k_i - 1) - 2|U_i| + 1}{2}.$$

It follows that

$$|E(G)| \leq \sum_{i=1}^{r} |E(G_i - U_i)| + \#\{e \in E(G) : e \cap (U_1 \cup \cdots \cup U_r) \neq \varnothing\}$$

$$\leq \sum_{i=1}^{r} \binom{2(k_i - 1) - 2|U_i| + 1}{2} + \binom{n}{2} - \binom{n - |U_1| - \cdots - |U_r|}{2}. \tag{2}$$

Now, consider the function $g(u_1, \ldots, u_r)$ defined by

$$g(u_1, \ldots, u_r) := \sum_{i=1}^{r} \binom{2(k_i - 1) - 2u_i + 1}{2} - \binom{n - u_1 - \cdots - u_r}{2}.$$

Claim 2.3. *Let $u_1, \ldots, u_r$ be such that $0 \leq u_i \leq k_i - 1$ for every $i \in [r]$. Then*

$$g(u_1, \ldots, u_r) < -\delta\binom{n}{2}.$$

Before proving Claim 2.3, let us complete the proof of Theorem 1.3 assuming this claim. Recall that $|U_i| \leq k_i - 1$ for every $i \in [r]$. Thus, by applying Claim 2.3 with $u_i = |U_i|$ ($i \in [r]$), we get that $g(|U_1|, \ldots, |U_r|) < -\delta\binom{n}{2}$. On the other hand, (2) states that $|E(G)| \leq \binom{n}{2} + g(|U_1|, \ldots, |U_r|)$, which contradicts our assumption that $|E(G)| \geq (1 - \delta)\binom{n}{2}$. Thus, in order to complete the proof it suffices to prove Claim 2.3.

Proof of Claim 2.3. It will be convenient to set $v_i := k_i - 1 - u_i$ for $i \in [r]$. Then $0 \le v_i \le k_i - 1$ for every $i \in [r]$. Note that the inequality $g(u_1, \ldots, u_r) < -\delta\binom{n}{2}$ is equivalent to having

$$h(v_1, \ldots, v_r) := \binom{n - \sum_{i=1}^r (k_i - 1) + \sum_{i=1}^r v_i}{2} - \sum_{i=1}^r \binom{2v_i + 1}{2} > \delta\binom{n}{2}. \tag{3}$$

For $1 \le i \le r$, observe that if we fix the values of $(v_j : j \in [r] \setminus \{i\})$ and let v_i vary, then the resulting function $h(v_1, \ldots, v_r)$ of v_i is a quadratic function in which the coefficient of v_i^2 is $-\frac{3}{2} < 0$. Therefore, this function is concave. It follows that for any choice of fixed values of $(v_j : j \in [r] \setminus \{i\})$, the minimum of $h(v_1, \ldots, v_r)$ over $0 \le v_i \le k_i - 1$ is obtained either at $v_i = 0$ or at $v_i = k_i - 1$, and is not obtained at any point in the open interval $(0, k_i - 1)$. We conclude that if $(v_1, \ldots, v_r)$ is a minimum point of $h(v_1, \ldots, v_r)$, then $v_i \in \{0, k_i - 1\}$ for every $1 \le i \le r$. So we see that in order to verify (3), it is enough to show that $h(v_1, \ldots, v_r) > \delta\binom{n}{2}$ for $v_1, \ldots, v_r$ satisfying $v_i \in \{0, k_i - 1\}$ for every $1 \le i \le r$.

Let $I \subseteq [r]$, and suppose that $v_i = k_i - 1$ for $i \in I$ and $v_i = 0$ for $i \in [r] \setminus I$. Then the value of $h(v_1, \ldots, v_r)$ is:

$$\binom{n - \sum_{i \in [r] \setminus I} (k_i - 1)}{2} - \sum_{i \in I} \binom{2k_i - 1}{2}$$

$$\ge \binom{\sum_{i \in I} (k_i - 1) + k + 1 + (r+1)\delta n}{2} - \sum_{i \in I} \binom{2k_i - 1}{2}$$

$$\ge \binom{\sum_{i \in I} (k_i - 1) + k}{2} + (k+1) \cdot (r+1)\delta n - \sum_{i \in I} \binom{2k_i - 1}{2}.$$

Here, the first inequality uses (1), and the second inequality follows from the fact that $\binom{x+y}{2} \ge \binom{x}{2} + xy$ for all $x, y \ge 0$. Now, since $k \ge (\frac{1}{r+1} - \delta)n - 1$ (as mentioned in the beginning of the proof) and $\delta \le \frac{1}{2(r+1)}$ (by assumption), we have

$$(k+1) \cdot (r+1)\delta n \ge \left(\frac{1}{r+1} - \delta\right)n \cdot (r+1)\delta n \ge \frac{n}{2(r+1)} \cdot (r+1)\delta n = \frac{\delta n^2}{2} > \delta\binom{n}{2}.$$

Thus, to establish Claim 2.3, it suffices to verify that

$$\binom{\sum_{i \in I} (k_i - 1) + k}{2} - \sum_{i \in I} \binom{2k_i - 1}{2} \ge 0. \tag{4}$$

Observe that for every $i \in I$, if we fix the values of $(k_j : j \in I \setminus \{i\})$ and consider the left-hand side of (4) as a one-variable function of k_i, then this function is quadratic and the coefficient of k_i^2 is $-3/2 < 0$. Thus, this function is concave. It follows that at a minimum point of the left-hand side of (4), we must have $k_i \in \{1, k\}$ for every $i \in I$ (recall that $k_i \le k$ for every $1 \le i \le r$). So let $J \subseteq I$, and suppose that $k_i = k$ for every $i \in J$ and $k_i = 1$ for every $i \in I \setminus J$. Setting $s := |J|$, we see that the left-hand side of (4) equals

$$\binom{s \cdot (k-1) + k}{2} - s \cdot \binom{2k - 1}{2} = \frac{(k-1)(s-1)}{2} \cdot ((s-1)k - s).$$

So it remains to show that $f(s) := (k-1)(s-1) \cdot ((s-1)k - s) \ge 0$ for every value of s. If $k = 1$ then $f(s) = 0$ for every s, so suppose that $k \ge 2$. Now, we have $f(0) = (k-1)k \ge 0$, $f(1) = 0$, and $f(s) \ge (s-1)k - s = (k-1)s - k \ge 2(k-1) - k \ge 0$ for every $s \ge 2$, as required. $\square$

With Claim 2.3 established, the proof of Theorem 1.3 is complete. □

It should be noted that a MathOverflow post due to F. Petrov [35] contains a derivation of the Cockayne–Lorimer result [9] using the Tutte–Berge formula in a similar manner to our proof of Theorem 1.3.

3   Large monochromatic linear forests in pseudorandom graphs

The goal of this section is to prove Theorem 1.5. In fact, we prove a stronger statement, namely Theorem 3.1 below. This theorem extends Theorem 1.5 to the more general setting of pseudorandom graphs, and will be used in the proof of Theorem 1.1.

Let us now introduce some definitions. For a pair of disjoint vertex-sets U, W in a graph, the **density** of (U, W) is defined as $d(U, W) := |E(U, W)|/(|U||W|)$. For $\gamma, p \in (0, 1]$, we say that $G = (V, E)$ is (γ, p)-**pseudorandom** if for any two disjoint $U, W \subseteq V$ with $|U|, |W| \geq \gamma|V|$ we have $|d(U, W) - p| \leq \gamma p$. We now recall the known fact that if $G = (V, E)$ is (γ, p)-pseudorandom then every set $U \subseteq V$ of size at least $2\gamma|V|$ satisfies

$$\left| \frac{|E(U)|}{\binom{|U|}{2}} - p \right| \leq \gamma p. \tag{5}$$

To see that (5) holds, take a random partition of U into two equal parts U_1, U_2 and observe that the expected value of $|E(U_1, U_2)|$ is $|E(U)| \cdot \binom{|U|}{2}^{-1} \cdot |U_1||U_2|$. On the other hand, we have $|d(U_1, U_2) - p| \leq \gamma p$ for every such choice of U_1, U_2. Therefore,

$$\left| \frac{|E(U)|}{\binom{|U|}{2}} - p \right| = \left| \frac{\mathbb{E}|E(U_1, U_2)|}{|U_1||U_2|} - p \right| \leq \gamma p,$$

as required.

Note that if G is a (γ, p)-pseudorandom graph on n vertices (for any $p \in (0, 1]$) then there exists an edge between any two disjoint sets of size at least γn.

The following is the main result of this section, and will play an important role in the proof of Theorem 1.1.

Theorem 3.1. *Let $r \geq 2$ be an integer and let $\varepsilon > 0$. Then there exist $\gamma = \gamma(r, \varepsilon)$ and $K = K(r, \varepsilon)$ such that the following holds. Let $G = (V, E)$ be a (γ, p)-pseudorandom graph for some $p \in (0, 1]$, and suppose $|V| = n$ is large enough (in terms of r, ε). Then, in any r-colouring of the edges of G there exists a path of length at least $(2/(r + 1) - \varepsilon)n$ in which all but at most K of the edges are of the same colour.*

The proof of Theorem 3.1 relies on (a "multicolour" version of) the well-known *sparse regularity lemma*, proved by Kohayakawa [26] and Rödl (see [10]), and later in a stronger form by Scott [37]. To state this result, we now introduce some additional definitions. A pair (U, W) of disjoint vertex-sets is called (δ, q)-**regular** if for all $U' \subseteq U, W' \subseteq W$ with $|U'| \geq \delta|U|$ and $|W'| \geq \delta|W|$ it holds that $|d(U', W') - d(U, W)| \leq \delta q$. An **equipartition** of a set is a partition in which the sizes of any two parts differ by at most 1 (to keep the presentation clean, we will ignore divisibility issues and just assume that all parts have the same size). Let $G_1, \ldots, G_r$

be graphs on the same vertex-set V of size n. An equipartition $\{V_1, \ldots, V_t\}$ of V is said to be (δ)-**regular** with respect to $(G_1, \ldots, G_r)$ if for all but at most $\delta\binom{t}{2}$ of the pairs (V_i, V_j), $1 \le i < j \le t$, it holds that for every $\ell \in [r]$, the pair (V_i, V_j) is (δ, q)-regular in G_ℓ, where $q := (|E(G_1)| + \cdots + |E(G_r)|)/\binom{n}{2}$. We are now ready to state the multicolour sparse regularity lemma from [37].

Theorem 3.2 (Multicolour sparse regularity lemma [37])**.** *For every $r, t_0 \ge 1$ and $\delta \in (0, 1)$ there exists $T = T(r, t_0, \delta)$ such that for every collection $G_1, \ldots, G_r$ of graphs on the same vertex-set V, there is an equipartition of V which is (δ)-regular with respect to $(G_1, \ldots, G_r)$, and has at least t_0 and at most T parts.*

Another tool we will use in the proof of Theorem 3.1 is the following simple lemma from [7] (see Lemma 4.4 there).

Lemma 3.3. *Let $n, k \ge 1$ be integers, and let F be a bipartite graph with sides X, Y of size n each. Suppose that there is an edge between every pair of sets $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| = |Y'| = k$. Then F contains a path of length at least $2n - 4k$.*

The proof of Lemma 3.3 proceeds by a careful analysis of the DFS algorithm, an idea which originated in [7] and has since been widely used in the study of paths in random and pseudorandom graphs (see also [28] and [30, Corollary 2.1]).

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. Let $r \ge 2$ and let $\varepsilon \in (0, 1)$. Fix $\delta > 0$ to be small enough so that $\delta < 1/(4r)$ and $\big(1/(r+1) - 2\delta\big) \cdot (2 - 4\delta) \ge 2/(r+1) - \varepsilon/2$ (for this second requirement, choosing $\delta \le \varepsilon/12$ suffices). Set $t_0 := 1/\delta$, and let $T = T(r, t_0, \delta)$ be as in Theorem 3.2. We will prove the theorem with $\gamma = \gamma(r, \varepsilon) := \varepsilon/(4T)$ and $K = K(r, \varepsilon) := T$.

Let $p \in (0, 1]$ and let G be a (γ, p)-pseudorandom graph on n vertices (for some sufficiently large n). Set $q := |E(G)|/\binom{n}{2}$, and note that by (5) we have $(1 - \gamma)p \le q \le (1 + \gamma)p$. Let $f : E(G) \to [r]$ be an r-colouring of the edges of G. For each $i \in [r]$, let G_i be the graph on $V(G)$ whose edges are the edges of G coloured by colour i. Let $\{V_1, \ldots, V_t\}$ be a (δ)-regular equipartition with respect to $(G_1, \ldots, G_r)$, where $t_0 \le t \le T$. Let H be the graph on $[t]$ in which $\{i, j\} \in E(H)$ if and only if (V_i, V_j) is (δ, q)-regular in G_ℓ for every $\ell \in [r]$. The definition of a (δ)-regular partition implies that $|E(H)| \ge (1 - \delta)\binom{|V(H)|}{2}$.

We now define a "reduced" edge-colouring of H. Let $\{i, j\}$ be an edge of H. Since $|V_i| = |V_j| = n/t \ge n/T \ge \gamma n$, we have $d_G(V_i, V_j) \ge (1 - \gamma)p \ge p/2$ (as G is (γ, p)-pseudorandom). Since $d_G(V_i, V_j) = d_{G_1}(V_i, V_j) + \cdots + d_{G_\ell}(V_i, V_j)$, there must be some $\ell \in [r]$ such that $d_{G_\ell}(V_i, V_j) \ge p/(2r)$. Colour the edge $\{i, j\}$ by colour ℓ (if there is more than one possible colour, choose one arbitrarily).

Since $|E(H)| \ge (1 - \delta)\binom{|V(H)|}{2}$, Corollary 1.4 implies that H contains a monochromatic matching of size at least $\lfloor\big(1/(r+1) - \delta\big)t\rfloor \ge \big(1/(r+1) - \delta\big)t - 1 \ge \big(1/(r+1) - 2\delta\big)t$ where the inequality holds because $t \ge t_0 = 1/\delta$. Suppose, without loss of generality, that this matching is in colour 1, and denote its edge-set by M. Fix any $e = \{i, j\} \in M$. Since $\{i, j\}$ is an edge of H coloured with colour 1, it must be the case that $d_{G_1}(V_i, V_j) \ge p/(2r)$ and that (V_i, V_j) is (δ, q)-regular in G_1. Then for every $V_i' \subseteq V_i$, $V_j' \subseteq V_j$ with $|V_i'| \ge \delta|V_i|$ and $|V_j'| \ge \delta|V_j|$ it holds that $d_{G_1}(V_i', V_j') \ge d_{G_1}(V_i, V_j) - \delta q \ge p/(2r) - \delta q \ge p/(2r) - \delta(1 + \gamma)p \ge p/(2r) - \delta \cdot 2p > 0$, where the last inequality holds due to our choice of δ. So we see that G_1 contains an edge between every pair of sets $V_i' \subseteq V_i$, $V_j' \subseteq V_j$ with $|V_i'| \ge \delta|V_i| = \delta n/t$ and $|V_j'| \ge \delta|V_j| = \delta n/t$.

By Lemma 3.3 with $k := \delta n/t$, the bipartite subgraph of G_1 with sides V_i and V_j contains a path P_e of length at least $(2 - 4\delta)n/t$.

Observe that the paths $(P_e : e \in M)$ are pairwise-disjoint (as M is a matching in H), and that the number of vertices covered by these paths is at least

$$|M| \cdot (2 - 4\delta)n/t \geq \big(1/(r+1) - 2\delta\big)t \cdot (2 - 4\delta)n/t \geq (2/(r+1) - \varepsilon/2)n,$$

where the last inequality uses our choice of δ.

Finally, put $k = |M|$, noting that $k \leq t \leq T$, and enumerate the paths $(P_e : e \in M)$ as $P_1, \ldots, P_k$. For each $1 \leq i \leq |M|$, let A_i, B_i denote the first, respectively last, γn vertices of P_i. Since G is (γ, p)-pseudorandom, there exists an edge $e_i = \{b_i, a_{i+1}\}$ between $b_i \in B_i$ and $a_{i+1} \in A_{i+1}$ for every $i = 1, \ldots, k-1$. Let a_1 be the first vertex of P_1 and let b_k be the last vertex of P_k. Let P be the path obtained by concatenating (parts of) the paths $P_1, \ldots, P_k$ using the edges $e_1, \ldots, e_{k-1}$, namely,

$$P = a_1 \xrightarrow{P_1} b_1 \xrightarrow{e_1} a_2 \xrightarrow{P_2} b_2 \xrightarrow{e_2} \cdots \xrightarrow{e_{k-1}} a_k \xrightarrow{P_k} b_k.$$

It is easy to see that

$$|P| \geq |P_1| + \cdots + |P_k| - (2k - 2) \cdot \gamma n \geq (2/(r+1) - \varepsilon/2)n - 2T\gamma n \geq (2/(r+1) - \varepsilon)n,$$

where in the last inequality we used our choice of γ. Moreover, all edges of P except for $e_1, \ldots, e_{k-1}$ have the same colour. As $k \leq T = K$, the path P satisfies all the required properties, completing the proof. $\qquad\square$

In view of Theorem 3.1, in order to obtain Theorem 1.5 it is enough to prove that random graphs (with sufficiently high edge density) are **whp** pseudorandom.

Lemma 3.4. *For every $\gamma > 0$ there exists $C = C(\gamma) > 0$ such that if $p \geq C/n$ then $G \sim G(n, p)$ is* **whp** *(γ, p)-pseudorandom.*

In the proof of Lemma 3.4 and in several other proofs in the next section we will make use of the following version of Chernoff bounds (see, e.g., in, [24, Chapter 2]).

Theorem 3.5 (Chernoff bounds). *Let $X = \sum_{i=1}^n X_i$, where $X_i \sim \mathsf{Bernoulli}(p_i)$ are independent, and let $\mu = \mathbb{E}X = \sum_{i=1}^n p_i$. Let $0 < \alpha < 1 < \beta$. Then*

$$\mathbb{P}(X \leq \alpha\mu) \leq \exp(-\mu(\alpha \log \alpha - \alpha + 1)),$$
$$\mathbb{P}(X \geq \beta\mu) \leq \exp(-\mu(\beta \log \beta - \beta + 1)).$$

Proof of Lemma 3.4. Note that we may assume $\gamma > 0$ is arbitrarily small. Write $V = V(G)$. Fix disjoint U, W with $|U|, |W| \geq \gamma n$ and write $x = |U||W|/n^2 \geq \gamma^2$. Note that $X := |E(U, W)|$ is a binomial random variable with xn^2 trials and success probability p. Thus by Theorem 3.5 there exists $c = c(\gamma) > 0$ such that

$$\mathbb{P}(|d(U, W) - p| \geq \gamma p) = \mathbb{P}(|X - pxn^2| \geq \gamma pxn^2) \leq 2\exp(-cpn^2).$$

Taking $C = C(\gamma)$ to be large enough so that $C > 2/c$, say, we obtain by the union bound that

$$\mathbb{P}(\exists U, W \subseteq V, \ |U|, |W| \geq \gamma n : \ |d(U, W) - p| \geq \gamma p) \leq 4^n \cdot e^{-2n} = o(1). \qquad\square$$

With Lemma 3.4, the proof of Theorem 1.5 is now complete.

4 Extending paths to Hamilton cycles

The goal of this section is to give a general machinery to "prepare" a random graph (above the hamiltonicity threshold) in a way that any path found in some large portion of the graph can be extended, **whp**, to a Hamilton cycle. We will then use this machinery to extend the path obtained in Theorem 3.1 to a Hamilton cycle, proving Theorem 1.1. Throughout this section, we assume that n is large enough whenever needed. In addition, as the statement in Theorem 1.1 is clearly monotone in p, we will conveniently assume throughout this section that $np - \log n - \log \log n \ll \log \log n$.

Lemma 4.1. *Let $\varepsilon > 0$, let $p = (\log n + \log \log n + \omega(1))/n$ and let $G \sim G(n,p)$. Then, **whp**, there exists a partition $V(G) = V^\star \cup V'$ with $|V^\star| \leq \varepsilon n$ for which every path $P \subseteq V'$ with $|V(P)| \leq 2n/3$ can be extended to a Hamilton cycle in G.*

The proof of Lemma 4.1 uses Pósa's rotation–extension technique [36]. For an overview of the rotation–extension technique, we refer the reader to [28]. The following is an immediate consequence of Pósa's lemma [36] (see also [28]).

Lemma 4.2. *Let G be a graph and let $P = v_0, \ldots, v_t$ be a longest path in G. Then there exists a set $R \subseteq V(P) \setminus \{v_0\}$ with the following properties:*

1. *For every $v \in R$ there is a path P' with $V(P') = V(P)$ and with endpoints v_0 and v.*

2. *$|N(R)| \leq 2|R| - 1$.*

Proof. Take R to be the set of endpoints of paths obtained from P via a sequence of rotations fixing v_0. Pósa's lemma [36] (see also [28]) states that $|N(R)| \leq 2|R| - 1$. $\square$

Recall that a non-edge $\{x, y\}$ of G is called a booster if adding $\{x, y\}$ to G creates a graph which is either Hamiltonian or whose longest path is longer than that of G. For a positive integer k and a positive real α we say that a graph $G = (V, E)$ is a (k, α)-**expander** if $|N(U)| \geq \alpha|U|$ for every set $U \subseteq V$ of at most k vertices. The following is a widely-used fact stating that $(k, 2)$-expanders have many boosters. For a proof, see e.g. [28].

Lemma 4.3. *Let G be a connected $(k, 2)$-expander which contains no Hamilton cycle. Then G has at least $(k + 1)^2/2$ boosters.*

We now move on to establish some useful properties satisfied **whp** by $G(n, p)$ (for p as in Lemma 4.7).

Lemma 4.4. *Let $\varepsilon > 0$ be sufficiently small, let $p = (\log n + \log \log n + \omega(1))/n$, and let $G \sim G(n, p)$. Then, **whp**,*

(P1) *$\delta(G) \geq 2$ and $\Delta(G) \leq 10 \log n$;*

(P2) *No vertex $v \in V(G)$ with $d(v) < \log n/10$ is contained in a 3- or a 4-cycle, and every two distinct vertices $u, v \in V(G)$ with $d(u), d(v) < \log n/10$ are at distance at least 5 apart;*

(P3) *Every set $U \subseteq V(G)$ of size at most $\varepsilon n/100$ spans at most $\varepsilon|U| \log n/10$ edges.*

(P4) *There exist disjoint sets $U_1, U_2 \subseteq V(G)$ with $|U_1|, |U_2| \leq \varepsilon n$ for which the following hold for every $v \in V(G)$:*

(a) *If $d(v) \geq \log n/10$ then $d(v, U_1), d(v, U_2) \geq \varepsilon \log n/100$;*

(b) *If $d(v) \leq \log n/10$ then v and all of its neighbours are in U_1.*

Proof of (P1). For the minimum degree see, e.g., [18]. For the maximum degree, since $d(v) \sim \mathsf{Bin}(n-1, p)$ we have

$$\mathbb{P}(d(v) \geq 10 \log n) \leq \binom{n}{10 \log n} p^{10 \log n} \leq \left(\frac{enp}{10 \log n} \right)^{10 \log n} \ll 1/n,$$

and the statement follows by the union bound. □

Proof of (P2). Write $V = V(G)$ and $\alpha = 1/10$. Let $1 \leq \ell \leq 4$ and let $P = (v_0, \ldots, v_\ell)$ be a sequence of $\ell + 1$ distinct vertices from V, where optionally $v_0 = v_\ell$. Suppose first that $v_0 \neq v_\ell$. Let $S_0 = V \setminus \{v_1, v_\ell\}$ and $S_\ell = V \setminus \{v_0, v_{\ell-1}\}$. Let $\mathcal{A}_P$ be the event that P is contained in G, and for $i = 0, \ell$ let $\mathcal{B}_i$ be the event that $d(v_i, S_i) \leq \alpha \log n$. By Theorem 3.5 we obtain that $\mathbb{P}(\mathcal{B}_i) \leq n^{-0.6}$. The events $\mathcal{A}_P, \mathcal{B}_0, \mathcal{B}_\ell$ are mutually independent, hence $\mathbb{P}(\mathcal{A}_P \wedge \mathcal{B}_0 \wedge \mathcal{B}_\ell) \leq p^\ell n^{-1.2}$. Let $\mathcal{A}$ be the event that there exists a path $P = v_0, \ldots, v_\ell$ with $\ell \in [4]$ in G such that $\mathcal{A}_P$ and $d(v_0), d(v_\ell) \leq \alpha \log n$. By the union bound, $\mathbb{P}(\mathcal{A}) \leq \sum_{\ell=1}^4 n^{\ell+1-1.2} p^\ell = o(1)$. The case $v_0 = v_\ell$ (which implies $\ell \in \{3, 4\}$) is similar. Let $S = V \setminus \{v_1, v_{\ell-1}\}$ and let $\mathcal{B}$ be the event $d(v_0, S) \leq \alpha \log n$. As before, $\mathbb{P}(\mathcal{B}) \leq n^{-0.6}$, and the events $\mathcal{A}_P, \mathcal{B}$ are independent, hence $\mathbb{P}(\mathcal{A}_p \wedge \mathcal{B}) \leq p^\ell n^{-0.6}$. Let $\mathcal{A}'$ be the event that there exists a cycle P of length $\ell \in \{3, 4\}$ such that $\mathcal{A}_P$ and $d(v_0) \leq \alpha \log n$. By the union bound, $\mathbb{P}(\mathcal{A}') \leq \sum_{\ell=3}^4 n^\ell p^\ell n^{-0.6} = o(1)$. □

Proof of (P3). For a given set $U \subseteq V(G)$ and for a given $k \geq 0$, the probability that $|E_G(U)| \geq k$ is at most

$$\binom{\binom{|U|}{2}}{k} \cdot p^k \leq \binom{|U|^2}{k} \cdot p^k \leq \left(\frac{e|U|^2 p}{k} \right)^k.$$

Hence, by the union bound, noting that $p \leq 2 \log n/n$, the probability that (P3) does not hold is at most

$$\sum_{t=1}^{\varepsilon n/100} \binom{n}{t} \cdot \left(\frac{et^2 p}{\varepsilon t \log n/10} \right)^{\varepsilon t \log n/10} \leq \sum_{t=1}^{\varepsilon n/100} \left(\frac{en}{t} \right)^t \cdot \left(\frac{60t}{\varepsilon n} \right)^{\varepsilon t \log n/10}$$

$$= \sum_{t=1}^{\varepsilon n/100} \left(\frac{60e}{\varepsilon} \cdot \left(\frac{60t}{\varepsilon n} \right)^{\varepsilon \log n/10 - 1} \right)^t$$

$$\leq \sum_{t=1}^{\varepsilon n/100} \left(\frac{60e}{\varepsilon} \cdot 0.6^{\Omega(\varepsilon \log n)} \right)^t$$

$$= \sum_{t=1}^{\varepsilon n/100} o(1)^t = o(1). \quad \square$$

Proof of (P4). The proof involves an application of the symmetric form of the Local Lemma (see, e.g., [2, Chapter 5]; a similar application appears in [23]). Write $V = V(G)$ and let $X = \{v \in V : d(v) \leq \log n/10\}$. We start by observing that X is typically small. Indeed, by Theorem 3.5 we have $\mathbb{P}(d(v) \leq \log n/10) \leq n^{-0.6}$, and by Markov's inequality $|X| \leq n^{0.5}$ **whp**. By the definition of X we have that $X^+ := X \cup N(X)$ satisfies $|X^+| \leq |X| \cdot \log n/10 \leq n^{0.6}$ **whp**.

From now on we fix G, assuming that $|X^+| \leq n^{0.6}$ and that G satisfies (P1) and (P2); these events happen **whp**. Let $\varepsilon' = 1/(\lceil 1/\varepsilon \rceil + 1)$ and note that for $\varepsilon < 1/2$ we have $\varepsilon/2 \leq \varepsilon' < \varepsilon$. Write $s = 1/\varepsilon'$, let $t = \lfloor n/s \rfloor \sim \varepsilon' n$, and let $A_1, \ldots, A_t, Z$ be a partitioning of the vertices of G into t "blobs" A_i of size s and an extra set Z with $|Z| \leq s$. For $j \in [t]$, let (x_j^1, x_j^2) be a uniformly chosen (ordered) pair of distinct vertices from A_j. For $i = 1, 2$ define $U_i' = \{x_j^i\}_{j=1}^t$. Clearly, $|U_1'| = |U_2'| = t$ and $U_1' \cap U_2' = \varnothing$. For every $v \in V \setminus X$, let $\mathcal{B}_v$ be the event that $d(v, U_i') < \varepsilon' \log n / 40$ for some $i = 1, 2$. For such v, let $L(v)$ be the set of blobs that contain neighbours of v, namely, $L(v) = \{A_i : N(v) \cap A_i \neq \varnothing\}$. For $j \in [t]$ write $n_j(v) = |N(v) \cap A_j|$, and note that $\sum_j n_j(v) \geq d(v) - s \geq \log n / 10 - s \geq \log n / 20$ (for n large enough). For $i = 1, 2$ and $j \in [t]$, let $\chi_j^i(v)$ be the indicator of the event that x_j^i is a neighbour of v, and note that $\mathbb{E}\chi_j^i(v) = \varepsilon' n_j(v)$. Observe that for $i = 1, 2$, $d(v, U_i') = \sum_j \chi_j^i(v)$, hence $\mathbb{E}[d(v, U_i')] = \varepsilon' \sum_j n_j(v) \geq \varepsilon' \log n / 20$. Thus, by Theorem 3.5, $\mathbb{P}(\mathcal{B}_v) \leq n^{-c}$ for some $c = c(\varepsilon) > 0$.

For two distinct vertices $u, v \in V \setminus X$ say that u, v are **related** if $L(u) \cap L(v) \neq \varnothing$. For a vertex $u \in V \setminus X$, let $R(u)$ be the set of vertices in $V \setminus X$ which are related to u, and note that $|R(u)| \leq s\Delta(G)^2$, which is, by (P1), at most $C \log^2 n$ for some $C = C(\varepsilon) > 0$. Note that $\mathcal{B}_u$ is mutually independent of the set of events $\{\mathcal{B}_v \mid v \in (V \setminus X) \setminus R(u)\}$. We now apply the symmetric form of the Local Lemma[1]: observing that $en^{-c} \cdot C \log^2 n < 1$ (for large enough n), we get that with positive probability, none of the events $(\mathcal{B}_v : v \in V \setminus X)$ occur, meaning that $d(v, U_i') \geq \varepsilon' \log n / 40 \geq \varepsilon \log n / 80$ for every $v \in V \setminus X$ and $i = 1, 2$. We choose U_1', U_2' to satisfy this. Now define $U_1 = U_1' \cup X^+$ and $U_2 = U_2' \setminus X^+$, and note that from the discussion above, $|U_1|, |U_2| \sim \varepsilon' n \leq \varepsilon n$. Let $v \in V \setminus X$. The fact that G satisfies (P2) implies that v has at most 1 neighbour in X^+. Thus, for every $v \in V \setminus X$ it holds that $d(v, U_1) \geq d(v, U_1') \geq \varepsilon \log n / 100$ and $d(v, U_2) \geq d(v, U_2') - 1 \geq \varepsilon \log n / 100$, as required. $\qquad \square$

In the proof of Lemma 4.1, we will argue that **whp** $G \sim G(n, p)$ is such that every subset $W \subseteq V(G)$ possessing certain properties induces a Hamiltonian graph. To this end, we will use the fact that given such a set W and a relatively sparse expander H on W which is a subgraph of G, it is highly likely that there is an edge e of G which is a booster with respect to H. This fact is established in Lemma 4.5 below. In the proof of Lemma 4.5 we will use the well-known and easy-to-show fact that if a graph H is a $(|V(H)|/4, 2)$-expander then H is connected. Indeed, if (by contradiction) H is not connected, then take a connected component X of size at most $|V(H)|/2$ and a set $U \subseteq X$ of size $\min\{|V(H)|/4, |X|\}$, and observe that $|N(U)| \leq |X| - |U| < 2|U|$, contradicting the assumption that H is a $(|V(H)|/4, 2)$-expander.

Lemma 4.5. *Let $c > 0$ be a sufficiently small absolute constant ($c = 10^{-5}$ suffices), let $p = (\log n + \log \log n + \omega(1))/n$ and let $G \sim G(n, p)$. Then, **whp**, G satisfies the following: for every $W \subseteq V(G)$ of size $|W| \geq 0.1n$ and for every $(|W|/4, 2)$-expander H on W which is a subgraph of G and has at most $cn \log n$ edges, if H is not Hamiltonian then G contains a booster with respect to H.*

Proof. We use a first moment argument. Evidently, the number of choices for the set W is at most 2^n. Let us fix a choice of W. For each t, the number of choices of H for which $|E(H)| = t$

---

[1] Note that in expectation there are $n^{\Omega(1)}$ vertices $v \in V \setminus X$ for which the event $\mathcal{B}_v$ occurs. Hence, it is not true that **whp** every vertex $v \in V \setminus X$ has high degree to both U_1' and U_2'. One can then try to fix the situation for the (relatively few) "unsatisfied" vertices by moving elements into U_1' and U_2' and between these sets. However, moving elements between U_1' and U_2' — which might be necessary if for example some $v \in V \setminus X$ has all of its neighbours in U_1' (and hence none in U_2') — can then affect the situation of other vertices. Seeing as the simple union-bound/alterations arguments do not work, we employ the Local Lemma.

is at most

$$\binom{\binom{|W|}{2}}{t} \leq \binom{n^2}{t} \leq \left(\frac{en^2}{t}\right)^t.$$

Now let H be a non-Hamiltonian $(|W|/4, 2)$-expander on W, and set $t := |E(H)|$. As mentioned above, H is connected. By Lemma 4.3, H has at least $(|W|/4)^2/2 = |W|^2/32 \geq n^2/3200$ boosters. Now, the probability that G contains H but no booster thereof is at most

$$p^t \cdot (1-p)^{n^2/3200} \leq p^t \cdot \left(1 - \frac{\log n}{n}\right)^{n^2/3200} \leq \left(\frac{2\log n}{n}\right)^t \cdot \exp(-n\log n/3200).$$

Summing over all choices of W and H, we see that the probability that the assertion of the lemma does not hold is at most

$$2^n \cdot \exp(-n\log n/3200) \cdot \sum_{t=1}^{cn\log n} \left(\frac{2en\log n}{t}\right)^t. \tag{6}$$

Setting $g(t) := (2en\log n/t)^t$, we note that $g'(t) = g(t) \cdot (\log(2en\log n/t) - 1) > 0$ for every t in the range of the sum in (6), assuming $c < 1$, say. Thus, this sum is not larger than

$$cn\log n \cdot (2e/c)^{cn\log n} = \exp((\log(2e/c) \cdot c + o(1))n\log n).$$

Now, if c is small enough so that $\log(2e/c) \cdot c < 1/3200$, we get that (6) tends to 0 as n tends to infinity. This completes the proof. $\qquad\square$

The following lemma states that a graph possessing certain simple properties is necessarily an expander. Statements of this type are fairly common in the study of Hamiltonicity of random graphs (see, e.g., [28]). For completeness, we include a proof.

Lemma 4.6. *Let $m, d \geq 1$ be integers and let H be a graph on $h \geq 4m$ vertices satisfying the following properties:*

1. *$\delta(H) \geq 2$;*

2. *No vertex $v \in V(H)$ with $d(v) < d$ is contained in a 3- or a 4-cycle, and every two distinct vertices $u, v \in V(H)$ with $d(u), d(v) < d$ are at distance at least 5 apart;*

3. *Every set $U \subseteq V(H)$ of size at most $5m$ contains at most $d|U|/10$ edges;*

4. *There is an edge between every pair of disjoint sets $U_1, U_2 \subseteq V(H)$ of size m each.*

Then H is an $(h/4, 2)$-expander.

Proof. Our goal is to show that for every $U \subseteq V(H)$ with $|U| \leq h/4$ it holds that $|N(U)| \geq 2|U|$. So let $U \subseteq V(H)$ be such that $|U| \leq h/4$. Suppose first that $|U| \geq m$. Since there evidently is no edge between U and $V(H) \setminus (U \cup N(U))$, it must be the case that $|V(H) \setminus (U \cup N(U)| < m$ by Item 4. So we have $|U \cup N(U)| > h - m$ and hence $|N(U)| > h - m - |U| \geq \frac{3}{4}h - m \geq \frac{1}{2}h \geq 2|U|$, as required. Here we used the assumption that $h \geq 4m$ as well as the fact that $|U| \leq h/4$.

Suppose now that $|U| \leq m$. Let X be the set of all $u \in U$ satisfying $d(u) < d$, and set $Y := U \setminus X$. We claim that $|N(Y)| \geq 4|Y|$. Suppose, for the sake of contradiction, that $|N(Y)| < 4|Y|$. Then $|Y \cup N(Y)| < 5|Y| \leq 5|U| \leq 5m$. On the other hand, the definition of Y implies that H has at least $d|Y|/2$ edges incident to vertices of Y. Since all of these edges

are contained in $Y \cup N(Y)$, we see that $Y \cup N(Y)$ contains at least $d|Y|/2 > d \cdot |Y \cup N(Y)|/10$ edges. But this stands in contradiction with Item 3. Thus, $|N(Y)| \geq 4|Y|$.

Next, note that by Item 2, every two elements of X are at distance at least 5; in particular, X is an independent set, and every two elements of X have disjoint neighbourhoods. Now Item 1 implies that $|N(X)| \geq 2|X|$.

Observe that each vertex of Y has at most one neighbour in $X \cup N(X)$, for otherwise there would be a 4-cycle containing an element of X or a pair of elements of X at distance at most 4, both of which are impossible due to Item 2. So we conclude that $|N(Y) \cap (X \cup N(X))| \leq |Y|$, and hence $|N(Y) \setminus (X \cup N(X))| \geq |N(Y)| - |Y|$. All in all, we get that

$$|N(U)| = |N(X) \setminus Y| + |N(Y) \setminus (X \cup N(X))| \geq |N(X)| - |Y| + |N(Y) \setminus (X \cup N(X))|$$
$$\geq 2|X| - |Y| + |N(Y)| - |Y| \geq 2|X| + 2|Y| = 2|U|,$$

as required. $\qquad\square$

The following lemma constitutes the main part of the proof of Lemma 4.1.

Lemma 4.7. *Let $\varepsilon > 0$. For $p = (\log n + \log\log n + \omega(1))/n$, the random graph $G \sim G(n,p)$ satisfies the following **whp**. Let $W \subseteq V(G)$ be such that $|W| \geq 0.1n$, and for every $v \in W$ it holds that $d(v, W) \geq \min\{d(v), \varepsilon \log n\}$. Then for every $w \in W$ there exists $Y \subseteq W$ with $|Y| \geq n/40$ such that for each $y \in Y$, there is a Hamilton path in $G[W]$ whose endpoints are w and y.*

Proof. We may and will assume ε is sufficiently small (it is enough to have $\varepsilon \leq \min\{1/10, c/2\}$, where c is the constant from Lemma 4.5). We will assume that the events defined in Lemma 3.4, Lemma 4.4 and Lemma 4.5 hold (this happens **whp**), and show that in this case, the assertion of Lemma 4.7 holds as well.

It will be convenient to set $d_0 := \varepsilon \log n$. Let $W \subseteq V(G)$ be as in the statement of Lemma 4.7. We select a random spanning subgraph H of $G[W]$ as follows. For each $v \in W$, if $d(v, W) < d_0$ then add to H all edges of $G[W]$ incident to v. Otherwise, namely if $d(v, W) \geq d_0$, then randomly select a set of d_0 edges of $G[W]$ incident to v and add these to H. Note that $|E(H)| \leq |W| \cdot d_0 \leq \varepsilon n \log n$. On the other hand, our assumption that $d(v, W) \geq \min\{d(v), \varepsilon \log n\}$ for every $v \in W$ implies that $\delta(H) \geq \min\{\delta(G), d_0\}$. Hence, as $d_0 \geq 2$ (for large enough n), we have $\delta(H) \geq 2$ by Property (P1) of Lemma 4.4.

We claim that with positive probability (in fact, **whp**), H is a $(|W|/4, 2)$-expander. In light of Lemma 4.6, it is sufficient to show that with positive probability, H satisfies Conditions 1–4 in that lemma. Here, we will choose the parameters of Lemma 4.6 as $d := d_0$ and $m := \varepsilon n/500$. We already showed that $\delta(H) \geq 2$ (which is Condition 1 in Lemma 4.6). Condition 2 holds because H is a subgraph of G and because the analogous statement holds for G, as guaranteed by Property (P2) in Lemma 4.4 (here we assume that $\varepsilon \leq 1/10$). Similarly, Condition 3 holds because H is a subgraph of G and due to Property (P3) in Lemma 4.4 (note that $5m = \varepsilon n/100$).

Let us now prove that Condition 4 holds. Let $U_1, U_2 \subseteq V(H) = W$ be disjoint sets satisfying $|U_1|, |U_2| = m = \varepsilon n/500$. Since G is (γ, p)-pseudorandom with $\gamma = \varepsilon/500$ (in fact, with $\gamma = o(1)$, see Lemma 3.4), we have

$$|E_G(U_1, U_2)| \geq (1 - \gamma)p \cdot |U_1||U_2| \geq \frac{|U_1||U_2| \log n}{2n} \geq \frac{\varepsilon^2 n \log n}{500000} = \Omega(n \log n). \qquad (7)$$

Now, let us bound (from above) the probability that $|E_H(U_1, U_2)| = 0$ (where the randomness is with respect to the choice of H). Recall that H is defined by choosing, for each $v \in W$, a random

set $E(v)$ of $\min\{d(v,W),d_0\}$ edges of $G[W]$ incident to v, with all choices made uniformly and independently, and letting $E(H) = \bigcup_{v \in W} E(v)$. Fix any $u_1 \in U_1$ with $d(u_1,U_2) \geq 1$, and let $\mathcal{A}_{u_1}$ be the event that there is no edge in $E(u_1)$ with an endpoint in U_2. Observe that if $d(u_1,W) < d_0$ then $\mathbb{P}(\mathcal{A}_{u_1}) = 0$, and otherwise

$$\mathbb{P}(\mathcal{A}_{u_1}) = \binom{d(u_1,W) - d(u_1,U_2)}{d_0} \Big/ \binom{d(u_1,W)}{d_0} = \prod_{i=0}^{d_0-1} \frac{d(u_1,W) - d(u_1,U_2) - i}{d(u_1,W) - i}$$

$$\leq \left(1 - \frac{d(u_1,U_2)}{d(u_1,W)}\right)^{d_0} \leq \left(1 - \frac{d(u_1,U_2)}{\Delta(G)}\right)^{d_0} \leq e^{-d(u_1,U_2) \cdot \frac{d_0}{\Delta(G)}} \leq e^{-\varepsilon d(u_1,U_2)/10}.$$

Here, in the last inequality we used Property (P1) in Lemma 4.4. Note that the events $(\mathcal{A}_{u_1} : u_1 \in U_1)$ are independent, and that if $E_H(U_1,U_2) = \varnothing$ then $\mathcal{A}_{u_1}$ occurred for every $u_1 \in U_1$ with $d(u_1,U_2) \geq 1$. It now follows that

$$\mathbb{P}(E_H(U_1,U_2) = \varnothing) \leq \exp\left(-\frac{\varepsilon}{10} \cdot \sum_{u_1 \in U_1} d(u_1,U_2)\right) = \exp\left(-\frac{\varepsilon}{10} \cdot |E_G(U_1,U_2)|\right) \leq e^{-\Omega(n \log n)},$$

where in the last inequality we used (7). By taking the union bound over all at most 2^{2n} choices of U_1, U_2, we see that with high probability, $E_H(U_1,U_2) \neq \varnothing$ for every pair of disjoint sets $U_1, U_2 \subseteq W$ of size m each.

Finally, we apply Lemma 4.6 to conclude that **whp** H is a $(|W|/4, 2)$-expander. From now on, we fix such a choice of H. Before establishing the assertion of the lemma, we first show that $G[W]$ is Hamiltonian. To find a Hamilton cycle in $G[W]$, we define a sequence of graphs H_i, $i \geq 0$, as follows. To begin, set $H_0 = H$. For each $i \geq 0$, if H_i is Hamiltonian then stop, and otherwise take a booster of H_i contained in $G[W]$ and add it to H_i to obtain H_{i+1}. That such a booster exists is guaranteed by Lemma 4.5, as we will always have $|E(H_i)| \leq |E(H)| + |W| \leq |E(H)| + n \leq \varepsilon n \log n + n \leq c/2 \cdot n \log n + n \leq c n \log n$, provided that ε is smaller than $c/2$, where c is the constant appearing in Lemma 4.5. Note also that H_i is a subgraph of $G[W]$ for each $i \geq 0$. Evidently, this process has to stop (because as long as H_i is not Hamiltonian, the maximum length of a path in H_i is longer than in H_{i-1}), thus showing that $G[W]$ must contain a Hamilton cycle, as claimed.

Now let $w \in W$. As $G[W]$ is Hamiltonian, there exists a Hamilton path P of $G[W]$ such that w is one of the endpoints of P. Evidently, P is a longest path in $G[W]$. Furthermore, note that $G[W]$ is a $(|W|/4, 2)$-expander because H, a subgraph of $G[W]$, is such an expander. By Lemma 4.2 (applied to $G[W]$), there exists a set $R \subseteq V(P) = W$ such that for every $y \in R$ there is a Hamilton path in $G[W]$ with endpoints w and y, and such that $|N_{G[W]}(R)| \leq 2|R| - 1$. Now, since $G[W]$ is a $(|W|/4, 2)$-expander, it must be the case that $|R| > |W|/4 \geq n/40$. So we see that the assertion of the lemma holds with $Y = R$. This completes the proof. $\square$

Proof of Lemma 4.1. For convenience we show the existence of a partition $V(G) = V^\star \cup V'$ with $|V^\star| \leq 2\varepsilon n$ instead of $|V^\star| \leq \varepsilon n$ (this clearly does not matter). We assume that G satisfies the properties detailed in Lemma 4.4, and that it is a (γ, p)-pseudorandom for $\gamma < 1/40$ and some $p \in (0,1)$, as guaranteed to happen **whp** by Lemma 3.4. Let U_1, U_2 be disjoint subsets of $V = V(G)$ satisfying (P4). Set $V^\star = U_1 \cup U_2$ and $V' = V \setminus V^\star$, and let $P \subseteq V'$ be a path with $|V(P)| \leq 2n/3$ and endpoints a_1, a_2. In particular, $|V^\star| \leq 2\varepsilon n$. Our goal is to extend P to a Hamilton cycle of G. Write $V'' = V' \setminus V(P)$, partition $V'' = V_1'' \cup V_2''$ as equally as
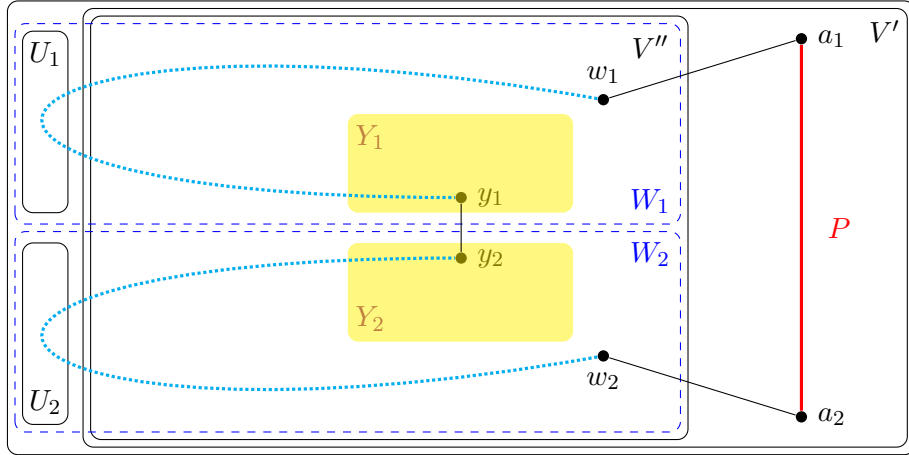
Figure 2: Outline of the proof of Lemma 4.1.

possible. For $i = 1, 2$, let $W_i = V_i'' \cup U_i$ and choose a neighbour w_i of a_i in W_i; this is possible since $d(a_i, U_i) \geq \varepsilon \log n/100$ by (P4). Note that $|W_i| \geq n/6$ and for every $v \in W_i$ it holds that $d(v, W_i) \geq \min\{d(v), \varepsilon \log n/100\}$, hence by Lemma 4.7 there exists a set $Y_i \subseteq W_i$ with $|Y_i| \geq n/40$ such that for every $y \in Y_i$ there is a Hamilton path spanning W_i from w_i to y. Since G is a (γ, p)-pseudorandom for $\gamma < 1/40$, it has an edge e between Y_1 and Y_2 with endpoints $y_i \in Y_i$, say. For $i = 1, 2$, denote by Q_{y_i} the Hamilton path between w_i and y_i. We now construct a Hamilton cycle of G as follows (as depicted in Fig. 2):

$$a_1 \to w_1 \xrightarrow{Q_{y_1}} y_1 \xrightarrow{e} y_2 \xrightarrow{Q_{y_2}} w_2 \to a_2 \xrightarrow{P} a_1. \qquad \square$$

We now put together Theorem 3.1 and Lemma 4.1 in order to prove Theorem 1.1.

Proof of Theorem 1.1. Let $r \geq 2$, $\varepsilon > 0$ and $p = (\log n + \log \log n + \omega(1))/n$, let $G \sim G(n, p)$ and consider an r-colouring of the edge set of G. Let γ be the constant obtained from Theorem 3.1 by plugging in r and ε. Let $V^\star \cup V'$ be the partition guaranteed **whp** by Lemma 4.1 which satisfies $n' = |V'| \geq (1 - \varepsilon)n$. By Lemma 3.4 we know that G is $(\gamma(1 - \varepsilon), p)$-pseudorandom (**whp**), hence $G' = G[V']$ is (γ, p)-pseudorandom. By Theorem 3.1 we know that there exists a path P in G' of length at most $2n'/(r + 1) \leq 2n/3$ having at least $(2/(r + 1) - \varepsilon)n' \geq (2/(r + 1) - 2\varepsilon)n$ edges of the same colour. By Lemma 4.1 we can, **whp**, extend P into a Hamilton cycle of G, still having at least $(2/(r + 1) - 2\varepsilon)n$ edges of the same colour. $\qquad \square$

5   Perfect matchings

We now sketch a proof of Theorem 1.2. The first observation is that with mild modifications of the proof of Lemma 4.1 we may prove a variant of the following form. Let $\varepsilon > 0$ and $p = (\log n + \omega(1))/n$. Then $G \sim G(n, p)$ **whp** admits a partition of its vertex set $V(G) = V^\star \cup V'$ with $|V^\star| \leq \varepsilon n$ such that **(a)** the set D_1 of vertices of degree 1 in G and its neighbourhood $N(D_1)$ are contained in V'; *and* **(b)** for every subset X of V' with $|X| \leq 2n/3$ and $D_1 \subseteq X$, the subgraph $G[V^\star \cup (V \setminus X)]$ contains a Hamilton path. We omit the proof details.

17

Having that lemma in hand, we proceed as follows. Let M_0 be the set of edges incident to vertices of D_1; there are, **whp**, $O(\log n)$ such edges, and they form, **whp**, a matching. As $G[V' \setminus V(M_0)]$ is (**whp**) (γ, p)-pseudorandom by Lemma 3.4, we know by Theorem 3.1 that it has an almost monochromatic path P of length $(2/(r+1) - \varepsilon)n$, from which we can extract a monochromatic matching of size at least $(1/(r+1) - \varepsilon')n$, for some $\varepsilon' > 0$. Add it to M_0, creating an almost monochromatic matching M_1 of size at least $(1/(r+1) - \varepsilon')n$. We now apply the lemma to find a Hamilton path in $G[V^\star \cup (V' \setminus V(M_1))]$, from which we extract a matching which completes M_1 into a perfect matching, in which at least $(1/(r+1) - \varepsilon')n$ edges are of the same colour. $\qquad\square$

References

[1] Miklós Ajtai, Janós Komlós and Endre Szemerédi, *First occurrence of Hamilton cycles in random graphs*, Cycles in graphs (Burnaby, B.C., 1982), 1985, pp. 173–178. MR821516 ↑1

[2] Noga Alon and Joel H. Spencer, **The probabilistic method**, Fourth Edition, Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016. MR3524748 ↑12

[3] Yahav Alon and Michael Krivelevich, *Random graph's hamiltonicity is strongly tied to its minimum degree*, Electronic Journal of Combinatorics **27** (2020), no. 1, Paper No. 1.30. ↑1

[4] Michael Anastos and Alan Frieze, *Pattern colored Hamilton cycles in random graphs*, SIAM Journal on Discrete Mathematics **33** (2019), no. 1, 528–545. MR3924610 ↑1

[5] József Balogh, Béla Csaba, Yifan Jing and András Pluhár, *On the discrepancies of graphs*, Electronic Journal of Combinatorics **27** (2020), no. 2, Paper No. 2.12. ↑2

[6] József Balogh, Béla Csaba, András Pluhár and Andrew Treglown, *A discrepancy version of the Hajnal–Szemerédi theorem*, Combinatorics, Probability and Computing **30** (2021), no. 3, 444–459. ↑2

[7] Ido Ben-Eliezer, Michael Krivelevich and Benny Sudakov, *The size Ramsey number of a directed path*, Journal of Combinatorial Theory. Series B **102** (2012), no. 3, 743–755. MR2900815 ↑9

[8] Béla Bollobás, *The evolution of random graphs*, Transactions of the American Mathematical Society **286** (1984), no. 1, 257–274. MR756039 ↑1

[9] Ernest J. Cockayne and Peter J. Lorimer, *The Ramsey number for stripes*, Journal of the Australian Mathematical Society. Series A **19** (1975), 252–256. MR0371733 ↑2, 3, 8

[10] David Conlon, *Combinatorial theorems relative to a random set*, Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV, 2014, pp. 303–327. MR3727614 ↑8

[11] Andrzej Dudek and Paweł Prałat, *On some multicolor Ramsey properties of random graphs*, SIAM Journal on Discrete Mathematics **31** (2017), no. 3, 2079–2092. MR3697158 ↑4

[12] Paul Erdős, Zoltán Füredi, Martin Loebl and Vera T. Sós, *Discrepancy of trees*, Studia Scientiarum Mathematicarum Hungarica **30** (1995), no. 1-2, 47–57. MR1341566 ↑2

[13] Paul Erdős and Alfréd Rényi, *On the evolution of random graphs*, A Magyar Tudományos Akadémia. Matematikai Kutató Intézetének Közleményei **5** (1960), 17–61. MR125031 ↑1

[14] Paul Erdős and Joel H. Spencer, *Imbalances in k-colorations*, Networks **1** (1971/72), 379–385. MR299525 ↑2

[15] Lisa Espig, Alan Frieze and Michael Krivelevich, *Elegantly colored paths and cycles in edge colored random graphs*, SIAM Journal on Discrete Mathematics **32** (2018), no. 3, 1585–1618. MR3825612 ↑1

[16] Agnieszka Figaj and Tomasz Łuczak, *The Ramsey number for a triple of long even cycles*, Journal of Combinatorial Theory. Series B **97** (2007), no. 4, 584–596. MR2325798 ↑4

[17] Alan Frieze, *Hamilton cycles in random graphs: a bibliography*, arXiv e-prints (January 2019), available at arXiv:1901.07139. ↑1

[18] Alan Frieze and Michał Karoński, **Introduction to random graphs**, Cambridge University Press, Cambridge, 2016. MR3675279 ↑12

[19] László Gerencsér and András Gyárfás, *On Ramsey-type problems*, Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Mathematica **10** (1967), 167–170. MR239997 ↑4

[20] Roman Glebov and Michael Krivelevich, *On the number of Hamilton cycles in sparse random graphs*, SIAM Journal on Discrete Mathematics **27** (2013), no. 1, 27–42. MR3032903 ↑1

[21] András Gyárfás, Miklós Ruszinkó, Gábor N. Sárközy and Endre Szemerédi, *Three-color Ramsey numbers for paths*, Combinatorica **27** (2007), no. 1, 35–69. MR2310787 ↑4

[22] András Gyárfás, Miklós Ruszinkó, Gábor N. Sárközy and Endre Szemerédi, *Corrigendum: "Three-color Ramsey numbers for paths" [Combinatorica 27 (2007), no. 1, 35–69; MR2310787]*, Combinatorica **28** (2008), no. 4, 499–502. MR2452848 ↑4

[23] Dan Hefetz, Michael Krivelevich and Tibor Szabó, *Sharp threshold for the appearance of certain spanning trees in random graphs*, Random Structures & Algorithms **41** (2012), no. 4, 391–412. MR2993127 ↑12

[24] Svante Janson, Tomasz Łuczak and Andrzej Rucinski, **Random graphs**, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR1782847 ↑10

[25] Charlotte Knierim and Pascal Su, *Improved bounds on the multicolor Ramsey numbers of paths and even cycles*, Electronic Journal of Combinatorics **26** (2019), no. 1, Paper No. 1.26, 17. MR3919617 ↑4

[26] Yoshiharu Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), 1997, pp. 216–230. MR1661982 ↑8

[27] János Komlós and Endre Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Mathematics **43** (1983), no. 1, 55–63. MR680304 ↑1

[28] Michael Krivelevich, *Long paths and hamiltonicity in random graphs*, Random graphs, geometry and asymptotic structure, 2016, pp. 4–27. ↑1, 9, 11, 14

[29] Choongbum Lee and Benny Sudakov, *Dirac's theorem for random graphs*, Random Structures & Algorithms **41** (2012), no. 3, 293–305. MR2967175 ↑1

[30] Shoham Letzter, *Path Ramsey number for random graphs*, Combinatorics, Probability and Computing **25** (2016), no. 4, 612–622. MR3506430 ↑4, 9

[31] László Lovász and Michael D. Plummer, **Matching theory**, AMS Chelsea Publishing, Providence, RI, 2009. Corrected reprint of the 1986 original. MR2536865 ↑5

[32] Tomasz Łuczak, $R(C_n, C_n, C_n) \leq (4 + o(1))n$, Journal of Combinatorial Theory. Series B **75** (1999), no. 2, 174–187. MR1676887 ↑4

[33] Richard Montgomery, *Hamiltonicity in random graphs is born resilient*, Journal of Combinatorial Theory. Series B **139** (2019), 316–341. MR4010194 ↑1

[34] Rajko Nenadov, Angelika Steger and Miloš Trujić, *Resilience of perfect matchings and Hamiltonicity in random graph processes*, Random Structures & Algorithms **54** (2019), no. 4, 797–819. MR3957368 ↑1

[35] Fedor Petrov, *A direct proof that every r-colored complete graph on $n = (r + 1)m - (r - 1)$ vertices has a monochromatic matching of size m? (answer)* (2018), https://mathoverflow.net/a/299813/46253. ↑8

[36] Lajos Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), no. 4, 359–364. MR389666 ↑1, 4, 11

[37] Alex Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combinatorics, Probability and Computing **20** (2011), no. 3, 455–466. MR2784637 ↑8, 9

[38] Benny Sudakov, *Robustness of graph properties*, Surveys in combinatorics 2017, pp. 372–408. MR3728112 ↑1

[39] Chuandong Xu, Hongna Yang and Shenggui Zhang, *A new proof on the Ramsey number of matchings*, arXiv e-prints (May 2019), available at arXiv:1905.08456. ↑3

LIOR GISHBOLINER

DEPARTMENT OF MATHEMATICS, ETH, ZÜRICH, SWITZERLAND

Email: lior.gishboliner@math.ethz.ch

Michael Krivelevich

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

Email: krivelev@tauex.tau.ac.il

Research supported in part by USA-Israel BSF grant 2018267 and by ISF grant 1261/17.

Peleg Michaeli

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

Email: peleg.michaeli@math.tau.ac.il

Research supported by ERC starting grant 676970 RANDGEOM and by ISF grants 1207/15 and 1294/19.