Hamilton cycles in highly connected expander graphs

Dan Hefetz * Michael Krivelevich † Tibor Szabó ‡

July 31, 2006

Abstract

In this paper we prove a sufficient condition for a graph to be hamiltonian. The condition is based on expansion and high connectivity and applies to relatively sparse graphs. It improves on previous results of similar nature. We then use it to prove several corollaries concerning cycles in graphs.

1 Introduction

A Hamilton cycle in a graph is a connected, spanning, 2-regular subgraph. A graph is called *hamiltonian* if it admits a Hamilton cycle. Hamiltonicity is one of the most central notions in Graph Theory, and many efforts have been devoted to obtain sufficient conditions for the existence of a Hamilton cycle (a "nice" necessary and sufficient condition should not be expected, as deciding whether a given graph admits a Hamilton cycle is known to be NP-complete). One of the first such results is the celebrated theorem of Dirac [7], 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. [9] 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 [5] proved that if $\kappa(G) \geq \alpha(G)$ (that is, the vertex connectivity of G is at least as large as the size of the largest independent set in G) then

*Department of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: dannyh@post.tau.ac.il. This paper is a part of the author's Ph.D. under the supervision of Prof. Michael Krivelevich.

†Department of Mathematics, 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 Israeli Science Foundation.

‡Institute of Theoretical Computer Science, ETH Zurich, CH-8092 Switzerland. Email: szabo@inf.ethz.ch.

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. However, there are graphs with as few as $\Theta(n^{3/2})$ edges to which it applies (one such example consists of $\sqrt{n}$ cliques on $\sqrt{n}$ vertices each, arranged in a cycle such that cliques $i \mod \sqrt{n}$ and $(i+1) \mod \sqrt{n}$ are connected by a perfect matching). When looking for sufficient conditions for hamiltonicity of sparse graphs, it is natural to look at random graphs with an appropriate edge probability. Many attempts where made to determine the threshold for hamiltonicity, starting with the breakthrough result of Posa [13]. Finally, Bollobás [2] proved that if $p = (\log n + \log\log n + \omega(1))/n$, where $\omega(1)$ tends to infinity with n arbitrarily slowly, then $G(n,p)$ is a.s. hamiltonian. Note that this is trivially best possible as if $p \leq (\log n + \log\log n - \omega(1))/n$ then a.s. there are vertices of degree one in $G(n,p)$. 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. [8, 12]). These are quite complicated as they rely on many properties of pseudo-random graphs. Furthermore, these conditions are not the most natural as hamiltonicity is a monotone increasing property, whereas pseudo-randomness is not. Our main result is natural and simple sufficient condition that is based on expansion and high connectivity. Before stating the result we explore the following properties of a graph $G = (V, E)$ where $|V| = n$:

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

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

Note that properties P1 and P2 together guarantee some expansion for every $S \subset V(G)$. 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 of size exactly $\frac{n}{dm}$ and so by property P1 expands at least to a size of $\frac{n}{m}$, that is it expands by a factor of at least 4130. Finally, if $|S| \geq \frac{n}{4130m}$ then $N(S) \geq (1 - o(1))n$ by property P2.

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 requirement $d \leq e^{\sqrt[3]{\log n}}$ makes sure (in particular) that $\frac{n}{4130m} = o(n)$ and so P2 is satisfied in a strong sense. Clearly, admitting a larger expansion cannot hurt us, but

it doesn't help us weaken P2 beyond a certain threshold. However, if we have a large expansion and want to make good use of it, we have the following variant of Theorem 1.1:

Theorem 1.2 *Let $12 \leq d \leq \sqrt{n}$ and let $G = (V, E)$ be a graph on n vertices satisfying the following two properties:*

P1' *For every $S \subseteq V$, if $|S| \leq \frac{n \log d}{1035 d \log n}$ 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 Theorems 1.1 and 1.2 is based on the ingenious rotation-extension technique, developed by Posa [13].

Theorem 1.1 was used by the authors (see [10]) to prove that Enforcer can win the $(1, q)$ 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 is presently the best known bound. A similar result can be obtained for Maker in the corresponding Maker-Breaker game. This falls short of the true value of the critical bias of $\frac{cn}{\log n}$ obtained by Beck (see [1]), but our proof is simpler and shorter. 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 1.3 *Let $G = (V, E)$ be a graph that satisfies properties P1 and P2; then G is Hamilton-connected.*

Remark. An immediate consequence of Theorem 1.3 is that for every $e \in E$ 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 that satisfies property P2, admits a cycle of "almost any length".

Theorem 1.4 *Let $G = (V, E)$, where $|V| = n$, be a graph, satisfying property P2; more precisely, for every disjoint subsets $A, B \subseteq V$ such that $|A|, |B| \geq n/t$, where $t = t(n)$, 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 $\frac{3n \log n}{t \log \log n} \leq k \leq n - 2n/t$.*

Remark. The upper bound on k in Theorem 1.4 is tight up to the constant factor 2 in the second order term, as is shown by the 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.

The following question appears in [9]:
If we have some sufficient condition for hamiltonicity, can it be used to prove the existence of a 2-factor with exactly k cycles for every $1 \leq k \leq f(n)$ for some function f which is as large as possible?

We prove that property P1 together with a strengthening of property P2 yield the desired 2-factor for almost every "reasonable" function f:

Theorem 1.5 *Let $G = (V, E)$, where $|V| = n$, be a graph and let $r = r(n)$ be a positive integer. Assume that G satisfies the following 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 $|N(S)| \geq d|S|$, where $d \geq 300r \log n$;*

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

Then G admits a 2-factor with exactly k cycles for every $1 \leq k \leq r$.

Theorem 1.1 (with minor changes) can be used to prove the following classic result (c.f. [2]).

Theorem 1.6 *$G(n, p)$, where $p = (\log n + \log \log n + \omega(1))/n$, is a.s. hamiltonian.*

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 $|A \setminus B| \leq |B \setminus A|$. In [4] it was proved that if $f(k) \geq 2(k+1)^2$ for every $k \in \mathbb{N}$ then G is hamiltonian for every $n \geq 3$. It was also 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 prove this conjecture for sufficiently large n:

Theorem 1.7 *If $G = (V, E)$, where $|V| = n$, is f-connected for $f(k) = 12 + k \log k$, 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 [6].

The rest of the paper is organized as follows: in Section 2 we prove Theorems 1.1 and 1.2, in Section 3 we prove it's corollaries: Theorems 1.3, 1.4, 1.5, 1.6 and 1.7.

2 Proof of the main result

Throughout this section we assume that G satisfies properties P1 and P2.

Proposition 2.1 *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. □

2.1 Constructing an initial long path

We use rotations and extensions in G to find a maximal path with large rotation endpoint sets, see for example [3], [8], [11], [12]. Let $P_0 = (v_1, v_2, \ldots, v_l)$ be a path of maximum length in G. If $1 \leq i \leq l-2$ and (v_l, v_i) is an edge of G then $P' = (v_1 v_2 \ldots v_i v_l v_{l-1} \ldots v_{i+1})$ is also of maximum length. It 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 can then, in general, rotate P' to get more maximum length paths.

Claim 2.2 *There exists a set $B(v_1)$ of at least $n/3$ vertices, such that for every $v \in B(v_1)$ there is a v_1v-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.*

Proof Let t_0 be the smallest integer such that $\left(\frac{d}{3}\right)^{t_0-2} > \frac{n}{md}$. Note that $t_0 \leq 2\frac{\log n}{\log d}$.

We prove that there exists 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 $0 \leq t \leq t_0$, every $v \in S_t$ is the endpoint of a path, obtainable from P_0 by t rotations with fixed endpoint v_1, such that for every $0 \leq i \leq t$, after the ith rotation the non-v_1-endpoint of the path is in S_i, and $|S_t| = \left(\frac{d}{3}\right)^t$ for every $t \leq t_0 - 3$, $|S_{t_0-2}| = \frac{n}{dm}$, $|S_{t_0-1}| = \frac{n}{4130m}$, and $|S_{t_0}| \geq n/3$.

First we construct the sets by induction on t. For $t = 0$, one can choose $S_0 = \{v_l\}$ and all requirements are trivially satisfied.

Induction step: let $0 \leq t < t_0 - 2$ and assume that the appropriate sets $S_0, \ldots, S_t$ were already constructed. We will now construct S_{t+1}. Let

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

Assume now that $v_i \in T$, $y \in S_t$ and $(v_i, y) \in E$. Then a v_1, y-path Q on l vertices can be produced from P_0 by t rotations such that all broken edges are in P_0 and the non-v_1-endpoint after the jth rotation is in S_j. Since $v_{i-1}, v_i, v_{i+1} \notin \bigcup_{j=0}^{t} S_j$, and each rotation breaks an edge of P_0 between a new endpoint and its neighbor along P_0, both edges (v_{i-1}, v_i) and (v_i, v_{i+1}) are present in Q (note that both are also edges of the original path P_0). Rotating Q with pivot v_i will make one of v_{i-1}, v_{i+1} an endpoint. Assume w.l.o.g. it is v_{i-1} and place it into the candidate set $\hat{S}_{t+1}$. The only other vertex that can place v_{i-1} into $\hat{S}_{t+1}$ is v_{i-2} (if it exists). Thus,

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

where the last inequality follows since $d \geq 12$. Clearly we can delete arbitrary elements of $\hat{S}_{t+1}$ to obtain S_{t+1} of size exactly $\left(\frac{d}{3}\right)^{t+1}$ if $t < t_0 - 3$ and of size exactly $\frac{n}{dm}$ if $t = t_0 - 3$. 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 similar methods. We define $\hat{S}_{t_0-1}$ similarly to the above. The calculation goes as follows:

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

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

To construct S_{t_0} we use a similar procedure. The difference is that we use the expansion guaranteed by property P2 instead of the one guaranteed by property P1. The calculation goes as follows.

$$\begin{aligned} |S_{t_0}| &\geq \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t_0-1})| - 3(1 + |S_1| + \ldots + |S_{t_0-2}| + |S_{t_0-1}|)\right) \\ &\geq \frac{n}{2}(1 - o(1)) - \frac{3}{2}\left(\frac{(d/3)^{t_0-3} - 1}{(d/3) - 1} + \frac{2n}{dm} + \frac{n}{4130m}\right) \\ &\geq \frac{n}{2}(1 - o(1)) - \left(\frac{n}{2m} + \frac{3n}{dm} + \frac{3n}{8260m}\right) \geq \frac{n}{3}, \end{aligned}$$

where the last inequality follows since $n/m = o(n)$. This concludes the proof of the claim. □

Note that in particular, $l = |V(P_0)| \geq n/3$.

Let $A_0 = B(v_1) \cup \{v_1\}$. Similarly to the above, for each $v \in B(v_1)$, we can construct a set $B(v)$, $|B(v)| \geq n/3$, of endpoints of maximum length paths with endpoint v, obtained from a maximum length path between v_1 and v by at most t_0 rotations.

In summary, for every $a \in A_0$, $b \in B(a)$ there is a maximum length path $P(a,b)$ joining a and b, which is obtainable from P_0 by at most $\rho = 2t_0 \leq \frac{4\log n}{\log d}$ rotations.

2.2 Finding many good initial rotations

Using an argument, similar to the one used in the proof of Claim 2.2, we can show (based on a condition similar to property P2) that any sufficiently long path, has many vertices such that if we start rotating from any of these vertices, we get a large number of endpoints.

Let H be a graph on l vertices. Let $P = (v_1, \ldots, v_l)$ be a spanning path in H. Assume that v_l has degree 1 in H. For a vertex $v_i \in V(H) \setminus \{v_l\}$ define: $S_0^{v_i} = \{v_l\}$, $S_1^{v_i} = \{v_{i+1}\}$, and for every $t \geq 2$, $S_t^{v_i} = \{v \in P : v \neq v_1$ and v is the endpoint of a path obtainable from P in $H \cup \{(v_i, v_l)\}$ by t rotations with fixed endpoint v_1 and without breaking the edge (v_{l-1}, v_l), where the first rotation is with pivot $v_i\}$.

A vertex $v_i \in V(P)$ is called a *bad initial pivot* (or simply a *bad vertex*) if $|\bigcup_{t\geq 1} S_t^{v_i}| < \frac{l}{43}$ and is called a *good initial pivot* (or a *good vertex*) otherwise. When rotating P starting with a good initial pivot, we are able to find a large number of endpoints. Let R be the set of all bad vertices of P.

Lemma 2.3 *Let H, P and R be as above. Assume that every two disjoint sets A, B of vertices of H of sizes $|A|, |B| \geq l/43$ are connected by an edge in H; then $|R| \leq 7l/43$.*

Proof Let $R = \{v_{i_1}, \ldots, v_{i_r}\}$. 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}.

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

where $ext(X)$ denotes the set containing the vertices of X together with their left and right neighbors on P. Clearly $|ext(X)| \leq 3|X|$.

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 inductively, such that for every t the following hold.

(a) $W_1 = \{v_{i_j+1}\}$;

(b) $W_t \subseteq \tilde{S}_t^{v_{i_j}}$;

(c) $|W_t| = 2|W_{t-1}|$;

(d) $W_t \cap \left(\cup_{s=1}^{t-1} W_s \cup X\right) = \emptyset$,

where $\tilde{S}_t^{v_{i_j}}$ contains those vertices y of $S_t^{v_{i_j}}$ for which a spanning path of $H \cup \{(v_{i_j}, v_l)\}$ ending at y can be produced from P by t rotations with fixed endpoint v_1, such that not only the initial pivot is v_{i_j} but even more generally, after the s^{th} rotation the new endpoint is in W_s, for every $s \leq t$.

As a start we set $W_1 = \{v_{i_j+1}\}$. Assume now that we have constructed $W_1, \ldots, W_t$ with properties $(a) - (d)$. If $|N_H(W_t) \setminus ext((\cup_{i=1}^{t} W_i \cup X))| > 5|W_t|$, then we create W_{t+1} with properties $(a) - (d)$, otherwise we finish the processing of v_{i_j} by updating U and X.

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

Let $v_i \in T$, $v_i \neq v_1, v_{l-1}$, and suppose that v_i is adjacent to $y \in W_t$. Recall, that by property (b) $y \in \tilde{S}_t^{v_{i_j}}$, so a path Q of $H \cup \{(v_{i_j}, v_l)\}$ ending at y can be produced from P by t rotations, such that for every $s \leq t$, after the s^{th} rotation the new endpoint is in W_s. The choice of v_i ensures that v_{i-1}, v_i and $v_{i+1} \notin \bigcup_{s=1}^{t} W_s$. Since each rotation breaks an edge of P between a new endpoint and its neighbor along P, both edges (v_{i-1}, v_i) and (v_i, v_{i+1}) are present in Q. Rotating Q with pivot v_i will place one of v_{i-1}, v_{i+1} in $\tilde{S}_{t+1}^{v_{i_j}}$. Denote this vertex by v_i'. We define $W_{t+1} = \{v_i' : v_i \in T\}$. Observe that besides v_i the only other vertex that can place v_i' into $\tilde{S}_{t+1}^{v_{i_j}}$ (and thus into W_{t+1}) is its other neighbor on the path P. Thus,

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

Deleting arbitrarily some vertices from W_{t+1} we can make sure that its cardinality is exactly $2|W_t|$. Property (b) is satisfied by the definition of W_{t+1}. Property (d) is satisfied because $v_i \notin ext\,(\cup_{s=1}^{t} W_s \cup X)$ ensures that none of its neighbors on P, in particular v_i', is a member of $(\cup_{s=1}^{t} W_s \cup X)$.

Property (c) ensures that $|\cup_{s=1}^{t} W_s|$ is strictly increasing so the processing of vertex v_{i_j} is bound to stop after a finite number of steps. In other words at some point we will have $|T| \leq 5|W_k|$ for some index k. Then we update U and X by adding W_k to U and adding $W_1 \cup \cdots \cup W_k \cup T$ to X. We have to check that the conditions of (1) hold. Observe that $|W_1 \cup \cdots \cup W_k| < 2|W_k|$, so the number of vertices added to X is at most seven times more than the number of vertices added to U. Also, property (d) and $U \subseteq X$ made sure

that W_k was disjoint from U, so indeed the property $|U| \geq |X|/7$ is still valid. The other conditions in (1) follow easily from the definition of the "new" U and X and from property (a). 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 $|\bigcup_{t\geq 1} S_t^{v_{i_j}}|$ 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 processing 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

We consider P_0 to be directed and divided into 2ρ segments $I_1, I_2, \ldots, I_{2\rho}$ of length at least $\lfloor |P_0|/2\rho \rfloor$ each, where $|P_0| \geq n/3$. As each $P(a,b)$ is obtained from P_0 by at most ρ rotations and every rotation breaks exactly one edge of P_0, the number of segments of P_0 which occur complete on this path, 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.

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 these 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 $\alpha = \alpha(\tau) = 1/9(4\tau)^{-\tau}$. 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 $\hat{A} = \{a \in P_0 : L$ contains at least $\alpha n/2$ pairs with a as first element$\}$. Then $|\hat{A}| \geq \alpha n/2$. 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. Both sub-sequences σ_0^1 and σ_0^2 have at least $\tau/2 \cdot n/(6\rho) \geq \frac{n}{96m}$ vertices. 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). Now we define the notion of good vertices in σ_0^1 and σ_0^2. For σ_0^1 construct a graph H_1 from the segments of σ_0^1 by adding a (virtual) vertex h_1 to $V(\sigma_0^1)$, joining h_1 to the first vertex of I_{i_1}, joining the last vertex of I_{i_j} to the first vertex of $I_{i_{j+1}}$ for every $1 \leq j < \tau/2$ and then adding the edges of G with both endpoints in the interior (that is, not endpoints) of segments of σ_0^1 to H_1. Then the segments of σ_0^1 with the edges linking them and h_1 form an oriented spanning path in H_1, starting at x. We define *good* vertices in σ_0^1 to be the vertices which are not endpoints of any segment of σ_0^1 and are good vertices of H_1 as defined in Section 2.2, with $l = \sum_{j=1}^{\tau/2} |I_{i_j}| + 1$. Due to property P2, Lemma 2.3 applies here, and so, since $\tau = o(|\sigma_0^1|)$, more than half 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 adding a (virtual) vertex h_2 to $V(\sigma_0^2)$, joining h_2 to the last vertex of I_{i_τ}, 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 then adding the 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 and h_2 form an oriented spanning path in H_2, starting at y. We define *good* vertices in σ_0^2 to be the vertices which are not endpoints of any segment of σ_0^2 and are good vertices of H_2 as defined in Section 2.2, with $l = \sum_{j=\tau/2+1}^{\tau} |I_{i_j}| + 1$. Due to property P2, Lemma 2.3 applies here, and so, since $\tau = o(|\sigma_0^2|)$, more than half of the vertices of σ_0^2 are good.

Since $|\hat{A}| \geq \alpha n/2 \geq \frac{n}{4130m}$ (which is why we get the upper bound on d in Theorem 1.1) and σ_0^1 has at least $|\sigma_0^1|/2 > \frac{n}{192m}$ good vertices, there is an edge from a vertex $\hat{a} \in \hat{A}$ to a good vertex in σ_0^1. Similarly, as $|\hat{B}(\hat{a})| \geq \alpha n/2$,there is an edge from some $\hat{b} \in \hat{B}(\hat{a})$ to a good vertex in σ_0^2. Consider the path $\hat{P} = P(\hat{a}, \hat{b})$ of maximum length connecting $\hat{a}$ and $\hat{b}$ and containing σ_0. The vertices x and y split this path into three sub-paths: P_1 from $\hat{a}$ to x, P_2 from y to $\hat{b}$ and P_3 from x to y. We will rotate P_1 with x as a fixed endpoint and P_2 with y as a fixed endpoint. We will show that the obtained endpoint sets V_1 and V_2 are sufficiently large. Then by property P2 there will be an edge of G between V_1 and V_2. Since we did not touch P_3, this edge closes a maximum path into a cycle, which is hamiltonian due to the connectivity of G.

Since there is an edge from $\hat{a}$ to a good vertex in σ_0^1, by the definition of a good vertex we can rotate P_1, starting from this edge, to get a set V_1 of at least $|\sigma_0^1|/43 > n/(4130m)$ endpoints. When doing this, we will treat the subpath that links $\hat{a}$ and the first vertex of I_{i_1} and each subpath that links two consecutive segments of σ_0^1, as single edges and ignore edges of G that are incident with an endpoint of some segment of σ_0^1 - like in H_1. This ensures that all rotations and broken edges are inside segments of σ_0^1 and so there is indeed

a path of the appropriate length from x to every vertex of V_1.

Similarly, since there is an edge from $\hat{b}$ to a good vertex in σ_0^2, we can rotate P_2, starting from this edge to get a set V_2 of at least $|\sigma_0^2|/43 > n/(4130m)$ endpoints. When doing this, we will treat the subpath that links $\hat{b}$ and the last vertex of I_{i_τ} and each subpath that links two consecutive segments of σ_0^2, as single edges and ignore edges of G that are incident with an endpoint of some segment of σ_0^2 - like in H_2. This ensures that all rotations and broken edges are inside segments of σ_0^2 and so there is indeed a path of the appropriate length from y to every vertex of V_2.

This concludes the proof of Theorem 1.1. □

The proof of Theorem 1.2 is almost the same. The only difference is that we take $\tau = 2$.

3 Corollaries

In this section we prove the corollaries of Theorem 1.1.

Proof of Theorem 1.3 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 $C = w_1 w_2 \dots w_n w_1$ 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 not 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 P_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 h_1 if $j = 1$) and b is the first vertex of I_{j+1} (or x if $j = \tau/2$).

□

Proof of Theorem 1.4 Fix some $\frac{3n\log n}{t\log\log n} \leq k \leq n - 2n/t$. Let $V_0 \subseteq V$ be an arbitrary subset of size $k + 2n/t$ and 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 $|N_{G[V_0 \setminus S_i]}(A_i)| < |A_i| \frac{\log n}{\log \log n}$, we update $S_{i+1} = S_i \cup A_i$. Let q be the smallest integer such that $|S_q| \geq n/t$ or $|N_{G[V_0 \setminus S_q]}(A)| \geq |A| \frac{\log n}{\log \log n}$ 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 $|A_i| \leq n/t$ for every $1 \leq i \leq q$, we have $|S_q| < 2n/t$. For every $1 \leq i \leq q$ we have $|N_{G[V_0 \setminus S_i]}(A_i)| < |A_i| \frac{\log n}{\log \log n}$ and so $|N_{G[V_0]}(S_q)| < |S_q| \frac{\log n}{\log \log n} + |S_q|$. But G satisfies property P2 and so $|N_{G[V_0]}(S_q)| > |V_0| - 2n/t \geq k \geq |S_q| \frac{\log n}{\log \log n} + |S_q|$, where the last inequality follows from our lower bound on k.

Let $U = V_0 \setminus S_q$ and let $r = |U| - k$. Note that $0 \leq r \leq 2n/t$. Let R be an r-subset of U drawn uniformly at random. We will prove that, with positive probability, $G[U \setminus R]$ satisfies property P1, that is:

P1 For every $A \subseteq U \setminus R$, if $|A| \leq n/t$ then $|N_{G[U \setminus R]}(A)| \geq 1/2 |A| \frac{\log n}{\log \log n}$.

Note that $G[U]$ satisfies a slightly stronger expansion condition, that is, for every $A \subseteq U$, if $|A| \leq n/t$ then $|N_{G[U]}(A)| \geq |A| \frac{\log n}{\log \log n}$.

Let $A \subseteq U \setminus R$ be any set of size $a \leq n/t$, then

$$\begin{aligned}
Pr[|N_{G[U \setminus R]}(A)| \leq \frac{a \log n}{2 \log \log n}] &\leq \binom{\frac{a \log n}{\log \log n}}{\frac{a \log n}{2 \log \log n}} \left(\frac{r}{|U|} \right)^{\frac{a \log n}{2 \log \log n}} \\
&\leq 2^{\frac{a \log n}{\log \log n}} \left(\frac{\frac{2n}{t}}{\frac{3n \log n}{t \log \log n}} \right)^{\frac{a \log n}{2 \log \log n}} \\
&\leq \left(\frac{8 \log \log n}{3 \log n} \right)^{\frac{a \log n}{2 \log \log n}} \\
&\leq \frac{1}{n} \binom{n}{a}^{-1}.
\end{aligned}$$

It follows by a union bound argument that

$$Pr[\text{ there exists an } A \subseteq U \setminus R \text{ such that } |A| \leq n/t \text{ but } N_{G[U \setminus R]}(A) < \frac{\log n}{2 \log \log n} |A|] < 1.$$

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

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

Proof of Theorem 1.5 We will use the following proposition that may be of independent interest.

Proposition 3.1 *Let $\varepsilon > 0$ and let $G = (V, E)$, where $|V| = n = n(\varepsilon)$, be a graph that satisfies property P1 with parameter $d \geq 3\varepsilon^{-2}k \log n$, where $k = k(n)$ is a positive integer. Then there exists a partition $V = \bigcup_{i=1}^{k} V_i$ such that for every $1 \leq i \leq k$:*

(i) $(1/k - \varepsilon)n \leq |V_i| \leq (1/k + \varepsilon)n$.

(ii) $G[V_i]$ *satisfies property P1 with parameter* $(1 - \varepsilon)d/k$.

Proof For every vertex $v \in V$, choose uniformly and independently at random an integer $i \in \{1, \ldots, k\}$ and then place v in V_i. Clearly, (i) holds almost surely. Let $A \subseteq V_i$, for some $1 \leq i \leq k$, be any set of size at most $\frac{n \log\log n \log d}{d \log n \log\log\log n}$. Using Chernoff's bound for the tail of the binomial distribution we get

$$Pr[|N_{G[V_i]}(A)| < (1 - \varepsilon)d/k \cdot |A|] \leq \exp\{-\varepsilon^2 d|A|/2k\} \leq n^{-3|A|/2}.$$

Hence, by a simple union bound argument, we obtain

$$\begin{aligned}
& Pr[\exists i \in \{1, \ldots, k\} \text{ and } \exists A \subseteq V_i \text{ such that } |N_{G[V_i]}(A)| < (1 - \varepsilon)d/k \cdot |A|] \\
\leq\ & k \cdot \sum_{a=1}^{(1+\varepsilon)n/k} \binom{(1/k + \varepsilon)n}{a} n^{-3a/2} \\
\leq\ & k \cdot \sum_{a=1}^{(1+\varepsilon)n/k} \left(\frac{3n}{ak} \cdot n^{-3/2}\right)^a \\
=\ & o(1).
\end{aligned}$$

□

Fix some $1 \leq k \leq r$. By Proposition 3.1 with $\varepsilon = 0.1$, one can obtain a partition $V = V_1 \cup \ldots \cup V_k$ such that $G[V_i]$ satisfies P1 for every $1 \leq i \leq k$. Furthermore, by definition, every $G[V_i]$ satisfies P2' as well and is therefore hamiltonian by Theorem 1.1. This yields the required 2-factor.

□

Proof of Theorem 1.6 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

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

Next, we deal with property P1. Since a.s. there are vertices of "low" degree in G, we cannot expect every "small" set to be expanded by a factor of d. Therefore, to handle this difficulty, we introduce some minor changes to the proof of Theorem 1.1. 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 $dist_G(u,v) \geq 241$.

(3) G satisfies a weak version of P1, that is, if $A \subseteq V\backslash SMALL$ and $|A| \leq \frac{n\log\log n\log d}{4130d\log n\log\log\log n}$ then $|N_G(A)| \geq 3d|A|$.

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

We will prove, that based on these properties, we can build initial long paths as in Claim 2.2 in Section 2.1 of Theorem 1.1. The proof is essentially the same; the main difference is that we change the definition of S_t for $t \geq 2$ to $S_t = \{v \in V(P_0) : v \neq v_1$, v is the endpoint of a path obtainable from P_0 by t steps, where each step consists of at most two rotations with fixed endpoint v_1 and all broken edges in $P_0\}$. That is, in order to substantially enlarge the set of endpoints in every step, we use up to two rotations per step, instead of just one. This increases the number of rotations used for building the initial path by a constant factor, and so has no real effect on the rest of the proof. The change for $t = 1$ is greater - we use up to 121 rotations in the first step. First we show how to build S_1. We rotate P_0 in order to obtain an endpoint set S_1 of size at least $(\log n)^{12}$. If this can be done, using at most 121 rotations then we are done, otherwise there must be some small vertex u in the set S_1 we were able to build. Let P_u denote the path of maximum length from v_1 to u. We start again, this time with P_u instead of P_0. By (1), u has at least one neighbor on P_u other than its predecessor. Thus we can rotate once and get a new endpoint w. By (2),

since $u \in SMALL$, we have $w \notin SMALL$ and we can rotate another 120 times, without ever reaching a small vertex. Thus, in each rotation we get at least $3d/3 = (\log n)^{0.1}$ new endpoints. Hence, after at most 121 rotations we obtain S_1 of size at least $(\log n)^{12}$.

We proceed to prove that $|S_t| \geq \max\{(\log n)^{12}, \left(\frac{d}{3}\right)^t\}$ by induction on t. Assume we have already built S_t, $(t \geq 1)$ such that $|S_t| \geq \max\{(\log n)^{12}, \left(\frac{d}{3}\right)^t\}$ and now wish to build S_{t+1}. If $|N(S_t)| \geq d|S_t|$ then we can proceed as in the proof of Claim 2.2; so assume otherwise. By (3), this must mean that $|S_t \cap SMALL| \geq 1/2|S_t|$. Let $S'_t = S_t \cap SMALL$. $|S'_t| = \omega(\log^{11} n)$ and so by (4), almost every vertex of S'_t has degree at least 12. By (3), no two small vertices have a common neighbor and so $|N(S'_t)| \geq (12 - o(1))|S'_t|$. As in the proof of Claim 2.2, we have $|T| \geq |S'_t|$. Now, by (2) we have $T \cap SMALL = \emptyset$ and so by (3) we can rotate once more to obtain $|N(T)| \geq 3d|T| \geq d|S_t|$. As in the proof of Claim 2.2, this will yield a set S_{t+1} of size at least $\left(\frac{d}{3}\right)^{t+1}$. Also, clearly $|S_{t+1}| \geq |S_t|$ which by the induction hypotheses is at least $(\log n)^{12}$. □

Proof of Theorem 1.7 Let $G = (V, E)$ be f-connected where $f(k) = 12 + k \log k$. Let $A \subseteq V$ be of size at most $\frac{n}{12m}$. Either $|A| > |V \setminus (A \cup N(A))|$ and so in particular $|N(A)| \geq 12|A|$, 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 $|N(A)| \geq f(|A|) \geq 12 + |A| \log |A| \geq 12$. 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 $P2$. By Theorem 1.1, G is hamiltonian for sufficiently large n.

□

Acknowledgements

We would like to thank Deryk Osthus for bringing to our attention the result and conjecture of [4].

References

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

[2] B. Bollobás, Random graphs, 2^{nd} ed. New York : Cambridge university press, 2001.

[3] 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.

[4] 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.

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

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

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

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

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

[10] D. Hefetz, M. Krivelevich and T. Szabó, Avoider-Enforcer games, submitted.

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

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

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