ON COVERING EXPANDER GRAPHS BY HAMILTON CYCLES

ROMAN GLEBOV, MICHAEL KRIVELEVICH, AND TIBOR SZABÓ

ABSTRACT. The problem of packing Hamilton cycles in random and pseudorandom graphs has been studied extensively. In this paper, we look at the dual question of covering all edges of a graph by Hamilton cycles and prove that if a graph with maximum degree Δ satisfies some basic expansion properties and contains a family of $(1-o(1))\Delta/2$ edge disjoint Hamilton cycles, then there also exists a covering of its edges by $(1+o(1))\Delta/2$ Hamilton cycles. This implies that for every $\alpha > 0$ and every $p \geq n^{\alpha-1}$ there exists a covering of all edges of $G(n, p)$ by $(1 + o(1))np/2$ Hamilton cycles asymptotically almost surely, which is nearly optimal.

1. INTRODUCTION

For an r-uniform hypergraph G and a family $\mathcal{F}$ of its subgraphs, we call a family $\mathcal{F}' \subset \mathcal{F}$ an $\mathcal{F}$-*decomposition* of G if every edge of G is contained in *exactly* one of the hypergraphs from $\mathcal{F}'$. We call a family $\mathcal{F}' \subset \mathcal{F}$ an $\mathcal{F}$-*packing* of G, if every edge of G is contained in *at most* one of the hypergraphs from $\mathcal{F}'$. Naturally, one tries to *maximize* the size of an $\mathcal{F}$-packing of G. The dual concept is that of an $\mathcal{F}$-*covering*: a family $\mathcal{F}'$ is called an $\mathcal{F}$-*covering* of G, if every edge of G is contained in *at least* one of the hypergraphs from $\mathcal{F}'$. Here the *minimum* size of an $\mathcal{F}$-covering of G is sought.

Decompositions, packings and coverings are in the core of combinatorial research (see [12] for a survey). One of the most famous problems in this area was the conjecture of Erdős and Hanani [8], dealing with the case when G is the complete r-uniform hypergraph on n vertices and $\mathcal{F}$ is the family of all k-cliques in G for some $k \geq r$. Clearly, if there was an $\mathcal{F}$-decomposition of G, its size would be $\binom{n}{r}/\binom{k}{r}$. Hence $\binom{n}{r}/\binom{k}{r}$ is an upper bound on the size of a largest $\mathcal{F}$-packing of G and a lower bound on the size of a smallest $\mathcal{F}$-covering of G. Erdős and Hanani conjectured both inequalities to be asymptotically tight for constant r and k, i.e., that the size of a largest $\mathcal{F}$-packing of G and the size of a smallest $\mathcal{F}$-covering of G are asymptotically equal to each other. Rödl [22] verified the conjecture by one of the first applications of the nibble method. Observe that in this setting, the two parts of the conjecture are trivially equivalent. The reason for this is that the size of the elements of the family $\mathcal{F}$ does not grow with n: from a packing $\mathcal{F}'$ of G of size $(1-\varepsilon)\binom{n}{r}/\binom{k}{r}$, one obtains a covering $\mathcal{F}''$ of G of size $(1 + \varepsilon\binom{k}{r})\binom{n}{r}/\binom{k}{r}$ by simply taking additionally one k-clique for every r-edge that was not contained in any clique from $\mathcal{F}'$.

In this paper we study a covering problem where the sets in $\mathcal{F}$ grow with n and the above equivalence is not entirely clear. Let $r = 2$, so our objects are usual graphs. For a graph G, we consider the family $\mathcal{H} = \mathcal{H}(G)$ of all Hamilton cycles of G. The corresponding concepts of decomposition, packing, and covering are called *Hamilton decomposition*, *Hamilton packing* and *Hamilton covering*, respectively.

The most well-known fact about Hamilton decompositions is the nearly folklore result of Walecki (see e.g. [2]), stating that for every odd n, the complete graph K_n has a Hamilton decomposition. In general, however, not many graphs are known to have a Hamilton decomposition; the interested reader is referred to [3].

Given that the minimum degree of a Hamilton cycle is 2, the maximum size of a Hamilton packing of a graph with minimum degree δ is $\lfloor \delta/2 \rfloor$. Interestingly, the random graph $G(n, p)$ seems to match this bound tightly. There has been an extensive research considering Hamilton packings of the random graph $G(n, p)$. A classic result of Bollobás [5] and Komlós and Szemerédi [16] states that as soon as the minimum degree of the random graph is 2, it contains a Hamilton cycle a.a.s. This result was extended by Bollobás and Frieze [7] who showed that we can replace 2 by $2k$ for any constant k and obtain a Hamilton packing of size k a.a.s. Frieze and Krivelevich [10] proved that for every constant and slightly subconstant p, $G \sim G(n, p)$ contains a packing of $(1 + o(1))\delta(G)/2$ Hamilton cycles a.a.s. They also conjectured that for every $p = p(n)$ there exists a Hamilton packing of $G \sim G(n, p)$ of size $\lfloor \delta(G)/2 \rfloor$ (and, in the case where $\delta(G)$ is odd, an additional (disjoint to the Hamilton cycles of the packing) matching of size $\lfloor n/2 \rfloor$) a.a.s. Frieze and Krivelevich [11] proved their conjecture as long as $p = (1 + o(1)) \ln n/n$, which was extended to the range of $p \leq 1.02 \ln n/n$ by Ben-Shimon, Krivelevich and Sudakov [4]. Meanwhile, Knox, Kühn and Osthus [14] extended the result from [10] to the range of $p = \omega(\ln n/n)$, and then proved the conjecture for $\ln^{50} n/n < p < 1 - n^{-1/4} \ln^9 n$ [15]. Very recently it was also proven by Krivelevich and Samotij [18] that there exists a positive constant $\varepsilon > 0$ such that for the range of $\ln n/n \leq p \leq n^{\varepsilon-1}$, $G \sim G(n, p)$ contains a Hamilton packing of size $\lfloor \delta(G)/2 \rfloor$ a.a.s., implying the conjecture in this range of p up to the existence of the additional matching.

To the best of our knowledge the dual concept of Hamilton covering of $G(n, p)$ has not been studied. Obviously, the size of any Hamilton covering of graph G is at least $\lceil \Delta(G)/2 \rceil$, where $\Delta(G)$ denotes the maximum degree of G. Recall that for $p = p(n) \gg \ln n/n$, $\Delta(G(n, p)) = (1 + o(1))np = \delta(G(n, p))$ a.a.s., hence the minimum size of a Hamilton cover and the maximum size of a Hamilton packing have a chance to be asymptotically equal. We prove that this, in fact, is the case for the range $p > n^{\alpha-1}$ where $\alpha > 0$ is an arbitrary small constant.

Theorem 1. *For any $\alpha > 0$, for $p \geq n^{\alpha-1}$ a.a.s. $G(n, p)$ can be covered by $(1 + o(1))np/2$ Hamilton cycles.*

1.1. **Pseudorandom setting.** Our argument will proceed in an appropriately chosen pseudorandom setting.

By the *neighborhood* $N(A)$ of a set A, we mean all the vertices outside A having at least one neighbor in A. Note that we explicitly exclude A from $N(A)$.

The following definition contains the most important notions of the paper.

Definition 2. *We say that a graph G has the* small expansion property $S(s,g)$ *with* expansion factor s *and* boundary g, *if for any set $A \subset V(G)$ of size $|A| \leq g$, the neighborhood of A satisfies $|N(A)| \geq |A|s$.*

We say that G has the large expansion property $L(l)$ *with* frame l, *if there is an edge between any two disjoint sets $A, B \subset V(G)$ of size $|A|, |B| \geq l$.*

We call a graph (s,g,l)-expander, if it satisfies properties $S(s,g)$ and $L(l)$.

We refer to an $\left(s, \frac{4n\ln s}{s\ln n}, \frac{n\ln s}{3000\ln n}\right)$-expander on n vertices briefly as an s-expander.

Notice that the expander-property is monotone in all three parameters, meaning that every (s,g,l)-expander is also an $(s-1,g,l)$-, an $(s,g-1,l)$- and an $(s,g,l+1)$-expander.

The proof of Theorem 1 is based on the following result.

Theorem 3. *For every constant $\alpha > 0$ there exists $n_0 = n_0(\alpha)$ such that for every $n \geq n_0$ and every $h > 0$, every n^α-expander graph G on n vertices with a Hamilton packing of size h has a Hamilton covering of size at most $h + \frac{28000(\Delta(G)-2h)}{\alpha^4}$.*

1.2. **Structure of the paper and outline of the proofs.** In Section 2, we prove Theorem 3 in the following main steps:

- in Lemma 5 we show that the small expansion property is "robust" in the sense that after deleting a small linear-size set of vertices from a graph satisfying S, we still have a large subgraph satisfying S with slightly worse parameters;
- in Lemma 7 we show that vertex-disjoint paths in an expander can be joined together without losing or gaining too many edges;
- in Lemma 8, Fact 9 and Lemma 10, we learn how to apply the rotation-extension technique, developed by Pósa [21], without losing too many important edges;
- Lemma 11 contains the main proof of the paper. There we show applying the previous technical statements that a small matching in an expander can mostly be covered by a Hamilton cycle of the same graph.
- After having digested the statement of Lemma 11 in Corollary 12, we prove Theorem 3 in 6 easy lines.

Section 3 contains the proof of Theorem 1. There, we first prove in Lemma 13 that $G(n,p)$ is a $\sqrt[5]{np}$-expander a.a.s., and prove Theorem 1 using this fact and the result from [14].

And finally, in Section 4, we give some concluding remarks leading to further open questions.

In general, we may drop floor and ceiling signs to improve the readability when they do not influence the asymptotic statements.

2. Proof of Theorem 3

The maximum over all pairs of vertices $x, y \in V(G)$ of the length of a shortest xy-path in a graph G is called the *diameter* of G and is denoted by $\mathrm{diam}(G)$. We start with an observation showing that graphs with appropriate expander properties have small diameter.

Observation 4. *Any n-vertex graph satisfying $S(s,g)$ and $L(l)$ for some s, g, l with $s > 1$ and $l \leq sg$ has diameter at most $2\ln n/\ln s + 3$.*

Proof Let G be a graph on n vertices. By the small expansion property of G, we know for every vertex $x \in V(G)$ that $|N(x)| \geq s$. In the following, for any $x \in V(G)$ and any $i \in \mathbb{N}$, let us denote by $B_i(x)$ the set of all vertices with distance at most i to x, i.e., $B_i(x)$ contains all those vertices $y \in V(G)$, for which there exists an xy-path of length at most i. Inductively, as long as for an $i \in \mathbb{N}$ and an $x \in V(G)$ it holds that $|B_i(x)| \leq g$, we obtain $|B_{i+1}(x)| \geq s^{i+1}$. Now, the last index i such that $|B_i(x)| \leq g$ is obviously at most $\lfloor \ln n / \ln s \rfloor - 1$, since $sg \leq n$. Hence, $|B_{\lfloor \ln n / \ln s \rfloor}(x)| \geq g$, and $|B_{\lfloor \ln n / \ln s \rfloor + 1}(x)| \geq sg \geq l$. Thus, for any two vertices $x, y \in V(G)$ we know that both sets $B_{\lfloor \ln n / \ln s \rfloor + 1}(x)$ and $B_{\lfloor \ln n / \ln s \rfloor + 1}(y)$ have at least l vertices. This guarantees that by property $L(l)$ of G, either these sets are not disjoint or there exists an edge between these sets. Each of these facts implies an x, y-path of length at most $2 \lfloor \ln n / \ln s \rfloor + 3$, and the observation follows. $\square$

The following lemma shows that if we have a graph satisfying the small expansion property and we remove an arbitrary subset of small size from the vertex set, then with the additional removal of a (possibly) even smaller subset we can recover some of the small expansion property again.

Lemma 5 (Induced expander lemma). *For any s, g, every graph $G = (V, E)$ satisfying $S(s, g)$ has the following induced expansion property. For every $D \subset V$ of size $|D| \leq \frac{gs}{4}$, there exists a set $Z \subset V$ of size $|Z| \leq \frac{2|D|}{s}$, such that the graph $G[V \setminus (D \cup Z)]$ satisfies $S(\frac{s}{2}, \frac{g}{2})$.*

Proof Let Z be a largest set in $V \setminus D$ among subsets of $V \setminus D$ of size at most g not satisfying the small expansion property with expansion factor $s/2$ in $G[V \setminus D]$, meaning that $|N(Z) \setminus D| < |Z|\frac{s}{2}$ (assuming there exists such a set; otherwise we are done by setting $Z = \emptyset$ in the statement of the lemma). We denote $U = V \setminus (D \cup Z)$ and remember that

$$|N(Z) \cap U| < \frac{|Z|s}{2}.$$

Thus, by the property $S(s, g)$ of G,

$$|D| + |N(Z) \cap U| \geq |N(Z)| \geq |Z|s,$$

implying that

$$|Z| \leq \frac{2|D|}{s} \leq \frac{g}{2}.$$

Assume now for the sake of contradiction that $G[U]$ does not satisfy $S(\frac{s}{2}, \frac{g}{2})$, i.e. there exists an $A \subset U$ with $|A| \leq \frac{g}{2}$ and $|N(A) \cap U| < |A|\frac{s}{2}$. Then the set $A \cup Z \subset V \setminus D$ satisfies both properties

$$|A \cup Z| \leq g$$

and

$$|N(A \cup Z) \cap U| \leq |N(A) \cap U| + |N(Z) \cap U| < \frac{|A \cup Z|s}{2},$$

contradicting the assumption that Z is a largest set in $V \setminus D$ with these properties. $\square$

The following concept allows us to join together paths in an appropriate way.

By the *k-end* of a path we mean the at most $2k$ vertices of this path's vertex set with distance at most $k-1$ to one of the two endpoints, whereas the endpoints have distance 0 to themselves and so are part of any k-end with $k \geq 1$. We call a path *non-trivial* if its length is at least one. Given a family $\mathcal{M}$ of non-trivial vertex-disjoint paths, we call a family $\mathcal{M}'$ of non-trivial vertex-disjoint paths a (d,k)-*extension of* $\mathcal{M}$, if we have that

- $\mu := |\mathcal{M}| - |\mathcal{M}'| \geq 0$,
- $|\bigcup_{P \in \mathcal{M}} E(P) \setminus \bigcup_{P' \in \mathcal{M}'} E(P')| \leq 2(k-1)\mu$, and
- $|\bigcup_{P' \in \mathcal{M}'} E(P') \setminus \bigcup_{P \in \mathcal{M}} E(P)| \leq (d+2)\mu$.

Informally speaking, we neither gain nor lose too many edges relative to the decrease in the number of paths, while passing from $\mathcal{M}$ to $\mathcal{M}'$. Behind the definition, in our mind lies the iterative algorithm which in each step either deletes a path of length less than $2k-1$ from the family $\mathcal{M}$ or joins the k-ends of two paths in $\mathcal{M}$ by paths of length at most $d+2$ (and deletes the k-ends). At the end, when one cannot do either, we have a size-minimal (d,k)-extension, the main object we use in our proofs. This intuition is to be formalized in Lemma 6.

To clarify the notation, when speaking about the *size* $|\mathcal{M}|$ of a family of non-trivial vertex-disjoint paths $\mathcal{M}$, we mean the number of these paths. Note that any extension of $\mathcal{M}$ has at most as many paths as $\mathcal{M}$. Notice that the definition of extension is transitive, i.e., if for some pair (d,k), $\mathcal{M}'$ is a (d,k)-extension of $\mathcal{M}$ and $\mathcal{M}''$ is a (d,k)-extension of $\mathcal{M}'$, then $\mathcal{M}''$ is a (d,k)-extension of $\mathcal{M}$. Furthermore, the relation is also reflexive, i.e., $\mathcal{M}$ is a (d,k)-extension of itself for any $k \geq 1$ and every $d \geq 0$. We say that $\mathcal{M}'$ is a size-minimum (d,k)-extension of $\mathcal{M}$, if every (d,k)-extension $\mathcal{M}''$ of $\mathcal{M}$ has size at least $|\mathcal{M}''| \geq |\mathcal{M}'|$. Notice that the transitivity and reflexivity of extensions imply that then $\mathcal{M}'$ is a size-minimum (d,k)-extension of itself.

We use the notation $V(\mathcal{M}) := \bigcup_{P \in \mathcal{M}} V(P)$ for the set of all vertices appearing in one of the paths of $\mathcal{M}$ and $E(\mathcal{M}) := \bigcup_{P \in \mathcal{M}} E(P)$ for the corresponding edge set.

For our applications we will mostly be interested in size-minimum extensions. The following lemma provides us with two basic properties of size-minimum extensions: namely that they do not contain very short paths and that there are no short paths between the ends of two distinct paths from such an extension.

Lemma 6. *Let G be a graph, $k \geq 1$ and $d \geq 0$ arbitrary integers and $\mathcal{M}$ a family of non-trivial vertex-disjoint paths, which is a size-minimum (d,k)-extension of itself. Then the following holds:*

1. *there exists no path in $\mathcal{M}$ of length less than $2k-1$ and*
2. *for every two distinct paths $P_1, P_2 \in \mathcal{M}$, for every vertex x in the k-end of P_1 and every vertex y in the k-end of P_2, for every $a \in N(x) \setminus V(\mathcal{M})$ and $b \in N(y) \setminus V(\mathcal{M})$, there exists no ab-path of length at most d in $G - V(\mathcal{M})$.*

Proof 1. Suppose for the sake of contradiction that there exists a path P of length at most $2k-2$ in $\mathcal{M}$. Delete this path from $\mathcal{M}$ and call the resulting family of non-empty vertex-disjoint paths $\mathcal{M}'$. Then since $|\mathcal{M}| - |\mathcal{M}'| = 1$, we obtain

$$|E(\mathcal{M}) \setminus E(\mathcal{M}')| = |E(P)| \leq 2k-2 = 2(k-1)\left(|\mathcal{M}| - |\mathcal{M}'|\right),$$

and

$$|E(\mathcal{M}') \setminus E(\mathcal{M})| = 0 \le (d+2)\left(|\mathcal{M}| - |\mathcal{M}'|\right).$$

Hence, $\mathcal{M}'$ is a (d, k)-extension of $\mathcal{M}$, contradicting the minimality of $\mathcal{M}$ as a (d, k)-extension of itself.

2. Suppose to the contrary that there exist two distinct paths $P_1, P_2 \in \mathcal{M}$ with a vertex x in the k-end of P_1, a vertex y in the k-end of P_2, and vertices $a \in N(x) \setminus V(\mathcal{M})$ and $b \in N(y) \setminus V(\mathcal{M})$ such that in $G - V(\mathcal{M})$, there is an ab-path of length at most d. Let us call this path P_3.

The vertex x splits P_1 into two subpaths, the shorter one has length at most $k-1$. Let us call the longer one P_x and construct P_y from P_2 analogously. By connecting the paths P_x, P_3 and P_y via the edges xa and by, we obtain a new path P'. Replace the two paths P_1 and P_2 in $\mathcal{M}$ by P' and call the resulting family of non-trivial vertex-disjoint paths $\mathcal{M}'$. Then $\mu = |\mathcal{M}| - |\mathcal{M}'| = 1$.

By construction $E(\mathcal{M})$ contains all but at most $k-1$ edges from each of P_1 and P_2, so

$$|E(\mathcal{M}) \setminus E(\mathcal{M}')| = |E(P_1) \cup E(P_2) \setminus E(P')| \le 2(k-1) = 2(k-1)\mu,$$

Furthermore, $E(\mathcal{M}')$ contains at most $|E(P') \setminus (E(P_1) \cup E(P_2))| \le d+2$ edges that are not contained in $E(\mathcal{M})$, so

$$|E(\mathcal{M}') \setminus E(\mathcal{M})| \le d+2 = (d+2)\mu.$$

In conclusion, $\mathcal{M}'$ is a (d, k)-extension of $\mathcal{M}$, contradicting the minimality of $\mathcal{M}$ as a (d, k)-extension of itself. $\square$

Lemma 7. *For every $k \ge 1$, in every n-vertex graph G satisfying $S(s, g)$ and $L(l)$ for some s, g and l with $sg \ge 4l$, $s \ge 18$, and n sufficiently large the following holds. For every family of non-trivial vertex-disjoint paths $\mathcal{M}$ on at most $|V(\mathcal{M})| \le \alpha gs/20$ vertices in G, where $\alpha = \ln s / \ln n$, there exists a $(6/\alpha, k)$-extension $\mathcal{M}'$ of $\mathcal{M}$ of size $|\mathcal{M}'| \le \frac{5|V(\mathcal{M})|}{2\alpha ks} + 1$.*

Proof   Let $\mathcal{M}'$ be a size-minimum $(6/\alpha, k)$-extension of $\mathcal{M}$. (One exists, since $\mathcal{M}$ is a $(6/\alpha, k)$-extension of itself.) Apply Lemma 5 with $D = V(\mathcal{M}')$ to find the corresponding sets Z and U (meaning that $|Z| \le 2|V(\mathcal{M}')|/s$, $U = V(G) \setminus (V(\mathcal{M}') \cup Z)$ and $G[U]$ satisfies $S(s/2, g/2)$). Lemma 5 can be applied since

$$\begin{aligned}
|D| &= |V(\mathcal{M}')| = |E(\mathcal{M}')| + |\mathcal{M}'| \\
&\le |E(\mathcal{M})| + (6/\alpha + 2)(|\mathcal{M}| - |\mathcal{M}'|) + |\mathcal{M}'| \\
&= |V(\mathcal{M})| + (6/\alpha + 1)(|\mathcal{M}| - |\mathcal{M}'|) \\
&< (1 + (6/\alpha + 1)/2)|V(\mathcal{M})| < 5|V(\mathcal{M})|/\alpha \le gs/4.
\end{aligned}$$

In the last line we used that the paths of $\mathcal{M}$ are non-trivial (implying that $|\mathcal{M}| \le |V(\mathcal{M})|/2$) and that $\alpha < 1$.

Let now x and y be two vertices from the k-ends of two distinct paths P_1 and $P_2 \in \mathcal{M}'$, respectively, and suppose each of them has a neighbor in U. Let $a \in U$ be a

neighbor of x and let $b \in U$ be a neighbor of y. Since $G[U]$ satisfies both $S(s/2, g/2)$ and $L(l)$ with $l \leq s/2 \cdot g/2$, Observation 4 implies that the diameter of $G[U]$ is at most

$$\mathrm{diam}(G[U]) \leq 2 \ln n / \ln(s/2) + 3 \leq 2 \frac{\ln n}{\frac{2}{3} \ln s} + 3 < \frac{6}{\alpha},$$

where the next to last inequality holds since $s \geq 8$. Hence, there exists an ab-path of length at most $\frac{6}{\alpha}$ in $G[U] \subseteq G[V(G) \setminus V(\mathcal{M})]$, contradicting the size-minimality of $\mathcal{M}'$ by Lemma 6.

Consequently, there can be at most one such path in $\mathcal{M}'$ that has a vertex in its k-end with a neighbor in U. Hence the k-ends of at least $|\mathcal{M}'| - 1$ paths have neighbors only in $V(\mathcal{M}') \cup Z$. By Lemma 6 each such path has length at least $2k - 1$ so we found a set S of $2k(|\mathcal{M}'| - 1)$ vertices such that S together with its neighborhood contains at most

$$
\begin{aligned}
|S| + |N(S)| \quad &\leq \quad |Z| + |V(\mathcal{M}')| \\
&< \quad 2|V(\mathcal{M}')|/s + |V(\mathcal{M}')| \leq 10|V(\mathcal{M}')|/9 \\
&< \quad \frac{10}{9}(|V(\mathcal{M})| + (6/\alpha + 1)|\mathcal{M}|) \\
&< \quad 5|V(\mathcal{M})|/\alpha < gs
\end{aligned}
$$

vertices. Here in the next to last inequality we use the fact that $\mathcal{M}$ contains no path of length 0, implying that $|\mathcal{M}| \leq |V(\mathcal{M})|/2$. Now, if S contained at least g elements, then taking a subset $S' \subseteq S$ of size g would lead to a contradiction via the small expansion property: $gs = |S'|s \leq |N(S') \cup S'| \leq |N(S) \cup S| < gs$. Hence $|S| < g$ and then the small expansion property implies that

$$2sk(|\mathcal{M}'| - 1) = s|S| \leq |N(S)| < 5|V(\mathcal{M})|/\alpha.$$

$\square$

The proof of the main theorem 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 or pseudorandom graphs (see for example [6], [9], [16], [19]).

Let G be a graph and let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path in G. If $1 \leq i \leq q - 2$ and (v_q, v_i) is an edge of G, then there exists a path $P' = (v_1 v_2 \ldots v_i v_q v_{q-1} \ldots v_{i+1})$ in G. 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 properties S and L to find a path on the same vertex set as P_0 with large rotation endpoint sets.

The next lemma is a slight strengthening of Claim 2.2 from [13] with a similar proof. It shows that in any graph having the small and large expansion properties for any path P_0 and its endpoint v_1 many other endpoints can be created by a small number of rotations with fixed endpoint v_1. In our setting we must also care about not breaking any of the edges from a small "forbidden" set F.

Lemma 8. *Let $G = (V, E)$ be a graph on n vertices that satisfies $S(s, g)$ and $L(l)$ with $s \geq 21$, $sg/3 > l$, and $l \leq n/24$. Let $P_0 = (v_1, v_2, \ldots, v_q)$ be a path in G and $F \subseteq E(P_0)$ with $|F| \leq s/24 - 1/2$. Denote by $B(v_1) \subset V(P_0)$ the set of all vertices $v \in V$ for which there is a v_1v-path on the vertex set $V(P_0)$ which can be obtained from P_0 by at most $3\frac{\ln n}{\ln s}$ rotations with fixed endpoint v_1 not breaking any of the edges of F. Then $B(v_1)$ satisfies one of the following properties:*

- *there exists a vertex $v \in B(v_1)$ with a neighbor outside $V(P_0)$, or*
- *$|B(v_1)| \geq n/3$.*

Proof   Assume $B(v_1)$ does not have the first property (i.e., for every $v \in B(v_1)$ it holds that $N(v) \subseteq V(P_0)$).

Let t_0 be the smallest integer such that $\left(\frac{s}{3}\right)^{t_0-2} \geq g$; note that $t_0 \leq \frac{\ln g}{\ln(s/3)} + 3 \leq 3\frac{\ln n}{\ln s}$, because $21 \leq s$.

We construct a sequence of sets $S_0, \ldots, S_{t_0} \subseteq B(v_1) \subseteq V(P_0) \setminus \{v_1\}$ of vertices, such that for every $0 \leq t \leq t_0$ 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 $0 \leq i < t$, the non-v_1-endpoint of the path after the ith rotation is contained in S_i. Moreover, $|S_t| = \left(\frac{s}{3}\right)^t$ for every $t \leq t_0 - 3$, $|S_{t_0-2}| = g$, $|S_{t_0-1}| = l$, and $|S_{t_0}| \geq n/3$. Furthermore, for any $1 \leq t \leq t_0$, $V(F) \cap S_t = \emptyset$, and hence no edge from F got broken in any of the rotations.

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 $0 < t \leq t_0 - 2$ 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

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

be the set of potential pivots for the tth rotation, and notice that $T \subset V(P_0)$ due to our assumption, since $T \subseteq N(S_{t-1})$ and $S_{t-1} \subseteq B(v_1)$. 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_1y-path Q can be obtained from P_0 by $t-1$ rotations not breaking any edge from F such that after the jth rotation, the non-v_1-endpoint is in S_j for every $0 \leq j \leq t-1$. 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 and also not in $V(F)$, 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 $Q \setminus F$.

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 the endpoint 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 we have

$$
\begin{aligned}
|\hat{S}_t| \;\geq\; & \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t-1})| - 3(1 + |S_1| + \ldots + |S_{t-1}| + 2|F|)\right) \\
\geq\; & \frac{s}{2}\left(\frac{s}{3}\right)^{t-1} - \frac{3}{2}\frac{(s/3)^t - 1}{s/3 - 1} - (s/8 - 3/2) \geq \left(\frac{s}{3}\right)^t,
\end{aligned}
$$

where in the third inequality we use the small expansion property for the set S_{t-1} (where $|S_{t-1}| \leq g$ by the definition of t_0) and the last inequality follows since $s \geq 21$. Clearly we can delete arbitrary elements of $\hat{S}_t$ to obtain S_t of size $\left(\frac{s}{3}\right)^t$ if $t \leq t_0 - 3$ and of size g 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 calculations are slightly different. If $g = 1$, then $t_0 - 1 = 1$, and analogously to the above calculation we obtain $\hat{S}_1$ with $|\hat{S}_1| \geq s/3 \geq l$. Otherwise, for $g \geq 2$, since $|N(S_{t_0-2})| \geq sg$, we have

$$
\begin{aligned}
|\hat{S}_{t_0-1}| \;\geq\; & \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t_0-2})| - 3(1 + |S_1| + \ldots + |S_{t_0-4}| + |S_{t_0-3}| + |S_{t_0-2}| + 2|F|)\right) \\
\geq\; & gs/2 - \frac{3}{2}\frac{(s/3)^{t_0-2} - 1}{s/3 - 1} - 3g/2 - (s/8 - 3/2) \\
\geq\; & gs/2 - 2 \cdot \left(\frac{s}{3}\right)^{t_0-3} - \frac{3}{2}g - (s/8 - 3/2) \\
\geq\; & gs/2 - 2g - \frac{3}{2}g - (s/8 - 3/2) \geq gs/3 > l,
\end{aligned}
$$

where the inequality in the last but one line and the last but one inequality follow since $s \geq 21$ and $g \geq 2$. We delete arbitrary elements of $\hat{S}_{t_0-1}$ to obtain S_{t_0-1} of size l.

For S_{t_0} the difference in the calculation comes from using the expansion guaranteed by the property L, rather than the property S. That is, we use the fact that $|N(S_{t_0-1})| \geq n - 2l$. Hence, we obtain

$$
\begin{aligned}
|S_{t_0}| \;\geq\; & \frac{1}{2}|T| \geq \frac{1}{2}\left(|N(S_{t_0-1})| - 3(1 + |S_1| + \ldots + |S_{t_0-2}| + |S_{t_0-1}| + 2|F|)\right) \\
\geq\; & \frac{n}{2} - l - 4g - \frac{3}{2}l - (s/8 - 3/2) \\
>\; & \frac{n}{3},
\end{aligned}
$$

where the last inequality follows since $4g \leq sg/3 \leq l$, $s \leq sg \leq 3l$ and $l \leq n/24$.

The set S_{t_0} is by construction a subset of $B(v_1)$, concluding the proof of the lemma. $\qquad\square$

Let H be a graph with a spanning path $P = (v_1, \ldots, v_m)$. For $2 \leq i < m$, let us define the auxiliary graph $H_i^+ = H_{v_i}^+$ by adding a vertex and two edges to H as follows: $V(H_i^+) = V(H) \cup \{w\}$, $E(H_i^+) = E(H) \cup \{(v_m, w), (v_i, w)\}$. Let $P_i = P_{v_i}$ be the spanning path of H_i^+ which we obtain from the path $P \cup \{(v_m, w)\}$ 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 $|S^{v_i}| < \frac{m}{43}$ 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.

Hefetz et al. [13] showed that H has many good initial pivots provided that property L is satisfied.

Fact 9 ([13] Lemma 2.3). *Let H be a graph satisfying $L(m/43)$ with a spanning path $P = (v_1, \ldots, v_m)$. Then*

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

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

With these statements in our toolbox, we can prove the following important technical lemma. It states that we can rotate a path until it can be extended, and still do not break too many of the important edges.

Lemma 10. *For every sufficiently large n and every $s = s(n)$ with $s \geq 21$, in every s-expander graph G on n vertices every path P_0 in G has the following property. For every pair of sets $F \subset F' \subseteq E(P_0)$ of at most $|F| \leq s/24 - 1/2$ and $|F'| \leq \frac{n \ln s}{9200 \ln n}$ edges of P_0, there exists a path P' in G between some $x, y \in V(P_0)$, such that $V(P') = V(P_0)$, $F \subset E(P')$, $|F' \setminus E(P')| \leq 6 \ln n / \ln s$, and G contains the edge $\{x, y\}$, or the set $\{x, y\}$ has neighbors outside P'.*

Proof   Assume for the sake of contradiction that the statement is not true. Let $P_0 = (v_1, v_2, \ldots, v_q)$, and let $A_0 = B(v_1) \subset V(P_0)$ be the set corresponding to P_0 and F as in Lemma 8, meaning 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 $t_0 = 3 \frac{\ln n}{\ln s}$ rotations with fixed endpoint v_1 not breaking any of the edges of F. Clearly, at most $3 \frac{\ln n}{\ln s}$ edges from F' were broken by the rotations, thus by our assumption every $v \in A_0$ has no neighbors outside P_0, hence by Lemma 8 we obtain $|A_0| \geq n/3$. For every $v \in A_0$ fix a $v_1 v$-path $P^{(v)}$ with the above properties and, again using our assumption and Lemma 8, construct sets $B(v)$, $|B(v)| \geq n/3$, of endpoints of paths with fixed endpoint v, obtained from the path $P^{(v)}$ by at most t_0 rotations not breaking any edge from F. To summarize, for every $a \in A_0$ and $b \in B(a)$ there is a path $P(a,b)$ joining a and b on the vertex set $V(P_0)$, which is obtainable from P_0 by at most $\varrho := 2t_0 = \frac{6 \ln n}{\ln s}$ rotations not breaking any of the edges from F. 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 undirected vertex disjoint segments $I_1, I_2, \ldots, I_{2\varrho}$ of length at least $\lfloor |V(P_0)|/2\varrho - 1 \rfloor$ each. As every $P(a,b)$ 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*. Although the segments themselves are undirected, they 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 tuples $\varsigma = (I_i, o_i, I_j, o_j)$, where I_i and I_j are unbroken segments of P_0, which occur in this order on $P(a,b)$,

and o_i and o_j denote their corresponding relative orientation. We call such a tuple ς *unbroken*, and say that $P(a, b)$ *contains* ς.

For a given unbroken tuple ς, we consider the set $C(\varsigma)$ of ordered pairs (a, b), $a \in A_0$, $b \in B(a)$, such that $P(a, b)$ contains ς.

The total number of unbroken tuples is at most $2^2(2\varrho)_2$. Any path $P(a, b)$ contains at least ϱ unbroken segments, and thus at least $\binom{\varrho}{2}$ unbroken tuples. The average, over unbroken tuples, of the number of pairs (a, b) such that $P(a, b)$ contains a given unbroken tuple is therefore at least

$$\frac{n^2}{9} \cdot \frac{\binom{\varrho}{2}}{2^2(2\varrho)_2} \geq 0.003n^2.$$

Thus, there is an unbroken tuple ς_0 and a set $C = C(\varsigma_0)$, $|C| \geq 0.003n^2$ of pairs (a, b), such that for each $(a, b) \in C$, the path $P(a, b)$ contains ς_0. Let $\hat{A} = \{a \in A_0 : C$ contains at least $0.003n/2$ pairs with a as first element$\}$. Since $|A_0|, |B(a)| \leq n$, we have $0.003n^2 \leq |C| \leq |\hat{A}|n + n \cdot \frac{0.003n}{2}$, entailing $|\hat{A}| \geq 0.003n/2$. For every $a \in \hat{A}$, let $\hat{B}(a) = \{b : (a, b) \in C\}$. Then, by the definition of $\hat{A}$, for every $a \in \hat{A}$ we have $|\hat{B}(a)| \geq 0.003n/2$.

For an unbroken tuple $\varsigma_0 = (I_i, o_i, I_j, o_j)$, we divide it into two oriented segments, $\varsigma_0^1 = \vec{I}_i$ and $\varsigma_0^2 = \vec{I}_j$, both of them maintaining the orientation in ς_0. Notice that for every $a \in \hat{A}$ and $b \in \hat{B}(a)$, in the path $P(a, b)$ the segment ς_0^1 comes before ς_0^2. For $i = 1, 2$, let us denote by $|\varsigma_0^i|$ the number of vertices in the segment ς_0^i. Then for both segments ς_0^1 and ς_0^2, we have that $|\varsigma_0^i| > n/(7\varrho)$. Let s_1 be the first vertex of ς_0^1, x be the last vertex of ς_0^1, and let y be the first vertex of ς_0^2 and s_2 be the last vertex of ς_0^2.

We construct a graph H_1 with $V(\varsigma_0^1)$ as vertex set. The edge set of H_1 is defined as follows. First, we add all edges of $G[V(\varsigma_0^1)]$, except for those that are incident with s_1, x or a vertex in $V(F')$. Further, we add all the edges from $E(\varsigma_0^1)$. Note that all the edges in H_1 are also edges of G. By its construction, ς_0^1 is a spanning path in H_1 starting at s_1 and ending at x. Let us denote the path reversed to ς_0^1 (spanning path in H_1, starting at x and ending at s_1) by P. We would like to apply Fact 9 to H_1 with $m = |V(\varsigma_0^1)|$ and P as the corresponding spanning path. The condition of the Fact holds since G satisfies property $L(l)$. Indeed, $l = \frac{n \ln s}{3000 \ln n}$, $m > n/(7\varrho) = \frac{n \ln s}{42 \ln n}$, the edges of H_1 differ from the edges of G only at $V(F')$ and at the endpoints of the segment ς_0^1, and $|V(F') \cup \{x, s_1\}| \leq \frac{n \ln s}{4600 \ln n} + 2$, implying $l + |V(F') \cup \{x, s_1\}| < \frac{n \ln s}{43 \cdot 42 \ln n} < m/43$. Notice that this are the lines justifying the choices of the integers 3000 and 9200 in the bounds for l and $|F'|$. Hence, H_1 satisfies $L(m/43)$, thus by Fact 9 at least a $\frac{36}{43}$-fraction of the vertices of H_1 are good.

For ς_0^2 we act similarly: construct a graph H_2 from ς_0^2 by adding all edges of G with both endpoints in the interior of ς_0^2 but not in $V(F')$ and edges from $E(\varsigma_0^2)$ to H_2. Then ς_0^2 forms an oriented spanning path in H_2, starting at y and ending at the last vertex s_2 of ς_0^2. Again, due to property L, Lemma 9 applies here, so at least a $\frac{36}{43}$-fraction of the vertices of H_2 are good.

Recall that s_1 is the first vertex of ς_0^1. Since $|\hat{A}| \geq 0.003n/2 > l + 1$ and H_1 has at least $\frac{36}{43}m > l + |V(F') \cup \{x, s_1\}|$ good vertices, there is an edge of G between a vertex $\hat{a} \in \hat{A} \setminus \{s_1\}$ and a good vertex $g_1 \in V(\varsigma_0^1) \setminus (V(F') \cup \{x, s_1\})$.

Similarly, as $|\hat{B}(\hat{a})| \geq 0.003n/2 > l+1$ and there are more than $l + |V(F') \cup \{y, s_2\}|$ 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 $g_2 \in V(\varsigma_0^2) \setminus (V(F') \cup \{y, s_2\})$.

Consider the path $P(\hat{a}, \hat{b})$ on the vertex set of P_0 connecting $\hat{a}$ and $\hat{b}$ and containing ς_0. The vertices x and y split this path into three sub-paths: 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, making sure that no edge from F' gets broken. We will show that the obtained endpoint sets V_1 and V_2 are sufficiently large (clearly, they are disjoint). Then by property L there will be an edge of G between V_1 and V_2. Since we did not touch R_3, this edge closes the path into a cycle, contradicting the assumption from the beginning of the proof.

First we construct the endpoint set V_1, the endpoint set V_2 can be constructed analogously. Recall the notation from Fact 9: 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 $|\varsigma_0^1|/43 > l$ 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 without breaking any edge from F'. 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 $V(\varsigma_0^1)$). Hence, the edge $(\hat{a}, g_1)$ is present in G, while we will consider the edge $(\hat{a}, s_1)$ *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 xz-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 to end up in an xz-path spanning $V(R_1)$. When performing these rotations, the subpath of R_1 that links $\hat{a}$ to s_1 corresponds to the artificial edge (w, s_1) 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 $H_{g_1}^+$, or (3) the broken edge is in F'. However, the construction of $H_{g_1}^+$ ensures that these problems will never occur. Indeed, in all three cases (1), (2) and (3) the pivot vertex has an artificial edge or an edge from F' incident with it, while having degree at least 3 (as all pivots). However, both endpoints of an artificial edge and both endpoints of edges from $F' \cap H_{g_1}^+$ have degree 2 in $H_{g_1}^+$ (for this last assertion we use the fact that $g_1 \notin \{x, s_1\} \cup V(F')$; 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 $V_1 = S^{g_1}$ containing all edges from $E(R_1) \cap F'$.

Similarly, since there is an edge from $\hat{b}$ to a good vertex g_2 in H_2, $g_2 \notin V(F')$, we can rotate R_2, starting from this edge to get a set $V_2 = S^{g_2}$ of at least l endpoints not

breaking any more edges from F'. In other words we have a spanning path of $G[V(R_2)]$ from y to every vertex of $V_2 = S^{g_2}$ containing the edges from $E(R_2) \cap F'$.

As we noted earlier, since by Fact 9 $|V_1|, |V_2| \geq \frac{m}{43} > l$, property $L(l)$ ensures that there is an edge between V_1 and V_2 in G, say $a'b' \in E(G)$ with $a' \in V_1$ and $b' \in V_2$. This contradicts our assumption, since the rotations we did to obtain $P(a', b')$ from P_0 did not break any edges from F and also all but at most $\varrho \leq 6 \ln n / \ln s$ edges from F' are on the path $P(a', b')$. □

We are now able to prove the main lemma, stating that for every matching there is a Hamilton cycle almost covering it. Notice that we use the same calligraphic letter $\mathcal{M}$ to denote a matching as for families of paths, since we are going to apply extension on the matching and hence we see it as a family of paths of length 1 each.

Lemma 11. *For every constant $\alpha \in (0, 1]$ and for every sufficiently large n the following holds. Let G be an n^α-expander graph on n vertices. For every matching $\mathcal{M}$ in G of size at most $|\mathcal{M}| \leq \alpha^3 n / 9200$ there exists a Hamilton cycle C in G with*

$$|E(\mathcal{M}) \setminus E(C)| \leq \left\lfloor \frac{1036|\mathcal{M}|}{\alpha^3 n^{\alpha/2}} \right\rfloor$$

.

Proof First we proceed inductively to construct a single path via (d, k)-extensions that contains most of the matching edges.

Using Lemma 7 with $s = n^\alpha$, $g = 4\alpha n^{1-\alpha}$, $l = \alpha n / 3000$, $k = 1$, and setting $d = 6/\alpha$, we find a $(d, 1)$-expansion $\mathcal{M}_2$ of $\mathcal{M}_1 = \mathcal{M}$ of size at most $\left\lfloor \frac{5|\mathcal{M}|}{\alpha n^\alpha} + 1 \right\rfloor$ containing all edges of $\mathcal{M}$.

For $i \geq 2$, given a family of vertex-disjoint non-trivial paths $\mathcal{M}_i$ of size

$$2 \leq |\mathcal{M}_i| \leq \left\lfloor \frac{45|\mathcal{M}|}{2\alpha^2 n^{i\alpha/2}} + 1 \right\rfloor$$

on at most

$$|V(\mathcal{M}_i)| \leq (d + 3)|\mathcal{M}|$$

vertices containing all but at most

$$(i - 2) \frac{45|\mathcal{M}|}{\alpha^2 n^{\alpha/2}}$$

edges from $\mathcal{M}$, we construct a size-minimum $(d, n^{(i-1)\alpha/2})$-extension $\mathcal{M}_{i+1}$ of $\mathcal{M}_i$. Then $\mathcal{M}_{i+1}$ satisfies the above properties by construction: it contains all but at most

$$\begin{aligned}
|E(\mathcal{M}) \setminus E(\mathcal{M}_{i+1})| &\leq |E(\mathcal{M}_i) \setminus E(\mathcal{M}_{i+1})| + |E(\mathcal{M}) \setminus E(\mathcal{M}_i)| \\
&\leq 2n^{(i-1)\alpha/2}(|\mathcal{M}_i| - 1) + (i - 2)\frac{45|\mathcal{M}|}{\alpha^2 n^{\alpha/2}} \\
&\leq (i - 1)\frac{45|\mathcal{M}|}{\alpha^2 n^{\alpha/2}}
\end{aligned}$$

edges from $\mathcal{M}$. Furthermore, since $\mathcal{M}_{i+1}$ was constructed from $\mathcal{M}$ by a series of (d, k)-extensions with varying k but fixed d, at most $d + 1$ vertices on average were added for every path removed, thus $\mathcal{M}_{i+1}$ has at most $(d + 3)|\mathcal{M}|$ vertices.

Finally, by Lemma 7 $\mathcal{M}_{i+1}$ has size at most

$$|\mathcal{M}_{i+1}| \leq \left\lfloor \frac{5|V(\mathcal{M}_i)|}{2\alpha n^{(i-1)\alpha/2} n^{\alpha}} + 1 \right\rfloor \leq \left\lfloor \frac{5(d+3)|\mathcal{M}|}{2\alpha n^{(i+1)\alpha/2}} + 1 \right\rfloor \leq \left\lfloor \frac{45|\mathcal{M}|}{2\alpha^2 n^{(i+1)\alpha/2}} + 1 \right\rfloor.$$

The family $\mathcal{M}_{last}$, where $last \leq 2/\alpha + 1 < 3/\alpha$ is the index we stop the induction with, contains only one path P. Let us apply Lemma 5 with $D = V(P)$. This is possible, since P contains at most $(d+3)|\mathcal{M}| < \alpha n = gs/4$ vertices. We obtain the corresponding sets Z and $U = V \setminus (D \cup Z)$ and conclude that the induced graph $G[U]$ satisfies the small expansion property $S(s/2, g/2)$. Theorem 2.5 from [13] states that for every choice of the expansion parameter r with $12 \leq r \leq \sqrt{n}$, every n-vertex graph G satisfying $S\left(r, \frac{n \ln r}{r \ln n}\right)$ and $L\left(\frac{n \ln r}{1035 \ln n}\right)$ is Hamiltonian. Hence, applying this statement to $G[U]$ with $s = n^{\alpha}$, $g = 4\alpha n/n^{\alpha}$ and $r = n^{\alpha}/2$, we see that $G[U]$ is Hamiltonian.

Furthermore, by Lemma 5 we know that

$$(1) \qquad\qquad |Z| \leq 2|V(P)|/n^{\alpha}.$$

Using the small expansion property of G, we obtain an edge between a vertex x in the $\lceil |\mathcal{M}|/n^{\alpha/2} \rceil$-end of P and a vertex y in U in the following way. Take a subset S of size $\min\{2 \lceil |\mathcal{M}|/n^{\alpha/2} \rceil, 4\alpha n^{1-\alpha}\}$ of the $\lceil |\mathcal{M}|/n^{\alpha/2} \rceil$-end of P. If $2 \lceil |\mathcal{M}|/n^{\alpha/2} \rceil > 4\alpha n^{1-\alpha} = g$, then by the small expansion property $|N(S)| \geq sg > |V(P) \cup Z|$. Otherwise, $|S| = 2 \lceil |\mathcal{M}|/n^{\alpha/2} \rceil \leq 4\alpha n^{1-\alpha} = g$, and hence $|N(S)| \geq 2n^{\alpha} \lceil |\mathcal{M}|/n^{\alpha/2} \rceil > |V(P) \cup Z|$.

The vertex x breaks the path P into two subpaths, one of which contains all but at most $|\mathcal{M}|/n^{\alpha/2}$ edges from P. Connecting this path via the edge xy with a Hamilton path in $G[U]$, we create a path R containing all but at most

$$(2) \qquad |\mathcal{M} \setminus E(R)| \leq \left\lfloor \frac{3}{\alpha} \cdot \frac{45|\mathcal{M}|}{\alpha^2 n^{\alpha/2}} \right\rfloor + \left\lfloor |\mathcal{M}|/n^{\alpha/2} \right\rfloor \leq \left\lfloor \frac{136|\mathcal{M}|}{\alpha^3 n^{\alpha/2}} \right\rfloor$$

edges from $\mathcal{M}$ and all vertices from U. Note that if $|\mathcal{M}| < \alpha^3 n^{\alpha/2}/136$ then $\left\lfloor \frac{136|\mathcal{M}|}{\alpha^3 n^{\alpha/2}} \right\rfloor = 0$, so that no edges of $\mathcal{M}$ are lost in forming R. Furthermore, notice that from (1) we have that R is missing only

$$(3) \qquad n - |E(R)| \leq |Z| + |\mathcal{M}|/n^{\alpha/2} + 1 < 2|\mathcal{M}|/n^{\alpha/2} + 1.$$

vertices to be Hamiltonian.

We aim to use Lemma 10 to rotate/extend R into a Hamilton cycle without losing many edges of $\mathcal{M}$ that are already on it.

For this we set $P_0 = R$ and $F' = F = \mathcal{M}$, in case $|\mathcal{M}| < \alpha^3 n^{\alpha/2}/136$ (and thus $\mathcal{M}$ is contained in R by (2)). Otherwise, if $|\mathcal{M}| \geq \alpha^3 n^{\alpha/2}/136$, let $F = \emptyset$ and $F' = \mathcal{M} \cap E(R)$.

We will now use Lemma 10 iteratively, in each step rotating/extending our current path with an edge until it is spanning and then closing it into a Hamilton cycle. Notice that Lemma 10 can be applied throughout the process since $|F'| \leq |\mathcal{M}| < \alpha n/9200$ and $|F| = o(n^{\alpha})$.

Consider the x, y-path P' arising from the application of Lemma 10 to F, F' and P_0. If one of the two vertices x or y has neighbors outside P', we can extend P' with one more edge to obtain a longer path $\hat{P}$ containing P'. We update for our iteration $P_0 := \hat{P}$ and $F' := \mathcal{M} \cap E(\hat{P})$ and start the iteration step again. Notice that in this step, the size of F' decreased by at most $6/\alpha$. If x and y have no neighbors outside P',

then there is a cycle C containing P'. If C is Hamilton, we stop the procedure since this is what we are aiming at. Otherwise, by the connectivity of G (guaranteed by properties S and L and stated implicitly in Observation 4), there is a vertex $w \in V(C)$ with a neighbor outside C, say $aw \in E(G)$, $a \in V(G) \setminus V(C)$. Notice that only one of the edges incident with w in C can be in F', since $F' \subset \mathcal{M}$ is a matching. Removing an edge incident with w in C which is not in F' and adding the edge aw, we obtain a path $\hat{P}$ of length $|\hat{P}| \geq |P'| + 1$ containing all edges from $F' \cap E(P')$. We update for our iteration $P_0 := \hat{P}$ and $F' := \mathcal{M} \cap E(\hat{P})$ and start the iteration step again. Notice that again, in this step, the size of F' decreased by at most $6/\alpha$.

Using (3), we see that after at most $n - |R| \leq 2|\mathcal{M}|/n^{\alpha/2} + 1$ steps the iteration ends and we obtain a Hamilton cycle C. If $|\mathcal{M}| < \alpha^3 n^{\alpha/2}/136$, then C contains all edges from $\mathcal{M}$. Otherwise, C contains all but at most

$$|\mathcal{M} \setminus E(C)| \leq |E(R) \setminus E(C)| + |\mathcal{M} \setminus E(R)| \leq \left(2|\mathcal{M}|/n^{\alpha/2} + 1\right) \cdot \frac{6}{\alpha} + \frac{136|\mathcal{M}|}{\alpha^3 n^{\alpha/2}}$$

$$= \frac{6}{\alpha} + \frac{12|\mathcal{M}|}{\alpha n^{\alpha/2}} + \frac{136|\mathcal{M}|}{\alpha^3 n^{\alpha/2}} < \frac{7}{\alpha} \cdot \frac{148|\mathcal{M}|}{\alpha^3 n^{\alpha/2}}$$

edges from $\mathcal{M}$, completing the proof of the lemma. $\square$

The following corollary condenses all the previous technical work. It states that every matching of an n^α-expander graph can be covered with a constant-size collection of Hamilton cycles.

Corollary 12. *For every constant $\alpha, 0 < \alpha \leq 1$, there exists n_0, such that for every $n \geq n_0$ the following holds. In every n^α-expander graph G on n vertices, for every matching M of G there exist at most $14000/\alpha^4$ Hamilton cycles such that M is contained in their union.*

Proof We start by splitting M into at most $\lceil 4600/\alpha^3 \rceil < 4601/\alpha^3$ matchings of size at most $\alpha^3 n/9200$ each. For every such matching M_1 we set $i = 1$ and perform the following iterative procedure:

- take a Hamilton cycle C_i covering as many of the edges of M_i as possible. If $M_i \subset E(C_i)$, then we found our covering and finish the procedure.
- otherwise, we set $M_{i+1} := M_i \setminus E(C_i)$ and remark that by Lemma 11

$$(4) \qquad |M_{i+1}| \leq \frac{1036|M_i|}{\alpha^3 n^{\alpha/2}} \leq \left(\frac{1036}{\alpha^3 n^{\alpha/2}}\right)^{i+1} |M_1|.$$

Update $i := i + 1$ and start again from the first iteration step.

From (4) we see that after at most $\lfloor 2/\alpha + 1 \rfloor < 3/\alpha$ iteration steps, we get a collection of at most $3/\alpha$ Hamilton cycles $C_1, C_2, \ldots$ covering M_1. Hence, there exist a collection of at most $\frac{3}{\alpha} \cdot \frac{4601}{\alpha^3}$ Hamilton cycles covering M, implying the statement of the lemma. $\square$

We are now able to prove Theorem 3.

Proof We start by taking h disjoint Hamilton cycles into our covering. Removing the union of these cycles from G, we are left with a graph H of maximum degree

exactly $\Delta(H) = \Delta(G) - 2h$. Using at most $2\Delta(H)$ colors we color the edges of H greedily, partitioning them into at most $2\Delta(H)$ matchings. By Corollary 12 for every of these matchings there exist $14000/\alpha^4$ Hamilton cycles covering it, completing the proof of the theorem.

$\square$

3. Proof of Theorem 1

In this section we derive the proof of Theorem 1 from Theorem 3 by checking that $G(n, p)$ is an s-expander a.a.s. with the appropriate choice of p and s. Notice that this choice of s is clearly not optimal, but suffices for our purposes.

Lemma 13. *For every constant α with $0 < \alpha < 1$ and every function $p = p(n) \geq n^{\alpha-1}$, $G(n, p)$ is a $\sqrt[5]{np}$-expander a.a.s.*

Proof   Let $s = \sqrt[5]{np}$.

First we prove that $G(n, p)$ has the small expansion property $S(s, \frac{4n \ln s}{s \ln n})$.

Let $A \subset V(G(n, p))$ be an arbitrary subset of size at most $|A| \leq \frac{4n \ln s}{s \ln n} \leq \frac{4n}{5s}$. The random variable $|N(A)|$ is the sum of the $n - |A|$ characteristic variables of the events $v \in N(A)$ for $v \in V \backslash A$. Hence for the expectation we obtain

$$\mathbf{E}[|N(A)|] = \sum_{v \in V \backslash A} \mathbf{Pr}[v \in N(A)] = (n - |A|)\left(1 - (1-p)^{|A|}\right) > (n - |A|)\frac{|A|p}{1 + |A|p}$$

$$\geq (1 + o(1))n\frac{|A|p}{1 + |A|p} \geq (1 + o(1))|A|s\frac{np/s}{1 + \frac{4}{5}np/s} = |A|s\left(\frac{5}{4} + o(1)\right).$$

Here we first used the simple fact that $(1-p)^{|A|} < \frac{1}{1+|A|p}$, then a couple of times that $|A| \leq \frac{4n}{5s}$. Since the elementary events $v \in N(A)$ that make up $|N(A)|$ are mutually independent the Chernoff bound (see e.g. Theorem A.1.13 in [1]) can be applied to estimate the probability that A is not expanding. We use the above estimate on $\mathbf{E}[|N(A)|]$ several times.

$$\mathbf{Pr}\left[|N(A)| \leq s|A|\right] < \exp\left[-\frac{(\mathbf{E}[|N(A)|] - |A|s)^2}{2\mathbf{E}[|N(A)|]}\right] < \exp\left(-\frac{\left(\left(\frac{1}{5} + o(1)\right)\mathbf{E}[|N(A)|]\right)^2}{2\mathbf{E}[|N(A)|]}\right)$$

$$= \exp\left(-\left(\frac{1}{50} + o(1)\right)\mathbf{E}[|N(A)|]\right) < \exp\left(-\left(\frac{1}{40} + o(1)\right)|A|s\right).$$

By the union bound the probability that $G(n, p)$ does not satisfy property $S\left(s, \frac{4n \ln s}{s \ln n}\right)$ is bounded by

$$\mathbf{Pr}\left[\exists A \subset V, |A| \leq \frac{4n}{5s} : |N(A)| \leq s|A|\right] < \sum_{a=1}^{n}\binom{n}{a}\exp\left[-\left(\frac{1}{40} + o(1)\right)as\right]$$

$$< \sum_{a=1}^{\infty}\left(n\exp\left[-0.01n^{\alpha/5}\right]\right)^a = o(1).$$

To complete the proof we show that $G(n,p)$ has the large expansion property $L\left(\frac{\alpha n}{15000}\right)$.

Let $A, B \subseteq V(G(n,p))$ be fixed subsets of size $|A|, |B| \geq \frac{\alpha n}{15000}$ with $A \cap B = \emptyset$. Then we have that

$$\mathbf{Pr}[\text{there are no edges between } A \text{ and } B] = (1-p)^{|A||B|} < \exp\left(-\frac{\alpha^2 n^2 p}{15000^2}\right).$$

Using union bound over all pairs of such disjoint sets $A, B \subseteq V(G(n,p))$, we get the desired probability

$$\mathbf{Pr}\left[G(n,p) \text{ satisfies } L\left(\frac{\alpha n}{15000}\right)\right] \geq 1 - 4^n \exp\left(-\frac{\alpha^2 n^2 p}{15000^2}\right) = 1 - o(1),$$

proving the lemma. $\qquad\square$

We are now able to prove Theorem 1 using Theorem 3.

Proof   Let p be in the range of the theorem. By Lemma 13 $G(n,p)$ is an $n^{\alpha/5}$-expander a.a.s. We know from [14] that there exists a packing of $(1 - o(1))np/2$ Hamilton cycles into $G(n,p)$. Finally, the maximum degree a.a.s. satisfies $\Delta(G(n,p)) = (1+o(1))np$. Hence, by Theorem 3, we obtain a covering of $G(n,p)$ by $(1-o(1))np/2 + 28000((1+o(1))np - 2(1-o(1))np/2)/(\alpha/5)^4 = (1+o(1))np/2$ Hamilton cycles, finishing the proof of the theorem. $\qquad\square$

4. Concluding remarks and open questions

In this paper we verified that the size of a largest Hamilton cycle packing and the size of a smallest Hamilton cycle covering are asymptotically equal a.a.s. in the random graph $G(n,p)$, provided $p \geq n^{\alpha-1}$ for an arbitrary constant $\alpha > 0$. Our result calls for at least two natural directions of possible improvement.

First of all, the only explanation why the lower bound on the edge probability needs to be at least $n^{\alpha-1}$ and the corresponding expansion factor s needs to be at least $n^{\alpha/5}$ is that for lower values of p and s most of our technical arguments would break down. We think these bounds on p and s are only artifacts of our proof. Since the minimum and maximum degrees of the random graph have to be asymptotically equal in order for the minimum Hamilton covering and the maximum Hamilton packing to be asymptotically of the same size, we need to assume that $p = \omega(\ln n)/n$. We strongly believe though that Theorem 1 holds already whenever $p = \omega(\ln n)/n$.

Conjecture 14. *For any $p = \omega(\ln n)/n$ the random graph $G(n,p)$ admits a covering of its edges with at most $(1 + o(1))np/2$ Hamilton cycles a.a.s.*

Even though [14] provides the corresponding packing result for this range, we were not able to extend our techniques to prove e.g. the analog of Lemma 11 for $\alpha = o(1)$, not to mention for $\alpha = (\ln \ln n + \omega(1))/\ln n$.

Another direction in which Theorem 1 could be tightened is to make the statement exact instead of approximate. The trivial lower bound on the size of a Hamilton

covering in terms of the maximum degree is $\lceil \Delta(G) \rceil$. In what range of p will this be tight.

Question 15. *In what range of p does there exist a Hamilton covering of $G \sim G(n,p)$ of size $\lceil \Delta(G)/2 \rceil$ a.a.s.?*

Recall that the analogous precise statement in terms of the minimum degree *is* true for Hamilton packings [15], [18]. To have a positive answer for the question, we clearly need to be above the Hamiltonicity threshold, but it is plausible that the statement is true immediately after that.

The question of covering the edges of a graph by Hamilton cycles can also be considered in the *pseudorandom* setup. A graph G is called an (n,d,λ)-graph if it is d-regular on n vertices and the second largest absolute value of its eigenvalues is λ. The concept of (n,d,λ)-graphs is a common way to formally express pseudorandomness, as (n,d,λ)-graphs with $\lambda = o(d)$ behave in many ways as random graphs are expected to do. (See, e.g., [20] for a general discussion on pseudorandom graphs and (n,d,λ)-graphs.) Theorem 2 from [10] implies that for (n,d,λ)-graph with $d = \Theta(n)$ and $\lambda = o(d)$ there exists a Hamilton packing of size $d/2 - 3\sqrt{\lambda n} = d/2 - o(d)$. The expansion properties of (n,d,λ)-graphs are also well-known. For example, it is stated in Section 3.1 of [17], it is stated for example that every (n,d,λ)-graph is an $\left(\frac{(d-2\lambda)^2}{3\lambda^2}, \frac{\lambda^2 n}{d^2} \right)$-expander (translated into our notation). This implies that for every $\alpha > 0$, every (n,d,λ)-graph with $\lambda \leq \frac{d}{2n^{\alpha/2}}$ is an n^α-expander. Hence, Theorem 3 implies that any such (n,d,λ)-graph has a Hamilton covering of size $d/2 + o(d)$, which is of course asymptotically best possible. It would be interesting to decide whether a similar statement holds for sparser pseudorandom graphs, maybe as sparse as $d = n^\epsilon$, for arbitrarily small $\epsilon > 0$.

A Hamilton cycle is a particular spanning structure of the complete graph, which can be used to decompose its edges. A further group of problems related to our result is to determine the typical sizes of a largest packing and of a smallest covering of various other spanning structures in the random graph. Here often the corresponding decomposition result for the complete graph is not known or is just conjectured. Still asymptotic packing and covering results would be of interest, for example for trees of bounded maximum degree.

REFERENCES

[1] N. Alon, J. Spencer, **The probabilistic method**, 3rd Edition, Wiley, New York, 2008. 3

[2] B. Alspach, The wonderful Walecki construction, Bulletin of the Institute of Combinatorics and its Applications, **52** (2008), 7–20. 1

[3] B. Alspach, J.-C. Bermond, D. Sotteau, Decompositions into cycles I: Hamilton decompositions, Cycles and rays (Montreal, PQ, 1987), Kluwer Academic Publishers, Dordrecht, 1990, 9–18. 1

[4] S. Ben-Shimon, M. Krivelevich, B. Sudakov, On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs, SIAM Journal on Discrete Mathematics **25** (2011), 1176–1193. 1

[5] B. Bollobás, The evolution of sparse graphs, Graph theory and combinatorics, Academic Press, London (1984), 35–57. 1

[6] B. Bollobás, T. I. Fenner, A. M. Frieze, An algorithm for finding Hamilton paths and cycles in random graphs, Combinatorica **7** (1987), 327–341. 2

[7] B. Bollobás, A. Frieze, On matchings and Hamiltonian cycles in random graphs, Random graphs '83 (Poznán, 1983), North-Holland Mathematics Studies, 118, North-Holland, Amsterdam (1985), 23–46. 1

[8] P. Erdős, H. Hanani, On a limit theorem in combinatorial analysis, Publicationes Mathematicae Debrecen **10** (1963), 10–13. 1

[9] A. Frieze, M. Krivelevich, Hamilton cycles in random subgraphs of pseudo-random graphs, Discrete Mathematics **256** (2002), 137–150. 2

[10] A. Frieze, M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, Journal of Combinatorial Theory Series B **94** (2005), 159–172. 1, 4

[11] A. Frieze, M. Krivelevich, On two Hamilton cycle problems in random graphs, Israel Journal of Mathematics **166** (2008), 221–234. 1

[12] Z. Füredi, Matchings and covers in hypergraphs, Graphs and Combinatorics **4** (1988), 115–206. 1

[13] D. Hefetz, M. Krivelevich, T. Szabó, Hamilton cycles in highly connected and expanding graphs, Combinatorica **29** (2009), 547–568. 2, 2, 9, 2

[14] F. Knox, D. Kühn, D. Osthus, Approximate Hamilton decompositions of random graphs, Random Structures & Algorithms **40** (2012), 133–149. 1, 1.2, 3, 4

[15] F. Knox, D. Kühn, D. Osthus, Edge-disjoint Hamilton cycles in random graphs, submitted. 1, 4

[16] J. Komlós, E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, Discrete Mathematics **43** (1983), 55–63. 1, 2

[17] M. Krivelevich, On the number of Hamilton cycles in pseudo-random graphs, Electronic Journal of Combinatorics, Volume **19** (2012), publication P25. 4

[18] M. Krivelevich, W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, SIAM Journal on Discrete Mathematics, to appear. 1, 4

[19] M. Krivelevich, B. Sudakov, Sparse pseudo-random graphs are Hamiltonian, Journal of Graph Theory **42** (2003), 17–33. 2

[20] M. Krivelevich, B. Sudakov, Pseudo-random graphs, In: More sets, graphs and numbers, E. Györi, G. O. H. Katona, L. Lovász, Eds., Bolyai Society Mathematical Studies **15**, Springer Verlag, 2006, 199–262. 4

[21] L. Pósa, Hamiltonian circuits in random graphs, Discrete Mathematics **14** (1976), 359–364. 1.2, 2

[22] V. Rödl, On a packing and covering problem, European Journal of Combinatorics, **6** (1985), 69–78. 1

[23] R. M. Wilson, Decomposition of a complete graph into subraphs isomorphic to a given graph, Utilitas Mathematica (1976), 647–695.

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany
E-mail address: `glebov@mi.fu-berlin.de`

School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel
E-mail address: `krivelev@post.tau.ac.il`

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany
E-mail address: `szabo@mi.fu-berlin.de`