Hamilton completion and the path cover number of sparse random graphs

Yahav Alon [*]   Michael Krivelevich [†]

January 28, 2024

Abstract

We prove that for every $\varepsilon > 0$ there is c_0 such that if $G \sim G(n, c/n)$, $c \geq c_0$, then with high probability G can be covered by at most $(1 + \varepsilon) \cdot \frac{1}{2} ce^{-c} \cdot n$ vertex disjoint paths, which is essentially tight. This is equivalent to showing that, with high probability, at most $(1 + \varepsilon) \cdot \frac{1}{2} ce^{-c} \cdot n$ edges can be added to G to create a Hamiltonian graph.

1 Introduction

Consider the binomial random graph model $G(n, p)$, in which every edge of the complete graph K_n is retained with probability p, independently of all other edges. A *Hamilton cycle* in a graph G is a simple cycle in G that covers the entire vertex set of G, and a *Hamiltonian graph* is a graph that contains a Hamilton cycle. A classical result by Komlós and Szemerédi [7], and independently by Bollobás [3], states that if $p = p(n) = (\log n + \log \log n + f(n))/n$ then

$$\lim_{n \to \infty} \mathbb{P}(G(n, p) \text{ is Hamiltonian}) = \begin{cases} 1 & \text{if } f(n) \to \infty; \\ 0 & \text{if } f(n) \to -\infty. \end{cases}$$

In light of this, one cannot expect a Hamilton cycle to exist in a random graph $G(n, p)$ when p is below the sharp threshold stated above. This invites a quantitative question: typically, how close is a random graph below the threshold to being Hamiltonian? To answer this question a measure of "distance" from Hamiltonicity should first be defined.

One example of such a measure is the maximum length of a cycle in the graph $L_{\max}(G)$, where if p is above the Hamiltonicity threshold then typically $L_{\max}(G(n, p)) = n$. Frieze [4] showed that if $c > 0$ is a large enough constant with respect to ε then typically

$$L_{\max}(G(n, c/n)) \geq (1 - (1 + \varepsilon)ce^{-c})n.$$

Up to the value of ε, this is a tight bound. This is due to the fact that, with high probability, there are $(c + 1 + o(1)) \cdot e^{-c} n$ vertices of degree less than 2 in $G(n, c/n)$, and such vertices cannot be a part of

---

[*]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel. Email: yahavalo@tauex.tau.ac.il.

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

any cycle. Following that, Anastos and Frieze [2] extended this result by showing that, in fact, there is a function f such that $L_{\max}(G(n, c/n))/n \xrightarrow{a.s.} f(c)$. The function $f(c)$ can be expressed as a series, with explicitly computable terms, the first few of which are computed in the paper.

In a recent paper Nenadov, Sudakov and Wagner [9] presented a general measure for the distance of a graph from a given property: deficiency. The *deficiency* of a graph G with respect to a graph property $\mathcal{P}$, denoted $\mathrm{def}(G, \mathcal{P})$, is equal to the smallest integer t such that $G * K_t$ has the property $\mathcal{P}$. Here, $G * H$ is the *join* of H and G, that is, the graph with vertex set $V(G) \uplus V(H)$ and edge set $E(G) \cup E(H) \cup (G \times H)$.

With respect to the property of Hamiltonicity, the notion of deficiency is equivalent to other natural measure of distance from Hamiltonicity: it is also equal to the distance in edges of a graph from the nearest Hamiltonian graph. In other words, the deficiency of G with respect to Hamiltonicity is equal to the smallest integer t such that there is a graph H with $e(H) = t$ and $G \cup H$ is Hamiltonian. Evidently, this is also equal to the *path cover number* of G, defined as follows.

Definition 1. *Let G be a graph. The* path cover number *of G, denoted by $\mu(G)$, is 0 if G is Hamiltonian, and otherwise it is the minimal number $k \in \mathbb{N}$ such that there is a covering of $V(G)$ by k vertex disjoint paths.*

Henceforth we will denote $\mathrm{def}(G, \mathcal{H})$, where $\mathcal{H}$ is the property of Hamiltonicity, as $\mu(G)$, as the two quantities are equal.

In the above mentioned paper Nenadov et al. presented a (tight) combinatorial result by presenting an inequality involving $|V(G)|, |E(G)|, \mu(G)$ and showing that it holds for every graph G, as well as finding examples of G of every size for which the inequality is an equality.

When random graphs are considered, an immediate bound on $\mu(G)$, similar to the bound on the maximum cycle length, can be derived from the fact that every path in a disjoint path covering of G contains at most two vertices of degree 1 or a single vertex of degree 0. From this we get that, with high probability for $G \sim G(n, c/n)$, one has $\mu(G) \geq (\frac{1}{2}c + 1 + o(1)) \cdot e^{-c}n$. Like in the case of the maximum cycle length, this trivial bound turns out to be tight, as shown in the following theorem, which is our result of this paper.

Theorem 1. *For every $\varepsilon > 0$ there is a positive $c_0 = c_0(\varepsilon)$ such that, if $G \sim G(n, c/n)$, $c \geq c_0$, then with high probability*

$$(1 - \varepsilon) \cdot \frac{1}{2}ce^{-c} \cdot n \leq \mu(G) \leq (1 + \varepsilon) \cdot \frac{1}{2}ce^{-c} \cdot n.$$

In Section 2 we introduce notation and definitions to be used later in the paper, as well as some auxiliary results required for our proof. In Section 3 we present our proof for Theorem 1.

2 Preliminaries

Throughout the paper we use the following graph theoretic notation. For a graph $G = (V, E)$ and two disjoint vertex subsets $U, W \subseteq V$, $E_G(U, W)$ denotes the set of edges of G adjacent to exactly one vertex from U and one vertex from V, and $e_G(U, W) = |E_G(U, W)|$. Similarly, $E_G(U)$ denotes the set of edges spanned by a subset U of V, and $e_G(U)$ its size. The (external) neighbourhood of a vertex subset U, denoted by $N_G(U)$, is the set of vertices in $V \setminus U$ adjacent to a vertex of U. The degree of a vertex $v \in V$,

denoted by $d_G(v)$, is the number of edges of G incident to v, and $d_G(v, U)$ is the number of vertices in U adjacent to v. $\Delta(G)$ is the maximum degree of a vertex in G, that is, $\Delta(G) = \max_{v \in V} d_G(v)$.

For $u, v \in G$ we let $\text{dist}_G(u, v)$ denote the distance in G between u and v, that is the length of a shortest path in G connecting them (or $\text{dist}_G(u, v) = \infty$ if u and v are not connected by any path), where $\text{dist}_G(v, v)$ is defined to be the minimal length of a cycle containing v (or $\text{dist}_G(v, v) = \infty$ if v is not contained in any cycle).

While using the above notation we occasionally omit G if the identity of the specific graph is clear from the context.

We suppress the rounding signs to simplify the presentation.

The following auxiliary results are required for our proof.

Lemma 2.1. (Bounds on binomial coefficients, see e.g. [5], Chapter 21.1) *The following inequalities hold for every $n \geq a, k \geq b$.*

1. $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$;

2. $\frac{\binom{n-a}{k-b}}{\binom{n}{k}} \leq \left(\frac{k}{n}\right)^b \cdot \left(\frac{n-k}{n-b}\right)^{a-b}$.

Lemma 2.2. (Bounds on binomial random variables, a corollary of Lemma 2.1) *Let $1 \leq k \leq n$ be integers, $0 < p < 1$, and let $X \sim \mathsf{Bin}(n, p)$. Then the following inequalities hold:*

1. $\mathbb{P}(X \geq k) \leq \left(\frac{enp}{k}\right)^k$;

2. $\mathbb{P}(X = k) \leq \left(\frac{enp}{k(1-p)}\right)^k \cdot e^{-np}$.

If, additionally, $k \leq np$, then

3. $\mathbb{P}(X \leq k) \leq (k+1) \cdot \left(\frac{enp}{k(1-p)}\right)^k \cdot e^{-np}$.

Lemma 2.3. (Chernoff bound for binomial lower tail, see e.g. [6]) *Let $X \sim \mathsf{Bin}(n, p)$ and $\delta > 0$. Then*

$$\mathbb{P}(X < (1-\delta)np) \leq \exp\left(-\frac{\delta^2 np}{2}\right).$$

Lemma 2.4. (see e.g. [5], Theorem 3.3) *Let $d \in \mathbb{N}$, $c > 0$ and let $G \sim G(n, c/n)$. Then, with high probability,*

$$\left| |\{v \in G \mid d(v) = d\}| - e^{-c} \cdot \frac{c^d}{d!} \cdot n \right| = o\left(n^{0.6}\right).$$

Definition 2.1. *The* connected 2-core *of a graph G, denoted $C_2(G)$, is the 2-core of the largest connected component of G. That is, the graph $C_2(G)$ is the maximum size subgraph of the largest connected component in which all vertices have degree at least 2. If G does not have a unique largest component then $C_2(G) = \emptyset$.*

Lemma 2.5. (see e.g. [5], Lemma 2.16) *Let $\varepsilon > 0$. Then there is $C = C(\varepsilon) > 1$ such that, if $c \geq C$, $x < 1$ is the unique solution in $(0, 1)$ to $xe^{-x} = ce^{-c}$, and $G \sim G(n, c/n)$, then, with high probability,*

$$|V(C_2(G))| = (1-x) \cdot \left(1 - \frac{x}{c} + o(1)\right) \cdot n.$$

Additionally, the solution x satisfies $x = ce^{-c} + c^2 e^{-2c} + o(e^{-2c})$. Therefore, with high probability,

$$|V(C_2(G))| \geq (1 - (c+1)e^{-c} - (1+\varepsilon)c^2 e^{-2c})n.$$

3

We will also use a strong variant of the Azuma-Hoeffding inequality due to Warnke. To introduce it, we first need to define a vertex exposure martingale.

Definition 2.2. *(vertex exposure martingale, see [1], Section 7.1) Let f be a graph theoretic function, and let $G' \sim G(n,p)$, where $V(G')$ is assumed to be $[n]$, and G is a graph on $[n]$. The vertex exposure martingale corresponding to G and f is the sequence of random variables $X_0, ..., X_n$, where*

$$X_i := \mathbb{E}\left[f(G') \mid \forall x,y \in [i] : \{x,y\} \in E(G') \leftrightarrow \{x,y\} \in E(G)\right].$$

Lemma 2.6. (an immediate consequence of [12], Theorem 1.2) *Let $G \sim G(n,p)$, let f be a graph theoretic function, and let $X_0, X_1, ..., X_n$ be the corresponding vertex exposure martingale. Further assume that there is a graph property $\mathcal{P}$ and a positive integer d such that, for every two graphs G_1, G_2 on V such that $E(G_1) \triangle E(G_2) \subseteq \{v\} \times V$ (the symmetric difference between $E(G_1)$ and $E(G_2)$) for some $v \in V$, the following holds:*

$$|f(G_1) - f(G_2)| \leq \begin{cases} d & \text{if } G_1, G_2 \in \mathcal{P}; \\ n & \text{otherwise.} \end{cases}$$

Then

$$\mathbb{P}\left((X_n \geq X_0 + t) \ \wedge \ (G \in \mathcal{P})\right) \leq \exp\left(-\frac{t^2}{2n(d+1)^2}\right).$$

3 Proof of Theorem 1

In this section we present a proof of Theorem 1. Recall that only the upper bound in the statement requires a proof. Throughout the proof we assume that c is large enough (with respect to ε) without stating this explicitly.

Proof outline: Let $V_1 \subseteq V(G)$ be the set of vertices of degree 1. We prove the claim by showing that for a large enough value of c, with high probability, if we add to G the edges of a (partial) matching on V_1, the resulting graph contains a cycle of length at least $(1 - \frac{1}{2}\varepsilon ce^{-c})n$. Since, with high probability, $|V_1| \leq (1+\frac{1}{2}\varepsilon)ce^{-c}n$, by removing the added matching from the cycle we get a set of at most $(1+\frac{1}{2}\varepsilon)\frac{1}{2}ce^{-c}n$ paths that cover at least $(1 - \frac{1}{2}\varepsilon ce^{-c})n$ vertices. The remaining vertices can now be covered by paths of length 0 to establish the result.

In Section 3.1 we study the structure of G, and show its typical properties that will help with proving the likely existence of a long cycle. We then describe a construction of an auxiliary graph G^* and a matching $M \subseteq G^*$ whose vertex set is contained in $N_G(V_1)$, such that if G^* contains a Hamilton cycle that contains all the edges of M then G along with the matching on V_1, obtained from M by translating an edge between two vertices of $N_G(V_1)$ to an edge between their neighbours in V_1, contains a sufficiently long cycle. In short, G^* is obtained by adding M, a matching on $N_G(V_1)$, to the subgraph induced by G on a subset of the vertices. The vertex set of G^* is obtained by removing vertices from $V(G)$ that may prove problematic, such as pairs of vertices with very small degree that are within a short distance from one another. This will enable us to claim that, with high probability, the obtained G^* has some good expansion properties, which will be useful in the next section.

Finally, in Section 3.2 we prove that indeed, with high probability, G^* contains such a Hamilton cycle. Here, we will employ adapted variations on methods used in the setting of finding Hamilton cycles in

expanding graphs, namely, rotations and extensions. For an overview on this subject we refer the reader to [8], Section 2.3.

3.1 Properties of G and its 2-core

Recall that V_1 denotes the set of vertices in G of degree 1. Define the following three subsets of $V(G)$.

$$
\begin{aligned}
\text{SMALL} &= \text{SMALL}(G) = \{v \in V(G) \mid d_G(v, V(G) \setminus N(V_1)) < c/1000\}; \\
\text{LARGE} &= \text{LARGE}(G) = \{v \in V(G) \mid d_G(v) > 20c\}; \\
\text{CLOSE} &= \text{CLOSE}(G) = \{v \in \text{SMALL} \mid \exists u \in \text{SMALL} : \text{dist}_G(v, u) \leq 4\}.
\end{aligned}
$$

Notice that $V_1 \subseteq \text{SMALL}$, and that CLOSE, among other elements, contains all the vertices of SMALL that belong to cycles of length 3 or 4. Informally, vertices of SMALL and CLOSE can potentially be problematic later in the proof, when G^* is constructed, and therefore will require special treatment. In light of this, it will be useful to show that these sets are typically not very big. This, among other typical properties of G, is proved in the following lemma.

Lemma 3.1. *With high probability G has the following properties:*

(P1) $(1 - n^{-0.4}) \cdot ce^{-c} \cdot n \leq |V_1| \leq (1 + n^{-0.4}) \cdot ce^{-c} \cdot n$;

(P2) $|SMALL| \leq e^{-0.9c} \cdot n$;

(P3) $|LARGE| \leq 10^{-6} \cdot n$;

(P4) $|CLOSE| \leq e^{-1.8c} \cdot n$;

(P5) *every $U \subseteq V(G)$ with $|U| \leq 10^{-5} \cdot n$ has $e(U) < 10^{-4}c \cdot |U|$;*

(P6) *every $U, W \subseteq V(G)$ disjoint with $|U| = 10^{-6} \cdot n$ and $|W| = \frac{1}{5}n$ satisfy: $e(U, W) \geq 10^{-7}c \cdot n$.*

Proof. For each of the given properties, we bound the probability that $G \sim G(n, p)$ fails to uphold it.

(P1). This is an immediate consequence of Lemma 2.4, with $d = 1$.

(P2). Since $|N(V_1)| \leq |V_1|$, assuming G has **(P1)**, the probability that $|\text{SMALL}| \geq s := e^{-0.9c}n$ is at most the probability that there is a set U of size s and a set W of size $t = 2ce^{-c} \cdot n$ with less than $\frac{c}{1000} \cdot s$ edges between U and $V(G) \setminus (U \cup W)$. By Lemma 2.1(1) and Lemma 2.2(2), the probability for this is at most

$$
\begin{aligned}
&\binom{n}{s} \cdot \binom{n}{t} \cdot \mathbb{P}\left(\text{Bin}\left(s(n - s - t), p\right) < \frac{cs}{1000}\right) \\
\leq\ & \left(\frac{en}{s}\right)^s \cdot \left(\frac{en}{t}\right)^t \cdot \frac{cs}{1000} \cdot \mathbb{P}\left(\text{Bin}\left(s(n - s - t), p\right) = \frac{c}{1000} \cdot s\right) \\
\leq\ & O(n) \cdot e^{0.91c \cdot s} \cdot e^{2c \cdot t} \left(\frac{1000es(n - s - t)p}{cs(1 - p)}\right)^{10^{-3}cs} \cdot e^{-s(n - s - t)p} \\
\leq\ & e^{0.92c \cdot s} \cdot (1001e)^{10^{-3}cs} \cdot e^{-0.99c \cdot s} \\
\leq\ & e^{-0.06cs} \\
=\ & o(1).
\end{aligned}
$$

(P3). If there are more than $10^{-6}n$ vertices with degree more than $20c$, then in particular there is a set $U \subseteq V(G)$ of size $10^{-6}n$ with $e(U) + e(U, V(G) \setminus U) \geq 10^{-5}cn$. By Lemma 2.1(1) and Lemma 2.2(1), the probability of this is at most

$$\binom{n}{10^{-6}n} \cdot \mathbb{P}\left(\mathsf{Bin}\left(\binom{10^{-6}n}{2} + 10^{-6} \cdot (1 - 10^{-6})n^2, p\right) \geq 10^{-5}cn\right)$$
$$\leq 2^n \cdot \left(\frac{e \cdot 10^{-6}n^2 \cdot p}{10^{-5}cn}\right)^{10^{-5}cn}$$
$$\leq 2^n \cdot \left(\frac{e}{10}\right)^{10^{-5}cn}$$
$$= o(1).$$

(P4). In order to bound the probability that G fails to satisfy **(P4)** we aim to apply Lemma 2.6 with $f(G) = |\mathrm{CLOSE}(G)|$ and $\mathcal{P}$ the property $\Delta(G) \leq \log n$. First, we bound $X_0 = \mathbb{E}[f(G)]$ from above.

Let Y_1 denote the random variable counting cycles in G of length 3 or 4, and Y_2 the random variable counting the number of paths in G of length at most 4, such that both their endpoints are in SMALL. Then, deterministically, $|\mathrm{CLOSE}| \leq Y_1 + 2Y_2$, and therefore $\mathbb{E}[\mathrm{CLOSE}] \leq \mathbb{E}[Y_1] + 2\mathbb{E}[Y_2]$. It is already known that $\mathbb{E}[Y_1] = O(1)$ (see e.g. [5], Theorem 5.4), so it remains to bound $\mathbb{E}[Y_2]$.

By the definition of SMALL, if $v \in$ SMALL then either v has a small total degree in G, say, at most $10^{-3}c + \delta$, or v has more than δ neighbour in $N(V_1)$. We utilize this fact for $\delta = 10^{-4}c$ to bound the expectation of Y_2.

For a given pair of vertices $u, v \in V(G)$ and a path P of length $\ell \leq 4$ between them, the probability that $P \subseteq G$ and $u, v \in$ SMALL is at most p^ℓ times the probability of the event: at least one of the vertices u, v has at least $10^{-4}c$ neighbours in $V(G) \setminus V(P)$ such that each of these neighbours has a neighbour of degree 1, or $e_G(\{u, v\}, V(G) \setminus V(P)) \leq 2(10^{-3} + 10^{-4})c)$. The probability of this event is at most

$$2 \cdot (n^2 p^2)^{c/10000} \cdot (1 - p)^{c(n-5)/10000} + \mathbb{P}\left(\mathsf{Bin}\left(2(n - \ell - 1), p\right) \leq 2(10^{-3} + 10^{-4})c\right)$$
$$\leq e^{-c^2/20000} + c \cdot (2000e)^{c/450} \cdot e^{-2c}$$
$$\leq e^{-1.95c}.$$

From this we get

$$\mathbb{E}[Y_2] \leq \sum_{\ell=1}^{4} n^{\ell+1} p^\ell e^{-1.95c} \leq c^4 e^{-1.9c} n,$$

and therefore $\mathbb{E}[|\mathrm{CLOSE}|] \leq \mathbb{E}[Y_1] + 2\mathbb{E}[Y_2] \leq e^{-1.85c}n$.

In order to apply Lemma 2.6 it remains to determine the parameter d. We claim that setting $d = 3\log^7 n$ satisfies the conditions of the lemma. Let G_1, G_2 be two graphs on V with maximum degree at most $\log n$, that differ from each other only in edges incident to $v \in V$. It suffices to show that $\mathrm{CLOSE}(G_1) \triangle \mathrm{CLOSE}(G_2)$ only contains vertices of distance at most 7 from v in either G_1 or G_2. Indeed, let u be (without loss of generality) in $\mathrm{CLOSE}(G_1) \setminus \mathrm{CLOSE}(G_2)$. This can occur for two reasons.

1. v is part of a path of length at most 4 in G_1 between u and another vertex, or a cycle of length at most 4 that contains u. This can only occur if $\mathrm{dist}_{G_1}(u,v) \le 3$.

2. u or a vertex of distance at most 4 from u are in $\mathrm{SMALL}(G_1) \setminus \mathrm{SMALL}(G_2)$. Call this vertex w (possibly $w = u$). Going from G_2 to G_1, this can happen because $\{v, w\} \in E(G_2)$, or because changing the edges of v caused another neighbour of w to be moved into $N_{G_1}(V_1(G_1))$. In any of these cases the distance between v and w is at most 3 in one of the graphs G_1 or G_2, and therefore the distance between u and v is at most 7 in that graph.

Now we get by Lemma 2.6

$$\mathbb{P}\left(|\mathrm{CLOSE}| > \mathbb{E}[|\mathrm{CLOSE}|] + n^{2/3}\right)$$
$$\le \quad \mathbb{P}\left(|\mathrm{CLOSE}| > \mathbb{E}[|\mathrm{CLOSE}|] + n^{2/3} \;\wedge\; \Delta(G) \le \log n\right) + \mathbb{P}\left(\Delta(G) > \log n\right)$$
$$\le \quad \exp\left(-\frac{n^{4/3}}{2n(3\log^7 n + 1)^2}\right) + n \cdot \binom{n}{\log n} \cdot \left(\frac{c}{n}\right)^{\log n}$$
$$= \quad o(1),$$

which implies that $|\mathrm{CLOSE}| \le e^{-1.8c}n$ with high probability.

(P5). By Lemma 2.1(1) and Lemma 2.2(1), the probability that there is a set $U \subseteq V(G)$ of size $k \le 10^{-5}n$ that contradicts **(P5)** is at most

$$\binom{n}{k} \cdot \mathbb{P}\left(\mathrm{Bin}\left(\binom{k}{2}, p\right) \ge 10^{-4}ck\right) \le \left(\frac{en}{k}\right)^k \cdot \left(\frac{10^4 ek^2 p}{2ck}\right)^{10^{-4}ck} \le \left(\frac{10^4 ek}{2n}\right)^{10^{-5}ck}.$$

Observe that if $k \le \log n$ this expression is of order $n^{-\Omega(1)}$. By the union bound the probability that G does not have **(P5)** is at most

$$\sum_{k=1}^{10^{-5}n} \left(\frac{10^4 ek}{2n}\right)^{10^{-5}ck} \le \log n \cdot n^{-\Omega(1)} + \sum_{k=\log n}^{10^{-5}n} \left(\frac{e}{20}\right)^{10^{-5}ck} = o(1).$$

(P6). Since there are at most 3^n ways to choose two disjoint subsets of $V(G)$, by the union bound and by Lemma 2.3 with parameters $m = 2 \cdot 10^{-7} \cdot n^2$ and $\delta = \frac{1}{2}$, the probability that there are disjoint $U, W \subseteq V(G)$ with $|U| = 10^{-6} \cdot n$ and $|W| = \frac{1}{5}n$ and less than $10^{-7}c \cdot n$ edges between them is at most $3^n \cdot e^{-\frac{1}{4} \cdot 10^{-7}cn} = o(1)$.

$\square$

Towards constructing G^*, following is the definition of the set $\mathrm{BAD} = \mathrm{BAD}(G) \subseteq V(G)$, whose vertices we would like to exclude from G^*. Let $X \subseteq V(G)$ be a minimum size set such that $d_G(v, \mathrm{SMALL} \cup X) \le 1$ for every $v \notin X$, let $Y := \{v \in V(G) \mid d_G(v) = 2, d_G(v, X) = 1\}$ and set

$$\mathrm{BAD} := X \cup Y.$$

Informally, excluding BAD from our constructed graph G^* (rather than just excluding SMALL) will ensure that the remaining vertices in the graph all have very few neighbours in G outside of $V(G^*)$.

Observe that X, and therefore BAD, are well defined: if $X_1, X_2 \subseteq V(G)$ satisfy $\forall v \notin X_i : d_G(v, \text{SMALL} \cup X_i) \leq 1$, $i = 1, 2$, then $\forall v \notin X_1 \cap X_2 : d_G(v, \text{SMALL} \cup (X_1 \cap X_2)) \leq 1$ also holds, and therefore there is a unique smallest set X that satisfies the condition.

Lemma 3.2. *With high probability $|BAD| \leq 3\sqrt{c}e^{-c}n$.*

Proof. We first show that, with high probability, $|X| \leq x := \sqrt{c}e^{-c}n$. Let A denote the complementing event, and observe the vertices of X by order of addition to X, according to the following process: while there is a vertex outside X with degree at least 2 into $\text{SMALL} \cup X$, add the first such vertex (according to some pre-determined order). Assume that A happened, that is, assume that $|X| \geq x$, and let X' be the set containing the first x vertices added to X. Then there is a subset $Z \subseteq \text{SMALL}$ such that $|Z| \leq 2x$ and $e(Z \cup X') \geq 2x$. By the definition of SMALL, this means that the degree of every vertex in Z into $V(G) \setminus N(V_1)$ is at most $\frac{c}{1000}$.

Therefore, assuming G has property **(P1)**, A is contained in the following event: there are sets $X', Z, S, T \subseteq V(G)$ such that the following hold:

1. $|X'| = x$, $|Z| \leq 2x$, $|T| = |S| \leq 2ce^{-c}n$;

2. $e(Z \cup X') \geq 2x$;

3. $e(Z, V(G) \setminus (X' \cup Z \cup S \cup T)) \leq \frac{c}{1000} \cdot |Z|$;

4. all the vertices in T have degree 1, and their unique neighbour is in S, so that no two vertices in T share their neighbour.

Here, $T = V_1$ and S is a set of size $|T|$ that contains $N(V_1)$.

For given sets X', Z, S, T of respective sizes x, z, s, s, where the set T is ordered, the probability that all the conditions above are satisfied is at most

$$p^s(1-p)^{s(n-2)} \cdot \mathbb{P}\left(\text{Bin}\left(\binom{3x}{2}, p\right) \geq 2x\right) \cdot \mathbb{P}\left(\text{Bin}\left(z(n-z-x-2s), p\right) \leq 10^{-3}cz\right) \tag{1}$$

We first use Lemma 2.2 to bound from above some of the terms in Equation (1), under the assumption $s \leq 2ce^{-c}n$.

First,

$$\mathbb{P}\left(\text{Bin}\left(\binom{3x}{2}, p\right) \geq 2x\right) \leq \left(\frac{9ex^2p}{4x}\right)^{2x} \leq e^{-1.9cx}.$$

Next, if $s \leq 2ce^{-c}n$ then

$$\mathbb{P}\left(\text{Bin}\left(z(n-z-x-2s), p\right) \leq 10^{-3}cz\right) \leq n \cdot \left(\frac{1000ez \cdot np}{cz \cdot (1-p)}\right)^{10^{-3}cz} \cdot e^{-0.99cz} \leq e^{-0.9cz}.$$

Altogether we get that the expected number of such sets is at most

$$\binom{n}{s} \cdot (np \cdot e^{-(n-2)p})^s \binom{n}{x}\binom{n}{z} \cdot \mathbb{P}\left(\text{Bin}\left(\binom{3x}{2}, p\right) \geq 2x\right) \cdot \mathbb{P}\left(\text{Bin}\left(z(n-z-x-2s), p\right) \leq 10^{-3}cz\right)$$

$$\leq \left(\frac{en}{x}\right)^x \cdot \left(\frac{en}{z}\right)^z \cdot \exp\left(2ce^{-c}n - 1.9cx - 0.9cz\right)$$

$$\leq \exp\left(e^{-c}n + 2ce^{-c}n - 0.7cx\right)$$

$$\leq e^{-cx/2};$$

where in the second inequality we used the fact that $\left(\frac{e \cdot e^{-c}n}{z}\right)^z \leq e^{e^{-c}n}$, since the expression is maximized when $z = e^{-c}n$.

Finally, summing over all $O(n^2) = o(e^{cx/2})$ possibilities for s, z, by the union bound, we get that the probability that sets $X', Z, S, T \subseteq V(G)$ satisfying the above-listed conditions exist, and therefore $\mathbb{P}(A)$, is of order $o(1)$.

To finish the proof we observe that every vertex in $Y \setminus X$ has a unique neighbour in X, and that for two vertices $u, v \in Y \setminus (\text{CLOSE} \cup X) \subseteq \text{SMALL} \setminus \text{CLOSE}$ these unique neighbours in X are distinct, since otherwise $\text{dist}(u, v) \leq 2$, a contradiction. It follows that $|Y \setminus X| \leq |X| + |\text{CLOSE}|$ and therefore, due to property **(P4)**, with high probability we have

$$|\text{BAD}| \leq |X| + |Y \setminus X| \leq 2|X| + |\text{CLOSE}| \leq 3\sqrt{c}e^{-c}n.$$

$\square$

Recall Definition 2.1 and let $C_2 = C_2(G)$. Now set

$$V(G^*) = V(C_2) \setminus (\text{CLOSE} \cup \text{BAD}),$$

set M to be a pairing on the vertex set $N_G(V_1 \setminus \text{CLOSE}) \cap V(G^*)$ by order, that is, the smallest vertex of $N_G(V_1 \setminus \text{CLOSE}) \cap V(G^*)$ is paired to the second smallest and so on, so that all vertices are paired except possibly the largest one, and set $E(G^*) = E_G(V(G^*)) \cup E(M)$. Observe that no two vertices in $V(M)$ are connected by an edge of $G^* \setminus M$. Indeed, if two such vertices are connected by an edge, then all their neighbours in V_1 must also be in CLOSE, which implies that these vertices are not in $N_G(V_1 \setminus \text{CLOSE})$, and therefore not in $V(M)$. Also observe that $d_{G^*}(v, V(G^*) \setminus V(M)) \geq 2$ for every $v \in V(G^*)$. Indeed, if $v \in V(G^*) \setminus \text{SMALL}$ then it cannot have more than one neighbour in $Y \setminus \text{CLOSE} \subseteq \text{SMALL} \setminus \text{CLOSE}$, since this implies that its (at least) two neighbours in this set belong to CLOSE, a contradiction. Furthermore, it cannot have more than two neighbours in $X \cup \text{CLOSE} \subseteq X \cup \text{SMALL}$, since every vertex outside of X has at most one neighbour in $X \cup \text{SMALL}$, and $V(G^*) \cap X = \emptyset$ since $X \subseteq \text{BAD}$. Overall, v cannot have more than three neighbour in $(Y \setminus \text{CLOSE}) \cup X \cup \text{CLOSE} = \text{BAD} \cup \text{CLOSE}$, and therefore we get

$$d_{G^*}(v, V(G^*) \setminus V(M)) \geq d_G(v, V(G) \setminus V(M)) - 3 \geq \frac{c}{1000} - 3 \geq 2.$$

On the other hand, if $v \in V(G^*) \cap \text{SMALL}$ then v has no neighbours in CLOSE, since otherwise it would be in CLOSE itself. Additionally, if v has degree exactly 2 into $V(G) \setminus V(M)$ then it no neighbours in BAD, since otherwise, by definition, it would be a member of $X \cup Y = \text{BAD}$. Finally, if v has degree more than 2 into $V(G) \setminus V(M)$ then it has no neighbour in Y since otherwise it is a member of CLOSE, and at most one neighbour in X, and therefore v has at most one neighbour in BAD. In each of these cases we conclude that $d_{G^*}(v, V(G^*) \setminus V(M)) \geq 2$.

Lemma 3.3. *With high probability* $|V(G^*)| \geq (1 - (1 + \frac{1}{4}\varepsilon) \cdot ce^{-c}) \cdot n$.

Proof. The lemma follows immediately from property **(P4)** in Lemma 3.1, from the fact that, with high probability, $|V(C_2)| \geq (1 - (1 + \frac{1}{8}\varepsilon) \cdot ce^{-c}) \cdot n$ due to Lemma 2.5 with $\frac{1}{8}\varepsilon$, and from Lemma 3.2. $\square$

Lemma 3.4. *With high probability* $|V(M)| \geq (1 - \frac{1}{4}\varepsilon) \cdot ce^{-c}n$.

9

Proof. Let $V_0 \subseteq V(G)$ be the set of isolated vertices in G. By Lemmas 2.4 and 2.5, with high probability

$$|V(G) \setminus (V_1 \cup V_0 \cup C_2)| \leq 2c^2 e^{-2c} n.$$

By this inequality, and by the high probability bounds on the sizes of BAD and CLOSE we get from **(P4)** and Lemma 3.2, we get

$$|V(M)| \geq |V_1| - |\text{CLOSE}| - |\text{BAD}| - |V(G) \setminus (V_1 \cup V_0 \cup C_2)| \geq (1 - \frac{1}{4}\varepsilon) \cdot ce^{-c} n.$$

$\square$

Recall that M is a (partial) pairing of $N_G(V_1)$. For every $v \in V(M)$ denote by v' the unique neighbour of v in V_1, and let $V_1^* := \{v' \mid v \in V(M)\} \subseteq V_1 \setminus \text{CLOSE}$. Define a matching M' on V_1^* by matching the pair v', u' for every $\{v, u\} \in E(M)$.

In the next section we show that, with high probability, G^* contains a Hamilton cycle that contains all the edges of M, and show that this implies Theorem 1.

3.2 M-Hamiltonicity of G^*

Call a path $P \subseteq G^*$ an M-*path* if for every edge $e \in M$, either $e \in E(P)$ or $e \cap V(P) = \emptyset$, and similarly define an M-*cycle*. Say that a graph containing M is M-*Hamiltonian* is it contains a Hamilton M-cycle. Note that, by its definition, a Hamilton M-cycle must contain all the edges of M. For a non-M-Hamiltonian subgraph $H \subseteq G$ we say that $e \notin E(H)$ is an M-*booster* with respect to H if $H \cup \{e\}$ is M-Hamiltonian, or a maximum length M-path contained in $H \cup \{e\}$ is strictly longer than a maximum length M-path contained in H. Finally, say that a graph Γ on $V(G^*)$ is an M-*expander* if $M \subseteq \Gamma$, and for every $U \subseteq V(G^*)$ with $|U| \leq \frac{1}{4}n$ the inequality $|N_\Gamma(U) \setminus V(M)| \geq 2|U|$ holds.

Lemma 3.5. *For every M-expander Γ on $V(G^*)$ such that Γ is not M-Hamiltonian, $\binom{V(G^*)}{2}$ contains at least $\frac{1}{32}n^2$ M-boosters with respect to Γ.*

A similar statement to this is made in [2], Section 3.5. For completeness, and since the authors of the aforementioned paper use some tools different than ours in their proof, we add our somewhat simpler version of a proof to this claim.

Proof. Let $\Gamma \supseteq M$ be an M-expander with no Hamilton M-cycle. Recall first the definition of a Pósa rotation of a path and its pivot (see [10]). Say $(v_0, ..., v_\ell) = P \subseteq \Gamma$ is a path. If $\{v_i, v_\ell\} \in E(\Gamma)$ then we say that the path $(v_0, ..., v_{i-1}, v_i, v_\ell, v_{\ell-1}, ..., v_{i+1})$ is obtained from P by a rotation with fixed point v_0 and pivot v_i.

Say that a rotation of an M-path P is M-*respecting* if the pivot is not a vertex in $V(M)$, and observe that if P' is obtained from P by an M-respecting rotation then P' is also an M-path. Indeed, no new vertices were added to the path in the process, and the unique removed edge cannot be an edge of M, since it is incident to the pivot vertex, which is not in $V(M)$. For a path P with v one of its endpoints, denote by $\text{END}_M(P, v)$ the set of all endpoints (other than v) of paths that can be obtained from P by a sequence of M-respecting rotations with fixed point v.

We now claim that if $P = (v_0, ..., v_\ell)$ is a maximal M-path in Γ then

$$N_\Gamma(\text{END}_M(P, v_0)) \setminus V(M) \subseteq N_P(\text{END}_M(P, v_0)).$$

10

Indeed, assume otherwise. Since P is assumed to be a maximal M-path, it must be that $N_\Gamma(\mathrm{END}_M(P, v_0)) \setminus V(M) \subseteq P$. Then there is $v_i \in N_\Gamma(\mathrm{END}_M(P, v_0)) \setminus V(M)$, and $u \in \mathrm{END}_M(P, v_i)$ such that $\{u, v_i\} \in E(\Gamma)$ and such that $v_{i-1}, v_{i+1} \notin \mathrm{END}_M(P, v_0)$. Let P' be an M-path from v_0 to u that can be obtained from P by a sequence of M-respecting rotations, and w be the successor of v_i on P'. Observe that, since $v_{i-1}, v_i, v_{i+1} \notin \mathrm{END}_M(P, v_0)$, it must be that $w \in \{v_{i-1}, v_{i+1}\}$. But since $v_i \notin V(M)$ and in particular $\{v_i, w\} \notin E(M)$, a rotation of P' with fixed point v_0 and pivot v_i is M-respecting, implying $w \in \mathrm{END}_M(P, v_0)$, a contradiction.

Now, since we assumed that Γ is M-expanding and $|N_P(U)| < 2|U|$ for every set $U \subset V(P)$ that includes an endpoint of P, this implies that if $P = (v_0, ..., v_\ell)$ is maximal then $|\mathrm{END}_M(P, v_0)| > n/4$.

To complete the proof, for $u \in \mathrm{END}_M(P, v)$ denote by P_u an M-path from v to u obtained by M-respecting rotations. We claim that if P is maximal and v is an endpoint of P then any edge of $\binom{V(G^*)}{2}$ between $u \in \mathrm{END}_M(P, v)$ and $\mathrm{END}_M(P_u, u)$ is an M-booster. Indeed, if e is such an edge then $P_u \cup \{e\}$ is a cycle of length $|V(P)|$. If P was a Hamilton M-path then we obtained a Hamilton M-cycle. Otherwise, since M-expansion implies connectivity, there is an outgoing edge $\{x, y\} \in e(\Gamma)$, where $x \in V(P)$ and $y \notin V(P)$. Since M is a matching, at least one of the edges incident to x in P_u is not in $E(M)$, which implies the existence of a strictly longer path in $\Gamma \cup \{e\}$. Since there are at least $\frac{1}{2} \cdot \left(\frac{n}{4}\right)^2 = \frac{1}{32}n^2$ such edges, this finishes the proof.

$\square$

Lemma 3.6. *With high probability, for every M-expander $\Gamma \subseteq G^*$ that is not M-Hamiltonian and has at most $\frac{c}{900}n$ edges, $E(G^*)$ contains an M-booster with respect to Γ.*

Proof. By Lemma 3.5, $\binom{V(G^*)}{2}$ contains at least $\frac{1}{32}n^2$ M-boosters with respect to any such expander. We use the union bound over all choices for the triple $\Gamma, V(G^*), V(M)$ to bound the probability that G^* contains such an expander but none of its M-boosters, here using the fact that $V(M) \subseteq V(G^*) \subseteq V(G)$, that is, the pair $V(M), V(G^*)$ defines a partition of $V(G)$ into three sets, and therefore there are at most 3^n choices for the pair $V(M), V(G^*)$. We get

$$\sum_{k=n}^{\frac{c}{900}n} 3^n \cdot \binom{\binom{n}{2}}{k} \cdot p^k \cdot (1-p)^{\frac{1}{32}n^2} \leq \sum_{k=n}^{\frac{c}{900}n} 3^n \cdot \left(\frac{en^2 p}{2k}\right)^k \cdot e^{-\frac{c}{32}n} \leq cn \cdot 3^n \cdot (450e)^{\frac{c}{900}n} \cdot e^{-\frac{c}{32}n} = o(1),$$

where in the last inequality we used the fact that $\left(\frac{en^2 p}{2k}\right)^k$ is increasing when $k \leq \frac{c}{900}n$, and therefore if $k \leq \frac{c}{900}n$ then this size is at most $(450e)^{\frac{c}{900}n}$.

$\square$

Lemma 3.7. *With high probability G^* contains an M-expander Γ_0 with at most $\frac{c}{950}n$ edges.*

Proof. We describe a construction of a random subgraph Γ_0 of G^* with at most $\frac{c}{950}n$ edges and prove that if G has properties **(P1)-(P6)** then the constructed subgraph is an M-expander with positive probability, which implies existence.

Construct a random subgraph Γ_0 of G^* as follows. For every $v \in V(G^*)$ set E_v to be $E_{G^*}(v, V(G^*) \setminus V(M))$ in the case $v \in \mathrm{SMALL}$, and otherwise set E_v to be a subset of $E_{G^*}(v, V(G^*) \setminus V(M))$ of size $\frac{c}{1000}$, chosen uniformly at random. The random subgraph Γ_0 is the G^*-subgraph with edge set $M \cup \bigcup_{v \in V(G^*)} E_v$.

Observe that, since the minimum degree of a vertex into $V(G^*) \setminus V(M)$ in G^* is at least 2, this is also true for a graph Γ_0 drawn this way, and that $e(\Gamma_0) \leq \frac{c}{1000}n + \frac{1}{2}n \leq \frac{c}{950}n$.

We bound from above the probability that Γ_0 contains a subset U with at most $n/4$ vertices with less than $2|U|$ neighbours in $V(G^*) \setminus V(M)$. Let $|U| = k \leq \frac{n}{4}$, and denote $U_1 = U \cap \mathrm{SMALL}, U_2 = U \setminus U_1$ and k_1, k_2 the sizes of U_1, U_2 respectively. Observe that $V(\Gamma_0) \cap \mathrm{CLOSE} = \emptyset$ implies that *(i)* every vertex in U_2 has at most one neighbour in $U_1 \cup N_G(U_1) \cup V(M) \subseteq \mathrm{SMALL} \cup N_G(\mathrm{SMALL})$, and therefore $|N_{\Gamma_0}(U_2) \cap (U_1 \cup N_{\Gamma_0}(U_1) \cup V(M))| \leq k_2$; and *(ii)* distinct vertices in $\mathrm{SMALL}(G)$ have non-intersecting neighbourhoods, and therefore $|N_{\Gamma_0}(U_1) \setminus V(M)| \geq 2k_1$.

First we show that if $k_2 \leq 2 \cdot 10^{-6}n$ then $|N_{\Gamma_0}(U) \setminus V(M)| \geq 2|U|$ with probability 1. Indeed, in this case $|N_{\Gamma_0}(U_2) \setminus V(M)| \geq 4k_2$. Assume otherwise, then $U_2 \cup (N_{\Gamma_0}(U_2) \setminus V(M))$ is a set of size at most $5k_2 \leq 10^{-5}n$ which spans at least $\frac{c}{1000} \cdot k_2 - e(U_2)$ edges. But by **(P5)**, U_2 spans at most $10^{-4}ck_2$ edges, and $U_2 \cup (N_{\Gamma_0}(U_2) \setminus V(M))$ spans at most $5 \cdot 10^{-4}ck_2$ edges, so we get

$$10^{-3}c \cdot k_2 - 10^{-4}ck_2 \leq 5 \cdot 10^{-4}ck_2,$$

a contradiction. Now

$$
\begin{aligned}
|N_{\Gamma_0}(U) \setminus V(M)| &\geq |N_{\Gamma_0}(U_1) \setminus (U_2 \cup V(M))| + |N_{\Gamma_0}(U_2) \setminus (N_{\Gamma_0}(U_1) \cup U_1 \cup V(M))| \\
&\geq 2k_1 - k_2 + 4k_2 - k_2 \\
&\geq 2k_1 + 2k_2 = 2|U|.
\end{aligned}
$$

Now assume that $2 \cdot 10^{-6}n \leq k_2 \leq \frac{1}{4}n$. We show that $|N_{\Gamma_0}(U) \setminus V(M)| \geq 2|U|$ for every $U \subseteq V(G^*)$ with $|U| \leq \frac{n}{4}$ with positive probability. Suppose otherwise, then $|V(G^*) \setminus (U \cup N_{\Gamma_0}(U) \cup V(M))| \geq \frac{1}{5}n$. In particular, by **(P3)** there are disjoint sets $U' \subseteq U_2 \setminus \mathrm{LARGE}$ and $W \subseteq V(G^*) \setminus (U \cup N_{\Gamma_0}(U) \cup V(M))$, of sizes $10^{-6}n$ and $\frac{1}{5}n$ respectively, such that $e_{\Gamma_0}(U', W) = 0$. Observe that by **(P6)**, $e_{G^*}(U', W) \geq e_G(U', W) \geq 10^{-7}cn$. For a given pair of subsets U', W, by Lemma 2.1(2) the probability for this is at most

$$
\begin{aligned}
\prod_{u \in U'} \mathbb{P}(e_{\Gamma_0}(u, W) = 0) &\leq \prod_{u \in U'} \frac{\binom{e_{G^*}(u, V(G^*) \setminus V(M)) - e_{G^*}(u, W)}{10^{-3}c}}{\binom{e_{G^*}(u, V(G^*) \setminus V(M))}{10^{-3}c}} \\
&\leq \prod_{u \in U'} e^{-\frac{c}{1000} \cdot \frac{e_{G^*}(u, W)}{d_{G^*}(u)}} \\
&\leq \exp\left(-\frac{c}{1000} \cdot \frac{e_{G^*}(U', W)}{20c}\right) \\
&\leq \exp\left(-10^{-12} \cdot cn\right)
\end{aligned}
$$

Since there are at most 3^n pairs of subsets U', W, by the union bound the probability that two subsets of this size with no edges between them in Γ_0 exist is at most $3^n \cdot \exp\left(-10^{-12} \cdot cn\right) = o(1)$, for large enough c. Consequently, the random subgraph Γ_0 is an M-expander with probability $1 - o(1)$, implying that G^* contains a sparse M-expander, as claimed.

$\square$

Lemma 3.8. *With high probability G^* is M-Hamiltonian.*

Proof. This is an immediate consequence of Lemmas 3.6, 3.7. By Lemma 3.7, with high probability, G^* contains an M-expander $\Gamma_0 \subseteq G^*$ with at most $\frac{c}{950}n$ edges. By Lemma 3.6, with high probability, if Γ_0 is not M-Hamiltonian then $E(G^*)$ contains an M-booster e_1 with respect to Γ_0, an M-booster e_2 with respect to $\Gamma_0 \cup \{e_1\}$ and so on. After adding at most n such M-boosters, the resulting subgraph is already M-Hamiltonian, and consequently so is G^*. $\qquad\square$

To finish our proof, we now claim the G^* being M-Hamiltonian with high probability implies Theorem 1. Indeed, if C is a Hamilton M-cycle in G^*, then since $E(G^*) \subseteq E(G \cup M)$ it is an M-cycle in $G \cup M$ of length $|V(G^*)|$. Therefore, the cycle C' obtained by replacing every edge $\{u, v\}$ of M with the 3-path (u, u', v', v) is contained in $G \cup M'$, and has length

$$|V(C')| = |V(C)| + |V(M')| = |V(G^*)| + |V(M)|,$$

which, by Lemmas 3.3 and 3.4, is at least

$$(1 - (1 + \varepsilon/4) \cdot ce^{-c}) \cdot n + (1 - \varepsilon/4) \cdot ce^{-c}n = \left(1 - \frac{1}{2}\varepsilon ce^{-c}\right) \cdot n.$$

As discussed earlier, by removing M' from this cycle we are left with at most $(1 + \frac{1}{2}\varepsilon)\frac{1}{2}ce^{-c}n$ vertex disjoint paths that cover at least $\left(1 - \frac{1}{2}\varepsilon ce^{-c}\right) \cdot n$ vertices, and therefore by using additional $\frac{1}{2}\varepsilon ce^{-c}n$ paths of length 0 to cover the remaining vertices we get a disjoint path covering of G with at most $(1 + \varepsilon)\frac{1}{2}ce^{-c}n$ paths, thus finishing the proof. $\qquad\square$

References

[1] N. Alon and J. H. Spencer, **The Probabilistic Method**, John Wiley & Sons, 2016.

[2] M. Anastos and A. Frieze, *A scaling limit for the length of the longest cycle in a sparse random graph*, Journal of Combinatorial Theory Series B 148 (2021), 184–208.

[3] B. Bollobás, *The evolution of sparse graphs*, Graph Theory and Combinatorics, Academic Press, London (1984), 35–57.

[4] A. Frieze, *On large matchings and cycles in sparse random graphs*, Discrete Mathematics 59 (1986), 243–256.

[5] A. Frieze and M. Karoński, **Introduction to random graphs**, Cambridge University Press, 2015.

[6] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association 58 (1963), 13–30.

[7] J. Komlós and E. Szemerédi, *Limit distributions for the existence of Hamilton circuits in a random graph*, Discrete Mathematics 43 (1983), 55–63.

[8] M. Krivelevich. Long paths and Hamiltonicity in random graphs. In: *Random Graphs, Geometry and Asymptotic Structure*, N. Fountoulakis and D. Hefetz, Eds., London Mathematical Society Student Texts 84, Cambridge University Press (2016), 4–27.

[9] R. Nenadov, B. Sudakov and A. Z. Wagner, *Completion and deficiency problems*, Journal of Combinatorial Theory Series B 145 (2020), 214–240.

[10] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics 14 (1976), 359–364.

[11] B. Sudakov. Robustness of graph properties. In: *Surveys in Combinatorics 2017*, London Mathematical Society Lecture Note Series 440, Cambridge University Press (2017), 372–408.

[12] L. Warnke, *On the method of typical bounded differences*, Combinatorics, Probability and Computing 25 (2016), 269–299.