On the trace of random walks on random graphs

Alan Frieze ∗  Michael Krivelevich †  Peleg Michaeli ‡  Ron Peled §

August 29, 2015

Abstract

We study graph-theoretic properties of the trace of a random walk on a random graph. We show that for any \( \varepsilon > 0 \) there exists \( C > 1 \) such that the trace of the simple random walk of length \( (1 + \varepsilon)n \ln n \) on the random graph \( G \sim G(n, p) \) for \( p > C \ln n/n \) is, with high probability, Hamiltonian and \( \Theta(\ln n) \)-connected. In the special case \( p = 1 \) (i.e. when \( G = K_n \)), we show a hitting time result according to which, with high probability, exactly one step after the last vertex has been visited, the trace becomes Hamiltonian, and one step after the last vertex has been visited for the \( k \)’th time, the trace becomes \( 2k \)-connected.

1 Introduction

Since the seminal study of Erdős and Rényi [13], random graphs have become an important branch of modern combinatorics. It is an interesting and natural concept to study for its own sake, but it has also proven to have numerous applications both in combinatorics and in computer science. Indeed, random graphs have been a subject of intensive study during the last 50 years: thousands of papers and at least three books [5, 15, 18] are devoted to the subject. The term random graph is used to refer to several quite different “models”, each of which is essentially a distribution over all graphs on \( n \) labelled vertices. Perhaps the two most famous models are the classical models \( G(n, m) \), obtained by choosing \( m \) edges uniformly at random among the \( \binom{n}{2} \) possible edges, and \( G(n, p) \), obtained by selecting each edge independently with probability \( p \). Other models are discussed in [15].

In this paper, we study a different model of random graphs — the (random) graph formed by the trace of a simple random walk on a finite graph. Given a base graph and a starting vertex, we select a vertex uniformly at random from its neighbours and move to this neighbour, then independently

∗Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, USA. Email: alan@random.math.cmu.edu. Research supported in part by NSF Grant ccf0502793.
†School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. E-mail: krivelev@post.tau.ac.il. Research supported in part by a USA-Israel BSF Grant and by a grant from Israel Science Foundation.
‡School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. E-mail: peleg.michaeli@math.tau.ac.il.
§School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. E-mail: peledron@post.tau.ac.il. Research supported in part by an IRG grant and by a grant from Israel Science Foundation.
select a vertex uniformly at random from that vertex’s neighbours and move to it, and so on. The sequence of vertices this process yields is a simple random walk on that graph. The set of vertices in this sequence is called the range of the walk, and the set of edges traversed by this walk is called the trace of the walk. The literature on the topic of random walks is vast; however, most effort was put into answering questions about the range of the walk, or about the distribution of the position of the walk at a fixed time. Examples include estimating the cover time (the time it takes a walk to visit all vertices of the graph) and the mixing time of graphs (see Lovász [21] for a survey). On the other hand, to the best of our knowledge, there are almost no works addressing questions about the trace of the walk. We mention here that there are several papers studying the subgraph induced by vertices that are not visited by the random walk (see for example Cooper and Frieze [10], and Černý, Teixeira and Windisch [7]). We also mention that on infinite graphs, several properties of the trace have been studied (for an example see [3]).

Our study focuses on the case where the base graph $G$ is random and distributed as $G(n,p)$. We consider the graph $\Gamma$ on the same vertex set $([n] = \{1, \ldots, n\})$, whose edges are the edges traversed by the random walk on $G$. A natural graph-theoretic question about $\Gamma$ is whether it is connected. A basic requirement for that to happen is that the base graph is itself connected. It is a well-known result (see [12]) that in order to guarantee that $G$ is connected, we must take $p > (\ln n + \omega(1))/n$. Given that our base graph is indeed connected, for the trace to be connected, the walk must visit all vertices. An important result by Feige [14] states that for connected graphs on $n$ vertices, this happens on average after at least $(1 - o(1))n \ln n$ steps. Cooper and Frieze [9] later gave a precise estimation for the average cover time of (connected) random graphs, directly related to how large $p$ is, in comparison to the connectivity threshold. In fact, it can be derived from their proof that if $p = \Theta(\ln n/n)$ and the length of the walk is at most $n \ln n$, then the trace is typically not connected.

It is thus natural to execute a random walk of length $(1 + \varepsilon)n \ln n$ on a random graph which is above the connectivity threshold by at least a large constant factor (which may depend on $\varepsilon$), and to ask what other graph-theoretic properties the trace has. For example, is it highly connected? Is it Hamiltonian? The set of visited vertices does not reveal much information about the global structure of the graph, so the challenge here is to gain an understanding of that structure by keeping track of the traversed edges. What we essentially show is that the trace is typically Hamiltonian and $\Theta(\ln n)$-vertex-connected. Our method of proof will be to show that the set of traversed edges typically forms an expander.

In the boundary case where $p = 1$, i.e. when the base graph is $K_n$, we prove a much more precise result. As the trace becomes connected exactly when the last vertex has been visited, and at least one more step is required for that last visited vertex to have degree 2 in the trace, one cannot hope that the trace would contain a Hamilton cycle beforehand. It is reasonable to expect however that this degree requirement is in fact the bottleneck for a typical trace to be Hamiltonian, as is the case in other random graph models. In this paper, we show a hitting time result according to which, with high probability, one step after the walk connects the subgraph (that is, one time step after the cover time), the subgraph becomes Hamiltonian. This result implies that the bottleneck to Hamiltonicity of the trace lies indeed in the minimum degree, and in that sense, the result is similar in spirit to the results of Bollobás [4], and of Ajtai, Komlós and Szemerédi [1]. We also extend this result for $k$-cover-time vs. minimum degree $2k$ vs. $2k$-vertex-connectivity, obtaining a result similar in spirit to the result of Bollobás and Thomason [6].
1.1 Notation and terminology

Let $G$ be a (multi)graph on the vertex set $[n]$. For two vertex sets $A, B \subseteq [n]$, we let $E_G(A, B)$ be the set of edges having one endpoint in $A$ and the other in $B$. If $v \in [n]$ is a vertex, we may write $E_G(v, B)$ when we mean $E_G(\{v\}, B)$. We denote by $N_G(A)$ the external neighbourhood of $A$, i.e., the set of all vertices in $[n] \setminus A$ that have a neighbour in $A$. Again, we may write $N_G(v)$ when we mean $N_G(\{v\})$. We also write $N_G^+(A) = N_G(A) \cup A$. The degree of a vertex $v \in [n]$ is denoted by $d_G(v)$. The simplified graph of $G$ is the simple graph $G'$ obtained by replacing each multiedge with a single edge having the same endpoints, and removing all loops. The simple degree of a vertex is its degree in the simplified graph; it is denoted by $d'_G(v) = d_{G'}(v)$. We let $\delta(G)$ and $\Delta(G)$ be the minimum and maximum simple degree ($d'$) of $G$. Let the edge boundary of a vertex set $S$ be the set of edges of $G$ with exactly one endpoint in $S$, and denote it by $\partial_G S$. If $v, u$ are distinct vertices of a graph $G$, the distance from $v$ to $u$ is defined to be the minimum length of a path from $v$ to $u$ (or $\infty$ if there is no such path); it is denoted by $d_G(v, u)$. If $v$ is a vertex, the ball of radius $r$ around $v$ is the set of vertices of distance at most $r$ from $v$; it is denoted by $B_G(v, r)$. In symbols:

$$B_G(v, r) = \{u \in [n] \mid d_G(v, u) \leq r\}.$$  

We will sometimes omit the subscript $G$ in the above notations if the graph $G$ is clear from the context.

A simple random walk on a graph $G$ of length $t$, starting at a vertex $v$, is denoted $X_t^v(G)$. When the graph is clear from the context, we may omit it and simply write $X_t^v$. When the starting vertex is irrelevant, we may omit it as well, writing $X_t$. The trace of a simple random walk on a graph $G$ of length $t$, starting at a vertex $v$, is the subgraph of $G$ having the same vertex set as $G$, whose edges are all edges traversed by the walk (including loops), counted with multiplicity (so it is in general a multigraph). We denote it by $\Gamma_t^v(G)$, $\Gamma_t^v$ or $\Gamma_t$, depending on the context.

For a positive integer $n$ and a real $p \in [0, 1]$, we denote by $G(n, p)$ the probability space of all (simple) labelled graphs on the vertex set $[n]$ where the probability of each such a graph, $G = ([n], E)$, to be chosen is $p^{|E|}(1 - p)^{\binom{n}{2} - |E|}$. For a random variable/graph $X$ and a probability space $\mathcal{P}$ we write $X \sim \mathcal{P}$ to denote the fact that $X$ has the same distribution as $\mathcal{P}$.

For the sake of simplicity and clarity of presentation, we do not make a particular effort to optimize some of the constants obtained in our proofs. We also omit floor and ceiling signs whenever these are not crucial. Most of our results are asymptotic in nature and whenever necessary we assume that $n$ is sufficiently large. We say that an event holds with high probability (whp) if its probability tends to 1 as $n$ tends to infinity.

1.2 Our results

Our first theorem states that if $G \sim G(n, p)$ with $p$ above the connectivity threshold by at least some large constant factor, and the walk on $G$ is long enough to traverse the expected number of edges required to make a random graph connected, then its trace is with high probability Hamiltonian and highly connected.
Theorem 1. For every $\varepsilon > 0$ there exist $C = C(\varepsilon) > 0$ and $\beta = \beta(\varepsilon) > 0$ such that for every edge probability $p = p(n) \geq C \cdot \frac{\ln n}{n}$ and for every $v \in [n]$, a random graph $G \sim G(n, p)$ is whp such that for $L = (1 + \varepsilon)n \ln n$, the trace $\Gamma^v_L(G)$ of a simple random walk of length $L$ on $G$, starting at $v$, is whp Hamiltonian and $(\beta \ln n)$-vertex-connected.

Our proof strategy will be as follows. First we prove that whp $G \sim G(n, p)$ satisfies some pseudo-random properties. Then we show that whp the trace of a simple random walk on any given graph, which satisfies these pseudo-random properties, has “good” expansion properties. Namely, it has two properties, one ensures expansion of “small” sets, the other guarantees the existence of an edge between any two disjoint “large” sets.

In the next two theorems we address the case of a random walk $X$ executed on the complete graph $K_n$, and we assume that the walk starts at a uniformly chosen vertex. Denote the number of visits of the random walk $X$ into a vertex $v$ by time $t$ (including the starting vertex) by $\mu_t(v)$. For a natural $k$, denote by $\tau^k_C$ the $k$-cover time of the graph by the random walk; that is, the first time $t$ for which each vertex in $G$ has been visited at least $k$ times. In symbols,

$$\tau^k_C = \min \{ t \mid \forall v \in [n], \mu_t(v) \geq k \}.$$

When $k = 1$ we simply write $\tau_C$ and call it the cover time of the graph. The objective of the following theorems is to prove that the trivial minimal requirements for Hamiltonicity and $k$-vertex-connectivity are in fact the bottleneck for a typical trace to have these properties.

Theorem 2. Denote by $\tau_H$ the hitting time, for the simple random walk on $K_n$, of the property of being Hamiltonian. Then whp $\tau_H = \tau_C + 1$.

Corollary 3. Assume $n$ is even. Denote by $\tau_{PM}$ the hitting time, for the simple random walk on $K_n$, of the property of admitting a perfect matching. Then whp $\tau_{PM} = \tau_C$.

Theorem 4. For every $k \geq 1$, denote by $\tau^k_\delta$ the hitting time, for the simple random walk on $K_n$, of the property of being spanning with minimum simple degree $k$, and denote by $\tau^k_\kappa$ the hitting time of the property of being spanning $k$-vertex-connected. Then whp

$$\tau^k_C = \frac{\tau^k_\delta}{2^{k-1}} = \frac{\tau^k_\kappa}{2^{k-1}},$$

$$\tau^k_C + 1 = \frac{\tau^k_\delta}{2^k} = \frac{\tau^k_\kappa}{2^k}.$$

1.3 Organization

The organization of the paper is as follows. In the next section we present some auxiliary results, definitions and technical preliminaries. In Section 3 we explore important properties of the random walk on a pseudo-random graph. In Section 4 we prove the Hamiltonicity and vertex-connectivity results for the trace of the walk on $G(n, p)$. In Section 5 we prove the hitting time results of the walk on $K_n$. We end by concluding remarks and proposals for future work in Section 6.
2 Preliminaries

In this section we provide tools to be used by us in the succeeding sections. We start by stating two versions of known bounds on large deviations of random variables, due to Chernoff [8] and Hoeffding [17], whose proofs can be found, e.g., in Chapter 2 of [18].

**Theorem 2.1** ([18], Theorem 2.1). Let $X \sim \text{Bin}(n,p)$, $\mu = np$, $a \geq 0$ and $\varphi(x) = (1+x) \ln(1+x) - x$ (for $x \geq -1$, or $\infty$ otherwise). Then the following inequalities hold:

\[ P(X \leq \mu - a) \leq \exp \left( -\mu \varphi \left( \frac{-a}{\mu} \right) \right) \leq \exp \left( -\frac{a^2}{2\mu} \right), \quad (1) \]

\[ P(X \geq \mu + a) \leq \exp \left( -\mu \varphi \left( \frac{a}{\mu} \right) \right) \leq \exp \left( -\frac{a^2}{2 \left( (\mu + a)/3 \right) \mu} \right). \quad (2) \]

**Theorem 2.2** ([18], Theorem 2.10). Let $N \geq 0$, and let $0 \leq K, n \leq N$ be natural numbers. Let $X \sim \text{Hypergeometric}(N,K,n)$, $\mu = \mathbb{E}(X) = nKN^{-1}$. Then, inequalities (1) and (2) hold.

The following is a trivial yet useful bound:

**Claim 2.3.** Suppose $X \sim \text{Bin}(n,p)$. The following bound holds:

\[ P(X \geq k) \leq \binom{n}{k} p^k \leq \left( \frac{enp}{k} \right)^k. \]

**Proof.** Think of $X$ as $X = \sum_{i=1}^{n} X_i$, where $X_i$ are i.i.d. Bernoulli tests with probability $p$. For any set $A \subseteq [n]$ with $|A| = k$, let $E_A$ be the event “$X_i$ have succeeded for all $i \in A$”. Clearly, $P(E_A) = p^k$. If $X \geq k$, there exists $A \subseteq [n]$ for which $E_A$. Thus, the union bound gives

\[ P(X \geq k) \leq \binom{n}{k} P(E_A) = \binom{n}{k} p^k \leq \left( \frac{enp}{k} \right)^k. \]

\[ \square \]

2.1 $(R,c)$-expanders

Let us first define the type of expanders we intend to use.

**Definition 2.4.** For every $c > 0$ and every positive integer $R$ we say that a graph $G = (V,E)$ is an $(R,c)$-expander if every subset of vertices $U \subseteq V$ of cardinality $|U| \leq R$ satisfies $|N_G(U)| \geq c|U|$.

Next, we state some properties of $(R,c)$-expanders.

**Claim 2.5.** Let $G = (V,E)$ be an $(R,c)$-expander, and let $S \subseteq V$ of cardinality $k < c$. Denote the connected components of $G \setminus S$ by $S_1, \ldots, S_t$, so that $1 \leq |S_1| \leq \ldots \leq |S_t|$. It follows that $|S_1| > R$.

**Proof.** Assume otherwise. Since any external neighbour of a vertex from $S_1$ must be in $S$, we have that

\[ c > k = |S| \geq |N(S_1)| \geq c|S_1| \geq c, \]

which is a contradiction. \[ \square \]
The following is a slight improvement of Lemma 5.1 in [2].

**Lemma 2.6.** For every positive integer $k$, if $G = ([n], E)$ is an $(R, c)$-expander such that $c \geq k$ and $R(c + 1) \geq \frac{1}{2} (n + k)$, then $G$ is $k$-vertex-connected.

**Proof.** Assume otherwise; let $S \subseteq [n]$ with $|S| = k - 1$ be a disconnecting set of vertices. Denote the connected components of $G \setminus S$ by $S_1, \ldots, S_t$, so that $1 \leq |S_1| \leq \ldots \leq |S_t|$ and $t \geq 2$. It follows from Claim 2.5 that $|S_1| > R$.

Take $A_i \subseteq S_i$ for $i \in [2]$ with $|A_i| = R$. Since any common neighbour of $A_1$ and $A_2$ must lie in $S$, it follows that

\[
|S_1 \cup S_2 \cup N(S_1) \cup N(S_2)| \\
\geq |N^+(A_1) \cup N^+(A_2)| \\
= |N^+(A_1)| + |N^+(A_2)| - |N(A_1) \cap N(A_2)| \\
\geq 2R(c + 1) - |S| \geq n + 1,
\]

which is a contradiction. \qed

The reason we study $(R, c)$-expanders is the fact that they entail some pseudo-random properties, from which we will infer the properties that are considered in this paper, namely, being Hamiltonian, admitting a perfect matching, and being $k$-vertex-connected.

### 2.2 Properties of random graphs

In the following rather technical lemma we establish properties of random graphs to be used later to prove Theorem 1.

**Theorem 2.7.** Let $C$ be a large enough constant, and let $C \leq \alpha = \alpha(n) \leq \frac{n}{\ln n}$. Let $p = p(n) = \alpha \cdot \frac{\ln n}{n}$, and let $G \sim G(n, p)$. Then, whp,

(P1) $G$ is connected,

(P2) For every $v \in [n]$, $|d_G(v) - \alpha \ln n| \leq 2\sqrt{\alpha \ln n}$; in particular, $\frac{5\alpha}{6} \ln n \leq d_G(v) \leq \frac{4\alpha}{3} \ln n$,

(P3) For every non-empty set $S \subseteq [n]$ with at most $0.8n$ vertices, $|\partial S| > \frac{|S| |S^c| \alpha \ln n}{2n}$,

(P4) For every large enough constant $K > 0$ such that for every non-empty set $A \subseteq [n]$ with $|A| = a$, the following holds:

- If $\frac{n}{\alpha \ln n} \leq a \leq \frac{n}{\ln n}$ then

  $$\left| E \left( A, \left\{ b \in N_G(A) \mid |E(b, A)| \geq \frac{K a \alpha \ln n}{n} \right\} \right) \right| \leq \frac{a \alpha \ln n}{K};$$

- If $a < \frac{n}{\alpha \ln n}$ then

  $$|E \left( (A, \{ b \in N_G(A) \mid |E(b, A)| \geq K \}) \right) | \leq \frac{a \alpha \ln n}{\ln K}.$$
(P5) For every set \( A \) with \( |A| = \frac{n(\ln \ln n)^{1.5}}{\ln n} \) there exist at most \( |A|/2 \) vertices \( v \) not in \( A \) for which \( |E(v, A)| \leq \frac{\alpha(\ln \ln n)^{1.5}}{2} \),

and, if \( \alpha < \ln^2 n \),

(P6) For every \( v \in [n] \), \( 0 < r \leq \frac{\ln n}{15 \ln \ln n} \), for every \( w \in N_G(B(v, r)) \), \( |E(w, B(v, r))| \leq 5 \).

Property (P1) is well-known, so we omit the proof here.

Proof of (P2). We note that \( d_G(v) \sim \text{Bin} \left( n - 1, p \right) \). Denote \( \mu = \mathbb{E}(d_G(v)) = (n - 1)p \). Fix a vertex \( v \in [n] \). Using Chernoff bounds (Theorem 2.1) we have that

\[
\mathbb{P} \left( d_G(v) \leq \alpha \ln n - 2\sqrt{\alpha \ln n} \right) \leq \mathbb{P} \left( d_G(v) \leq \mu - \sqrt{3\alpha \ln n} \right) \leq \exp \left( -\frac{3\alpha \ln^2 n}{2\mu} \right) = o \left( n^{-1} \right)
\]

and that

\[
\mathbb{P} \left( d_G(v) \geq \alpha \ln n + 2\sqrt{\alpha \ln n} \right) \leq \mathbb{P} \left( d_G(v) \geq \mu + 2\sqrt{\alpha \ln n} \right) \leq \exp \left( -\frac{4\alpha \ln^2 n}{2\left( \mu + 2\sqrt{3\alpha \ln n} \right)} \right) \leq \exp \left( -\frac{4\alpha \ln^2 n}{3.5\alpha \ln n} \right) = o \left( n^{-1} \right).
\]

The union bound over all vertices \( v \in [n] \) yields the desired result. For large enough \( \alpha \), we also ensure that for every \( v \in [n] \),

\[
\frac{5\alpha}{6} \ln n \leq d_G(v) \leq \frac{4\alpha}{3} \ln n.
\]

Proof of (P3). Fix a set \( S \subseteq [n] \) with \( 1 \leq |S| = s \leq 0.8n \). We note that \( |\partial S| \sim \text{Bin} \left( s(n - s), p \right) \), thus by Theorem 2.1 we have that

\[
\mathbb{P} \left( |\partial S| \leq \frac{1}{2} s(n - s)p \right) \leq \exp \left( -\frac{1}{8} s(n - s)p \right).
\]

Let \( F_s \) be the event “\( \exists S, |S| = s \) such that \( |\partial S| \leq \frac{1}{2} s(n - s)p \)”. The union bound gives

\[
\mathbb{P} (F_s) \leq \binom{n}{s} \exp \left( -\frac{1}{8} s(n - s)p \right) \leq \exp \left( s \left( 1 + \ln n - \ln s - \frac{1}{8} (n - s)p \right) \right) \leq \exp \left( s \left( 1 + \ln n - \ln s - \frac{1}{40} \ln n \right) \right) = o \left( n^{-1} \right)
\]

for large enough \( \alpha \). Finally, let \( F \) be the event “\( \exists S, 1 \leq |S| \leq 0.8n \), for which \( F_s \) holds”. The union bound gives

\[
\mathbb{P} (F) \leq 0.8n \cdot \mathbb{P} (F_s) = o(1).
\]

\[\square\]
Proof of (P4). Fix $A$ with $|A| = a$, and suppose first that $\frac{n}{\ln n} \leq a \leq \frac{n}{\ln n}$. Let 

$$B_0 = \{ b \in N_G(A) \mid |E(b, A)| \geq Kap \},$$

and let $K$ be a constant to be determined later. For a vertex $b \notin A$, the random variable $|E(b, A)|$ is binomially distributed with $a$ trials and success probability $p$. Thus, using Claim 2.4 we have that

$$\mathbb{P} ( |E(b, A)| \geq Kap ) \leq \left( \frac{e}{K} \right)^{Kap} \leq e^{-K},$$

for large enough $K$. Thus $|B_0|$ is stochastically dominated by a binomial random variable with $n$ trials and success probability $e^{-K}$. It follows again by Claim 2.4 that $\mathbb{P} ( |B_0| > 3e^{-K}n ) \leq c^n$ for some $0 < c = c(K) < 1$. Since $a \leq n/\ln n$, $n(a) = o(e^{-n})$ Thus,

$$\mathbb{P} \left( \exists A, |A| = a : |E(A, B_0)| > \frac{anp}{K} \right) \leq \left( \binom{n}{a} \right) \left( c^n + \mathbb{P} \left( |E(A, B_0)| > \frac{anp}{K} \mid |B_0| < 3e^{-K}n \right) \right) \leq o \left( n^{-1} \right) + \left( \binom{n}{a} \right) \left( \frac{3ae^{-K}n}{anp/K} \right) p^{anp/K} \leq o \left( n^{-1} \right) + \left( \binom{n}{a} \right) \left( \frac{3ae^{-K}n}{anp/K} \right) p^{anp/K} \leq o \left( n^{-1} \right) + e^{4Kae^{-K} \cdot (9Ke^{-K})^{anp/K}} = o \left( n^{-1} \right),$$

for large enough $K$. Now suppose $a \leq \frac{n}{\alpha \ln n}$. Let 

$$B_0 = \{ b \in N_G(A) \mid |E(b, A)| \geq K \}. $$

From (P2) we know that the number of edges going out from $A$ is whp at most $4anp/3$. Given that, $|B_0| \leq 2anp/K$. Let $F_a$ be the event “there exists $A$, $|A| = a$, such that $|E(A, B_0)| > \frac{anp}{K}$”. Thus,

$$\mathbb{P} ( F_a \mid \Delta(G) \leq 4anp/3 ) \leq \binom{n}{a} \left( \frac{n}{2anp/K} \right)^{\frac{2a2n^2/K}{anp/\ln K}} p^{anp/\ln K} \leq \left( \frac{3K}{ap} \right)^{\frac{3n}{2anp/K}} \left( \frac{2eap \ln K}{K} \right)^{anp/\ln K} + o(1) = \left( \frac{3K}{ap} \right)^{\frac{3n}{2anp/K}} \left( \frac{2eap \ln K}{K} \right)^{1/\ln K}^{anp} = o \left( n^{-1} \right),$$

for large enough $K$. Taking the union bound over all cardinalities $1 \leq a \leq n/\ln n$ implies that the claim holds whp in both cases. 

Proof of (P5). Fix a set $A \subseteq [n]$. We say that a vertex $v \notin A$ is bad if $|E(v, A)| \leq \frac{1}{2} |A|p$. Since $|E(v, A)| \sim \text{Bin} (|A|, p)$, Chernoff bounds give that the probability that $v$ is bad with respect to $A$ is at most $\exp \left( -|A|p/8 \right)$.

Suppose $|A| = \Lambda = \frac{n(\ln n)^{1.5}}{\ln n}$, and let $U_A$ be the set of bad vertices with respect to $A$. We now show that $U_A$ is typically not too large. To this end, note that $|U_A|$ is stochastically dominated by a
binomial random variable with \( n \) trials and success probability \( \exp (-\Lambda p/8) \). Thus, using Chernoff bounds again, we have that

\[
P(|U_A| \geq \Lambda/2) \leq \binom{n}{\Lambda/2} \exp (-\Lambda^2 p/16).
\]

The probability that there exists \( A \) of cardinality \( \Lambda \) whose set of bad vertices is of cardinality at least \( \Lambda/2 \) is thus at most

\[
P(\exists A : |A| = \Lambda, |U_A| \geq \Lambda/2) \leq \binom{n}{\Lambda} \binom{n}{\Lambda/2} \exp (-\Lambda^2 p/16)
\]

\[
\leq \left( \frac{en}{\Lambda} \right)^{2\Lambda} \exp (-\Lambda^2 p/16)
\]

\[
\leq \exp \left( 3\Lambda \ln (n/\Lambda) - \frac{\Lambda^2}{2} \right)
\]

\[
\leq \exp \left( -\frac{n}{\ln n} (\ln n)^2.9 \right) = o(1).
\]

Noting that \( \Lambda p = \alpha (\ln \ln n)^{1.5} \), the claim follows.

\[
\square
\]

**Proof of (P6).** Assume \( \alpha < \ln^2 n \). Let \( \lambda = \frac{\ln n}{15 \ln \ln n} \). For \( v \in [n] \), \( 0 < r \leq \lambda \), let \( A(v, r) \) be the event “\( \exists w \in N(B(v, r)) \) for which \( |E(w, B(v, r))| > 5 \)”. Fix \( v, r \), and expose a BFS tree \( T \), rooted at \( v \), of depth \( r \). We note that it follows from (P2) that the number of leaves of \( T \) is at most \( (C \ln^3 n)^r \) for some constant \( C \). Now fix a vertex \( w \notin B(v, r) \). For each leaf \( \ell \in T \), the probability that \( w \) is a neighbour of \( \ell \) is \( p \), independently of all other leaves. In addition, any vertex in \( B(v, r) \cap N(w) \) is a leaf in \( T \). It follows that \( |E(w, B(v, r))| \) is stochastically dominated by a binomial random variable with \( (C \ln^3 n)^r \) trials and success probability \( p \). It follows by Claim 2.3 that

\[
P(|E(w, B(v, r))| > 5) \leq \left( \frac{e (C \ln^3 n)^{r+1}}{5n} \right)^5 = o(n^{-3})
\]

thus for every \( 0 < r \leq \lambda \)

\[
P(A(v, r)) = o\left(n^{-2}\right).
\]

For \( v \in [n] \), let \( A(v) \) be the event “\( \exists 0 < r \leq \lambda \) for which \( A(v, r) \) holds”. The union bound gives

\[
P(A(v)) \leq \lambda \cdot \max_r P(A(v, r)) = o\left(n^{-1.5}\right),
\]

and

\[
P(\exists v : A(v)) \leq n \cdot P(A(v)) = o\left(n^{-0.5}\right),
\]

and that completes the proof.

\[
\square
\]

For \( \alpha = \alpha(n) > 0 \), a graph for which (P1),…,(P5) hold (and (P6) as well, if \( \alpha < \ln^2 n \)) will be called \( \alpha \)-pseudo-random.
2.3 Properties of random walks

Throughout this section, \( G \) is a graph with vertex set \([n]\), having properties (P1), (P2) and (P3) for some \( \alpha > 0 \), and \( X \) is a \( \frac{1}{2} \)-lazy simple random walk on \( G \), starting at some arbitrary vertex \( v_0 \).

By \( \frac{1}{2} \)-lazy we mean that it stays put with probability \( \frac{1}{2} \) at each time step, and moves to a random neighbour otherwise. Our purpose in this section is to show that \( X \) “mixes well”, in a sense that will be further clarified below. To this end, we shall need some preliminary definitions and notations.

The transition rate of \( X \) from \( u \) to \( v \) is the probability

\[
p_{uv} = \mathbb{P}(X_{t+1} = v \mid X_t = u) = \mathbb{P}(X_1 = v \mid X_0 = u).
\]

For \( k \in \mathbb{N} \) we similarly denote

\[
p_{uv}^{(k)} = \mathbb{P}(X_{t+k} = v \mid X_t = u) = \mathbb{P}(X_k = v \mid X_0 = u)
\]

We note that the stationary distribution of \( X \) is given by

\[
\pi_v = \frac{d(v)}{\sum_{u \in [n]} d(u)} = \frac{d(v)}{2 |E|},
\]

and for every subset \( S \subseteq [n] \),

\[
\pi_S = \sum_{v \in S} \pi_v.
\]

The total variation distance between \( X_t \) and the stationary distribution is

\[
d_{TV}(X_t, \pi) = \sup_{S \subseteq [n]} |\mathbb{P}(X_t \in S) - \pi_S|,
\]

and as is well-known, we have that

\[
d_{TV}(X_t, \pi) = \frac{1}{2} \sum_{v \in [n]} |\mathbb{P}(X_t = v) - \pi_v|.
\]

Now, let \((Y_t)_{t \geq 0}\) be the stationary walk on \( G \); that is, the \( \frac{1}{2} \)-lazy simple random walk for which for every \( v \in [n] \), \( \mathbb{P}(Y_0 = v) = \pi_v \). We note for later use that there exists a standard coupling of \( X, Y \) under which for every \( t \),

\[
\mathbb{P}(\exists s \geq t \mid X_s \neq Y_s) \leq d_{TV}(X_t, \pi).
\]

Our goal is therefore to find not too large \( t \)'s for which that total variation distance is very small. That is, we wish to bound the \( \varepsilon \)-mixing time of \( X \), which is given by

\[
\tau(\varepsilon) = \min \{ t \geq 0 \mid \forall s \geq t, \ d_{TV}(X_s, \pi) < \varepsilon \}.
\]

A theorem of Jerrum and Sinclair ([19]) will imply that the \( \varepsilon \)-mixing time of \( X \) is indeed small. Their bound uses the notion of conductance: the conductance of a cut \((S, S^c)\), with respect to \( X \), is defined as

\[
\varphi_X(S) = \frac{\sum_{v \in S, w \in S^c} \pi_v P_{vw}}{\min(\pi_S, \pi_{S^c})}
\]
which can be equivalently written in our case as
\[
\varphi_X(S) = \frac{2 |\partial S|}{\min \left( \sum_{v \in S} d(v), \sum_{w \in S^c} d(w) \right)}.
\]

The conductance of \( G \) is defined as
\[
\Phi_X(G) = \min_{S \subseteq [n]} \varphi_X(S).
\]

**Claim 2.8.** Let \( \pi_{\text{min}} = \min_v \pi_v \). For every \( \varepsilon > 0 \),
\[
\tau(\varepsilon) \leq \frac{2 \Phi_X(G)}{\Phi_X(G)} \left( \ln \left( \frac{1}{\pi_{\text{min}}} \right) + \ln \left( \frac{1}{2\varepsilon} \right) \right).
\]

**Proof.** Let
\[
\tau'(\varepsilon) = \min \left\{ t \geq 0 \mid \forall s \geq t, \ u, v \in [n], \ \left| \frac{p_{uv}^{(s)} - \pi_v}{\pi_v} \right| < \varepsilon \right\}
\]
be the \( \varepsilon \)-uniform mixing time of \( X \). Corollary 2.3 in [19] implies that
\[
\tau'(\varepsilon) \leq 2 \Phi_X(G) \left( \ln \left( \frac{1}{\pi_{\text{min}}} \right) + \ln \left( \frac{1}{\varepsilon} \right) \right).
\]
Let \( t = \tau'(2\varepsilon) \); thus, for all \( s \geq t \), \( u, v \in [n] \), \( \left| p_{uv}^{(s)} - \pi_v \right| < 2\varepsilon \pi_v \). Fix \( s \geq t \). We have that
\[
d_{TV}(X_s, \pi) = \frac{1}{2} \sum_{v \in [n]} |\mathbb{P}(X_s = v) - \pi_v| \leq \frac{2\varepsilon}{2} \sum_{v \in [n]} \pi_v = \varepsilon,
\]
thus \( \tau(\varepsilon) \leq t = \tau'(2\varepsilon) \) and the claim follows. \( \square \)

**Corollary 2.9.** For \( \varepsilon > 0 \), \( \tau(\varepsilon) \leq 200 \ln(n/\varepsilon) \).

**Proof.** We note that due to (P2), for every \( v \in [n] \),
\[
\pi_v \geq \frac{5}{8n},
\]
thus for every \( S \subseteq [n] \) with \( 0 < \pi_S \leq 1/2 \) we have that
\[
\frac{1}{2} \geq \pi_S \geq |S| \cdot \frac{5}{8n},
\]
hence \( 0 < |S| \leq \frac{4}{5} n \). Thus, according to (P3),
\[
\Phi_X(G) = \min_{\substack{S \subseteq [n] \ 0 < \pi_S \leq 1/2}} \varphi(S)
\]
\[
\geq \min_{\substack{S \subseteq [n] \ 0 < |S| \leq 4n/5}} \frac{2|\partial S|}{\sum_{v \in S} d(v)}
\]
\[
\geq \min_{\substack{S \subseteq [n] \ 0 < |S| \leq 4n/5}} \frac{2 \cdot |S||S^c|\alpha \ln n}{2n |S| \cdot \frac{4}{3} \alpha \ln n} \geq \frac{1}{10}.
\]
Plugging this into Theorem 2.8 we have
\[ \tau(\varepsilon) \leq 200 \left( \ln \left( \frac{8n}{5} \right) + \ln \left( \frac{1}{2\varepsilon} \right) \right) \leq 200 \ln \left( \frac{n}{\varepsilon} \right). \]

The following is an immediate corollary of the above:

**Corollary 2.10.** Let \( F_t = \sigma(X_0, \ldots, X_t) \) and let \( b = 400 \ln n \). Conditioned on \( F_t \), there exists a coupling of \((X_{t+b+s})_{s \geq 0}\) and \((Y_s)_{s \geq 0}\) under which
\[ P \left( \exists s' \geq s \mid X_{t+b+s'} \neq Y_{s'} \right) \leq \frac{1}{n}. \]

### 3 Walking on a pseudo-random graph

In order to prove Theorem 1, we will prove that the trace \( \Gamma = \Gamma^v_L(G) \) is whp a “good” expander, in the sense that it satisfies the following two properties:

- **(E1)** There exists \( \beta > 0 \) such that every set \( A \subseteq [n] \) of cardinality \( |A| \leq \frac{n}{\ln n} \) satisfies \( |N_\Gamma(A)| \geq |A| \cdot \beta \ln n \);
- **(E2)** There is an edge of \( \Gamma \) between every pair of disjoint subsets \( A, B \subseteq [n] \) satisfying \( |A|, |B| \geq \frac{n(\ln \ln n)^{1.5}}{\ln n} \).

**Theorem 3.1.** For every \( \varepsilon > 0 \) there exist \( C = C(\varepsilon) > 0 \) and \( \beta = \beta(\varepsilon) > 0 \) such that for every edge probability \( p = p(n) \geq C \cdot \frac{\ln n}{n} \) and for every \( v \in [n] \), a random graph \( G \sim G(n,p) \) is whp such that for \( L = (1 + \varepsilon)n \ln n \), the trace \( \Gamma^v_L(G) \) of a simple random walk of length \( L \) on \( G \), starting at \( v \), has whp the properties (E1) and (E2).

It will be convenient the consider a slight variation of this theorem, in which the random walk is lazy:

**Theorem 3.2.** For every \( \varepsilon > 0 \) there exists \( C = C(\varepsilon) \) such that if \( \alpha = \alpha(n) \geq C \) and \( G \) is a \( \alpha \)-pseudo-random graph on the vertex set \([n]\), \( v_0 \in [n] \) and \( L_2 = \lfloor (2 + \varepsilon)n \ln n \rfloor \), then the trace \( \Gamma = \Gamma^v_{L_2}(G) \) of a \( \frac{1}{2} \)-lazy random walk of length \( L_2 \) on \( G \), starting at \( v_0 \), has whp the properties (E1) and (E2).

Before proving this theorem, we show that Theorem 3.1 is a simple consequence of it.

**Proof of Theorem 3.1.** Let \( \varepsilon > 0 \) and let \( L = \lfloor (1 + \varepsilon)n \ln n \rfloor \), \( L_2 = \lfloor (2 + \varepsilon)n \ln n \rfloor \). Choose \( C \) large enough so that by Theorem 2.7 \( G \) is whp \( \alpha \)-pseudo-random, and for which Theorem 3.2 holds. Let \( X \) be the \( \frac{1}{2} \)-lazy random walk of length \( L_2 \) on \( G \), starting at \( v_0 \), and define
\[ R = |\{0 < t \leq L_2 \mid X_t \neq X_{t-1}\}|. \]
We note that by standard deviation results for binomial random variables, \( \mathbb{P}(R > L) \) tends to 0 as \( n \) grows to infinity. Denote by \( \Gamma_{L_2}^\ell \) the trace of that walk, and let \( P \) be a monotone graph property which \( \Gamma_{L_2}^\ell \) satisfies \textbf{whp}. Given \( R \), the trace \( \Gamma_{L_2}^\ell \) has the same distribution as the trace of the non-lazy walk \( \Gamma_R \). Thus:

\[
\mathbb{P}(\Gamma_L \in P) \geq \mathbb{P}(\Gamma_{L_2}^\ell \in P, R \leq L) = 1 - o(1).
\]

As (E1) and (E2) are both monotone, and since \( G \) is \textbf{whp} \( \alpha \)-pseudo-random, the claim holds using Theorem 3.2.

Thus, in what follows, \( G \) is a \( \alpha \)-pseudo-random graph on the vertex set \([n]\), \( X \) is a \( \frac{1}{2} \)-lazy simple random walk on \( G \) starting at some fixed vertex \( v_0 \), and \( Y \) is the \( \frac{1}{2} \)-lazy simple random walk on \( G \), starting at random vertex sampled from the stationary distribution of \( X \).

The rest of this section is organised as follows. In the first subsection we show that \textbf{whp}, every vertex is visited at least a logarithmic number of times. In the second and third subsections we use that fact to conclude that “small” vertex sets typically expand “well”, and that large vertex sets are typically connected, by that proving that the trace \textbf{whp} satisfies (E1) and (E2).

### 3.1 Number of visits

Define

\[
\nu(v) = |\{0 < t \leq L_2 \mid X_t = v, X_{t+1} \neq v\}|, \quad v \in [n].
\]

**Theorem 3.3.** There exists \( \rho > 0 \) such that \textbf{whp}, for every \( v \in [n] \), \( \nu(v) \geq \rho \ln n \).

In order to prove this theorem, we first introduce a number of definitions and lemmas. Recall that a \textit{supermartingale} is a sequence \( M(0), M(1), \ldots \) of random variables such that each conditional expectation \( \mathbb{E}(M(t+1) \mid M(0), \ldots, M(t)) \) is at most \( M(t) \). Given such a sequence, a \textit{stopping rule} is a function from finite histories of the sequence into \{0, 1\}, and a \textit{stopping time} is the minimum time in which some stopping rule is satisfied (that is, equals 1). For two integers \( s, t \), let \( s \wedge t = \min\{s, t\} \).

Let \( \lambda = \frac{\ln n}{15 \ln \ln n} \), and for every \( v \in [n] \) let \( F_t^v \) be the event “\( Y_t = v \) or \( d_G(Y_t, v) > \lambda \)” (recall that for two vertices \( u, v \), \( d_G(u, v) \) denotes the distance from \( u \) to \( v \) in \( G \)).

**Lemma 3.4.** Suppose \( \alpha < \ln^2 n \). For \( v \in [n] \), define the process

\[
\mathcal{M}^v(t) = \left( \frac{10}{\alpha \ln n} \right)^{d_G(Y_t, v)}.
\]

Let \( S = \min\{t \geq 0 \mid F_t^v \} \) be a stopping time; then \( \mathcal{M}^v(t \wedge S) \) is a supermartingale. In particular, for every \( v_0 \in [n] \),

\[
\mathbb{P}(Y_S = v \mid Y_0 = v_0) \leq \left( \frac{10}{\alpha \ln n} \right)^{d_G(v_0, v)}.
\]

**Proof.** For a vertex \( w \in [n] \), denote

\[
p_{\leftarrow}(w) = \mathbb{P}(d_G(Y_1, v) < d_G(Y_0, v) \mid Y_0 = w),
\]

\[
p_{\rightarrow}(w) = \mathbb{P}(d_G(Y_1, v) > d_G(Y_0, v) \mid Y_0 = w).
\]
We note that for \( y \leq x \leq 1, \frac{y}{x} + x - y \leq 1 \). Thus, for \( a, b \) \( > 0 \) for which \( \frac{p_{\leftarrow}(w)}{p_{\rightarrow}(w)} \leq \frac{a}{b} \leq 1 \),

\[
E \left( \left( \frac{a}{b} \right)^{d_G(Y_1,v)} \mid Y_0 = w \right) - \left( \frac{a}{b} \right)^{d_G(w,v)}
\]

\[
= \left( \frac{a}{b} \right)^{d_G(w,v)} \left( (p_{\leftarrow}(w) \frac{b}{a} + p_{\rightarrow}(w) \frac{a}{b}) + (1 - p_{\leftarrow}(w) - p_{\rightarrow}(w)) \right) - 1
\]

\[
= \left( \frac{a}{b} \right)^{d_G(w,v)} p_{\rightarrow}(w) \left( \frac{bp_{\leftarrow}(w)}{ap_{\rightarrow}(w)} + \frac{a}{b} - \frac{p_{\leftarrow}(w)}{p_{\rightarrow}(w)} - 1 \right) \leq 0.
\]

Let \( w \) be such that \( 0 < d_G(v, w) \leq \lambda \). Since \( \alpha < \ln^2 n \), \( G \) satisfies (P6). Considering that and (P2) we have that

\[
a := \frac{5}{2(\alpha - 2\sqrt{\alpha}) \ln n} \geq \frac{5}{2d_G(w)} \geq p_{\leftarrow}(w),
\]

\[
b := \frac{\alpha \ln n}{4(\alpha - 2\sqrt{\alpha}) \ln n} \leq \frac{1}{2} - \frac{5}{2(\alpha - 2\sqrt{\alpha}) \ln n} \leq \frac{d(w) - 5}{2d(w)} \leq p_{\rightarrow}(w),
\]

and as \( \frac{p_{\leftarrow}(w)}{p_{\rightarrow}(w)} \leq \frac{a}{b} \leq 1 \) and \( \frac{a}{b} = \frac{10}{\alpha \ln n} \), we have that \( M^v(t \wedge S) \) is a supermartingale. In addition, for every \( t \),

\[
\left( \frac{10}{\alpha \ln n} \right)^{d_G(Y_0,v)} = M^v(0)
\]

\[
\geq E \left( M^v(t \wedge S) \mid Y_0 \right)
\]

\[
= \sum_{w \in [n]} \left( \frac{10}{\alpha \ln n} \right)^{d_G(w,v)} \cdot P \left( Y_{t \wedge S} = w \mid Y_0 \right)
\]

\[
\geq P \left( Y_{t \wedge S} = v \mid Y_0 \right).
\]

As this is true for every \( t \geq 0 \), and since \( S \) is \textbf{whp} finite, it follows that for every \( v_0 \in [n] \),

\[
P \left( Y_S = v \mid Y_0 = v_0 \right) \leq \left( \frac{10}{\alpha \ln n} \right)^{d_G(v_0,v)}
\]

\textbf{whp}.

Let

\[
T = \ln^2 n.
\]

For a walk \( W \) on \( G \), let \( I_W(v, t) \) be the event “\( W_t = v \) and \( W_{t+1} \neq v \)”. Our next goal is to estimate the probability that \( I_Y(v, t) \) occurs for some \( 1 \leq t < T \), given that \( I_Y(v, 0) \) occurred.

\textbf{Lemma 3.5}. For every vertex \( v \in [n] \) we have that

\[
P \left( \bigcup_{1 \leq t < T} I_Y(v, t) \mid I_Y(v, 0) \right) \leq \ln^{-1/2} n.
\]
Proof. Fix $v \in [n]$. Define the following sequence of stopping times: $U_0 = 0$, and for $i \geq 1$,

$$U_i = \min \{ t > U_{i-1} \mid F^v_t \}.$$ 

Then,

$$\mathbb{P} \left( \bigcup_{1 \leq i < T} I_Y(v, t) \mid I_Y(v, 0) \right) \leq \mathbb{P} (Y_{U_1} = v \mid I_Y(v, 0)) + \sum_{i=2}^{T-1} \mathbb{P} (Y_{U_i} = v \mid Y_{U_{i-1}} \neq v, I_Y(v, 0)).$$

Now, if $\alpha < \ln^2 n$, Lemma 3.4 and the Markov property imply that

$$\mathbb{P} (Y_{U_1} = v \mid I_Y(v, 0)) \leq \frac{10}{\alpha \ln n}, \quad \mathbb{P} (Y_{U_i} = v \mid Y_{U_{i-1}} \neq v, I_Y(v, 0)) \leq \left( \frac{10}{\alpha \ln n} \right)^\lambda \quad (i \geq 2),$$

so

$$\mathbb{P} \left( \bigcup_{1 \leq i < T} I_Y(v, t) \mid I_Y(v, 0) \right) \leq \frac{10}{\alpha \ln n} + T \cdot \left( \frac{10}{\alpha \ln n} \right)^\lambda = \frac{10}{\alpha \ln n} + \ln^2 n \cdot o \left( n^{-1/20} \right) \leq \frac{20}{\alpha \ln n} \leq \ln^{-1/2} n.$$ 

Now consider the case $\alpha \geq \ln^2 n$. The number of exits from $v$ at times $1, \ldots, T-1$ is at most the number of enters to $v$ at times $1, \ldots, T-2$ plus 1. Recalling (P2), at any time $i \in [T-3]$, the probability to enter $v$ at time $i+1$ is at most $1/d_G(X_i) \leq \ln^{-2.5} n$. Thus, the number of exits from $v$ is stochastically dominated by a binomial random variable with $T-1$ trials and success probability $\ln^{-2.5} n$. Thus,

$$\mathbb{P} \left( \bigcup_{1 \leq i < T} I_Y(v, t) \mid I_Y(v, 0) \right) \leq \mathbb{P} \left( \sum_{i=1}^{T-1} 1_{I_Y(v, t)} \geq 1 \right) \leq (T-1) \ln^{-2.5} n \leq \ln^{-1/2} n,$$

and the claim follows.

For a walk $W$ on $G$ and a vertex $v \in [n]$, let $M_W(v) = \sum_{i=0}^{T-1} 1_{I_W(v, t)}$.

**Lemma 3.6.** For every vertex $v \in [n]$, $\mathbb{P} (M_Y(v) \geq 1) \geq \frac{T}{2n} (1 - 6\alpha^{-1/2})$.

**Proof.** Fix $v \in [n]$. It follows from (P2) that

$$\mathbb{E} (M_Y(v)) = \frac{T \cdot \pi_v}{2} \geq \frac{T \cdot d_v}{4|E(G)|} \geq \frac{T \cdot (1 - 2\alpha^{-1/2})}{2n \cdot (1 + 2\alpha^{-1/2})} \geq \frac{T}{2n} \cdot \left( 1 - 5\alpha^{-1/2} \right)$$

and the claim follows. □
(for large enough $\alpha$). Thus by Lemma 3.5,

$$
\frac{T}{2n} \left(1 - 5\alpha^{-1/2}\right) \leq \mathbb{E}(M_Y(v)) \\
= \sum_{i=1}^{\infty} i \mathbb{P}(M_Y(v) = i) \\
\leq \mathbb{P}(M_Y(v) = 1) \sum_{i=1}^{\infty} i \left(\ln^{-1/2} n\right)^{i-1} \\
= \mathbb{P}(M_Y(v) = 1) \left(1 - \left(\ln^{-1/2} n\right)^{-2}\right).
$$

It follows that

$$
\mathbb{P}(M_Y(v) \geq 1) \geq \mathbb{P}(M_Y(v) = 1) \geq \frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right).
$$

\[ \square \]

**Corollary 3.7.** Let $t \geq 0$ and let $b = b(n) = 400 \ln n$. Conditioned on $F_t$, for every vertex $v \in [n], \mathbb{P}\left(M_{X^{(k)}_{s+b}+s}(v) \geq 1\right) \geq \frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right) - \frac{1}{n}.

**Proof.** According to Lemma 3.6 and Corollary 2.10,

$$
\mathbb{P}\left(M_{X^{(k)}_{s+b}+s}(v) \geq 1\right) \geq \mathbb{P}\left(M_{Y_{s+b}+s}(v) \geq 1\right) - \frac{1}{n} \geq \frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right) - \frac{1}{n}.
$$

\[ \square \]

**Proof of Theorem 3.3.** Consider dividing the $L_2$ steps of the walk $X$ into segments of length $T + 1$ with “buffers” of length $b = b(n) = 400 \ln n$ between them. Formally, the $k$'th segment is the walk

$$
(X_{s+b}^{(k)})_{s=0}^T = (X_{(k-1)(T+1+b)+b+s})_{s=0}^T.
$$

It follows from Corollary 3.7 that independently between the segments, for a given $v,

$$
\mathbb{P}(M_{X^{(k)}_{s+b}+s}(v) \geq 1) \geq \frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right) - \frac{1}{n}.
$$

Thus, $\nu(v)$ stochastically dominates a binomial random variable with $\lfloor L_2/(T + 1 + b) \rfloor$ trials and success probability $\frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right) - \frac{1}{n}$. Let $\mu = \lfloor L_2/(T + 1 + b) \rfloor \cdot \left(\frac{T}{2n} \left(1 - 6\alpha^{-1/2}\right) - \frac{1}{n}\right)$. We note that

$$
\left(1 + \frac{\varepsilon}{2} - \varepsilon'\right) \ln n \leq \mu \leq \left(1 + \frac{\varepsilon}{2}\right) \ln n
$$

where $\varepsilon'$ can be chosen to be arbitrarily small, given that $\alpha$ is large enough. Thus, Chernoff bound (Theorem 2.1) gives

$$
\mathbb{P}(\nu(v) < \rho \ln n) \leq \exp\left(-\mu \varphi\left(-\frac{\mu - \rho \ln n}{\mu}\right)\right) \\
\leq \exp\left(-\left(1 + \frac{\varepsilon}{2} - \varepsilon'\right) \ln n \cdot \varphi\left(-\frac{1 + \frac{\varepsilon}{2} - \varepsilon' - \rho}{1 + \frac{\varepsilon}{2}}\right)\right)
$$

16
and as $\varphi(x) \to 1$ with $x \to -1$, for small enough $\rho, \varepsilon'$ this is at most $n^{-(1+\varepsilon/3)}$. The union bound gives that for small enough $\rho > 0$,

$$\mathbb{P}(\exists v \in [n] : \nu(v) < \rho \ln n) \leq n \cdot \mathbb{P}(\nu(v) < \rho \ln n) = o(1),$$

and that concludes the proof. \qed

### 3.2 Expansion of small sets

**Theorem 3.8.** There exists $\beta > 0$ such that whp for every set $A \subseteq [n]$ with $|A| = a \leq n/\ln n$, it holds that $|N_1(A)| \geq \beta \cdot a \ln n$.

**Proof.** Let $K > 0$ be a constant guaranteed by (P4). Suppose first that $A$ is such that $a \geq \frac{n}{\alpha \ln n}$. Let

$$B_0 = \left\{ b \in N_G(A) \mid |E(b, A)| \geq \frac{K\alpha \ln n}{n} \right\},$$

and let

$$A_0 = \left\{ v \in A \mid |E(v, B_0)| \geq \frac{2\alpha \ln n}{K} \right\}.$$

According to (P4), $|A_0| \leq \frac{a}{2}$. Let $A_1 = A \setminus A_0$. For a vertex $v \in A_1$, let $\gamma(v)$ count the number of moves the walk has made from $v$ to $B_0$ in the first $\rho \ln n$ exits from $v$. Let $R$ be the event “$\forall v \in [n], \nu(v) \geq \rho \ln n$”, and recall that $\mathbb{P}(R) = 1 - o(1)$ by Theorem 3.3. Thus, by Claim 2.3,

$$\mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2}\right) \leq \mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2} \mid R\right) (1 + o(1)) \leq \left(\frac{\rho \ln n}{\rho \ln n/2}\right) \left(\frac{|E(v, B_0)|}{d(v)}\right)^{\rho \ln n/2} (1 + o(1)) \leq \left(\frac{5}{K}\right)^{\rho \ln n/2}.$$

Let $A_2 \subseteq A_1$ be the set of vertices $v$ with $\gamma(v) \geq \rho \ln n/2$. We have that

$$\mathbb{P}\left(|A_2| \geq \frac{a}{4}\right) \leq \left(\frac{a}{\alpha}\right)^{a/4} \mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2}\right)^{a/4} \leq \left(\frac{6}{K}\right)^{a \rho \ln n/8}.$$

Note that for large enough $K$, $\binom{n}{a} \left(\frac{6}{K}\right)^{a \rho \ln n/8} = o\left(n^{-1}\right)$. Set $A_3 = A_1 \setminus A_2$. Fix $B_3 \subseteq N_G(A) \setminus B_0$ with $|B_3| \leq a\beta \ln n$. For $v \in A_3$, let $p_v$ be the probability that a walk which exits $v$ will land in $B_3$. We have that for every $v \in A_3$,

$$p_v \leq \frac{|E(v, B_3)|}{\frac{\rho}{\alpha} \ln n}.$$

Assuming that $|A_3| \geq \frac{a}{4}$,

$$\frac{1}{|A_3|} \sum_{v \in A_3} p_v \leq \frac{5}{a} \cdot \frac{|E(A_3, B_3)|}{\alpha \ln n} \leq \frac{5K\alpha \beta \ln n}{n},$$

17
and by the AM/GM inequality, making sure $\beta = \beta(K)$ is small enough, we have that

$$\prod_{v \in A_3} p_v \leq \left(\frac{5K\alpha \beta \ln n}{n}\right)^{a/4}.$$ 

Thus, taking the union bound,

$$\mathbb{P}(\exists A, |A| = a : |N_{G}(A)| \leq a\beta \ln n) \leq \sum_{A, |A| = a} \mathbb{P}(|N_{G}(A)| \leq a\beta \ln n)|$$

$$\leq \left(\binom{n}{a}\right) \mathbb{P}\left(|A_2| \geq \frac{a}{4}\right)$$

$$+ \left(\binom{n}{a}\right) \mathbb{P}\left(\exists B, |B| = a\beta \ln n, N_{G}(A) \subseteq B \mid |A_3| \geq \frac{a}{4}\right)$$

$$\leq \left(\binom{n}{a}\right) \left(\frac{6}{K}\right)^{a\beta \ln n/8} + \left(\binom{n}{a}\right) \left(\frac{n}{a\beta \ln n}\right) \left(\prod_{v \in A_3} p_v\right)^{\rho \ln n/2}$$

$$\leq o(n^{-1}) + \left(\frac{3n}{a\beta \ln n}\right)^{a\beta \ln n/8} \left(\frac{5K\alpha \beta \ln n}{n}\right)^{a\rho \ln n/8}$$

$$= o(n^{-1}) + \left(\frac{3n}{a\beta \ln n}\right)^{\beta} \left(\frac{5K\alpha \beta \ln n}{n}\right)^{\rho/8} \left(\frac{\alpha \beta \ln n}{\ln K}\right)^{\ln n},$$

and we may take $\beta > 0$ to be small enough so that expression will tend to 0 faster than $1/n$.

Now consider the case where $a < \frac{n}{\alpha \ln n}$. Let

$$B_0 = \{b \in N_{G}(A) \mid |E(b, A)| \geq K\},$$

and

$$A_0 = \left\{v \in A \mid |E(v, B_0)| \geq \frac{2\alpha \ln n}{\ln K}\right\}. $$

According to (P4), $|A_0| \leq \frac{a}{2}$. Let $A_1 = A \setminus A_0$. Define $\gamma(v)$ for vertices from $A_1$ as in the first case. It follows that

$$\mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2}\right) \leq \mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2} \mid R\right) (1 + o(1))$$

$$\leq \left(\frac{\rho \ln n}{\rho \ln n/2}\right) \left(\frac{|E(v, B_0)|}{d(v)}\right)^{\rho \ln n/2} (1 + o(1)) \leq \left(\frac{5}{\ln K}\right)^{\rho \ln n/2}. $$

Let $A_2 \subseteq A_1$ be the set of vertices $v$ with $\gamma(v) \geq \rho \ln n/2$. We have that

$$\mathbb{P}\left(|A_2| \geq \frac{a}{4}\right) \leq \left(\frac{a}{4}\right) \mathbb{P}\left(\gamma(v) \geq \frac{\rho \ln n}{2}\right)^{a/4} \leq \left(\frac{6}{\ln K}\right)^{a\rho \ln n/8}. $$

Note that for large enough $K$, $\binom{n}{a} \left(\frac{6}{\ln K}\right)^{a\rho \ln n/8} = o(n^{-1})$. Set $A_3 = A_1 \setminus A_2$. Fix $B_3 \subseteq N_{G}(A) \setminus B_0$ with $|B_3| \leq a\beta \ln n$. For $v \in A_3$, let $p_v$ be the probability that a walk which exits $v$ will land in $B_3$. We have that for every $v \in A_3$,

$$p_v \leq \frac{|E(v, B_3)|}{\frac{6}{\alpha \ln n}}.$$
Assuming that $|A_3| \geq \frac{a}{4}$,

$$\frac{1}{|A_3|} \sum_{v \in A_3} p_v \leq \frac{5}{a} \frac{|E(A_3, B_3)|}{\alpha \ln n} \leq \frac{5K\beta}{\alpha},$$

and by the AM/GM inequality (again, making sure $\beta = \beta(K)$ is small enough) we have that

$$\prod_{v \in A_3} p_v \leq \left(\frac{5K\beta}{\alpha}\right)^{a/4}.$$

Recalling (P2), we notice that $|N_G(A)| \leq \frac{\alpha}{3} a \alpha \ln n$. Thus, taking the union bound,

$$\mathbb{P}(\exists A, |A| = a : |N_\Gamma(A)| \leq a\beta \ln n) \leq \left(\frac{\alpha}{a}\right) \left(\frac{6}{\ln K}\right)^{a\rho \ln n/8} + \left(\frac{\alpha}{a}\right) \left(\frac{4\alpha \ln n}{a \beta \ln n}\right) \left(\prod_{v \in A_3} p_v\right)^{\frac{\rho \ln n}{2}} \leq \alpha (n^{-1}) + \left(\frac{en}{a}\right)^a \left(\frac{4\alpha}{\beta}\right)^{a\beta \ln n} \left(\frac{5K\beta}{\alpha}\right)^{a\rho \ln n/8} = \alpha (n^{-1}) + \left(\frac{en}{a}\right)^{1/\ln n} \left(\frac{4\alpha}{\beta}\right)^{\beta} \left(\frac{5K\beta}{\alpha}\right)^{\rho/8} \left(\frac{\alpha}{a}\right)^{a\rho \ln n/8},$$

and we may take $\beta > 0$ to be small enough so that the last expression will tend to 0 faster than $1/n$. Taking the union bound over all cardinalities $1 \leq a \leq n/\ln n$ implies that the claim holds whp in both cases.

### 3.3 Edges between large sets

**Theorem 3.9.** With high probability, there is an edge of $\Gamma$ between every pair of disjoint subsets $A, B \subseteq [n]$ satisfying $|A|, |B| \geq \frac{n(\ln \ln n)^{1.5}}{\ln n}$.

**Proof.** For each vertex $v \in [n]$ and integer $k \geq 0$, let $x_v^k \sim \mathcal{U}(N_G(v))$, independently of each other. Let $\nu_t(v)$ be the number of exits from vertex $v$ by the time $t$. Think of the random walk $X_t$ as follows:

$$X_{t+1} = \mathcal{U}\left(\left\{X_t, x_{X_t}^{\nu_t(X_t)}\right\}\right).$$

That is, with probability $1/2$, the walk stays, and with probability $1/2$ it goes to a uniformly chosen vertex from $N_G(v)$, independently from all previous choices.

Let $\Lambda = \frac{n(\ln \ln n)^{1.5}}{\ln n}$, and fix two disjoint $A, B$ with $|A| = |B| = \Lambda$. Denote

$$B' = \left\{v \in B \mid |E(v, A)| \geq \frac{\Lambda \alpha \ln n}{2n}\right\},$$

and recall that according to property (P5) of an $\alpha$-pseudo-random graph, $|B'| \geq \frac{\Lambda}{2}$. Let $E_{B,A}$ be the event “the walk has exited each of the vertices in $B$ at least $\rho \ln n$ times, but has not traversed an
edge from $B$ to $A$”, and for $v \in B'$, $k > 0$, let $F_{v,k,A}$ be the event “$\forall 0 \leq i < k$, $x_i^v \notin A$”. Clearly, $E_{B,A} \subseteq \bigcap_{v \in B'} F_{v,\rho \ln n, A}$, and $F_{v,\rho \ln n, A}$ are mutually independent, hence,

$$\mathbb{P}(E_{B,A}) \leq (\mathbb{P}(F_{v,\rho \ln n, A}))^{A/2}.$$ 

Let $E$ be the event “there exist two disjoint sets $A, B$ such that the walk $X$ has exited each of the vertices of $B$ at least $\rho \ln n$ times, but has not traversed an edge from $B$ to $A$”. Since for every $v \in B'$, according to (P2),

$$\frac{|A \cap N_G(v)|}{|N_G(v)|} \geq \frac{\Lambda \alpha \ln n}{2 \rho \ln n} \geq 3 \Lambda, \frac{8n}{3 \Lambda \ln n} \geq \Lambda,$$

we have that

$$\mathbb{P}(E) \leq \left(\frac{n}{A}\right)^2 \left(\mathbb{P}(F_{v,\rho \ln n, A})\right)^{A/2} \leq \left(\frac{e\Lambda}{A}\right)^2 \left(1 - \frac{3 \Lambda \rho \ln n}{8n}\right) \leq \exp\left(2 \Lambda \ln (e\Lambda / A) - \frac{3 \rho \Lambda^2 \ln n}{16n}\right) \leq \exp\left(3 \cdot \frac{n}{\ln n} (\ln \ln n)^{2.5} - \frac{n}{\ln n} (\ln \ln n)^{2.9}\right) = o(1).$$

Finally, let $E'$ be the event “there exist two disjoint sets $A, B$ such that the walk $X$ has not traversed an edge from $B$ to $A$”. We have that

$$\mathbb{P}(E') = \mathbb{P}(E', \forall v \in B : \nu(v) \geq \rho \ln n) + \mathbb{P}(E', \exists v \in B : \nu(v) < \rho \ln n) = \mathbb{P}(E) + o(1) = o(1),$$

and that completes the proof. \hfill \Box

## 4 Hamiltonicity and vertex connectivity

This short section is devoted to the proof of Theorem 1, which is a simple corollary of the results presented in the previous sections. In addition to these results, we will use the following Hamiltonicity criterion by Hefetz et al:

**Lemma 4.1 ([16], Theorem 1.1).** Let $12 \leq d \leq e^{3/\ln n}$ and let $G$ be a graph on $n$ vertices satisfying properties (Q1), (Q2) below:

1. **(Q1)** For every $S \subseteq [n]$, if $|S| \leq \frac{n \ln n \ln n \ln d}{\ln \ln n \ln \ln n}$, then $|N(S)| \geq d|S|$;
2. **(Q2)** There is an edge in $G$ between any two disjoint subsets $A, B \subseteq [n]$ such that $|A|, |B| \geq \frac{n \ln n \ln n}{41 \ln n \ln \ln n \ln n}$.

Then $G$ is Hamiltonian, for sufficiently large $n$. 


Proof of Theorem 1. Noting that Theorem 3.1 follows from Theorem 3.8 and Theorem 3.9, and setting \( d = \ln^{1/2} n \) in the above lemma, we see that its conditions are typically met by the trace \( \Gamma_L \). \( L = (1 + \varepsilon)n \ln n \), with much room to spare actually. Hence \( \Gamma_L \) is whp Hamiltonian.

Theorem 3.1 states also that \( \Gamma_L \) is whp \((\frac{n}{\ln n}, \beta \ln n)\)-expander, for some \( \beta > 0 \), and in addition, there is an edge connecting every two disjoint sets with cardinality at least \( \frac{n(\ln \ln n)^{1.5}}{\ln n} \).

Set \( k = \beta \ln n \), and suppose to the contrary that under these conditions, \( \Gamma_L^k \) is not \( k \)-connected. Thus, there is a cut \( S \subseteq [n] \) with \( |S| \leq k - 1 \) such that \( [n] \setminus S \) can be partitioned into two sets, \( A, B \), with no edge connecting them. Without loss of generality, assume \( |A| \leq |B| \). If \( |A| < \beta n - (k - 1) \), take \( A_0 \subseteq A \) with \( |A_0| = \min \{ |A|, \frac{n}{\ln n} \} \), so \( N (A_0) \subseteq A \cup S \) but \( |N (A_0)| \geq \beta \ln n |A_0| > |A \cup S| \), a contradiction. Otherwise, \( |A|, |B| \geq \beta n - (k - 1) \geq n (\ln \ln n)^{1.5} / \ln n \), thus there is an edge connecting the two sets, again a contradiction.

5 Hitting time results

From this point on, a lazy random walk on \( K_n \) is a walk which starts at a uniformly chosen vertex, and at any given step, stays at the current vertex with probability \( 1/n \). Of course, this does not change matters much, and the random walk of the theorem, including its cover time, can be obtained from the lazy walk by simply ignoring loops. Considering the lazy version makes things much more convenient, however; indeed, observe that for any \( t \geq 0 \), the modified random walk is equally likely to be located at any of the vertices of \( K_n \) after \( t \) steps, regardless of its history. Hence, for any \( t \), if we look at the trace graphs \( \Gamma^t \) and \( \Gamma^t_o \) formed by the edges (including loops) traversed by the lazy walk at its odd, respectively even, steps, they are mutually independent, and the graph formed by them is distributed as \( G'(n, m) \) with \( m = \lfloor t/2 \rfloor \) and \( m = \lceil t/2 \rceil \), respectively, where \( G'(n, m) \) is the random (multi)graph formed by drawing independently \( m \) edges (with replacement) from all possible edges (and loops) of the complete graph \( K_n \). Note that whenever \( m = o \left( n^2 \right) \), the probability of a given edge to appear in \( G'(n, m) \) is \( \frac{2m}{n^2} (1 + o(1)) \).

Let now for \( k \geq 1 \),

\[
\begin{align*}
t_{-}^{(k)} &= n(\ln n + (k - 1) \ln \ln n - \ln \ln \ln n), \\
t_{+}^{(k)} &= n(\ln n + (k - 1) \ln \ln n + \ln \ln \ln n).
\end{align*}
\]

We may as well just write \( t_{-} \) or \( t_{+} \), when \( k \) is clear or does not matter. The following is a standard result on the coupon collector problem:

**Theorem 5.1** (Proved in [11]). For every \( k \geq 1 \), whp \( t_{-}^{(k)} < \tau_{C}^{k} < t_{+}^{(k)} \).

To ease notations, we shall denote \( \Gamma_{+} = \Gamma_{t_{+}^{(k)}} \) and similarly \( \Gamma_{-} = \Gamma_{t_{-}^{(k)}} \). We add a superscript \( o \) or \( e \) to consider the odd, respectively even, steps only.

We note that the trace of our walk is typically not a graph, but rather a multigraph. However, that fact does not matter much, as the multiplicity of the edges of that multigraph is typically well bounded, as the following lemma shows:

**Lemma 5.2.** With high probability, the multiplicity of any edge of \( \Gamma_{+} \) is at most 4.
Proof. Suppose the multiplicity of an edge \( e \) in \( \Gamma_+ \) is greater than 4; in that case, its multiplicity in \( \Gamma_+^o \) or in \( \Gamma_+^e \) is at least 3. As \( \Gamma_+^o \sim G(n, [\epsilon/2]) \), we have that the probability for that to happen is \( O(t^3/n^3) = o(n^{-2}) \). Applying the union bound over all possible edges gives the desired result for the odd case (and the even case is identical). \( \square \)

5.1 \( k \)-connectivity

Clearly, if a given vertex has been visited at most \( k - 1 \) times, or has been visited \( k \) times without exiting the last time, its degree in the trace is below \( 2k - 1 \) or \( 2k \) respectively, hence \( \tau^k_C \leq \tau^{2k-1}_\delta \) and \( \tau^k_C + 1 \leq \tau^{2k}_\delta \); furthermore, if some vertex has a (simple) degree less than \( m \), then removing all of its neighbours from the graph will disconnect it, hence it is not \( m \)-vertex-connected, thus \( \tau^m_\delta \leq \tau^m_k \).

To prove Theorem 4 it therefore suffices to prove the following two claims:

Claim 5.3. For any constant integer \( k \geq 1 \), whp \( \tau^k_C \geq \tau^{2k-1}_\delta \) and \( \tau^k_C + 1 \geq \tau^{2k}_\delta \).

Claim 5.4. For any constant integer \( m \geq 1 \), whp \( \tau^m_\delta \geq \tau^m_k \).

5.1.1 The set SMALL

To argue about the relation between the number of visits of a vertex and its degree, we would wish to limit the number of loops and multiple edges incident to a vertex. This can be easily achieved for small degree vertices, which are the only vertices that may affect the minimum degree anyway. This gives motivation for the following definition.

Denote \( d_0 = \lfloor \delta_0 \ln n \rfloor \) for a small constant \( \delta_0 \) to be chosen later. Let

\[
\text{SMALL} = \left\{ v \in [n] \mid d_{\Gamma_+^o}(v) < d_0 \right\}
\]

be the set of all small degree vertices of \( \Gamma_+^o \).

Lemma 5.5. With high probability, no vertex in SMALL is incident to a loop or to a multiple edge in \( \Gamma_+ \).

Proof. Let \( L^i_v \) be the event “\( v \) is incident to a loop in \( \Gamma_+ \) which is the \( i \)th step of the random walk”. Fix a vertex \( v \) and assume it is incident to a loop in \( \Gamma_+ \). Take \( i \) such that the \( i \)th step of \( X_t, e_i \), is a loop incident to \( v \) (that is, \( X_{i-1} = X_i = v \)). Let \( G' \) be the graph obtained from \( \Gamma_+^o \) by removing \( e_j \) for \( i-1 \leq j \leq i+1 \). It is clear then that \( G' \) is distributed like \( G'(n, m) \) with \( t_+ / 2 - 2 \leq m \leq t_+ / 2 - 1 \), and it is independent of the event \( L^i_v \). Hence in \( G' \), \( d(v) \) is distributed as Bin \( (m, 2/(n+1)) \). We begin by estimating the probability that a given vertex is in SMALL. We will use Chernoff bounds (Theorem 2.1) for that. Set \( \mu = 2m/n+1 \) and \( a = \mu - d_0 \). Note that

\[
\mu = \ln n(1 + o(1)),
\]

and

\[
a = (1 - \delta_0)(1 + o(1)).
\]
Thus, by taking $\delta_0$ small enough, we can make $a/\mu$ arbitrarily close to 1, and as $\varphi(x) \nearrow 1$ with $x \searrow 1$, for small enough $\delta_0$ we have that $\varphi(-a/\mu) \geq 0.95$. We conclude that

$$
\mathbb{P}(d_{G'}(v) < d_0) = \mathbb{P}(d_{G'}(v) < \mu - a) \\
\leq \exp \left( -\mu \varphi \left( -\frac{a}{\mu} \right) \right) \\
\leq \exp(-0.9 \ln n) = n^{-0.9}.
$$

Noting that $\mathbb{P}(L_i^v) = n^{-2}$, we can apply the union bound over all vertices and over all potential places for loops at that vertex to obtain the following upper bound for the existence of a vertex from SMALL which is incident to a loop:

$$
\mathbb{P} \left( \exists v \in [n], i \in t_+ : L_i^v, v \in \text{SMALL} \right) \leq n \cdot t_+ \cdot \mathbb{P}(L_i^v, d_{G'}(v) < d_0) \\
= n \cdot t_+ \cdot \mathbb{P}(L_i^v) \mathbb{P}(d_{G'}(v) < d_0) \\
\leq n \cdot t_+ \cdot n^{-0.9} \cdot n^{-2} = o(1).
$$

Using a similar method, we can show that \textbf{whp} there is no vertex in SMALL which is incident to a multiple edge in $\Gamma_+$, and this completes the proof.

The next corollary follows from the proof of the above lemma and Markov’s inequality.

**Corollary 5.6.** With high probability, $|\text{SMALL}| \leq n^{0.2}$.

**Lemma 5.7.** With high probability, for every pair of disjoint vertex subsets $U, W \subseteq [n]$ of size $|U| = |W| = n/\ln^{1/2} n$, $\Gamma^o$ has at least $0.5n$ edges between $U$ and $W$.

**Proof.** We note that $|E_{\Gamma^o}(U, W)|$ is distributed according to $\text{Bin} \left( \lfloor t_- / 2 \rfloor, p \right)$ where $p = \frac{n^2}{\ln n} \left( \frac{n+1}{2} \right)^{-1}$.

As $p > 1.9/\ln n$, using the Chernoff bounds we have that

$$
\mathbb{P} \left( \left| E_{\Gamma^o}(U, W) \right| < 0.5n \right) \leq \mathbb{P} \left( \text{Bin} \left( \lfloor t_- / 2 \rfloor, 1.9/\ln n \right) < 0.5n \right) \\
\leq \mathbb{P} \left( \text{Bin} \left( n \ln n/1.9, 1.9/\ln n \right) \leq n - 0.5n \right) \leq e^{-0.1n},
$$

thus by the union bound

$$
\mathbb{P} \left( \exists U, W : \left| E_{\Gamma^o}(U, W) \right| < 0.5n \right) \leq \left( \frac{n}{\ln^{1/2} n} \right)^2 e^{-0.1n} \\
\leq (e^2 \ln n)^{n/\ln^{1/2} n} e^{-0.1n} \\
\leq \exp \left( \frac{n}{\ln^{1/2} n} \left( 2 + \ln \ln n - 0.1n \right) \right) = o(1).
$$

\hfill \Box
5.1.2 Extending the trace

Now, assuming the edges of the random walk are \( \{e_i \mid i \geq 1\} \), define

\[
\Gamma_* = \Gamma_o - \big\{ e_i \mid 1 \leq i \leq \tau C + 1, e_i \cap \text{SMALL} \neq \emptyset \big\}.
\]

**Corollary 5.8.** *With high probability, \( \delta(\Gamma_*) \geq 2k \).*

**Proof.** Let \( v \) be a vertex. If \( v \notin \text{SMALL} \) then \( d(v) \geq d_0 \) hence \( \text{whp} \) \( d'(v) \geq (d_0 - 8)/4 \geq 2k \) (according to Lemma 5.2). On the other hand, if \( v \in \text{SMALL} \), and is not the first vertex of the random walk, then \( \text{whp} \) it was entered and exited at least \( k \) times in the first \( \tau C + 1 \) steps of the random walk. By the definition of \( \Gamma_* \), all of these enters and exits are in \( E(\Gamma_*) \). Since \( \text{whp} \) none of these vertices is incident to loops or multiple edges, the minimum degree of the set \( \text{SMALL} \) is at least \( 2k \).

Noting that \( \text{whp} \) the first vertex of the random walk is not in \( \text{SMALL} \) we obtain the claim. \( \square \)

We note that by deleting the edge \( e_{\tau C + 1} \) from \( \Gamma_* \) its minimum degree cannot drop by more than one, so Claim 5.3 follows from Corollary 5.8.

**Lemma 5.9.** *With high probability, \( \Delta(\Gamma_*) \leq 6 \ln n \).*

**Proof.** Fix a vertex \( v \); its degree in \( \Gamma_* \) is at most its degree in \( \Gamma_o - \) in addition to its degree in \( \Gamma_e - \). Since its degree in \( \Gamma_e - \) is distributed according to a Binomial distribution with \( \lceil t_- / 2 \rceil \) trials and success probability \( 2 / (n + 1) \), we may use Chernoff bounds to conclude

\[
P\left( d_{\Gamma_o}^e(v) > 3 \ln n \right) \leq P\left( \text{Bin} \left( \frac{n \ln n}{1.9}, \frac{2}{n} \right) > 3 \ln n \right) \leq \exp \left( -\frac{2}{1.9} \ln n \varphi \left( \frac{3 - \frac{2}{1.9}}{1.9} \right) \right) < n^{-1.1}.
\]

Similarly one can derive \( P\left( d_{\Gamma_+}^e(v) > 3 \ln n \right) < n^{-1.1} \). Since \( d'(v) \leq d_{\Gamma_o}^e(v) + d_{\Gamma_+}^e(v) \) we have that \( P\left( d'(v) > 6 \ln n \right) < n^{-1.09} \). The union bound over all vertices gives \( P(\Delta(\Gamma_*) > 6 \ln n) = o(1) \), as we have wished to show. \( \square \)

**Lemma 5.10.** *Fix \( \ell \geq 1 \). With high probability there is no non-empty path of length at most \( \ell \) in \( \Gamma_* \) such that both of its (possibly identical) endpoints lie in \( \text{SMALL} \).*

**Proof.** Fix \( \ell \geq 1 \) and \( P = (v_0, \ldots, v_\ell) \), a path of length \( \ell \). Suppose first that \( v_0 \neq v_\ell \). Let \( A \) be the event \( P \subseteq E(\Gamma_+) \). For every \( \ell \)-tuple \( T \in [\ell] \), let \( A_T \) be the event "\( \forall j \in [\ell], e_{T(j)} = \{v_{j-1}, v_j\}\)". and let \( i(T) \) be the minimal number of integer intervals whose union is the set of elements from \( T \).
We have that
\[ P(A, v_0, v_\ell \in \text{SMALL}) \leq \sum_{T \in [t_+]^\ell} P(A_T, v_0, v_\ell \in \text{SMALL}) \]
\[ = \sum_{r=1}^\ell \sum_{T \in [t_+]^\ell} P(A_T, v_0, v_\ell \in \text{SMALL}). \]

For every \( T \in [t_+]^\ell \), let
\[ I_T = \{ i \in [t_-] : i \text{ is odd, } \exists j \in T : |i - j| \leq 1 \}, \]
and for a vertex \( v \in [n] \), let \( d_{i_T}(v) \) be the degree of \( v \) in the graph formed by the edges \( \{e_i : i \in I_T\} \).
Let \( D_T(v) \) be the event \( d_{i_T}(v) \leq d_0 \). Clearly, \( D_T(v_0) \) and \( D_T(v_\ell) \) are independent of the event \( A_T \).
Moreover, if \( v \in \text{SMALL} \) then \( D_T(v) \), and as there is exactly one edge of \( K_n \) connecting \( v_0 \) with \( v_\ell \), conditioning on the event \( D_T(v_0) \) cannot increase the probability of the event \( D_T(v_\ell) \) by much:
\[ P(D_T(v_0), D_T(v_\ell)) \leq P(D_T(v_0), D_T(v_\ell) | \{v_0, v_\ell\} \notin I_T) \]
\[ = P(D_T(v_0) | \{v_0, v_\ell\} \notin I_T) P(D_T(v_\ell) | \{v_0, v_\ell\} \notin I_T) \]
\[ \leq P(D_T(v_0)) P(D_T(v_\ell)) \frac{1}{P(\{v_0, v_\ell\} \notin I_T)} \]
\[ = P(D_T(v_0)) P(D_T(v_\ell)) (1 + o(1)) \leq n^{-1.7}, \]
here we have used the same reasoning as in Lemma 5.5, and the fact that \(|I_T| = (1 + o(1))n \ln n/2\).
Thus, for a fixed \( T \),
\[ P(A_T, v_0, v_\ell \in \text{SMALL}) \leq P(A_T, D_T(v_0), D_T(v_\ell)) \]
\[ = P(A_T) P(D_T(v_0), D_T(v_\ell)) \]
\[ = P(A_T) \cdot n^{-1.7}. \]

Now, given that \( i(T) = r \) (1 \( \leq r \leq \ell \)), the probability of \( A_T \) is at most \( n^{-(\ell + r)} \). It may be 0, in case \( T \) is not feasible, and otherwise there are exactly \( \ell + r \) times where the walk is forced to be at a given vertex (the walk has to start each of the \( r \) intervals at a given vertex, and to walk according to the intervals \( \ell \) steps in total), and the probability for each such restriction is \( 1/n \). The number of \( T \)'s for which \( i(T) = r \) is \( O((t_+)^r) \) (choose \( r \) points from \([t_+]\) to be the starting points of the \( r \) intervals; then for every \( j \in [\ell] \) there are at most \( r \ell \) options for \( T(j) \)). Noting that the number of paths of length \( P \) is \( n^{\ell + 1} \), the union bound gives
\[ P(\exists P : A, v_0, v_\ell \in \text{SMALL}) \leq n^{\ell + 1} \sum_{r=1}^\ell \frac{O((t_+)^r)}{n^{\ell + r}} \cdot n^{-1.7} \]
\[ \leq n^{-0.7} \sum_{r=1}^\ell O(\ln^r n) < n^{-0.6}. \]

Minor changes to the above argument show that the claim holds for paths with identical endpoints as well.
Lemma 5.11. With high probability, every vertex set \( U \) with \( |U| \leq n/\ln^{1/2} n \) spans at most \( 2|U| \cdot \ln^{3/4} n \) edges (counting multiple edges and loops) in \( \Gamma_* \).

Proof. Fix \( U \subseteq [n] \) with \( |U| = u \leq n/\ln^{1/2} n \). Let \( e^o(U) \) and \( e^e(U) \) be the number of edges (including multiple edges and loops) spanned by \( U \) in \( \Gamma_*^u \) and \( \Gamma_*^e \) respectively. Note that \( e^o(U) \) is binomially distributed with \( \lceil t_+ / 2 \rceil \) trials and success probability \( (u + 1) / (n + 1) \). Thus, using Claim 2.3 we have that

\[
\mathbb{P}(e^o(U) > u \ln^{3/4} n) \leq \left( \frac{e t_+ (u + 1)}{2^{\lceil t_+ / 2 \rceil} u \ln^{3/4} n} \right)^u \leq \left( \frac{e \ln^{1/4} n u}{n} \right)^u .
\]

The union bound over all choices of \( U \) yields

\[
\mathbb{P}(\exists U, |U| \leq n/\ln^{1/2} n, e^o(U) \geq |U| \ln^{3/4} n) \leq \sum_{u=1}^{n/\ln^{1/2} n} \binom{n}{u} \left( \frac{e \ln^{1/4} n u}{n} \right)^u \leq \sum_{u=1}^{n/\ln^{1/2} n} \left( \frac{e n \ln^{1/4} n}{u} \right)^u = o(1),
\]

We now split the sum into two:

\[
\sum_{u=1}^{\ln n} \left( \frac{e n \ln^{1/4} n u}{u} \right)^{\ln^{3/4} n} \leq \ln n \cdot e n \ln^{5/4} n = o(1),
\]

and

\[
\sum_{u=\ln n}^{n/\ln^{1/2} n} \left( \frac{e n \ln^{1/4} n u}{u} \right)^{\ln^{3/4} n} = \sum_{u=\ln n}^{n/\ln^{1/2} n} \left( e \left( \frac{u}{n} \right) \right)^{\ln^{3/4} n - 1} \left( \frac{e \ln^{1/4} n u}{n} \right)^u \leq n \left( e \left( \frac{1}{\ln^{1/2} n} \right) \right)^{\ln^{3/4} n - 1} \left( e \ln^{1/4} n \right)^{\ln^{3/4} n} = o(1).
\]

As the same bound applies for \( x = e \), the union bound over \( x \in \{o, e\} \) concludes the claim (noting that \( \Gamma_* \subseteq \Gamma_+ \)). \qed

Lemma 5.12. With high probability, for every pair of disjoint vertex sets \( U, W \) with \( |U| \leq n/\ln^{1/2} n \) and \( |W| \leq |U| \cdot \ln^{1/4} n \), it holds that \( |E_{\Gamma_*}(U, W)| \leq 2|U| \ln^{0.9} n \).

Proof. For \( U, W \subseteq [n] \), \( |U| \leq n/\ln^{1/2} n \), \( |W| \leq |U| \ln^{1/4} n \), let \( e^o(U, W) \) (\( e^e(U, W) \)) be the number of edges in \( \Gamma_*^o \) (in \( \Gamma_*^e \)) between \( U \) and \( W \). For \( x \in \{o, e\} \), let \( A^x(U, W) \) be the event “\( e^x(U, W) \geq |U| \ln^{0.9} n \)”, and let \( A^x \) be the event “\( \exists U, W, |U| \leq n/\ln^{1/2} n \), \( |W| \leq |U| \ln^{1/4} n \), \( A^x(U, W) \)”.  

26
Fix $U, W$ with $|U| = u \leq n/\ln^{1/2}n$ and $|W| = w \leq u\ln^{1/4}n$. Note that $e^o(U, W)$ is binomially distributed with $[t_+/2]$ trials and success probability $uw/(n+1)$. Thus, using Claim 2.3 we have that

\[
\mathbb{P}(e^o(U, W) > u\ln^{0.9}n) \leq \left(\frac{e^{-1}uw}{2(n+1)u\ln^{0.9}n}\right)^{u\ln^{0.9}n} \leq \left(\frac{e\ln^{0.1}n}{n}\right)^{u\ln^{0.9}n}.
\]

The union bound over all choices of $U, W$ yields

\[
\mathbb{P}(A^o) \leq \sum_{u=1}^{n/\ln^{1/2}n} \sum_{w=1}^{u\ln^{1/4}n} \binom{n}{u} \binom{n}{w} \left(\frac{e\ln^{0.1}n}{n}\right)^{u\ln^{0.9}n} \leq e^{\ln^{0.1}n/\ln^{1/2}n} \left(\frac{e\ln^{0.35}n}{n}\right)^{u\ln^{0.9}n} \leq e^{\ln^{0.35}n}\left(\frac{e\ln^{0.35}n}{n}\right)^{u\ln^{0.9}n} \ln^{0.9}n \leq \sum_{u=1}^{\ln n} u \ln^{1/4}n \left(e^{-1/2}n\right)^{u\ln^{0.9}n\ln^{1/4}n} \leq \ln^{1/4}n \cdot e^{\ln^{0.9}n\ln^{1/4}n} = o(1),
\]

and

\[
\sum_{u=1}^{n/\ln^{1/2}n} e^{1/2\ln n} \left(e^{-1/2}n\right)^{u\ln^{0.9}n\ln^{1/4}n} \leq n^2 \ln^{0.9}n \ln^{1/4}n e^{1/2\ln n} = o(1).
\]

As the same bound applies for $x = e$, the union bound over $x \in \{o, e\}$ concludes the claim (noting that $\Gamma_o \subseteq \Gamma_+$).

We will need the following lemma, according to which not too many edges were added by extending the trace, when we will prove the Hamiltonicity of the trace:

**Lemma 5.13.** With high probability, $|E(\Gamma_o) \setminus E(\Gamma_o^o)| \leq n^{0.4}$. 

27
Proof. From Lemma 5.6 it follows that \(\text{wph } |\text{SMALL}| \leq n^{0.2}\). From Lemma 5.9 it follows that \(\text{wph } \Delta(\Gamma_*) \leq 6 \ln n\). From Lemma 5.2 it follows that \(\text{wph } d_{\Gamma_*}(v) \leq 24 \ln n\) for every \(v \in \text{SMALL}\). We conclude that the number of edges in \(\Gamma_*\) with at least one end in \(\text{SMALL}\) is \(\text{wph}\) at most \(n^{0.2} \cdot 24 \ln n < n^{0.4}\), and the claim follows by the definition of \(\Gamma_*\). 

5.1.3 Sparsifying the extension

We may use the results of lemmas 5.7—5.12 to show that \(\Gamma_*\) is a (very) good expander. This, together with Lemma 2.6, will imply that \(\Gamma_*\) is \(2k\)-connected. However, in order to later show that \(\Gamma_*\) is Hamiltonian, we wish to show it contains a much sparser expander, which is still good enough to guarantee high connectivity.

To obtain this, we assume \(\Gamma_*\) has the properties guaranteed by these lemmas, and sparsify \(\Gamma_*\) randomly as follows: for each vertex \(v\), if \(v \in \text{SMALL}\), define \(E(v)\) to be all edges incident to \(v\); otherwise let \(E(v)\) be a uniformly chosen subset of size \(d_0\) of all edges incident to \(v\). Let \(\Gamma_0\) be the spanning subgraph of \(\Gamma_*\) whose edge set is the union of \(E(v)\) over all vertices \(v\).

Lemma 5.14. With high probability (over the choices of \(E(v)\)), for every pair of disjoint vertex sets \(U, W \subseteq [n]\) of size \(|U| = |W| = n/\ln^{1/2} n\), \(\Gamma_0\) has at least one edge between \(U\) and \(W\).

Proof. Let \(U, W \subseteq [n]\) with \(|U| = |W| = n/\ln^{1/2} n\). From Lemma 5.7 it follows that in \(\Gamma = \Gamma^2\) there are at least \(0.5n\) edges between \(U\) and \(W\). If there is a vertex \(v \in U \cap \text{SMALL}\) with an edge into \(W\), we are done, so we can assume that there is no such. Let \(U' = U \setminus \text{SMALL}\); thus, \(|E_{\Gamma}(U', W)| \geq 0.5n\).

Fix a vertex \(u \in U'\). Let \(X_u\) be the number of edges between \(u\) and \(W\) in \(\Gamma\) that fall into \(E(u)\). \(X_u\) is a random variable, distributed according to Hypergeometric \((d_\Gamma(u), |E_{\Gamma}(u, W)|, d_0)\). According to Theorem 2.2, the probability that \(X_u = 0\) may be bounded from above by

\[
\exp\left(-\frac{|E_{\Gamma}(u, W)| \cdot d_0}{2d_\Gamma(u)}\right),
\]

which, according to Lemmas 5.2 and 5.9, may be bounded from above by

\[
\exp\left(-\frac{|E_{\Gamma}(u, W)| \cdot d_0}{50 \ln n}\right).
\]

Hence, the probability that there is no vertex \(u \in U\) from which there exists an edge to \(W\) can be bounded from above by

\[
\prod_{u \in U'} \exp\left(-\frac{d_0}{50 \ln n} \cdot |E_{\Gamma}(u, W)|\right) = \exp\left(-\frac{d_0}{50 \ln n} \cdot |E_{\Gamma}(U', W)|\right) = \exp(-\Theta(n)).
\]

Union bounding over all choices of \(U, W\), we have that the probability that there exists such a pair of sets with no edge between them is at most

\[
\left(\frac{n}{n/\ln^{1/2} n}\right)^2 e^{-\Theta(n)} \leq \exp\left(\frac{n}{\ln^{1/2} n} (2 + \ln \ln n) - \Theta(n)\right) = o(1).
\]
Lemma 5.15. \( \delta(\Gamma_0) \geq 2k \).

Proof. This follows from Corollary 5.8, since we have not removed any edge incident to a vertex from SMALL and since any other vertex is incident to at least \( d_0 \) edges. \( \square \)

Lemma 5.16. With high probability (over the choices of \( E(v) \)) \( \Gamma_0 \) is a \( \left( \frac{n}{2k+2}, 2k \right) \)-expander, with at most \( d_0 n \) edges.

Proof. Since by definition \( |E(v)| \leq d_0 \) for every \( v \in [n] \), it follows immediately that \( |E(\Gamma_0)| \leq d_0 n \).

Let \( S \subseteq [n] \) with \( |S| \leq n/(2k + 2) \). Denote \( S_1 = S \cap \text{SMALL} \) and \( S_2 = S \setminus \text{SMALL} \). Consider each of the following cases:

**In case** \( |S_2| \geq n/\ln^{1/2} n \): From Lemma 5.14 it follows that the set of all non-neighbours of \( S_2 \) (in \( \Gamma_0 \)) is of cardinality less than \( n/\ln^{1/2} n \). Thus

\[
|N_{\Gamma_0}(S)| \geq n - n/\ln^{1/2} n - |S| \geq \frac{(2k + 1)n}{2k + 2} - n/\ln^{1/2} n \geq \frac{2kn}{2k + 2} \geq 2k|S|.
\]

**In case** \( |S_2| < n/\ln^{1/2} n \): From Lemma 5.15, together with Lemma 5.10, it follows that \( |N_{\Gamma_0}(S_1)| \geq 2k|S_1| \). From Lemma 5.11 it follows that \( S_2 \) spans at most \( 2|S_2| \cdot \ln^{3/4} n \) edges in \( \Gamma_0 \). Consequently,

\[
|\partial_{\Gamma_0} S_2| \geq d_0 |S_2| - 2|E_{\Gamma_0}(S_2)| > |S_2|(d_0 - 4 \ln^{3/4} n) \geq 3|S_2| \cdot \ln^{0.9} n,
\]

hence, by Lemma 5.12 it holds that \( |N_{\Gamma_0}(S_2)| > |S_2| \cdot \ln^{1/4} n \). Finally, by Lemma 5.10 we obtain that for each \( u \in S_2 \), \( |N_{\Gamma_0}(S_1) \cap N_{\Gamma_0}^+(u)| \leq 1 \), hence

\[
|N_{\Gamma_0}(S_1) \cap N_{\Gamma_0}^+(S_2)| \leq |S_2|,
\]

and thus

\[
|N_{\Gamma_0}(S_1) \setminus N_{\Gamma_0}^+(S_2)| \geq 2k|S_1| - |S_2|.
\]

Similarly, for each vertex in \( S_2 \) has at most one neighbour in \( S_1 \), thus

\[
|N_{\Gamma_0}(S_2) \setminus S_1| \geq |N_{\Gamma_0}(S_2)| - |S_2| > |S_2| \cdot \ln^{0.2} n.
\]

To summarize, we have that

\[
|N_{\Gamma_0}(S)| = |N_{\Gamma_0}(S_1) \setminus N_{\Gamma_0}^+(S_2)| + |N_{\Gamma_0}(S_2) \setminus S_1| \\
\geq 2k|S_1| - |S_2| + |S_2| \cdot \ln^{0.2} n \\
\geq 2k(|S_1| + |S_2|) = 2k|S|.
\]

\( \square \)
Since $\Gamma_0$ is whp an $(R, c)$-expander (with $R(c + 1) = \frac{n(2k+1)}{2k+2} \geq \frac{n}{2} + k$), we have that $\Gamma_*$ is such, and from Lemma 2.6 we conclude it is $2k$-vertex-connected. Claim 5.4 follows for even values of $m$. We have already shown (in Claim 5.3) that $\tau_{2k-1}^\delta + 1 = \tau_C^k + 1 = \tau_\kappa^{2k}$. Hence, using what we have just shown we have that $\tau_{2k-1}^\delta + 1 = \tau_\kappa^{2k}$. Since removing an edge may decrease connectivity by not more than 1, it follows that $\tau_{2k-1}^\delta \geq \tau_\kappa^{2k-1}$.

That concludes the proof of Claim 5.4 and of Theorem 4.

5.2 Hamiltonicity

We start by describing the background and tools needed for our proof.

**Definition 5.17.** Given a graph $G$, a non-edge $e = \{u, v\}$ of $G$ is called a booster if adding $e$ to $G$ creates a graph $G'$, which is either Hamiltonian or whose maximum path is longer than that of $G$.

Note that technically every non-edge of a Hamiltonian graph $G$ is a booster by definition. Boosters advance a graph towards Hamiltonicity when added; adding sequentially $n$ boosters clearly brings any graph on $n$ vertices to Hamiltonicity.

**Lemma 5.18.** Let $G$ be a connected non-Hamiltonian $(R, 2)$-expander. Then $G$ has at least $\frac{(R+1)^2}{2}$ boosters.

The above is a fairly standard tool in Hamiltonicity arguments for random graphs, based on the so called Pósa rotation-extension technique [22]. Its proof can be found, e.g., in Chapter 8.2 of [5]. We have proved in Lemma 5.16, for $k = 1$, that $\Gamma_*$ (and thus the trace $\Gamma_{\tau_{\kappa C+1}}$) typically contains a sparse $(\frac{n}{4}, 2)$-expander $\Gamma_0$. We can obviously assume $\Gamma_0$ does not contain loops or multiple edges. Expanders are not necessarily Hamiltonian themselves, but they are extremely helpful in reaching Hamiltonicity as there are many boosters relative to them by Lemma 5.18. We will thus start with $\Gamma_0$ and will add to it boosters repeatedly to bring it to Hamiltonicity. Note that those boosters should come from within. This is taken care of by the following lemma.

**Lemma 5.19.** With high probability there does not exist a non-Hamiltonian $(\frac{n}{4}, 2)$-expander $H \subseteq \Gamma_*$ such that $|E(H)| \leq d_0n + n$, $|E(H) \setminus E(\Gamma_\circ)| \leq n^{0.4}$ and $\Gamma_\circ$ does not contain a booster with respect to $H$.

**Proof.** For a non-Hamiltonian $(\frac{n}{4}, 2)$-expander $H$ with $m \leq d_0n + n$ edges, let $H_o = H \cap \Gamma_\circ$ and $H_e = H \setminus H_o$ be two (random) subgraphs of $H$. Denote by $\mathcal{B}(H)$ the set of boosters with respect to $H$. At the first stage we will choose $H$. For that, we first choose how many edges $H$ has (at most $d_0n + n$) and call that quantity $i$, then we choose the edges themselves between the edges of $K_n$. At the second stage we will choose $H_e$. For that, we first choose how many of $H$’s edges are not in $\Gamma_\circ$ (at most $n^{0.4}$) and call that quantity $j$, then we choose the edges themselves between the edges of $H$. At the third stage, we require all of $H_o$’s edges to appear in $\Gamma_\circ$. For that, we first choose for each edge of $H_o$ a time in which it was traversed, then we actually require that edge to be traversed on that time.
Finally, we wish to bound the probability that given all of the above choices, $\Gamma_0$ does not contain a booster with respect to $H$. For that, let $T$ be the set of times in which edges from $H$ were traversed, and define $I_T$ as in the proof of Lemma 5.10. Note that

$$|I_T| \geq \frac{t}{2} - 2|E(H)| \geq \frac{t}{2} - 3d_0n \geq \frac{t}{3},$$

and observe that every edge traversed in $\Gamma_0$ at one of the times in $I_T$ is chosen uniformly at random, and independently of all previous choices, from all $\binom{n+1}{2}$ possible edges (including loops). Since $H$ is a $\left(\frac{n}{2}, 4\right)$-expander, it is connected, hence by Lemma 5.18, $|B(H)| \geq n^2/32$, and it follows that for $t \in I_T$,

$$\mathbb{P}(e_t \in B(H)) \geq \frac{n^2}{32} \cdot \left(\frac{n+1}{2}\right)^{t-i} \geq \frac{1}{17}.$$

To summarize,

$$\mathbb{P}(\exists H : B(H) \cap E(\Gamma_0) = \emptyset) \leq \sum_{i \leq d_0n} \left(\begin{array}{c} n \\ i \end{array}\right) \sum_{j \leq n^{0.4}} \left(\begin{array}{c} i \\ j \end{array}\right) \left(\frac{t}{2}\right)^{i-j} \left(\frac{n+1}{2}\right)^{-(i-j)} \prod_{t \in I_T} \mathbb{P}(e_t \notin B(H))$$

$$\leq \sum_{i \leq 2d_0n} \left(\frac{3n^2}{2i}\right)^i n^{0.4} (2d_0n)^{n^{0.4}} \left(\frac{t}{n^2}\right)^{-j} \left(\frac{2}{n^2}\right)^i \left(\frac{16}{17}\right)^{t/3}$$

$$\leq \left(\frac{16}{17}\right)^{t/3} \sqrt{n} \sum_{i \leq 2d_0n} \left(\frac{3n^2}{2i}\right)^i \left(\frac{2}{n^2}\right)^i$$

$$\leq \left(\frac{16}{17}\right)^{t/4} \sum_{i \leq 2d_0n} \left(\frac{3t_+}{i}\right)^i.$$

Let $f(x) = (3t_-/x)^2$. In the interval $(0, 3t_+)$, $f$ gets its maximum at $3t_+/e$, and is unimodal. Recalling that $d_0 = \lfloor \delta_0 \ln n \rfloor$, we choose $\delta_0 > 0$ to be sufficiently small so that in the interval $(0, 2d_0n)$, $f$ is strictly increasing. Thus

$$\mathbb{P}(\exists H : B(H) \cap E(\Gamma_0) = \emptyset) \leq \left(\frac{16}{17}\right)^{t/4} 2d_0n \left(\frac{3t_+}{2d_0n}\right)^{2d_0n}$$

$$\leq \exp\left(\frac{t_-}{4} \ln \left(\frac{16}{17}\right) + 2d_0n \ln \left(\frac{3}{\delta_0}\right)\right)$$

$$\leq \exp\left(n \ln n \left(\frac{\ln(16/17)}{4} + 2\delta_0 \ln \left(\frac{3}{\delta_0}\right)\right)\right),$$

and choosing $\delta_0 > 0$ to be sufficiently small, we can make sure this expression is vanishing with growing $n$. \hfill \Box

Now all ingredients are in place for our final argument. We first state that \textbf{whp} the graph $\Gamma_*$ contains a sparse $\left(\frac{n}{4}, 2\right)$-expander $\Gamma_0$, as delivered by Lemma 5.16. We set $H_0 = \Gamma_0$, and as long as $H_i$ is not Hamiltonian, we seek for a booster from $\Gamma_0$ relative to it; once such a booster $b$ is found, we add it to the graph and set $H_{i+1} = H_i + b$. This iteration is repeated less than $n$ times. It cannot get stuck as otherwise we would get graph $H_i$ for which the following hold:
• $H_i$ is a non-Hamiltonian $(\frac{n}{4}, 2)$-expander (as $H_0 \subseteq H_i$)

• $|E(H_i)| \leq d_0 n + n$ (as $|E(\Gamma_0)| \leq d_0 n$)

• $|E(H_i) \setminus E(\Gamma_0)| \leq n^{0.4}$ (follows from Lemma 5.13)

• $\Gamma_0$ does not contain a booster with respect to $H_i$

and by Lemma 5.19, with high probability, such $H_i$ does not exist.

This shows that $\Gamma_{\tau_C + 1}$ is whp Hamiltonian; since $\delta(\Gamma_{\tau_C}) = 1$, $\tau_H = \tau_C + 1$, and the proof of Theorem 2 is complete.

### 5.3 Perfect Matching

Assume $n$ is even. Since $\delta(\Gamma_{\tau_C - 1}) = 0$, in order to prove Corollary 3 it suffices to show that $\tau_{PM} \leq \tau_C$. Indeed, our proof above shows that whp $\Gamma_{\tau_C}$ contains a Hamilton path. Taking every second edge of that path, including the last edge, yields a matching of size $n/2$, thus whp $\Gamma_{\tau_C}$ contains a matching of that size, and Corollary 3 follows.

### 6 Concluding remarks

We have investigated several important graph properties (minimum degree, vertex-connectivity, Hamiltonicity) of the trace of a long-enough random walk on a dense-enough random graph, showing that in the relevant regimes, the trace behaves much like a random graph with a similar density. In the special case of a complete graph, we have shown a hitting time result, which is similar to standard results about random graph processes.

However, the two models are, in some aspects, very different. For example, an elementary result from random graphs states that the threshold for the appearance of a vertex of degree 2 is $n^{-3/2}$, whereas the expected density of the trace of the walk on $K_n$, at the moment the maximum degree reaches 2, is of order $n^{-2}$ (as it typically happens after two steps). It is therefore natural to ask for which graph properties (Planarity? Containment of fixed subgraphs?), and in which regimes, the two models are alike.

Further natural questions inspired by our results include asking for the properties of the trace of the walk in different random environments, such as random regular graphs, or in deterministic environments, such as $(n, d, \lambda)$-graphs and other pseudo-random graphs (see [20] for a survey). We have decided to leave these questions for a future research.

A different direction would be to study the directed trace. Consider the set of directed edges traversed by the random walk. This induces a random directed (multi)graph, and we may ask, for example: is it true that when walking on the complete graph, typically one step after covering the graph we achieve a directed Hamilton cycle?
References


