Hamilton cycles in highly connected and expanding graphs

Dan Hefetz [*] Michael Krivelevich [†] Tibor Szabó [‡]

December 13, 2008

Abstract

In this paper we prove a sufficient condition for the existence of a Hamilton cycle, which is applicable to a wide variety of graphs, including relatively sparse graphs. In contrast to previous criteria, ours is based on two properties only: one requiring expansion of "small" sets, the other ensuring the existence of an edge between any two disjoint "large" sets. We also discuss applications in positional games, random graphs and extremal graph theory.

1 Introduction

A Hamilton cycle in a graph G is a cycle passing through all vertices of G. A graph is called *Hamiltonian* if it admits a Hamilton cycle. Hamiltonicity is one of the most central notions in Graph Theory, and much effort has been devoted to obtain sufficient conditions for the existence of a Hamilton cycle (an effective necessary and sufficient condition should not be expected however, as deciding whether a given graph contains a Hamilton cycle is known to be NP-complete). In this paper we will mostly concern ourselves with establishing a sufficient condition for Hamiltonicity which is applicable to a wide class of sparse graphs.

[*]Institute of Theoretical Computer Science, ETH Zurich, CH-8092 Switzerland, and School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: dan.hefetz@inf.ethz.ch. This paper is a part of the author's Ph.D. thesis written under the supervision of Prof. Michael Krivelevich.

[†]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: krivelev@post.tau.ac.il. Research supported in part by a USA-Israeli BSF grant and a grant from the Israel Science Foundation.

[‡]Department of Mathematics and Statistics, McGill University, Montreal QC H3A 2K6, Canada. Email: szabo@math.mcgill.ca.

One of the first Hamiltonicity results is the celebrated theorem of Dirac [10], which asserts that if the minimum degree of a graph G on n vertices is at least $n/2$ then G is Hamiltonian. Since then, many other sufficient conditions that deal with dense graphs, were obtained (see e.g. [14] for a comprehensive reference). However, all these conditions require the graph to have $\Theta(n^2)$ edges, whereas for a Hamilton cycle only n edges are needed. Chvátal and Erdős [8] proved that if $\kappa(G) \geq \alpha(G)$ (that is, the vertex connectivity of G is at least as large as the size of a largest independent set in G), then G is Hamiltonian. Note that if G is a d-regular graph, then $\kappa(G) \leq d$ and $\alpha(G) \geq \frac{n}{d+1}$; hence the Chvátal-Erdős criterion cannot be applied if $d \leq c\sqrt{n}$ for an appropriate constant c.

When looking for sufficient conditions for the Hamiltonicity of sparse graphs, it is natural to consider random graphs with an appropriate edge probability. Erdős and Rényi [11] raised the question of what the threshold probability of Hamiltonicity in random graphs is. After a series of efforts by various researchers, including Korshunov [17] and Pósa [21], the problem was finally solved by Komlós and Szemerédi [18] and independently by Bollobás [5], who proved that if $p \geq (\log n + \log\log n + \omega(1))/n$, where $\omega(1)$ tends to infinity with n arbitrarily slowly, then $G(n,p)$ is almost surely Hamiltonian. Note that this is best possible since for $p \leq (\log n + \log\log n - \omega(1))/n$ there are vertices of degree at most one in $G(n,p)$ almost surely. An even stronger result was obtained by Bollobás [5]. He proved that for almost every graph process, the hitting time of being Hamiltonian is exactly the same as the hitting time of having minimum degree 2, that is, with probability tending to 1, the very edge which increases the minimum degree to 2, also makes the graph Hamiltonian.

The next natural step is to look for Hamilton cycles in relatively sparse pseudo-random graphs. During the last few years, several such sufficient conditions were found (see e.g. [13, 19]). These routinely rely on many properties of pseudo-random graphs and are thus quite complicated. Furthermore, one can argue that these conditions are not the most natural, as Hamiltonicity is a monotone increasing property, whereas pseudo-randomness is not. Our main result is a natural and simple (at least on the qualitative level) sufficient condition for Hamiltonicity which is based on expansion and high connectivity. Before stating our result, we introduce and discuss the two key graph properties involved. As usual, the notation $N(S)$ stands for the *external neighborhood* of S, that is, $N(S) = \{v \in V \setminus S : \exists u \in S,\ (u,v) \in E\}$. Let $G = (V,E)$ be a graph where $|V| = n$ and let $d = d(n)$ be a parameter. Consider the following two properties:

P1 For every $S \subset V$, if $|S| \leq \frac{n \log\log n \log d}{d \log n \log\log\log n}$ then $|N(S)| \geq d|S|$;

P2 There is an edge in G between any two disjoint subsets $A, B \subseteq V$ such that $|A|, |B| \geq \frac{n \log\log n \log d}{4130 \log n \log\log\log n}$.

From now on, for the sake of convenience, we denote

$$m = m(n,d) = \frac{\log n \cdot \log\log\log n}{\log\log n \cdot \log d}\ .$$

Let us give an informal interpretation of the above conditions. Condition P1 guarantees *expansion*: every sufficiently small vertex subset (of size $|S| \leq \frac{n}{dm}$) expands by a factor of d. Condition P2 is what can be classified as a *high connectivity* condition of some sort: every two disjoint subsets $A, B \subseteq V$ which are relatively large (of size $|A|, |B| \geq \frac{n}{4130m}$) are connected by at least one edge. Note that properties P1 and P2 together guarantee some expansion for *every* $S \subset V(G)$ of size $o(n)$. Indeed, if $|S| \leq \frac{n}{dm}$ then $|N(S)| \geq d|S|$ by property P1. If $\frac{n}{dm} < |S| < \frac{n}{4130m}$ (assuming $d > 4130$) then S contains a subset S_0 of size exactly $\frac{n}{dm}$ that, by property P1, expands at least to a size of $\frac{n}{m}$. Hence, S expands by a factor of at least 4129. Finally, if $|S| \geq \frac{n}{4130m}$ then $N(S) \geq (1 - o(1))n$ as, by property P2, the number of vertices of $V \setminus S$ that do not have any neighbor in S is strictly less than $\frac{n}{4130m}$.

We can now state our main result:

Theorem 1.1 *Let $12 \leq d \leq e^{\sqrt[3]{\log n}}$ and let G be a graph on n vertices satisfying properties P1, P2 as above; then G is Hamiltonian, for sufficiently large n.*

The lower bound on d in the theorem above can probably be somewhat improved through a more careful implementation of our arguments. As for the upper bound $d \leq e^{\sqrt[3]{\log n}}$, it is a mere technicality; one expects that proving Hamiltonicity when d is larger should in fact be easier. The requirement $d \leq e^{\sqrt[3]{\log n}}$ makes sure (in particular) that $\frac{n}{4130m} = o(n)$ and so P2 is a non-trivial condition. We can obtain a sufficient condition for Hamiltonicity, similar to that of Theorem 1.1, and applicable to graphs with larger values of $d = d(n)$ as well; more details are given in Section 2.4.

Let $\bar{d}$ be the average degree of G. Obviously, P1 can only be valid for $d \leq \bar{d}$. On the other hand, G contains an independent set of size $\Theta(n/\bar{d})$. Hence, in order to apply property P2, one must require $cn/\bar{d} \leq n/m$. These two inequalities entail that the applicability of Theorem 1.1 is limited to graphs whose average degree $\bar{d}$ is at least $(\log n)^{1-o(1)}$.

It is instructive to observe that neither P1 nor P2 is enough to guarantee Hamiltonicity by itself, without relying on its companion property (unless of course they degenerate to something trivial). Indeed, for property P1 observe that the complete bipartite graph $K_{n,n+1}$ is a very strong expander locally, yet obviously it does not contain a Hamilton cycle. As for property P2, the graph G formed by a disjoint union of a clique of size $n - \frac{n}{4130m} + 1$ and $\frac{n}{4130m} - 1$ isolated vertices clearly meets P2, but is obviously quite far from being Hamiltonian. Thus, P1 and P2 complement each other in an essential way.

Next, we discuss several applications of our main result. Theorem 1.1 was first used by the authors (see [15]) to address a problem of Beck [4]: it is proved that Enforcer can win the $(1 : q)$ Avoider-Enforcer Hamilton cycle game, played on the edges of K_n, for every $q \leq \frac{cn \log\log\log\log n}{\log n \log\log\log n}$ where c is an appropriate constant. This upper bound on q was subsequently improved in [20] to $\frac{(1-o(1))n}{\log n}$ which provides an affirmative answer to Beck's question. The proof also relies on Theorem 1.1. Using the same approach, it was also proved in [20] that Maker can win the $(1:q)$ Maker-Breaker Hamilton cycle game, played on the edges of K_n, for every $q \leq \frac{(\log 2-o(1))n}{\log n}$ improving the best known bound of Beck [3] by a factor of 27. In [16], Theorem 1.1 was used to prove that Maker can win the $(1:1)$ Maker-Breaker Hamilton cycle game, played on the edges of the random graph $G(n,p)$, for every $p \geq \frac{\log n+(\log\log n)^c}{n}$, where c is a sufficiently large constant. Clearly, this is asymptotically tight. [A brief background: both Maker-Breaker and Avoider-Enforcer games mentioned above are played on the edge set of the complete graph K_n. In every move, Maker (resp. Avoider) claims one unoccupied edge and Breaker (resp. Enforcer) responds by claiming q unoccupied edges. The game ends when all edges have been claimed by one of the players. In the Maker-Breaker Hamiltonicity game Maker wins if he creates a Hamilton cycle, otherwise Breaker wins. In the Avoider-Enforcer version, Avoider wins if he avoids creating a Hamilton cycle by the end of the game, otherwise Enforcer wins. More details can be found in [4].] Recently, Alon and Nussboim [2] and Frieze, Vempala, and Vera [12] used Theorem 1.1 to prove the Hamiltonicity of k-wise independent, and of log-concave random graphs, respectively.

In this paper we prove several other corollaries of Theorem 1.1. A graph $G=(V,E)$ is called *Hamilton-connected* if for every $u,v \in V$ there is a Hamilton path in G from u to v.

Theorem 1.2 *Let $G=(V,E)$ be a graph on n vertices that satisfies properties P1 and P2. Then G is Hamilton-connected, for sufficiently large n.*

Remark. An immediate consequence of Theorem 1.2 is that for every edge $e \in E$ there is a Hamilton cycle of G that includes e.

A graph G is called *pancyclic* if it admits a cycle of length k for every $3 \leq k \leq n$. We prove that a graph which satisfies property P2 is "almost pancyclic".

Theorem 1.3 *Let $G=(V,E)$, where $|V|=n$ is sufficiently large, be a graph, satisfying property P2; more precisely, for every disjoint subsets $A,B \subseteq V$ such that $|A|,|B| \geq n/t$, where $t=t(n) \geq 2$, there is an edge between a vertex of A and a vertex of B. Then G admits a cycle of length exactly k for every $\frac{8n\log n}{t\log\log n} \leq k \leq n-3n/t$.*

Remark. The upper bound on k in Theorem 1.3 is tight up to a constant factor in the second order term, as shown by a disjoint union of $K_{n+1-n/t}$ and $n/t-1$ isolated vertices. On the other hand, we believe that the lower bound can be improved to $\frac{c\log n}{\log t}$ for some constant c. Methods recently introduced by Verstraëte [23] and by Sudakov and Verstraëte [22] can possibly be used to establish this conjecture.

Theorem 1.1 (with minor changes to its proof) can be used to prove the following classic result (see [18], [5]).

Theorem 1.4 *The random graph $G(n,p)$, where $p = (\log n + \log\log n + \omega(1))/n$, is almost surely Hamiltonian.*

Our proof technique has the potential to be applied in other settings and models of random graphs. One major difference from previous proofs is that the pseudo-random properties on which we rely are monotone increasing.

Let $G = (V, E)$, where $|V| = n$, and let $f : \mathbb{Z}^+ \to \mathbb{R}$. A pair (A, B) of proper subsets of V is called a *separation* of G if $A \cup B = V$ and there are no edges in G between $A \setminus B$ and $B \setminus A$. The graph G is called f-connected if $|A \cap B| \geq f(|A \setminus B|)$, for every separation (A, B) of G with $|A \setminus B| \leq |B \setminus A|$. Brandt, Broersma, Diestel, and Kriesell [7] proved that if $f(k) \geq 2(k+1)^2$ for every $k \in \mathbb{N}$, then G is Hamiltonian for every $n \geq 3$. They conjectured that there exists a function f which is linear in k and is enough to ensure Hamiltonicity. Using Theorem 1.1, we can get quite close to proving this conjecture for sufficiently large n:

Theorem 1.5 *If $G = (V, E)$, where $|V| = n$, is f-connected for $f(k) = k \log k + O(1)$, then it is Hamiltonian for sufficiently large n.*

For the sake of simplicity and clarity of presentation, we do not make a particular effort to optimize the constants obtained in theorems we prove. We also omit floor and ceiling signs whenever these are not crucial. All of our results are asymptotic in nature and whenever necessary we assume that n is sufficiently large. Throughout the paper, log stands for the natural logarithm. We say that some event holds *almost surely*, or a.s. for brevity, if the probability it holds tends to 1 as n tends to infinity. Our graph-theoretic notation is standard and follows that of [9].

The rest of the paper is organized as follows: in Section 2 we prove and discuss Theorem 1.1, in Section 3 we prove its corollaries: Theorems 1.2, 1.3, 1.4 and 1.5.

2 Proof of the main result

The proof of Theorem 1.1 is based on the ingenious rotation-extension technique, developed by Pósa [21], and applied later in a multitude of papers on Hamiltonicity (mostly of random graphs). Our proof technique borrows some technical ideas from the paper of Ajtai, Komlós and Szemerédi [1].

Before diving into the fine details of the proof, we would like to compare our Hamiltonicity criterion and its proof with its predecessors. Several previous papers, including [1], [13] and [19], state, explicitly or implicitly, sufficient conditions for Hamiltonicity which are, in principle, applicable to sparse graphs. Criteria of this sort are carefully tailored to work for random or pseudo-random graphs, and are therefore rather complicated and not always the most natural. In particular, they are often fragile in the sense that they can be violated by adding more edges to the graph – a somewhat undesirable feature considering that Hamiltonicity is a monotone increasing property. Our criterion, given in Theorem 1.1 is (on a qualitative level, at least) easily comprehensible, monotone increasing, and potentially can be applied to a very wide class of graphs. As for our proof, due to the relative simplicity of the conditions we use, the argument is perhaps more involved than some of the previous proofs; there are however similarities. A novel ingredient, relying heavily on Property P2, is the part presented in Section 2.2 (finding many good initial rotations).

In order to be able to refer to the proof of our criterion while proving some of its corollaries, we break the proof into four parts, each time indicating which property is needed for which part.

Proposition 2.1 *Let G satisfy properties P1 and P2. Then G is connected.*

Proof If not, let C be the smallest connected component of G. Then by P1, $|C| > \frac{n}{m}$, but then by P2, $E(C, V \setminus C) \neq \emptyset$ – a contradiction. $\square$

2.1 Constructing an initial long path

In this subsection we show that a graph which satisfies some expansion properties (that is, property P1 and some expansion of larger sets, implied by property P2) contains a long path, and even more, it has many paths of maximum length starting at the same vertex.

Let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path of maximum length in G. If $1 \leq i \leq q-2$ and (v_q, v_i) is an edge of G, then $P' = (v_1 v_2 \ldots v_i v_q v_{q-1} \ldots v_{i+1})$ is also of maximum length. P' is called a *rotation* of P_0 with *fixed endpoint* v_1 and *pivot* v_i. The edge (v_i, v_{i+1}) is called the *broken* edge of the rotation. We say that the segment $v_{i+1} \ldots v_q$ of P_0 is reversed in P'.

In case the new endpoint, v_{i+1}, has a neighbor v_j such that $j \notin \{i, i+2\}$, then we can rotate P' further to obtain more paths of maximum length. We use rotations together with property P1 to find a path of maximum length with large rotation endpoint sets (see for example [6], [13], [18], [19]).

Claim 2.2 *Let $G = (V, E)$ be a graph on n vertices that satisfies property P1 (with parameter d, $12 \leq d \leq \sqrt{n}$), and moreover any subset of V of size $n/4130m$ has at least $n - o(n)$ external neighbors. Let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path of maximum length in G. Then there exists a set $B(v_1) \subseteq V(P_0)$ of at least $n/3$ vertices, such that for every $v \in B(v_1)$ there is a v_1, v-path of maximum length which can be obtained from P_0 by at most $\frac{2 \log n}{\log d}$ rotations with fixed endpoint v_1. In particular $|V(P_0)| \geq n/3$.*

Proof Let t_0 be the smallest integer such that $\left(\frac{d}{3}\right)^{t_0-2} > \frac{n}{md}$; note that $t_0 \le 2\frac{\log n}{\log d}$, because $12 \le d \le \sqrt{n}$.

We construct a sequence of sets $S_0, \ldots, S_{t_0} = B(v_1) \subseteq V(P_0)\setminus\{v_1\}$ of vertices, such that for every $0 \le t \le t_0$ and every $v \in S_t$, v is the endpoint of a path which can be obtained from P_0 by a sequence of t rotations with fixed endpoint v_1, such that for every $0 \le i < t$, the non-v_1-endpoint of the path after the ith rotation is contained in S_i. Moreover, $|S_t| = \left(\frac{d}{3}\right)^t$ for every $t \le t_0 - 3$, $|S_{t_0-2}| = \frac{n}{dm}$, $|S_{t_0-1}| = \frac{n}{4130m}$, and $|S_{t_0}| \ge n/3$.

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

Let now t be an integer with $0 < t \le t_0 - 2$ and assume that the sets $S_0, \ldots, S_{t-1}$ with the appropriate properties have already been constructed. We will now construct S_t. Let

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

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

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

$$\begin{aligned} |\hat{S}_t| &\ge \frac{1}{2}|T| \ge \frac{1}{2}\left(|N(S_{t-1})| - 3(1 + |S_1| + \ldots + |S_{t-1}|)\right) \\ &\ge \frac{d}{2}\left(\frac{d}{3}\right)^{t-1} - \frac{3}{2}\frac{(d/3)^t - 1}{d/3 - 1} \ge \left(\frac{d}{3}\right)^t \end{aligned}$$

where the last inequality follows since $d \ge 12$. Clearly we can delete arbitrary elements of $\hat{S}_t$ to obtain S_t of size exactly $\left(\frac{d}{3}\right)^t$ if $t \le t_0 - 3$ and of size exactly $\frac{n}{dm}$ if $t = t_0 - 2$. So the proof of the induction step is complete and we have constructed the sets $S_0, \ldots, S_{t_0-2}$.

To construct S_{t_0-1} and S_{t_0} we use the same technique as above, only the calculation is slightly different. Since $|N(S_{t_0-2})| \ge d \cdot \frac{n}{dm}$, we have

$$\begin{aligned}
|\hat{S}_{t_0-1}| &\geq \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t_0-2})| - 3(1+|S_1|+\ldots+|S_{t_0-4}|+|S_{t_0-3}|+|S_{t_0-2}|)\right) \\
&\geq \frac{n}{2m} - \frac{3}{2}\left(\frac{(d/3)^{t_0-3}-1}{(d/3)-1} + 2\frac{n}{dm}\right) \geq \frac{n}{2m} - \frac{3}{2}\cdot\left(\frac{d}{3}\right)^{t_0-3} - 3\frac{n}{dm} \\
&\geq \frac{n}{2m} - \frac{3}{2}\cdot\frac{n}{dm} - 3\frac{n}{dm} \geq \frac{n}{4130m},
\end{aligned}$$

where the last inequality follows since $d \geq 12$.

For S_{t_0} the difference in the calculation comes from using the expansion guaranteed by the weak P2-type property of the claim, rather than property P1. That is, we use the fact that $|N(S_{t_0-1})| \geq n - o(n)$. Hence, we have

$$\begin{aligned}
|S_{t_0}| &\geq \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t_0-1})| - 3(1+|S_1|+\ldots+|S_{t_0-2}|+|S_{t_0-1}|)\right) \\
&\geq \frac{n}{2}(1-o(1)) - \frac{3}{2}\left(\frac{(d/3)^{t_0-3}-1}{(d/3)-1} + \frac{2n}{dm} + \frac{n}{4130m}\right) \\
&\geq \frac{n}{2}(1-o(1)) - \frac{3}{2}\left(\frac{3n}{dm} + \frac{n}{4130m}\right) \\
&\geq \frac{n}{3},
\end{aligned}$$

where the last inequality follows since $m \geq 3$ and $d \geq 12$.

The set S_{t_0} can be chosen to be $B(v_1)$; it satisfies all the requirements of Claim 2.2. Note that since $S_{t_0} \subseteq V(P_0)$, we have $|V(P_0)| \geq n/3$. This concludes the proof of the claim. $\Box$

Remark Note that, although we do not need it here, the rotations which create these paths always break an edge of the original path P_0.

2.2 Finding many good initial rotations

In this subsection we prove an auxiliary lemma, which will be used in the next subsection to conclude the proof of Theorem 1.1.

Let H be a graph with a spanning path $P = (v_1, \ldots, v_l)$. For $2 \leq i < l$, let us define the auxiliary graph $H_i^+ = H_{v_i}^+$ by adding a vertex and two edges to H as follows: $V(H_i^+) = V(H) \cup \{w\}$, $E(H_i^+) = E(H) \cup \{(v_l, w), (v_i, w)\}$. Let $P_i = P_{v_i}$ be the spanning path of H_i^+ which we obtain from the path $P \cup \{(v_l, w)\}$ by rotating with pivot v_i. Note that the endpoints of P_i are v_1 and v_{i+1}.

For a vertex $v_i \in V(H)$, let S^{v_i} be the set of those vertices of $V(P) \setminus \{v_1\}$, which are endpoints of a spanning path of H_i^+ obtained from P_i by a series of rotations with fixed endpoint v_1.

A vertex $v_i \in V(P)$ is called a *bad initial pivot* (or simply a *bad vertex*) if $|S^{v_i}| < \frac{l}{43}$ and is called a *good initial pivot* (or a *good vertex*) otherwise. We can rotate P_i and find a large number of endpoints, provided that v_i is a good initial pivot.

We will show that H has many good initial pivots provided that a certain condition, similar to property P2, is satisfied.

Lemma 2.3 *Let H be a graph with a spanning path $P = (v_1, \dots, v_l)$. Assume that every two disjoint sets A, B of vertices of H of sizes $|A|, |B| \geq l/43$ are connected by an edge. Then*

$$|R| \leq 7l/43,$$

where $R = R(P) \subseteq V(P)$ is the set of bad vertices.

Proof We will create a set $U \subseteq V(H)$, whose size is at least $|R|/7$, but it is not "too large" and it does not expand enough, that is, $|U \cup N_H(U)| \leq 21|U|$. This in turn will imply that the set R of bad vertices cannot be big.

Let $R = \{v_{i_1}, \dots, v_{i_r}\}$. We *process* the vertices of R one after the other. We will maintain subsets U and X of $V(H)$ where initially $U = X = \emptyset$. Whenever we finish processing a vertex of R we update the sets U and X. The following properties will hold after the processing of v_{i_j}.

$$U \subseteq X\,, \quad N_H(U) \subseteq ext(X)\,, \quad |U| \geq \frac{1}{7}|X|\,, \quad \{v_{i_1+1}, \dots, v_{i_j+1}\} \subseteq X\,, \tag{1}$$

where for every $Y \subseteq V(P)$, $ext(Y)$ denotes the set containing the vertices of Y together with their left and right neighbors on P. Clearly $|ext(Y)| \leq 3|Y|$.

Suppose the current vertex to process is v_{i_j}. If $v_{i_j+1} \in X$, then we do not change U and X and so the conditions of (1) trivially hold by induction.

Otherwise, we will create sets $W_t \subseteq S^{v_{i_j}}$ inductively, such that for every t the following properties hold.

(*a*) $W_t \subseteq S_t^{v_{i_j}}$;

(*b*) $|W_t| = 2^t$;

(*c*) $W_t \cap \left(\cup_{s=0}^{t-1} W_s \cup X\right) = \emptyset$, where $S_t^{v_{i_j}}$ is the set of vertices y of $S^{v_{i_j}}$ for which a spanning path of $H_{i_j}^+$ ending at y can be obtained from P_{i_j} by a sequence of t rotations with fixed endpoint v_1, such that after the sth rotation the new endpoint is contained in W_s, for every $s < t$.

We begin by setting $W_0 = \{v_{i_j+1}\}$. Conditions (a) and (b) trivially hold, for condition (c) note that $v_{i_j+1} \notin X$.

Assume now that we have constructed $W_0, \ldots, W_t$ with properties $(a) - (c)$. If $|N_H(W_t) \setminus ext\,(\cup_{i=0}^t W_i \cup X)\,| > 5|W_t|$, then we create W_{t+1} with properties $(a) - (c)$, otherwise we finish the processing of v_{i_j} by updating U and X.

Let $T_t = N_H(W_t) \setminus ext\,(\cup_{i=0}^t W_i \cup X)$ and assume first that $|T_t| > 5|W_t|$. We use an argument similar to the one used in Claim 2.2 to create W_{t+1} with properties $(a) - (c)$.

Let $v_i \in T_t \setminus \{v_1, v_l\}$, and suppose that v_i is adjacent to $y \in W_t$. Recall, that by property (a), a spanning path Q of $H_{i_j}^+$ ending at y can be obtained from P_{i_j} by t rotations, such that for every $s < t$, after the sth rotation the new endpoint is in W_s. Since the vertices v_{i-1}, v_i and $v_{i+1} \notin \bigcup_{s=0}^t W_s$, they are not endpoints after any of these t rotations. Each rotation breaks an edge incident with the new endpoint, hence both edges (v_{i-1}, v_i) and (v_i, v_{i+1}) of the original path P_{i_j} must be present in Q. Rotating Q with pivot v_i will break one of them. Such a rotation also turns one of v_{i-1} and v_{i+1} into an endpoint, and as such, into an element of $S_{t+1}^{v_{i_j}}$. Denote this vertex by v_i'. We define $W_{t+1} = \{v_i' : v_i \in T_t\}$. We say that v_i' *is placed in* W_{t+1} *by* v_i. Observe that besides v_i the only other vertex that can place v_i' in W_{t+1} is its other neighbor on the path P_{i_j}. Thus,

$$|W_{t+1}| \geq \left\lceil \frac{1}{2}(|T_t| - 2) \right\rceil \geq 2|W_t|.$$

By deleting arbitrarily some vertices from W_{t+1}, we can make sure that its cardinality is exactly $2|W_t|$. Properties (a) and (b) are then clearly satisfied. Property (c) is satisfied because by the definition of T_t we have $v_i \notin ext\,(\cup_{s=0}^t W_s \cup X)$ and so none of its neighbors on P_{i_j}, in particular v_i', is an element of $(\cup_{s=0}^t W_s \cup X)$.

Property (b) ensures that $|W_t|$ is strictly increasing, so the processing of the vertex v_{i_j} is bound to reach a point in which $|T_k| \leq 5|W_k|$ for some index k. At that point we update U and X by adding W_k to U and adding $W_0 \cup W_1 \cup \cdots \cup W_k \cup T_k$ to X. We have to check that the conditions of (1) hold.

Observe that $|W_0 \cup \cdots \cup W_k| < 2|W_k|$, so the number of vertices added to X is at most seven times more than the number of vertices added to U. Also, property (c) and $U \subseteq X$ made sure that W_k was disjoint from U, so indeed the property $|U| \geq |X|/7$ remains valid. The other conditions in (1) follow easily from the definition of the "new" U and X. Hence the processing of v_{i_j} is complete.

Claim $|U| \leq l/43$.

Proof Assume the contrary and let j be the smallest index, such that $|U| > l/43$ after

the processing of v_{i_j}.
Observe that $|U| \leq 2l/43$. Indeed, after the processing of v_{i_j} the set U received at most $|S^{v_{i_j}}|$ vertices, which is at most $l/43$, due to the fact that v_{i_j} is a bad vertex. We thus have $l/43 < |U| \leq 2l/43$, $U \subseteq X$, $N_H(U) \subseteq ext(X)$ and $|ext(X)| \leq 3|X| \leq 21|U|$. Then $|V(P) \setminus ext(X)| \geq l/43$, and there are no edges of H between U and $V(P) \setminus ext(X)$. This contradicts our assumption on H. □

To conclude the proof of the lemma we note that after having processed all vertices of R, we have $R^+ := \{v_{i_1+1}, \ldots, v_{i_r+1}\} \subseteq X$ and $|U| \geq |X|/7$ by (1). Since $|U| \leq l/43$, it follows that $|R| = |R^+| \leq 7l/43$. □

2.3 Closing the maximal path

Lemma 2.4 *Let G be a connected graph that satisfies property P2 with parameter $d \leq e^{\sqrt[3]{\log n}}$. Let the conclusion of Claim 2.2 be also true for G, that is, for every path $P_0 = (v_1, v_2, \ldots, v_q)$ of maximum length in G there exists a set $B(v_1) \subseteq V(P_0)$ of at least $n/3$ vertices, such that for every $v \in B(v_1)$ there is a v_1, v-path of maximum length which can be obtained from P_0 by at most $t_0 = \frac{2\log n}{\log d}$ rotations with fixed endpoint v_1. Then G is Hamiltonian.*

Proof We will prove that there exists a path of maximum length which can be closed into a cycle. This, together with the connectedness of G, implies that the cycle is Hamiltonian. To find such a path of maximum length we will create two disjoint sets of vertices, large enough to satisfy property P2, such that between any two vertices (one from each set) there is a path of maximum length.

Let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path of maximum length in G, and let $A_0 = B(v_1)$. For every $v \in A_0$ fix a v_1, v-path $P^{(v)}$ of maximum length and, using our assumption, construct sets $B(v)$, $|B(v)| \geq n/3$, of endpoints of maximum length paths with fixed endpoint v, obtained from the path $P^{(v)}$ by at most t_0 rotations. To summarize, for every $a \in A_0$ and $b \in B(a)$ there is a maximum length path $P(a,b)$ joining a and b, which is obtainable from P_0 by at most $\rho := 2t_0 = \frac{4\log n}{\log d}$ rotations. Moreover, this clearly entails $|V(P_0)| \geq n/3$.

We consider P_0 to be directed from v_1 to v_q and divided into 2ρ consecutive vertex disjoint segments $I_1, I_2, \ldots, I_{2\rho}$ of length at least $\lfloor |V(P_0)|/2\rho \rfloor$ each. As each $P(a,b)$ is obtained from P_0 by at most ρ rotations, and every rotation breaks at most one edge of P_0, the number of segments of P_0 which also occur as segments of $P(a,b)$, although perhaps reversed, is at least ρ. We say that such a segment is *unbroken*. These segments have an absolute orientation given to them by P_0, and another, relative to this one, given to them by $P(a,b)$, which we consider to be directed from a to b. We consider sequences $\sigma = I_{i_1}, I_{i_2}, \ldots, I_{i_\tau}$ of unbroken segments of P_0, which occur in this order on $P(a,b)$, where σ also specifies the relative orientation of each segment. We call such a sequence σ a *τ-sequence*, and say that $P(a,b)$ *contains* σ. Note that $0 < \tau \leq \rho$ is a parameter whose exact value will be determined later.

For a given τ-sequence σ, we consider the set $L(\sigma)$ of ordered pairs (a,b), $a \in A_0,\ b \in B(a)$, such that $P(a,b)$ contains σ.

The total number of τ-sequences is $2^\tau(2\rho)_\tau$. Any path $P(a,b)$ contains at least ρ unbroken segments, and thus at least $\binom{\rho}{\tau}$ τ-sequences. The average, over τ-sequences, of the number of pairs (a,b) such that $P(a,b)$ contains a given τ-sequence is therefore at least

$$\frac{n^2}{9} \cdot \frac{\binom{\rho}{\tau}}{2^\tau(2\rho)_\tau} \geq \alpha n^2,$$

where $\alpha = \alpha(\tau) = \frac{1}{9}(4\tau)^{-\tau}$ for $0 < \tau \leq \rho$. Thus, there is a τ-sequence σ_0 and a set $L = L(\sigma_0)$, $|L| \geq \alpha n^2$ of pairs (a,b), such that for each $(a,b) \in L$, the path $P(a,b)$ contains σ_0. Let $\hat{A} = \{a \in A_0 : L$ contains at least $\alpha n/2$ pairs with a as first element$\}$. Since $|A_0|, |B(a)| \leq n$, we have $\alpha n^2 \leq |L| \leq |\hat{A}|n + n \cdot \frac{\alpha n}{2}$, entailing $|\hat{A}| \geq \alpha n/2$. For each $a \in \hat{A}$, let $\hat{B}(a) = \{b : (a,b) \in L\}$. Then, by the definition of $\hat{A}$, for each $a \in \hat{A}$ we have $|\hat{B}(a)| \geq \alpha n/2$.

Let $\tau = \frac{\log\log n}{2\log\log\log n}$ and let $\sigma_0 = (I_{i_1}, I_{i_2}, \ldots, I_{i_\tau})$. We divide σ_0 into two sub-sequences, $\sigma_0^1 = (I_{i_1}, \ldots, I_{i_{\tau/2}})$ and $\sigma_0^2 = (I_{i_{\tau/2+1}}, \ldots, I_{i_\tau})$ where both sub-sequences maintain the order and orientation of the segments of σ_0. For $i = 1, 2$, let us denote by $|\sigma_0^i|$ the number of vertices in the union of segments of σ_0^i. Then for both sub-sequences σ_0^1 and σ_0^2, we have that $|\sigma_0^i| \geq \tau/2 \cdot n/(6\rho) = \frac{n}{96m}$. Let x be the last vertex of $I_{i_{\tau/2}}$, and let y be the first vertex of $I_{i_{\tau/2+1}}$ (in the orientation given by σ_0).

For σ_0^1 construct a graph H_1 with $\cup_{j=1}^{\tau/2} V(I_{i_j})$ as vertex set. The edge set of H_1 is defined as follows. First, we add all edges of G, except for those that are incident with a vertex in EP_1, where EP_1 denotes the set of endpoints of the paths $I_{i_1}, \ldots, I_{i_{\tau/2}}$. For an endpoint z of the path I_{i_j}, we add the edge that connects z to its neighbor in I_{i_j}. Note that this edge is also an edge of G. Finally, we join by an edge the last vertex of I_{i_j} to the first vertex of $I_{i_{j+1}}$ for every $1 \leq j < \tau/2$. These edges might or might not be in G, in any case, we refer to them as *artificial* edges. By its construction, H_1 contains a spanning path P starting at x and ending at the first vertex s_1 of I_{i_1}, that links the vertices of the oriented path segments $I_{i_1}, \ldots, I_{i_{\tau/2}}$ in reverse order. We would like to apply Lemma 2.3 to H_1 with $l = \sum_{j=1}^{\tau/2} |V(I_{i_j})|$. The condition of the lemma holds since G satisfies property P2. Indeed, the edges of H differ from the edges of G only at the endpoints of the segments I_{i_j}, and $|EP_1| = \tau = o(|V(H_1)|)$. Hence, at least a $\frac{36}{43}$-fraction of the vertices of σ_0^1 are good.

For σ_0^2 we act similarly: construct a graph H_2 from the segments of σ_0^2 by joining the first vertex of I_{i_j} to the last vertex of $I_{i_{j-1}}$ for every $\tau/2 + 1 < j \leq \tau$ and adding all edges of G with both endpoints in the interior of segments of σ_0^2 to H_2. Then the segments of σ_0^2 with the edges linking them form an oriented spanning path in H_2, starting at y and ending at the last vertex s_2 of I_{i_τ}. Again, due to property P2, Lemma 2.3 applies here, so at least a

$\frac{36}{43}$-fraction of the vertices of σ_0^2 are good.

Recall that s_1 is the first vertex of I_{i_1}. Since $|\hat{A}| \geq \alpha n/2 \geq \frac{n}{4130m} + 1$ (which is why we get the upper bound on d in Theorem 1.1) and H_1 has at least $\frac{36}{43} \cdot \frac{n}{96m}$ good vertices, there is an edge of G between a vertex $\hat{a} \in \hat{A} \setminus \{s_1\}$ and a good vertex $g_1 \in \cup_{j=1}^{\tau/2} V(I_{i_j})$, such that $g_1 \notin EP_1$ (the last assertion follows, since the number τ of endpoints is $o(n/m)$).

Similarly, as $|\hat{B}(\hat{a})| \geq \alpha n/2$ and there are at least $\frac{36}{43} \cdot \frac{n}{96m}$ good vertices in H_2, there is an edge from some $\hat{b} \in \hat{B}(\hat{a}) \setminus \{s_2\}$ to a good vertex $g_2 \in \cup_{\tau/2+1}^{\tau} V(I_{i_j})$, such that g_2 is not the endpoint of any segment of σ_0^2 (here s_2 denotes the last vertex of I_{i_τ}).

Consider the path $P(\hat{a}, \hat{b})$ of maximum length in G connecting $\hat{a}$ and $\hat{b}$ and containing σ_0. The vertices x and y split this path into three subpaths: R_1 from $\hat{a}$ to x, R_2 from y to $\hat{b}$ and R_3 from x to y. We will rotate R_1 with x as a fixed endpoint and R_2 with y as a fixed endpoint. We will show that the obtained endpoint sets V_1 and V_2 are sufficiently large (clearly, they are disjoint). Then by property P2 there will be an edge of G between V_1 and V_2. Since we did not touch R_3, this edge closes a maximum path into a cycle, which is Hamiltonian due to the connectivity of G.

First we construct the endpoint set V_1, the endpoint set V_2 can be constructed analogously. Recall the notation from Subsection 2.2: Let $H_{g_1}^+$ denote the graph we obtain from H_1 by adding the extra vertex w and the edges (w, g_1) and (w, s_1). The spanning path of $H_{g_1}^+$ obtained by rotating $P \cup \{(w, s_1)\}$ with fixed endpoint x at pivot g_1 is denoted by P_{g_1}. By the definition of a good vertex, the set S^{g_1}, of vertices which are endpoints of a spanning path of $H_{g_1}^+$ that can be obtained from P_{g_1} by a sequence of rotations with fixed endpoint x, has at least $|\sigma_0^1|/43 > n/(4130m)$ vertices.

We claim that also in G, any vertex in S^{g_1} can be obtained as an endpoint by a sequence of rotations of R_1 with fixed endpoint x. The role of the vertex w will be played by $\hat{a}$ in G (note that we made sure that $\hat{a} \neq s_1$, so $\hat{a}$ is not contained in the union of the segments.) Hence, the edge (w, g_1) is present in G, while we will consider the edge (w, s_1) *artificial*.

For any endpoint $z \in S^{g_1}$ there is a sequence of pivots, such that performing the sequence of rotations with fixed endpoint x at these pivots results in an x, z-path spanning $H_{g_1}^+$. We claim that in $G[V(R_1)]$ it is also possible to perform a sequence of rotations with the exact same pivot sequence and eventually end up in an x, z-path spanning $V(R_1)$. When performing these rotations, the subpath of R_1 that links $\hat{a}$ to the first vertex of I_{i_1} corresponds to the artificial edge (w, s_1) in $H_{g_1}^+$ and each subpath that links two consecutive segments of σ_0^1 corresponds to the appropriate artificial edge in $H_{g_1}^+$.

Problems in performing these rotations in G could arise if a rotation is called for where (1) the pivot is connected to the endpoint of the current spanning path via an artificial edge of $H_{g_1}^+$: this rotation might not be possible in G as this edge might not exist in G, or (2) the broken edge is artificial: after such a rotation in G the endpoint of the new spanning path might be different from the one we have after performing the same rotation in $H_{g_1}^+$.

However, the construction of $H_{g_1}^+$ ensures that these problems will never occur. Indeed, in both cases (1) and (2) the pivot vertex has an artificial edge incident with it, while having degree at least 3 (as all pivots). However, both endpoints of an artificial edge have degree 2 in $H_{g_1}^+$ (for this last assertion we use the fact that $g_1 \notin EP_1$; this is important as g_1 is the first pivot.)

Hence we have ensured that there is indeed a spanning path of $G[V(R_1)]$ from x to every vertex of $V_1 = S^{g_1}$.

Similarly, since there is an edge from $\hat{b}$ to a good vertex g_2 in H_2, we can rotate R_2, starting from this edge to get a set $V_2 = S^{g_2}$ of at least $n/(4130m)$ endpoints. In other words we have a spanning path of $G[V(R_2)]$ from y to every vertex of $V_2 = S^{g_2}$.

As we noted earlier, property P2 ensures that there is an edge between V_1 and V_2, which closes a maximal path of G into a cycle and then the Hamiltonicity of G follows from its connectedness. This concludes the proof of Lemma 2.4 and consequently, that of Theorem 1.1. □

2.4 Hamiltonicity with larger expansion

As we have mentioned, our Hamiltonicity criterion can be extended to handle graphs with a larger expansion than that postulated in Theorem 1.1 ($d \leq e^{\sqrt[3]{\log n}}$). In particular, using very similar arguments, we can prove the following statement.

Theorem 2.5 *Let $12 \leq d \leq \sqrt{n}$ and let G be a graph on n vertices satisfying the following two properties:*

P1' *For every $S \subset V$, if $|S| \leq \frac{n \log d}{d \log n}$ then $|N(S)| \geq d|S|$;*

P2' *There is an edge in G between any two disjoint subsets $A, B \subseteq V$ such that $|A|, |B| \geq \frac{n \log d}{1035 \log n}$.*

Then G is Hamiltonian, for sufficiently large n.

The proof of Theorem 2.5 is almost identical to that of Theorem 1.1 given above. The only notable difference is that here one can take $\tau = 2$ in the proof of Lemma 2.4.

3 Corollaries

In this section we prove the aforementioned corollaries of Theorem 1.1.

Proof of Theorem 1.2 Let $G_{uv} = (V, E \cup \{(u, v)\})$; clearly G_{uv} satisfies properties P1 and P2 and is therefore Hamiltonian by Theorem 1.1. Let $C = w_1 w_2 \ldots w_n w_1$ be a Hamilton cycle in G_{uv}. If (u, v) is an edge of C, remove it to obtain the desired path in G. Otherwise, assuming that $u = w_i$ and $v = w_j$, add (u, v) to $E(C)$ and remove (u, w_{i+1}) and (v, w_{j+1}), where all indices are taken modulo n, to obtain a Hamilton path of G_{uv} that contains the edge (u, v); denote this path by P. We will close P into a Hamilton cycle that includes (u, v); removing this edge will result in the required path. The building of the cycle will be done as in the proof of Theorem 1.1 Section 2.3, with P as P_0, while making sure that (u, v) is never broken. The proof is essentially the same, except for the following minor changes:

1. When dividing P into 2ρ segments, we will make sure that (u, v) is in one of the segments; denote it by I_j.
2. When considering τ-sequences, we will restrict ourselves to those that include I_j.
3. Assume without loss of generality that $I_j \in \sigma_0^1$. When building H_1 (and later, when rotating R_1 according to the model of H_1) we will ignore I_j, that is, we will replace it by a single edge (a, b) where a is the last vertex of I_{j-1} (or $\hat{a}$ if $j = 1$) and b is the first vertex of I_{j+1} (or x if $j = \tau/2$).

□

Proof of Theorem 1.3

Fix some $\frac{8n \log n}{t \log \log n} \leq k \leq n - 3n/t$. Let $V_0 \subseteq V$ be an arbitrary subset of size $k + 3n/t$. We construct a sequence of subsets $S_i \subseteq V_0$. First, let $S_0 = \emptyset$. As long as $|S_i| < n/t$ and there exists a set $A_i \subseteq V_0 \setminus S_i$ such that $|A_i| \leq n/t$ but $|N_{G[V_0 \setminus S_i]}(A_i)| < |A_i| \frac{4 \log n}{\log \log n}$, we define $S_{i+1} := S_i \cup A_i$. Let q be the smallest integer such that either $|S_q| \geq n/t$ or $|N_{G[V_0 \setminus S_q]}(A)| \geq |A| \frac{4 \log n}{\log \log n}$ for every $A \subseteq V_0 \setminus S_q$ of size at most n/t. We claim that $|S_q| < n/t$. Indeed assume for the sake of contradiction that $|S_q| \geq n/t$. Since we halt the process as soon as this occurs, and $|A_{q-1}| \leq n/t$, we have $|S_q| < 2n/t$. For every $0 \leq i \leq q - 1$ we have $|N_{G[V_0 \setminus S_i]}(A_i)| < |A_i| \frac{4 \log n}{\log \log n}$ and so $|N_{G[V_0]}(S_q)| < |S_q| \frac{4 \log n}{\log \log n}$. On the other hand, the fact that G satisfies property P2 together with our lower bound on k implies $|N_{G[V_0]}(S_q)| > |V_0| - n/t - |S_q| \geq |V_0| - 3n/t \geq k \geq |S_q| \frac{4 \log n}{\log \log n}$, a contradiction.

Hence, $|S_q| < n/t$ and so, for $U = V_0 \setminus S_q$, $G[U]$ satisfies an expansion condition similar to **P1**, that is, for every $A \subseteq U$, if $|A| \leq n/t$ then $|N_{G[U]}(A)| \geq 4|A| \frac{\log n}{\log \log n}$.

In the following we prove that with positive probability the induced subgraph of G on a random k-element subset of U also satisfies a condition similar to **P1**. Let K be a k-subset of U drawn uniformly at random. We will prove that, with positive probability, $G[K]$ satisfies the following:

P1 For every $A \subseteq K$, if $|A| \leq n/t$ then $|N_{G[K]}(A)| \geq 2|A|\frac{\log n}{\log\log n}$.

Let $r = |U| - k$. Note that $0 \leq r \leq 3n/t$ and moreover $r < |U|$. Let $A \subseteq U$ be any set of size $a \leq n/t$, then, as was noted above, $|N_{G[U]}(A)| \geq 4a\frac{\log n}{\log\log n}$. Let $N_0 \subseteq N_{G[U]}(A)$ be an arbitrary subset of size $4a\frac{\log n}{\log\log n}$. If $A \subseteq K$ and $|N_{G[K]}(A)| \leq 2a\frac{\log n}{\log\log n}$, then K misses at least $2a\frac{\log n}{\log\log n}$ vertices from N_0. This can occur with probability at most

$$\begin{aligned}
\frac{\binom{|N_0|}{\frac{2a\log n}{\log\log n}}\binom{|U|-\frac{2a\log n}{\log\log n}}{r-\frac{2a\log n}{\log\log n}}}{\binom{|U|}{r}} &\leq \binom{\frac{4a\log n}{\log\log n}}{\frac{2a\log n}{\log\log n}}\left(\frac{r}{|U|}\right)^{\frac{2a\log n}{\log\log n}} \\
&\leq 2^{\frac{4a\log n}{\log\log n}}\left(\frac{\frac{3n}{t}}{\frac{8n\log n}{t\log\log n}}\right)^{\frac{2a\log n}{\log\log n}} \\
&= \left(\frac{3\log\log n}{2\log n}\right)^{\frac{2a\log n}{\log\log n}}.
\end{aligned}$$

Note that the latter bound is $o(\frac{1}{n})$ for $a = 1$, and $o(\frac{1}{n}\binom{n}{a}^{-1})$ for every $a \geq 2$.

It follows by a union bound argument that

$$Pr\left[\text{there exists an } A \subseteq K \text{ such that } |A| \leq n/t \text{ but } N_{G[K]}(A) < \frac{2\log n}{\log\log n}|A|\right] = o(1).$$

Hence, there exists a k-subset X of U such that for every $A \subseteq X$, if $|A| \leq n/t$ then $|N_{G[X]}(A)| \geq \frac{2\log n}{\log\log n}|A|$. Moreover, if A, B are disjoint subsets of V, and $|A|,|B| \geq \frac{k\log\log k\log\left(\frac{2\log n}{\log\log n}\right)}{4130\log k\log\log\log k} \geq n/t$ then there is an edge between a vertex of A and a vertex of B.

Thus $G[X]$ satisfies the conditions of Theorem 1.1 with $|V| = k$ and $d = \frac{2\log n}{\log\log n}$ and is therefore Hamiltonian. It follows that G admits a cycle of length exactly k.

□

Proof of Theorem 1.4 Let $G = G(n,p) = (V,E)$ and let $d = (\log n)^{0.1}$. We begin by showing that a.s. G satisfies property P2 with respect to d. Indeed

$$
\begin{aligned}
Pr[G \not\models P2] &\leq \left(\frac{n}{\frac{n \log\log n \log d}{4130 \log n \log\log\log n}}\right)^2 \left(1 - \frac{\log n + \log\log n + \omega(1)}{n}\right)^{\left(\frac{n \log\log n \log d}{4130 \log n \log\log\log n}\right)^2} \\
&\leq \left(\frac{4130 e \log n \log\log\log n}{0.1(\log\log n)^2}\right)^{\frac{0.2 n (\log\log n)^2}{4130 \log n \log\log\log n}} \\
&\quad \times \exp\left\{-\frac{\log n + \log\log n + \omega(1)}{n} \cdot \frac{0.01 n^2 (\log\log n)^4}{4130^2 (\log n)^2 (\log\log\log n)^2}\right\} \\
&= o(1).
\end{aligned}
$$

Next, we deal with property P1. Since a.s. there are vertices of "low" degree in G, we cannot expect every "small" set to expand by a factor of d. Therefore, to handle this difficulty, we introduce some minor changes to the proof of Theorem 1.1, in fact only to the part included in Claim 2.2. First of all, note that a.s. G is connected (this fact replaces Proposition 2.1). Let $SMALL = \{u \in V : d_G(u) \leq (\log n)^{0.2}\}$ denote the set of all vertices of G that have a "low" degree. The vertices in $SMALL$ will be called *small vertices*. Standard calculations show that a.s. G satisfies the following properties:

(1) $\delta(G) \geq 2$.

(2) For every $u \neq v \in SMALL$ we have $dist_G(u, v) \geq 250$, where $dist_G(u, v)$ is the number of edges in a shortest path between u and v in G.

(3) G satisfies a weak version of P1, that is, if $A \subseteq V \setminus SMALL$ and $|A| \leq \frac{n \log\log n \log d}{d \log n \log\log\log n}$ then $|N_G(A)| \geq 3d|A|$.

(4) The number of vertices of degree at most 11 in G is $O(\log^{11} n)$.

We will prove that, based on these properties, we can build initial long paths as in Claim 2.2; this will conclude our proof of Theorem 1.4, as in Subsections 2.2 and 2.3 we did not rely on property P1. The argument is essentially the same as in Claim 2.2; the main difference is that we will use roughly twice as many rotations to create the eventual endpoint set of size $n/3$. This extra factor of two has no real effect on the rest of the proof.

Suppose first that the initial path of maximum length P_0 is such that, while creating the sets $S_0, S_1, \ldots, S_{120}$ as we did in the proof of Claim 2.2, no vertex from $\cup_{i=0}^{119} S_i$ is a small vertex. Then, by (3), like in the proof of Claim 2.2, after the ith rotation there are exactly $(3d/3)^i = (\log n)^{0.1i}$ new endpoints in S_i. Therefore, after 120 rotations we will have an endpoint set S_{120} with $(\log n)^{12}$ elements.

Suppose now that there is a vertex $u \in S_j \cap SMALL$ for some $0 \leq j \leq 119$. Let P_u denote a path of maximum length from v_1 to u (which can be obtained from P_0 by at most 119 rotations). At this point we ignore the endpoint sets S_i, $i \leq j$ created so far and restart creating them. The first rotation is somewhat special. By property (1), u has at least one neighbor on P_u other than its predecessor. Thus we can rotate P_u once and obtain a v_1, w-path P_w of maximum length, such that w is at distance two from a small vertex. Then we create new endpoint sets $S_1, \ldots, S_{120}$ with P_w as the initial path. Note that property (2) implies $w \notin SMALL$. Since a new endpoint is always at distance at most two from the old endpoint, we can rotate another 120 times without ever creating an endpoint which is a small vertex. Thus, property (3) applies, and after the ith rotation (not including the one that turned w into an endpoint), $i \leq 120$, there are exactly $(3d/3)^i = (\log n)^{0.1i}$ new endpoints in S_i. Hence, after these 120 new rotations we obtain a set S_{121} of size exactly $(\log n)^{12}$. Altogether we used up to 240 rotations.

In the following we will prove that the endpoint sets we build grow by a multiplicative factor of $d/3$ every at most *two* rotations.

We will prove by induction on t that there exist endpoint sets $S_{121}, S_{122}, \ldots$ such that for every $t \geq 122$, either $|S_t| = \frac{d}{3}|S_{t-1}|$ or $|S_t| = |S_{t-1}| = \frac{d}{3}|S_{t-2}|$.

We will show that this implies $\sum_{i=0}^{t} |S_i| \leq \frac{4}{3}|S_t|$ if $|S_t| = \frac{d}{3}|S_{t-1}|$, provided that n is sufficiently large.

For the base case we just have to note that $\sum_{i=0}^{121} |S_i| \leq \frac{4}{3}|S_{121}|$. Suppose we have already built S_t for some $t \geq 121$ such that $\sum_{i=0}^{t} |S_i| \leq \frac{4}{3}|S_t|$ and now wish to build S_{t+1}. We will proceed as in the proof of Claim 2.2.

Assume first that $|N(S_t)| \geq d|S_t|$. Then, as in the proof of Claim 2.2

$$|\hat{S}_{t+1}| \geq \frac{1}{2}(d|S_t| - 3 \cdot \frac{4}{3}|S_t|) = \frac{d-4}{2}|S_t|.$$

Hence, a subset $S_{t+1} \subseteq \hat{S}_{t+1}$ with $|S_{t+1}| = \frac{d}{3}|S_t|$ can be selected.

Assume now that $|N(S_t)| < d|S_t|$. By (3), this must mean that for $S'_t := S_t \cap SMALL$ we have $|S'_t| \geq \frac{2}{3}|S_t|$. Since $|S'_t| \gg \log^{11} n$, property (4) implies that almost every vertex of S'_t has degree at least 12. By (3), no two small vertices have a common neighbor, so $|N(S'_t)| \geq (12 - o(1))|S'_t| \geq (8 - o(1))|S_t|$. As in the proof of Claim 2.2, we have

$$|\hat{S}_{t+1}| \geq \frac{1}{2}(|N(S'_t)| - 3 \cdot |\cup_{i=0}^{t} S_i|) \geq \frac{1}{2}((8 - o(1))|S_t| - 3 \cdot \frac{4}{3}|S_t|) \geq |S_t|.$$

Hence we can select a subset $S_{t+1} \subseteq \hat{S}_{t+1}$ such that $|S_{t+1}| = |S_t|$. Note that, since we only used vertices from S'_t for further rotation, all the new endpoints in S_{t+1} are at distance two from a small vertex. It follows by property (2) that $S_{t+1} \cap SMALL = \emptyset$. Hence $|N(S_{t+1})| \geq 3d|S_{t+1}|$ by property (3), which implies that after the next rotation we will have

$$|\hat{S}_{t+2}| \geq \frac{1}{2}(3d|S_{t+1}| - 3(\frac{4}{3}|S_t| + |S_{t+1}|)) = \frac{3d-7}{2}|S_t|.$$

Hence, a subset $S_{t+2} \subseteq \hat{S}_{t+2}$ with $|S_{t+2}| = \frac{d}{3}|S_t|$ can be selected.

For the last rotation, our calculations are identical to the ones in Claim 2.2 as those depend on the expansion properties implied by condition P2.

In conclusion, we created an endpoint set $B(v_1)$ of size at least $n/3$ such that for every $v \in B(v_1)$ there is a v_1, v-path of maximum length which can be obtained from P_0 by at most $240 + \frac{4 \log n}{\log d}$ rotations with fixed endpoint v_1. □

Proof of Theorem 1.5

Let $G = (V, E)$ be f-connected where $f(k) = 12e^{12} + k \log k$. We prove that G satisfies conditions P1 and P2 with $d = 12$ and apply Theorem 1.1 to conclude that G is Hamiltonian for sufficiently large n. Let $A \subseteq V$ be of size at most $\frac{n}{12m}$. Either $|A| > |V \setminus (A \cup N(A))|$ and so in particular $|N(A)| \geq 12|A|$ if n is sufficiently large, or the pair $(A \cup N(A), V \setminus A)$ is a separation of G with $|A| \leq |V \setminus (A \cup N(A))|$ and so by our assumption $|(A \cup N(A)) \cap (V \setminus A)| = |N(A)| \geq f(|A|) \geq 12e^{12} + |A| \log |A| \geq 12|A|$. It follows that G satisfies property $P1$ with $d = 12$. Let A, B be two disjoint subsets of V such that $|B| \geq |A| \geq \frac{n}{4130m}$. Assume for the sake of contradiction that there is no edge in G between A and B; hence $(V \setminus B, V \setminus A)$ is a separation of G. By our assumption $|(V \setminus A) \cap (V \setminus B)| = |V \setminus (A \cup B)| \geq f(|A|) \geq |A| \log |A| > n$. This is clearly a contradiction and so G satisfies property $P2$ with $d = 12$. □

Acknowledgements

We would like to thank Deryk Osthus for suggesting to use our criterion to address the conjecture of [7]. We would also like to thank the anonymous referee for his helpful comments.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, The first occurrence of Hamilton cycles in random graphs, *Annals of Discrete Mathematics* **27** (1985), 173–178.

[2] N. Alon and A. Nussboim, k-wise independent random graphs, 2008 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008), 813–822.

[3] J. Beck, Random graphs and positional games on the complete graph, *Annals of Discrete Math.* **28** (1985) 7–13.

[4] J. Beck, **Combinatorial Games: Tic-Tac-Toe Theory**, Cambridge University Press, 2008.

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

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

[7] S. Brandt, H. Broersma, R. Diestel and M. Kriesell, Global connectivity and expansion: long cycles and factors in f-connected graphs, *Combinatorica* **26** (2006), 17–36.

[8] V. Chvátal and P. Erdős, A note on Hamiltonian circuits, *Discrete Math.* **2** (1972), 111–113.

[9] R. Diestel, **Graph Theory**, Springer New-York, 2^{nd} ed. 1999.

[10] G. Dirac, Some theorems on abstract graphs, *Proc. London Math. Society* **2** (1952), 69–81.

[11] P. Erdős and A. Rényi, On the evolution of random graphs, *Bull. Inst. Statist. Tokyo* **38** (1961), 343–347.

[12] A. Frieze, S. Vempala, and J. Vera, Logconcave random graphs, 2008 40th Annual ACM Symposium on Theory of Computing (STOC 2008), 779–788.

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

[14] R. J. Gould, Advances on the Hamiltonian problem - a survey, *Graphs and Combinatorics* **19** (2003), 7–52.

[15] D. Hefetz, M. Krivelevich and T. Szabó, Avoider-Enforcer games, *Journal of Combinatorial Theory, Ser. A.* **114** (2007) 840–853.

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

[17] A. D. Korshunov, Solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs, *Soviet Math. Dokl.* **17** (1976), 760–764.

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

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

[20] M. Krivelevich and T. Szabó, Biased positional games and small hypergraphs with large covers, *Electronic Journal of Combinatorics* **15**(1) (2008), publ. R70.

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

[22] B. Sudakov and J. Verstraëte, Cycle lengths in sparse graphs, *Combinatorica* **28** (2008), 357–372.

[23] J. Verstraëte, On arithmetic progressions of cycle lengths in graphs, *Combinatorics, Probability and Computing* **9** (2000), 369–373.