

## ON THE RESILIENCE OF HAMILTONICITY AND OPTIMAL PACKING OF HAMILTON CYCLES IN RANDOM GRAPHS\*

SONNY BEN-SHIMON<sup>†</sup>, MICHAEL KRIVELEVICH<sup>‡</sup>, AND BENNY SUDAKOV<sup>§</sup>

**Abstract.** Let  $\mathbf{k} = (k_1, \dots, k_n)$  be a sequence of  $n$  integers. For an increasing monotone graph property  $\mathcal{P}$  we say that a base graph  $G = ([n], E)$  is  $\mathbf{k}$ -resilient with respect to  $\mathcal{P}$  if for every subgraph  $H \subseteq G$  such that  $d_H(i) \leq k_i$  for every  $1 \leq i \leq n$  the graph  $G - H$  possesses  $\mathcal{P}$ . This notion naturally extends the idea of the *local resilience* of graphs recently initiated by Sudakov and Vu. In this paper we study the  $\mathbf{k}$ -resilience of a typical graph from  $\mathcal{G}(n, p)$  with respect to the Hamiltonicity property, where we let  $p$  range over all values for which the base graph is expected to be Hamiltonian. Considering this generalized approach to the notion of resilience our main result implies several corollaries which improve on the best known bounds of Hamiltonicity related questions. For one, it implies that for every positive  $\varepsilon > 0$  and large enough values of  $K$ , if  $p > \frac{K \ln n}{n}$ , then with high probability the local resilience of  $\mathcal{G}(n, p)$  with respect to being Hamiltonian is at least  $(1 - \varepsilon)np/3$ , improving on the previous bound for this range of  $p$ . Another implication is a result on optimal packing of edge-disjoint Hamilton cycles in a random graph. We prove that if  $p \leq \frac{1.02 \ln n}{n}$ , then with high probability a graph  $G$  sampled from  $\mathcal{G}(n, p)$  contains  $\lfloor \frac{\delta(G)}{2} \rfloor$  edge-disjoint Hamilton cycles, extending the previous range of  $p$  for which this was known to hold.

**Key words.** random graphs, resilience, Hamiltonicity

**AMS subject classifications.** 05C80, 05C45

**DOI.** 10.1137/110821299

**1. Introduction.** A *Hamilton cycle* in a graph is a simple cycle that traverses all vertices of the graph. The study of whether a graph contains a Hamilton cycle, or the *Hamiltonicity* graph property, became one of the main themes of graph theory from the very beginning. Deciding whether a graph is *Hamiltonian* was one of the first problems that were proved to be NP-COMplete, giving some insight on the elusiveness of this problem. Although the computational problem of determining in “reasonable” time whether a graph is Hamiltonian seems hopeless, the road does not need to stop there, as there are many other natural and interesting questions about this property. For one, given a prescribed number of edges  $m$ , are most graphs on  $n$  vertices with  $m$  edges Hamiltonian? Hence, the study of Hamiltonicity of *random graphs* seems like a natural approach to pursue along the path to understanding this intricate property.

Fixing  $m$  and selecting uniformly at random a graph on  $n$  vertices with  $m$  edges was indeed the original random graph model introduced by Erdős and Rényi, but the most widely studied random graph model is the binomial random graph,  $\mathcal{G}(n, p)$ . In this model we start with  $n$  vertices, labeled, say, by  $1, \dots, n$ , and select a graph on these  $n$  vertices by going over all  $\binom{n}{2}$  pairs of vertices, deciding independently with probability  $p$  for a pair to be an edge. The product probability space nature gives this model a greater

---

\*Received by the editors January 18, 2011; accepted for publication April 29, 2011; published electronically July 28, 2011.

<http://www.siam.org/journals/sidma/25-3/82129.html>

<sup>†</sup>School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel (sonny.benshimon@cs.tau.ac.il). This author’s research was partially supported by a Farajun Foundation Fellowship.

<sup>‡</sup>School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel (krivelev@tau.ac.il). This author’s research was supported in part by a USA-Israel BSF grant and by a grant from the Israel Science Foundation.

<sup>§</sup>Department of Mathematics, University of California, Los Angeles, CA 90005 (bsudakov@math.ucla.edu). This author’s research was supported in part by NSF CAREER award DMS-0812005 and by USA-Israeli BSF grant.

appeal than the original one, but they are indeed very much related and, in a sense, equivalent (see monographs [7] and [16] for a thorough introduction to the subject of random graphs). We note that we will sometimes abuse the notation and use  $\mathcal{G}(n, p)$  to denote both the distribution on the graphs just described and a random sample from this distribution; which of the two should be clear from the context.

We denote by  $\mathcal{HAM}$  the graph property of having a Hamilton cycle. One of the cornerstone results in the theory of random graphs is that of Bollobás [6] and of Komlós and Szemerédi [18], who proved that if  $G \sim \mathcal{G}(n, p)$  for  $p \geq \frac{\ln n + \ln \ln n + \omega(1)}{n}$  (where  $\omega(1)$  is any function tending to infinity with the number of vertices,  $n$ ), then with high probability (w.h.p.)<sup>1</sup>  $G \in \mathcal{HAM}$ . It is fairly easy to show (see, e.g., [7, Chapter 3]) that if  $p \leq \frac{\ln n + \ln \ln n - \omega(1)}{n}$ , then w.h.p.  $G$  contains at least one vertex of degree smaller than 2 (and hence is not Hamiltonian). This idea not only gives a very precise range of  $p$  for which the typical graph from  $\mathcal{G}(n, p)$  is Hamiltonian, but it also suggests that the bottleneck for  $\mathcal{HAM}$  (at least for small values of  $p$ ) stems from the vertices of low degree.

**1.1. Types of resilience.** Let  $\mathcal{P}$  be a monotone increasing graph property (i.e., a family of graphs on the same vertex set that is closed under the addition of edges and isomorphism). Counting the minimal number of edges one needs to remove from a base graph  $G$  in order to obtain a graph not in  $\mathcal{P}$  may, arguably, seem like one of the most natural questions to consider. Indeed, this notion, which is now commonly denoted by the *global resilience* of  $G$  with respect to  $\mathcal{P}$  (or the *edit distance* of  $G$  with respect to  $\bar{\mathcal{P}}$ ) is one of the fundamental questions in extremal combinatorics. This field can be traced back to the celebrated theorem of Turán [24] which states (in this terminology) that the complete graph  $K_n$  on  $n$  vertices has global resilience  $\frac{n}{2}(\frac{n}{r} - 1)$  (assuming  $r$  divides  $n$ ) with respect to containing a copy of  $K_{r+1}$ .

For some graph properties of a global nature, such as  $\mathcal{HAM}$  or being connected, the removal of all edges incident to a vertex of minimum degree is enough to destroy them, hence supplying a trivial upper bound on the global resilience of any graph with respect to these properties. For such properties the notion of global resilience does not seem to convey what one would expect from such a distance measure. One would like to gain some control of the number of edges incident to a single vertex that can be removed. To pursue this approach, for a given base graph  $G = ([n], E)$  one would like to gain a better understanding of the possible degree sequences of subgraphs  $H$  of  $G$  for which the graph  $G - H$  possesses the property  $\mathcal{P}$ .

Let  $\mathbf{a} = (a_1, \dots, a_n)$  and  $\mathbf{b} = (b_1, \dots, b_n)$  be two sequences of  $n$  numbers. We write  $\mathbf{a} \leq \mathbf{b}$  if  $a_i \leq b_i$  for every  $1 \leq i \leq n$ . Given a graph  $G$  on vertex set  $[n]$ , we denote its degree sequence by  $\mathbf{d}_G = (d_G(1), \dots, d_G(n))$ .

**DEFINITION 1.1.** Let  $G = ([n], E)$  be a graph. Given a sequence  $\mathbf{k} = (k_1, \dots, k_n)$  and a monotone increasing graph property  $\mathcal{P}$ , we say that  $G$  is  $\mathbf{k}$ -resilient with respect to the property  $\mathcal{P}$  if for every subgraph  $H \subseteq G$  such that  $\mathbf{d}_H \leq \mathbf{k}$ , we have  $G - H \in \mathcal{P}$ .

It was Sudakov and Vu [23] who initiated the systematic study of such a notion, albeit stated a little differently. In their original work the object of study was the minimum value of the maximum degree of a non- $\mathbf{k}$ -resilient sequence. They coined this parameter as the *local resilience* of a graph with respect to  $\mathcal{P}$ . We will use the following notation to denote this parameter:

$$r_\ell(G, \mathcal{P}) = \min\{r: \exists H \subseteq G \text{ such that } \Delta(H) = \text{rand } G \setminus H \notin \mathcal{P}\}.$$

---

<sup>1</sup>In this paper, we say that a sequence of events  $\mathcal{A}_n$  in a random graph model occurs w.h.p. if the probability of  $\mathcal{A}_n$  tends to 1 as the number of vertices  $n$  tends to infinity.

So, for local resilience, there is a uniform constraint on the number of deletions of edges incident to a single vertex. Although not explicitly, the study of local resilience lies in the heart of previous results in classical graph theory. In fact, one of the cornerstone results in the study of  $\mathcal{HAM}$  is the classical theorem of Dirac (see, e.g., [12, Theorem 10.1.1]), which states (in this terminology) that  $K_n$  has local resilience  $n/2$  with respect to  $\mathcal{HAM}$ . As has already been pointed out in [23], there seems to be a duality between the global or local nature of the graph property at hand and the type of resilience that is more natural to consider. More specifically, global resilience seems to be a more appropriate notion for studying *local* properties (e.g., containing a copy of  $K_k$ ), whereas for global properties (e.g., being Hamiltonian), the study of local resilience appears to be more natural.

This study of resilience has gained popularity, and in a relatively short period of time quite a few research papers have studied this and related distance notions [23, [14], [11], [19], [3], [2], [5], [9], [20], [15], [4]. This evolving body of research explored the resilience with respect to many graph properties, where the base graph of focus was mainly the binomial random graph  $\mathcal{G}(n, p)$  and graphs from families of pseudorandom graphs. In one of the subsequent papers of Dellamonica et al. [11], a more refined version of local resilience was considered. Consider a graph  $G$  with degree sequence  $\mathbf{d}$ . The authors of [11] defined the local resilience of  $G$  with respect to  $\mathcal{P}$  as the maximal value of  $0 \leq \alpha \leq 1$  for which a graph  $G$  is  $(\alpha\mathbf{d})$ -resilient with respect to  $\mathcal{P}$ . This notion is a little more robust than the original definition of local resilience as it can also deal with degree sequences of irregular graphs. The term  $\mathbf{k}$ -resilience is a further generalization on the same theme. The main motivation for studying this more general notion is that it allows us to give a unified approach to several problems, as will be exposed.

**1.2. Previous work.** One of the first local resilience problems Sudakov and Vu [23] considered was the local resilience problem of  $\mathcal{G}(n, p)$  with respect to  $\mathcal{HAM}$ . They provided both an upper bound and a lower bound for this parameter for almost the whole range of  $p$ . First, note that by the threshold probability for minimum degree 2, the local resilience parameter becomes interesting only for  $p \geq \frac{\ln n + \ln \ln n + o(1)}{n}$  (as for lower values of  $p$ , it is by definition equal to zero for non-Hamiltonian graphs, which w.h.p. is the case in this range). As an upper bound they proved that for every  $0 \leq p \leq 1$  w.h.p.

$$(1.1) \quad r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) \leq \frac{np}{2}(1 + o(1)).$$

Maybe more importantly, the lower bound they proved states that for every  $\delta, \varepsilon > 0$ , if  $p \geq \frac{\ln^{2+\delta} n}{n}$ , then w.h.p.  $r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) \geq \frac{np}{2}(1 - \varepsilon)$ , which essentially settles the problem for this range of  $p$ . Note that in fact this can be viewed as a far reaching generalization of Dirac's theorem. Frieze and Krivelevich in [14] studied this problem for the range of  $p$  "shortly after"  $\mathcal{G}(n, p)$  becomes Hamiltonian w.h.p., but the lower bound they obtained in this range is weaker. They proved that there exist absolute constants  $\alpha, C > 0$  such that for every  $p \geq \frac{C \ln n}{n}$  w.h.p.  $r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) \geq \alpha np$ . Recently, the authors in [5] improved on the above by showing that for every  $\varepsilon > 0$  there exists an absolute constant  $C > 0$  such that if  $p \geq \frac{C \ln n}{n}$ , then w.h.p.  $r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) \geq \frac{np}{6}(1 - \varepsilon)$ . It is plausible that w.h.p.  $r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) = \frac{np}{2}(1 \pm o(1))$  as soon as  $p \gg \frac{\log n}{n}$ , but the above-mentioned results still leave a gap to fill. In this work we make some progress on this front, but, alas, we are unable to close the gap completely.

A related question is the number of edge-disjoint Hamilton cycles one can have in a graph. Nash–Williams [21] asserted that Dirac's sufficient condition for a Hamilton cycle in fact guarantees at least  $\lfloor \frac{5n}{224} \rfloor$  edge-disjoint Hamilton cycles. Quite recently,

Christofides et al. [10], answering one of Nash–Williams’ original conjectures asymptotically, proved that minimum degree  $(\frac{1}{2} + o(1))n$  is sufficient for the existence of  $\lfloor \frac{n}{8} \rfloor$  edge-disjoint Hamilton cycles in a graph. When considering random graphs, the current knowledge about the packing of edge-disjoint Hamilton cycles is even more satisfactory. Bollobás and Frieze [8] showed that for every fixed  $r$ , if  $p \geq \frac{\ln n + (2r-1) \ln \ln n + \omega(1)}{n}$ , the minimal  $p$  for which  $\delta(\mathcal{G}(n, p)) \geq 2r$ , one can typically find  $r$  edge-disjoint Hamilton cycles in  $\mathcal{G}(n, p)$ . Kim and Wormald [17] established a similar result for random  $d$ -regular graphs (for fixed  $d$ ), proving that such graphs typically contain  $\lfloor \frac{d}{2} \rfloor$  edge-disjoint Hamilton cycles. The previous statements are of course the best possible but invite the natural question of what happens when the minimum degree is allowed to grow with  $n$ . Denote by  $\mathcal{H}_\delta$  the property of a graph  $G$  to contain  $\lfloor \delta(G)/2 \rfloor$ -edge-disjoint Hamilton cycles. Frieze and Krivelevich in [14], extending [8], showed that if  $p \leq \frac{(1+o(1)) \ln n}{n}$ , then w.h.p.  $\mathcal{G}(n, p) \in \mathcal{H}_\delta$ . In [13] Frieze and Krivelevich conjectured that this property is in fact typical for the whole range of  $p$ .

CONJECTURE 1.2 (Frieze and Krivelevich [13]). *For every  $0 \leq p(n) \leq 1$ , w.h.p.  $\mathcal{G}(n, p)$  has the  $\mathcal{H}_\delta$  property.*

In this paper, we are able to extend the range of  $p$  for which Conjecture 1.2 holds, but we cannot resolve the conjecture completely.

**1.3. Our results.** As previously mentioned, in this work we explore the notion of  $\mathbf{k}$ -resilience of random graphs with respect to  $\mathcal{HAM}$ . To state our main result we need the following notation. Let  $G = ([n], E)$  be a graph. For every positive  $t$  we denote by  $\mathcal{D}_t = \mathcal{D}_t(G) = \{v \in [n] : d_G(v) < t\}$  the subset of vertices of degree less than  $t$ . Denote by  $\mathbf{d} = (d_1, \dots, d_n)$  the degree sequence of  $G$ . For every constant  $\varepsilon > 0$  and  $t > 0$  we define the (not necessarily integral) sequence  $\tilde{\mathbf{d}}(t, \varepsilon) = (\tilde{d}_1, \dots, \tilde{d}_n)$  as follows:

1.  $\tilde{d}_v = d_v - 2$  for every  $v \in \mathcal{D}_t(G)$ ;
2.  $\tilde{d}_v = d_v(\frac{1}{3} - \varepsilon)$  for every  $v \in V_1 = V \setminus \mathcal{D}_t(G)$ .

THEOREM 1. *For every  $\varepsilon > 0$  and  $p \geq \frac{\ln n + \ln \ln n + \omega(1)}{n}$ , w.h.p.  $G = ([n], E) \sim \mathcal{G}(n, p)$  with degree sequence  $\mathbf{d}$  is  $\tilde{\mathbf{d}}(\frac{np}{100}, \varepsilon)$ -resilient with respect to  $\mathcal{HAM}$ .*

Note that Theorem 1 essentially covers the whole range of relevant values of  $p$ . Moreover, if  $p$  is such that there exists some  $\varepsilon > 0$  for which w.h.p.  $\delta(\mathcal{G}(n, p)) - 2 \leq \frac{np}{100}(\frac{1}{3} - \varepsilon)$ , then the result gives an exact local resilience for this range. On the one hand, every graph  $G$  satisfies  $r_\ell(G, \mathcal{HAM}) \leq \delta(G) - 1$ , as in order to leave the graph Hamiltonian after the deletion of edges, all degrees in the resulting graph must be at least 2. On the other hand, Theorem 1 guarantees that in this range of  $p$  if  $G \sim \mathcal{G}(n, p)$ , then w.h.p. removing any subgraph  $H \subseteq G$  of maximum degree  $\Delta(H) \leq \delta(G) - 2$  leaves a Hamiltonian graph.

THEOREM 2. *If  $\frac{\ln n + \ln \ln n + \omega(1)}{n} \leq p \leq \frac{1.02 \ln n}{n}$  and  $G \sim \mathcal{G}(n, p)$ , then w.h.p.  $r_\ell(G, \mathcal{HAM}) = \delta(G) - 1$ .*

*Proof.* As was previously mentioned, in light of Theorem 1, it is enough to prove that for the given range of  $p$  w.h.p.  $\delta(\mathcal{G}(n, p)) - 2 \leq \frac{np}{100}(\frac{1}{3} - \varepsilon)$  for some  $\varepsilon > 0$ . We use a basic result in the theory of random graphs due to Bollobás (see, e.g., [7, Chapter 3]), which asserts that if  $\binom{n-1}{k} p^k (1-p)^{n-1-k} = \omega(\frac{1}{n})$ , then w.h.p.  $\delta(\mathcal{G}(n, p)) \leq k$ . Note that if  $k$  is such that w.h.p.  $\delta(\mathcal{G}(n, p)) \leq k$ , then this is also true for every  $p' \leq p$  due to monotonicity. Setting  $k = \frac{\ln n}{300}$ , it suffices to prove that for  $p = \frac{1.02 \ln n}{n-k}$ , w.h.p.  $\delta(\mathcal{G}(n, p)) \leq k \leq \frac{np}{100}(\frac{1}{3} - \frac{1}{160}) + 2$ , which follows from

$$\begin{aligned}
\binom{n-1}{k} p^k (1-p)^{n-1-k} &\geq (1-o(1)) \left( \frac{e(n-k)p}{k} \right)^k \cdot e^{-(n-k)p} \\
&\geq (1-o(1)) n^{\frac{1+\ln 306}{300}} \cdot n^{-1.02} \\
&= \omega\left(\frac{1}{n}\right).
\end{aligned}$$

This completes the proof of the theorem.  $\square$

Comparing Theorem 2 with the upper bound from (1.1), we see that the local resilience drops from being equal to one less than the minimal degree in the beginning of the range to being equal to less than roughly half of it as  $p$  grows. This is due to the fact that when  $p$  is small enough the appearances of vertices whose degree is much smaller than the average degree create a bottleneck for  $\mathcal{HAM}$  (and many other graph properties).

Not all is lost when  $p$  becomes larger. Our main result implies that taking  $p$  to be large enough such that w.h.p. there are no vertices of degree less than  $\frac{np}{100}$ , the removal of almost one third of the incident edges at every vertex from a typical random graph will leave a graph that is Hamiltonian. Again, using straightforward calculations, which we omit, the following result is readily established.

**THEOREM 3.** *For every  $\varepsilon > 0$  there exists a constant  $C = C(\varepsilon) > 0$  such that if  $p \geq \frac{C \ln n}{n}$ , then w.h.p.*

$$r_\ell(\mathcal{G}(n, p), \mathcal{HAM}) \geq \frac{np}{3}(1 - \varepsilon).$$

Note that Theorem 3 in fact improves on the best known results of [5] for the local resilience of  $\mathcal{G}(n, p)$  with respect to  $\mathcal{HAM}$  in this range of  $p$ .

Lastly we show how from Theorem 2 one can obtain the existence of an optimal packing of Hamilton cycles in a typical random graph.

**THEOREM 4.** *For every  $p \leq \frac{1.02 \ln n}{n}$ , w.h.p.  $\mathcal{G}(n, p)$  has the property  $\mathcal{H}_\delta$ .*

*Proof.* First note that in light of the result of Frieze and Krivelevich [14], we may assume that  $p \geq (1 + o(1)) \frac{\ln n}{n}$ . We claim that in the range  $(1 + o(1)) \frac{\ln n}{n} \leq p \leq \frac{1.02 \ln n}{n}$  w.h.p. one can sequentially remove a Hamilton cycle from a typical graph  $G \sim \mathcal{G}(n, p)$  for  $\lfloor \delta(G)/2 \rfloor - 1$  rounds, leaving the graph Hamiltonian. Indeed, assume that the assertion of Theorem 2 holds; then the removal of  $0 \leq i \leq \lfloor \delta(G)/2 \rfloor - 1$  edge-disjoint Hamilton cycles from  $G$  is a removal of a  $2i$ -regular subgraph from  $G$ , and therefore Theorem 2 asserts that the resulting graph must be Hamiltonian. The removal of the last Hamilton cycle concludes the proof.  $\square$

Although the improvement of Theorem 4 relative to the previous best known bound on  $p$  of Frieze and Krivelevich [14] may seem quite insignificant, it should be stressed that the method used in [14] cannot be made to work for  $p = (1 + \varepsilon) \frac{\ln n}{n}$  for any fixed  $\varepsilon > 0$ , so the improvement presented here is of a more qualitative nature. Alas, the methods presented here also cannot be extended much further, as the degree sequence of the random graph becomes more balanced, causing  $\mathcal{D}_{\frac{np}{100}}$  to be the empty set. So, Conjecture 1.2 remains open.

**1.4. Organization.** The rest of the paper is organized as follows. We start with section 2, where we state all the needed preliminaries that are used throughout the proofs of our results. Section 3 is devoted to showing why a graph with pseudorandom properties is in fact  $\mathbf{k}$ -resilient to  $\mathcal{HAM}$ , and section 4 is dedicated to proving that all of the random-like properties needed in the previous section appear w.h.p. in  $\mathcal{G}(n, p)$ .

Section 5 is devoted to the proof of the main result of this paper—namely, Theorem 1—and we conclude the paper with some final remarks and open questions in section 6

**2. Preliminaries.** In this section we provide the necessary background information needed in the course of the proofs of the main results of this paper.

**2.1. Notation.** Given a graph  $G = (V, E)$ , the *neighborhood*  $N_G(U)$  of a subset  $U \subseteq V$  of vertices is the set of vertices defined by  $N_G(U) = \{v \notin U : \exists u \in U. \{v, u\} \in E\}$ , and the degree of a vertex  $v$  is  $d_G(v) = |N_G(\{v\})|$ . We denote by  $E_G(U)$  the set of edges of  $G$  that have both endpoints in  $U$ , and by  $e_G(U)$  we denote its cardinality. Similarly, for two disjoint subsets of vertices  $U$  and  $W$ ,  $E_G(U, W)$  denotes the set of edges with an endpoint in  $U$  and the other in  $W$ , and  $e_G(U, W)$  denotes its cardinality. We will sometimes refer to  $e_G(\{u\}, W)$  by  $d_G(u, W)$ . We use the usual notation of  $\Delta(G)$  and  $\delta(G)$  to denote the respective maximum and minimum degrees in  $G$ . We say that  $H$  is a *spanning subgraph* of  $G$  (or simply a subgraph, as all the subgraphs we consider will be spanning) and write  $H \subseteq G$  if the graph  $H = (V, F)$  has the same vertex set as  $G$  and its edge set satisfies  $F \subseteq E$ . We will denote by  $\ell(G)$  the length of a longest path in  $G$ .

Let  $R < n$  be positive integers and  $f: [R] \rightarrow \mathbb{R}^+$ . We say that a graph  $G = (V, E)$  on  $n$  vertices is an  $(R, f)$ -*expander* if every  $U \subseteq V$  of cardinality  $|U| \leq R$  satisfies  $|N_G(U)| \geq f(|U|) \cdot |U|$ . When  $f$  is a constant function equal to some  $\beta > 0$ , we say that  $G$  is an  $(R, \beta)$ -*expander*. When a function  $f: A \rightarrow \mathbb{R}^+$  satisfies  $f(a) \geq c$  for any  $a \in A$ , where  $c \geq 0$  is a constant, we simply write  $f \geq c$ .

*Remark 2.1.* Note that if  $G = (V, E)$  is an  $(R, f)$ -expander, then every  $H = (V, F)$  for  $F \supseteq E$  is also an  $(R, f)$ -expander.

The main research interest of this paper is the asymptotic behavior of some properties of graphs, when the graph is sampled from some probability measure  $\mathcal{G}$  over a set of graphs on the same vertex set  $[n]$ , and the number of vertices,  $n$ , grows to infinity. Therefore, from now on and throughout the rest of this work, when needed we will always assume  $n$  to be large enough. We use the usual asymptotic notation. For two functions of  $n$ ,  $f(n)$  and  $g(n)$ , we denote  $f = O(g)$  if there exists a constant  $C > 0$  such that  $f(n) \leq C \cdot g(n)$  for large enough values of  $n$ ,  $f = o(g)$  or  $f \ll g$  if  $f/g \rightarrow 0$  as  $n$  goes to infinity,  $f = \Omega(g)$  if  $g = O(f)$ , and  $f = \Theta(g)$  if both  $f = O(g)$  and  $g = O(f)$ .

Throughout the paper we will need to employ bounds on large deviations of random variables. We will mostly use the following well-known bound on the lower and the upper tails of the binomial distribution due to Chernoff (see, e.g., [1, Appendix A]).

**THEOREM 2.2** (Chernoff bounds). *Let  $X \sim \mathbf{Bin}(n, p)$ ; then for every  $\alpha > 0$*

1.  $\Pr[X > (1 + \alpha)np] < \exp(-np((1 + \alpha) \ln(1 + \alpha) - \alpha))$ ;
2.  $\Pr[X < (1 - \alpha)np] < \exp(-\frac{\alpha^2 np}{2})$ .

It will sometimes be more convenient to use the following bound on the upper tail of the binomial distribution.

**LEMMA 2.3.** *If  $X \sim \mathbf{Bin}(n, p)$  and  $k \gg np$ , then  $\Pr[X \geq k] \leq \binom{n}{k} p^k \leq (enp/k)^k$ .*

Lastly, we stress that throughout this paper we may omit floor and ceiling values when these are not crucial to avoid cumbersome exposition.

**3. From pseudorandomness to Hamiltonicity.** This section will provide all the necessary steps to show why a graph that possesses some random-like properties must be Hamiltonian. First, we show that a pseudorandom graph must contain an expander subgraph (Lemma 3.2). We will also insist this expander subgraph is quite sparse in relation to the original graph, where the sparsity requirement on the expander subgraph will play a major role in the proof of Theorem 1. As a second step we show how

the expansion of small sets of vertices is a combinatorial property that is quite resourceful toward proving Hamiltonicity, where the main tool used to achieve this is the celebrated Pósa rotation-extension technique (Lemma 3.6).

**3.1. From pseudorandomness to expansion.** We start by defining a family of graphs with desired pseudorandom properties.

DEFINITION 3.1. *We say that a graph  $G_1 = (V_1, E_1)$  on  $n_1$  vertices is  $(n_1, d, \beta)$ -quasi-random if it satisfies the following properties:*

- (P0)  $\delta(G_1) \geq \frac{d}{150}$ ;
- (P1) every  $U \subseteq V_1$  of cardinality  $|U| < n_1^{0.11} \ln n_1$  satisfies  $e_{G_1}(U) \leq d^{0.13}|U|$ ;
- (P2) every  $U \subseteq V_1$  of cardinality  $|U| < 12\beta n_1$  satisfies  $e_{G_1}(U) \leq 50\beta d|U|$ ;
- (P3) every two disjoint subsets  $U, Z \subseteq V_1$ , where  $|U| = \beta n_1$  and  $|Z| = (\frac{1}{3} - 27\beta)n_1$ , satisfy  $e_{G_1}(U, Z) \geq n_1 \ln \ln n_1$ .

The goal will be to find subgraphs of  $(n_1, d, \beta)$ -quasi-random graphs that satisfy certain expansion properties. To define this family of subgraphs we introduce the following “expansion” function. Given an integer  $n_1$ , we let  $f_\beta: [\beta n_1] \rightarrow \mathbb{R}^+$  denote the function defined by

- (Q1)  $f_\beta(t) = (\ln n_1)^{0.8}$  for every integer  $1 \leq t < n_1^{0.1}$ ;
- (Q2)  $f_\beta(t) = 11$  for every integer  $n_1^{0.1} \leq t < \beta n_1$ ;
- (Q3)  $f_\beta(\beta n_1) = \frac{2(1+39\beta)}{3\beta}$ .

The following lemma guarantees that every  $(n_1, d, \beta)$ -quasi-random graph contains a sparse  $(\beta n_1, f_\beta)$ -expander subgraph.

LEMMA 3.2. *For every constant  $0 < \beta < \frac{1}{15 \cdot 10^4}$ , there exists an integer  $n_0 = n_0(\beta) > 0$  such that if  $G_1$  is an  $(n_1, d, \beta)$ -quasi-random graph for some  $n_1 \geq n_0$  and  $d \geq \ln n_1$ , then  $G_1$  contains a  $(\beta n_1, f_\beta)$ -expander subgraph  $\Gamma$  satisfying  $e(\Gamma) \leq 10^6 \beta e(G_1)$ .*

*Proof.* Pick every edge of  $G_1 = (V_1, E_1)$  to be an edge of  $\Gamma$  with probability  $\gamma = 15 \cdot 10^4 \beta$  independently of all other choices. Our goal is to prove that  $\Gamma$  is an  $(\beta n_1, f_\beta)$ -expander with positive probability.

First, we analyze the minimum degree of  $\Gamma$ . The degree of every vertex  $v \in V_1$  in  $\Gamma$  is binomially distributed,  $d_\Gamma(v) \sim \mathbf{Bin}(d_{G_1}(v), \gamma)$ , with median at least  $\lfloor \gamma \delta(G_1) \rfloor$ . Therefore  $\Pr[d_\Gamma(v) \geq \lfloor \gamma \delta(G_1) \rfloor] \geq 1/2$ . We choose  $n_1$  to be large enough so that  $\lfloor \gamma \delta(G_1) \rfloor \geq \frac{\gamma d}{200}$ ; then by property P0 and the fact that the degrees in  $\Gamma$  of every two vertices are positively correlated, we have that

$$\begin{aligned} \Pr \left[ \delta(\Gamma) \geq \frac{\gamma d}{200} \right] &\geq \Pr[\forall v \in V_1. d_\Gamma(v) \geq \lfloor \gamma \delta(G_1) \rfloor] \\ &\geq \prod_{v \in V_1} \Pr[d_\Gamma(v) \geq \lfloor \gamma \delta(G_1) \rfloor] \\ &\geq 2^{-n_1}, \end{aligned}$$

using the FKG inequality (see, e.g., [1, Chapter 6]).

Under the assumption that  $\delta(\Gamma) \geq \frac{\gamma d}{200}$ , we show that  $\Gamma$  must satisfy Q1—namely that every  $U \subseteq V_1$  of cardinality  $|U| < n_1^{0.1}$  satisfies  $|N_\Gamma(U)| \geq |U|(\ln n_1)^{0.8}$ . Let  $U \subseteq V_1$  be some subset of cardinality  $|U| < n_1^{0.1}$ , and assume that  $|N_\Gamma(U)| < |U|(\ln n_1)^{0.8}$ . Denote by  $W = U \cup N_\Gamma(U)$ ; then by our assumption  $|W| < |U|(1 + (\ln n_1)^{0.8}) \leq |U|(\ln n_1)^{0.81} \ll n_1^{0.11} \ln n_1$ . Property P0 of  $G_1$  implies both  $e_\Gamma(U) \leq e_{G_1}(U) \leq d^{0.13}|U|$  and  $e_\Gamma(W) \leq e_{G_1}(W) \leq d^{0.13}|W| < |U|d^{0.13}(\ln n_1)^{0.81}$ . On the other hand,  $e_\Gamma(W) \geq \delta(\Gamma) \cdot |U| - e_\Gamma(U) \geq \frac{\gamma d}{200} \cdot |U| - e_{G_1}(U) > |U|d^{0.13}(\frac{\gamma(\ln n_1)^{0.87}}{200} - 1) > |U|d^{0.13}(\ln n_1)^{0.81}$ , which is a contradiction.

Property **Q2** of  $\Gamma$  will follow the exact same lines, again under the assumption that  $\delta(\Gamma) \geq \frac{\gamma d}{200}$ . Let  $U \subseteq V_1$  be some subset of cardinality  $n_1^{0.1} \leq |U| < \beta n_1$ , and assume that  $|N_\Gamma(U)| < 11|U|$ . Denote by  $W = U \cup N_\Gamma(U)$ ; then by our assumption  $|W| < 12|U| < 12\beta n_1$ . Property **P2** of  $G_1$  implies both  $e_\Gamma(U) \leq e_{G_1}(U) \leq 50\beta d|U|$  and  $e_\Gamma(W) \leq e_{G_1}(W) \leq 50\beta d|W| < 600\beta d|U|$ . On the other hand,  $e_\Gamma(W) \geq \delta(\Gamma) \cdot |U| - e_\Gamma(U) \geq \frac{\gamma d}{200} \cdot |U| - e_{G_1}(U) \geq 750\beta d|U| - 50\beta d|U| > 600\beta d|U|$ , which is a contradiction.

We proceed to show that  $\Gamma$  satisfies the bound on its number of edges and property **Q3** with probability greater than  $1 - 2^{-n_1}$ . This will imply that there exists a subgraph  $\Gamma$  as stated by the lemma with positive probability. Consider the number of edges in  $\Gamma$ . Clearly,  $e(\Gamma) \sim \mathbf{Bin}(e(G_1), \gamma)$  and by property **P0** we have  $e(G_1) \geq \frac{n_1 \delta(G_1)}{2} \geq \frac{n_1 d}{300} \geq \frac{n_1 \ln n_1}{300}$ . Using Lemma 2.3, it follows that

$$\Pr[e(\Gamma) > 10^6 \beta e(G_1)] \leq \left( \frac{e \cdot e(G_1) \cdot \gamma}{10^6 \beta e(G_1)} \right)^{10^6 \beta e(G_1)} = o(2^{-n_1}).$$

Fix a subset  $U \subseteq V$  of cardinality  $|U| = \beta n_1$ . Assume that for every subset of vertices  $Z \subseteq V_1$  of cardinality  $|Z| = (\frac{1}{3} - 27\beta)n_1$  and disjoint from  $U$ , we have that  $e_\Gamma(U, Z) > 0$ ; then  $|N_\Gamma(U)| \geq (\frac{2}{3} + 26\beta)n_1 = \frac{2(1+39\beta)}{3\beta}|U|$ . Therefore, in order to prove that  $\Gamma$  satisfies **Q3** with the required probability, we prove that  $e_\Gamma(U, Z) > 0$  for every pair of disjoint subsets of vertices  $U, Z \subseteq V_1$  of cardinality  $|U| = \beta n_1$ , and  $|Z| = (\frac{1}{3} - 27\beta)n_1$  with probability at least  $1 - 2^{-n_1}$ . By property **P3** it follows that  $e_{G_1}(U, Z) > n_1 \ln \ln n_1$ , and therefore  $e_\Gamma(U, Z)$  is stochastically dominated by  $\mathbf{Bin}(n_1 \ln \ln n_1, \gamma)$ . This implies that

$$\Pr[e_\Gamma(U, Z) = 0] < (1 - \gamma)^{n_1 \ln \ln n_1} < \exp(-\gamma n_1 \ln \ln n_1).$$

Now, by going over all possible pairs of subsets  $U$  and  $Z$ , as there are at most  $4^{n_1}$  of those, we have that the probability that  $\Gamma$  does not satisfy property **Q3** is at most  $\exp(-\gamma n_1 \ln \ln n_1) \cdot 4^{n_1} = o(2^{-n_1})$ , which completes the proof.  $\square$

**3.2. From expansion to Hamiltonicity: Pósa’s rotation-extension technique.** In order to describe the relevant connection between Hamiltonicity and expanders, we require the notion of *boosters*.

**DEFINITION 3.3.** *For every graph  $G$  we say that a nonedge  $\{u, v\} \notin E(G)$  is a booster with respect to  $G$  if  $G + \{u, v\}$  is Hamiltonian or  $\ell(G + \{u, v\}) > \ell(G)$ . For any vertex  $v \in V$  we denote by*

$$(3.1) \quad B_G(v) = \{w \notin N_G(v) \cup \{v\} : \{v, w\} \text{ is a booster}\}.$$

The following simple lemma describes a sufficient condition for a graph  $G$  to be such that deleting the edges of some subgraph  $H \subseteq G$  leaves a Hamiltonian graph. Note that a crucial point in the lemma is the existence of another graph that acts as a “backbone” by providing enough boosters.

**LEMMA 3.4.** *Let  $G = (V, E)$  be a graph and let  $H \subseteq G$  be some subgraph of it. If there exists a subgraph  $\Gamma \subseteq G - H$  such that for every  $E' \subseteq E(G - H - \Gamma)$  of cardinality  $|E'| \leq n$  there exists a vertex  $v \in V$  (which may depend on  $E'$ ) satisfying  $|N_{G-(\Gamma+E')}(v) \cap B_{\Gamma+E'}(v)| > d_H(v)$ , then  $G - H$  is Hamiltonian.*

*Proof.* Assume  $\Gamma \subseteq G - H$  is as required by the lemma; then we prove there exists an edge set  $F \subseteq E(G - H - \Gamma)$  such that the graph  $\Gamma + F$  must be Hamiltonian, which in turn implies that  $G - H$  is Hamiltonian. Start with  $F_0 = \emptyset$ . Assume that  $F_i$  is a



subset of  $0 \leq i \leq n$  edges of  $E(G - H - \Gamma)$ . If the graph  $\Gamma_i = \Gamma + F_i \subseteq G - H$  is Hamiltonian, we are done. Otherwise, by the assumption of the lemma, there exists a vertex  $v_i \in V$  such that  $|N_{G-\Gamma_i}(v_i) \cap B_{\Gamma_i}(v_i)| > d_H(v_i)$ , and hence there exists at least one neighbor of  $v_i$  in  $G$ , which we denote by  $w_i$ , such that the pair  $\{v_i, w_i\}$  is still an edge in  $G - H$  and is a booster with respect to  $\Gamma_i$  (hence not an edge of  $\Gamma_i$ ). It follows that the graph  $\Gamma_i + \{v_i, w_i\}$  is Hamiltonian or  $\ell(\Gamma_i + \{v_i, w_i\}) > \ell(\Gamma_i)$ . Finally, set  $F_{i+1} = F_i \cup \{\{v_i, w_i\}\}$ . Note that there must exist an integer  $i_0 \leq n$  such that  $\Gamma_{i_0}$  is Hamiltonian, as the length of a path cannot exceed  $n - 1$ .  $\square$

We now describe and apply a crucial technical tool, originally developed by Pósa [22], which lies in the foundation of many Hamiltonicity results of random and pseudorandom graphs. This technique, which has come to be known as Pósa’s rotation-extension, relies on the following basic operation on a longest path in a graph. We use the following two definitions.

**DEFINITION 3.5.** *Let  $G = (V, E)$  be a graph, and let  $P = (v_0, v_1, \dots, v_\ell)$  be a longest path in  $G$ . If  $\{v_i, v_\ell\} \in E$  for some  $0 \leq i \leq \ell - 2$ , then an elementary rotation of  $P$  along  $\{v_i, v_\ell\}$  is the construction of a new longest path  $P' = P - \{v_i, v_{i+1}\} + \{v_i, v_\ell\} = (v_0, v_1, \dots, v_i, v_\ell, v_{\ell-1}, \dots, v_{i+1})$ . We say that the edge  $\{v_i, v_{i+1}\}$  is broken by this rotation.*

Given a fixed  $\beta > 0$  and integer  $n$ , we define a family  $\mathcal{L} = \mathcal{L}(n, \beta)$  of graphs with vertex set  $V = [n]$  to which these elementary rotations will be applied. The graphs in  $\mathcal{L}$  all entail some pseudorandom properties. Now, a graph  $G = (V, E) \in \mathcal{L}(n, \beta)$  if its vertex set can be partitioned as  $V = V_1 \cup D$  such that

- (L1)**  $|D| \leq n^{0.09}$ ;
- (L2)**  $d(u, V_1) \geq 2$  for every  $u \in D$ ;
- (L3)** there is no path in  $G$  of length at most  $\frac{2 \ln n}{3 \ln \ln n}$  with both (possibly identical) endpoints in  $D$ ;
- (L4)**  $G_1 = G[V_1]$  is a  $(\beta n_1, f_\beta)$ -expander on  $n_1 = |V_1|$  vertices.

Using elementary rotations we proceed to show that any  $G \in \mathcal{L}(n, \beta)$  must be Hamiltonian or that the subset of vertices with “large”  $B_G(v)$  must also be large. Our proof uses ideas similar to those found in [5].

**LEMMA 3.6.** *For every fixed  $\beta > 0$  there exists an integer  $n_0(\beta) = n_0 > 0$  such that if  $n \geq n_0$ , then every graph  $G = (V, E) \in \mathcal{L}(n, \beta)$  is Hamiltonian or must satisfy  $|\{v \in V_1 : |B_G(v)| \geq n/3 + \beta n\}| \geq n/3 + \beta n$ .*

**Remark 3.7.** We remark that property **L4** implicitly implies an upper bound on  $\beta$  for  $\mathcal{L}(n, \beta)$  to be nonempty. Indeed  $\frac{2(1+39\beta)}{3} \leq 1 - \beta$  must hold, which implies  $\beta \leq \frac{1}{81}$ .

*Proof.* First, we claim that under the assumptions on  $G$ , it must be connected. In order to prove this, we begin by showing that  $G_1$  is connected. Indeed, assume otherwise and let  $W \subseteq V_1$  be a connected component of cardinality  $|W| \leq n_1/2$ . Properties **Q1**, **Q2**, and **Q3** imply that every subset of at most  $\beta n_1$  vertices has a nonempty neighbor set; hence we can further assume that  $|W| > \beta n_1$ . Let  $W' \subseteq W$  be of cardinality  $\beta n_1$ ; then  $|N_G(W')| > 2n_1/3 > n_1/2$  by property **Q3** and hence cannot be contained in  $W$ , a contradiction. Property **L2** guarantees that every vertex in  $D$  is connected to some vertex in  $V_1$ , so adding the vertices of  $D$  to the graph  $G_1$  leaves the graph connected.

Take a longest path  $P = (v_0, \dots, v_\ell)$  in  $G$ . By the assumption on  $G$  we can clearly assume that  $\ell \geq 3$ . Since  $P$  is a longest path,  $N_G(v_0) \cup N_G(v_\ell) \subseteq P$ . We first claim that we can choose such a path  $P$  with both of its endpoints in  $V_1$ . So, assume that at least one of the endpoints of  $P$ ,  $v_0$ , and  $v_\ell$  is in  $D$ . If  $v_0$  and  $v_\ell$  are neighbors, then  $G$  must contain a cycle of length  $\ell(G)$ . This implies that  $G$  is Hamiltonian, as otherwise  $\ell(G) < n$ , and since  $G$  is connected, there is an edge emitting out of this cycle, creating

a path of length  $\ell(G) + 1$  in  $G$ , which is a contradiction. We can thus assume  $v_\ell$  and  $v_0$  are not neighbors. Assume without loss of generality (w.l.o.g.) that  $v_\ell \in D$ ; then by property **L2** the vertex  $v_\ell$  must have a neighbor other than  $v_{\ell-1}$ . As all the neighbors of  $v_\ell$  must lie on  $P$ , we denote this neighbor by  $v_i$ , where  $1 \leq i \leq \ell - 2$ . Performing an elementary rotation of  $P$  along  $\{v_i, v_\ell\}$  results in the path  $P' = (v_0, v_1, \dots, v_i, v_\ell, v_{\ell-1}, \dots, v_{i+1})$  with  $v_{i+1}$  as an endpoint. By Property **L2** of  $D$ , the vertex  $v_{i+1} \notin D$ , since it is at distance at most 2 from  $v_\ell \in D$ . Denote the resulting longest path which terminates at  $v_{i+1}$  by  $P'$ . If  $v_0 \in D$  as well, then we can also perform a rotation on  $P'$  keeping  $v_{i+1}$  fixed. Note that in order to do so we can assume  $v_0$  and  $v_{i+1}$  are non-neighbors, as otherwise our graph is Hamiltonian as previously claimed. We can thus assume that our initial longest path  $P$  has indeed two endpoints in  $V_1$ .

We set with foresight  $r_0 = \frac{\ln n_1}{7 \ln \ln n_1}$ , and  $r$  to be the minimal integer for which  $\beta n_1 \leq 2^{r-r_0} \cdot T^{r_0} < 2\beta n_1$ , where  $T = \ln^{0.7} n_1$ . For every  $0 \leq i \leq r$  we will find a set  $S_i \subseteq V_1$  that is composed of (not necessarily all) endpoints of longest paths in  $G$  obtained by performing a series of  $i$  elementary rotations starting from  $P$  while keeping the endpoint  $v_0$  fixed such that for every  $j < i$  in the  $j$ th rotation the non- $v_0$  endpoint lies in  $S_j \subseteq V_1$ . We construct this sequence of sets  $\{S_i\}_{i=0}^r$  as follows. We start with  $S_0 = \{v_\ell\}$  (which, by our assumption, is not in  $D$ ). For every  $i > 0$ , let  $U_i = \{v_k \in N_G(S_i) : v_{k-1}, v_k, v_{k+1} \notin \bigcup_{j=0}^i S_j\}$ . Let  $v_k \in U_i$ , and  $x \in S_i$  be such that  $\{x, v_k\} \in E$ , and denote by  $Q$  the longest path from  $v_0$  to  $x$  obtained from  $P$  by  $i$  elementary rotations leaving  $v_0$  fixed. By the definition of  $U_i$ , none of the vertices  $v_{k-1}, v_k, v_{k+1}$  is an endpoint of one of the sequence of longest paths starting from  $P$  and yielding  $Q$ ; hence both edges  $\{v_{k-1}, v_k\}$  and  $\{v_k, v_{k+1}\}$  were not broken and are therefore present in  $Q$ . Rotating  $Q$  along the edge  $\{x, v_k\}$  will make one of the vertices  $\{v_{k-1}, v_{k+1}\}$  an endpoint in the resulting path. Assume w.l.o.g. that it is  $v_{k-1}$ , and hence add it to the set  $S'_{i+1}$ . Note that the vertex  $v_{k-1}$  can also be added to the set  $S'_{i+1}$  if the vertex  $v_{k-2}$  is in  $U_i$ ; therefore

$$(3.2) \quad |S'_{i+1}| \geq \frac{1}{2} |U_i| \geq \frac{1}{2} \left( |N_G(S_i)| - 3 \sum_{j=0}^i |S_j| \right).$$

Having defined the set  $S'_{i+1}$  for every  $i$ , we proceed to demonstrate how the set  $S_{i+1}$  is chosen as an appropriate subset of  $S'_{i+1}$ . This is done according to the value of  $i$ . First, for  $0 \leq i < r_0$  the set  $S_i$  is chosen to be of cardinality  $T^i$ , where we prove we can do so inductively. Clearly, our assumption on  $S_0$  proves that the base of the induction holds. Assuming we can choose the subsets  $S_j \subseteq S'_j$  for all  $j \leq i < r_0 - 1$ , we prove we can choose a subset  $S_{i+1} \subseteq S'_{i+1} \cap V_1$ . We have by our induction hypothesis that  $|S_i| = T^i < T^{r_0} \leq n_1^{0.1}$ ; hence for every  $i < r_0$  property **Q1** implies  $|N_G(S_i)| \geq T^{i+1.1}$ , and therefore by (3.2)

$$|S'_{i+1}| \geq \frac{1}{2} \left( T^{i+1.1} - 3 \left( \frac{T^{i+1} - 1}{T - 1} \right) \right) > T^{i+1} + 1.$$

Every vertex in  $S'_{i+1}$  is at distance at most  $2i$  from  $v_\ell$ ; therefore, every two vertices in  $S'_{i+1}$  are at distance at most  $4i < 4r_0 \leq \frac{4 \ln n_1}{7 \ln \ln n_1} < \frac{2 \ln n}{3 \ln \ln n}$  from each other. Property **L2** implies there can be at most one vertex from  $D$  in this set. Removing this vertex if it exists, we can set  $S_{i+1}$  to be any subset of  $S'_{i+1}$  of  $T^{i+1}$  vertices from  $V_1$ . In particular  $|S_{r_0}| = T^{r_0} \gg n^{0.09} \geq |D|$ . Next, for  $r_0 \leq i < r$  we will construct the sets  $S_i$  to be of cardinality  $R \cdot 2^{i-r_0} \gg |D|$ , where  $R = T^{r_0}$ . Properties **Q1** and **Q2** imply  $|N_G(S_i)| \geq 11 \cdot |S_i| \geq 10|S_i| + 2|D|$ , and therefore by (3.2),

$$\begin{aligned} |S'_{i+1}| &\geq \frac{1}{2} \left( 10 \cdot R \cdot 2^{i-r_0} + 2|D| - 3 \left( \sum_{j=0}^{r_0-1} T^i + \sum_{j=r_0}^i R \cdot 2^{j-r_0} \right) \right) \\ &\geq |D| + 5R \cdot 2^{i-r_0} - \frac{3}{2} (T^{r_0} + R(2^{i+1-r_0} - 1)) \\ &= |D| + R \cdot 2^{i-r_0+1}. \end{aligned}$$

We can therefore set  $S_{i+1}$  to be any subset of  $S'_{i+1} \cap V_1$  of cardinality  $R \cdot 2^{i-r_0}$ . For the sake of convenience, we replace  $S_r$  by some arbitrary subset of it of cardinality  $\beta n_1$  and denote this new set by  $S_r$ . Finally, we construct similarly the set  $S'_{r+1}$ . By property **Q3** we have  $|N_G(S_r)| \geq (1 + 39\beta)2n_1/3$ . Just as in the previous cases, by (3.2),

$$\begin{aligned} |S'_{r+1}| &\geq \frac{1}{2} \left( \frac{2n_1}{3} (1 + 39\beta) - 3 \left( \sum_{j=0}^{r_0-1} T^i + \sum_{j=r_0}^r R \cdot 2^{j-r_0} \right) \right) \\ &\geq n_1 \left( \frac{1}{3} + 13\beta \right) - \frac{3}{2} T^{r_0} - \frac{3}{2} R \cdot 2^{r-r_0+1} \\ &\geq n_1 \left( \frac{1}{3} + 13\beta \right) - o(n_1) - 6\beta n_1 \\ &> \frac{n_1}{3} + 6\beta n_1 > \frac{n}{3} + \beta n. \end{aligned}$$

Assume  $S'_{r+1} \cap N_G(v_0) \neq \emptyset$ ; then clearly  $G$  must contain a cycle of length  $\ell(G)$ . Assume that  $G$  is not Hamiltonian, i.e., that  $\ell(G) < n$ . As  $G$  is connected, there is an edge emitting out of this cycle, creating a path of length  $\ell(G) + 1$  in  $G$ , which is a contradiction; hence  $G$  must be Hamiltonian in this case. On the other hand, if  $S'_{r+1} \cap N_G(v_0) = \emptyset$ , then  $S'_{r+1} \subseteq B_G(v_0)$ . Now take any endpoint  $u_0$  in  $S'_{r+1}$  and take a longest path  $P'$  starting from  $u_0$  (which must exist since all vertices of  $S'_{r+1}$  are endpoints of longest paths starting in  $v_0$ ) and repeat the same argument, while rotating  $P'$  and keeping  $u_0$  fixed. This way we obtain the desired set  $|B_G(u_0)|$  of  $n/3 + \beta n$  endpoints for every  $u_0 \in S'_{r+1}$ , thus completing the proof.  $\square$

**4. Structural properties of  $\mathcal{G}(n, p)$ .** We start with a very simple claim regarding the number of edges in the binomial random graph model  $\mathcal{G}(n, p)$ .

*Claim 4.1.* For every  $p \geq \frac{\ln n}{n}$ , w.h.p.  $\frac{n^2 p}{4} \leq e(\mathcal{G}(n, p)) \leq n^2 p$ .

*Proof.* This is a simple application of Theorem 2.2. Clearly,  $e(G) \sim \text{Bin}(\binom{n}{2}, p)$ ; then  $\Pr[e(G) > n^2 p] < \Pr[e(G) \geq 2 \frac{n}{2} \cdot p] \leq \exp(-\frac{(n-1)\ln n}{6}) = o(1)$ . Similarly,  $\Pr[e(G) < \frac{n^2 p}{4}] \leq \Pr[e(G) < 0.51 \frac{n}{2} p] \leq \exp(\frac{0.49^2(n-1)\ln n}{4}) = o(1)$ .  $\square$

The following is (a special case of) a well-known property of the binomial random graph model. Specifically, this addresses the problem of finding the threshold probability for a random graph distributed according to  $\mathcal{G}(n, p)$  to have minimum degree at least 2 (see, e.g., [7]).

**THEOREM 4.2.** *Let  $h(n) = \omega(1)$  be any function that grows arbitrarily slowly with  $n$ ; then*

1. *if  $p \geq \frac{\ln n + \ln \ln n + h(n)}{n}$ , then w.h.p.  $\delta(\mathcal{G}(n, p)) \geq 2$ ;*
2. *if  $p \leq \frac{\ln n + \ln \ln n - h(n)}{n}$ , then w.h.p.  $\delta(\mathcal{G}(n, p)) < 2$ .*

Recall that given a graph  $G$  we defined the set  $\mathcal{D}_t(G) = \{v \in V : d_G(v) < t\}$ . We prove some structural properties of the set  $\mathcal{D}_t(G)$ , where  $G$  is sampled from the random graph model  $\mathcal{G}(n, p)$ .

*Claim 4.3.* For every  $p \geq \frac{\ln n}{n}$  and integer  $t \leq \frac{np}{100}$ , w.h.p.  $|\mathcal{D}_t(\mathcal{G}(n, p))| \leq n^{0.09}$ .

*Proof.* Let  $G \sim \mathcal{G}(n, p)$ ; then setting  $t_0 = \frac{np}{100}$ , we can bound the probability that a vertex is in  $\mathcal{D}_t(G)$  as follows.

$$\begin{aligned} \Pr[d_G(v) < t] &\leq \Pr[\mathbf{Bin}(n-1, p) < t_0] \\ &\leq \sum_{i=0}^{t_0-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} \\ &\leq t_0 \cdot \binom{n-1}{t_0} p^{t_0} (1-p)^{n-1-t_0} \\ &\leq t_0 \cdot \left(\frac{e(n-1)p}{t_0}\right)^{t_0} e^{-p(n-1-t_0)} \\ &\leq \exp\left(-np + p + \frac{np^2}{100} + \frac{np}{100}(1 + \ln 100) + \ln \frac{np}{100}\right) \\ &< e^{-0.92np} \\ &\leq n^{-0.92}. \end{aligned}$$

This implies that  $\mathbf{E}[\mathcal{D}_t(\mathcal{G}(n, p))] \leq n^{0.08}$ , and the claim follows from Markov's inequality.  $\square$

*Claim 4.4.* For every  $p \geq \frac{\ln n}{n}$  and integer  $t \leq \frac{np}{100}$ , w.h.p  $G \sim \mathcal{G}(n, p)$  does not contain a nonempty path of length at most  $\frac{2 \ln n}{3 \ln \ln n}$  such that both of its (possibly identical) endpoints lie in  $\mathcal{D}_t(G)$ .

*Proof.* Setting  $t_0 = \frac{np}{100}$ , we prove the claim for two distinct endpoints in  $\mathcal{D}_{t_0}(G)$ , and for paths of length  $r$ , where  $2 \leq r \leq \frac{2 \ln n}{3 \ln \ln n}$ . The other cases (i.e., identical endpoints or  $r = 1$ ) are similar and a little simpler. Fix two vertices  $u, w \in V(G)$  and let  $P = (u = v_0, \dots, v_r = w)$  be a sequence of vertices of  $V(G)$ , where  $2 \leq r \leq \frac{2 \ln n}{3 \ln \ln n}$ . Denote by  $\mathcal{A}_P$  the event  $\{v_i, v_{i+1}\} \in E(G)$  for every  $0 \leq i \leq r-1$ , and by  $\mathcal{B}_{u,w}$  the event that both  $u$  and  $w$  are elements of  $\mathcal{D}_{t_0}(G)$ . Clearly,  $\Pr[\mathcal{A}_P] = p^r$ ; then

$$\Pr[\mathcal{B}_{u,w} \wedge \mathcal{A}_P] = p^r \cdot \Pr[\mathcal{B}_{u,w} | \mathcal{A}_P].$$

Let  $X_{u,w}$  denote the random variable that counts the number of edges in  $G$  incident with  $u$  or  $w$ , disregarding the pairs  $\{u, v_1\}$ ,  $\{v_{r-1}, w\}$ , and  $\{u, w\}$ . We can therefore bound  $\Pr[\mathcal{B}_{u,w} | \mathcal{A}_P] \leq \Pr[X_{u,w} < 2t_0 - 2]$  and as  $X_{u,w} \sim \mathbf{Bin}(2n-6, p)$ , it follows that

$$\begin{aligned} \Pr[X_{u,w} < 2t_0 - 2] &\leq \sum_{i=0}^{2t_0-2} \binom{2n-6}{i} p^i (1-p)^{2n-6-i} \\ &\leq \frac{np}{50} \binom{2n-6}{2t_0-2} p^{2t_0-2} (1-p)^{2n-4-2t_0} \\ &\leq \frac{np}{50} \cdot \left(\frac{e(2n-6)p}{2t_0-2}\right)^{2t_0} e^{-p(2n-4-2t_0)} \\ &\leq \exp\left(-2np + p\left(4 + \frac{np}{50}\right) + \frac{np}{50}(2 + \ln 100) + \ln \frac{np}{50}\right) \\ &< e^{-1.8np}. \end{aligned}$$

Fixing the two endpoints  $u, w$ , the number of such sequences is at most  $(r-1)! \binom{n}{r-1} \leq n^{r-1}$ . Applying a union bound argument over all such pairs of vertices

and possible sequences connecting them, we conclude that the probability that there exists a path in  $G$  of length  $r \leq \frac{2 \ln n}{3 \ln \ln n}$ , connecting two vertices of  $\mathcal{D}_t(G)$ , is at most

$$\begin{aligned} \sum_{r=1}^{\frac{2 \ln n}{3 \ln \ln n}} \binom{n}{2} \cdot n^{r-1} \cdot p^r \cdot e^{-1.8np} &\leq \sum_{r=1}^{\frac{2 \ln n}{3 \ln \ln n}} \frac{n^{r+1}}{2} \cdot \frac{\ln^r n}{n^r} \cdot n^{-1.8} \\ &\leq \frac{\ln n}{3 \ln \ln n} \cdot n^{-0.8} \cdot (\ln n)^{\frac{2 \ln n}{3 \ln \ln n}} \\ &= o(1), \end{aligned}$$

where the first inequality follows by noting that the expression on the left-hand side decreases as  $p$  grows; hence we can replace  $p$  by  $\frac{\ln n}{n}$ . This completes the proof of the claim.  $\square$

Given a graph  $G$  and some  $t > 0$  (which may depend on  $n$  and  $p$ ), we denote by  $G_1(t) = G_1 = (V_1, E_1)$  the graph  $G_1 = G[V \setminus \mathcal{D}_t(G)]$  and by  $n_1$  its number of vertices; i.e.,  $|V_1| = |V \setminus \mathcal{D}_t(G)| = n_1$ . The following lemma, which contains the main technical result of this section, implies that if  $p \geq \frac{\ln n}{n}$ , then the removal of the vertices of low degree from a typical graph  $G$  sampled from  $\mathcal{G}(n, p)$  leaves a graph  $G_1$  which is robust in the following sense: The deletion of almost a third of the edges at each vertex of  $G_1$  leaves a graph with some strong pseudorandom properties.

**LEMMA 4.5.** *For every fixed  $\varepsilon > 0$ , there exists a small enough constant  $\beta_0 = \beta_0(\varepsilon) > 0$  such that for every  $0 < \beta \leq \beta_0$ ,  $p \geq \frac{\ln n}{n}$ , and  $t = \frac{np}{100}$ , if  $G \sim \mathcal{G}(n, p)$ , then w.h.p. for any subgraph  $H_1 \subseteq G_1(t) = G_1$  such that  $\mathbf{d}_{H_1} \leq (\frac{1}{3} - \varepsilon)\mathbf{d}_G$ , the graph  $G_1 - H_1$  is  $(n_1, np, \beta)$ -quasi-random.*

*Proof.* First, by Claim 4.4 we can assume that in  $G$  every vertex of  $V_1$  has at most one neighbor in  $\mathcal{D}_t(G)$ ; hence for every  $v \in V_1$  we have  $d_{G_1}(v) \geq d_G(v) - 1$ , and therefore  $\delta(G_1) \geq t - 1$ . It follows that  $\delta(G_1 - H_1) > (\frac{2}{3} + \varepsilon)\delta(G_1) > \frac{2t}{3} = \frac{np}{150}$ . Second, using Claim 4.3 we can, and will, assume that  $n_1 \geq n - n^{0.09} = n(1 - o(1))$ . The rest of the properties, and hence the proof of the lemma, will be a simple consequence from the following series of claims. We stress that throughout we will not compute  $\beta$  explicitly, but we will assume it is small enough as a function of  $\varepsilon$  for the arguments to go through.

**Claim 4.6.** W.h.p. for every  $U \subseteq V$  of cardinality  $|U| \leq n^{0.11} \ln n$ , it holds that  $e_{G_1 - H_1}(U) \leq e_G(U) \leq (np)^{3/25} |U|$ .

*Proof.* Fixing such a subset of vertices  $U$  of cardinality  $u \leq n^{0.11} \ln n$ , we have that  $e_G(U) \sim \mathbf{Bin}(\binom{u}{2}, p)$  and since  $e(\binom{u}{2}, p) \leq u(np)^{3/25} \cdot n^{-1/200}$  we have by Lemma 2.3 that

$$\begin{aligned} \Pr[e_G(U) > (np)^{3/25} u] &< \left( \frac{e(\binom{u}{2}, p)}{(np)^{3/25} u} \right)^{(np)^{3/25} u} \leq \exp\left(-\frac{(np)^{3/25} \cdot u \cdot \ln n}{200}\right) \\ &\leq \exp\left(-\frac{(\ln n)^{1.12} \cdot u}{200}\right). \end{aligned}$$

To upper bound the probability of the existence of a subset of vertices for which the assertion of the claim does not hold, we apply the union bound over all possible sets  $U$  of cardinality  $u \leq n^{0.11} \ln n$ :

$$\begin{aligned} \sum_{u=1}^{n^{0.11} \ln n} \binom{n}{u} \exp\left(-\frac{(\ln n)^{1.12} \cdot u}{200}\right) &\leq \sum_{u=1}^{n^{0.11} \ln n} \exp\left(u \cdot \left(\ln \frac{en}{u} - \frac{(\ln n)^{1.12}}{200}\right)\right) \\ &\leq n^{0.11} \ln n \cdot \exp(-(\ln n)^{1.1}) = o(1). \quad \square \end{aligned}$$

*Claim 4.7.* W.h.p. every  $U \subseteq V$  of cardinality  $|U| \leq 12\beta n$  satisfies  $e_{G_1-H_1}(U) \leq e_G(U) \leq 50\beta np|U|$ .

*Proof.* Very similarly to the proof of Claim 4.6, we fix a subset of vertices  $U$  of cardinality  $1 \leq u \leq 12\beta n$ ; then noting that  $e_G(U) \sim \mathbf{Bin}(\binom{u}{2}, p)$ , and since  $\binom{u}{2}p \leq 50\beta npu$ , we have by Lemma 2.3 that

$$\Pr[e_G(U) > 50\beta npu] < \left(\frac{e^u p}{50\beta npu}\right)^{50\beta npu} \leq \exp\left(-50\beta npu \ln \frac{100\beta n}{eu}\right).$$

To upper bound the probability of the existence of a subset of vertices for which the assertion of the claim does not hold, we apply the union bound over all possible sets  $U$  of cardinality  $1 \leq u \leq 12\beta n$ :

$$\begin{aligned} & \sum_{u=1}^{12\beta n} \binom{n}{u} \exp\left(-50\beta npu \ln \frac{100\beta n}{eu}\right) \\ & \leq \sum_{u=1}^{12\beta n} \exp\left(u \cdot \left(\ln \frac{en}{u} - 50\beta np \ln \frac{100\beta n}{eu}\right)\right) \\ & \leq \sum_{u=1}^{12\beta n} \exp\left(u \cdot \ln \frac{en}{u} \cdot (1 - 50\beta \ln n) - 50\beta \ln n \ln \beta\right) \\ & = o(1). \quad \square \end{aligned}$$

*Claim 4.8.* W.h.p. every two disjoint subsets  $U, Z \subseteq V_1$ , where  $|U| = \beta n_1$  and  $|Z| = n_1(\frac{1}{3} - 27\beta)$  satisfy  $e_{G_1-H_1}(U, Z) \geq n_1 \ln \ln n_1$ .

*Proof.* Fix two such disjoint subsets of vertices  $U, Z \subseteq V_1$  of the required cardinalities and note that

$$\begin{aligned} e_{G_1-H_1}(U, Z) &= e_{G_1}(U, Z) - e_{H_1}(U, Z) \\ &\geq e_G(U, Z) - \sum_{v \in U} d_{H_1}(v) \\ &\geq e_G(U, Z) - \left(\frac{1}{3} - \varepsilon\right) \sum_{v \in U} d_G(v). \end{aligned}$$

As  $e_G(U, Z) \sim \mathbf{Bin}(|U| \cdot |Z|, p)$ , we can apply Theorem 2.2 item 2 and get that

$$\Pr[e_G(U, Z) < \left(\frac{1}{3} - 28\beta\right) n_1 \cdot |U| \cdot p] \leq e^{-\Theta(n_1^2 p)} = o(4^{-n_1}),$$

where the hidden constants in the exponent above are functions of  $\beta$  alone. Next, we prove that the random variable  $X(G) = \sum_{v \in U} d_G(v)$  is very likely not to deviate much from its expectation. To achieve this we resort to the Azuma–Hoeffding inequality for martingales of bounded variance (see, e.g., [1, Theorem 7.4.3]). Note that for any two graphs on the vertex set  $V$  that differ by a single edge, their value of  $X$  can change by at most 2, and  $\mathbf{E}[X(G)] = |U|(n-1)p < (1+o(1))|U|n_1p$ . In order to divulge the value of  $X$  one only needs to expose the pairs of vertices that have at least one endpoint in  $U$ . This implies that the total variance of the martingale is at most  $\beta n_1^2 p(1-p)$ , and hence

$$\Pr[X(G) > (1 + \beta)|U|n_1p] \leq e^{-\Theta(n_1^2 p)} = o(4^{-n_1}),$$

where the hidden constants in the exponent above are, again, functions of  $\beta$  alone. Recalling that  $\beta$  can be made small enough with respect to  $\varepsilon$  and that  $|U|n_1p \gg n_1 \ln \ln n_1$ , we have that

$$\left(\frac{1}{3} - 28\beta\right) > \left(\frac{1}{3} - \varepsilon\right)(1 + \beta) + \frac{n_1 \ln \ln n_1}{|U|n_1p},$$

and hence  $\Pr[e_{G_1-H_1}(U, Z) < n_1 \ln \ln n_1] = o(4^{-n_1})$ . By applying the union bound over all pairs of subsets of vertices  $U$  and  $Z$ , the proof of the claim is completed.  $\square$

Recalling Definition 3.1 completes the proof of the lemma.  $\square$

We can now present the main result of this section which based on the above is readily established. In what follows let  $G = (V, E)$ , with vertex set  $V = [n]$ , be sampled from  $\mathcal{G}(n, p)$  and let  $t = \frac{np}{100}$ . We recall that for every  $\varepsilon > 0$  if  $\mathbf{d} = (d_1, \dots, d_n)$  is the degree sequence of  $G$ , then  $\tilde{\mathbf{d}}(t, \varepsilon) = (\tilde{d}_1, \dots, \tilde{d}_n)$  is the sequence defined by  $\tilde{d}_v = d_v - 2$  for every  $v \in \mathcal{D}_t(G)$  and  $\tilde{d}_v = d_v(1/3 - \varepsilon)$  for every  $v \in V_1 = V \setminus \mathcal{D}_t(G)$ .

**COROLLARY 4.9.** *For every fixed  $\varepsilon > 0$ , there exists a small enough constant  $\beta_0 = \beta_0(\varepsilon) > 0$  such that for every  $0 < \beta \leq \beta_0$  and  $p \geq \frac{\ln n + \ln \ln n + \omega(1)}{n}$ , the following holds. W.h.p. every subgraph  $H \subseteq G$  of degree sequence  $\mathbf{d}_H \leq \tilde{\mathbf{d}}(t, \varepsilon)$  is such that the graph  $G - H$  contains a subgraph  $\Gamma_0 \in \mathcal{L}(n, \beta)$  that spans at most  $2 \cdot 10^6 \beta n^2 p$  edges. Moreover, adding to  $\Gamma_0$  any subset of edges  $E_0 \subseteq E$  results in a graph in  $\mathcal{L}(n, \beta)$  and the partition  $V = \mathcal{D}_t(G) \cup V_1$  guarantees that  $\Gamma_0 + E_0$  is in  $\mathcal{L}(n, \beta)$ .*

*Proof.* Fix  $\varepsilon > 0$ , let  $\beta_0$  be as guaranteed by Lemma 4.5, let  $0 < \beta \leq \beta_0$ , and set  $t = \frac{np}{100}$ . Fix a subgraph  $H \subseteq G$  with degree sequence  $\mathbf{d}_H \leq \tilde{\mathbf{d}}(t, \varepsilon)$ , denote by  $G_1 = G_1(t) = (V_1, E_1)$ , let  $H_1 = H[V_1]$ , and set  $n_1 = |V_1|$ . We can assume that  $G$  satisfies the following properties:

1.  $\delta(G - H) \geq 2$  (Theorem 4.2).
2.  $|\mathcal{D}_t(G)| \leq n^{0.09}$  (Claim 4.3).
3. There is no path of length at most  $\frac{2 \ln n}{3 \ln \ln n}$  with both (possibly identical) endpoints in  $|\mathcal{D}_t(G)|$  (Claim 4.4).
4.  $G_1 - H_1$  contains a  $(\beta n_1, f_\beta)$ -expander spanning subgraph  $\Gamma$  with at most  $10^6 \beta n^2 p$  edges (Lemmas 4.5 and 3.2, where Claim 4.1 can be used to bound the number of edges in  $e(G_1)$ ).

To get the graph  $\Gamma_0$ , we add to the graph  $\Gamma$  the set of vertices  $D = \mathcal{D}_t(G)$  with all of its incident edges from  $G - H$ . Note that  $\Gamma_0 \in \mathcal{L}(n, \beta)$  (using the partition  $V = V_1 \cup D$ ) and that  $e(\Gamma_0) = e(\Gamma) + e_{G-H}(D, V_1) \leq 10^6 \beta n^2 p + t \cdot n^{0.1} < 2 \cdot 10^6 \beta n^2 p$  as claimed.

Consider any subset of edges  $E_0 \subseteq E(G)$ . As the vertices of  $D$  form an independent set, all edges of  $E_0$  that are not in  $\Gamma_0$  must have at least one endpoint in  $V_1$ . The addition cannot create a path of length at most  $\frac{2 \ln n}{3 \ln \ln n}$  between endpoints in  $D$  since no such path exists in  $G$ . The addition of any edge from  $E_G(D, V_1)$  to  $\Gamma_0$  can only increase the degree of every vertex from  $D$ , and, finally, the addition of any edge from  $E_G(V_1)$  to  $\Gamma_0$  clearly leaves the induced subgraph on  $V_1$  as a  $(\beta n_1, f_\beta)$ -expander (as this is a monotone increasing graph property), and therefore the same partition of the vertex set  $V = V_1 \cup D$  also implies that  $\Gamma_0 + E_0 \in \mathcal{L}(n, \beta)$ .  $\square$

**5. Proof of Theorem 1.** We can now provide the full proof of the main result of this paper—namely, the proof of Theorem 1. The road we take to achieve this is to show that given a typical graph  $G$  sampled from  $\mathcal{G}(n, p)$ , no matter how  $H \subseteq G$  is chosen (given that it satisfies the conditions on its degree sequence), not only will the graph  $G - H$  contain a sparse expander subgraph  $\Gamma$ , it will also have as edges enough boosters with respect to  $\Gamma$  as to transform it into a Hamiltonian graph.

*Proof of Theorem 1.* Fix  $\varepsilon > 0$ , set  $\beta = \beta(\varepsilon)$  to be a sufficiently small constant such that the assertion of Corollary 4.9 holds, and let  $G \sim \mathcal{G}(n, p)$  be with degree sequence  $\mathbf{d}$ . If  $G$  does not satisfy the conclusion of Corollary 4.9, we say that  $G$  is *corrupted* and denote by  $\mathcal{A}_G$  this event of probability  $o(1)$ .

Assume, then, that  $G$  is *not* corrupted and not  $\tilde{\mathbf{d}} = \tilde{\mathbf{d}}(\frac{np}{100}, \varepsilon)$ -resilient to  $\mathcal{HAM}$ ; i.e., there exists a subgraph  $H_0$  with degree sequence  $\mathbf{d}_H \leq \mathbf{d}$  for which  $G - H_0$  is not Hamiltonian. Corollary 4.9 implies that  $G - H_0$  contains a subgraph  $\Gamma_0 \in \mathcal{L}(n, \beta)$  that spans at most  $2 \cdot 10^6 \beta n^2 p$  edges, and moreover, adding to it any subset of edges from  $E_1 \subseteq E(G)$  results in a graph  $\Gamma_0 + E_1 \in \mathcal{L}(n, \beta)$ . Lemma 3.4 implies that there must exist a set  $E_0 \subseteq E(G - H - \Gamma_0) \subseteq E(G)$  of at most  $n$  edges for which  $|N_{G - (\Gamma_0 + E_0)}(v) \cap \mathcal{B}_{\Gamma_0 + E_0}(v)| \leq \tilde{d}_v$  for every vertex  $v \in V$ . As  $|E_0| \ll e(\Gamma_0)$  from the above, we conclude that  $\Gamma_2 = \Gamma_0 + E_0 \in \mathcal{L}'(n, \beta) = \{\Gamma \in \mathcal{L}(n, \beta) : |E(\Gamma)| \leq 10^7 \beta n^2 p\}$ .

Corollary 4.9 guarantees that  $V = V_1 \cup \mathcal{D}_i(G)$  is a partition of the vertex set for which  $\Gamma_2$  satisfies the properties of  $\mathcal{L}(n, \beta)$ . Lemma 3.6 implies that the set  $A = \{v \in V_1 : |B_{\Gamma_2}(v)| \geq n(1/3 + \beta)\}$  must satisfy  $|A| \geq n(1/3 + \beta)$ . Let  $A_0 \subseteq A$  be a subset of cardinality  $|A_0| = \frac{n}{3}$ .

So, in fact, for noncorrupted  $G$  we will resort to bounding the event that there exists a graph  $\Gamma_2 \in \mathcal{L}'$  that is contained in  $G$  for which  $|N_{G - \Gamma_2}(v) \cap \mathcal{B}_{\Gamma_2}(v)| \leq \tilde{d}_v = (\frac{1}{3} - \varepsilon) \tilde{d}_v$  for every vertex  $v \in A_0$ . One should note that from the independence of the appearance of edges in the  $\mathcal{G}(n, p)$  model, given some  $\Gamma_2 \in \mathcal{L}'$ , the events  $[\Gamma_2 \subseteq G]$  and  $[|N_{G - \Gamma_2}(v) \cap \mathcal{B}_{\Gamma_2}(v)| \leq \tilde{d}_v]$  are independent as they stem from the appearance of disjoint sets of edges, i.e.,  $e(\Gamma_2)$  and  $\binom{V}{2} \setminus e(\Gamma_2)$ , respectively. Putting it all together yields that the probability that  $G$  is not  $\tilde{\mathbf{d}}$ -resilient to  $\mathcal{HAM}$  is upper bounded by

$$\begin{aligned} & \Pr[\mathcal{A}_G] + \Pr[\exists \Gamma_2 \in \mathcal{L}'. (\Gamma_2 \subseteq G) \wedge (\forall v \in V. |N_{G - \Gamma_2}(v) \cap \mathcal{B}_{\Gamma_2}(v)| \leq \tilde{d}_v) | \bar{\mathcal{A}}_G] \\ & \leq o(1) + \frac{1}{\Pr[\bar{\mathcal{A}}_G]} \sum_{\Gamma_2 \in \mathcal{L}'} \Pr[\Gamma_2 \subseteq G] \cdot \Pr[\forall v \in V. |N_{G - \Gamma_2}(v) \cap \mathcal{B}_{\Gamma_2}(v)| \leq \tilde{d}_v] \\ & \leq o(1) + (1 + o(1)) \sum_{\Gamma_2 \in \mathcal{L}'} p^{e(\Gamma_2)} \times \\ (5.1) \quad & \Pr \left[ \sum_{v \in A_0} |N_{G - \Gamma_2}(v) \cap \mathcal{B}_{\Gamma_2}(v)| \leq \left(\frac{1}{3} - \varepsilon\right) (2e_G(A_0) + e_G(A_0, V \setminus A_0)) \right]. \end{aligned}$$

Taking into account that we are using a union bound argument by summing over all graphs  $\Gamma_2 \in \mathcal{L}'$  (and there may be an exponential number of those), it is left to show that every summand in the above expression is exponentially small. We define the following random variable (which depends on the choice of  $\Gamma_2$ ). Let

$$X = X(G) = \sum_{v \in A_0} |N_{G - \Gamma_2}(v) \cap B_{\Gamma_2}(v)|,$$

whose expectation satisfies

$$\mathbf{E}[X] = \sum_{v \in A_0} \mathbf{E}[|N_{G - \Gamma_2}(v) \cap B_{\Gamma_2}(v)|] = p \cdot \sum_{v \in A_0} \mathbf{E}[|B_{\Gamma_2}(v)|] \geq \frac{n^2 p}{9}.$$

Note that for any two graphs on the vertex set  $V$  that differ by a single edge, their value of  $X$  can change by at most 2 (1 for every endpoint of the edge); hence we can apply the Azuma–Hoeffding inequality for martingales of bounded variance



(see, e.g., [1, Theorem 7.4.3]) to prove that  $X$  is concentrated around its expectation. In the process of “exposing” the edges of the graph, it suffices to expose only the pairs with an endpoint in  $A_0$  which are nonedges of  $\Gamma_2$ . This implies that the total variance of the martingale is upper bounded by  $\frac{n^2}{3}p(1-p)$ , and hence

$$(5.2) \quad \Pr \left[ X(G) \leq \frac{n^2 p}{9} (1 - \varepsilon) \right] \leq \exp(-\Theta(\varepsilon^2 n^2 p)).$$

On the other hand, by a standard application of Theorem 2.2 (the Chernoff bound), it follows that

$$(5.3) \quad \Pr \left[ \exists U \subseteq V: |U| = \frac{n}{3} \wedge 2e(U) + e(U, V \setminus U) > \frac{n^2 p}{3} (1 + \varepsilon) \right] = \exp(-\Theta(\varepsilon^2 n^2 p)).$$

Using the fact that  $\frac{n^2 p}{9} (1 - \varepsilon) \geq \frac{n^2 p}{3} (\frac{1}{3} - \varepsilon)(1 + \varepsilon)$ , it follows from (5.2) and (5.3) that

$$\Pr \left[ X \leq \left( \frac{1}{3} - \varepsilon \right) (2e_G(A_0) + e_G(A_0, V \setminus A_0)) \right] \leq \exp(-\Theta(\varepsilon^2 n^2 p)).$$

Returning to our upper bound on the probability that  $G$  is not  $\tilde{\mathbf{d}}$ -resilient to  $\mathcal{HAM}$ , we set  $\mu = 10^7 \beta$ . Plugging in the above in (5.1), the probability is upper bounded by

$$\begin{aligned} & o(1) + (1 + o(1)) \sum_{m=1}^{\mu n^2 p} \binom{\binom{n}{2}}{m} \cdot p^m \cdot \exp(-\Theta(\varepsilon^2 n^2 p)) \\ & \leq o(1) + (1 + o(1)) \sum_{m=1}^{\mu n^2 p} \left( \frac{en^2 p}{2m} \right)^m \cdot \exp(-\Theta(\varepsilon^2 n^2 p)) \\ & \leq o(1) + \exp \left( \Theta \left( \mu n^2 p \ln \frac{1}{\mu} \right) - \Theta(\varepsilon^2 n^2 p) \right) = o(1), \end{aligned}$$

where the transition from the first to the second inequality follows from the fact that  $\left(\frac{en^2 p}{m}\right)^m$  is increasing with  $m$  for the given range, and that  $\beta$  (and  $\mu$ ) can be chosen small enough with respect to  $\varepsilon$ . This completes the proof of the theorem.  $\square$

**6. Concluding remarks and further research directions.** This work is yet another building block in the recently initiated research area of resilience of graph properties. The generalized approach allowed us to tackle in a uniform way two different problems regarding Hamiltonicity. The main motivation for studying resilience of the type considered in this work is the ability to have refined control over vertices of small degree that create the main obstacle for Hamiltonicity when  $p$  is in the low end of the range. As  $p$  grows, the degree sequence of the graph becomes more and more balanced, so this approach does not seem to be suitable (or even necessary) for larger values of  $p$ . Although in this work some progress has been made on the two fronts (the local resilience of random graphs with respect to  $\mathcal{HAM}$  and optimal packing of Hamilton cycles in random graphs), there are still gaps to fill in order to settle the two main questions considered in this paper, and we believe new ideas will be needed to resolve them completely.

## REFERENCES

- [1] N. ALON AND J. H. SPENCER, *The Probabilistic Method*, Wiley–Intersci. Ser. Discrete Math. Optim., 3rd ed., John Wiley & Sons, New York, 2008.
- [2] N. ALON AND B. SUDAKOV, *Increasing the chromatic number of a random graph*, J. Comb., 1 (2010), pp. 345–356.
- [3] J. BALOGH, B. CSABA, AND W. SAMOTLIJ, *Local resilience of almost spanning trees in random graphs*, Random Structures Algorithms, 38 (2011), pp. 121–139.
- [4] J. BALOGH, C. LEE, AND W. SAMOTLIJ, *Corrádi and Hajnal’s Theorem for Sparse Random Graphs*, preprint, arxiv: <http://arxiv.org/abs/1011.5443>.
- [5] S. BEN-SHIMON, M. KRIVELEVICH, AND B. SUDAKOV, *Local resilience and Hamiltonicity maker–breaker games in random regular graphs*, Combin. Probab. Comput., 20 (2011), pp. 173–211.
- [6] B. BOLLOBÁS, *Almost all regular graphs are Hamiltonian*, Eur. J. Combin., 4 (1983), pp. 94–106.
- [7] B. BOLLOBÁS, *Random Graphs*, Cambridge University Press, Cambridge, UK, 2001.
- [8] B. BOLLOBÁS AND A. FRIEZE, *On matchings and Hamiltonian cycles in random graphs*, in Random Graphs (Poznań 1983), Ann. Discrete Math. 28, North-Holland, Amsterdam, 1985, pp. 23–46.
- [9] J. BÖTTCHER, Y. KOHAYAKAWA, AND A. TARAZ, *Almost Spanning Subgraphs of Random Graphs after Adversarial Edge Removal*, preprint, arxiv: <http://arxiv.org/abs/1003.0890>.
- [10] D. CHRISTOFIDES, D. KÜHN, AND D. OSTHUS, *Edge-disjoint Hamilton Cycles in Graphs*, preprint, arxiv: <http://arxiv.org/abs/1011.5443>.
- [11] D. DELLAMONICA, Y. KOHAYAKAWA, M. MARCINISZYN, AND A. STEGER, *On the resilience of long cycles in random graphs*, Electron. J. Combin., 15 (2008), p. R32.
- [12] R. DIESTEL, *Graph Theory*, Grad. Texts in Math. 173, 3rd ed., Springer-Verlag, New York, 2005.
- [13] A. FRIEZE AND M. KRIVELEVICH, *On packing Hamilton cycles in  $\epsilon$ -regular graphs*, J. Combin. Theory Ser. B, 94 (2005), pp. 159–172.
- [14] A. FRIEZE AND M. KRIVELEVICH, *On two Hamiltonian cycle problems in random graphs*, Israel J. Math., 166 (2008), pp. 221–234.
- [15] H. HUANG, C. LEE, AND B. SUDAKOV, *Bandwidth theorem for sparse graphs*, J. Combin. Theory Ser. B, to appear.
- [16] S. JANSON, T. ŁUCZAK, AND A. RUCIŃSKI, *Random Graphs*, Wiley–Intersci. Ser. Discrete Math. Optim., John Wiley & Sons, New York, 2000.
- [17] J. H. KIM AND N. WORMALD, *Random matchings which induce Hamilton cycles and Hamiltonian decompositions of random regular graphs*, J. Combin. Theory Ser. B, 81 (2001), pp. 20–44.
- [18] J. KOMLÓS AND E. SZEMERÉDI, *Limit distributions for the existence of Hamilton circuits in a random graph*, Discrete Math., 43 (1983), pp. 55–63.
- [19] M. KRIVELEVICH, C. LEE, AND B. SUDAKOV, *Resilient pancyclicity of random and pseudorandom graphs*, SIAM J. Discrete Math., 24 (2010), pp. 1–16.
- [20] C. LEE AND W. SAMOTLIJ, *Pancyclic Subgraphs of Random Graphs*, preprint, arxiv: <http://arxiv.org/abs/1005.5716>.
- [21] C. S. J. A. NASH-WILLIAMS, *Edge-disjoint Hamiltonian circuits in graphs with vertices of large valency*, in Studies in Pure Mathematics, L. Minski, ed., Academic Press, London, 1971, pp. 157–183.
- [22] L. PÓSA, *Hamiltonian circuits in random graphs*, Discrete Math., 14 (1976), pp. 359–364.
- [23] B. SUDAKOV AND V. H. VU, *The local resilience of random graphs*, Random Structures Algorithms, 33 (2008), pp. 409–433.
- [24] P. TURÁN, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok, 48 (1941), pp. 436–452.