Combinatorics, Probability and Computing (2021), **30**, pp. 17–36
doi:10.1017/S0963548320000346

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Dirac's theorem for random regular graphs

Padraig Condon[1], Alberto Espuny Díaz[2], António Girão[1,*], Daniela Kühn[1] and Deryk Osthus[1,†]

[1]School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK and [2]Institut für Mathematik, Technische Universität Ilmenau, 98693 Ilmenau, Germany
*Corresponding author. Email: giraoa@bham.ac.uk

(Received 30 April 2019; revised 23 June 2020; first published online 28 August 2020)

Abstract

We prove a 'resilience' version of Dirac's theorem in the setting of random regular graphs. More precisely, we show that whenever d is sufficiently large compared to $\varepsilon > 0$, a.a.s. the following holds. Let G' be any subgraph of the random n-vertex d-regular graph $G_{n,d}$ with minimum degree at least $(1/2 + \varepsilon)d$. Then G' is Hamiltonian.

This proves a conjecture of Ben-Shimon, Krivelevich and Sudakov. Our result is best possible: firstly the condition that d is large cannot be omitted, and secondly the minimum degree bound cannot be improved.

2020 MSC Codes: Primary 05C80; Secondary 05C35, 05C45

1. Introduction

The study of Hamiltonicity has been at the core of graph theory for the past few decades. A graph G is said to be *Hamiltonian* if it contains a cycle which covers all of the vertices of G, and this is called a *Hamilton cycle*. It is well known that the problem of determining whether a graph is Hamiltonian is NP-complete, and thus most results about Hamiltonicity deal with sufficient conditions which guarantee this property. One of the most well-known examples is due to Dirac, who proved that any graph G on $n \geqslant 3$ vertices with minimum degree at least $n/2$ is Hamiltonian.

1.1 Hamilton cycles in random graphs

The search for Hamilton cycles in various models of random graphs has also been a driving force in the development of this theory. The classical binomial model $G_{n,p}$, in which each possible edge is added to an n-vertex graph with probability p independently of the other edges, has seen many results in this direction. In particular, Komlós and Szemerédi [23] showed that $p = \log n/n$ is the 'sharp' threshold for the existence of a Hamilton cycle. This can be strengthened to obtain the following hitting time result. Consider a random graph process as follows: given a set of n vertices, add each of the $\binom{n}{2}$ possible edges, one by one, by choosing the next edge uniformly at random among those that have not been added yet. In this setting, Ajtai, Komlós and Szemerédi [1] and

†This project has received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 786198, D. Kühn and D. Osthus). The research leading to these results was also partially supported by the EPSRC, grants EP/N019504/1 (A. Girão and D. Kühn) and EP/S00100X/1 (D. Osthus), as well as the Royal Society and the Wolfson Foundation (D. Kühn).

Bollobás [10] independently proved that a.a.s. the resulting graph becomes Hamiltonian as soon as its minimum degree is at least 2.

The search for Hamilton cycles in other random graph models has proved more difficult. In this paper we will deal with random *regular* graphs: given $n, d \in \mathbb{N}$ such that $d < n$ and nd is even, $G_{n,d}$ is chosen uniformly at random from the set of all d-regular graphs on n vertices. The study of this model is often more challenging than that of $G_{n,p}$ due to the fact that the presence and absence of edges in $G_{n,d}$ are correlated. Several different techniques have been developed to deal with this model, such as the configuration model (see Section 3.3) or edge-switching techniques. Robinson and Wormald [33] proved that $G_{n,3}$ is a.a.s. Hamiltonian, and later extended this result to $G_{n,d}$ for any fixed $d \geqslant 3$ [34]. This is in contrast to $G_{n,p}$, where the average degree must be logarithmic in n to ensure Hamiltonicity. These results were later generalized by Cooper, Frieze and Reed [14] and Krivelevich, Sudakov, Vu and Wormald [26] for the case when d is allowed to grow with n, up to $d \leqslant n-1$. Many further results can be found in the recent survey of Frieze [16].

1.2 Local resilience

More recently, several extremal results have been translated to random graphs via the concept of local resilience. The *local resilience* of a graph G with respect to some property $\mathcal{P}$ is the maximum number $r \in \mathbb{N}$ such that, for all $H \subseteq G$ with $\Delta(H) < r$, the graph $G \setminus H$ satisfies $\mathcal{P}$. We say that G is r-*resilient* with respect to a property $\mathcal{P}$ if the local resilience of G is greater than r. The systematic study of local resilience was initiated by Sudakov and Vu [36], and the subject has seen a lot of research since.

Note that Dirac's theorem can be restated in this terminology to say that the local resilience of the complete graph K_n with respect to Hamiltonicity is $\lfloor n/2 \rfloor$. This concept of local resilience then naturally suggests a generalization of Dirac's theorem to random graphs. In the binomial model, Lee and Sudakov [27] showed that, for any constant $\varepsilon > 0$, if $p \geqslant C \log n/n$ and C is sufficiently large, then a.a.s. $G_{n,p}$ is $(1/2 - \varepsilon)np$-resilient with respect to Hamiltonicity. This improved on earlier bounds [7, 8, 17, 36]. Very recently, Montgomery [30] and independently Nenadov, Steger and Trujić [32], proved a hitting time result for the local resilience of $G_{n,p}$ with respect to Hamiltonicity. In a different direction, Condon, Espuny Díaz, Kim, Kühn and Osthus [12] considered 'resilient' versions of Pósa's theorem and Chvátal's theorem for $G_{n,p}$.

The resilience of random regular graphs with respect to Hamiltonicity is less understood. Ben-Shimon, Krivelevich and Sudakov [7] proved that, for large (but constant) d, a.a.s. $G_{n,d}$ is $(1-\varepsilon)d/6$-resilient with respect to Hamiltonicity. They conjectured that the true value should be closer to $d/2$.

Conjecture 1.1 (Ben-Shimon, Krivelevich and Sudakov [7])**.** *For every $\varepsilon > 0$ there exists an integer $D = D(\varepsilon) > 0$ such that, for every fixed integer $d > D$, the local resilience of $G_{n,d}$ with respect to Hamiltonicity a.a.s. lies in the interval* $((1/2 - \varepsilon)d, (1/2 + \varepsilon)d)$.

They also suggested studying the same problem when d is allowed to grow with n. In this direction, Sudakov and Vu [36] showed that, for any fixed $\varepsilon > 0$ and for any (n, d, λ)-graph G (*i.e.* a d-regular graph on n vertices whose second largest eigenvalue in absolute value is at most λ) with $d/\lambda > \log^2 n$, we have that G is $(1/2 - \varepsilon)d$-resilient with respect to Hamiltonicity. This, together with a result of Krivelevich, Sudakov, Vu and Wormald [26] and recent results of Cook, Goldstein and Johnson [13] and Tikhomirov and Youssef [37] about the spectral gap of random regular graphs, implies that for $\log^4 n \ll d \leqslant n-1$, a.a.s. $G_{n,d}$ is $(1/2 - \varepsilon)d$-resilient with respect to Hamiltonicity. One can extend this to $d \gg \log n$ by combining a result of Kim and Vu [22] on joint distributions of binomial random graphs and random regular graphs with the result of Lee and Sudakov [27] about the resilience of $G_{n,p}$ with respect to Hamiltonicity.

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

The study of local resilience has not been restricted to Hamiltonicity. Other properties that have been considered include the containment of perfect matchings [12, 32], directed Hamilton cycles [15, 18, 31], cycles of all possible lengths [25], kth powers of cycles [35], bounded degree trees [5], triangle factors [6] and bounded degree graphs [2, 20].

1.3 New results

In this paper we completely resolve Conjecture 1.1, as well as its extension to d growing slowly with n (recall that the case when $d \gg \log n$ is covered by earlier results). This can be seen as a version of Dirac's theorem for random regular graphs. Our main result gives the lower bound in Conjecture 1.1.

Theorem 1.2. *For every $\varepsilon > 0$ there exists D such that, for every $D < d \leqslant \log^2 n$, the random graph $G_{n,d}$ is a.a.s. $(1/2 - \varepsilon)d$-resilient with respect to Hamiltonicity.*

While we do not try to optimize the dependence of D on ε, we remark that D in Theorem 1.2 can be taken to be polynomial in ε^{-1}. This is essentially best possible in the sense that Theorem 1.2 fails if $d \leqslant (2\varepsilon)^{-1}$.

Theorem 1.3. *For any odd $d > 2$, the random graph $G_{n,d}$ is not a.a.s. $(d-1)/2$-resilient with respect to Hamiltonicity.*

Our proof also shows that $G_{n,d}$ is not a.a.s. $(d-1)/2$-resilient with respect to the containment of a perfect matching. Moreover, one can adapt the proof of Theorem 1.3 to show that, for every even d, the random graph $G_{n,d}$ is not a.a.s. $d/2$-resilient with respect to Hamiltonicity (or the containment of a perfect matching). It would also be interesting to obtain bounds on the resilience for small d. In particular, here are some questions.

(i) Given any fixed even d, determine whether the graph $G_{n,d}$ is a.a.s. $(d/2-1)$-resilient with respect to Hamiltonicity.
(ii) What is the likely resilience of $G_{n,4}$ with respect to Hamiltonicity or the containment of perfect matchings? Is a graph obtained from $G_{n,4}$ by removing any matching a.a.s. Hamiltonian?

Finally, we observe (as is well known) that the upper bound of $(1/2 + \varepsilon)d$ in Conjecture 1.1 follows easily from edge distribution properties of random regular graphs. Indeed, we note that for every $\varepsilon > 0$ there exists a constant D such that for every $D \leqslant d \leqslant \log^2 n$, a.a.s. the graph $G = G_{n,d}$ has the property that between any two disjoint sets A, B of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$, respectively, the number of edges in $G[A, B]$ is a.a.s. bounded from above by $(1/2 + \varepsilon/2)nd/2$ (see Proposition 4.2). Now let A, B be a maximum cut in G. Thus $e_G(a, B) \geqslant d/2$ for all $a \in A$, and similarly for all $b \in B$. If $|A| \neq |B|$, then by deleting the edges in $G[A] \cup G[B]$, the remaining graph is not Hamiltonian since it forms an unbalanced bipartite graph. If $|A| = |B|$, then by the above property, there must exist a vertex $x \in A$ such that $e_G(x, B) \leqslant (1/2 + \varepsilon/2)d$. Let $A' \coloneqq A \setminus \{x\}$ and $B' \coloneqq B \cup \{x\}$. As before, by deleting the edges in $G[A'] \cup G[B']$, we obtain a graph which is not Hamiltonian.

1.4 Organization of the paper

The remainder of the paper is organized as follows. In Section 2 we give a sketch of the proof of Theorem 1.2. In Section 3 we collect notation, some probabilistic tools, and observations about the configuration model. Section 4 is devoted to proving different edge-distribution and expansion properties of random regular graphs and their subgraphs, and the proof of Theorem 1.2 is given in Section 5, using all the techniques that have been introduced before. Finally, we prove Theorem 1.3 in Section 6.

2. Outline of the proof of Theorem 1.2

Consider $G = G_{n,d}$. Let $H \subseteq G$ be such that $\Delta(H) \leqslant (1/2 - \varepsilon)d$ and let $G' \coloneqq G \setminus H$. We will prove that G' contains a 'sparse' spanning subgraph R which has strong edge expansion properties. These properties will then be used to provide a lower bound on the number of edges in G whose addition would make R Hamiltonian, or increase the length of a longest path in R (such edges are commonly called 'boosters': see *e.g.* [24]). We then argue that some of these edges must in fact be retained when passing to G'. We then add such edges to R and iterate the above process (at most n times) until R becomes Hamiltonian.

More specifically, as a preliminary step we 'thin' the graph G', that is, we take a subgraph $R \subseteq G'$ with $\Delta(R) \leqslant \delta d$, for some $\delta \ll \varepsilon$. As described above, we consider a longest path in R and then argue that it can be extended via edges in $G' \setminus R$. The fact that R is relatively 'sparse' with respect to G' will be important when calculating union bounds over all graphs R of this type, at a later stage in the proof (this idea was introduced by Ben-Shimon, Krivelevich and Sudakov [7]).

Given many paths of maximum length and with different end-points in R, it follows that there will be many edges whose addition will increase the length of a longest path (or make R Hamiltonian). A theorem of Pósa implies that graphs with strong expansion properties will indeed contain many of such paths. These expansion properties are captured by the notion of a 3-expander (see Definition 4.1). Therefore we wish to show that our thinned graph R can be chosen to be a 3-expander. This is one point where working with the random graph $G_{n,d}$ proves more difficult than working with $G_{n,p}$, due to the fact that the appearance of edges in $G_{n,d}$ is correlated.

The next step is to provide a lower bound on the number of edges whose addition to R would increase the length of a longest path (or make R Hamiltonian). Here we further develop an approach of Montgomery [30] who, instead of considering single edges that would bring R closer to being Hamiltonian, considered 'booster' edge pairs whose addition would yield the same result. For example, if R is connected and P is a longest path in R with end-points x and y, and ab is an edge of P (with b closer to y on P), then $\{ya, xb\}$ is a booster pair. The main advantage of considering such pairs of edges is that it results in a much larger set of boosters for R. More precisely, we show the existence of another thinned graph $F \subseteq G' \setminus R$ for which each booster we consider is of the form $\{e, e'\}$, where $e \in E(F)$ and $e' \in E(G')$ (see Corollary 5.4).

Finally, we can complete the proof of the main theorem by iteratively adding booster pairs to the thinned graph R, increasing the length of a maximum path in each step until R becomes Hamiltonian. Two points are important here as to why we can iterate this process. First, proving the existence of boosters (see Lemma 5.5) involves a union bound over all pairs of thinned graphs R and F. To bound this efficiently, we need that both R and F are relatively 'sparse' with respect to G'. But in each step we only add two booster edges to R, so it remains sparse. Secondly, we take special care to ensure that no vertex is contained in too many of the boosters we add to R, ensuring that its degree in successive iterations remains small. This process terminates after at most n iterations, resulting in a graph $R' \subseteq G'$ which is Hamiltonian.

3. Preliminaries

3.1 Notation

For $n \in \mathbb{N}$, we denote $[n] \coloneqq \{1, \ldots, n\}$. Given any set S, we denote

$$S^{(2)} \coloneqq \{\{s_1, s_2\} : s_1, s_2 \in S, s_1 \neq s_2\}.$$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

The parameters which appear in hierarchies are chosen from right to left. That is, whenever we claim that a result holds for $0 < a \ll b \leqslant 1$, we mean that there exists a non-decreasing function $f\colon [0, 1) \to [0, 1)$ such that the result holds for all $a > 0$ and all $b \leqslant 1$ with $a \leqslant f(b)$. We will not compute these functions explicitly.

Throughout this paper, the word *graph* will refer to a simple, undirected graph. Whenever the graphs are allowed to have parallel edges or loops, we will refer to these as *multigraphs*. Given any (multi)graph $G = (V, E)$ and sets $A, B \subseteq V$, we will denote the (multi)set of edges of G spanned by A as $E_G(A)$, and the (multi)set of edges of G having one end-point in A and one end-point in B as $E_G(A, B)$. The number of such edges will be denoted by $e_G(A)$ and $e_G(A, B)$, respectively. We will also write $e(G)$ for $e_G(V)$. Given two (multi)graphs G_1 and G_2 on the same vertex set V, we write $G_1 + G_2 \coloneqq (V, E(G_1) \cup E(G_2))$, where the union represents set union for graphs and multiset union for multigraphs. When G_1 and G_2 are graphs, we write $G_1 \setminus G_2 \coloneqq (V, E(G_1) \setminus E(G_2))$. Given any vertex $v \in V$, we will denote the set of vertices which are adjacent to v in G by $N_G(v)$. We define $N_G(A) \coloneqq \bigcup_{v \in A} N_G(v)$. The *degree* of vertex v in a multigraph G is

$$d_G(v) \coloneqq |\{e \in E(G)\colon v \in e\}| + |\{e \in E(G)\colon e = vv\}|$$

(*i.e.* each loop at v contributes two to $d_G(v)$). We denote $\Delta(G) \coloneqq \max_{v \in V} d_G(v)$ and $\delta(G) \coloneqq \min_{v \in V} d_G(v)$. The (multi)graph G is said to be *d-regular* for some $d \in \mathbb{N}$ if all vertices have degree d. Given a multigraph G on $[n]$, we refer to the vector $\mathbf{d} = (d_G(1), \ldots, d_G(n))$ as its *degree sequence*. In general, a vector $\mathbf{d} = (d_1, \ldots, d_n)$ with $d_i \in \mathbb{Z}_{\geqslant 0}$ for all $i \in [n]$ is called *graphic* if there exists a graph on n vertices with degree sequence $\mathbf{d}$ (note that, as long as $\sum_{i=1}^{n} d_i$ is even, there is always a multigraph with degree sequence $\mathbf{d}$). Given a graph G and a real number $\alpha > 0$, let $\mathcal{H}_\alpha(G)$ be the collection of all spanning subgraphs $H \subseteq G$ for which $d_H(v) \leqslant \alpha d_G(v)$, for all $v \in V(G)$.

We will use $\mathcal{G}_{n,d}$ to denote the set of all d-regular graphs on vertex set $[n]$, and $G_{n,d}$ will denote a graph chosen from $\mathcal{G}_{n,d}$ uniformly at random. Whenever we use this notation, we implicitly assume that nd is even. In more generality, given a graphic degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$, we will denote the collection of all graphs on vertex set $[n]$ with degree sequence $\mathbf{d}$ by $\mathcal{G}_{n,\mathbf{d}}$, and $G_{n,\mathbf{d}}$ will denote a graph chosen from $\mathcal{G}_{n,\mathbf{d}}$ uniformly at random.

We use *a.a.s.* as an abbreviation for *asymptotically almost surely*. Given a sequence of events $\{\mathcal{E}_n\}_{n \in \mathbb{N}}$, whenever we claim that $\mathcal{E}_n$ holds a.a.s., we mean that the probability that $\mathcal{E}_n$ holds tends to 1 as n tends to infinity. For the purpose of clarity, we will ignore rounding issues when dealing with asymptotic statements. By abusing notation, given $p \geqslant 0$ and $n \in \mathbb{N}$, we write $\mathrm{Bin}(n, p)$ for the binomial distribution with parameters n and $\min\{p, 1\}$.

3.2 Probabilistic tools

We will need the following Chernoff bound (see *e.g.* [21, Corollary 2.3]).

Lemma 3.1. *Let X be the sum of n independent Bernoulli random variables and let $\mu \coloneqq \mathbb{E}[X]$. Then, for all $0 \leqslant \delta \leqslant 1$, we have that $\mathbb{P}[|X - \mu| \geqslant \delta\mu] \leqslant 2e^{-\delta^2\mu/3}$.*

The following bound will also be used repeatedly (see *e.g.* [3, Theorem A.1.12]).

Lemma 3.2. *Let $X \sim \mathrm{Bin}(n, p)$, and let $\beta > 1$. Then $\mathbb{P}[X \geqslant \beta np] \leqslant (e/\beta)^{\beta np}$.*

Given any sequence of random variables $X = (X_1, \ldots, X_n)$ taking values in a set A and a function $f\colon A^n \to \mathbb{R}$, for each $i \in [n] \cup \{0\}$ define $Y_i \coloneqq \mathbb{E}[f(X) \mid X_1, \ldots, X_i]$. The sequence $Y_0, \ldots, Y_n$ is called the *Doob martingale* for f. All the martingales that appear in this paper will be of this form. To deal with them, we will need the following version of the well-known Azuma–Hoeffding inequality.

Lemma 3.3 (Azuma's inequality [4, 19]). *Let $X_0, X_1, \ldots$ be a martingale and suppose that $|X_i - X_{i-1}| \leqslant c_i$ for all $i \in \mathbb{N}$. Then, for any $n, t \in \mathbb{N}$,*

$$\mathbb{P}[|X_n - X_0| \geqslant t] \leqslant 2\exp\left(\frac{-t^2}{2\sum_{i=1}^{n} c_i^2}\right).$$

Finally, the Lóvasz local lemma will be useful. Let $\mathfrak{E} \coloneqq \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m\}$ be a collection of events. A *dependency graph* for $\mathfrak{E}$ is a graph H on vertex set $[m]$ such that, for all $i \in [m]$, $\mathcal{E}_i$ is mutually independent of $\{\mathcal{E}_j : j \neq i, j \notin N_H(i)\}$, that is, if $\mathbb{P}[\mathcal{E}_i] = \mathbb{P}[\mathcal{E}_i \mid \bigwedge_{j\in J} \mathcal{E}_j]$ for all $J \subseteq [m] \setminus (N_H(i) \cup \{i\})$. We will use the following version of the local lemma (it follows from [3, Lemma 5.1.1], for example).

Lemma 3.4 (Lóvasz local lemma). *Let $\mathfrak{E} \coloneqq \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m\}$ be a collection of events and let H be a dependency graph for $\mathfrak{E}$. Suppose that $\Delta(H) \leqslant d$ and $\mathbb{P}[\mathcal{E}_i] \leqslant p$ for all $i \in [m]$. If $ep(d+1) \leqslant 1$, then*

$$\mathbb{P}\left[\bigwedge_{i=1}^{m} \overline{\mathcal{E}_i}\right] \geqslant (1 - ep)^m.$$

3.3 The configuration model

We will work with the *configuration model* introduced by Bollobás [9], which can be used to sample d-regular graphs uniformly at random. In more generality, it can be used to produce graphs with any given graphic degree sequence $\mathbf{d}$. The process to generate such graphs is as follows.

Given $n \in \mathbb{N}$ and a degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$ with $m \coloneqq \sum_{i=1}^{n} d_i$ even, consider a set of m vertices labelled as x_{ij} for $i \in [n]$ and $j \in [d_i]$. For each $i \in [n]$, we call the set $\{x_{ij} : j \in [d_i]\}$ the *expanded set* of i. Similarly, for any $X \subseteq [n]$, we call the set $\{x_{ij} : i \in X, j \in [d_i]\}$ the *expanded set* of X. Choose uniformly at random a perfect matching M covering the expanded set of $[n]$. Then obtain a multigraph $\varphi(M) = ([n], E)$ by letting E be the following multiset: for each edge $e \in M$, consider its end-points $e = x_{ij}x_{k\ell}$, for some $i, k \in [n]$, $j \in [d_i]$ and $\ell \in [d_k]$, and add ik to E (if $i = k$, this adds a loop to E).

When we consider a multigraph G obtained via this configuration model, this will be denoted by $G \sim \mathcal{C}_{n,\mathbf{d}}$. In particular, when we obtain a d-regular multigraph via the configuration model, we will denote this by $G \sim \mathcal{C}_{n,d}$. We refer to the possible perfect matchings on the expanded set of $[n]$ as *configurations*, and we will denote a configuration obtained uniformly at random by $M \sim \mathcal{C}^*_{n,\mathbf{d}}$. By abusing notation, we will sometimes also use $\mathcal{C}^*_{n,\mathbf{d}}$ to denote the set of all configurations with parameters n and $\mathbf{d}$. In order to easily distinguish the setting of graphs from that of configurations, we will call the elements of the expanded sets *points*, and each element in a configuration will be called a *pairing*.

The above process may produce a multigraph with loops and/or multiple edges. However, if $\mathbf{d}$ is a graphic degree sequence, then, when conditioning on the resulting multigraph being simple, the configuration model yields a graph $G \in \mathcal{G}_{n,\mathbf{d}}$ chosen uniformly at random. The following proposition bounds the probability that this happens, and can be proved similarly to (part of) a result of Cooper, Frieze and Reed [14, Lemma 7] (see [11] for details). It will be useful when analysing the distribution of edges in $G_{n,d}$ via the configuration model.

Proposition 3.5. *Let $0 < \delta < 1/10$. Let $d \leqslant \log^2 n$ be a positive integer and let R be a graph on vertex set $[n]$ with degree sequence $\mathbf{d}' = (d_1, \ldots, d_n)$ such that $d_i < \delta d$ for all $i \in [n]$. Let $\mathbf{d} \coloneqq (d - d_1, \ldots, d - d_n)$ and let $F \sim \mathcal{C}_{n,\mathbf{d}}$. Then, if n is sufficiently large,*

$$\mathbb{P}[R + F \text{ is simple}] \geqslant e^{-3d^2}.$$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

Note that, by choosing R to be the empty graph on n vertices, we obtain a lower bound on the probability that the multigraph obtained by a random configuration is simple.

When studying the configuration model, it will be useful to consider the following process to generate $M \sim \mathcal{C}^*_{n,\mathbf{d}}$. Let $\mathbf{d} = (d_1, \ldots, d_n)$ and suppose that $m := \sum_{i=1}^n d_i$ is even. Label the points of the expanded set of $[n]$ in any arbitrary order, $x_1, \ldots, x_m$, and identify them naturally with the set $[m]$. Start with an empty set of pairings M_0. Inductively, for each $i \in [m]$, if i is covered by M_{i-1}, let $M_i := M_{i-1}$; otherwise, choose a point $j \in [m] \setminus (V(M_{i-1}) \cup \{i\})$ uniformly at random and define $M_i := M_{i-1} \cup \{ij\}$. We sometimes refer to M_i as the ith *partial configuration*. Finally, let $M := M_m$. It is clear that the resulting configuration M is generated uniformly at random, independently of the labelling of the expanded set of $[n]$.

We will often be interested in bounding the number of edges in $G_{n,d}$ between two sets of vertices. For this, it will be useful to consider binomial random variables that stochastically dominate the number of edges. We formalize this via the following lemma.

Lemma 3.6. *Let $n, d \in \mathbb{N}$ with $d < n$, and let $\delta \in [0, 1)$. Let $\mathbf{d} = (d_1, \ldots, d_n)$ with $\sum_{i=1}^n d_i$ even be such that $(1-\delta)d \leqslant d_i \leqslant d$ for all $i \in [n]$. Let $G \sim \mathcal{C}_{n,\mathbf{d}}$ and let $A, B \subseteq [n]$ be any (not necessarily disjoint) sets of vertices such that $2|A| < (1-\delta)n$. Then the random variable $e_G(A, B)$ is stochastically dominated by a random variable*

$$X \sim \mathrm{Bin}\left(\sum_{a\in A} d_a, |B|/((1-\delta)n - 2|A|)\right).$$

Proof. Let $t := \sum_{a\in A} d_a$. Let $\mathcal{X}$, A' and B' be the expanded sets of $[n]$, A and B, respectively. Label the points of $\mathcal{X}$ so that all the points in A' come first, that is, $A' = \{x_1, \ldots, x_t\}$. Generate a random configuration $M \sim \mathcal{C}^*_{n,\mathbf{d}}$ following this labelling. Then $e_G(A, B)$ is the number of pairings in M with one end-point in A' and the other in B', and we will estimate the probability that each pairing added to M contributes to $e_G(A, B)$.

First, note that all pairings added after M_t do not contribute to $e_G(A, B)$, as they do not have an end-point in A'. For each $i \in [t]$, define an indicator random variable X_i which takes value 1 if $M_i \neq M_{i-1}$ and $e = x_i y \in M_i \setminus M_{i-1}$ is such that $y \in B'$, and 0 otherwise, so that $e_G(A, B) = \sum_{i\in[t]} X_i$. Observe that, in the above process, the bound

$$\mathbb{P}[X_i = 1 \mid M_i \neq M_{i-1}] \leqslant \frac{|B|}{(1-\delta)n - 2|A|}$$

holds for all $i \in [t]$, since at every step of the process there are at most $|B|d$ points available in B' and at least $(1-\delta)nd - 2|A|d$ points available in $\mathcal{X} \setminus (V(M_{i-1}) \cup \{x_i\})$. On the other hand, $\mathbb{P}[X_i = 1 \mid M_i = M_{i-1}] = 0$, so given $M_0, M_1, \ldots, M_{i-1}$, each X_i is stochastically dominated by a Bernoulli random variable Y_i with parameter $|B|/((1-\delta)n - 2|A|)$. By summing over all $i \in [t]$, we conclude that $e_G(A, B)$ is stochastically dominated by $X \sim \mathrm{Bin}(t, |B|/((1-\delta)n - 2|A|))$. □

4. On the existence of a sparse 3-expander

Definition 4.1. An n-vertex graph G is called a 3-*expander* if it is connected and, for every $S \subseteq [n]$ with $|S| \leqslant n/400$, we have $|N_G(S)| \geqslant 3|S|$.

In order to give bounds on the distribution of edges in $G_{n,d}$, we will use an edge-switching technique first introduced by McKay and Wormald [29]. We consider the following switching.

Definition 4.2. Let $G = (V, E)$ and $G' = (V, E')$ be two multigraphs on the same vertex set such that $|E| = |E'|$. We write $G \sim G'$ if there exist $u_1u_2, v_1v_2 \in E$ such that $E' = (E \setminus \{u_1u_2, v_1v_2\}) \cup \{u_1v_1, u_2v_2\}$.

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

The following lemma bounds the probability that certain variables on configurations deviate from their expectation.

Lemma 4.1. *Let $\mathbf{d} = (d_1, \ldots, d_n)$ be a degree sequence with $d_i \leqslant \log^2 n$ for all $i \in [n]$, and such that $\sum_{i=1}^n d_i$ is even. Let $\Delta := \max_{i\in[n]}\{d_i\}$. Let $c > 0$ and let X be a random variable on $\mathcal{C}^*_{n,\mathbf{d}}$ such that, for every pair of configurations $M \sim M'$, we have $|X(M) - X(M')| \leqslant c$. Then, for all $\varepsilon > 0$,*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geqslant \varepsilon\mathbb{E}[X]] \leqslant 2\exp\left(-\frac{\varepsilon^2\mathbb{E}[X]^2}{2\Delta nc^2}\right).$$

Proof. Let $m := \sum_{i=1}^n d_i$. Fix any labelling $x_1, \ldots, x_m$ of the expanded set of $[n]$. Let $M \sim \mathcal{C}^*_{n,\mathbf{d}}$ be generated following this labelling. Let the partial configurations of M be $M_0, \ldots, M_m$. For each $i \in [m] \cup \{0\}$, let

$$Y_i(M) := \mathbb{E}[X(M) \mid M_i] = \mathbb{E}[X(M) \mid M_0, \ldots, M_i].$$

It follows that the sequence $Y_0(M), Y_1(M), \ldots, Y_m(M)$ is a Doob martingale, where $Y_0(M) = \mathbb{E}[X]$ and $Y_m(M) = X(M)$. We will now show that the differences of this martingale are bounded by c.

For any $i \in [m]$, if $M_i = M_{i-1}$, then $Y_i(M) = Y_{i-1}(M)$ and there is nothing to prove, so assume that $M_i \neq M_{i-1}$, that is, when generating the ith partial configuration, the ith point does not lie in any of the previous pairings. For each $j \in [m] \setminus (V(M_{i-1}) \cup \{i\})$, let $\mathcal{M}_j$ be the set of configurations which contain M_{i-1} as well as ij. It is easy to see that, for each $k \in [m] \setminus (V(M_{i-1}) \cup \{i\})$, there is a bijection $g_{j,k}$ between $\mathcal{M}_j$ and $\mathcal{M}_k$ so that $g_{j,k}(M') \sim M'$ for all $M' \in \mathcal{M}_j$. Fix $j \in [m] \setminus (V(M_{i-1}) \cup \{i\})$, let $N := |\mathcal{M}_j|$ and label the configurations in $\mathcal{M}_j$ as $M_{j,1}, \ldots, M_{j,N}$. For all $k \in [m] \setminus (V(M_{i-1}) \cup \{i, j\})$, label $\mathcal{M}_k$ by $M_{k,\ell} := g_{j,k}(M_{j,\ell})$ for each $\ell \in [N]$. By assumption, we have $|X(M_{j,\ell}) - X(M_{k,\ell})| \leqslant c$ for all distinct $j, k \in [m] \setminus (V(M_{i-1}) \cup \{i\})$ and $\ell \in [N]$. Using this, it is easy to conclude that $|Y_i(M) - Y_{i-1}(M)| \leqslant c$.

The statement now follows by Lemma 3.3. □

The following proposition implies that the distribution of edges in $G_{n,d}$ behaves roughly as in a binomial random graph $G_{n,d/n}$, even after conditioning on the containment of some 'sparse' subgraph.

Proposition 4.2. *For every $0 < \varepsilon \leqslant 1/2$ there exists $\delta > 0$ such that the following holds. Let $d \leqslant \log^2 n$ be a positive integer and let $G = G_{n,d}$. Let R be a graph on vertex set $[n]$ with $\Delta(R) < \delta d$. Moreover, let $A \subseteq [n]$ and, for each $a \in A$, let $Z_a \subseteq [n]^{(2)} \setminus E(R)$ be a collection of edges incident to a such that $z := \sum_{a\in A} |Z_a|$ satisfies $z > \varepsilon n^2$. Then*

$$\mathbb{P}\left[\left|\sum_{a\in A} |Z_a \cap E(G)| - \frac{zd}{n}\right| \geqslant \varepsilon\frac{zd}{n} \mid R \subseteq G\right] \leqslant e^{-(\varepsilon/10)^4 nd}.$$

Proof. Let $0 < \delta \ll \varepsilon$. For each $i \in [n]$, let $d_i := d - d_R(i) > (1-\delta)d$, and let $\mathbf{d} := (d_1, \ldots, d_n)$. Let $M \sim \mathcal{C}^*_{n,\mathbf{d}}$ and let $F = \varphi(M)$, so that $F \sim \mathcal{C}_{n,\mathbf{d}}$ and $R + F$ is a d-regular multigraph. By Lemma 3.6, for each $a \in A$, the random variable $Y_a \sim \mathrm{Bin}(d_a, (n - |Z_a|)/((1-\delta)n - 2))$ stochastically dominates $e_F(a, [n] \setminus (V(Z_a) \setminus \{a\}))$. Let $Z(F) := \sum_{a\in A} |Z_a \cap E(F)|$.

Note that $\mathbb{E}[Y_a] < (1+\varepsilon^3)d_a(n - |Z_a|)/n$ for all $a \in A$. It then follows that

$$\mathbb{E}[|Z_a \cap E(F)|] \geqslant d_a - \mathbb{E}[Y_a] \geqslant (1+\varepsilon^3)|Z_a|d_a/n - \varepsilon^3 d_a.$$

Therefore we have $\mathbb{E}[Z(F)] \geqslant (1-\varepsilon^2)zd/n$. Now, note that $|Z(F) - Z(F')| \leqslant 8$ when $F \sim F'$. Let $Z' : \mathcal{C}^*_{n,\mathbf{d}} \to \mathbb{Z}$ be such that $Z'(M) = Z(F)$ whenever $\varphi(M) = F$. It follows that $|Z'(M) - Z'(M')| \leqslant 8$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

when $M \sim M'$. Moreover, $\mathbb{E}[Z'(M)] = \mathbb{E}[Z(F)]$. Therefore we can apply Lemma 4.1 to obtain

$$\mathbb{P}\left[Z'(M) \leqslant (1-\varepsilon)\frac{zd}{n}\right] \leqslant 2e^{-\varepsilon^4 nd/512}.$$

By definition, the same bound holds for $Z(F)$. It now follows from Proposition 3.5 that

$$\mathbb{P}\left[Z(F) \leqslant (1-\varepsilon)\frac{zd}{n} \mid R+F \text{ is simple}\right] \leqslant 2e^{3d^2} e^{-\varepsilon^4 nd/512}. \tag{4.1}$$

By a similar argument we can show that

$$\mathbb{P}\left[Z(F) \geqslant (1+\varepsilon)\frac{zd}{n} \mid R+F \text{ is simple}\right] \leqslant 2e^{3d^2} e^{-\varepsilon^4 nd/512}. \tag{4.2}$$

The result follows by combining (4.1) and (4.2). □

Lemma 4.3. *For every $0 < \delta < 10^{-5}$ there exists $D \in \mathbb{N}$ such that, for any $D < d \leqslant \log^2 n$, we have that a.a.s. the random graph $G_{n,d}$ satisfies the following properties.*

(i) *For every $S \subseteq [n]$ with $\delta^2 d \leqslant |S| \leqslant 5\delta^2 n$, we have $e_{G_{n,d}}(S) \leqslant \delta d|S|/25$.*
(ii) *For every $S \subseteq [n]$ with $5\delta^2 n \leqslant |S| \leqslant n/100$, we have $e_{G_{n,d}}(S) \leqslant d|S|/25$.*

Proof. Let $1/D \ll \delta$. For any $D \leqslant d \leqslant \log^2 n$, let $G \sim \mathcal{C}_{n,d}$. For each $S \subseteq [n]$ such that $\delta^2 d \leqslant |S| \leqslant 5\delta^2 n$ and any multigraph F on $[n]$, let $g(S, F)$ be the event that $e_F(S) \leqslant \delta d|S|/25$. It follows by Lemma 3.6 that the variable $e_G(S)$ is stochastically dominated by $Y \sim \text{Bin}(d|S|, 5|S|/(4n))$. We let $\hat{\mathbb{P}}$ denote the probability measure associated with the configuration model and let $\mathbb{P}$ be the measure associated with the space of (simple) d-regular graphs. Therefore, by Lemma 3.2 we have

$$\hat{\mathbb{P}}[\overline{g(S,G)}] \leqslant \hat{\mathbb{P}}[e_G(S) \geqslant \delta d|S|/25] = \hat{\mathbb{P}}[e_G(S) \geqslant (4\delta n/(125|S|))5d|S|^2/(4n)] < (|S|/en)^{2|S|}.$$

To see this last inequality, note that

$$\begin{aligned}
\left(\left(\frac{125e|S|}{4\delta n}\right)^{d\delta/50} \frac{en}{|S|}\right)^{2|S|} &= \left(\left(\frac{125e|S|}{4\delta n}\right)^{d\delta/50-1} \frac{125e|S|}{4\delta n}\frac{en}{|S|}\right)^{2|S|} \\
&\leqslant \left(\frac{125e^2}{4\delta}(1000\delta e)^{\delta d/50-1}\right)^{2|S|} \\
&< 1.
\end{aligned}$$

It follows by Proposition 3.5 that

$$\begin{aligned}
\mathbb{P}\left[\bigvee_{\substack{S\subseteq[n]\\ \delta^2 d\leqslant|S|\leqslant 5\delta^2 n}} \overline{g(S, G_{n,d})}\right] &= \hat{\mathbb{P}}\left[\bigvee_{\substack{S\subseteq[n]\\ \delta^2 d\leqslant|S|\leqslant 5\delta^2 n}} \overline{g(S,G)} \mid G \text{ is simple}\right] \\
&\leqslant e^{3d^2} \sum_{\substack{S\subseteq[n]\\ \delta^2 d\leqslant|S|\leqslant 5\delta^2 n}} \hat{\mathbb{P}}[\overline{g(S,G)}] \\
&\leqslant e^{3d^2} \sum_{i=\delta^2 d}^{5\delta^2 n} \binom{n}{i}\left(\frac{i}{en}\right)^{2i} \\
&= o(1).
\end{aligned}$$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

Thus property (i) in the statement holds with probability $1-o(1)$. Similarly, we can show that property (ii) also holds with probability $1-o(1)$. □

Proposition 4.4. *For every $0<\delta<10^{-5}$ there exists $D\in\mathbb{N}$ such that, for any $D<d\leqslant\log^2 n$, we have that a.a.s. the random graph $G=G_{n,d}$ satisfies the following properties.*

(i) *Let $R\subseteq G$ be a spanning subgraph with $\delta(R)>\delta d$. Then, for every $S\subseteq[n]$ with $|S|\leqslant\delta^2 n$, we have $|N_R(S)|\geqslant 3|S|$.*
(ii) *For every $S,S'\subseteq[n]$ with $\delta^2 n\leqslant|S|\leqslant|S'|\leqslant 3|S|\leqslant 3n/400$, we have $e_G(S,S')\leqslant d|S|/5$.*

Proof. Let $1/D\ll\delta$ and condition on the statement of Lemma 4.3 holding, which occurs a.a.s. We first prove (i). For each $S\subseteq[n]$ such that $|S|<\delta^2 d$, the fact that every vertex has degree at least δd ensures that $|N_R(S)|\geqslant\delta d>3\delta^2 d$. Now let $S\subseteq[n]$ with $\delta^2 d\leqslant|S|\leqslant\delta^2 n$. Suppose $|N_R(S)|<3|S|$. Let $Y\subseteq[n]$ be such that $|Y|=3|S|$ and $N_R(S)\subseteq Y$. We have by Lemma 4.3(i) that

$$4|S|\delta d/25\geqslant e_G(S\cup Y)\geqslant e_R(S\cup Y)\geqslant e_R(S,Y)\geqslant|S|\delta d-e_G(S)>|S|\delta d/2,$$

a contradiction. The result follows.

In order to prove (ii), let $S\subseteq[n]$ with $\delta^2 n\leqslant|S|\leqslant n/400$. Suppose there exists $S'\subseteq[n]$ with $|S|\leqslant|S'|\leqslant 3|S|$ and such that $e_G(S,S')>d|S|/5$. We have by Lemma 4.3 that

$$4|S|d/25\geqslant e_G(S\cup S')\geqslant e_G(S,S')>d|S|/5,$$

a contradiction. The result follows. □

Proposition 4.5. *For every $0<\delta<10^{-5}$ there exists $D\in\mathbb{N}$ such that, for any $D<d\leqslant\log^2 n$, we have that a.a.s. the random graph $G=G_{n,d}$ has the following property. Let $H\in\mathcal{H}_{1/2}(G)$ and let $G':=G\setminus H$. Then there exists a spanning graph $R\subseteq G'$ such that $\Delta(R)<\delta d$ and, for every $S\subseteq[n]$ with $|S|\leqslant n/400$, we have that $|N_R(S)|\geqslant 3|S|$.*

Proof. Let $1/D\ll\delta$ and let $\hat\delta:=\delta/8$. Condition on the event that the statements of Lemma 4.3 and Proposition 4.4 hold with $\hat\delta$ playing the role of δ, which happens a.a.s. Suppose G satisfies these events and $H\in\mathcal{H}_{1/2}(G)$, and let $G':=G\setminus H$. We now construct a suitable R for this G'. Consider a random subgraph R of G' where each edge is chosen independently and uniformly at random with probability $4\hat\delta$. Consider the following events.

($\mathcal{G}1$) For all $v\in[n]$, we have $\hat\delta d<d_R(v)<8\hat\delta d$.
($\mathcal{G}2$) For every $S\subseteq[n]$ with $|S|\leqslant n/400$, we have $|N_R(S)|\geqslant 3|S|$.

Note that, if both ($\mathcal{G}1$) and ($\mathcal{G}2$) hold, then R is a subgraph which satisfies the properties in the statement of the lemma.

For each $v\in[n]$, let $\mathcal{A}_v$ be the event that $d_R(v)\notin(\hat\delta d,8\hat\delta d)$. By Lemma 3.1, we have

$$\mathbb{P}[\mathcal{A}_v]<4e^{-\hat\delta d/6}$$

for all $v\in[n]$. Observe that G' is itself a dependency graph for $\{\mathcal{A}_v\}_{v\in[n]}$, and it has degree at most d. By Lemma 3.4, it follows that

$$\mathbb{P}[R\text{ satisfies }(\mathcal{G}1)]=\mathbb{P}\left[\bigwedge_{v\in[n]}\overline{\mathcal{A}_v}\right]\geqslant(1-12e^{-\hat\delta d/6})^n\geqslant 2^{-n}.$$

Next, for $S,S'\subseteq[n]$, let $g(S,S')$ be the event that $N_R(S)\subseteq S'$. Let ($\mathcal{G}3$) be the event that, for no pair of subsets $S,S'\subseteq[n]$ with $S'\subseteq N_{G'}(S)$ and $\hat\delta^2 n\leqslant|S|\leqslant|S'|\leqslant 3|S|\leqslant 3n/400$, the event $g(S,S')$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

occurs. We have by Proposition 4.4(ii) and Lemma 4.3 that

$$e_{G'}(S,[n]\setminus S') \geqslant d|S|/2 - e_{G'}(S,S') - e_{G'}(S) \geqslant d|S|/2 - d|S|/5 - d|S|/25 \geqslant d|S|/5.$$

Therefore we have

$$\mathbb{P}[g(S,S')] \leqslant (1-4\hat{\delta})^{d|S|/5} \leqslant e^{-4\hat{\delta}d|S|/5} \leqslant 2^{-4n}.$$

A union bound implies that

$$\mathbb{P}[R \text{ fails to satisfy } (\mathcal{G}3)] \leqslant 2^{2n}2^{-4n} < 2^{-n}.$$

Therefore there exists an instance of R which satisfies both ($\mathcal{G}1$) and ($\mathcal{G}3$) simultaneously. Furthermore, since R satisfies ($\mathcal{G}1$), it follows by Proposition 4.4(i) that, for every $S \subseteq [n]$ with $|S| \leqslant \hat{\delta}^2 n$, we have that $|N_R(S)| \geqslant 3|S|$. Combining this with ($\mathcal{G}3$) we see that R also satisfies ($\mathcal{G}2$). Thus R is a subgraph of the desired form. □

Proposition 4.6. *For every $\varepsilon > 0$ there exists $D > 0$ such that, for any $D < d \leqslant \log^2 n$, we have that a.a.s. the random graph $G = G_{n,d}$ has the following property. Let $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ and let $G' := G \setminus H$. Let $R \subseteq G'$ be a spanning graph such that, for every $S \subseteq [n]$ with $|S| \leqslant n/400$, we have $|N_R(S)| \geqslant 3|S|$. Then there exists a spanning 3-expander $R' \subseteq G'$ with $e(R') \leqslant e(R) + 400$.*

Proof. Let $1/D \ll \varepsilon$. We are first going to prove that a.a.s. G' is connected. Note that a.a.s., for any $A, B \subseteq [n]$ with $|A| = n/400$ and $|B| = (1/2 - \varepsilon/10)n$, we have $\sum_{a\in A} e_G(a,B) > (1/2-\varepsilon/5)|A|d$. Indeed, this follows by an application of Proposition 4.2 with $R := \varnothing$ and Z_a being the star with centre a whose leaves are all the vertices in $B \setminus \{a\}$. We now claim that, for any $A \subseteq [n]$ with $|A| \geqslant n/400$, we have

$$|N_{G'}(A)| \geqslant (1/2+\varepsilon/10)n. \tag{4.3}$$

To see this, note that if there exists $A \subseteq [n]$ with $|A| \geqslant n/400$ and $|N_{G'}(A)| < (1/2+\varepsilon/10)n$, then we may take subsets $A' \subseteq A$ with $|A'| = n/400$ and $B \subseteq [n]$ with $|B| = (1/2-\varepsilon/10)n$ such that $e_{G'}(A',B) = 0$. However, we have already noted that, for such A' and B, we have

$$\sum_{a\in A'} e_G(a,B) \geqslant (1/2-\varepsilon/5)|A'|d.$$

It follows that there exists $a \in A'$ with $e_G(a,B) > (1/2-\varepsilon/5)d$ and therefore $e_{G'}(a,B) > 0$. Thus no such A and B exist.

In particular, (4.3) implies that G' is connected. Indeed, assume that G' is not connected and let $A \subsetneq [n]$ be a (connected) component of size $|A| \leqslant n/2$. We must have that $|N_{G'}(A)| \leqslant |A|$, but (4.3) and the statement hypotheses imply that $|N_{G'}(A)| > |A|$, a contradiction.

Finally, note that R consists of at most 400 components, since each connected component has order at least $n/400$. Since G' is connected, we may choose a set $E \subseteq E(G')$ with $|E| \leqslant 400$ such that the graph $R' := ([n], E(R) \cup E)$ is connected, and thus is a spanning 3-expander. □

Lemma 4.7. *For every $\varepsilon > 0$ and $0 \leqslant \delta \leqslant 10^{-5}$, there exists $D > 0$ such that, for any $D < d \leqslant \log^2 n$, we have that a.a.s. the random graph $G = G_{n,d}$ has the following property. Let $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ and let $G' := G \setminus H$. Then there exists a spanning 3-expander $R \subseteq G'$ with $\Delta(R) < \delta d$.*

Proof. Let $1/D \ll \delta, \varepsilon$ and condition on the statements of Propositions 4.5 and 4.6 both holding with $\delta/2$ playing the role of δ, which happens a.a.s. By Proposition 4.5 we may find a spanning subgraph $R' \subseteq G'$ with $\Delta(R') < \delta d/2$ and such that, for all $S \subseteq [n]$ with $|S| \leqslant n/400$, we have $|N_{R'}(S)| \geqslant 3|S|$. Then, by Proposition 4.6, we may find a spanning 3-expander $R \subseteq G'$ with $\Delta(R) < \Delta(R') + 400 < \delta d$. □

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

5. Finding many boosters

The following proposition provides an upper bound on the expected number of 'thin' subgraphs that $G_{n,d}$ contains.

Proposition 5.1. *Let $1/n \ll 1/d, \delta \ll 1$, where $n, d \in \mathbb{N}$, and let $G = G_{n,d}$. Let $\mathcal{R}$ be a family of graphs on vertex set $[n]$ with $e(R) \leqslant \delta dn$ for all $R \in \mathcal{R}$. Then*

$$\sum_{R\in\mathcal{R}} \mathbb{P}[R \subseteq G] \leqslant e^{2\delta dn \log(1/\delta)}.$$

Proof. For each $R \in \mathcal{R}$, let X_R be an indicator random variable where $X_R(G) = 1$ if and only if $R \subseteq G$. Let $X_{\mathcal{R}} := \sum_{R\in\mathcal{R}} X_R$. Then $\mathbb{E}[X_{\mathcal{R}}] = \sum_{R\in\mathcal{R}} \mathbb{P}[R \subseteq G]$. Moreover, note that we always have

$$X_{\mathcal{R}} \leqslant \sum_{i=1}^{\delta dn} \binom{dn/2}{i} \leqslant e^{2\delta dn \log(1/\delta)},$$

and therefore

$$\sum_{R\in\mathcal{R}} \mathbb{P}[R \subseteq G] = \mathbb{E}[X_{\mathcal{R}}] \leqslant e^{2\delta dn \log(1/\delta)},$$

as desired. □

The following result can easily be proved using 'Pósa rotations' (see *e.g.* [24]).

Lemma 5.2. *Let R be a 3-expander and let P be a longest path in R, with end-point v. Then there exists a set $A \subseteq V(P)$ with $|A| > n/10^4$ such that, for each $a \in A$, there exists a path P_a in R with end-points v and a, and such that $V(P_a) = V(P)$.*

Definition 5.1 (booster)**.** Let H be a graph and let $E \subseteq V(H)^{(2)}$. Let $F := (V(H), E)$. We call E a *booster for H* if the graph $H + F$ contains a longer path than H does, or if $H + F$ is Hamiltonian.

We will often be interested in the case where E consists of a single edge $e \notin E(H)$. In this case we refer to e as a *booster* for H.

Given any path P with end-points u and v, assume an orientation on its edges (from u to v, say). Given any vertex $x \in V(P) \setminus \{v\}$, we call the vertex that follows x in this orientation its *successor*, and we denote this by $\mathrm{suc}_P(x)$.

Lemma 5.3. *For all $0 < \varepsilon < 1/10^5$ there exist $\delta, D > 0$ such that, for $D \leqslant d \leqslant \log^2 n$, the random graph $G = G_{n,d}$ a.a.s. satisfies the following.*

Let $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ and let $G' := G \setminus H$. Let $R \subseteq G'$ be a spanning 3-expander with $\Delta(R) \leqslant 2\delta d$, and let $S \subseteq [n]$ with $|S| \leqslant \delta n$. Then there exists a set $V_R \subseteq [n]$ with $|V_R| \geqslant n/10^4$ with the following property. For each $v \in V_R$, there exists a set $U_v \subseteq [n]$ with $|U_v| \geqslant (1/2 + \varepsilon/8)n$ such that, for each $u \in U_v$, there exists a set $E_{v,u}$ as follows:

(a) *$E_{v,u} \subseteq E((G' \setminus R)[[n] \setminus S])$ with $|E_{v,u}| \geqslant 50/(\varepsilon\delta)$,*
(b) *$\{uv, e\}$ is a booster for R for every $e \in E_{v,u}$,*
(c) *$E_{v,u_1} \cap E_{v,u_2} = \varnothing$ for all $u_1 \neq u_2$.*

Proof. Let $1/D \ll \delta \ll \varepsilon < 1/10^5$. Let $\mathcal{R}$ be the set of all n-vertex 3-expander graphs R on $[n]$ with $\Delta(R) \leqslant 2\delta d$. It follows by Lemma 5.2 that, for each $R \in \mathcal{R}$, there exists a set $V_R \subseteq [n]$ of size $|V_R| \geqslant n/10^4$ such that for every $v \in V_R$ there exists a longest path in R terminating at v.

For each $R \in \mathcal{R}$, $v \in V_R$ and $S \subseteq [n]$ with $|S| \leqslant \delta n$, let $f(R, S, v)$ be the event that, for every $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ such that $R \subseteq G'$, there exists a set of vertices $U_v \subseteq [n]$ with $|U_v| \geqslant (1/2 + \varepsilon/8)n$ and such that for each $u \in U_v$ there exists a set $E_{v,u}$ satisfying (a)–(c). With this definition, the probability p^* that the assertion in the lemma fails is bounded by

$$p^* \leqslant \sum_{S \subseteq [n]\colon |S| \leqslant \delta n} \sum_{R \in \mathcal{R}} \sum_{v \in V_R} \mathbb{P}[\overline{f(R, S, v)} \mid R \subseteq G]\, \mathbb{P}[R \subseteq G]. \tag{5.1}$$

For fixed $R \in \mathcal{R}$, $v \in V_R$ and $S \subseteq [n]$ with $|S| \leqslant \delta n$, we shall now estimate $\mathbb{P}[\overline{f(R, S, v)} \mid R \subseteq G]$. Let P be a longest path in R with end-point v. As R is a 3-expander, by Lemma 5.2 there must exist a set $A \subseteq V(P) \setminus S$ of size $|A| = \varepsilon n/20$ such that, for each $a \in A$, there is a longest path P_a in R starting at v and ending at a with $V(P_a) = V(P)$ (if there is more than one such path, fix one arbitrarily). Assume that each P_a is oriented from v to a. Let $B \coloneqq [n] \setminus (A \cup S \cup \{v\})$. For each $u \in B \cap V(P)$, let $X_u \coloneqq \{ab\colon a \in A, b \in B, u = \mathrm{suc}_{P_a}(b)\}$. Observe that $\{uv, ab\}$ is a booster for R for any $ab \in X_u$. Clearly $|X_u| \leqslant |A|$ and $X_u \cap X_{u'} = \varnothing$ for all distinct $u, u' \in B \cap V(P)$. Furthermore, for each $u \in B \setminus V(P)$, let $X_u \coloneqq \{au\colon a \in A\}$. Note that $au \in X_u$ is a booster since its inclusion would result in a longer path in R. We shall now show that, for most vertices $u \in B$, there is a 'large' set of boosters, that is, X_u is 'large'. We will then use this to show that many of these boosters must lie in $G' \setminus R$.

For every $a \in A$, there are at least $|V(P)| - 2|A| - 2|S| - 2$ vertices $b \in V(P)$ such that neither b nor its successor on P_a belong to $A \cup S \cup \{v\}$. It follows that

$$\left| \bigcup_{u \in B \cap V(P)} X_u \right| \geqslant |A|(|V(P)| - 2|A| - 2|S| - 2).$$

We also have

$$\left| \bigcup_{u \in B \setminus V(P)} X_u \right| = |A|(n - |V(P) \cup S|).$$

Therefore the following holds:

$$\begin{aligned} \left| \bigcup_{u \in B} X_u \right| &\geqslant |A|(|V(P)| - 2|A| - 2|S| - 2) + |A|(n - |V(P) \cup S|) \\ &\geqslant |A|(n - 2|A| - 3|S| - 2) \\ &\geqslant (1 - \varepsilon/9)|A|n. \end{aligned}$$

For each $u \in B$, let $Y_u \coloneqq X_u \setminus E(R)$. It follows that

$$\left| \bigcup_{u \in B} Y_u \right| \geqslant (1 - \varepsilon/9)|A|n - e(R) \geqslant (1 - \varepsilon/8)|A|n.$$

For each $a \in A$, let Z_a be the set of edges in $\bigcup_{u \in B} Y_u$ with a as an end-point. It is easy to see that

$$\sum_{a \in A} |Z_a| = \left| \bigcup_{a \in A} Z_a \right| = \left| \bigcup_{u \in B} Y_u \right| \geqslant (1 - \varepsilon/8)|A|n.$$

Now consider the following events.

($\mathcal{F}1$)

$$\left| \bigcup_{a \in A} (Z_a \cap E(G)) \right| \geqslant (1 - \varepsilon/4)|A|d.$$

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

($\mathcal{F}2$) For any $U \subseteq B$ with $|U| \leqslant (1/2+\varepsilon/8)n$ we have

$$\left|\bigcup_{u\in U} Y_u \cap E(G)\right| < (1/2+\varepsilon/4)|A|d.$$

From two applications of Proposition 4.2 we obtain

$$\mathbb{P}[\mathcal{F}_1 \wedge \mathcal{F}_2 \mid R \subseteq G] \geqslant 1 - e^{-(\varepsilon/500)^4 dn}.$$

To finish the proof we must show that if $\mathcal{F}_1 \wedge \mathcal{F}_2$ holds, then $f(R,S,v)$ also holds. Consider any $G \in \mathcal{G}_{n,d}$ which satisfies both $\mathcal{F}_1$ and $\mathcal{F}_2$ and such that $R \subseteq G$. Fix any $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ such that $R \subseteq G' = G \setminus H$. For each $u \in B$, let $E_u := Y_u \cap E(G')$. As we have seen above, for each $e \in E_u$, the set $\{uv, e\}$ is a booster for R. Furthermore, none of the end-vertices of e lies in S, by construction. Let $U \subseteq B$ be the set of vertices $u \in B$ for which $|E_u| \geqslant 50/(\varepsilon\delta)$. Observe that, by $\mathcal{F}_2$, if

$$\left|\bigcup_{u\in U} Y_u \cap E(G)\right| \geqslant (1/2+\varepsilon/4)|A|d,$$

then $|U| \geqslant (1/2+\varepsilon/8)n$. But

$$\begin{aligned}
\left|\bigcup_{u\in U} Y_u \cap E(G)\right| &\geqslant \left|\bigcup_{u\in U} E_u\right| \\
&= \left|\bigcup_{u\in B} Y_u \cap E(G)\right| - \left|\bigcup_{u\in B} Y_u \cap E(H)\right| - \sum_{u\in B\setminus U} |E_u| \\
&\geqslant \left|\bigcup_{a\in A} Z_a \cap E(G)\right| - \left|\bigcup_{a\in A} Z_a \cap E(H)\right| - \frac{50}{\varepsilon\delta}|B \setminus U| \\
&\overset{(\mathcal{F}1)}{\geqslant} \left(1-\frac{\varepsilon}{4}\right)|A|d - \sum_{a\in A}\left(\frac{1}{2}-\varepsilon\right)d_G(a) - \frac{50}{\varepsilon\delta}n \\
&\geqslant \left(1-\frac{\varepsilon}{4}\right)|A|d - \left(\frac{1}{2}-\varepsilon\right)|A|d - 10^3\frac{|A|d}{\varepsilon^2 d\delta} \\
&\geqslant \left(\frac{1}{2}+\frac{\varepsilon}{4}\right)|A|d.
\end{aligned}$$

Hence, by $\mathcal{F}_2$ we have that $|U| \geqslant (1/2+\varepsilon/8)n$, as we wanted to show. Since H was arbitrary, it follows that $f(R,S,v)$ holds. Thus

$$\mathbb{P}[f(R,S,v) \mid R \subseteq G] \geqslant \mathbb{P}[\mathcal{F}_1 \wedge \mathcal{F}_2 \mid R \subseteq G] \geqslant 1 - e^{-(\varepsilon/500)^4 dn}.$$

We can now use this bound in equation (5.1) to obtain

$$p^* \leqslant 2^n n e^{-(\varepsilon/500)^4 dn} \sum_{R\in\mathcal{R}} \mathbb{P}[R \subseteq G] \leqslant 2^n n e^{-(\varepsilon/500)^4 dn} e^{2\delta dn \log(1/\delta)} = o(1),$$

where the second inequality follows from Proposition 5.1. This shows that the statement in the lemma holds a.a.s. □

Definition 5.2. Given graphs H and H' with $V(H) = V(H') = V$ and $E(H) \cap E(H') = \varnothing$, we say H has *ε-many boosters with help from H'* if there are at least $\varepsilon|V|$ vertices $v \in V$ for which there exists a set $U_v \subseteq V \setminus \{v\}$ of size at least $(1/2+\varepsilon)|V|$ with the property that for every $u \in U_v$ there exists $e \in E(H')$ so that $\{uv, e\}$ is a booster for H. We call uv the *primary edge* and we call e the *secondary edge*.

Corollary 5.4. *For all $0 < \varepsilon < 1/10^5$ there exist $\delta, D > 0$ such that, for $D \leqslant d \leqslant \log^2 n$, the random graph $G = G_{n,d}$ a.a.s. satisfies the following.*

Let $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ and let $G' := G \setminus H$. Let $R \subseteq G'$ be a spanning 3-expander with $\Delta(R) \leqslant 2\delta d$, and let $S \subseteq [n]$ with $|S| \leqslant \delta n$. Then there exists some subgraph $F \subseteq G' \setminus R$ satisfying $\Delta(F) \leqslant 2\delta d$, such that R has $\varepsilon/16$-many boosters with help from F, with the property that the set of secondary edges is vertex-disjoint from S.

Proof. Let $1/D \ll \delta \ll \varepsilon < 1/10^5$. Condition on the event that G satisfies all the properties in the statement of Lemma 5.3, which happens a.a.s. Let H, G', R, S be as in the statement of Corollary 5.4. By Lemma 5.3, we may find a set $V_R \subseteq [n]$ of size $|V_R| \geqslant n/10^4$ such that, for each $v \in V_R$, there exists a set $U_v \subseteq [n]$ with $|U_v| \geqslant (1/2 + \varepsilon/8)n$ such that, for each $u \in U_v$, there exists a set $E_{v,u} \subseteq E((G' \setminus R)[[n] \setminus S])$ with $|E_{v,u}| \geqslant 50/(\varepsilon\delta)$ and such that, for every $e \in E_{v,u}$, $\{uv, e\}$ is a booster for R, and such that $E_{v,u_1} \cap E_{v,u_2} = \varnothing$ for all $u_1 \neq u_2$. Note that each such edge e is vertex-disjoint from S, by construction.

Let F be a random subgraph of $G' \setminus R$ where every edge in $G' \setminus R$ is chosen independently at random with probability $\delta/2$. For each $v \in V_R$, let $U'_v \subseteq U_v$ be the set of vertices $u \in U_v$ for which $E_{v,u} \cap E(F) \neq \varnothing$. For every $u \in U_v$, we have

$$\mathbb{P}[u \notin U'_v] \leqslant (1 - \delta/2)^{50/(\varepsilon\delta)} \leqslant e^{-25/\varepsilon} \leqslant \varepsilon/32.$$

Let $\mathcal{A}$ be the event that $|U'_v| \geqslant (1/2 + \varepsilon/16)n$ for every $v \in V_R$. Since for different $u \in U_v$ the sets $E_{v,u}$ are disjoint, by Lemma 3.1 we have

$$\mathbb{P}[|U'_v| \leqslant (1/2 + \varepsilon/16)n] \leqslant \mathbb{P}[|U'_v| \leqslant (1 - \varepsilon/16)|U_v|] \leqslant e^{-\varepsilon^2 n/10^6}$$

for each $v \in V_R$. Therefore

$$\mathbb{P}[\overline{\mathcal{A}}] \leqslant n e^{-\varepsilon^2 n/10^6} \leqslant e^{-\varepsilon^3 n}.$$

Now, let $\mathcal{B}$ be the event that $\Delta(F) \leqslant 2\delta d$. For each $v \in [n]$, let $\mathcal{B}_v$ be the event that $d_F(v) > 2\delta d$. By Lemma 3.2, we have

$$\mathbb{P}[B_v] < e^{-\delta d/8}$$

for all $v \in [n]$. Now observe that $G' \setminus R$ is itself a dependency graph for $\{\mathcal{B}_v\}_{v \in [n]}$, and every vertex in this graph has degree at most d. It follows by Lemma 3.4 that

$$\mathbb{P}[\mathcal{B}] = \mathbb{P}\left[\bigwedge_{v \in [n]} \overline{\mathcal{B}_v}\right] \geqslant (1 - e^{1-\delta d/8})^n \geqslant e^{-\varepsilon^4 n} > \mathbb{P}[\overline{\mathcal{A}}].$$

Therefore the probability that both events $\mathcal{A}$ and $\mathcal{B}$ occur is strictly positive, implying that there exists some $F \subseteq G' \setminus R$ satisfying the required properties. □

We have now shown that a.a.s. if the random graph $G_{n,d}$ contains a sparse 3-expander subgraph R after deleting some $H \in \mathcal{H}_{1/2-\varepsilon}(G_{n,d})$, then $G' = G_{n,d} \setminus H$ must also have a sparse subgraph F with the property that R has 'many' boosters with help from F. Our next goal is to prove that some primary edge of these boosters must actually be present in G'.

Lemma 5.5. *For all $0 < \varepsilon < 1/10^5$ there exist $\delta, D > 0$ such that, for $D \leqslant d \leqslant \log^2 n$, the random graph $G = G_{n,d}$ satisfies the following a.a.s. Let $S \subseteq [n]$ with $|S| \leqslant \delta n$ and let $R, F \subseteq G$ be two spanning edge-disjoint subgraphs such that*

(P1) $\Delta(R), \Delta(F) \leqslant 2\delta d$,

(P2) *R has ε-many boosters with help from F, such that every secondary edge is vertex-disjoint from S.*

Then, for any $H \in \mathcal{H}_{1/2}(G)$, the graph $G' := G \setminus H$ contains an edge e for which there exists some edge $e' \in E(F)$ with the property that $\{e, e'\}$ is a booster for R, and such that $V(\{e, e'\}) \cap S = \varnothing$.

Proof. Let $1/D \ll \delta \ll \varepsilon < 1/10^5$. Let $\mathcal{P}$ be the set of all triples (R, F, S) where R and F are edge-disjoint graphs on $[n]$ which satisfy (P1) and (P2) and $S \subseteq [n]$ with $|S| \leqslant \delta n$.

Fix a triple $(R, F, S) \in \mathcal{P}$. For every $x \in [n]$, let V_x be the set of vertices $v \in [n] \setminus (S \cup \{x\})$ for which there exists some edge $e \in E(F)$ such that none of the end-vertices of e lies in S and $\{xv, e\}$ is a booster for R. Let

$$X' := \{x \in [n] : |V_x| \geqslant (1/2 + 3\varepsilon/4)n\}.$$

By assumption on the triple (R, F, S) and using Definition 5.2, we must have that $|X'| \geqslant \varepsilon n$.

Let $X := X' \setminus S$. Let $f(R, F, S)$ be the event that

$$\sum_{x \in X} e_{G\setminus(R+F)}(x, V_x) \geqslant (1+\varepsilon)d|X|/2.$$

It follows by Proposition 4.2 that

$$\mathbb{P}[\overline{f(R, F, S)} \mid R + F \subseteq G] \leqslant e^{-(\varepsilon/30)^4 dn}. \tag{5.2}$$

It follows that the probability that $(R + F \subseteq G) \wedge \overline{f(R, F, S)}$ for some triple $(R, F, S) \in \mathcal{P}$ is at most

$$\begin{aligned}\sum_{(R,F,S)\in\mathcal{P}} \mathbb{P}[\overline{f(R, F, S)} \mid R + F \subseteq G]\, \mathbb{P}[R + F \subseteq G] &\overset{(5.2)}{\leqslant} 2^n e^{-(\varepsilon/30)^4 dn} \sum_{\substack{K\subseteq K_n\\ e(K)\leqslant 2\delta dn}} \mathbb{P}[K \subseteq G] \\ &\leqslant 2^n e^{-(\varepsilon/30)^4 dn} e^{4\delta dn \log 1/(2\delta)} \\ &\leqslant e^{-(\varepsilon/50)^4 dn},\end{aligned}$$

where the second inequality follows by Proposition 5.1.

We conclude that a.a.s. for all $(R, F, S) \in \mathcal{P}$ with $R + F \subseteq G$ we have

$$\sum_{x\in X}\left(e_{G\setminus(R+F)}(x, V_x) - \frac{1}{2}d_G(x)\right) \geqslant (1+\varepsilon)\frac{d|X|}{2} - \frac{d|X|}{2} > 0.$$

Hence there must exist some $x \in X$ with $e_{G\setminus(R+F)}(x, V_x) > d/2$. Therefore, for any $H \in \mathcal{H}_{1/2}(G)$, there is some vertex $x \in X$ such that $e_{G'}(x, V_x) \geqslant e_G(x, V_x) - d_H(x) > 0$. That is, there must be some $v \in N_{G'}(x) \cap V_x$. By the definition of V_x, there is some $e \in E(F)$ such that $\{xv, e\}$ is a booster for R. Furthermore, by construction, we have $V(\{xv, e\}) \cap S = \varnothing$, and this completes the proof of the lemma. □

Armed with the previous lemmas, we are now in a position to complete the proof of Theorem 1.2.

Proof of Theorem 1.2. Let $1/D \ll \delta \ll \varepsilon < 1/10^5$ be such that Corollary 5.4 holds for ε, Lemma 5.5 holds for $\varepsilon/16$ and Lemma 4.7 holds for ε. Condition on each of these holding.

Let $H \in \mathcal{H}_{1/2-\varepsilon}(G)$ and let $G' := G \setminus H$. By Lemma 4.7, there exists a subgraph $R \subseteq G'$ which is a spanning 3-expander with $\Delta(R) \leqslant \delta d$.

Let $R_0 := R$. We now proceed recursively as follows: for each $i \in [n]$, choose $e_{i,1}, e_{i,2} \in E(G')$ such that $\{e_{i,1}, e_{i,2}\}$ is a booster for R_{i-1}, and let $R_i := R_{i-1} + e_{i,1} + e_{i,2}$. In order to show that there exist such $e_{i,1}, e_{i,2}$ for all $i \in [n]$, consider the following. Assume that R_{i-1} satisfies $\Delta(R_{i-1}) \leqslant 2\delta d$. Let $S_i \subseteq [n]$ be the set of vertices $v \in [n]$ with $d_{R_{i-1}}(v) \geqslant 2\delta d - 1$. For all $i \in [n]$ we have $|S_i| \leqslant 2e(R_{i-1} \setminus R_0)/(\delta d - 1) \leqslant 4n/(\delta d - 1) < \delta n$. By applying Corollary 5.4 with S_i, R_{i-1} playing the roles of S and R, respectively, there exists some subgraph $F_i \subseteq G' \setminus R_{i-1}$ such that R_{i-1} has $(\varepsilon/16)$-many

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

boosters with help from F_i, where each secondary edge is vertex-disjoint from S_i. Furthermore, we have $\Delta(F_i) \leqslant 2\delta d$. Therefore, by applying Lemma 5.5, there are some $e_{i,1}, e_{i,2} \in E(G')$ such that $\{e_{i,1}, e_{i,2}\}$ is a booster for R_{i-1} and where $V(\{e_{i,1}, e_{i,2}\}) \cap S_i = \varnothing$. It follows that $\Delta(R_i) \leqslant 2\delta d$.

By the end of this process, we have added n boosters to R to obtain $R_n \subseteq G'$. Therefore R_n, and hence G', is Hamiltonian. □

6. Graphs of small degree with low resilience

In this section we prove Theorem 1.3. For this, we will require a crude bound on the number of edges spanned by any set of $n/2$ vertices in $G_{n,d}$. To achieve this, we shall make use of the following result, which follows from a theorem of McKay [28] (see *e.g.* [38]). We let $\alpha(G)$ denote the size of a maximum independent set in G.

Theorem 6.1. *For every fixed $d \geqslant 3$, a.a.s. we have that $\alpha(G_{n,d}) \leqslant 0.46n$.*

Lemma 6.1. *For every fixed $d \geqslant 3$, a.a.s. we have that $e_{G_{n,d}}(A) > n/100$ for all $A \subseteq [n]$ with $|A| = \lfloor n/2 \rfloor$.*

Proof. By Theorem 6.1, every set of size $n/2$ must span at least $n/100$ edges, as otherwise it would contain an independent set of size $n/2 - n/50 > 0.46n$. □

Alternatively, this lemma can be proved directly using a switching argument.

In order to prove Theorem 1.3 we will use a switching argument. Given a graph $G \in \mathcal{G}_{n,d}$ and any integer $\ell \in [d]$, let $u \in [n]$ and let $\Lambda^+_{u,\ell} = (e_1, \ldots, e_\ell, f_1, \ldots, f_\ell)$ be an ordered set of 2ℓ edges from $E(G)$ such that $e_i = uv_i$ with $v_i \neq v_j$ for all $i \neq j$, and $\{f_i \colon i \in [\ell]\}$ is a set of independent edges such that, for each $i \in [\ell]$, the distance between f_i and e_i is at least 2. We call $\Lambda^+_{u,\ell}$ a (u, ℓ)-*switching configuration*. For each $i \in [\ell]$, choose an orientation of f_i and write $f_i = x_iy_i$, where f_i is oriented from x_i to y_i. Let

$$\lambda^+_{u,\ell} \coloneqq \{e_1, \ldots, e_\ell, f_1, \ldots, f_\ell\},$$

$$\Lambda^-_{u,\ell} \coloneqq (uy_1, \ldots, uy_\ell, x_1v_1, \ldots, x_\ell v_\ell),$$

$$\lambda^-_{u,\ell} \coloneqq \{uy_1, \ldots, uy_\ell, x_1v_1, \ldots, x_\ell v_\ell\}.$$

We say that the graph

$$G' \coloneqq ([n], (E(G) \setminus \lambda^+_{u,\ell}) \cup \lambda^-_{u,\ell}) \in \mathcal{G}_{n,d}$$

is obtained from G by a *u-switching of type ℓ*. Observe that, given such a setting, we also have that G is obtained from G' by a u-switching of type ℓ, interchanging the roles of $\Lambda^+_{u,\ell}$ and $\Lambda^-_{u,\ell}$.

Proof of Theorem 1.3. Fix any odd $d > 2$. Let $\hat{\mathcal{G}}_{n,d} \subseteq \mathcal{G}_{n,d}$ be the collection of graphs for which the statement of Lemma 6.1 holds. We have by Lemma 6.1 that $|\hat{\mathcal{G}}_{n,d}| = (1 - o(1))|\mathcal{G}_{n,d}|$. Let $\mathcal{G}'_{n,d} \subseteq \hat{\mathcal{G}}_{n,d}$ be the collection of all graphs $G \in \hat{\mathcal{G}}_{n,d}$ which are not $(d-1)/2$-resilient with respect to Hamiltonicity. Let $p \coloneqq |\mathcal{G}'_{n,d}|/|\hat{\mathcal{G}}_{n,d}|$. We will prove that p is bounded from below by a positive constant which does not depend on n.

For each $G \in \hat{\mathcal{G}}_{n,d}$, consider a maximum cut M with parts A_M and B_M, where $|A_M| \leqslant |B_M|$ (thus $M = E_G(A_M, B_M)$). By abusing notation, we also use M to denote the bipartite graph $G[A_M, B_M]$. Note that for all $x \in A_M$ we have $d_M(x) > d/2$, as otherwise we could move x from A_G to B_G to obtain a larger cut; similarly, $d_M(y) > d/2$ for all $y \in B_M$.

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

Given $G \in \hat{\mathcal{G}}_{n,d}$, suppose there exists a maximum cut M for G such that $|A_M| < |B_M|$. Let

$$H \coloneqq ([n], E_G(A_M) \cup E_G(B_M)).$$

It is then clear that $M = G \setminus H$ is not Hamiltonian, as it is an unbalanced bipartite graph. Furthermore, we have that $\Delta(H) \leqslant (d-1)/2$, so we conclude that G is not $(d-1)/2$-resilient with respect to Hamiltonicity and thus $G \in \mathcal{G}'_{n,d}$. (Below we will use that the same conclusion holds if there is *any* cut M of G such that $|A_M| < |B_M|$, $d_M(x) > d/2$ for all $x \in A_M$, and $d_M(y) > d/2$ for all $y \in B_M$.) Therefore, for every $G \in \hat{\mathcal{G}}_{n,d} \setminus \mathcal{G}'_{n,d}$, we have that $|A_M| = |B_M|$ for every maximum cut M of G.

For each $G \in \hat{\mathcal{G}}_{n,d} \setminus \mathcal{G}'_{n,d}$, fix a maximum cut M_G of G which partitions $[n]$ into A_G and B_G. Then, for each $x \in A_G$, there exists $k \in [\lceil d/2 \rceil]$ such that $d_{M_G}(x) = \lfloor d/2 \rfloor + k$. Let $\ell \in [\lceil d/2 \rceil]$ be such that there exist at least $(1-p)|\hat{\mathcal{G}}_{n,d}|/d$ graphs $G \in \hat{\mathcal{G}}_{n,d} \setminus \mathcal{G}'_{n,d}$ with the property that there are at least $n/(2d)$ vertices $x \in A_G$ with $d_{M_G}(x) = \lfloor d/2 \rfloor + \ell$. Let $D \coloneqq \lfloor d/2 \rfloor + \ell$. Denote the collection of all such graphs G by Ω.

For each $G \in \Omega$ and for each $x \in A_G$ such that $d_{M_G}(x) = D$, we consider all possible x-switchings of type D where the (x, D)-switching configuration $\Lambda^+_{x,D} = (e_1, \ldots, e_D, f_1, \ldots, f_D)$ satisfies that $\{e_1, \ldots, e_D\} = E_{M_G}(x, B_G)$ and $\{f_1, \ldots, f_D\} \subseteq E_G(A_G)$. We say that any $G' \in \mathcal{G}_{n,d}$ that can be obtained from G by such an x-switching of type D is obtained by an *out-switching* from G, and we call $\Lambda^+_{x,D}$ an *out-switching configuration*. Let Ω' denote the set of all graphs $G' \in \mathcal{G}_{n,d}$ that can be obtained by out-switchings from some graph $G \in \Omega$. In particular, note that for each G' obtained from G by an out-switching, we may define $A' \coloneqq A_G \setminus \{x\}$ and $B' \coloneqq [n] \setminus A'$, so that $|A'| < |B'|$, $e_{G'}(u, B') > d/2$ for all $u \in A'$, and $e_{G'}(v, A') > d/2$ for all $v \in B'$, which means, as observed previously, that G' is not $(d-1)/2$-resilient with respect to Hamiltonicity. Therefore $\Omega' \subseteq (\mathcal{G}_{n,d} \setminus \hat{\mathcal{G}}_{n,d}) \cup \mathcal{G}'_{n,d}$ and $\Omega \cap \Omega' = \varnothing$.

To show that Ω' is large, we consider an auxiliary bipartite graph Γ with parts Ω and Ω'. We place an edge between $G \in \Omega$ and $G' \in \Omega'$ if G' is obtained from G by an out-switching. First let $G \in \Omega$. We will now provide a lower bound on the number of out-switchings from G. Since $G \in \Omega$, by construction there are at least $n/(2d)$ vertices $x \in A_G$ such that $d_{M_G}(x) = D$. For each such x, the number of out-switching configurations is given by the different choices for the edges in $(e_1, \ldots, e_D, f_1, \ldots, f_D)$, chosen sequentially. There are $D!$ choices for $(e_1, \ldots, e_D)$. For all $i \in [D]$, as each of the f_i has to be independent from the previously chosen edges, at distance at least 2 from e_i, and spanned by A_G, by Lemma 6.1 we conclude that the number of choices for f_i is at least $n/100 - 4d^2$ Finally, once the out-switching configuration is given, there are 2^D possible switchings, one for each possible orientation of the set of edges $\{f_i \colon i \in [D]\}$; on the other hand, $D!$ different out-switching configurations result in the same outcome G'. We conclude that

$$d_\Gamma(G) \geqslant \frac{n}{2d} 2^D \left(\frac{n}{100} - 4d^2 \right)^D. \tag{6.1}$$

Now consider any $G' \in \Omega'$. It is easy to see that

$$d_\Gamma(G') \leqslant n 2^D \binom{d}{D} \left(\frac{nd}{2} \right)^D. \tag{6.2}$$

Therefore, by double-counting the edges in Γ, from (6.1) and (6.2), we have

$$|\Omega| \leqslant 2d \binom{d}{D} \left(\frac{101d}{2} \right)^D |\Omega'|.$$

It follows that there exists a constant p which does not depend on n, for which a p fraction of the graphs in $\hat{\mathcal{G}}_{n,d}$ are not $(d-1)/2$-resilient with respect to Hamiltonicity. The result follows. □

https://doi.org/10.1017/S0963548320000346 Published online by Cambridge University Press

References

[1] Ajtai, M., Komlós, J. and Szemerédi, E. (1985) First occurrence of Hamilton cycles in random graphs. In *Cycles in Graphs (Burnaby, BC, 1982)*, Vol. 115 of North-Holland Mathematics Studies, pp. 173–178. North-Holland.

[2] Allen, P., Böttcher, J., Ehrenmüller, J. and Taraz, A. (2020) The bandwidth theorem in sparse graphs. *Adv. Combin.* **2020**. doi:10.19086/aic.12849

[3] Alon, N. and Spencer, J. H. (2016) *The Probabilistic Method*, Wiley Series in Discrete Mathematics and Optimization. Wiley.

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

[5] Balogh, J., Csaba, B. and Samotij, W. (2011) Local resilience of almost spanning trees in random graphs. *Random Struct. Algorithms* **38** 121–139.

[6] Balogh, J., Lee, C. and Samotij, W. (2012) Corrádi and Hajnal's theorem for sparse random graphs. *Combin. Probab. Comput.* **21** 23–55.

[7] Ben-Shimon, S., Krivelevich, M. and Sudakov, B. (2011) Local resilience and Hamiltonicity maker-breaker games in random regular graphs. *Combin. Probab. Comput.* **20** 173–211.

[8] Ben-Shimon, S., Krivelevich, M. and Sudakov, B. (2011) On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs. *SIAM J. Discrete Math.* **25** 1176–1193.

[9] Bollobás, B. (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316.

[10] Bollobás, B. (1984) The evolution of sparse graphs. In *Graph Theory and Combinatorics (Cambridge, 1983)*, pp. 35–57. Academic Press.

[11] Condon, P., Espuny Díaz, A., Girão, A., Kühn, D. and Osthus, D. (2019) Dirac's theorem for random regular graphs. arXiv:1903.05052

[12] Condon, P., Espuny Díaz, A., Kim, J., Kühn, D. and Osthus, D. (2019) Resilient degree sequences with respect to Hamilton cycles and matchings in random graphs. *Electron. J. Combin.* **26** 4.54.

[13] Cook, N., Goldstein, L. and Johnson, T. (2018) Size biased couplings and the spectral gap for random regular graphs. *Ann. Probab.* **46** 72–125.

[14] Cooper, C., Frieze, A. and Reed, B. (2002) Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combin. Probab. Comput.* **11** 249–261.

[15] Ferber, A., Nenadov, R., Noever, A., Peter, U. and Škorić, N. (2017) Robust Hamiltonicity of random directed graphs. *J. Combin. Theory Ser. B* **126** 1–23.

[16] Frieze, A. (2019) Hamilton cycles in random graphs: a bibliography. arXiv:1901.07139

[17] Frieze, A. and Krivelevich, M. (2008) On two Hamilton cycle problems in random graphs. *Israel J. Math.* **166** 221–234.

[18] Hefetz, D., Steger, A. and Sudakov, B. (2016) Random directed graphs are robustly Hamiltonian. *Random Struct. Algorithms* **49** 345–362.

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

[20] Huang, H., Lee, C. and Sudakov, B. (2012) Bandwidth theorem for random graphs. *J. Combin. Theory Ser. B* **102** 14–37.

[21] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience.

[22] Kim, J. H. and Vu, V. H. (2004) Sandwiching random graphs: universality between random graph models. *Adv. Math.* **188** 444–469.

[23] Komlós, J. and Szemerédi, E. (1983) Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.* **43** 55–63.

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

[25] Krivelevich, M., Lee, C. and Sudakov, B. (2010) Resilient pancyclicity of random and pseudorandom graphs. *SIAM J. Discrete Math.* **24** 1–16.

[26] Krivelevich, M., Sudakov, B., Vu, V. H. and Wormald, N. C. (2001) Random regular graphs of high degree. *Random Struct. Algorithms* **18** 346–363.

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

[28] McKay, B. D. (1987) Independent sets in regular graphs of high girth. *Ars Combin.* **23** 179–185.

[29] McKay, B. D. and Wormald, N. C. (1990) Uniform generation of random regular graphs of moderate degree. *J. Algorithms* **11** 52–67.

[30] Montgomery, R. (2019) Hamiltonicity in random graphs is born resilient. *J. Combin. Theory Ser. B* **139** 316–341.

[31] Montgomery, R. (2020) Hamiltonicity in random directed graphs is born resilient. *Combin. Probab. Comput.* doi:10.1017/S0963548320000140.

[32] Nenadov, R., Steger, A. and Trujić, M. (2019) Resilience of perfect matchings and Hamiltonicity in random graph processes. *Random Struct. Algorithms* **54** 797–819.

[33] Robinson, R. W. and Wormald, N. C. (1992) Almost all cubic graphs are Hamiltonian. *Random Struct. Algorithms* **3** 117–125.
[34] Robinson, R. W. and Wormald, N. C. (1994) Almost all regular graphs are Hamiltonian. *Random Struct. Algorithms* **5** 363–374.
[35] Škorić, N., Steger, A. and Trujić, M. (2018) Local resilience of an almost spanning k-cycle in random graphs. *Random Struct. Algorithms* **53** 728–751.
[36] Sudakov, B. and Vu, V. H. (2008) Local resilience of graphs. *Random Struct. Algorithms* **33** 409–433.
[37] Tikhomirov, K. and Youssef, P. (2019) The spectral gap of dense random regular graphs. *Ann. Probab.* **47** 362–419.
[38] Wormald, N. C. (1999) Models of random regular graphs. In *Surveys in Combinatorics (Canterbury, 1999)*, Vol. 267 of London Mathematical Society Lecture Note Series, pp. 239–298. Cambridge University Press.

Cite this article: Condon P, Espuny Díaz A, Girão A, Kühn D and Osthus D (2021). Dirac's theorem for random regular graphs. *Combinatorics, Probability and Computing* **30**, 17–36. https://doi.org/10.1017/S0963548320000346