ON THE NUMBER OF HAMILTON CYCLES IN SPARSE RANDOM GRAPHS

ROMAN GLEBOV AND MICHAEL KRIVELEVICH

ABSTRACT. We prove that the number of Hamilton cycles in the random graph $G(n,p)$ is $n!p^n(1+o(1))^n$ a.a.s., provided that $p \geq \frac{\ln n + \ln \ln n + \omega(1)}{n}$. Furthermore, we prove the hitting-time version of this statement, showing that in the random graph process, the edge that creates a graph of minimum degree 2 creates $\left(\frac{\ln n}{e}\right)^n (1+o(1))^n$ Hamilton cycles a.a.s.

1. INTRODUCTION

The goal of this paper is to estimate the number of Hamilton cycles in the random graph $G(n,p)$. To be more formal, we show that the number of Hamilton cycle is asymptotically almost surely, or a.a.s. for brevity, concentrated around the expectation up to a factor $(1+o(1))^n$, provided the minimum degree is at least 2.

It is well known (see e.g. [5]) that the minimum degree of $G(n,p)$ is a.a.s. at most one for $p \leq \frac{\ln n + \ln \ln n - \omega(1)}{n}$, and therefore $G(n,p)$ contains no Hamilton cycle in this range of p a.a.s. Komlós and Szemerédi [23] and Korshunov [24] were the first to show that this bound is tight, i.e., $G(n,p)$ is a.a.s. Hamiltonian for every $p \geq \frac{\ln n + \ln \ln n + \omega(1)}{n}$. Bollobás [6] and independently Ajtai, Komlós, and Szemerédi [1] proved the hitting time version of the above statement, showing that in the random graph process, the very edge that increases the minimum degree to two also makes the graph Hamiltonian a.a.s.

There exists a rich literature about hamiltonicity of $G(n,p)$ in the range when it is a.a.s. Hamiltonian. Recent results include packing and covering problems (see e.g. [13], [21], [22], [26], [15], and [18]), local resilience (see e.g. [34], [14], [4] and [28]) and Maker-Breaker games ([33], [16], [3], and [11]). In this paper, we are interested in estimating the typical number of Hamilton cycles in a random graph when it is a.a.s. Hamiltonian. Several recent results about Hamiltonicity ([22],[26], [34],[11]) can be used to show fairly easily that $G(n,p)$ with $p = p(n)$ above the threshold for Hamiltonicity contains typically many, or even exponentially many Hamilton cycles. Here we aim however for (relatively) accurate bounds.

Using linearity of expectation we immediately see that the expected value of the number of Hamilton cycles in $G(n,p)$ is $\frac{(n-1)!}{2}p^n$. As the common intuition for random graphs may suggest, we expect the random variable to be concentrated around its

mean, perhaps after some normalization (it is easy to see that the above expressions for the expectation become exponentially large in n already for p inverse linear in n).

The reality appears to confirm this intuition – to a certain extent. Denoting by X the number of Hamilton cycles in $G(n, p)$, we immediately obtain $X < \left(\frac{np}{e}\right)^n$ a.a.s. by Markov's inequality. Janson [20] considered the distribution of X for $p = \Omega\left(1/\sqrt{n}\right)$ and proved that X is log-normal distributed, implying that $X = \left(\frac{np}{e}\right)^n (1 + o(1))^n$ a.a.s. It is instructive to observe that assuming $p = o(1)$, the distribution of X is in fact concentrated way below its expectation, in particular implying that $X/\mathbf{E}(X) \xrightarrow{p} 0$. For random graphs of density $p = o\left(n^{-1/2}\right)$ not much appears to be known about the asymptotic behavior of the number of Hamilton cycles in corresponding random graphs. We nevertheless mention the result of Cooper and Frieze [8], who proved that in the random graph process typically at the very moment the minimum degree becomes two, not only the graph is Hamiltonian but it has $(\log n)^{(1-o(1))n}$ Hamilton cycles.

Our main result is the following theorem, which can be interpreted as an extension of Janson's results [20] to the full range of $p(n)$.

Theorem 1. *Let $G \sim G(n, p)$ with $p \geq \frac{\ln n + \ln\ln n + \omega(1)}{n}$. Then the number of Hamilton cycles is $n! p^n (1 - o(1))^n$ a.a.s.*

Improving the main result of [8], we also show the following statement.

Theorem 2. *In the random graph process, at the very moment the minimum degree becomes two, the number of Hamilton cycles becomes $(\ln n/e)^n (1 - o(1))^n$ a.a.s.*

We continue with a short overview of related results for other models of random and pseudorandom graphs. For the model $G(n, M)$ of random graphs with n vertices and M edges, notice the result of Janson [20] showing in particular that for the regime $n^{3/2} \ll M \leq 0.99\binom{n^2}{m}$, the number of Hamilton cycles is indeed concentrated around its expectation. The situation appears to change around $M = \Theta\left(n^{3/2}\right)$, where the asymptotic distribution becomes log-normal instead. Notice also that the number of Hamilton cycles is more concentrated in $G(n, M)$ compared to $G(n, p)$; this is not surprising as $G(n, M)$ is obtained from $G(n, p)$ by conditioning on the number of edges of G being exactly equal to M, resulting in reducing the variance.

For the probability space of random regular graphs, it is the opposite case of very sparse graphs that is relatively well understood. Janson [19], following the previous work of Robinson and Wormald [31], [32], described the asymptotic distribution of the number of Hamilton cycles in a random d-regular graph $G(n, d)$ for a *constant* $d \geq 3$. The expression obtained is quite complicated, and we will not reproduce it here. For the case of growing degree $d = d(n)$, the result of Krivelevich [25] on the number of Hamilton cycles in (n, d, λ)-graphs in addition to known eigenvalue results for $G_{n,d}$ imply an estimation on the number of Hamilton cycles in $G_{n,d}$ with a superpolylogarithmic lower bound on d.

For an overview of these results as well as of the corresponding results in pseudorandom settings, we refer the interested reader to [25].

1.1. **Definitions and notation.** The random oriented graph $\vec{G}(n, p)$ is obtained from $G(n, p)$ by randomly giving an orientation to every edge (every of the two possible

directions with probability $1/2$). Notice that whenever we use the notation $\vec{G}$ for an oriented graph, there exists an underlying non-oriented graph obtained by omitting the orientations of the edges of $\vec{G}$; it is denoted by G. Making the notation consistent, when omitting the vector arrow above an oriented graph, we refer to the underlying non-oriented graph.

Given a graph G, we denote by $h(G)$ the number of Hamilton cycles in G. We call a spanning 2-regular subgraph of G a *2-factor*. Notice that every connected component of a 2-factor is a cycle. We denote by $f(G, s)$ the number of 2-factors in G with exactly s cycles. Similarly, a 1-factor of an oriented graph $\vec{G}$ is a spanning 1-regular subgraph, i.e., a spanning subgraph with all in- and outdegrees being exactly one. Analogously, the number of 1-factors in $\vec{G}$ with exactly s cycles is denoted by $f\left(\vec{G}, s\right)$. For the purposes of our proofs, we relax the notion of a 2-factor and call a spanning subgraph $H \subseteq G$ an *almost 2-factor* of G if H is a collection of vertex-disjoint cycles and at most $|V(G)| / \ln^2(|V(G)|)$ isolated vertices. We denote the number of almost 2-factors of G containing exactly s cycles by $f'(G, s)$. Similarly to the notation for non-oriented graphs, we call an oriented subgraph $\vec{H}$ of $\vec{G}$ an *almost 1-factor* of $\vec{G}$ if $\vec{H}$ is a 1-regular oriented graph on at least $|V(\vec{G})| - |V(\vec{G})| / \ln^2(|V(\vec{G})|)$ vertices. The number of almost 1-factors of $\vec{G}$ with exactly s cycles is denoted by $f'\left(\vec{G}, s\right)$.

As usual, in a graph G for a vertex $x \in V(G)$ we denote by $d_G(x) := |N_G(x)|$ its *degree*, i.e., the size of its neighborhood. We denote by $\delta(G)$ and respectively $\Delta(G)$ its minimum and maximum degrees. For a set $S \subseteq V(G)$, we denote by $N_G(S)$ the set of all vertices outside S having a neighbor in S. Whenever the underlying graph is clear from the context we might omit the graph from the index. Similarly, in an oriented graph $\vec{G}$ for a vertex $x \in V\left(\vec{G}\right)$ we call $d_{in,\vec{G}}(x) := \left|\left\{y \in V\left(\vec{G}\right) : yx \in E\left(\vec{G}\right)\right\}\right|$ the *indegree* of x and $d_{out,\vec{G}}(x) := \left|\left\{z \in V\left(\vec{G}\right) : xz \in E\left(\vec{G}\right)\right\}\right|$ the *outdegree* of x. We denote by $\delta_{in}\left(\vec{G}\right)$, $\Delta_{in}\left(\vec{G}\right)$, $\delta_{out}\left(\vec{G}\right)$, and $\Delta_{out}\left(\vec{G}\right)$ the minimum and maximum in- and outdegrees of $\vec{G}$.

In a graph G for two sets $A, B \subseteq V(G)$ we denote by $e_G(A, B)$ the number of edges incident with both sets. In an oriented graph $\vec{G}$, for two sets $A, B \subseteq V(G)$ the notation $e_{\vec{G}}(A, B)$ stands for the number of edges going from a vertex in A to a vertex in B. We write $e_G(A) := e_G(A, A)$ and $e_{\vec{G}}(A) := e_{\vec{G}}(A, A)$ for short. Similarly to the degrees, whenever the underlying graph is clear from the context we might omit the graph from the index.

To simplify the presentation, we omit all floor and ceiling signs whenever these are not crucial. Whenever we have a graph on n vertices, we suppose its vertex set to be $[n]$.

1.2. **Outline of the proofs.** In Section 2, the lower bounds for Theorems 1 and 2 are proven in the following steps.

- In Lemma 4 we show using the permanent of the incidence matrix that under certain pseudorandom conditions, an oriented graph contains sufficiently many oriented 1-factors.

- In Lemma 6 we prove that the random oriented graph $\vec{G}(n,p)$ a.a.s. contains a large subgraph with all in- and outdegrees being concentrated around the expected value. This subgraph then satisfies one of the conditions of Lemma 4.
- In Lemma 7 we show that the random graph $G(n,p)$ contains many almost 2-factors a.a.s. In the proof, we orient the edges of $G(n,p)$ randomly and apply Lemma 4 to the subgraph with almost equal degrees whose existence is guaranteed by Lemma 6 a.a.s.
- In Lemma 8 we prove that most of these almost 2-factors have few cycles a.a.s.
- We then call a graph p-expander if it satisfies certain expansion properties and show in Lemma 10 that in the random graph process, the graph $G(n,p)$ has these properties in a strong way.
- Lemma 11 shows that in any graph having the p-expander properties and minimum degree 2, for any path P_0 and its endpoint v_1 many other endpoints can be created by a small number of rotations with fixed endpoint v_1.
- Lemma 12 contains the main technical statement of the paper. Its states that in a graph satisfying certain pseudorandom conditions, for almost every almost 2-factor F with few components, there exists a Hamilton cycle with a small Hamming distance from F. The proof is a straightforward use of Lemma 11.
- The proofs of Theorems 1 and 2 are completed with a double counting argument. On the one hand, by Lemma 8 there exist many almost 2-factors with few cycles a.a.s. Furthermore, for each of these almost 2-factors there exists a Hamilton cycle with small Hamming distance from it a.a.s. by Lemma 12. On the other hand, for each Hamilton cycle, there are not many almost 2-factors with few cycles having a small Hamming distance from it. Hence, the number of Hamilton cycles is strongly related to the number of almost 2-factors with few cycles, finishing the proof.

1.3. **Chernoff bounds.** In many estimations we will have to bound the probability for a random variable to deviate far from its expectation. For this aim, we extensively use Chernoff bounds. We decided to state them explicitly in the following lemma, see e.g. Appendix A of [2].

Lemma 3. *Let X be a binomially distributed random variable with parameters n and p. Then the following is true.*

- *For every $\varepsilon > 0$ there exists a $c = c(\varepsilon) > 0$ such that $\mathbf{Pr}(|X - np| > \varepsilon np) < 2e^{-cnp}$.*
- *For $a > 2np$, $\mathbf{Pr}(X > a) < \left(\frac{enp}{a}\right)^a$.*

2. The proofs

Let $G \sim G(n,p)$. Since $\mathbf{E}(h(G)) = (n-1)!p^n/2$, we obtain

$$h(G) < \ln n (n-1)!p^n/2 < \left(\frac{np}{e}\right)^n$$

a.a.s., using just Markov's inequality. Thus, for the remainder of the section we are only interested in the lower bound on the typical number of Hamilton cycles in the random graph.

We know from [21] and using e.g. the results from [22] and [26] that in $G \sim G(n, p)$ there are at least $\left(\frac{\lfloor \delta(G)/2 \rfloor}{n} \right)^n n!$ 2-factors a.a.s. We now want to give an a.a.s. lower bound on the number of 2-factors in G, and we want to do it within a multiplicative error term of at most $2^{o(n)}$ from the "truth", basically deleting the 2 from the denominator in the above expression in the case $p \gg \ln n / n$, and replacing the term $\lfloor \delta(G)/2 \rfloor$ by asymptotically np.

We first prove a pseudo-random technical statement that will give us the desired inequality once we show that G (or a large subgraph of it) satisfies the pseudo-random conditions. The proof is based on the permanent method as used in [12].

Lemma 4. *Let $r = r(n) = \omega(\ln \ln n)$, and let $\vec{G}$ be an oriented graph on n vertices satisfying the following (pseudo-random) conditions:*

- *$\delta_{in}(\vec{G}), \delta_{out}(\vec{G}), \Delta_{in}(\vec{G}), \Delta_{out}(\vec{G}) \in (r - 4r/\ln \ln n, r + 4r/\ln \ln n)$*
- *for any two sets $A, B \subset V(\vec{G})$ of size at most $|A|, |B| \leq 0.6n$, there are at most $0.8r\sqrt{|A||B|}$ edges going from A to B.*

Then $\vec{G}$ contains at least $\left(\frac{r - 100r/\ln \ln n}{e} \right)^n$ oriented 1-factors, provided that n is sufficiently large.

Proof Create an auxiliary bipartite graph G' from $\vec{G}$ in the following way: take two copies X and Y of the vertex set $[n]$ by doubling each vertex $v \in [n]$ into $v_X \in X$ and $v_Y \in Y$. We put a (non-oriented) edge $uv \in E(G')$ between vertices $u_X \in X$ and $v_Y \in Y$ if $\vec{uv} \in E(\vec{G})$ is an edge oriented from u to v in $\vec{G}$. We observe a one-to-one correspondence between oriented 1-factors in $\vec{G}$ and perfect matchings in G'.

In order to use the permanent to obtain a lower bound on the number of perfect matchings of G', we need a (large) spanning regular subgraph of G'. Its existence is guaranteed by the following claim.

Claim 5. *G' contains a spanning regular subgraph G'' with regularity at least $d = r - 100r/\ln \ln n$.*

Proof Applying the Ore-Ryser theorem [29] we see that the statement of the claim is true provided that for every $Y' \subseteq Y$ we have

$$d|Y'| \leq \sum_{x \in X} \min\{d, e_{G'}(x, Y')\}.$$

Suppose to the contrary that this contrition does not hold, i.e., there exists a $Y' \subseteq Y$ s.t.

$$d|Y'| > \sum_{x \in X} \min\{d, e_{G'}(x, Y')\}.$$

We examine the number of edges incident to Y' that can be deleted from G' without disturbing the right hand side of the above inequality. Formally, we denote it by $c = \sum_{x \in X} \max\{0, e_{G'}(x, Y') - d\}$. Notice that

$$c = e_{G'}(X, Y') - \sum_{x \in X} \min\{d, e_{G'}(x, Y')\} > e_{G'}(X, Y') - d|Y'|$$

as supposed above.

Since $(r - 4r/\ln\ln n)|Y'| \le \delta(G')|Y'| \le e_{G'}(X, Y') < d|Y'| + c$, we obtain

$$c > \frac{96r}{\ln\ln n}|Y'|.$$

On the other hand, denoting by X' the set of vertices that have at least d neighbors in Y', and noticing that $\Delta(G') \le r + 4r/\ln\ln n$, we obtain

$$c \le \frac{104r}{\ln\ln n}|X'|.$$

Hence,

(1)
$$|X'| > 0.9|Y'|.$$

Notice that by the choices of Y' and X', we have

(2)
$$d|Y'| > \sum_{x \in X} \min\{d, e_{G'}(x, Y')\} = d|X'| + e_{G'}(X \setminus X', Y').$$

For the number of edges between $Y \setminus Y'$ and $X \setminus X'$ we see that

$$(r + 4r/\ln\ln n)|Y \setminus Y'| \ge e_{G'}(X \setminus X', Y \setminus Y') = e_{G'}(X \setminus X', Y) - e_{G'}(X \setminus X', Y')$$

$$\overset{(2)}{>} \delta(G')|X \setminus X'| - d(|Y'| - |X'|)$$

(3)
$$\ge \frac{96r}{\ln\ln n}|X \setminus X'| + (r - 100r/\ln\ln n)|Y \setminus Y'|,$$

leading to

(4)
$$|X \setminus X'| < 1.1|Y \setminus Y'|.$$

Furthermore, notice that by (2) it holds that

(5)
$$|X'| < |Y'|.$$

We prove the claim by case analysis.

- If $|Y'| \le n/2$, we obtain for the number of edges between X' and Y'

$$e_{G'}(X', Y') \overset{\text{Choice of } X'}{\ge} d|X'| \overset{(1)}{>} 0.9d\sqrt{|X'||Y'|} > 0.8r\sqrt{|X'||Y'|},$$

  contradicting the second condition of the lemma.
- If $|Y'| > n/2$, then again by the definition of X' we obtain $e_{G'}(X', Y') \ge d|X'|$, leading to

$$e_{G'}(X \setminus X', Y') \overset{(2)}{<} d(|Y'| - |X'|) = d(|X \setminus X'| - |Y \setminus Y'|) \overset{(4)}{<} 0.1d|Y \setminus Y'| \overset{(5)}{<} 0.1d|X \setminus X'|.$$

  Thus, using the fact that $\delta(G') > d$, we see that

$$e_{G'}(X \setminus X', Y \setminus Y') \ge 0.9d|X \setminus X'| \overset{(5)}{>} 0.8r\sqrt{|X \setminus X'| \cdot |Y \setminus Y'|},$$

  again contradicting the same condition of the lemma, since now both $X \setminus X'$ and $Y \setminus Y'$ have size less than $0.6n$ by (4).

$\square$

We observe that the number of perfect matchings in G'' equals the permanent of the incidence matrix of G''. Hence the result of Egorychev [9] and Falikman [10] on the conjecture of van der Waerden implies that the number of perfect matchings in G'' is at least $d^n n!/n^n > \left(\frac{d}{e}\right)^n$. $\square$

In order to use Lemma 4, we first prove the a.a.s. existence of a large subgraph of $\vec{G}(n,p)$ satisfying the degree-conditions of Lemma 4 a.a.s.

Lemma 6. *Let $\vec{G} \sim \vec{G}(n,p)$ with $p \geq \ln n/n$. Then there exists a set $V' \subseteq [n]$ of at least $n - n/\ln^2 n$ vertices of $\vec{G}$ such that the graph $\vec{C} := \vec{G}[V']$ satisfies $\delta_{in}(\vec{C}), \delta_{out}(\vec{C}), \Delta_{in}(\vec{C}), \Delta_{out}(\vec{C}) \in \left(\frac{np - 3np/\ln\ln n}{2}, \frac{np + np/\ln\ln n}{2}\right)$ a.a.s.*

Proof We observe using Lemma 3 that for $p \gg \ln n (\ln\ln n)^2/n$ the statement holds for $V' = [n]$ a.a.s. Hence, from now on we assume $np = O(\ln n (\ln\ln n)^2)$.

Let L be the set of all vertices whose in- or outdegree is at most $\frac{np - np/\ln\ln n}{2} + 1$. For every $y \in [n]$, we can estimate using Lemma 3

$$\mathbf{Pr}(y \in L) = \exp\left(-\Omega\left(\ln n/(\ln\ln n)^2\right)\right).$$

Thus, by Markov's inequality we obtain

$$|L| \leq \ln n \cdot \mathbf{E}(|L|) = \ln n \cdot (n-1)\exp\left(-\Omega\left(\ln n/(\ln\ln n)^2\right)\right)$$

(6)
$$= n\exp\left(-\Omega\left(\ln n/(\ln\ln n)^2\right)\right)$$

a.a.s.

Fix an arbitrary vertex $x \in [n]$. We denote

$$L_x = \left\{y \in [n] \setminus \{x\} : d_{in,\vec{G}-x}(y) \leq \frac{np - np/\ln\ln n}{2} \text{ or } d_{out,\vec{G}-x}(y) \leq \frac{np - np/\ln\ln n}{2}\right\}.$$

Notice that $L_x \subseteq L$, and thus (6) bounds $|L_x|$ as well.

Since for every $y \in [n] \setminus \{x\}$ the events "$xy \in E(G)$" and "$y \in L_x$" are independent, we obtain using Lemma 3 again

$$\mathbf{Pr}\left[\left(|N_G(x) \cap L_x| \geq \frac{np}{2\ln\ln n}\right) \mid \left(|L_x| = n\exp\left(-\Omega\left(\ln n/(\ln\ln n)^2\right)\right)\right)\right]$$

(7)
$$\leq \exp\left(-\frac{np}{2\ln\ln n}\Omega\left(\ln n/(\ln\ln n)^2\right)\right) = o(1/n).$$

Similarly, we let R be the set of all vertices whose in- or outdegree is at least $\frac{np + np/\ln\ln n}{2} - 1$ and obtain

(8)
$$|R| \leq n\exp\left(-\Omega\left(\ln n/(\ln\ln n)^2\right)\right)$$

a.a.s.

We define analogously

$$R_x = \left\{y \in [n] \setminus \{x\} : d_{in,\vec{G}-x}(y) \geq \frac{np + np/\ln\ln n}{2} - 1 \text{ or } d_{out,\vec{G}-x}(y) \geq \frac{np + np/\ln\ln n}{2} - 1\right\},$$

and observe analogously to (7) that $R_x \subseteq R$ and

(9)
$$\mathbf{Pr}\left[\left(|N_G(x) \cap R_x| \geq \frac{np}{2 \ln \ln n}\right) \mid \left(|R_x| = n \exp\left(-\Omega\left(\ln n/(\ln \ln n)^2\right)\right)\right)\right] = o(1/n).$$

We denote by V' the set of all vertices from $[n]$ whose in- an outdegrees in $\vec{G}$ lie in $\left(\frac{np-np/\ln\ln n}{2}, \frac{np+np/\ln\ln n}{2}\right)$. Notice that $[n] \setminus V' \subseteq L_x \cup R_x$ for every $x \in [n]$. Hence, we see that $|V'| > n - \frac{n}{\ln^2 n}$ a.a.s. by (6) and (8). Furthermore, from (7) and (9) we obtain that all in- an outdegrees in $\vec{G}[V']$ lie in $\left(\frac{np-3np/\ln\ln n}{2}, \frac{np+np/\ln\ln n}{2}\right)$ a.a.s., completing the proof of the lemma. $\qquad\square$

From now on, whenever we have n and p chosen, we denote
$$d = d(n,p) = np - 100np/\ln\ln n\,.$$

In the following lemma, we show that the random graph contains a.a.s. many 2-factors.

Lemma 7. *The random graph $G \sim G(n,p)$ with $p \geq \ln n/n$ satisfies*
$$\sum_{s \in [n/3]} 2^s f'(G,s) \geq d^{-n/\ln^2 n}(d/e)^n$$
a.a.s.

Proof   In order to use Lemma 4, we orient G at random to obtain $\vec{G}$ (as always in this paper, for every edge each of the two possible orientations gets probability $1/2$ independently of the choices of other edges).

First we show that the second condition of Lemma 4 holds a.a.s. for $\vec{G}$ with the intuitive choice $r = np/2$. Since the maximum degree of G is at most $3np$ a.a.s. (see e.g. [5]), we obtain that in G a.a.s. for any two sets A and B with $|A| > 100|B|$ the number of edges between them is at most $3np|B| < 0.4np\sqrt{|A||B|}$. Hence, $e_{\vec{G}}(A,B) \leq 0.4np\sqrt{|A||B|}$ and $e_{\vec{G}}(B,A) \leq 0.4np\sqrt{|A||B|}$. Thus, we are left with the case of sets A and B of sizes $|A| \leq 100|B|$ and $|B| \leq 100|A|$.

For small disjoint sets, we obtain using Lemma 3

$$\mathbf{Pr}\left(\exists A', B' \subset [n],\, A' \cap B' = \emptyset,\, |A'||B'| \leq \frac{n^2}{\ln\ln n},\, |A| \leq 100|B| \leq 10^4|A| : e_{\vec{G}}(A', B') \geq 0.4np\sqrt{|A'||B'|}\right)$$

$$\leq \sum_{a,b=o(n),\, a=\Theta(b)} \binom{n}{a}\binom{n-a}{b} \exp\left(-\Omega\left(np\sqrt{ab}\ln\left(\frac{np\sqrt{ab}}{pab}\right)\right)\right)$$

$$\leq \sum_{a,b=o(n)} \left(\frac{ne}{a}\right)^a \left(\frac{ne}{b}\right)^b \exp\left(-\Omega\left(a\ln n\ln\left(\Omega\left(\frac{n}{a}\right)\right)\right) - \Omega\left(b\ln n\ln\left(\Omega\left(\frac{n}{b}\right)\right)\right)\right)$$

$$\leq \sum_{a,b=o(n)} \exp\left(a\ln\left(\frac{ne}{a}\right) + b\ln\left(\frac{ne}{b}\right) - \Omega\left(a\ln\left(\frac{n}{a}\right)\ln n\right) - \Omega\left(b\ln\left(\frac{n}{b}\right)\ln n\right)\right)$$

$$= o(1).$$

Similarly, for large disjoint sets we obtain using Lemma 3

$$\mathbf{Pr}\left(\exists A', B' \subset [n],\ A' \cap B' = \emptyset,\ |A'||B'| > \frac{n^2}{\ln \ln n},\ |A|, |B| \leq 0.6n:\ e_{\vec{G}}(A', B') \geq 0.4np\sqrt{|A'||B'|}\right)$$

$$\leq \sum_{a,b \leq n,\, ab > \frac{n^2}{\ln \ln n}} \binom{n}{a}\binom{n-a}{b} \exp\left(-\Omega\left(abp\right)\right)$$

$$\leq 4^n \exp\left(-\Omega\left(\frac{n \ln n}{\ln \ln n}\right)\right) = o(1).$$

Hence, a.a.s. for every pair of disjoint sets A' and B', the number of edges going from A' to B' satisfies

$$(10) \qquad\qquad e_{\vec{G}}(A', B') < 0.4np\sqrt{|A'||B'|}.$$

Analogously, we see that a.a.s. for every $M \subseteq [n]$ of size at most $0.6n$,

$$(11) \qquad\qquad e_G(M) < 0.4np|M|.$$

Thus, a.a.s. for every $A, B \subset [n]$ of size $|A|, |B| \leq 0.6n$, the number of edges going from A to B in $\vec{G}$ is bounded by

$$e_{\vec{G}}(A, B) = e_{\vec{G}}(A \setminus B, B \setminus A) + e_G(A \cap B) \overset{(10),(11)}{<} 0.4np\sqrt{|A \setminus B||B \setminus A|} + 0.4np|A \cap B|$$

$$\leq 0.4np\sqrt{|A||B|},$$

establishing that the second condition of Lemma 4 holds a.a.s. for every subgraph of $\vec{G}$.

Hence, by Lemma 6 the graph $G \sim G(n,p)$ a.a.s. is such that for a random orientation $\vec{G}$, there a.a.s. exists a vertex set $V' \subseteq [n]$ of size at least $n - n/\ln^2 n$ such that the induced subgraph $\vec{G}[V']$ satisfies the conditions of Lemma 4 with $r = np/2$. Applying Lemma 4 to this induced subgraph, we obtain

$$\sum_{s \in [n/3]} f\left(\vec{G}[V'], s\right) \geq \left(\frac{d}{2e}\right)^{n - n/\ln^2 n}$$

a.a.s. Thus, we obtain

$$\sum_{s \in [n/3]} \mathbf{E}\left(f'\left(\vec{G}, s\right)\right) \geq (1 - o(1))\left(\frac{d}{2e}\right)^{n - n/\ln^2 n}$$

a.a.s., where the expectation is taken over the random choice of orienting the edges of G, the process creating $\vec{G}$ from G.

On the other hand, when we orient the edges, an almost 2-factor of G with exactly s cycles becomes an almost 1-factor of $\vec{G}$ with probability at most $2^{\frac{n}{\ln^2 n} - n + s}$, implying

$$\sum_{s \in [n/3]} 2^s f'(G, s) \geq \sum_{s \in [n/3]} 2^{n - \frac{n}{\ln^2 n}} \mathbf{E}\left(f'\left(\vec{G}, s\right)\right).$$

Putting these two facts together, we obtain

$$\sum_{s\in[n/3]} 2^s f'(G,s) \geq (1+o(1)) \left(\frac{d}{e}\right)^{n-n/\ln^2 n} \geq d^{-n/\ln^2 n}(d/e)^n$$

a.a.s., completing the proof of the lemma. $\qquad\square$

We show now that there are typically many almost 2-factors in G with a small number of cycles. We denote

$$s^* = s^*(n) = \frac{n}{\ln n \sqrt{\ln\ln n}}.$$

Lemma 8. *For every $p \geq \ln n/n$, the random graph $G \sim G(n,p)$ satisfies*

$$\sum_{s=1}^{s^*} f'(G,s) \geq (np/e)^n (1-o(1))^n$$

a.a.s.

Proof   By Lemma 7 we know that $\sum_{s\in[n/3]} 2^s f'(G,s) \geq d^{-n/\ln^2 n}(d/e)^n$ a.a.s.

We show now that the contribution of almost 2-factors with too many cycles is negligible. We use the estimate (5) of [21]: in the random graph $H \sim G(n',p)$, for every $s \geq \ln n'$,

$$\mathbf{E}(f(H,s)) \leq \frac{(n'-1)!\,(\ln n')^{s-1}\,p^{n'}}{(s-1)!2^s}.$$

We obtain

$$\sum_{s=s^*}^{n/3} \mathbf{E}(2^s f'(G,s)) \leq \sum_{\ell\leq n/\ln^2 n} \binom{n}{\ell} \sum_{s=s^*}^{n/3} \frac{n!\,(\ln n)^s\,p^{n-\ell}}{s!}$$

$$\leq n! p^n \left(\frac{n}{p}\right)^{n/\ln^2 n} \sum_{s=s^*}^{n/3} \left(\frac{s}{e\ln n}\right)^{-s}$$

$$= (d/e)^n e^{O(n/\ln\ln n)} \left(\frac{s^*}{e\ln n}\right)^{-s^*}$$

$$= o\left(d^{-n/\ln^2 n}\,(d/e)^n\right).$$

Hence, using this estimate together with Markov's inequality, we see that the number of almost 2-factors of G with at most s^* cycles is

$$\sum_{s=1}^{s^*} f'(G,s) \geq \frac{1}{2}2^{-s^*} d^{-n/\ln^2 n}\,(d/e)^n = (np/e)^n (1-o(1))^n$$

a.a.s. $\qquad\square$

The next technicality we need to prove in order to be ready to prove the main theorem is the expansion of $G(n,p)$.

To collect all but one expansion properties that we need, we make the following definition.

Definition 9. *We call a graph G with the vertex set $[n]$ a p-expander, if there exists a set $D \subset [n]$ such that G and D satisfy the following properties:*

- *$|D| \leq n^{0.09}$.*
- *The graph G does not contain a non-empty path of length at most $\frac{2 \ln n}{3 \ln \ln n}$ such that both of its (possibly identical) endpoints lie in D.*
- *For every set $S \subset [n] \setminus D$ of size $|S| \leq \frac{1}{p}$, its external neighborhood satisfies $|N(S)| \geq \frac{np}{1000}|S|$.*

The following lemma shows that these properties are pseudo-random.

Lemma 10. *Consider the two-round expansion of the random graph and fix $G \sim G(n,p)$ with $\ln n/n \leq p \leq 1 - 2 \ln \ln n/n$ and $G \subseteq \hat{G} \sim G(n,\hat{p})$ with $\hat{p} = p + 2 \ln \ln n/n$. Then it is a.a.s. true that every graph G' satisfying $G \subseteq G' \subseteq \hat{G}$ is a p-expander.*

Proof   We first expose G and fix $D = \{v \in [n] : d_G(v) < np/100\}$ to be the set of all vertices of G with degree less than $np/100$ in G. Since for a fixed set D the second property is decreasing and the third property is increasing, it suffices to prove the second statement for $\hat{G}$ and the third statement for G.

The first property is satisfied by Claim 4.3 from [4] a.a.s. The second property can be proven to hold in $\hat{G}$ a.a.s. similarly to Claim 4.4 from [4] (there it is proven to hold for G a.a.s.)

For the third property, assume to the contrary that there exists a set $S \subset [n] \setminus D$ of size at most $|S| \leq \frac{1}{p}$ such that its external neighborhood in G satisfies $|N_G(S)| < \frac{np}{1000}|S|$. By the definition of D, the number of edges incident to S in G is

$$e_G(S, N_G(S) \cup S) \geq |S|np/200.$$

But Lemma 3 tells us that

$$\mathbf{Pr}\left(\exists A, B \subseteq [n], |A| \leq \frac{1}{p}, |B| < \frac{np}{1000}|A| : e_G(A, B \cup A) \geq |A|np/200\right)$$

$$< \sum_{A,B \subset [n], |A| \leq 1/p, |B| < |A|np/1000} \left(e \cdot \frac{\mathbf{E}\left(|e_G(A, B \cup A)|\right)}{|A|np/200}\right)^{|A|np/200}$$

$$< \sum_{A,B \subset [n], |A| \leq 1/p, |B| < |A|np/1000} \left(\frac{200e|A||A \cup B|p}{|A|np}\right)^{|A|np/200}$$

$$< \sum_{a \leq 1/p} anp \binom{n}{a}\binom{n}{\frac{anp}{1000}}\left(\frac{3ap}{5}\right)^{anp/200}$$

$$< \sum_{a \leq 1/p} anp \left(\frac{3ap}{5}\right)^{anp/400} = o(1),$$

providing that the third property holds in G a.a.s. $\qquad\square$

The proof of the next lemma is based on the ingenious rotation-extension technique, developed by Pósa [30], and applied later in a multitude of papers on Hamiltonicity, mostly of random or pseudorandom graphs (see for example [7], [12], [23], [27]).

Let G be a graph and let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path in G. If $1 \leq i \leq q-2$ and (v_q, v_i) is an edge of G, then there exists a path $P' = (v_1 v_2 \ldots v_i v_q v_{q-1} \ldots v_{i+1})$ in G with the same set of vertices. The path P' is called a *rotation* of P_0 with *fixed endpoint* v_1 and *pivot* v_i. The edge (v_i, v_{i+1}) is called the *broken* edge of the rotation. We say that the segment $v_{i+1} \ldots v_q$ of P_0 is *reversed* in P'. In case the new endpoint v_{i+1} has a neighbor v_j such that $j \notin \{i, i+2\}$, then we can rotate P' further to obtain more paths of the same length. We will use rotations together with the expansion properties from Lemma 10 and the necessary minimum degree condition to find a path on the same vertex set as P_0 with large rotation endpoint sets.

The next lemma shows that in any graph having the p-expander property and minimum degree 2, for any path P_0 and its endpoint v_1, after a small number of rotations with fixed endpoint v_1, we either create many other endpoints or extend the path. Its proof has certain similarities to the proofs of Lemma 8 from [15] and of Claim 2.2 from [17].

Lemma 11. *Let n be a sufficiently large integer and G be an n-vertex p-expander with minimum degree $\delta(G) \geq 2$ and $np \geq \ln n$. Let P_0 be a v_1w-path in G. Denote by $B(v_1) \subset V(P_0)$ the set of all vertices $v \in V$ for which there is a v_1v-path on the vertex set $V(P_0)$ which can be obtained from P_0 by at most $3\frac{\ln n}{\ln(np)}$ rotations with fixed endpoint v_1. Then $B(v_1)$ satisfies one of the following properties:*

- *there exists a vertex $v \in B(v_1)$ with a neighbor outside $V(P_0)$, or*
- *$|B(v_1)| \geq n/3000$.*

Proof Assume that $B(v_1)$ does not have the first property (i.e., for every $v \in B(v_1)$ it holds that $N(v) \subseteq V(P_0)$).

Let t_0 be the smallest integer such that $\left(\frac{np}{3000}\right)^{t_0-1} \geq \frac{1}{p}$; note that $t_0 \leq 2\frac{\ln n}{\ln(np)}$. Since G is a p-expander, there is a corresponding vertex set D as in Definition 9.

At the first step, we find a neighbor $u \notin D \cup N(D)$ of w that is not a neighbor of w along P_0. Its existence is guaranteed since w has at least two neighbors along P_0, and by the second p-expansion property, at most one of them can have a neighbor in D. We rotate the initial path P_0 with pivot u and call the resulting path $P' = (v_1, \ldots, v_q)$. Notice that this way, v_q is guaranteed not to belong to D.

We construct a sequence of sets $S_0, \ldots, S_{t_0} \subseteq B(v_1) \setminus D \subseteq V(P_0) \setminus \{v_1\}$ of vertices, such that for every $0 \leq t \leq t_0$ and every $v \in S_t$, v is the endpoint of a path which can be obtained from P' by a sequence of t rotations with fixed endpoint v_1, such that for every $0 \leq i < t$, the non-v_1-endpoint of the path after the ith rotation is contained in S_i. Moreover, $|S_t| = \left(\frac{np}{3000}\right)^t$ for every $t \leq t_0 - 2$, $|S_{t_0-1}| = \frac{1}{p}$, and $|S_{t_0}| \geq \frac{n}{3000}$.

We construct these sets by induction on t. For $t = 0$, one can choose $S_0 = \{v_q\}$ and all requirements are trivially satisfied.

Let now t be an integer with $0 < t \leq t_0 - 1$ and assume that the sets $S_0, \ldots, S_{t-1}$ with the appropriate properties have already been constructed. We will now construct S_t. Let

$$T = \left\{ v_i \in N(S_{t-1}) : v_{i-1}, v_i, v_{i+1} \notin \bigcup_{j=0}^{t-1} S_j \cup D \right\}$$

be the set of potential pivots for the tth rotation, and notice that $T \subset V(P_0)$ due to our assumption, since $T \subseteq N(S_{t-1})$ and $S_{t-1} \subseteq B(v_1)$. Assume now that $v_i \in T$, $y \in S_{t-1}$ and $(v_i, y) \in E(G)$. Then, by the induction hypothesis, a $v_1 y$-path Q can be obtained from P' by $t - 1$ rotations such that after the jth rotation, the non-v_1-endpoint is in S_j for every $0 \leq j \leq t - 1$. Each such rotation breaks an edge which is incident with the new endpoint, obtained in that rotation. Since v_{i-1}, v_i, v_{i+1} are not endpoints after any of these $t - 1$ rotations, both edges (v_{i-1}, v_i) and (v_i, v_{i+1}) of the original path P' must be unbroken and thus must be present in Q.

Hence, rotating Q with pivot v_i will make either v_{i-1} or v_{i+1} an endpoint (which of the two, depends on whether the unbroken segment $v_{i-1} v_i v_{i+1}$ is reversed or not after the first $t - 1$ rotations). Assume without loss of generality that the endpoint is v_{i-1}. We add v_{i-1} to the set $\hat{S}_t$ of new endpoints and say that v_i *placed* v_{i-1} *in* $\hat{S}_t$. The only other vertex that can place v_{i-1} in $\hat{S}_t$ is v_{i-2} (if it exists).

Observe now that if $t < 0.1 \ln n / \ln \ln n$, the distance between any vertex from S_{t-1} and v_q is at most $2t - 2 < 0.2 \ln n / \ln \ln n$ by the way the sets were constructed. Hence, between any two vertices from $N(S_{t-1}) \cup N(N(S_{t-1}))$, there is a path of length at most $0.5 \ln n / \ln \ln n$. Thus at most one vertex from D can be in $N(S_{t-1}) \cup N(N(S_{t-1}))$. On the other hand, it $t \geq 0.1 \ln n / \ln \ln n$, then $|D| \leq n^{0.09} = o(|S_{t-1}|) = o(|N(S_{t-1})|)$. Thus, in both cases $|D \cap (N(S_{t-1}) \cup N(N(S_{t-1})))| = o(|N(S_{t-1})|)$.

Combining all this information together, we obtain

$$\begin{aligned}
|\hat{S}_t| &\geq \frac{1}{2} |T| \\
&\geq \frac{1}{2} \left(|N(S_{t-1})| - 3(1 + |S_1| + \ldots + |S_{t-1}| + |D \cap (N(S_{t-1}) \cup N(N(S_{t-1})))|) \right) \\
&\geq \left(\frac{np}{3000} \right)^t .
\end{aligned}$$

Clearly we can delete arbitrary elements of $\hat{S}_t$ to obtain S_t of size $\left(\frac{np}{3000} \right)^t$ if $t \leq t_0 - 2$ and of size $\frac{1}{p}$ if $t = t_0 - 1$. So the proof of the induction step is complete and we have constructed the sets $S_0, \ldots, S_{t_0 - 1}$.

To construct S_{t_0} we use the same technique as above, only the calculations are slightly different.

$$\begin{aligned}
|\hat{S}_{t_0}| &\geq \frac{1}{2} |T| \\
&\geq \frac{1}{2} \left(|N(S_{t_0-1})| - 3(1 + |S_1| + \ldots + |S_{t_0-2}| + |S_{t_0-1}| + |D \cap (N(S_{t-1}) \cup N(N(S_{t-1})))|) \right) \\
&\geq n/3000.
\end{aligned}$$

The set $S_{t_0} := \hat{S}_{t_0}$ is by construction a subset of $B(v_1)$, and the number of rotations needed to make any of its vertices an endpoint of the current path is at most $t_0 + 1$, concluding the proof of the lemma. $\square$

The proof of the following lemma relies on the final part of the proof of Theorem 1 from [25] and uses Lemma 11. It shows that under certain pseudorandom conditions in a graph G for every almost 2-factor, after adding few random edges, there exists a Hamilton cycle within a small Hamming distance from it a.a.s.

Lemma 12. *Let G be a connected n-vertex p-expander with minimum degree 2 and S be a set of vertices of G of size $|S| = o(n)$ such that there exist at least n non-edges in G not incident to S. Let F be an almost 2-factor of G with at most s^* cycles. Choose n non-edges $e_1, \ldots, e_n$ of G i.a.r. under the condition that none of them is incident to S and denote by G' the (random) graph obtained from G by turning them into edges. Then, if it is a.a.s. true that every graph $G \subseteq \hat{G} \subseteq G'$ is a p-expander, then G' a.a.s. contains a Hamilton cycle H with Hamming distance at most $17s^* \ln n / \ln(np)$ from F.*

Proof   Fix an arbitrary component $C \subseteq F$. Since G is connected, there exists an edge in G connecting a vertex $v \in V(C)$ and $y \notin V(C)$ - unless of course C is already Hamiltonian. We denote by C' the component of y in F. Opening C up by deleting an edge of C incident to v (no need to do so if C is just one isolated vertex), we get a path P. We append the edge vy to P, go through it to C', and if C' is a cycle, then we open it up by deleting an edge of C' incident to y to get a longer path P' and repeat the argument. If at some point there are no edges between the endpoints of the current path P'' and other components from F, then we can fix one endpoint x of P'' and rotate P'' using Lemma 11 to extend it outside or to obtain a set $B(x)$ of size at least $|B(x)| \geq n/3000$ of potential other endpoints. For every vertex $z \in B(x)$, we can rotate the resulting path fixing z as one endpoint to obtain a set $A(z)$ of size at least $|A(z)| \geq n/3000$ of potential other endpoints or to extend the path outside. If the path still cannot be extended outside and we can still not close it to a cycle, we have a set E' of at least $10^{-8} n^2$ non-edges of G not incident to S, so that turning any of them into an edge would close the path to a cycle. We add pairs $e_1, e_2, \ldots$ to $E(G)$, until one of them falls inside E'. Notice that for every $i \in [n]$, the pair e_i falls into E' with at least some constant positive probability. This means that considering events "$e_i \in E'$", every event has probability $\Theta(1)$ regardless of the previous events. Notice that in a successful round, the number of components gets reduced or a Hamilton cycle is created, since the edge that appeared in E' closed the path into a cycle or extended the path directly. To reduce the number of components by one, we do at most $\ln n / \ln(np)$ rotations by Lemma 11, therefore increasing the Hamming distance from F by at most $4 + 12 \ln n / \ln(np)$. Since it is enough to have $s^* + n/\ln^2 n$ successful events to obtain a Hamilton cycle, the expected number of needed turns of non-edges into edges is at most $O(1) \cdot (s^* + n/\ln^2 n) = o(n)$. Hence the n additional edges suffice to create a Hamilton cycle H from F by Markov's inequality a.a.s., replacing at most

$8 \ln n / \ln(np)$ edges for every component of F. Thus, the Hamming distance between F and H is at most $2 \cdot 8 \frac{\ln n}{\ln(np)}(s^* + n/\ln^2 n) \leq 17s^* \ln n/\ln(np)$. $\qquad \square$

We are now ready to prove Theorem 1.

Proof   Notice that only the lower bound is of interest for us. We expose G in two rounds.

We choose a function $p_1 = p_1(n)$ such that $\frac{\ln n + \ln\ln n + \omega(1)}{n} \leq p_1 \leq p - \frac{\omega(1)}{n}$, $p_1 \leq 1 - 2\ln\ln n/n$, and $p_1 = (1 - o(1))p$. In the first round, we expose $G_1 \sim G(n, p_1)$. We determine $D := \{v \in [n] : d_{G_1}(v) < np_1/100\}$.

In the second round, we expose the binomial random graph G_2 by including every edge from $K_n \setminus G_1$ into $E(G_2)$ with probability $p_2 := \frac{p - p_1}{1 - p_1}$. Since $(1 - p_1)(p_2) = 1 - p$, we obtain a graph $G := G_1 \cup G_2 \sim G(n, p)$. We know by Lemma 10 that a.a.s. every graph between G_1 and G including them both is a p_1-expander. Furthermore, notice that the expected number of edges in G_2 is $\binom{n}{2}p_2 \geq \binom{n}{2}(1 - p_1) = \omega(n)$, hence a.a.s. at least n additional random edges appeared in the second round of expansion by Markov's inequality. Since these edges were chosen i.a.r., and G_1 is a.a.s. connected with minimum degree at least 2 (see e.g. [5]), the conditions of Lemma 12 are satisfied for G_1 and the first n edges exposed in the second round with $S = \emptyset$.

Now, we put all we know together:

- By Lemma 8 we obtain $\sum_{s=1}^{s^*} f'(G_1, s) \geq (np_1/e)^n (1 - o(1))^n$ a.a.s.
- For every almost 2-factor F of G_1 with at most s^* cycles, there a.a.s. exists a Hamilton cycle in G with Hamming distance at most $k := 17s^* \ln n/\ln(np_1) = \frac{17n}{\ln(np_1)\sqrt{\ln\ln n}} = o(n)$ from F by Lemma 12.
- On the other hand, for every Hamilton cycle H in G, to obtain an almost 2-factor of G_1 of distance at most k from H, we can first delete at most k edges of H, thus obtaining a collection of at most k paths. These paths should then be tailored into an almost 2-factor, and the choices here are for each of the at most $2k$ endpoints of the paths to be connected to one of its $\Delta(G_1)$ neighbors in G_1 or to stay isolated. Thus, there are at most $\binom{n}{k}(\Delta(G_1) + 1)^{2k}$ almost 2-factors of G with Hamming distance at most k from H.
- Hence, by double counting almost 2-factors of G with at most s^* cycles, we obtain

$$h(G) \geq \frac{\sum_{s=1}^{s^*} f'(G', s)}{\binom{n}{k}(\Delta(G_1) + 1)^{2k}} \geq (np_1/e)^n \frac{(1 - o(1))^n}{2^{o(n)}(4\ln(np_1))^{2k}} = (np/e)^n (1 - o(1))^n$$

a.a.s. $\qquad \square$

To strengthen the result of Cooper and Frieze [8], we now prove Theorem 2.

Proof   The proof goes along the argument of Theorem 1, but now we expose the graph in three rounds. We first expose $G_1 \sim G(n, \ln n/n)$ and fix the set D of vertices of degree at most $\ln n/100$. Notice that similarly to the argument in the proof of Lemma 10, Claim 4.3 from [4] implies that $|D| \leq n^{0.09}$. In the second round of exposure, in addition to G_1 we expose those edges that are incident to D one by one, until the minimum degree becomes two; the resulting graph is called G'. In the third

round of exposure we consider the binomial random graph G_2 by including every edge of $K_n \setminus G_1$ not incident to D with probability $p_2 := \frac{\ln \ln n}{2n}$.

Let us denote by G the graph obtained by stopping the random graph process at the moment the minimum degree becomes two. Notice first that since in the random graph process $\delta \left(G \left(n, \frac{\ln n + 0.5 \ln \ln n}{n} \right) \right) = 1$ a.a.s., we obtain $G' \cup G_2 \subseteq G$ a.a.s., where by the union of two graphs with vertex sets $[n]$ we denote the graph on the same vertex set where the union is taken over the edge sets. Furthermore, observe that since in the random graph process $\delta \left(G \left(n, \frac{\ln n + 2 \ln \ln n}{n} \right) \right) \geq 2$ a.a.s., we obtain $G' \cup G_2 \subseteq G \left(n, \frac{\ln n + 2 \ln \ln n}{n} \right)$ a.a.s. (The two statements above can be found e.g. in [5].) Finally, G' is connected a.a.s. because of the expansion properties and the fact that the edge set between two linearly large sets is not empty (see e.g. [17]).

Since $p_2 = \omega(1/n)$ and $|D| = o(n)$ a.a.s., we obtain $|E(G_2)| = \omega(n)$ a.a.s. Furthermore, these edges are random under the only conditions of being non-edges of G_1 and being not incident to D. Hence, the conditions of Lemma 12 are satisfied for G' and the first n edges exposed in the third round.

Thus, following the lines of the proof of Theorem 1, we obtain the desired estimate.
□

3. Concluding remarks

In this paper we have proven that for any value of the edge probability $p = p(n)$, for which the random graph $G \, G(n, p)$ is a.a.s. Hamiltonian, the number of Hamilton cycles in G is $n! p^n (1 + o(1))^n$ a.a.s., thus being asymptotically equal to the expected value – up to smaller order exponential terms. Of course, it would be very nice to extend Janson's result [20] to smaller values of p and to understand more accurately the distribution of the number of Hamilton cycles in relatively sparse random graphs. However, given that the machinery used in [20] is rather involved, and the result (limiting distribution) is somewhat surprising, this will not necessarily be an easy task.

Our bound on the number of Hamilton cycles in $G(n, p)$ can be used to bound the number of perfect matching similarly to [25]. Let $m(G)$ denote the number of perfect matchings in the graph G. Since every Hamilton cycle is a union of two perfect matchings, we obtain $h(G) \leq \binom{m(G)}{2}$. Hence, for $G \sim G(n, p)$ the a.a.s. lower bound on $h(G)$ from Theorem 1 provides the a.a.s. lower bound $m(G) \geq (np/e)^{n/2}(1 - o(1))^n$. Since the upper bound is easily obtained from the expected value by Markov's inequality similarly to the first paragraph of Section 2, we have $m(G) = (np/e)^{n/2}(1 - o(1))^n$. The corresponding hitting time statement is obtained by a straightforward modification of the proof of Theorem 2: In the random graph process, the edge that makes the graph connected a.a.s. creates $(\ln n/e)^{n/2}(1 - o(1))^n$ perfect matchings.

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, First occurrence of Hamilton cycles in random graphs, Cycles in graphs (Burnaby, B.C., 1982), North-Holland Mathematical Studies **115**, North-Holland, Amsterdam (1985), 173–178. 1

[2] N. Alon, J. Spencer, **The probabilistic method**, 3rd Edition, Wiley, New York, 2008. 1.3

[3] S. Ben-Shimon, A. Ferber, D. Hefetz, M. Krivelevich, Hitting time results for Maker-Breaker games, Random Structures & Algorithms **41** (2012), 23–46. 1

[4] S. Ben-Shimon, M. Krivelevich, B. Sudakov, On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs, SIAM Journal on Discrete Mathematics **25** (2011), 1176–1193. 1, 2, 2

[5] B. Bollobás, **Random graph theory**, 2nd Edition, Cambridge University Press, Cambridge, 2001. 1, 2, 2

[6] B. Bollobás, The evolution of sparse graphs, Graph Theory and Combinatorics, Academic Press, London (1984), 35–57. 1

[7] B. Bollobás, T. I. Fenner, A. Frieze, An algorithm for finding Hamilton paths and cycles in random graphs, Combinatorica **7** (1987), 327–341. 2

[8] C. Cooper, A. Frieze, On the number of Hamilton cycles in a random graph, Journal of Graph Theory **13** (1989), 719–735. 1, 1, 2

[9] G. Egorychev, The solution of the Van der Waerden problem for permanents, Doklady Akademii Nauk SSSR **258** (1981), 1041–1044. 2

[10] D. Falikman, A proof of the Van der Waerden problem for permanents of a doubly stochastic matrix, Matematicheskie Zametki **29** (1981), 931–938. 2

[11] A. Ferber, R. Glebov, M. Krivelevich, A. Naor, Biased spanning games on random boards, in preparation. 1

[12] A. Frieze, M. Krivelevich, Hamilton cycles in random subgraphs of pseudo-random graphs, Discrete Mathematics **256** (2002), 137–150. 2, 2

[13] A. Frieze, M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, Journal of Combinatorial Theory Series B **94** (2005), 159-172. 1

[14] A. Frieze, M. Krivelevich, On two Hamilton cycle problems in random graphs, Israel Journal of Mathematics **166** (2008), 221–234. 1

[15] R. Glebov, M. Krivelevich, T. Szabó, On covering expander graphs by Hamilton cycles, Random Structures & Algorithms, to appear. 1, 2

[16] D. Hefetz, M. Krivelevich, M. Stojaković, T. Szabó, A sharp threshold for the Hamilton cycle Maker-Breaker game, Random Structures & Algorithms **34** (2009), 112–122. 1

[17] D. Hefetz, M. Krivelevich, T. Szabó, Hamilton cycles in highly connected and expanding graphs, Combinatorica **29** (2009), 547–568. 2, 2

[18] D. Hefetz, D. Kühn, J. Lapinskas, D. Osthus, Optimal covers with Hamilton cycles in random graphs, preprint. 1

[19] S. Janson, Random regular graphs: asymptotic distributions and contiguity, Combinatorics, Probability and Computing **4** (1995), 369–405. 1

[20] S. Janson, The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph, Combinatorics, Probability and Computing **3** (1994), 97–126. 1, 1, 3

[21] F. Knox, D. Kühn, D. Osthus, Approximate Hamilton decompositions of random graphs, Random Structures & Algorithms **40** (2012), 133–149 1, 2, 2

[22] F. Knox, D. Kühn, D. Osthus, Edge-disjoint Hamilton cycles in random graphs, preprint. 1, 2

[23] J. Komlós, E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, Discrete Mathematics **43** (1983), 55–63. 1, 2

[24] A. Korshunov, A solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs, Soviet Mathematics. Doklady, **17** (1976), 760–764. 1

[25] M. Krivelevich, On the number of Hamilton cycles in pseudo-random graphs, Electronic Journal of Combinatorics, Volume **19** (2012), publication P25. 1, 2, 3

[26] M. Krivelevich, W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, SIAM Journal on Discrete Mathematics **26** (2012), 964–982. 1, 2

[27] M. Krivelevich, B. Sudakov, Sparse pseudo-random graphs are Hamiltonian, Journal of Graph Theory **42** (2003), 17–33. 2

[28] C. Lee and B. Sudakov, Dirac's theorem for random graphs, Random Structures & Algorithms, **41** (2012), 293–305. 1

[29] O. Ore, **Theory of Graphs**, American Mathematics Society College Publishers, Providence, RI, USA 1962. 2

[30] L. Pósa, Hamiltonian circuits in random graphs, Discrete Mathematics **14** (1976), 359–364. 2

[31] R. Robinson, N. Wormald, Almost all cubic graphs are Hamiltonian, Random Structures & Algorithms **3** (1992), 117–125. 1

[32] R. Robinson, N. Wormald, Almost all regular graphs are Hamiltonian, Random Structures & Algorithms **5** (1994), 363–374. 1

[33] M. Stojaković, T. Szabó, Positional games on random graphs, Random Structures & Algorithms **26** (2005), 204–223. 1

[34] B. Sudakov, V. H. Vu, The local resilience of random graphs, Random Structures & Algorithms **33** (2008), 409–433. 1

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany

  E-mail address: `glebov@mi.fu-berlin.de`

School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

  E-mail address: `krivelev@post.tau.ac.il`