Sparse Pseudo-Random Graphs are Hamiltonian

Michael Krivelevich[1]* and Benny Sudakov[2,3]

[1]DEPARTMENT OF MATHEMATICS
RAYMOND AND BEVERLY SACKLER
FACULTY OF EXACT SCIENCES
TEL AVIV UNIVERSITY
TEL AVIV 69978, ISRAEL
E-mail: krivelev@post.tau.ac.il

[2]DEPARTMENT OF MATHEMATICS
PRINCETON UNIVERSITY
PRINCETON, NJ 08544
E-mail: bsudakov@math.princeton.edu

[3]INSTITUTE FOR ADVANCED STUDY
PRINCETON, NJ 08540
E-mail: bsudakov@math.princeton.edu

Received August 12, 2001; Revised May 17, 2002

DOI 10.1002/jgt.10065

Abstract: In this article we study Hamilton cycles in sparse pseudo-random graphs. We prove that if the second largest absolute value λ of an eigenvalue of a d-regular graph G on n vertices satisfies

Contract grant sponsor: USA-Israeli BSF (for M.K.); Contract grant sponsor: Israel Science Foundation; Contract grant sponsor: Bergmann Memorial Award (for M.K.); Contract grant sponsor: NSF; Contract grant numbers: DMS-0106589, CCR-9987845 (to B.S.); Contract grant sponsor: State of New Jersey.
*Correspondence to: Michael Krivelevich, School of Mathematical Sciences, Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
E-mail: krivelev@post.tau.ac.il

$$\lambda \leq \frac{(\log\log n)^2}{1000\,\log n\,(\log\log\log n)}\,d$$

and n is large enough, then G is Hamiltonian. We also show how our main result can be used to prove that for every $c > 0$ and large enough n a Cayley graph $X(G,S)$, formed by choosing a set S of $c\log^5 n$ random generators in a group G of order n, is almost surely Hamiltonian. J Graph Theory 42: 17–33, 2003

Keywords: *Hamiltonian cycles; pseudo-random graphs; graph eigenvalues*

1. INTRODUCTION

A *Hamilton cycle* in a graph is a cycle passing through all the vertices of this graph. A graph is called *Hamiltonian* if it has at least one Hamilton cycle. The notion of Hamilton cycles is one of the most central in modern Graph Theory, and many efforts have been devoted to obtain sufficient conditions for Hamiltonicity. One of the oldest such results is the celebrated theorem of Dirac [13], who showed that if the minimal degree of graph G on n vertices is at least $n/2$, then G contains a Hamilton cycle. This result is only one example of a vast majority of known sufficient conditions for Hamiltonicity that mainly deal with graphs which are fairly dense. On the other hand, it appears that not too much is known about Hamilton cycles in relatively sparse graphs. Here we would like to propose one such sufficient condition that works also in the sparse case. Our condition is not based on degree or density conditions, rather it has to do with what is usually called *pseudo-randomness*.

Pseudo-random graphs can be informally described as graphs whose edge distribution resembles closely that of a truly random graph $G(n,p)$ of the same edge density. Pseudo-random graphs have been a subject of intensive study during the last two decades (see, e.g., [27], [28], [9], [26], [2], [22]).

In this article, we restrict our attention to pseudo-random regular graphs. This will enable us to use the powerful and well developed machinery of Spectral Graph Theory (see [8]) to connect between the eigenvalues of a graph and its edge distribution. Some definitions are in place here. Let $G = (V,E)$ be a graph with vertex set $V = \{1,\ldots,n\}$. The adjacency matrix $\mathbf{A} = \mathbf{A}(G)$ is an n-by-n $0,1$-matrix whose entry $\mathbf{A}_{ij}$ is 1 whenever $(i,j) \in E(G)$, and is 0 otherwise. As $\mathbf{A}$ is a real symmetric matrix, all its eigenvalues are real. We thus denote the eigenvalues of $\mathbf{A}$, usually also called the eigenvalues of the graph G itself, by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. In case G is a d-regular graph, it follows from the Perron-Frobenius Theorem that $\lambda_1 = d$ and $|\lambda_i| \leq d$ for all $2 \leq i \leq n$. Let now $\lambda = \lambda(G) = \max\{|\lambda_i(G)| : i = 2,3,\ldots,n\}$. The parameter λ is usually called *the second eigenvalue of G*.

It is well known that the larger the so called spectral gap (i.e., difference between d and λ) is, the more closely the edge distribution of G approaches that of a random graph $G(n, d/n)$. We will cite relevant quantitative results later, for now we just state that the value of λ will serve us as the measure of pseudo-randomness.

The subject of this study is to show that under certain conditions pseudo-randomness ensures the existence of a Hamilton cycle. The connection between pseudo-randomness and Hamiltonicity has been already explored in several articles ([27], [15], [16]). Note that if the vertex degree $d = d(n)$ satisfies $d = n^a$, where a is a constant close to 1, then already a very weak condition on the second eigenvalue λ guarantees Hamiltonicity. This follows immediately from the well known theorem of Chvátal and Erdős [11]. They proved that if the connectivity of graph G is at least as large as the size of maximal independent set, then G contains a Hamilton cycle. The independence number of a d-regular graph G on n vertices with the second eigenvalue λ can be bounded from above by $\lambda n/d$ (see Theorem 2.1 below), and the connectivity is $d(1 - o(1))$ if $\lambda \ll d$. Plugging these two estimates into the Chvátal-Erdős theorem, one immediately gets a sufficient condition for Hamiltonicity. This approach has been used by Thomason in [27], and recently the authors together with Vu and Wormald [23] applied it to show that almost surely random regular graphs of high degree contain a Hamilton cycle.

Here we state and prove a sufficient condition for a Hamilton cycle in a pseudo-random regular graph G, which does not require its degree d to be a power of $|V(G)|$. Specifically, we prove the following theorem.

Theorem 1.1. *Let $G = (V, E)$ be a d-regular graph on n vertices and with second largest by absolute value eigenvalue λ. If n is large enough and*

$$\lambda \leq \frac{(\log\log n)^2}{1000 \log n (\log\log\log n)}\, d\,,$$

then G is Hamiltonian.

It is known[1] that $\lambda(G) = \Omega(\sqrt{d})$ for $d \leq n/2$. Therefore, for pseudo-random graphs with the best possible order of magnitude of $\lambda = \lambda(d, n)$, our result starts working already when the degree d is only polylogarithmic in n, i.e., for quite sparse pseudo-random graphs.

It should be stressed that the definition of pseudo-random graphs used in this study is rather restrictive and applies only to *regular* graphs. It seems plausible, however, that our techniques can be used to prove Hamiltonicity of almost regular graphs (i.e., graphs in which all degrees are very close to an average degree d)

[1]It is easy to see that if $d \leq n/2$, then $\lambda(G) = \Omega(\sqrt{d})$. Indeed, let $\mathbf{A}$ be the adjacency matrix of G, then the the trace of $\mathbf{A}^2$ satisfies

$$nd = 2|E(G)| = tr(\mathbf{A}^2) = \sum_i \lambda_i^2 \leq d^2 + (n-1)\lambda^2(G) \leq dn/2 + (n-1)\lambda^2(G).$$

Therefore, $\lambda^2(G) \geq d/2$ and $\lambda(G) = \Omega(\sqrt{d})$.

with a large enough eigenvalue gap. Proving Hamiltonicity in other models of sparse pseudo-random graphs, for example in the so called jumbled graphs introduced by Thomason in [27], appears to be an interesting and challenging task.

The rest of this article is organized as follows. In the next section we summarize some useful quantitative results on the edge distribution of pseudo-random regular graphs, which we use later in the proof. In Section 3, we present the proof of our main theorem. In Section 4, we indicate how our main result can be used to prove that for every $c > 0$ a Cayley graph $X(G, S)$, formed by choosing a set S of $c \log^5 n$ random generators in a group G of order n, is almost surely Hamiltonian. The last section of the article is devoted to concluding remarks and discussion of relevant open problems.

We close this section with some conventions and notation. A graph G is called (n, d, λ)-*graph* if it is d-regular, has n vertices and the second eigenvalue of G equals to λ. For a subset of vertices $U \subset V(G)$ we denote by $N(U)$ the set of all vertices in $V - U$ adjacent to some vertex in U. We also denote by $e(U)$ the number of edges spanned by U. If log has no suffix, it denotes the natural logarithm. Throughout the article, we omit occasionally the floor and ceilings signs for the sake of convenience. We will also make no serious attempt to optimize our absolute constants.

2. PROPERTIES OF PSEUDO-RANDOM GRAPHS

In this section we gather quantitative results on the edge distribution in pseudo-random regular graphs, to be used later in the proof. Essentially all of them are easy corollaries of the following well-known result whose proof can be found, inter alia, in Chapter 9 of a monograph of Alon and Spencer [4].

Theorem 2.1. *Let $G = (V, E)$ be an (n, d, λ)-graph. Then*:

- *For every two subsets $B, C \subset V$ the number of edges of G with one endpoint in B and the other in C satisfies*:

$$\left| e(B, C) - \frac{|B||C|d}{n} \right| \leq \lambda \sqrt{|B||C|\left(1 - \frac{|B|}{n}\right)} \; ; \tag{1}$$

- *For every subset $B \subseteq V$,*

$$\left| e(B) - \frac{|B|^2 d}{2n} \right| < \frac{\lambda |B|}{2}. \tag{2}$$

In the propositions below we will assume that $G = (V, E)$ is an (n, d, λ)-graph with $\lambda < d/2$.

Proposition 2.1. *Every subset $V_0 \subset V$ of cardinality $|V_0| \leq \frac{\lambda n}{d}$ spans at most $\lambda |V_0|$ edges.*

Proof. By (2),

$$e(V_0) \leq \frac{|V_0|^2 d}{2n} + \frac{\lambda |V_0|}{2} \leq \frac{\lambda n}{d}\frac{d|V_0|}{2n} + \frac{\lambda |V_0|}{2}$$
$$= \frac{\lambda |V_0|}{2} + \frac{\lambda |V_0|}{2} = \lambda |V_0|.$$

Proposition 2.2. *For every subset $V_0 \subset V$ of cardinality $|V_0| \leq \frac{\lambda^2 n}{d^2}$,*

$$|N(V_0)| > \frac{(d-2\lambda)^2}{3\lambda^2}|V_0|.$$

Proof. Denote $N(V_0) = U$. Then by Proposition 2.1, $e(V_0) \leq \lambda|V_0|$. As the degree of every vertex in V_0 is d, we obtain:

$$e(V_0, U) \geq d|V_0| - 2e(V_0) \geq d|V_0| - 2\lambda|V_0|.$$

On the other hand, it follows from (1) that

$$e(V_0, U) < \frac{|V_0||U|d}{n} + \lambda\sqrt{|V_0||U|}.$$

The above inequalities imply:

$$\frac{|V_0||U|d}{n} + \lambda\sqrt{|V_0||U|} > (d-2\lambda)|V_0|. \tag{3}$$

If $|U| \leq \frac{(d-2\lambda)^2}{3\lambda^2}|V_0|$, then

$$\frac{|V_0||U|d}{n} + \lambda\sqrt{|V_0||U|} \leq \frac{|V_0|^2(d-2\lambda)^2 d}{3\lambda^2 n} + \frac{\lambda|V_0|(d-2\lambda)}{\sqrt{3}\lambda}$$
$$\leq \frac{\lambda^2 n}{d^2}\cdot\frac{(d-2\lambda)^2 d|V_0|}{3\lambda^2 n} + \frac{(d-2\lambda)|V_0|}{\sqrt{3}}$$
$$= \frac{(d-2\lambda)^2|V_0|}{3d} + \frac{(d-2\lambda)|V_0|}{\sqrt{3}}$$
$$< \frac{(d-2\lambda)|V_0|}{3} + \frac{(d-2\lambda)|V_0|}{\sqrt{3}} < (d-2\lambda)|V_0|$$

— a contradiction to (3).

Proposition 2.3. *For every subset $V_0 \subseteq V$ of cardinality $|V_0| > \frac{\lambda^2 n}{d^2}$, $|N(V_0)| > \frac{n}{2} - |V_0|$.*

Proof. Set $U = V \backslash (V_0 \cup N(V_0))$. Then, clearly $e(V_0, U) = 0$. On the other hand, by (1), $e(V_0, U) \geq \frac{|V_0||U|d}{n} - \lambda\sqrt{|V_0||U|(1 - \frac{|U|}{n})}$. Therefore, $\frac{|V_0||U|d}{n} \leq \lambda \sqrt{|V_0||U|(1 - \frac{|U|}{n})}$, implying:

$$\frac{|U|}{\left(1 - \frac{|U|}{n}\right)} \leq \frac{\lambda^2 n^2}{d^2 |V_0|} < n\,,$$

and thus $|U| < n/2$. Hence $|N(V_0)| = |V| - |V_0| - |U| > n/2 - |V_0|$.

Proposition 2.4. *If disjoint subsets $U_1, U_2 \subset V(G)$ are not connected by an edge in G, then $|U_1|\,|U_2| < \lambda^2 n^2/d^2$.*

Proof. By (1), $0 = e(U_1, U_2) > |U_1||U_2|d/n - \lambda\sqrt{|U_1||U_2|}$. Therefore, $|U_1||U_2|d/n < \lambda\sqrt{|U_1||U_2|}$, and the claim follows.

Proposition 2.5. *G is connected.*

Proof. If G is disconnected, then G has a connected component V_0 of size $|V_0| \leq n/2$. As $N(V_0) = \emptyset$, it follows from Proposition 2.3 that $|V_0| \leq \lambda^2 n/d^2$. This contradicts Proposition 2.2 as $|N(V_0)| \geq \frac{(d-2\lambda)^2 |V_0|}{3\lambda^2} > 0$.

3. PROOF OF THE MAIN THEOREM

We first show that a longest path in an (n, d, λ)-graph G, satisfying the conditions of Theorem 1, has a length linear in n. Then we show that some path of a maximal length in G can be closed to a cycle, which easily implies the Hamiltonicity of G due to its connectivity, provided by Proposition 2.5. Our approach relies on the so called rotation-extension technique, invented by Posa in [25] and applied in several subsequent articles on Hamiltonicity of random graphs [21], [6], [16]. Our notation and proof methodology are quite similar to those of [16].

A. Constructing an Initial Long Path

Let $P_0 = (v_1, v_2, \ldots, v_l)$ be a longest path in G. If $1 \leq i < l$ and $(v_i, v_l) \in E(G)$, then the path $P' = (v_1, v_2, \ldots, v_i, v_l, v_{l-1}, \ldots, v_{i+1})$ is also of maximal length. We say that P' is a *rotation* of P with *fixed endpoint* v_1, *pivot* v_i and *broken edge* (v_i, v_{i+1}) (the reason for the last term being the fact that (v_i, v_{i+1}) is deleted from the edge set of P to get P'). We can then rotate P' in a similar fashion to get a new path P'' of the same length, and so on.

For $t \geq 0$, let $S_t = \{v \in V(P_0) \backslash \{v_1\}$: v is the endpoint of a path obtainable from P_0 by at most t rotations with fixed endpoint v_1, and all broken edges in $P_0\}$. Obviously, the sequence of sets $\{S_t : t \geq 0\}$ forms a family of nested sets, as $S_t \subseteq S_{t+1}$ for all $t \geq 0$. Notice that due to the maximality of P_0 all edges incident to a vertex from S_t for some $t \geq 0$, have their second endpoint inside P_0.

Proposition 3.1. *For all* $t \geq 0$,

$$|S_{t+1}| \geq \frac{1}{2}|N(S_t)| - \frac{3}{2}|S_t|.$$

Proof. Let

$$T = \{i \geq 2 \ : \ v_i \in N(S_t), v_{i-1}, v_i, v_{i+1} \notin S_t\}.$$

Obviously, $|T| \geq |N(S_t)| - 3|S_t|$. Consider a vertex $v_i \in V(P_0)$ with $i \in T$. Then, v_i has a neighbor $x \in S_t$. This means that there exists a path Q with x as an endpoint, obtainable from P_0 by at most t rotations. Observe that if during the repeated rotation process an edge from P_0 gets deleted, then one of its vertices should be an endpoint of some rotation of P_0. As $v_{i-1}, v_i, v_{i+1} \notin S_t$, both edges (v_{i-1}, v_i) and (v_i, v_{i+1}) are still present in Q. Rotating Q with a pivot v_i and one of the edges (v_{i-1}, v_i), (v_i, v_{i+1}) as a broken edge (which of these two edges is chosen depends on their order along Q) will put one of v_{i-1}, v_{i+1} in S_{t+1}. Suppose for example it is v_{i-1}. The only other vertex which can cause v_{i-1} to be put into S_{t+1} is v_{i-2}. Therefore,

$$|S_{t+1}| \geq \frac{1}{2}|T| \geq \frac{1}{2}|N(S_t)| - \frac{3}{2}|S_t|. \qquad \blacksquare$$

Let

$$t_0 = \left\lceil \frac{\log n - 2\log(d/\lambda)}{2\log(d/\lambda) - 7} \right\rceil + 2;$$
$$\rho = 2t_0.$$

By Proposition 2.2, as long as $|S_t| \leq \lambda^2 n/d^2$ we get, $|N(S_t)| \geq (d-2\lambda)^2|S_t|/(3\lambda^2)$ and thus $|S_{t+1}| \geq (d-2\lambda)^2|S_t|/(6\lambda^2) - \frac{3}{2}|S_t|$. It can easily be proven by induction that in this case $|S_{j+1}|/|S_j| \geq (d-2\lambda)^2/(7\lambda^2)$ for all $j \leq t$. This implies that after at most

$$\frac{\log \dfrac{\lambda^2 n}{d^2}}{\log \dfrac{(d-2\lambda)^2}{7\lambda^2}} \leq t_0 - 2$$

steps we get $|S_t| > \lambda^2 n/d^2$. One additional step together with application of Proposition 2.3 give us:

$$|S_{t+1}| \geq \frac{1}{2}|N(S_t)| - \frac{3}{2}|S_t| \geq \frac{1}{2}\left(\frac{n}{2} - |S_t|\right) - \frac{3}{2}|S_t|$$
$$= \frac{n}{4} - 2|S_t| \geq \frac{n}{4} - 2|S_{t+1}|.$$

Hence $|S_{t+1}| \geq \frac{n}{12}$. This, by Proposition 2.4, implies in turn that $|N(S_{t+1})| = n - o(n)$. Applying Proposition 2.3 once again, we obtain:

$$|S_{t+2}| \geq \frac{1}{2}|N(S_{t+1})| - \frac{3}{2}|S_{t+1}| \geq \frac{1}{2}(n - o(n)) - \frac{3}{2}|S_{t+2}|,$$

and therefore $|S_{t+2}| \geq (1 - o(1))n/5 > n/6$.

Let $B(v_1) = S_{t_0}$, $A_0 = B(v_1) \cup \{v_1\}$. For each $v \in B(v_1)$ we can use the above argument to show the existence of a set $B(v)$, $|B(v)| \geq n/6$, of endpoints of maximum length paths with endpoint v. Note that each endpoint in $B(v)$ was obtained by at most $t_0 + t_0 = 2t_0$ rotations. As clearly $B(v) \subseteq V(P_0)$ for each such v, we get in particular that $l \geq n/6$, and thus P_0 has a linear length.

To summarize, for each $a \in A_0$, $b \in B(a)$ there is a maximum length path $P(a, b)$ joining a and b and obtainable from P_0 by at most $\rho = 2t_0$ rotations.

B. Closing a Maximal Path to a Hamilton Cycle

We consider the path P_0 to be directed and divided into 2ρ disjoint segments $I_1, \ldots, I_{2\rho}$, all of length at least $\lfloor |V(P_0)|(2\rho) \rfloor \geq \lfloor n/(12\rho) \rfloor$. Notice that each path $P(a, b)$ as above is obtained from P_0 by at most ρ rotations and therefore contains at least ρ of the segments I_i untouched (but possibly traversed in the opposite direction). We call each such segment *unbroken* in $P(a, b)$. These segments have an absolute orientation induced by P_0, and another, relative to this by $P(a, b)$ which we consider directed from a to b.

Let

$$k = 2\max\left\{1, \left\lceil \frac{400\lambda\rho}{d} \right\rceil\right\}.$$

We consider sequences $\sigma = I_{i_1}, \ldots, I_{i_k}$ of k unbroken segments of P_0 which occur in this order in $P(a, b)$, where σ specifies not only the order of segments in $P(a, b)$ but also their relative orientation. We say then that $P(a, b)$ *contains* σ. Note that as $P(a, b)$ has at least ρ unbroken segments I_{i_j}, $P(a, b)$ contains at least $\binom{\rho}{k}$ sequences σ.

For a given σ we denote by $L(\sigma)$ the set of all pairs $a \in A_0, b \in B(a)$, for which the path $P(a, b)$ contains σ.

The total number of possible sequences σ is at most $(2\rho)_k 2^k$. Therefore, by averaging we obtain that there exists a sequence σ_0 for which

$$|L(\sigma_0)| \geq \frac{n^2}{64} \frac{\binom{\rho}{k}}{(2\rho)_k 2^k} > \frac{n^2}{64} \left(\frac{\rho - k}{2\rho - k}\right)^k \frac{1}{k!2^k}.$$

It is easy to check that $k \leq \rho/2$. Then, $(\rho - k)/(2\rho - k) \geq 1/3$, and it follows that there exists a sequence σ_0 for which $|L(\sigma_0)| \geq n^2/(64k!6^k)$. Fix such a sequence.

Denote

$$\alpha = \frac{1}{64k!6^k}.$$

Let $\hat{A} = \{a \in A_0 : L(\sigma_0)$ contains at least $\alpha n/2$ pairs with a as the first element$\}$. Then $|\hat{A}| \geq \alpha n/2$. For each $a \in \hat{A}$ let $\hat{B}(a) = \{b \in B(a) : (a, b) \in L(\sigma_0)\}$. The definition of $\hat{A}$ guarantees that $|\hat{B}(a)| \geq \alpha n/2$.

Let C_1 be the union of the first $k/2$ segments of σ_0, in the fixed order and with the fixed relative orientation in which they occur along *any* of the paths $P(a, b)$, $(a, b) \in L(\sigma_0)$. Let C_2 be the union of the last $k/2$ segments of σ_0. Notice that

$$|C_i| \geq \frac{k}{2} \left\lfloor \frac{n}{12\rho} \right\rfloor \geq \frac{400\lambda\rho}{d} \left\lfloor \frac{n}{12\rho} \right\rfloor > \frac{32\lambda n}{d}. \tag{4}$$

Given a path P_0 and a set $S \subseteq V(P_0)$, a vertex $v \in S$ is called an *interior point* of S with respect to P_0 if both neighbors of v along P_0 lie in S. The set of all interior points of S will be denoted by $int(S)$.

Proposition 3.2. *The set C_1 contains a subset C_1' with $|int(C_1')| \geq nk/(48\rho)$ so that every vertex $v \in C_1'$ has at least 14λ neighbors in $int(C_1')$.*

Proof. We start with $C_1' = C_1$ and as long as there exists a vertex $v_j \in C_1'$ for which $d_{int(C_1')}(v_j) < 14\lambda$, we delete v_j and repeat. If this procedure continued for $r = |C_1|/7$ steps then we get a subset $R = \{v_1, \ldots, v_r\}$ so that $|int(C_1')| \geq |int(C_1)| - 3r = (1 - o(1))|C_1| - 3r > |C_1|/2$ and $e(R, int(C_1')) \leq 14\lambda r = 14\lambda|C_1|/7$. But according to (1) and (4),

$$\begin{aligned} e(R, int(C_1')) &\geq \frac{|R||int(C_1')|d}{n} - \lambda\sqrt{|R||int(C_1')|} \\ &\geq \frac{|C_1|}{7}\frac{|C_1|}{2}\frac{d}{n} - \lambda\sqrt{\frac{|C_1|}{7}\frac{|C_1|}{2}} = \frac{|C_1|^2 d}{14n} - \frac{\lambda|C_1|}{\sqrt{14}} \\ &\geq \frac{32\lambda n}{d}\frac{|C_1|d}{14n} - \frac{\lambda|C_1|}{\sqrt{14}} = \lambda|C_1|\left(\frac{16}{7} - \frac{1}{\sqrt{14}}\right) \\ &> \frac{14\lambda|C_1|}{7} \end{aligned}$$

— a contradiction. ■

Obviously, an analogous statement holds for C_2 as well. We fix the obtained sets C_1' and C_2'.

Proposition 3.3. *There is a vertex $\hat{a} \in \hat{A}$ connected by an edge to $int(C_1')$.*

Proof. Recall that $|\hat{A}| \geq \alpha n/2$ and $|int(C_1')| \geq nk/(48\rho)$. Therefore, by Proposition 2.4, the claim will follow if we will show that $(\alpha n)(nk/\rho) \gg \lambda^2 n^2/d^2$, or (substituting the value of α) $d^2/(\lambda^2\rho) \gg (k-1)!6^k$.

Consider first the case when $400\lambda\rho/d \geq 1$. In this case,

$$\begin{aligned} k &= 2\frac{(1+o(1))400\lambda\rho}{d} = \frac{800\lambda}{d}\frac{(1+o(1)\log n}{\log(d/\lambda)} \\ &\leq \frac{(1+o(1))800\log n}{\frac{1000\log n(\log\log\log n)}{(\log\log n)^2}\cdot\log\log n} \\ &= \frac{(1+o(1))0.8\log\log n}{\log\log\log n}, \end{aligned}$$

and thus $(k-1)!6^k < (\log n)^{0.9}$. On the other hand,

$$\begin{aligned} \frac{d^2}{\lambda^2\rho} &\geq \frac{d^2}{\lambda^2}\frac{\log(d/\lambda)}{(1+o(1))\log n} \\ &\geq \frac{\log^2 n\,(\log\log\log n)^2}{(\log\log n)^4}\frac{\log\log n}{(1+o(1))\log n} > \frac{(1+o(1))\log n}{(\log\log n)^3} \\ &\gg (\log n)^{0.9}, \end{aligned}$$

as required.

In the second case, when $400\lambda\rho/d \leq 1$, we get $k = 2$, and then the expression $(k-1)!6^k$ is an absolute constant, while $d^2/(\lambda^2\rho) = (d/\lambda)(d/(\lambda\rho)) \geq 400d/\lambda \to \infty$. The proposition follows. ∎

Repeating the same argument, mutatis mutandis, gives:

Proposition 3.4. *There exists a vertex $\hat{b} \in \hat{B}(\hat{a})$ connected by an edge to $int(C_2')$.*

Let now x be a vertex separating C_1' and C_2' along $P(\hat{a}, \hat{b})$. We consider two half paths P_1 and P_2 obtained by splitting $P(\hat{a}, \hat{b})$ at x. Our idea is as follows: rotating each P_i while keeping x as a fixed point and using vertices in $int(C_i')$ as pivots, we wish to achieve the situation where the corresponding endpoint sets V_1, V_2 are large enough. Then, Proposition 2.4 will show that there is an edge between V_1 and V_2. This edge closes a path of maximal length to a cycle. As G is connected by Proposition 2.5, any non-Hamilton cycle can be extended to a path covering some additional vertices. Therefore, the assumption about the maximality of P_0 implies that P_0 is a Hamilton path, and thus the above created cycle is Hamilton as well.

Consider P_1. Let $T_i = \{v \in C_1' \backslash \{x\} : v$ is the endpoint of a path obtainable from P_1 by i rotations with fixed endpoint x, all pivots in $int(C_1')$ and all broken edges in $P_1\}$.

Proposition 3.5. *There exists an i for which* $|T_i| \geq \lambda n/d$.

Proof. It is enough to prove that there exists a sequence of sets $U_i \subseteq T_i$ such that $|U_i| = 1$ and $|U_{i+1}| = 2|U_i|$, as long as $|U_i| < \lambda n/d$. Note that according to Proposition 3.4 $\hat{a}$ has a neighbor in $int(C_1')$, and therefore $T_1 \neq \emptyset$. Note also that if we perform a rotation at a vertex from $int(C_1')$ and a broken edge in P_1, then the resulting endpoint is in C_1'.

Suppose we have found sets $U_1, \ldots, U_i$ as stated above, and still $|U_i| < \lambda n/d$. Similarly to the proof of Proposition 3.1 one has that

$$|T_{i+1}| \geq \frac{1}{2}|N(U_i) \cap int(C_1')| - \frac{3}{2}\sum_{j=1}^{i}|U_j|.$$

As $\sum_{j=1}^{i}|U_j| < 2|U_i|$, the claim will follow if we will prove that $|N(U_i) \cap int(C_1')| \geq 10|U_i|$. Since $U_i \subset C_1'$, every vertex $u \in U_i$ has at least 14λ neighbors in $int(C_1')$. Therefore, the number of edges with one endpoint in U_i and another one in $int(C_1')$ is at least $14\lambda|U_i|$. Set $W_i = N(U_i) \cap int(C_1')$. If $|W_i| < 10|U_i|$, then by (1) one has:

$$\begin{aligned} e(U_i, W_i) &\leq \frac{|U_i||W_i|d}{n} + \lambda\sqrt{|U_i||W_i|} = \frac{10|U_i|^2 d}{n} + \sqrt{10}\lambda|U_i| \\ &< \frac{\lambda n}{d}\,\frac{10|U_i|d}{n} + \sqrt{10}\lambda|U_i| < 14\lambda|U_i| \end{aligned}$$

— a contradiction. Therefore $|W_i| \geq 10|U_i|$, as required. ∎

Hence, the set V_1 of endpoints of all rotations of P_1 has cardinality $|V_1| \geq \lambda n/d$. As by Proposition 3.5, $\hat{b}$ has a neighbor in $int(C_2')$, the same argument can be carried out for P_2 to show that the set V_2 of endpoints of its rotations has at least $\lambda n/d$ vertices as well. Then, by Proposition 2.4 there is an edge connecting V_1 and V_2 and thus closing a Hamilton cycle. This completes the proof of Theorem 1.1. ∎

4. ON HAMILTONICITY OF RANDOM CAYLEY GRAPHS

The (undirected) Cayley graph $X(G,S)$ of a group G with respect to a set S of elements in the group (generators) is a graph with vertex set G, whose edge set is the set of all unordered pairs $\{\{g, gs\} : s \in S\}$. This is obviously a regular graph of degree $|S \cup S^{-1}| \leq 2|S|$.

The question of Hamiltonicity of Cayley graphs has drawn quite an amount of attention of many researchers over the years. It is enough to mention that a survey article by Curran and Gallian [12] on Hamiltonicity of Cayley graphs has eighty nine references. Much of the focus of the research has been centered around the following conjecture.

Conjecture 4.1. *Every connected Cayley graph with more than* 2 *vertices is Hamiltonian.*

So far only special cases of the above conjecture have been proven, the most important of them undoubtedly being the case when the group G is Abelian (see, e.g., [24], Chapter 12, Problem 17 for a proof).

Given the apparent difficulty in proving this conjecture, it is quite natural to prove it for the case where the set of generators S is chosen *at random* according to some probability distribution. A result of this type has been obtained by Meng and Huang [20], who proved that almost all Cayley graphs of a group G are Hamiltonian (the asymptotic parameter here is the order $n = |G|$ of the group G). The proof proceeds by showing that almost all Cayley graphs of the group G with n elements are two-connected and d-regular for $d \geq n/3$, and then by invoking a result of Jackson [19], according to which the last two conditions are sufficient to guarantee Hamiltonicity. As the last sentence indicates, a random set S of generators almost surely has size linear in n, thus resulting in quite dense graphs.

As it turns out, the main result of this article can be used to show that the Cayley graph $X(G, S)$ is almost surely Hamiltonian for a random set S of generators of much smaller size. To do so, we first apply an approach of Alon and Roichman [3] to bound the eigenvalue gap of such a graph, and then invoke our Theorem 1.1 to prove Hamiltonicity. Here is a result and an outline of its proof.

Theorem 4.1. *Let G be a group of order n. Then for every $c > 0$ and large enough n a Cayley graph $X(G, S)$, formed by choosing a set S of $c \log^5 n$ random generators in G, is almost surely Hamiltonian.*

Proof. Let λ be the second largest by absolute value eigenvalue of $X(G, S)$. Note that the Cayley graph $X(G, S)$ is d regular for $d \geq c \log^5 n$. Therefore, to prove Hamiltonicity of $X(G, S)$, by Theorem 1.1, it is enough to show that almost surely $\lambda/d \leq O(\log n)$. We will briefly sketch how this can be done using an approach of Alon and Roichman [3], referring an interested reader to their article for more details.

It is enough to show that the expected value $\mathbf{E}(\lambda/d)$ is bounded by $O(\log^{-1} n)$. Then one can finish the proof by considering an appropriate martingale, which shows that λ is concentrated around its expectation as it is done in [3]. Let $\mathbf{A}$ be the adjacency matrix of $X(G, S)$ and let $\mathbf{B} = \frac{1}{d}\mathbf{A}$. Then, it is easy to see that for every natural number m we have

$$\lambda \leq (Tr(\mathbf{A}^{2m}) - d^{2m})^{1/2m}.$$

This, by Jensen's inequality, implies:

$$\mathbf{E}\big(\lambda/d\big) \leq (\mathbf{E}(Tr(B^{2m})) - 1)^{1/2m}.$$

Denote by P_{2m} the probability of a walk of length $2m$ in our Cayley graph to be closed. Since a Cayley graph is vertex transitive $\mathbf{E}(Tr(B^{2m})) = n\mathbf{E}(P_{2m})$, and hence:

$$\mathbf{E}\big(\lambda/d\big) \leq (n\mathbf{E}(P_{2m}) - 1)^{1/2m}. \tag{5}$$

Next we need the following lemma.

Lemma 4.1.

$$\mathbf{E}(P_{2m}) \leq 2^{2m}\big(2m/c\log^5 n\big)^m + 1/n + O\big(m/n^2\big).$$

Proof. As in [7] and [3] we consider a dynamic process for choosing a random set S and a random walk on $X(G,S)$. This is done as follows.

(a) We choose in the free group $F_{c\log^5 n}$ (generated by $c\log^5 n$ distinct letters and their inverses) a random word of length $2m$.
(b) We assign to each letter an element of the group G at random.

It is easy to see that this process is equivalent to the one in which a random Cayley graph $X(G,S)$ with $|S| = c\log^5 n$ is chosen first and a random walk of length $2m$ in it is chosen afterwards.

In order to obtain an upper bound for $\mathbf{E}(P_{2m})$, we estimate the probabilities of the following two events whose union includes the event that our walk of length $2m$ is closed.

(A) There is no letter such that the total number of appearances of this letter together with its inverse in this word is exactly one.
(B) (A) does not hold, but after the assignment of the chosen elements in the group G to the corresponding letters the word is reduced to the unity.

Obviously $\mathbf{E}(P_{2m}) \leq \mathbf{Pr}(\mathbf{A}) + \mathbf{Pr}(\mathbf{B})$.

First we estimate $\mathbf{Pr}(\mathbf{A})$. Let W be the word of length $2m$ which satisfies the conditions of $(\mathbf{A})$. Clearly, the number of distinct symbols (letter and its inverse are the same symbol) that appear in W is at most m. We expose the letters of W in the following order. First, we expose the subset consisting of the first occurrence of each symbol that appears in the word. Second, we expose the other letters. For each letter in the second subset, the probability that it equals to the letter or the inverse of the letter which has appeared in the first subset is at most $2m/c\log^5 n$. The number of possibilities to place the first subset is at most 2^{2m}. Hence $\mathbf{Pr}\ (\mathbf{A}) \leq 2^{2m}\big(2m/c\log^5 n\big)^m$.

Next we bound $\mathbf{Pr}(\mathbf{B})$. Let τ be a symbol (a letter or its inverse) that appears only once in the word. We expose the assignments of all the letters except that of τ. Denote by $x(\tau)$ the assignment of τ. The event whose probability we wish to estimate is now the event $gx(\tau)h = 1$ where g, h are some known elements in G. The probability that $x(\tau)$ solves this equation is at most $1/(n-2m) = 1/n + O(m/n^2)$. Therefore, $\mathbf{Pr}(\mathbf{B}) \le 1/n + O(m/n^2)$. This implies the assertion of the lemma. ■

To finish the proof of the Theorem 4.1, fix now $m = \frac{\log n}{2\log\log n}$. Combining Lemma 4.1 with (5) we obtain,

$$\begin{aligned}\mathbf{E}(\lambda/d) &\le (n\mathbf{E}(P_{2m}) - 1)^{1/2m} \le \left(n\left(2^{2m}\left(2m/c\log^5 n\right)^m + 1/n + O\left(m/n^2\right)\right) - 1\right)^{1/2m}\\ &\le n^{1/2m}\left(2^{2m}\left(2m/c\log^5 n\right)^m\right)^{1/2m} + \left(n\left(1/n + O\left(m/n^2\right)\right) - 1\right)^{1/2m}\\ &\le 2n^{1/2m}\left(\frac{2m}{c\log^5 n}\right)^{1/2} + O\left((m/n)^{1/2m}\right) = O\left(\log^{-1} n\right).\end{aligned}$$

This completes the proof of the theorem. ■

Remark. The bound of Theorem 4.1 is probably far from being tight. It is well known that the Cayley graph $X(G,S)$ of a group G of order n with respect to a random set S of $O(\log n)$ generators is almost surely connected. Therefore, Conjecture 4.1 suggests that already $O(\log n)$ random generators are enough to guarantee Hamiltonicity of $X(G,S)$.

5. CONCLUDING REMARKS

In this study we provide a sufficient condition for the existence of a Hamilton cycle in pseudo-random graphs. A distinctive feature of our result, Theorem 1.1, is that it connects between spectral properties of a graph and Hamiltonicity. This connection has a potential to be useful in proving Hamiltonicity of certain classes of regular graphs as sometimes is much easier to bound the eigenvalue gap than to show the existence of a Hamilton cycle directly.

In particular, our result can be used to give another proof that for $d > \log^2 n$, the random d-regular graph on n vertices $G_{n,d}$ is almost surely Hamiltonian. Indeed, by Theorem 1.1, it is enough to show that the second eigenvalue of $G_{n,d}$ satisfies $\lambda < d(\log\log n)/\log n$. For $d \ll n^{1/2}$ this can be done directly in the so called configuration model, using an approach of Kahn and Szemerédi from [14]. For $d \gg n^{1/2}$, this follows from the recent results of the authors together with Vu and Wormald [23]. In the intermediate range $d \sim \sqrt{n}$, one can for example prove first that almost surely the number of copies of C_4 and C_6 in $G_{n,d}$ is asymptotically equal to the expected number of copies of these two cycles in the binomial random graph $G(n, d/n)$, and then derive the required upper bound on λ by considering the trace of $\mathbf{A}^6$, where $\mathbf{A}$ is the adjacency matrix of $G_{n,d}$.

Another feature of our argument is that it is in fact *algorithmic*. Indeed, it is easy to see that our argument works even if in Section 3.1 instead of a longest path in G we will consider any path P_0 which is maximum by inclusion. Such a path can be found efficiently by the greedy algorithm. Therefore, our proof provides a polynomial time algorithm for finding a Hamilton cycle in a graph satisfying the conditions of Theorem 1.1. This may be quite valuable especially taking into account the notorious difficulty of the Hamilton cycle problem, both theoretically and practically (see, e.g., [18]). Notice also that this observation combined with the remark above about the second eigenvalue of a random d-regular graph $G_{n,d}$, provides an algorithm for finding almost surely a Hamilton cycle in $G_{n,d}$ for $d > \log^2 n$. This solves a problem posed by Frieze and McDiarmid ([17], Research Problem 3) for this range of degrees d.

Our bound is not known to be tight. In fact, we suspect that a much stronger result should be true. We propose the following conjecture.

Conjecture 5.1. *There exists a positive constant C such that for large enough n, any (n, d, λ)-graph that satisfies $d/\lambda > C$ contains a Hamilton cycle.*

This conjecture is closely related to another well known problem. The *toughness* $t(G)$ of a graph G is the largest real t so that for every positive integer $x \geq 2$ one should delete at least tx vertices from G in order to get an induced subgraph of it with at least x connected components. G is t-tough if $t(G) \geq t$. This parameter was introduced by Chvátal in [10], where he observed that Hamiltonian graphs are 1-tough and conjectured that t-tough graphs are Hamiltonian for large enough t. He even suggested that $t = 2$ should be enough, but this was recently refuted in [5]. On the other hand, one can show (see, e.g., [1]) that if G is an (n, d, λ)-graph, then the toughness of G satisfies $t(G) > \Omega(d/\lambda)$. Therefore, the conjecture of Chvátal implies our conjecture.

ACKNOWLEDGMENT

We would like to thank two anonymous referees for helpful comments and N. Alon and I. Pak for stimulating discussions on random Cayley graphs.

REFERENCES

[1] N. Alon, Tough Ramsey graphs without short cycles, J Algebraic Combinatorics 4 (1995), 189–195.

[2] N. Alon, M. Krivelevich, and B. Sudakov, List coloring of random and pseudo-random graphs, Combinatorica 19 (1999), 453–472.

[3] N. Alon and Y. Roichman, Random Cayley graphs and expanders, Random Structures and Algorithms 5 (1994), 271–284.

[4] N. Alon and J. H. Spencer, The probabilistic method, 2nd Ed., Wiley, New York, 2000.

[5] D. Bauer, H. J. Broersma, and H. J. Veldman, Not every 2-tough graph is Hamiltonian, Discrete Appl Math 99 (2000), 317–321.

[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] A. Broder and E. Shamir, On the second eigenvalue of random regular graphs, Proc 28th IEEE FOCS, IEEE (1987), 286–294.

[8] F. R. K. Chung, Spectral graph theory, Regional Conference Series in Mathematics 92, Amer Math Soc Providence, 1997.

[9] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, Combinatorica 9 (1989), 345–362.

[10] V. Chvátal, Tough graphs and Hamiltonian circuits, Discrete Math 5 (1973), 215–218.

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

[12] S. J. Curran and J. A. Gallian, Hamiltonian cycles and paths in Cayley graphs and digraphs—a survey, Discrete Math 156 (1996), 1–18.

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

[14] J. Friedman, J. Kahn, and E. Szemerédi, On the second eigenvalue in random regular graphs, Proc of 21th ACM STOC (1989), 587–598.

[15] A. M. Frieze, On the number of perfect matchings and Hamilton cycles in ε-regular non-bipartite graphs, Electronic J Combinatorics Vol. 7 (2000), publ. R57.

[16] A. M. Frieze and M. Krivelevich, Hamilton cycles in random subgraphs of pseudo-random graphs, Discrete Mathematics, in press.

[17] A. Frieze and C. McDiarmid, Algorithmic theory of random graphs, Random Structures and Algorithms 10 (1997), 5–42.

[18] M. Garey and D. Johnson, Computers and intractability: A guide to the theory of NP-completeness, W. H. Freeman and Co., San Francisco, 1979.

[19] B. Jackson, Hamiltonian cycles in regular 2-connected graphs, J Combin Theory Ser B 29 (1980), 27–46.

[20] J. Meng and Q. Huang, Almost all Cayley graphs are Hamiltonian, Acta Math Sinica N. S. 12 (1996), 151–155.

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

[22] M. Krivelevich, B. Sudakov, and T. Szabó, Triangle factors in pseudo-random graphs, submitted.

[23] M. Krivelevich, B. Sudakov, V. Vu, and N. Wormald, Random regular graphs of high degree, Random Structures and Algorithms 18 (2001), 346–363.

[24] L. Lovász, Combinatorial problems and exercises, North Holland, Amsterdam, 1979.

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

[26] M. Simonovits and V. T. Sós, Hereditary extended properties, quasi-random graphs and not necessarily induced subgraphs, Combinatorica 17 (1997), 577–596.

[27] A. Thomason, Pseudorandom graphs, In: Random graphs'85 (Poznań 85), North-Holland Math Stud 144, North Holland, Amsterdam, 1987, pp. 307–331.

[28] A. Thomason, Random graphs, strongly regular graphs and pseudorandom graphs, In: Surveys in Combinatorics 1987, London Math Soc Lecture Note Ser 123, Cambridge Univ Press, Cambridge, 1987, pp. 173–195.