Crux and long cycles in graphs

John Haslegrave* Jie Hu† Jaehoon Kim ‡ Hong Liu* Bingyu Luan§
Guanghui Wang§

5th July 2021

Abstract

We introduce a notion of the *crux* of a graph G, measuring the order of a smallest dense subgraph in G. This simple-looking notion leads to some generalisations of known results about cycles, offering an interesting paradigm of 'replacing average degree by crux'. In particular, we prove that *every* graph contains a cycle of length linear in its crux.

Long proved that every subgraph of a hypercube Q^m (resp. discrete torus C_3^m) with average degree d contains a path of length $2^{d/2}$ (resp. $2^{d/4}$), and conjectured that there should be a path of length $2^d - 1$ (resp. $3^{d/2} - 1$). As a corollary of our result, together with isoperimetric inequalities, we close these exponential gaps giving asymptotically optimal bounds on long paths in hypercubes, discrete tori, and more generally Hamming graphs.

We also consider random subgraphs of C_4-free graphs and hypercubes, proving near optimal bounds on lengths of long cycles.

1 Introduction

The study on the existence of long cycles in graphs has a rich history. A celebrated result of Dirac [8] states that every graph G on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$ contains a Hamiltonian cycle. However, any graph satisfying Dirac's condition is dense, having $\Theta(n^2)$ edges. A natural line of work is to consider how long a cycle we can ensure in a well-connected *sparse* graph.

1.1 Motivations

A folklore result on cycles finds in any cyclic graph G a cycle of length linear in its average degree, i.e. $\Omega(d(G))$. Indeed, remove low-degree vertices to obtain a subgraph H with $\delta(H) \geq d(G)/2$ and then greedily extend a path to find a cycle in H of length at least $\delta(H) + 1$. This linear in average degree bound is the best we could hope for, as the graph G might be disjoint union of cliques. It seems intuitive that better bounds can be obtained if we step away from such examples. This motivates the following notion of the *crux* of a graph; it measures the order of the smallest subgraph of G which retains a fraction of average degree of G.

*Mathematics Institute, University of Warwick, UK. Email: {j.haslegrave, h.liu.9}@warwick.ac.uk. J.H. and H.L. were supported by the UK Research and Innovation Future Leaders Fellowship MR/S016325/1.

†Laboratoire Interdisciplinaire des Sciences du Numérique, Université Paris-Saclay, France. Email: hujie@lri.fr.

‡Department of Mathematical Sciences, KAIST, South Korea. Email: jaehoon.kim@kaist.ac.kr. J.K. was supported by the POSCO Science Fellowship of POSCO TJ Park Foundation and by the KAIX Challenge program of KAIST Advanced Institute for Science-X.

§School of Mathematics, Shandong University, China. Email: byluan@mail.sdu.edu.cn, ghwang@sdu.edu.cn. B.L. and G.W. were supported by Natural Science Foundation of China (11871311, 11631014) and seed fund program for international research cooperation of Shandong University.

Definition 1.1 (Crux)**.** For a constant $\alpha \in (0,1)$, a subgraph $H \subseteq G$ is an α*-crux* if $d(H) \geq \alpha \cdot d(G)$. Define the α*-crux function*, $c_\alpha(G)$, of G to be the order of a minimum α-crux in G, that is,

$$c_\alpha(G) = \min\{|H| : H \subseteq G \text{ and } d(H) \geq \alpha \cdot d(G)\}.$$

Note that trivially we have $c_\alpha(G) > \alpha \cdot d(G)$, $c_\alpha(G) > c_{\alpha'}(G)$ for $\alpha > \alpha'$, and that if $H \subseteq G$ with $d(H) \geq d(G)/2$ then $c_{2\alpha}(H) \geq c_\alpha(G)$.

In this paper, we investigate the following 'replacing average degree by crux' paradigm.

Question A. Suppose we have a result guaranteeing the existence of a certain substructure whose size is a function of $d(G)$ (or $\delta(G)$). Under what circumstances can we replace $d(G)$ (or $\delta(G)$) with $c_\alpha(G)$?

Positive instances for the above question would lead to improvements on embedding problems for graph classes whose crux size is much larger than their average degree.

Example B. There are many natural classes of graphs having $c_\alpha(G)$ much larger than $d(G)$. Some specific classes are graphs with geometric structure, such as hypercubes Q^m and Hamming graphs $H(m,r)$, which are Cartesian products of m complete graphs K_r:

$$c_\alpha(Q^m) \geq 2^{\alpha m}, \qquad c_\alpha(H(m,r)) \geq r^{\alpha m};\text{[1]} \tag{1}$$

$K_{s,t}$-free graphs with $s,t \geq 2$, which satisfy $c_\alpha(G) \geq \frac{(\alpha d(G))^{s/(s-1)}}{2t}$; and blow-ups of constant degree expanders.

Let us first see an example demonstrating a positive answer to Question A.

Example C. A classical result of Komlós and Szemerédi [22] and of Bollobás and Thomason [6] says that every graph G contains a topological clique of order $\Omega(\sqrt{d(G)})$. In incoming work [17], it is proved that every graph G contains a topological clique of order $\Omega(\sqrt{c_\alpha(G)}/(\log c_\alpha(G))^{O(1)})$. Since $c_\alpha(G) = \Omega(d^2(G))$ when G is a C_4-free graph, this implies Mader's conjecture that C_4-free graphs contains topological cliques of order linear in its average degree, up to polylogarithmic factors [35]. (Actually, Liu and Montgomery [31] have demonstrated that Mader's conjecture is true.)

Let us see another motivating question regarding cycles in expanders, i.e. graphs in which vertex subsets expand to large neighbourhoods. Originally introduced for network design, expanders, apart from being a central notion in graph theory, also have close interplay with other areas of mathematics and theoretical computer science, see e.g. the comprehensive survey of Hoory, Linial and Wigderson [16]. The type of expanders hitherto studied usually have constant expansion, i.e. are linear expanders. We consider here instead expanders with sublinear expansion, introduced by Komlós and Szemerédi in the 90s [21, 22]. We defer the formal definition of sublinear expanders to Section 2.2. This notion of sublinear expanders has been proved to be a powerful tool for embedding sparse graphs, playing an essential role in the recent resolutions of several long-standing conjectures that were previously out of reach, see e.g. [11, 14, 17, 19, 31, 32, 34]. It would therefore be useful to study these sublinear expanders.

Cycle lengths in linear expanders have been well studied, see e.g. [12, 25]. In particular, Krivelevich [25] proved that every linear expander contains a cycle of length linear in its order. What about sublinear expanders? Note that we *cannot* necessarily find a linear-size cycle, unlike the linear expander case. For example, $K_{n,\frac{n}{\log^2 n}}$ is a sublinear expander, but any cycle must take half its vertices from the smaller part, and consequently has length sublinear in the total number of vertices. However, in the case of $K_{n,\frac{n}{\log^2 n}}$ we can instead consider a subexpander

[1]See Propositions 2.5 and 2.7.

$H = K_{n',n'}$, where $n' = \frac{n}{\log^2 n}$, which has the average degree about half of $K_{n,\frac{n}{\log^2 n}}$. Now this subexpander H does have a cycle of length linear in the order of H.

Does such phenomenon always occur? That is, is it true that if we cannot find a linear-size cycle in a sublinear expander G, then we can find within G a subgraph H, with about the same average degree as G, that has a cycle of length linear in the order of H? We shall see shortly that this is *indeed the case*.

1.2 Crux and cycles

Our first result finds a cycle of length linear in crux size in generic graphs, extending the aforementioned folklore result of cycles linear in average degree and giving a positive instance to Question A.

Theorem 1.2. *Let $0 < \alpha < 1$. Then every graph G contains a cycle of length at least*

$$\frac{1-\alpha}{16000} \cdot c_\alpha(G).$$

It is worth mentioning that the above statement for $\alpha < 1/2$ can be deduced using a variant of the classical Pósa's lemma [36] that if sets up to size k expands linearly, then there is a cycle of length $\Omega(k)$. To see this, first pass to a subgraph H with $\delta(H) \geq d(G)/2$; clearly $|H| \geq c_{1/2}(G) \geq c_\alpha(G)$. Then every set $X \subseteq V(H)$ of size $\Theta(c_\alpha(G))$ must expand linearly, for otherwise $H[X \cup N_H(X)]$ has average degree almost $d/2$ while having smaller order than $c_\alpha(G)$, a contradiction. Such argument, however, cannot push α beyond $1/2$ as we cannot guarantee the minimum degree of a graph to be larger than half of its average degree.

Remark D. The value of Theorem 1.2 is that we can take $\alpha = 1 - o(1)$, which is needed to close the exponential gaps in the applications below, see Corollaries 1.4 and 1.5. The idea of getting the whole range $0 < \alpha < 1$ is to pass to an expander subgraph with different expansion threshold t to have better expansions for large sets.

We have the following corollary on cycles in sublinear expanders. The above example of imbalanced complete bipartite graph shows that both terms in the bound below are best possible up to multiplicative constants.

Corollary 1.3. *Let $0 < \alpha < 1$, $0 < \varepsilon \leq \frac{1-\alpha}{500}$, $t \geq 1$ and suppose $n \geq 150t$. Then every n-vertex (ε, t)-expander G contains a cycle of length*

$$\max\left\{ \frac{\varepsilon}{32} c_\alpha(G) \ , \ \frac{\varepsilon n}{1200 \log^2 n} \right\}.$$

1.3 Application to Long's conjecture

Long [33, Conjecture 8.9] conjectured that any subgraph of the hypercube Q^m that has average degree d contains a path of length at least $2^d - 1$. He gave a weaker bound of a path of length at least $2^{d/2} - 1$, by passing to a subgraph of minimum degree at least $d/2$. A similar conjecture for discrete tori C_3^m was made in the same paper. Long proved that every subgraph of C_3^m that has average degree at least d contains a path of length at least $2^{d/4} - 1$, and he conjectured [33, Conjecture 8.3] that the correct bound should be $3^{d/2} - 1$. Both conjectures, if true, would be best possible by considering sub-hypercubes or sub-torus.

Using Theorem 1.2 and isoperimetric inequalities (1), we immediately close the above exponential gaps and settle both conjectures asymptotically. It would be interesting to see if stability method can be combined to obtain exact results.

Corollary 1.4. *Every subgraph of the hypercube with average degree d contains a cycle of length*

$$2^{d-o(d)}.$$

Proof. Fix arbitrary $0 < \varepsilon < 1$ and let $H \subseteq Q^m$ be a subgraph with $d(H) = d$. By the definition of crux and (1), we have $c_{1-\varepsilon}(H) \geq c_{(1-\varepsilon)\frac{d}{m}}(Q^m) \geq 2^{(1-\varepsilon)d}$. Then by Theorem 1.2, H contains a cycle of length at least $\frac{\varepsilon}{16000} 2^{(1-\varepsilon)d}$ as desired. □

The same proof applies also to Hamming graphs. The case $r = 3$ below covers discrete tori.

Corollary 1.5. *Every subgraph of the Hamming graph $H(m, r)$ with average degree d contains a cycle of length*

$$r^{\frac{d}{r-1} - o(d)}.$$

1.4 Random subgraphs of a given graph

Our next positive instances of Question A concern long cycles in random subgraphs of a given graph. For a given finite graph G and a real $p \in [0, 1]$, let G_p be a random subgraph of G obtained by taking each edge independently with probability p. Analysis of G_p can be used to demonstrate the robustness of a graph G with respect to a graph property $\mathcal{P}$, see e.g. [27, 28]. If G is the complete graph K_n, then G_p is simply the Erdős–Rényi binomial random graph $G(n, p)$. We say an event happens *asymptotically almost surely* (a.a.s.) or *with high probability* (w.h.p.) in $G(n, p)$ if its probability tends to 1 as $n \to \infty$.

Long paths, cycles and Hamiltonicity in $G(n, p)$ have been intensively studied, see e.g. [1, 3, 4, 5, 13, 23, 20, 30, 36]. In particular, Frieze [13] proved that for large C, w.h.p. $G(n, C/n)$ has a cycle of length at least $(1 - (1 - o_C(1))Ce^{-C})n$. Krivelevich, Lee and Sudakov [28] extended these classical results of long paths and cycles in $G(n, p)$ to random subgraphs G_p, where G has large minimum degree. For long cycles, they proved that given a graph G with minimum degree k, if $pk \to \infty$, then w.h.p. G_p contains a cycle of length at least $(1 - o(1))k$. Riordan [37] subsequently gave a shorter proof, and Ehard and Joos [9] further improved the error term. Krivelevich and Samotij [29] later considered graphs without a fixed bipartite subgraph H; in the case of C_4-free G with $\delta(G) \geq k$, they showed that w.h.p. G_p, $p = \frac{1+\varepsilon}{k}$, contains a cycle of length $\Omega_\varepsilon(k^2)$. We give a short proof for random subgraphs of C_4-free graphs with $p = \omega(\frac{1}{k})$. Note that the constant 1 below is best possible, as there are C_4-free graphs with minimum degree k and order $(1 + o(1))k^2$.

Theorem 1.6. *Let G be a C_4-free graph with minimum degree k. Suppose that $pk \to \infty$ as $k \to \infty$. Then w.h.p. G_p contains a cycle of length at least $(1 - o(1))k^2$.*

Random subgraphs of the hypercube are also well studied, see e.g. [2, 7, 15]. For hypercubes, we obtain the following near linear bound. It would be interesting to prove a linear bound. While this paper was being prepared, Erde, Kang and Krivelevich [10] proved Theorem 1.7 with a better error term $\Omega(\frac{2^m}{m^3 \log^3 m})$.

Theorem 1.7. *Let Q^m be the m-dimensional hypercube. If $p = \frac{1+\varepsilon}{m}$, where $\varepsilon > 0$, then w.h.p. Q^m_p contains a cycle of length $\frac{2^m}{4m^{32}} = 2^{(1-o(1))m}$.*

Organisation. The rest of the paper is organised as follows. Section 2 contains some necessary tools needed in our proofs. In Section 3, we give the proofs of Theorem 1.2 and Corollary 1.3. We prove Theorems 1.6 and 1.7 in Section 4. Concluding remarks are given in Section 5.

2 Preliminaries

For $a, b \in \mathbb{N}$ with $a < b$, let $[a] := \{1, \ldots, a\}$ and $[a, b] := \{a, a + 1, \ldots, b\}$. If we claim that a result holds whenever we have $c \gg d > 0$, it means that there exists a positive function f such that the result holds as long as $c > f(d)$. We write log for the natural logarithm.

Given a graph G, denote its order and size by $|G|$ and $e(G)$ respectively, and its average degree $2e(G)/|G|$ by $d(G)$. For a vertex subset $U \subseteq V(G)$, write $N_G(U) := \{v \in V(G) \setminus U : v \text{ has a neighbour in } U\}$ for its external neighbourhood; write ∂U for the edge boundary of U, that is, $E_G(U, V(G) \setminus U)$; and write $G - U = G[V(G) \setminus U]$ for the subgraph induced on $V(G) \setminus U$.

2.1 Depth First Search

We will need Depth First Search (DFS), which is a graph exploration algorithm that visits all the vertices of an input graph. It may be summarised as follows. We maintain a searching stack S, initially empty, set of unexplored vertices U, initially $V(G)$, and set of explored vertices X, initially empty, as well as a spanning subgraph F, initially empty. At each step, if S is empty but U is not, remove an arbitrary vertex of U and push it onto S. If the top vertex of S has a neighbour in U, remove such a neighbour, push it onto S, and add the corresponding edge to F. If the top vertex of S has no neighbour in U, then pop it from S and add it to X. Stop when $X = V(G)$.

We will use the following straightforward properties of S, U and X which hold throughout the process.

- The stack S induces a path in G.
- There is no edge of G between U and X.

2.2 Sublinear expanders

For $\varepsilon > 0$ and $t > 0$, let $\rho(x)$ be the function

$$\rho(x) = \rho(x, \varepsilon, t) := \begin{cases} 0 & \text{if } x < t/5, \\ \varepsilon / \log^2(15x/t) & \text{if } x \geq t/5, \end{cases} \tag{2}$$

where, when it is clear from context, we will not write the dependency of $\rho(x)$ on ε and t.

Definition 2.1 (Sublinear expander)**.** A graph G is an (ε, t)-*expander* if for any subset $X \subseteq V(G)$ of size $t/2 \leq |X| \leq |V(G)|/2$, we have $|N_G(X)| \geq \rho(|X|) \cdot |X|$.

Although sublinear expanders have a weaker expansion property, they have a key advantage over linear expanders in that any graph contains a sublinear expander subgraph that, furthermore, is almost as dense as the original graph, as shown by Komlós and Szemerédi [21, 22]. We shall use the following strengthening by Haslegrave, Kim and Liu [14].

Lemma 2.2 ([14])**.** *Let $C > 30, \varepsilon \leq 1/(10C), t > 0, d > 0$ and $\rho(x) = \rho(x, \varepsilon, t)$ as in (2). Then every graph G with $d(G) = d$ has a subgraph H such that H is an (ε, t)-expander, $d(H) \geq (1-\delta)d$ and $\delta(H) \geq d(H)/2$, where $\delta := \frac{C\varepsilon}{\log 3}$.*

The following lemma shows the key property of sublinear expanders that we will utilise. It roughly says that in a sublinear expander, we can connect two sets X_1, X_2 using a short path while avoiding another set W as long as W is a bit smaller than X_1, X_2. Although in many applications the bound on the length of such a path will be important, in this paper all we shall actually need is the existence of a path avoiding a certain set.

Lemma 2.3 (Small diameter lemma [22, Corollary 2.3])**.** *If G is an n-vertex (ε, t)-expander, then for any two vertex sets X_1, X_2 each of size at least $x \geq t/2$, and a vertex set W of size at most $\rho(x)x/4$, there exists a path in $G - W$ between X_1 and X_2 of length at most $\frac{2}{\varepsilon} \log^3(\frac{15n}{t})$.*

2.3 Isoperimetry

To find long cycles in subgraphs of hypercubes and Hamming graphs, we will need the following isoperimetric result.

Theorem 2.4 ([18, Theorem 1]). *Every $U \subseteq V(Q^m)$ satisfies $|\partial U| \geq |U| \cdot \log_2(2^m/|U|)$.*

A bound on the order of a subgraph with average degree d then immediately follows.

Proposition 2.5. *Every subgraph G of Q^m with average degree d has at least 2^d vertices.*

Proof. By Theorem 2.4, $|\partial V(G)| \geq |G| \cdot \log_2(2^m/|G|)$. Since $2|E(G)| + |\partial V(G)| = m|G|$, we have $|E(G)| = d \cdot |G|/2 \leq |G| \cdot \log_2|G|/2$. Hence, $|G| \geq 2^d$. □

A similar result for Hamming graphs holds.

Proposition 2.6 ([38, Proposition 2]). *Every subgraph G of the Hamming graph $H(m, r)$ has at most $(r-1)|G| \cdot \log_r|G|/2$ edges.*

Consequently, in such a graph $d(G) \leq (r-1)\log_r|G|$, giving the following corollary.

Proposition 2.7. *Every subgraph G of $H(m, r)$ with average degree d has at least $r^{\frac{d}{r-1}}$ vertices.*

3 Cycles of length linear in crux

3.1 Proof of Theorem 1.2

Theorem 3.1 ([24, Theorem 1]). *Let $k > 0, t \geq 2$ be integers. Let G be a graph on more than k vertices, satisfying:*

$$|N_G(W)| \geq t, \text{ for every } W \subseteq V(G) \text{ with } k/2 \leq |W| \leq k.$$

Then G contains a cycle of length at least $t+1$.

Proof of Theorem 1.2. Let $\delta = 1 - \alpha$ and take $C = 40, \varepsilon = \frac{\delta}{500}$, so $\delta > \frac{C\varepsilon}{\log 3}$. Write $n_c = c_\alpha(G)$ and let $H \subseteq G$ be a subgraph that is an $(\varepsilon, n_c/2)$-expander, guaranteed by Lemma 2.2. Then $d(H) \geq (1-\delta)d(G)$, by the definition of the crux, we have $n_H := |H| \geq n_c$. Set $K = \frac{n_H}{n_c} \geq 1$.

As $\rho(x)x$ is increasing in x and $K \geq 1$, by the expansion property of H, every set of size $n_H/4 \leq x \leq n_H/2$ has an external neighbourhood of size at least

$$\rho\Big(\frac{n_H}{4}\Big)\frac{n_H}{4} = \frac{\varepsilon n_H}{4\log^2(\frac{15 n_H/4}{n_c/2})} = \frac{\varepsilon K n_c}{4\log^2(15K/2)} \geq \frac{\varepsilon}{32} \cdot n_c.$$

We may assume that $\frac{\varepsilon}{32} \cdot n_c \geq 2$, for otherwise we can take a single edge as a degenerate cycle. Then by Theorem 3.1, the graph H, hence also G, contains a cycle of length at least $\frac{\varepsilon}{32} n_c = \frac{1-\alpha}{16000} c_\alpha(G)$. □

3.2 Proof of Corollary 1.3

A cycle of length $\frac{\varepsilon}{32} c_\alpha(G)$ follows from the proof of Theorem 1.2. The 2nd term $\varepsilon n/(1200 \log^2 n)$ follows from expansion property of sublinear expander and Theorem 3.1. We give a direct proof for completeness.

First, as $\varepsilon < 1/500$, the conditions on n imply that $n/300 \geq t/2$, that $\varepsilon n/(1200 \log^2 n) \leq (n/300) \cdot \rho(n/300)/4$, and that $\varepsilon n/(1200 \log^2 n) \leq n/300$.

Consequently, if there is a path of length $n/100$, then we are done, because after removing the middle $\varepsilon n/(1200 \log^2 n)$ vertices of the path, there is still a short path avoiding the middle part connecting the two halves by Lemma 2.3. This gives a cycle containing the middle $\varepsilon n/(1200\log^2 n)$ vertices of the path. So assume that such a path does not exist.

We run DFS until some point where $|X| = n/3$. Since the stack S always induces a path in G, we have $|S| < n/100$, and so $|U| > 0.65n$. By Lemma 2.3 and the fact that S is a cut between X and U, we have $|S| > 0.3n \cdot \rho(0.3n)/4 > \varepsilon n/(1200\log^2 n)$. Let P_1 be the path induced by S at that point and set $i = 2$. Now continue running DFS. Whenever a new vertex is added to S, call the new path P_i and increment i. Do this until $i = n/3$. By the same reasoning throughout this process we have $\varepsilon n/(1200\log^2 n) < |S| < n/100$, and in particular the lower bound implies the first $\varepsilon n/(1200\log^2 n)$ vertices of the path never change. Thus we have a set of $n/3$ paths with a long common first section and different endpoints.

Now consider the largest common first section P. This corresponds to the point between P_1 and $P_{n/3}$ where S is smallest (and equals P). Fix X and U corresponding to their values at that point. Again, P is a cut between X and U, both of which have size at least $0.32n$. Let P' be the subpath of P consisting of the final $\varepsilon n/(1200\log^2 n)$ vertices, and u be the same endpoint of P' and P. Since $|P| = |S| > \varepsilon n/(1200\log^2 n)$, we have $V(P) \setminus V(P') \neq \varnothing$.

Suppose without loss of generality (if not, exchange X and U) more than half of the paths $P_1, \ldots, P_{n/3}$ come before this point. This means their endpoints are in X; let Y be the set of these endpoints, giving $|Y| \geq 0.16n$. For any vertex in Y, there is a path to u which lies entirely in X. Let $Z = U \cup V(P) \setminus V(P')$. Then Z has size more than $0.32n$. By Lemma 2.3, there exists a short path in $G - V(P')$ connecting Y and Z. Indeed, as there are no edges between U and X, the short path connects Y and $V(P) \setminus V(P')$. This gives a cycle containing P' with desired length.

4 Random subgraphs

4.1 Long cycles in random subgraphs of C_4-free graphs

We prove Theorem 1.6 by adapting Riordan's proof [37]. Fix $0 < \varepsilon < 1/10$ and let $C = 10/\varepsilon$. It suffices to show that w.h.p. G_p contains a cycle of length at least $(1 - 20\varepsilon)k^2$.

Consider a DFS forest T of G_p, leaving edges *unrevealed* if they are not needed in the exploration. To be precise, when checking whether the top vertex v of the stack has a neighbour in U, we list the edges $E_G(v, U)$ and reveal whether each in turn is in G_p until either we find such an edge or exhaust the list. We write n for the order of G and $Q \subseteq G$ for the subgraph consisting of all unrevealed edges. Throughout the process, each edge in Q is present in G_p independently with probability p; in particular this means that for any given set of εk edges of Q, w.h.p. at least one is present since $\varepsilon kp \to \infty$. Furthermore, every edge of Q joins two vertices in T one of which is an ancestor of the other; see [37, Lemma 2]. We consider each component of T to be rooted at the first vertex to be added to the stack S, and we consider the *descendants* $D(v)$ of a vertex v to be the set of vertices w such that the path from w to the root contains v (note in particular that $v \in D(v)$).

Note that we are done if the following holds for any subset R:

$$\sum_{v\in R} \big|\{u : uv \in Q, (1 - 20\varepsilon)k^2 \leq d_T(u, v) < \infty\}\big| \geq \varepsilon k, \tag{3}$$

where $d_T(u, v)$ is the distance in T, since then w.h.p. at least one of these εk edges is present, say uv, and creates a cycle of length at least $(1 - 20\varepsilon)k^2$ together with the path in T from u to v. Thus we assume from now on that (3) is not true for any set R.

A vertex is *full* if it has at least $(1 - \varepsilon)k$ incident edges in Q. Standard concentration inequalities show that w.h.p. at most $2n/p = o(kn)$ edges are revealed in the whole process; and so w.h.p. all but $o(n)$ vertices are *full*.

Claim 4.1. For any set A of Ck full vertices, we have $|N_Q(A)| \geq (1 - 5\varepsilon)k^2$.

Proof. Consider the bipartite graph $H = Q[A, B]$ consisting of the unrevealed edges between A and B where $B = N_Q(A)$. Note that $G[A]$ is a C_4-free graph with Ck vertices, hence it contains at most $C^2k^{1.5} < \varepsilon^2k^2$ edges. Then, as the vertices in A are full, H contains at least $(1 - \varepsilon - \varepsilon^2)Ck^2$ edges.

If $\sum_{v\in B} \binom{d_H(v)}{2} > \binom{|A|}{2} = \binom{Ck}{2}$, then there exists a pair of vertices in A having two common neighbours, a contradiction to the C_4-freeness of G. Hence, by convexity of the function $f(x) = \binom{x}{2}$, we have

$$\binom{Ck}{2} \geq \sum_{v\in B} \binom{d_H(v)}{2} \geq |B|\binom{(1-\varepsilon-\varepsilon^2)Ck^2/|B|}{2} \geq (1-3\varepsilon)\Big(\frac{C^2k^4}{2|B|} - \frac{Ck^2}{2}\Big).$$

As $C > 10/\varepsilon$, this yields that $|B| \geq (1 - 4\varepsilon)(1 - \frac{1}{C+1})k^2 \geq (1 - 5\varepsilon)k^2$. □

We say that a vertex is *poor* if it has at most εk^2 descendants, and *rich* otherwise. We wish to show that at most $o(n)$ vertices are poor. In [37] where we aim for a cycle of length $(1 - o(1))k$, this is immediate, since if v is both poor and full then $\{v\}$ satisfies (3), but this does not translate to our setting. Consequently establishing that there are few poor vertices is the main difficulty in extending the proof.

Lemma 4.2. *If* (3) *does not hold for any set R, then $o(n)$ vertices are poor.*

Proof. Let W be a subset of children of some vertex v and write $R(W) = \bigcup_{w\in W} D(w)$. Suppose $2Ck \leq |R(W)| \leq \varepsilon k^2$. If at least Ck vertices in $R(W)$ are full, then by Claim 4.1, we may choose $(1-5\varepsilon)k^2$ neighbours of those Ck vertices via edges of Q. In particular, at least $(1-6\varepsilon)k^2$ such neighbours are not in $R(W)$ and must be ancestors of v; since it has at most one ancestor at each distance, at least εk^2 of them are at distance at least $(1 - 7\varepsilon)k^2$ from $R(W)$, and so (3) holds for $R(W)$. Thus in any such vertex subset $R(W)$, at least half the vertices are not full.

Note that if v is rich but all its children are poor then we may divide its children into disjoint subsets $W_1, \dots, W_r, L$ such that each of $R(W_1), \dots, R(W_r)$ have size between $2Ck$ and εk^2 and $R(L)$ has size less than $2Ck$. Thus, noting that $\sum_i |R(W_i)| \geq \varepsilon k^2 - 2Ck$, almost half of the descendants of v are not full, and so the total number of descendants of such vertices v is $o(n)$.

Similarly, suppose that v has some poor children and some rich children. If the total number of descendants of poor children of v is at least $2Ck$, then we can argue as before to show that at least half of these descendants are not full.

Grouping the remaining poor vertices by their nearest rich ancestor v, each group P_v has size less than $2Ck$. We aim to associate each of these poor vertices y with $\omega(1)$ other vertices Z_y, in such a way that the Z_ys are disjoint; this would imply that $\bigcup_v P_v = o(n)$. However, we will not necessarily cover all vertices this way; we will only fail to cover vertices from groups which are mostly not full, ensuring that there are few of them.

Write A for set of rich vertices v satisfying $P_v \neq \varnothing$ but $P_u = \varnothing$ for each ancestor u of v. Choose an arbitrary vertex $v \in A$ and follow a path (in T) downwards from v, using only rich vertices, until one of the following happens.

1. The total size of $\bigcup_w P_w$ for w on the path is at least $2Ck$, or
2. the last vertex on the path, x has no rich children.

Write P for the union of all P_w where w is on the path.

In case 2, the total size of P is at most $2Ck$. However, $|D(x)| \geq \varepsilon k^2$ (all these descendants have already been counted above and so are not in any P_w). Thus we may choose Z_y for $y \in P$ disjointly from $D(x)$ with $|Z_y| \geq \lfloor \varepsilon k/(2C)\rfloor$.

In case 1, the total size of P is at most $4Ck$. If at least half of the vertices in P are full, we may find, using Claim 4.1, $(1 - 5\varepsilon)k^2$ distinct vertices adjacent to vertices in P by unrevealed edges. All such vertices must be either inside P, on the path, or ancestors of v. If the length of the path is at most εk^2 then at least $(1-6\varepsilon)k^2$ of the vertices must be ancestors of v, and so (3) is satisfied as above. Thus the length of the path is at least εk^2 and we can choose each Z_y for $y \in P$ from the path ensuring that $|Z_y| \geq \lfloor \varepsilon k/(4C) \rfloor$. If the majority of vertices in P are not full then we do not define these Z_ys.

In either case, we now set all P_w along the path to $\varnothing$, update A, and continue. Note that all paths chosen during this process are disjoint, since if x is an ancestor of a vertex in a chosen path then P_x is empty either at the time the path is chosen (by choice of start vertex, if x is not on the path) or immediately afterwards (if x is also on the path), so no subsequent path can start at x. Thus all the Z_y chosen have size $\omega(1)$ and are disjoint as required. Additionally, at most $o(n)$ vertices y do not have associated Z_y, since at least half of any such vertices are not full. Consequently, there are $o(n)$ poor vertices. □

A path in a rooted tree is *vertical* if it does not contain the root as an interior vertex. Define a vertex as *light* if $\left|D_{\leq(1-10\varepsilon)k^2}(v)\right| \leq (1-9\varepsilon)k^2$, where $D_{\leq i}(v) \subseteq D(v)$ are the descendants within distance i. Let H be the set of non-light vertices. The proof of the following lemma is the same as [37, Lemmas 5, 6] up to slight changes in the parameters.

Lemma 4.3. *Suppose that T contains $o(n)$ poor vertices and $Y \subseteq V(T)$ satisfies $|Y| = o(n)$. Then T contains a vertical path P of length $2Ck^2$, containing at most $\varepsilon^2 k^2$ vertices in $Y \cup H$.*

We are now ready to complete the proof. We are done if any set satisfies (3), so assume not. Then Lemmas 4.2 and 4.3 ensure the long path described above exists. Write Z for the set of vertices on the path which are both full and light. We order Z according to height on the path, and will consider blocks of Ck consecutive vertices of Z in this ordering. By Lemma 4.3, there are at most $\varepsilon^2 k^2$ vertices on the path which are not in Z, so the total distance on the path between the top and bottom vertices of any such block is at most $\varepsilon^2 k^2 + Ck < \varepsilon k^2$. By Claim 4.1, the vertices in a block have at least $(1-5\varepsilon)k^2$ distinct neighbours by unrevealed edges.

We consider the vertex closest to the root to be the top vertex of the block. Since the top vertex of the block is light, at most $(1-9\varepsilon)k^2$ vertices are below this vertex but within distance $(1-10\varepsilon)k^2$. Since (3) does not apply, at most εk neighbours are more than $(1-11\varepsilon)k^2$ below the bottom vertex of the block. Recalling the bound on the length of a block, and noting that descendants of any vertex in the block are descendants of the top vertex, any neighbours below fall into one of these two categories, so at least $4\varepsilon k^2 - \varepsilon k$ neighbours are above the top vertex of the block.

Taking V_0 to be the bottom Ck vertices of Z we know that these have εk neighbours at least distance $4\varepsilon k^2 - 2\varepsilon k \geq 3\varepsilon k^2$ above the top vertex of V_0, so w.h.p. we can find a $v_0 \in V_0$ and u_0 at least this distance above, connected by an edge which is present. Then we choose V_1 to be the highest Ck vertices in Z below u_0 and continue. Note that those Ck vertices are disjoint as $3\varepsilon k^2 > Ck + \varepsilon^2 k^2$. Note that we go up at least $3\varepsilon k^2$ steps from the top vertex of V_0 to u_0 and down at most $\varepsilon^2 k^2$ steps from u_0 to the top of V_1. Since $0 < \varepsilon < 1/10$ and $d(v_0, u_0) < k^2$ (for otherwise we have a length-k^2 cycle), and the path has length $2Ck^2$, w.h.p. we may continue in this way to find overlapping 'chords' $v_i u_i$ for $0 \leq i \leq C$. Since $d(u_i v_{i+2}) \geq 3\varepsilon k^2 - 2\varepsilon^2 k^2 - Ck > \varepsilon k^2$, w.h.p. there is a cycle of length at least $C\varepsilon k^2 \geq k^2$ consisting of these chords together with the sections of the path $v_0 \cdots v_1$ and $u_i \cdots v_{i+2}$ for $0 \leq i \leq C-2$, and $u_{C-1} \cdots u_C$.

4.2 Long cycles in random subgraphs of hypercubes

To prove Theorem 1.7, we use concentration of the size of the giant component to show that w.h.p. there is no small separators. This idea is not new and appeared earlier in the work of Krivelevich, Lubetzky and Sudakov [26]. To carry out this argument, we need a result relating separability of graphs to separator size; we first give the necessary definitions.

Definition 4.4. Given a graph $G = (V, E)$ on n vertices, a vertex set $S \subseteq V$ is called a *separator* if there is a partition $V = A \cup B \cup S$ of the vertex set of G such that G has no edges between A and B, and $|A|, |B| \leq 2n/3$.

Definition 4.5. Let s, t be positive integers. A graph G is (s, t)*-separable* if there exists a vertex subset $S \subseteq V(G)$ such that $|S| \leq s$ and every component of $G - S$ has at most t vertices.

Lemma 4.6. *Let G be a graph with n vertices and fix $t, s > 0$. If G is not $(\frac{4n^2}{st}, t)$-separable, then G has a subgraph H such that $|H| \geq t$ and H has no separator with size at most $\frac{1}{s}|H|$.*

Proof. Suppose that every subgraph H of G with at least t vertices has a separator with size at most $\frac{1}{s}|H|$. Then G has a separator S such that $|S| \leq \frac{1}{s}|G|$ and $V(G) \setminus S = X_1 \dot\cup X_2$ with $|X_1|, |X_2| \leq \frac{2n}{3}$ and $e_G(X_1, X_2) = 0$. For each X_i $(i \in \{1, 2\})$, if $|X_i| \geq t$, then $G[X_i]$ has a separator S_i such that $|S_i| \leq \frac{1}{s}|X_i|$ and $X_i \setminus S_i = X_{i1} \dot\cup X_{i2}$ with $|X_{i1}|, |X_{i2}| \leq \frac{2|X_i|}{3} \leq (\frac{2}{3})^2 n$ and $e_G(X_{i1}, X_{i2}) = 0$. For each X_{ij} $(i, j \in \{1, 2\})$, if $|X_{ij}| \geq t$, then $G[X_{ij}]$ has a separator S_{ij} such that $|S_{ij}| \leq \frac{1}{s}|X_{ij}|$ and $X_{ij} \setminus S_{ij} = X_{ij1} \dot\cup X_{ij2}$ with $|X_{ij1}|, |X_{ij2}| \leq \frac{2|X_{ij}|}{3} \leq (\frac{2}{3})^3 n$ and $e_G(X_{ij1}, X_{ij2}) = 0$. We repeat this to obtain $S_{ijk}, X_{ijk1}, X_{ijk2}$ $(i, j, k \in \{1, 2\})$ and so on. Assume that this process stops when $S_{i_1 i_2 i_3 \ldots i_\ell}, X_{i_1 i_2 i_3 \ldots i_{\ell+1}}$ are obtained, i.e. each $X_{i_1 i_2 i_3 \ldots i_{\ell+1}}$ has size less than t. For each $k \leq \ell + 1$ let $\mathcal{A}^k = \{i_1 \ldots i_k : X_{i_1 \ldots i_k} \text{ is defined}\}$.

As $t \leq |X_{i_1 i_2 i_3 \ldots i_\ell}| \leq (\frac{2}{3})^\ell n$, we know that $\ell \leq \log_{3/2}(n/t)$. Let $S^0 = S$ and for $1 \leq k \leq \ell$, $S^k = \bigcup_{i_1 \ldots i_k \in \mathcal{A}^k} S_{i_1 i_2 i_3 \ldots i_k}$. Then

$$|S^k| \leq \sum_{i_1 \ldots i_k \in \mathcal{A}^k} \frac{1}{s}|X_{i_1 i_2 i_3 \ldots i_k}| \leq 2^k \cdot \frac{1}{s} \cdot \left(\frac{2}{3}\right)^k n \leq \left(\frac{4}{3}\right)^k \cdot \frac{n}{s}.$$

Let $S^* = \bigcup_{0 \leq k \leq \ell} S^k$. Then $|S^*| \leq 3 \cdot (\frac{4}{3})^{\ell+1} \cdot \frac{n}{s} \leq \frac{4n^2}{st}$ and every component in $G - S^*$ has size less than t. Hence, G is $(\frac{4n^2}{st}, t)$-separable, a contradiction. $\square$

By taking $s = 4\psi(n)^3$ and $t = \frac{n}{\psi(n)}$, where $\psi(n) = n^{o(1)}$, we have the following corollary.

Corollary 4.7. *If G is not $(\frac{n}{\psi(n)^2}, \frac{n}{\psi(n)})$-separable, then G has a subgraph H such that $|H| \geq \frac{n}{\psi(n)}$ and H has no separator with size at most $\frac{1}{4\psi(n)^3}|H|$.*

Write $\mathcal{C}_1(G)$ for the largest component in a graph G. Let $p = (1 + 4\theta)/m$ for some small $\theta > 0$, and set $p' = (1 - \theta)p > (1 + 2\theta)/m$. Write $p_1 = (1 + \theta)/m$ and choose $p_2 \geq \theta/m$ such that $(1 - p_1)(1 - p_2) = 1 - p'$ and $n = 2^m$. We assume that $\varepsilon = 4\theta$ is given and m is sufficiently large.

Claim 4.8. There exists $c = c(\theta)$ satisfying the following: $\mathbb{P}[|\mathcal{C}_1(Q^m_{p'})| \geq c2^m] \geq 1 - \exp(-n/m^{14})$.

Proof. We prove this in two steps. The first step (clustering) is performed in $Q^m_{p_1}$, and we deduce that w.h.p. $\Omega(2^m)$ vertices are contained in components of size at least m^4 and most of vertices are adjacent to at least one such a component. For the second step (sprinkling), we mainly follow the sprinkling process in [15, Section 1.3]: add the edges of $Q^m_{p_2}$ and show that they can connect many of the clusters of size at least $\sqrt{m}$ into a giant cluster of size $\Theta(2^m)$.

Step 1. Let $V = V(Q^m)$. Let the random variable $B = B(Q^m)$ be the set of vertices in $Q^m_{p_1}$ that belong to a component of order at least m^4. By the main theorem in [2], there exists $c_1 = c_1(\theta/3)$ such that for any $q \geq (1 + \theta/3)/m$,

$$\mathbb{P}[\mathcal{C}_1(Q^m_q) > 12c_1 2^m] \geq 1 - c_1. \tag{4}$$

Since $c_1 2^m > m^4$, it follows that $\mathbb{E}[|B|] \geq 6c_1 2^m$.

For a vertex $v \in V(Q^m)$, we can find vertices $v_1, \dots, v_{\theta m/3} \in N_{Q^m}(v)$ and vertex-disjoint subhypercubes $Q_1, \dots, Q_{\theta m/3}$ of dimension $(1-\theta/3)m$ in Q^m where each Q_i contains v_i.

Note that conditioning on the existence of a component of size $12c_1|Q^m|$ in Q^m_q, the probability that such a component contains a specific vertex v is at least $12c_1$ as Q^m is vertex-transitive. Hence, the equation (4) (with $(1-\theta/3)m$ playing the role of m) implies that the vertex v_i belongs to a component of size $12c_1|Q_i| \geq m^4$ in $(Q_i)_{p_1}$ with probability at least $12c_1(1-c_1) \geq c_1$. As $Q_1, \dots, Q_{\theta m/3}$ are disjoint subgraphs of Q^m, those events are mutually independent. Moreover, if one such an event happens, then we have $v \in N_{Q^m}[B]$, where we write $N_{Q^m}[B] = B \cup N_{Q^m}(B)$. Hence, we have

$$\mathbb{E}[|V \setminus N_{Q^m}[B]|] = \sum_{v \in V} \mathbb{P}[v \notin N_{Q^m}[B]] \leq (1-c_1)^{\theta m/3} \cdot 2^m \leq \frac{2^m}{m^2}.$$

Enumerate edges of Q^m as $e_1, e_2, \dots, e_{m2^{m-1}}$; let I_i be the indicator random variable that $e_i \in E(Q^m_{p_1})$ and let $\mathcal{F}_i$ be the σ-algebra generated by $(I_j)_{j \leq i}$. Consider the edge-exposure martingale $X_0, X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ with

$$X_i = \mathbb{E}[|B| : \mathcal{F}_i] \text{ and } Y_i = \mathbb{E}[|V \setminus N_{Q^m}(B)| : \mathcal{F}_i].$$

Note that changing one I_i makes $|B|$ by at most $2m^4$ and $|N_Q[B]|$ by at most $2m^5$, since any vertex for which e_i is critical is in a component of order less than m^4 in $Q^m_{p_1} - e_i$ containing exactly one endpoint of e_i, and such a component has at most m^5 neighbours in Q^m. Thus the martingale is $2m^4$-Lipschitz and $2m^5$-Lipschitz respectively, and by Azuma's inequality we have

$$\mathbb{P}\left[|B| < 3c_1 2^m\right] \leq \mathbb{P}\left[|B| < \mathbb{E}[|B|] - 3c_1 2^m\right] \leq \exp\Big(-\frac{9c_1^2 2^{2m}}{2(2m^4)^2 \cdot m2^{m-1}}\Big) \leq \exp\Big(-\frac{2^m}{m^{10}}\Big),$$

$$\begin{aligned}\mathbb{P}\left[|V \setminus N_{Q^m}[B]| > \frac{2^{m+1}}{m}\right] &\leq \mathbb{P}\left[|V \setminus N_{Q^m}[B]| > \mathbb{E}\left[|V \setminus N_{Q^m}[B]|\right] + \frac{2^m}{m}\right] \\ &\leq \exp\Big(-\frac{2^{2m}/m^2}{2(2m^5)^2 \cdot m2^{m-1}}\Big) \leq \exp\Big(-\frac{2^m}{m^{13}}\Big).\end{aligned}$$

Step 2. From Step 1, we have $|B| \geq 3c_1 2^m$ and $|V \setminus N_{Q^m}[B]| \leq 2^{m+1}/m$ with sufficiently high probability. We say that *sprinkling fails* when these high probability events happen but $|\mathcal{C}_1(Q^m_{p_1} \cup Q^m_{p_2})| \leq c_1 2^m$. If sprinkling fails, then we can partition $B = C \dot\cup D$ such that $|C|, |D| \geq c_1 2^m$, each of C and D is a union of components in $Q^m_{p_1}$, and any C-D path in Q^m has an edge erased in $Q^m_{p_2}$. Since every component of $Q^m_{p_1}$ meeting B has size at least m^4, the number of partitions meeting the second condition is at most $2^{2^m/m^4}$.

By the isoperimetric inequality for the hypercube and Menger's Theorem, for a particular partition $C \dot\cup D$ with $|C|, |D| \geq c_1 2^m$ there exist at least $\frac{c_1}{100\sqrt{m}} \cdot 2^m$ internally vertex-disjoint C-D paths in Q^m.

Take such a collection $\mathcal{P}$ of paths with the minimum total sum of lengths. Note that a path in $\mathcal{P}$ has at most four vertices in $N_{Q^m}[B]$. Indeed, if a vertex u_i in the path $u_1 u_2 \dots u_s$ with $u_1 \in C, u_s \in D$ and $3 \leq i \leq s-2$ has a neighbour w in $B = C \cup D$, then either the path $u_1 \dots u_i w$ or the path $w u_i u_{i+1} \dots u_s$ can replace the path $u_1 \dots u_s$ in $\mathcal{P}$ to contradict the minimality of $\mathcal{P}$. Hence, at most $|V(Q^m) \setminus N_{Q^m}[B]| \leq 2^{m+1}/m$ paths in $\mathcal{P}$ have length at least 4 and at least $\frac{c_1}{100\sqrt{m}} \cdot 2^m - \frac{2^{m+1}}{m} \geq \frac{c_1}{200\sqrt{m}} 2^m$ paths have length at most 3. Hence, the probability that all such paths have an edge erased in $Q^m_{p_2}$ is at most

$$\Big(1 - (\frac{\theta}{m})^3\Big)^{\frac{c_1 2^m}{200\sqrt{m}}} < \exp\Big(-\frac{1}{2}(\frac{\theta}{m})^3 \cdot \frac{c_1 2^m}{200\sqrt{m}}\Big) < 2^{-2^{m+2}/m^4}.$$

Consequently the probability that sprinkling fails is at most

$$2^{2^m/m^4} \cdot 2^{-2^{m+2}/m^4} \leq \exp(-2^m/m^4).$$

By the above two steps, we obtain that

$$\mathbb{P}[|\mathcal{C}_1(Q^m_{p'})| \geq c_1 n] \geq 1 - \exp\big(- 2^m/m^{14}\big). \qquad \square$$

Proof of Theorem 1.7. Let $G = Q^m$. Note that $G_{p'}$ can be obtained by deleting edges in G_p with probability θ independently. Let $\mathcal{A}$ be the event that G_p is $(n/m^{16}, n/m^8)$-separable and $\mathcal{B}$ be the event that $|\mathcal{C}_1(G_{p'})| < n/m^8$. Assume that $\mathcal{A}$ occurs. Then we have a vertex subset S with size at most n/m^{16} such that every component of $G - S$ has at most n/m^8 vertices. If all edges between S and $G - S$ are deleted when passing from Q^m_p to $Q^m_{p'}$, then $\mathcal{B}$ happens. This deletion of all edges between S and $G - S$ happens with probability at least $\theta^{|S|m} \geq \theta^{n/m^{15}}$. Hence, $\mathbb{P}[\mathcal{B}] \geq \mathbb{P}[\mathcal{A}] \cdot \theta^{n/m^{15}}$. However, $\mathbb{P}[\mathcal{B}] \leq \exp(-n/m^{14})$ by Claim 4.8. Thus we have $\mathbb{P}[\mathcal{A}] \leq \exp(-n/m^{14}) \cdot \theta^{-n/m^{15}} = o(1)$.

By Corollary 4.7, w.h.p. G_p has a subgraph H such that $|H| \geq n/m^8$ and H has no separator with size at most $|H|/(4m^{24})$. Thus we have $N_H(W) \geq |H|/(4m^{24}) \geq n/(4m^{32})$ for any $W \subseteq V(H)$ with $|H|/3 \leq |W| \leq 2|H|/3$. Applying Theorem 3.1 we obtain that H, and so also G_p, has a cycle of length at least $n/(4m^{32}) = 2^{(1-o(1))m}$. $\square$

5 Concluding remarks

In this paper, we define the crux of a graph recording the order of the smallest dense patch of a graph, and study the 'replacing average degree by crux' paradigm. As a first example, we finds in generic graphs cycles of length linear in its crux size and apply it to address two conjectures of Long regarding long paths in subgraphs of hypercubes and Hamming graphs. As the crux of a C_4-free graph is quadratic in its average degree, and the crux of a hypercube is exponential in its dimension, Theorems 1.6 and 1.7 are two more examples of this paradigm. It would be interesting to see more results of this form.

Acknowledgement

We would like to thank Michael Krivelevich for bringing [26] to our attention.

Note added before submission. Theorem 1.7 has been proved independently by Erde, Kang and Krivelevich [10] with better cycle length $\Omega(\frac{2^m}{m^3 \log^3 m})$.

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, The longest path in a random graph. *Combinatorica*, 1, (1981), 1–12.

[2] M. Ajtai, J. Komlós, E. Szemerédi, Largest random component of a k-cube. *Combinatorica*, 2, (1982), 1–7.

[3] M. Ajtai, J. Komlós, E. Szemerédi, First occurence of Hamilton cycles in random graphs. *North-Holland Mathematics Studies*, 115(C), (1985), 173–178.

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

[5] B. Bollobás, T. Fenner, A. Frieze, Long cycles in sparse random graphs. *Graph Theory and Combinatorics* (Cambridge, 1983), (1984), 59–64.

[6] B. Bollobás, A. Thomason, Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *European Journal of Combinatorics*, 19, (1998), 883–887.

[7] P. Condon, A. E. Díaz, A. Girão, D. Kühn, D. Osthus, Hamiltonicity of random subgraphs of the hypercube. *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, (2021), 889–898.

[8] G. A. Dirac, Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 2, (1952), 69–81.

[9] S. Ehard, F. Joos, Paths and cycles in random subgraphs of graphs with large minimum degree. *Electronic Journal of Combinatorics*, 25(2), (2018), P2.31.
[10] J. Erde, M. Kang, M. Krivelevich, Expansion, long cycles, and complete minors in supercritical random subgraphs of the hypercube. arXiv preprint, arXiv:2106.04249.
[11] I. Gil Fernández, J. Kim, Y. Kim, H. Liu, Nested cycles with no geometric crossings. arXiv preprint, arXiv:2104.04810.
[12] L. Friedman, M. Krivelevich, Cycle lengths in expanding graphs. *Combinatorica*, 41, (2021), 53–74.
[13] A.M. Frieze, On large matchings and cycles in sparse random graphs. *Discrete Mathematics*, 59(3), (1986), 243–256.
[14] J. Haslegrave, J. Kim, H. Liu, Extremal density for sparse minors and subdivisions. *International Mathematics Research Notices*, to appear.
[15] R. van der Hofstad, A. Nachmias, Hypercube percolation. *Journal of the European Mathematical Society*, 19, (2017), 725–814.
[16] S. Hoory, N. Linial, A. Wigderson, Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43, (2006), 439–561.
[17] S. Im, J. Kim, Y. Kim, H. Liu, Topological cliques in sublinear expanders. *preprint*.
[18] P. Keevash, E. Long, A stability result for the cube edge isoperimetric inequality. *Journal of Combinatorial Theory, Series A*, 155, (2018), 360–375.
[19] J. Kim, H. Liu, M. Sharifzadeh, K. Staden, Proof of Komlós's conjecture on Hamiltonian subsets, *Proceedings of the London Mathematical Society*, 115 (5), (2017), 974–1013.
[20] J. Komlós, E. Szemerédi, Limit distribution for the existence of Hamilton cycles in random graphs. *Discrete Mathematics*, 43, (1983), 55–63.
[21] J. Komlós, E. Szemerédi, Topological cliques in graphs. *Combinatorics, Probability and Computing*, 3, (1994), 247–256.
[22] J. Komlós, E. Szemerédi, Topological cliques in graphs II. *Combinatorics, Probability and Computing*, 5, (1996), 79–90.
[23] M. Krivelevich, Long paths and Hamiltonicity in random graphs. *Random graphs, geometry and asymptotic structure*, 84, (2016), 1.
[24] M. Krivelevich, Long cycles in locally expanding graphs, with applications. *Combinatorica*, 39, (2019), 135–151.
[25] M. Krivelevich, Expanders - how to find them, and what to find in them. *Surveys in Combinatorics*, 456, (2019), 115–142.
[26] M. Krivelevich, E. Lubetzky, B. Sudakov, Asymptotics in percolation on high-girth expanders. *Random Structures & Algorithms*, (2020), 1–21.
[27] M. Krivelevich, C. Lee, B. Sudakov, Robust Hamiltonicity of Dirac graphs. *Transactions of the American Mathematical Society*, 366(6), (2014), 3095–3130.
[28] M. Krivelevich, C. Lee, B. Sudakov, Long paths and cycles in random subgraphs of graphs with large minimum degree. *Random Structures & Algorithms*, 46, (2015), 320–345.
[29] M. Krivelevich, W. Samotij, Long paths and cycles in random subgraphs of H-free graphs. *Electronic Journal of Combinatorics*, 21(1), (2014), P1.30.
[30] C. Lee, B. Sudakov, Dirac's theorem for random graphs. *Random Structure & Algorithms*, 41(3) (2012), 293–305.
[31] H. Liu, R.H. Montgomery, A proof of Mader's conjecture on large clique subdivisions in C_4-free graphs. *Journal of the London Mathematical Society*, 95(1), (2017), 203–222.
[32] H. Liu, R.H. Montgomery, A solution to Erdős and Hajnal's odd cycle problem. arXiv preprint, arXiv:2010.15802.
[33] E. Long, Long paths and cycles in subgraphs of the cube. *Combinatorica*, 33, (2013), 395–428.
[34] H. Liu, G. Wang, D. Yang, Clique immersion in graphs without fixed bipartite graph. arXiv preprint, arXiv:2011.10961.
[35] W. Mader, An extremal problem for subdivisions of K_5^-. *Journal of Graph Theory*, 30, (1999), 261–276.
[36] L. Pósa, Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14, (1976), 359–364.
[37] O. Riordan, Long cycles in random subgraphs of graphs with large minimum degree. *Random Structure & Algorithms*, 45, (2014), 764–767.
[38] R. Squier, B. Torrence, A. Vogt, The number of edges in a subgraph of a Hamming graph. *Applied Mathematics Letters*, 14, (2001), 701–705.