The Support of Open Versus Closed Random Walks

Thomas Sauerwald ✉
University of Cambridge, UK

He Sun ✉
University of Edinburgh, UK

Danny Vagnozzi ✉
University of Edinburgh, UK

Abstract

A *closed* random walk of length ℓ on an undirected and connected graph $G = (V, E)$ is a random walk that returns to the start vertex at step ℓ, and its properties have been recently related to problems in different mathematical fields, e.g., geometry and combinatorics (Jiang et al., Annals of Mathematics '21) and spectral graph theory (McKenzie et al., STOC '21). For instance, in the context of analyzing the eigenvalue multiplicity of graph matrices, McKenzie et al. show that, with high probability, the support of a closed random walk of length $\ell \geqslant 1$ is $\Omega(\ell^{1/5})$ on any bounded-degree graph, and leaves as an open problem whether a stronger bound of $\Omega(\ell^{1/2})$ holds for any regular graph.

First, we show that the support of a closed random walk of length ℓ is at least $\Omega(\ell^{1/2}/\sqrt{\log n})$ for any regular or bounded-degree graph on n vertices. Secondly, we prove for every $\ell \geqslant 1$ the existence of a family of bounded-degree graphs, together with a start vertex such that the support is bounded by $O(\ell^{1/2}/\sqrt{\log n})$. Besides addressing the open problem of McKenzie et al., these two results also establish a subtle separation between *closed* random walks and *open* random walks, for which the support on any regular (or bounded-degree) graph is well-known to be $\Omega(\ell^{1/2})$ for all $\ell \geqslant 1$. For irregular graphs, we prove that even if the start vertex is chosen uniformly, the support of a closed random walk may still be $O(\log \ell)$. This rules out a general polynomial lower bound in ℓ for all graphs. Finally, we apply our results on random walks to obtain new bounds on the multiplicity of the second largest eigenvalue of the adjacency matrices of graphs.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains

Keywords and phrases support of random walks, eigenvalue multiplicity

Digital Object Identifier 10.4230/LIPIcs.ICALP.2023.103

Category Track A: Algorithms, Complexity and Games

Funding *He Sun*: EPSRC Early Career Fellowship (EP/T00729X/1)
Danny Vagnozzi: EPSRC Early Career Fellowship (EP/T00729X/1)

1 Introduction

A random walk on a graph is a Markov chain in which, starting from some vertex of an undirected graph $G = (V, E)$, the walk moves to one of the neighbors of the current vertex according to the transition matrix of G. As a fundamental stochastic process, random walks have been employed to model numerous mathematical and physical processes. In computer science, random walks have been widely applied in designing randomized and distributed algorithms. Classical examples range from algorithms for satisfiability, deciding connectivity to approximating the volume of convex bodies. The vast majority of research on random walks focuses on "open" random walks, as opposed to *closed* random walks, which are random walks of fixed length ℓ conditioned on being at the start vertex at step ℓ.

50th International Colloquium on Automata, Languages, and Programming (ICALP 2023).
Editors: Kousha Etessami, Uriel Feige, and Gabriele Puppis; Article No. 103; pp. 103:1–103:21
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Following up on an earlier work by Jiang, Tidor, Yao, Zhang and Zhao [11], McKenzie, Rasmussen and Srivastava [16] develop a more systematic study of closed random walks on finite graphs. In particular, they study the support of closed random walks, which is the number of distinct vertices visited by a closed random walk of some length ℓ. As in [11], they then leverage their proven lower bounds on the support of closed random walks to upper bound the eigenvalue multiplicities via the trace method of graph matrices. One of the main ingredients in [16] are general lower bounds on the support of a random walk for any connected graph. For instance, for any bounded-degree graph, a closed random walk is shown to have support at least $\Omega(\ell^{1/5})$. In the same work, they also ask for sharper bounds:

> **Open Question 3 ([16]):** Let $d > 1$ be a fixed integer. Does there exist an $\alpha > 1/5$ such that for every connected d-regular graph G on n vertices and every vertex u of G, a closed random walk[1] of length $2\ell < n$ rooted at u has support $\Omega(\ell^\alpha)$ in expectation? Is $\alpha = 1/2$ true? Does such a bound hold for simple random walks in general?

Note that the constant $\alpha = 1/2$ is a natural target, since any open random walk on a regular graph has support $\Omega(\ell^{1/2})$ (cf. [3, 7]), and this is matched by the n-cycle. Furthermore, also for closed random walks on n-cycles as well as the continuous analogue called Brownian bridges, the support can be shown to be $\Theta(\ell^{1/2})$. In fact, McKenzie et al. [16] states that "*we know of no example where the answer is* $o(\ell^{1/2})$".

In this paper, we address the Open Question 3 of McKenzie et al. [16]. Our first result proves a lower bound of almost $\Omega(\ell^{1/2})$, provided the random walk is sufficiently long. Here we use X^t to denote the vertex that a (lazy) random walk visits in time t, $\mathsf{supp}_{\mathbf{P}}(\ell)$ to denote the number of distinct vertices that a lazy random walk of length ℓ visits, and $\mathbf{P}$ to denote the associated transition matrix (see Section 2 for more on notation).

▶ **Theorem 1.1** (informal version of Theorem 3.1). *Consider any connected, n-vertex graph $G = (V, E)$ with minimum degree δ and maximum degree Δ. Then, for a lazy random walk of length $\ell = O((\Delta/\delta) \cdot n^2 \log n)$ and any vertex $u \in V$, it holds that*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^\ell = X^0 = u\right] = \Omega\left(\frac{\ell^{1/2}}{\sqrt{(\Delta/\delta) \cdot \log n}}\right).$$

This theorem shows that, for any regular or bounded-degree graph, the support of a closed random walk of length ℓ is at least $\Omega\left(\ell^{1/2}/\sqrt{\log n}\right)$; this result improves the lower bound of $\Omega\left(\ell^{1/5}\right)$ from [16, Theorem 1.3] whenever $\ell \geqslant (\log n)^{5/3}$. Apart from the $\ell^{1/2}$-term in our lower bound, one may wonder about the $\sqrt{\log n}$-term, which intuitively does not seem tight. However, we can construct a family of graphs to demonstrate that this $\sqrt{\log n}$-term is needed, establishing that our lower bound is tight up to constant factors. Our upper bound result is summarized as follows:

▶ **Theorem 1.2** (informal version of Theorem 4.1). *The following statements hold:*

- *For any $\ell = \Omega((\log n)^{7/2})$, there exists a family of bounded-degree n-vertex graph $G = (V, E)$ such that a lazy random walk of length ℓ starting at some vertex r satisfies*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^\ell = X^0 = r\right] = O\left(\frac{\ell^{1/2}}{\sqrt{\log n}}\right).$$

[1] The original formulation in [16] is stated for a randomly chosen closed walk, which has the same distribution as a closed random walk since G is regular.

Table 1 Overview of the lower and upper bounds on the support of closed and open random walk. Results highlighted in green are from this work. The lower bound of $\Omega(\ell^{1/5})$ holds for bounded-degree graphs. Note that, while our lower bounds hold for all such graphs and all ℓ, the upper bounds only hold for a specific graph (family) which depends on ℓ and ℓ may be additionally restricted.

Graph	Closed Random Walk		Open Random Walk	
	Lower Bound	Upper Bound	Lower B.	Upper B.
reg./bound. deg.	$\Omega(\ell^{1/5})$ [16]	$O(\ell^{5/14}), \ell \leqslant (\log n)^{7/2}$	$\Omega(\ell^{1/2})$ [7]	$O(\ell^{1/2})$
	$\Omega(\ell^{1/2}/\sqrt{\log n})$	$O(\ell^{1/2}/\sqrt{\log n}), \ell \geqslant (\log n)^{7/2}$		
arbitrary	–	$O(\log \ell), \ell = \Theta(\log n)$	$\Omega(\ell^{1/3})$ [7]	–

- *For any $\ell = O((\log n)^{7/2})$, there exists a family of bounded-degree n-vertex graph $G = (V, E)$ such that a lazy random walk of length ℓ starting at some vertex r satisfies*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^\ell = X^0 = r\right] = O\left(\ell^{5/14}\right).$$

This result shows that the potential lower bound of $\Omega(\ell^{1/2})$ mentioned in [16] (Open Question 3) does not hold in general. In fact, on certain graphs the support of a closed lazy random walk can be as small as $O(\ell^{5/14})$; therefore the "right" exponent must be between $1/5$ and $5/14$, which is a strong separation from the exponent $1/2$ for open random walks. In contrast, when assuming a suitable lower bound on ℓ, the support of closed random walks is $\Theta(\ell^{1/2}/\sqrt{\log n})$, which is nearly the $\Theta(\ell^{1/2})$ bound for open random walks. Table 1 lists the known upper and lower bounds for closed and open random walks; the interplay of these bounds for regular and bounded-degree graphs is further illustrated in Figure 1.

We now proceed to study closed random walks, with a focus on (highly) irregular graphs. McKenzie et al. [16] proves that the support of a randomly chosen closed walk can be as small as $O(\log \ell)$, if starting from a specific vertex (note that a *randomly chosen closed walk* will have a different distribution to a *closed random walk*, unless the graph is regular). Here we provide a similar upper bound for the support of a closed random walk on a family of irregular graphs. Interestingly, this upper bound even holds if we assume that the start vertex is chosen uniformly at random. This lower bound also establishes an "exponential discrepancy" on the support of closed random walks versus open random walks on irregular graphs: while the support of open random walks is known to be at least $\ell^{1/3}$ (cf. [3]) for any graph and any start vertex, one can construct graphs for which the support is only $O(\log \ell)$ for closed random walks.

▶ **Theorem 1.3** (informal version of Theorem 4.5)**.** *There exists a family of connected, n-vertex graphs $G = (V, E)$ such that a lazy random walk of some length $\ell = \Theta(\log n)$ that starts from a vertex chosen uniformly at random satisfies*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0\right] = O(\log \ell).$$

Consequently, there is a vertex $r \in V$ such that

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = r\right] = O(\log \ell).$$

As side results, we apply our lower bounds on the support of closed random walks to obtain new eigenvalue multiplicity bounds for certain classes of graphs. We state one result below, and an additional result is presented in Section 5.

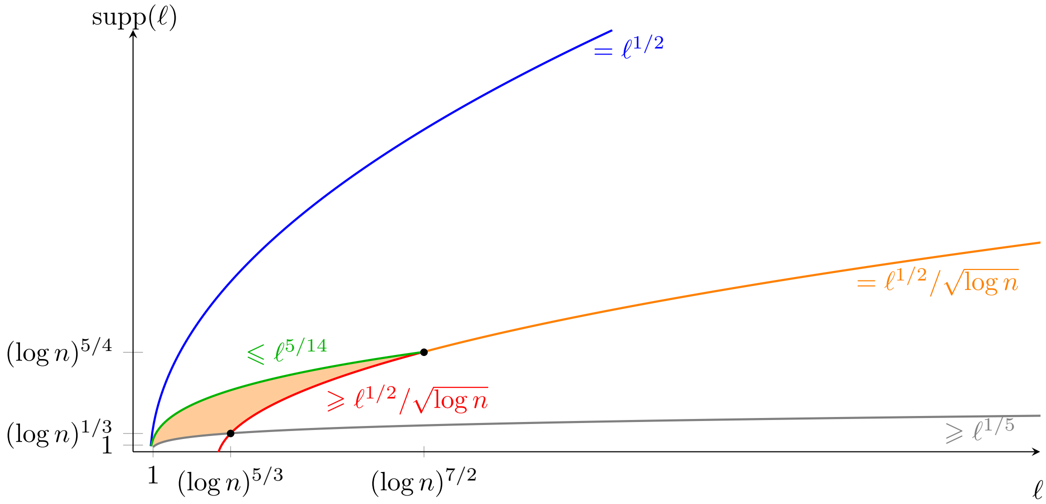

Figure 1 Comparison of the (worst case) support of open random walks to that of closed random walks for bounded-degree graphs. We prove for any $\ell = \Omega((\log n)^{7/2})$ (and $\ell \leqslant n^{1/5}$) that the asymptotic worst case bound for closed random walks equals $\ell^{1/2}/\sqrt{\log n}$, while for $\ell = O((\log n)^{7/2})$ the correct asymptotic bound is confined in the orange area. Since the lower bound $\ell^{1/2}/\sqrt{\log n}$ is tight for any $\ell \geqslant (\log n)^{7/2}$ but becomes trivial for $\ell \leqslant \log n$, we can infer that there must be a phase transition in the interval $[(\log n)^{5/3}, (\log n)^{7/2}]$.

▶ **Theorem 1.4** (informal version of Theorem 5.1)**.** *Consider any connected, n-vertex graph $G = (V, E)$ with minimum degree δ and maximum degree Δ, such that its second largest eigenvalue λ of $\mathbf{P}$ satisfies $|1-\lambda| = O(\frac{\delta}{\Delta} \cdot \frac{1}{\log^4 n})$. Then, the number of eigenvalues of $\mathbf{P}$ in the range $\left[\left(1 - \frac{\delta}{32c\Delta \cdot \log^5 n}\right) \cdot \lambda, \lambda\right]$ is at most $O\left(\frac{n}{\log n}\right)$.*

Our eigenvalue multiplicity results are in general incomparable with the ones in [16], and have their own features. For instance, our eigenvalue multiplicity bound above is based on the spectral gap condition of $\mathbf{P}$; as such, this result brings a new connection between the eigenvalue multiplicity and the eigenvalue distribution, the relationship of which is informally established in spectral graph theory through the high-order Cheeger inequalities [12].

1.1 Further Related Work

There is a plethora of works on random walks on graphs, which often revolves around quantities such as mixing times, hitting times and cover times [14]. In particular, properties of short random walks such as their support or return probabilities have found numerous applications in the analysis of randomness amplification [10], space-efficient graph exploration with random walks [2, 7] and the voter model [17]. These concepts have been also successfully applied to algorithmic tasks such as estimating network sizes and densities [4], load balancing [19], information spreading [8], property testing [6] and clustering [20]. In addition to these applications, many of the random walk quantities have close connections to other mathematical areas, such as geometry, group theory, electrical networks and spectral graph theory (see [13, 14] for more details).

More closely related to this work, Benjamini, Izkovsky and Kesten [5] analyze the support of closed random walks on various finite and infinite graphs, with a focus on vertex-transitive and Cayley graphs. In particular, for high-girth expander graphs, they prove that the support of closed random walks is linear in their length. While there are some studies of closed random walks on finite or infinite graphs with special geometry or symmetries as well as studies of Brownian bridges in continuous space, much less is known about the support of closed random walks on finite graphs (without additional assumptions on their symmetry or geometry).

One specific motivation for studying closed random walks is the relation to the second eigenvalue multiplicity of the normalized adjacency matrix of graphs, as established recently in [11, 15, 16]. By examining the support of closed random walks of short length, it is shown that, for any connected graph G of maximum degree Δ, the second eigenvalue multiplicity of G's normalized adjacency matrix is $\widetilde{O}\left(n \cdot \Delta^{7/5} / \log^{1/5} n\right)$, where the notation $\widetilde{O}(\cdot)$ suppresses $\operatorname{poly}\log\log(n)$ terms. Haiman, Schildkraut, Zhang and Zhao [9] show the existence of infinitely many connected 18-regular graphs G on n vertices with the second largest eigenvalue multiplicity at least $n^{2/5} - 1$, and the existence of infinitely many connected n-vertex graphs with maximum degree 4 and second eigenvalue multiplicity at least $\sqrt{n/\log_2 n}$.

1.2 Organization

The remaining part of the paper is organized as follows. Section 2 introduces our notation and provides some basic lemmas used in this work. We derive our lower bound on the support of closed random walks in Section 3, and the proofs of our two upper bound results are presented in Section 4. Finally, we employ our random walk results to analyze the eigenvalue multiplicity problem in Section 5. We summarize our results and point to some open questions in Section 6.

2 Definitions and Preliminaries

All graphs in this paper will be undirected. For any vertex $u \in V$ of a graph $G = (V, E)$, the degree of u is denoted by $\deg(u)$; the maximum and minimum degrees of G are denoted by Δ and δ, respectively. For any $u \in V$ and $\ell \in \mathbb{N}$, let $B_{\leqslant \ell/2}(u) \triangleq \{v \in V \colon \operatorname{dist}(u, v) \leqslant \ell/2\}$. For any integer k, let $[k] \triangleq \{1, \ldots, k\}$.

We use $\mathbf{Q}$ to represent the transition matrix of a non-lazy random walk in G defined by $\mathbf{Q}_{u,v} = \frac{1}{\deg(u)}$ if $\{u, v\} \in E(G)$ and $\mathbf{Q}_{u,v} = 0$ otherwise. We use $\mathbf{P}$ to represent the lazy random walk matrix of G, where $\mathbf{P}_{u,u} = \frac{1}{2}$ for all $u \in V$, $\mathbf{P}_{u,v} = \frac{1}{2\deg(u)}$ if $\{u, v\} \in E(G)$, and $\mathbf{P}_{u,v} = 0$ otherwise. We use $\mathbf{A}$ to represent the adjacency matrix of G, and $\mathbf{D}$ to represent the diagonal matrix of degrees of G. For any matrix $\lozenge \in \{\mathbf{P}, \mathbf{Q}\}$ of size $n \times n$, the eigenvalues of $\lozenge$ are denoted by $\lambda_1(\lozenge) \geqslant \ldots \geqslant \lambda_n(\lozenge)$. We define $M_{\lozenge}[x, y]$ to be the number of eigenvalues of the matrix $\lozenge$ in the interval $[x, y]$.

For any (non-)lazy random walk that starts from a vertex (or possibly distribution over vertices) $X^0 \in V$, we use X^t to denote the vertex that the random walk reaches at step t for any $t \geqslant 0$, and define

$$p^t_{u,v} \triangleq \mathbf{Pr}\left[\, X^t = v \mid X^0 = u \,\right]$$

to be the probability that a random walk that starts from u is located at v after t steps; if the start vertex is deterministic and clear from the context, we sometimes omit the conditioning on $X^0 = u$. Further, we write $X^0 \sim \mathcal{U}$ if the start vertex of the random walk is chosen uniformly at random from V. We define the support of a random walk of length ℓ to be

$$\mathsf{supp}_{\lozenge}(\ell) \triangleq \left|\left\{X^i \mid i \leqslant \ell\right\}\right|,$$

where we use $\diamondsuit \in \{\mathbf{P}, \mathbf{Q}\}$ to distinguish a lazy random walk from a non-lazy one. It is well-known that, for a lazy random walk with loop probability $1/2$ on a connected graph G, it holds for all $u, v \in V$ that $\lim_{t\to\infty} p_{u,v}^t = \pi(v)$, where $\pi \in \mathbb{R}^n_{\geqslant 0}$ is the stationary distribution defined by $\pi(u) \triangleq \frac{\deg(u)}{2\cdot|E(G)|}$ for any $u \in V$.

We list two lemmas used in our analysis. Our first lemma allows us to translate bounds on the support from lazy random walks to non-lazy ones (and vice versa) at the cost of a small constant factor.

▶ **Lemma 2.1.** *For any graph $G = (V, E)$ and any fixed $\ell \geqslant 0$, it holds that $\mathsf{supp}_{\mathbf{Q}}(\ell)$ is stochastically larger than $\mathsf{supp}_{\mathbf{P}}(\ell)$. Moreover, we have for any $x \geqslant 0$ that*

$$\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(4\cdot\ell) \geqslant x\,\right] \geqslant \frac{1}{2}\cdot\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{Q}}(\ell) \geqslant x\,\right],$$

and thus

$$\mathbf{E}\left[\,\mathsf{supp}_{\mathbf{P}}(4\cdot\ell)\,\right] \geqslant \frac{1}{2}\cdot\mathbf{E}\left[\,\mathsf{supp}_{\mathbf{Q}}(\ell)\,\right].$$

The next lemma gives a lower bound on the return probability of a random walk. While there are a number of results upper bounding the return probability of a random walk (e.g., [13, 14, 18]),to the best of our knowledge much less is known in terms of lower bounds.

▶ **Lemma 2.2.** *For any connected, n-vertex graph $G = (V, E)$ and a lazy random walk with transition matrix $\mathbf{P}$, it holds for any vertex $u \in V$ and step $t \geqslant 0$ that*

$$p_{u,u}^t \geqslant \pi(u) = \frac{\deg(u)}{2|E|} \geqslant \frac{1}{n^2}.$$

The same also holds for non-lazy random walks with transition matrix $\mathbf{Q}$, if additionally t is even. Furthermore, if G has minimum degree δ and maximum degree Δ, we also have for any $t \geqslant 2$,

$$p_{u,u}^t \geqslant \frac{\delta^t}{\Delta^{t+1}}\cdot\frac{1}{|B_{\leqslant t/2}(u)|},$$

and the same inequality holds for the transition matrix $\mathbf{Q}$ if $t \geqslant 2$ is even.

3 A Lower Bound on the Support of Closed Random Walks

This section provides lower bounds on the support of closed random walks, in particular, we will prove Theorem 1.1. We first present a more detailed formulation of Theorem 1.1, in which all the hidden constants are stated precisely.

▶ **Theorem 3.1.** *Consider any connected, n-vertex graph $G = (V, E)$ with minimum degree δ and maximum degree Δ. Then there is a constant $c \geqslant 1$ independent of n, such that a random walk of length $\ell \leqslant 512\frac{\Delta}{\delta}cn^2\log n$ satisfies for any $u \in V$ that*

$$\mathbf{E}\left[\mathsf{supp}_{\diamondsuit}(\ell)\;\middle|\; X^\ell = X^0 = u\right] \geqslant \frac{1}{2}\cdot\left\lfloor\sqrt{\frac{1}{576c}\cdot\frac{\delta}{\Delta}\cdot\frac{\ell}{\log n}}\right\rfloor,$$

where $\diamondsuit \in \{\mathbf{P}, \mathbf{Q}\}$ (in the case of $\diamondsuit = \mathbf{Q}$, the length ℓ needs additionally to be even). Furthermore, for any $\mu \in [1, \ell]$ satisfying $\ell \leqslant 32\frac{\Delta}{\delta}cn^2\mu$, it holds that

$$\mathbf{Pr}\left[\mathsf{supp}_{\diamondsuit}(\ell) \leqslant \left\lfloor\sqrt{\frac{1}{64c}\cdot\frac{\delta}{\Delta}\cdot\frac{\ell}{\mu}}\right\rfloor\;\middle|\; X^\ell = X^0 = u\right] \leqslant (5/8)^{\mu/2}\cdot\frac{1}{p_{u,u}^\ell},$$

where ℓ is required to be even if $\diamondsuit = \mathbf{Q}$.

To examine the significance of this result, notice that it is shown in [16, Theorem 1.3] that

$$\mathbf{Pr}\left[\mathsf{supp}_{\mathbf{Q}}(\ell) \leqslant s \;\middle|\; X^{\ell} = u\right] \leqslant \exp\left(-\frac{\ell}{130\Delta^7 s^4}\right),$$

if $s \leqslant \frac{1}{4}\left(\frac{\ell}{2\Delta^7 \log \Delta}\right)^{1/5}$ and ℓ is even. In comparison to their result, our bound is not affected by the density, but only by the degree ratio Δ/δ. Regarding the expected support, it follows that for bounded-degree graphs, the first statement in Theorem 3.1 improves on [16, Theorem 1.3] for moderately longer walks, i.e., $\ell \geqslant (\log n)^{5/3}$, whereas it is worse for $\ell \leqslant (\log n)^{5/3}$.

Next, we present the lemmas needed to prove Theorem 3.1. The first lemma (Lemma 3.3) lower bounds the support of *open* random walks, and relies on the following result by Feige [7] bounding the expected time until a certain number of distinct vertices are visited; similar results are also shown in [3].

▶ **Lemma 3.2** ([7, Theorem 4])**.** *Consider any connected, n-vertex graph $G = (V, E)$ with minimum degree δ and maximum degree Δ. For any $s \in [1, n]$, let $T(s)$ be the time until a random walk visits s distinct vertices. Then there is a constant $c \geqslant 1$ (independent of s and n), such that for any lazy random walk and any vertex $u \in V$,*

$$\mathbf{E}\left[T(s) \mid X^0 = u\right] \leqslant c \cdot \left(s + \frac{s^2}{\delta} \cdot \min\{s, \Delta\}\right).$$

▶ **Lemma 3.3.** *Consider any connected, n-vertex graph $G = (V, E)$, and a random walk of length ℓ for some $\ell \triangleq 32 \cdot \lceil c \cdot \frac{\Delta}{\delta} \cdot s^2 \rceil$, where $1 \leqslant s \leqslant n$ is any integer. Then, there is some constant $c \geqslant 1$ (independent of n and ℓ), such that it holds for any start vertex $u \in V$ that*

$$\mathbf{Pr}\left[\mathsf{supp}_{\Diamond}(\ell) \geqslant s \mid X^0 = u\right] \geqslant \frac{3}{8},$$

where $\Diamond \in \{\mathbf{P}, \mathbf{Q}\}$.

Proof. First of all, we note that the constant c involved in this result is the constant from Lemma 3.2. With this, we first prove the result for non-lazy random walks, i.e., for $\Diamond = \mathbf{Q}$. Let

$$\widetilde{\ell} \triangleq 8 \cdot \left\lceil c \cdot \frac{\Delta}{\delta} \cdot s^2 \right\rceil.$$

Recall that $T(s)$ is the stopping time until a walk has visited s different vertices. By Lemma 3.2, we have

$$\mathbf{E}\left[T(s) \;\middle|\; X^0 = u\right] \leqslant c \cdot \left(s + \frac{\Delta}{\delta} s^2\right) \leqslant 2 \cdot \left\lceil c \cdot \frac{\Delta}{\delta} s^2 \right\rceil = \frac{1}{4} \cdot \widetilde{\ell}.$$

By Markov's inequality, it holds that

$$\mathbf{Pr}\left[T(s) \geqslant \widetilde{\ell} \;\middle|\; X^0 = u\right] \leqslant \mathbf{Pr}\left[T(s) \geqslant 4 \cdot \mathbf{E}\left[T(s)\right] \;\middle|\; X^0 = u\right] \leqslant \frac{1}{4}.$$

Note that $T(s) \leqslant \widetilde{\ell}$ is equivalent to $\mathsf{supp}(\widetilde{\ell}) \geqslant s$, and this gives us that

$$\mathbf{Pr}\left[\mathsf{supp}(\widetilde{\ell}) \geqslant s \;\middle|\; X^0 = u\right] \geqslant \frac{3}{4} \geqslant \frac{3}{8}, \tag{1}$$

which completes the proof in case of $\Diamond = \mathbf{Q}$. For $\Diamond = \mathbf{P}$, the statement follows immediately from (1), and the second statement of Lemma 2.1, since $\ell = 4 \cdot \widetilde{\ell}$. ◀

We remark that the bound presented in Lemma 3.3 is essentially tight: if one takes a random walk of length ℓ on a path (or cycle), then by the Central Limit Theorem the probability that the walk visits at least $\varepsilon \cdot \sqrt{\ell}$ vertices can be upper bounded by $1-\delta$ for some $\delta = \delta(\varepsilon) > 0$; in particular, this probability can be bounded independently of ℓ.

Proof of Theorem 3.1. Fix an arbitrary start vertex $u \in V$ as $X^0 = u$. We split the random walk of length ℓ into consecutive sections of length $\ell' \triangleq \lceil \ell/(9\log n) \rceil$. Without loss of generality, we assume that $\ell \geqslant 576 \frac{\Delta}{\delta} c \log n$, since otherwise the statement holds trivially. Given $\ell \geqslant 576 \frac{\Delta}{\delta} c \log n$ (and $c \geqslant 1$), we have $\ell' \leqslant \ell/(8\log n)$. Hence, it would take a random walk (at least) $8 \log n$ sections before reaching step ℓ. Next we define the integer

$$\gamma \triangleq \left\lfloor \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \ell'} \right\rfloor,$$

with c being the constant from Lemma 3.2. We make the following observations about the range of γ:

- Since $\ell' \leqslant \ell/(8\log n)$ and by the precondition $\ell \leqslant 512 \frac{\Delta}{\delta} c n^2 \log n$, we have

$$\gamma \leqslant \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \frac{\ell}{8\log n}} \leqslant n.$$

- Similarly, since $\ell' \geqslant \ell/(9\log n)$ and $\ell \geqslant 576 \frac{\Delta}{\delta} c \log n$, we have

$$\gamma \geqslant \left\lfloor \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \frac{\ell}{9\log n}} \right\rfloor \geqslant 1.$$

In conclusion, γ is an integer between 1 and n. Notice that the definition of γ implies

$$\gamma \leqslant \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \ell'},$$

and thus $\ell' \geqslant 64c \cdot \frac{\Delta}{\delta} \cdot \gamma^2$. Since $\gamma \geqslant 1$, we have $c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \geqslant 1$ and

$$c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \geqslant \frac{1}{2} \left\lceil c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \right\rceil.$$

This implies that

$$\ell' \geqslant 64c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \geqslant 32 \cdot \left\lceil c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \right\rceil.$$

We now apply Lemma 3.3 (with $s = \gamma$) and conclude

$$\mathbf{Pr}\left[\mathsf{supp}_{\Diamond}(\ell') \geqslant \gamma \;\middle|\; X^0 = u \right] \geqslant \mathbf{Pr}\left[\mathsf{supp}_{\Diamond}\left(32 \cdot \left\lceil c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \right\rceil\right) \geqslant \gamma \;\middle|\; X^0 = u \right] \geqslant \frac{3}{8},$$

which holds for any start vertex $u \in V$. Therefore, by considering the at least $8\log n$ consecutive sections of length ℓ' each, and using the Markov property we have

$$\begin{aligned} \mathbf{Pr}\left[\mathsf{supp}_{\Diamond}(\ell) < \gamma \;\middle|\; X_0 = u \right] &\leqslant \left(\max_{v \in V} \mathbf{Pr}\left[\mathsf{supp}_{\Diamond}(\ell') < \gamma \;\middle|\; X_0 = v \right] \right)^{8\log n} \leqslant \left(\frac{5}{8}\right)^{8\log n} \\ &\leqslant n^{-3}, \end{aligned}$$

since $(5/8)^8 \leqslant \mathrm{e}^{-3}$.

By Lemma 2.2, it holds for a lazy random walk (i.e., $\Diamond = \mathbf{P}$) starting with any $u \in V$ and integer $\ell \geqslant 0$ that

$$\mathbf{Pr}\left[X^\ell = u \;\middle|\; X^0 = u \right] \geqslant \pi(u) \geqslant \frac{\delta}{\Delta \cdot n};$$

we also know that the same statement holds for a non-lazy random walk (i.e., $\Diamond = \mathbf{Q}$) and an even value of ℓ. Hence,

$$\begin{aligned}
\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) < \gamma \;\middle|\; X^0 = X^\ell = u \right] &= \frac{\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) < \gamma \cap X^0 = X^\ell = u \right]}{\mathbf{Pr}\left[X^0 = X^\ell = u \right]} \\
&\leqslant \frac{\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) < \gamma \right]}{\mathbf{Pr}\left[X^\ell = u \;\middle|\; X^0 = u \right]} \\
&\leqslant \frac{n^{-3}}{n^{-2}} = n^{-1}. \qquad (2)
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathbf{E}\left[\mathsf{supp}_\Diamond(\ell) \;\middle|\; X^0 = X^\ell = u \right] &\geqslant \gamma \cdot \mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) \geqslant \gamma \;\middle|\; X^0 = X^\ell = u \right] \\
&\geqslant \gamma \cdot (1 - n^{-1}) \\
&\geqslant \frac{1}{2} \cdot \left\lfloor \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \frac{\ell}{9 \log n}} \right\rfloor = \frac{1}{2} \cdot \left\lfloor \sqrt{\frac{1}{576c} \cdot \frac{\delta}{\Delta} \cdot \frac{\ell}{\log n}} \right\rfloor,
\end{aligned}$$

which proves the first statement.

We proceed to the proof of the second statement, which essentially uses the same argument as before but with different parameters. First note that we may assume $\ell \geqslant 64c \cdot \frac{\Delta}{\delta} \cdot \mu$, since otherwise the statement holds trivially. In particular, this implies $\ell/\mu \geqslant 1$, so if we split the random walk of length ℓ into consecutive sections of length $\ell' \triangleq \lceil \ell/\mu \rceil$, it holds that $\ell' \leqslant 2\ell/\mu$. Hence there are at least $\mu/2$ consecutive sections before reaching step ℓ. We define

$$\gamma \triangleq \left\lfloor \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \ell'} \right\rfloor,$$

and rearranging this implies

$$\ell' \geqslant 64c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \geqslant 32 \cdot \left\lceil c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \right\rceil.$$

Again, we examine the range of γ:

- Since $\ell' \leqslant 2\ell/\mu$ and by the precondition $\ell \leqslant 32\frac{\Delta}{\delta} cn^2\mu$, we have $\gamma \leqslant n$.
- Since $\ell' \geqslant \ell/\mu$ and $\ell \geqslant 64c \cdot \frac{\Delta}{\delta} \cdot \mu$, we have $\gamma \geqslant 1$.

In conclusion, γ is an integer between 1 and n. By Lemma 3.3 (with $s = \gamma$), it holds that

$$\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell') \geqslant \gamma \;\middle|\; X^0 = u \right] \geqslant \mathbf{Pr}\left[\mathsf{supp}_\Diamond\left(32 \left\lceil c \cdot \frac{\Delta}{\delta} \cdot \gamma^2 \right\rceil \right) \geqslant \gamma \;\middle|\; X^0 = u \right] \geqslant \frac{3}{8},$$

and, as in the proof of the first statement,

$$\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) \leqslant \gamma \;\middle|\; X^0 = u \right] \leqslant (5/8)^{\mu/2}.$$

Finally, we apply the same argument as in (2) to conclude that

$$\mathbf{Pr}\left[\mathsf{supp}_\Diamond(\ell) \leqslant \gamma \;\middle|\; X^0 = X^\ell = u \right] \leqslant (5/8)^{\mu/2} \cdot \frac{1}{p^\ell_{u,u}}. \qquad \blacktriangleleft$$

4 Upper Bounds on the Support of Closed Random Walks

This section studies upper bounds on the support of closed random walks, by examining certain "worst-case" graphs. The section is structured as follows: we first study a family of bounded-degree graphs, and give the proof of Theorem 1.2 in Section 4.1. We present a more formal statement of Theorem 1.3, and prove the statement in Section 4.2.

4.1 Proof of Theorem 1.2

We first present a more detailed formulation of Theorem 1.2, in which all the hidden constants are stated precisely.

▶ **Theorem 4.1.** *There is a constant $C \geqslant 1$, such that for any pair of integers β being a power of 2 and ℓ with $C \leqslant \ell \leqslant \beta^{1/5}$ the following holds: there is a connected, n-vertex graph G satisfying $n \in [2\beta + 1, 2\beta + \beta^{1/10} - 1]$, $\Delta = 3$, and some vertex $r \in V(G)$, such that it holds for a random walk of length ℓ that starts at r that*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^{\ell} = X^{0} = r \right] \leqslant 3 \cdot \ell^{1/2-\varepsilon},$$

where $\varepsilon \triangleq \min\left(1/2 \cdot \log(1/16 \cdot \log_{12} \beta)/\log \ell, 1/7\right)$.

We first construct the family of graphs G used in the proof. Given two integers $p \geqslant 1$ and $D \geqslant 1$, our constructed graph $G = G[p, D]$ is based on the following two graphs:

- let $G_1 = (V_1, E_1)$ be a path graph of length $p + 1$, where $V_1 = \{v_0, v_1, \ldots, v_p\}$ and $E_1 = \{\{v_i, v_{i+1}\} \mid 0 \leqslant i < p\}$;
- let $G_2 = (V_2, E_2)$ be a complete binary tree over $D \geqslant 1$ levels labelled $0, 1, \ldots, D-1$. Hence, G_2 has $2^{D+1} - 1$ vertices.

We set the root of the binary tree to be vertex $z = v_0$, and let G be the union of the graphs G_1 and G_2; see Figure 2 for an illustration of our considered graph. Notice that G has $n = p + (2^{D+1} - 1)$ vertices, as vertex $z = v_0$ appears in both G_1 and G_2. Since all vertices of graph G have degree one, two or three, we have $\Delta = 3$.

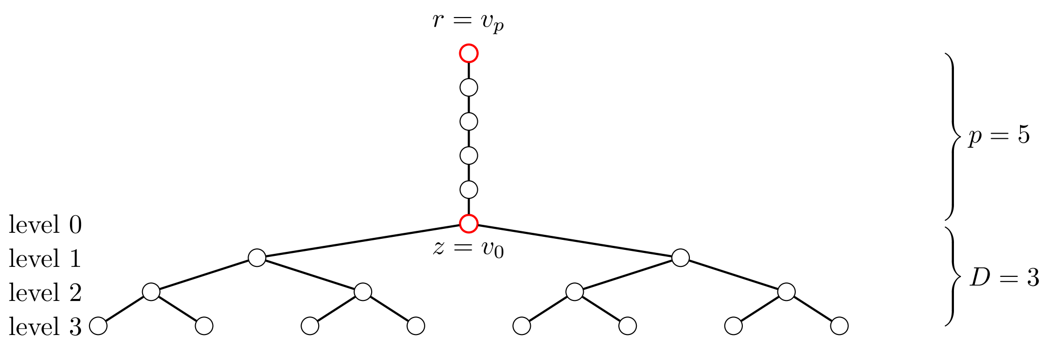

Figure 2 The construction of the graph $G = G[p, D]$ for $p = 5$, $D = 3$. We have $2^4 - 1 = 15$ vertices on the binary tree, and the total number of vertices of G is $n = 5 + 16 - 1 = 20$.

Now we explain the intuition behind our construction. We study a closed random walk from r of length p^{2+c} for some suitably small constant $c > 0$. On one hand, if the closed random walk remains only on the path, then the support of this walk is at most $p \ll (p^{2+c})^{1/2}$. On the other hand, once the random walk leaves the path, the walk is likely to "get lost" in the binary tree, and the probability for the walk to return to $r = v_p$ is very small. Hence, if we sample from the space of all closed random walks, there is a strong bias towards those walks that never leave the path (and thus have necessarily small support).

Now we turn this intuition into a formal proof. Recall that the *level* of a vertex in the binary tree is the distance to the root vertex z, and recall that the maximum level is equal to the depth D.

▶ **Lemma 4.2.** *Consider a lazy random walk on $G = G[p, D]$ starting at r. For any integer $\ell \geqslant 0$, the following holds with probability at least $1 - \exp(-\ell/288)$: if the random walk makes at least ℓ transitions within the binary tree, it reaches at least once a vertex which is at level $\min(D, \ell/12)$.*

This lemma is quite intuitive, as there is a strong drift on the binary tree to increase the distance to the root, and one can exploit this using Hoeffding's inequality. Next, we present a simple fact of random walks on binary trees.

▶ **Lemma 4.3.** *Consider a lazy random walk in a complete binary tree with levels $0, 1, \ldots, D$ starting at a vertex which has distance $k \in [1, D]$ from the root. Then the probability that the walk reaches the root within ℓ steps is upper bounded by $\ell \cdot 2^{-k}$.*

Lemmas 4.2 and 4.3 together establish the intuitive fact that, once the random walks makes sufficiently many transitions in the binary tree, it is unlikely to return to the root of the tree within a small number of steps.

Next, we consider a lazy random walk on a path with vertices $0, 1, \ldots, p$, starting from p, with the special property that the random walk "gets killed" once it reaches the other endpoint 0. The following lemma lower bounds the probability of the random walk "surviving" after $\gamma \cdot p^2$ steps, i.e., the probability that a random walk does not reach the other endpoint before step $\gamma \cdot p^2$.

▶ **Lemma 4.4.** *Consider a lazy random walk on the integers $\{0, 1, \ldots, p\}$, starting at vertex p, such that the random walk gets killed after reaching vertex 0. More precisely, we define the following $p \times p$ matrix $\mathbf{R}$:* [2]

$$\mathbf{R}_{i,j} = \begin{cases} \frac{1}{2} & \textit{if } j = i \in \{1, \ldots, p\}, \\ \frac{1}{2} & \textit{if } i = p,\ j = p-1, \\ \frac{1}{4} & \textit{if } j = i-1,\ 1 < i < p, \\ \frac{1}{4} & \textit{if } j = i+1,\ 1 \leqslant i < p. \end{cases}$$

Let $r^t_{p,\cdot}$ be the probability distribution[3] *of this t-step random walk, when starting at vertex p. Then, it holds for any integer $\gamma \geqslant 1$ that*

$$r^{\gamma \cdot p^2}_{p,p} \geqslant \frac{1}{2p} \cdot 12^{-8 \cdot \gamma}.$$

With the previous lemmas at hand, we are now ready to prove Theorem 4.1.

---

[2] This matrix is called a "substochastic" matrix [1, Section 3.6.5, page 95]. Note that this is *not* a transition matrix, since from state 1 with probability 1/4 the walk gets killed.

[3] Since the random walk gets killed at vertex 0, $\|r^t_{p,\cdot}\|_1$ may generally not be equal to 1, but it is upper bounded by 1.

Proof of Theorem 4.1. Given the two integers ℓ and β being a power of 2, we will now instantiate a graph $G = G[p, D]$. Firstly, the length of the path is $p \triangleq \ell^\delta$ and $\delta \triangleq 1/2 - \varepsilon$; recall that

$$\varepsilon = \min\left(1/2 \cdot \log(1/16 \cdot \log_{12}\beta)/\log \ell, 1/7\right).$$

Secondly, the depth of the binary tree is $D \triangleq \log_2 \beta$ (so in turn, the binary tree has $2^{\log_2(\beta)+1} - 1$ vertices). Hence, the total number of vertices in G is

$$n = \ell^\delta + 2^{\log_2(\beta)+1} - 1,$$

which is at least $2\beta + 1$ and at most $2\beta + \beta^{1/10}$, as $\ell \leqslant \beta^{1/5}$.

Our objective is to show that the expected support of a closed random walk of length ℓ starting at vertex r is at most $3 \cdot \ell^\delta = 3 \cdot \ell^{1/2-\varepsilon}$. We first upper bound the probability of the event that a random walk (not necessarily closed) visits at least $2\ell^\delta + 1$ vertices in ℓ steps and then returns to r.

In the following, let us define the following events and stopping time:

1. The event $\mathcal{A} := \left\{\mathsf{supp}(\ell) \geqslant 2\ell^\delta + 1\right\}$, meaning the random walk of length ℓ visits at least $2\ell^\delta + 1$ vertices.
2. The event $\mathcal{B} := \left\{X^\ell = r\right\}$, meaning the random walk is at the start vertex at time ℓ.
3. The event $\mathcal{C}$ which occurs if the random walk of length ℓ makes at least ℓ^δ transitions on the binary tree.
4. The stopping time τ, which is the number of transitions on the binary tree until a vertex at level $\min\{D, \ell^\delta/12\}$ is reached for the first time.
5. The event $\mathcal{C}(\tau)$ which occurs if the random walk of length ℓ makes at least τ transitions on the binary tree.

In order to visit at least $2\ell^\delta+1$ vertices, the walk needs to visit at least $2\ell^\delta+1-(\ell^\delta+1) = \ell^\delta$ vertices on the tree (excluding the vertex $z = v_0$), since there are only $\ell^\delta + 1$ vertices on the path. Hence we have

$$\mathcal{A} \subseteq \mathcal{C}. \tag{3}$$

Furthermore, we have by Lemma 4.2 (applied to a random walk on the binary tree with ℓ^δ transitions),

$$\mathbf{Pr}\left[\,\tau \leqslant \ell^\delta\,\right] \geqslant 1 - \exp(-\ell^\delta/288). \tag{4}$$

Furthermore, let $T(\tau)$ be the time-step of the random walk on G when the τ-th transition on the binary tree is made; so, $T(\tau) \geqslant \tau$. Then,

$$\mathbf{Pr}\left[\,X^\ell = r \;\middle|\; \mathfrak{F}^{T(\tau)}, T(\tau) \leqslant \ell\right] \leqslant \ell \cdot \max\left(2^{-D}, 2^{-\ell^\delta/12}\right), \tag{5}$$

since by Lemma 4.3, the random walk starting from a vertex at level level $\min\{D, \ell^\delta/12\}$ in the binary tree does not even reach $z = v_0$ within ℓ additional steps (and therefore cannot reach the vertex r at step ℓ). By combining the last three inequalities,

$$
\begin{aligned}
&\mathbf{Pr}\left[\mathcal{A}\cap\{X^\ell=r\}\right]\\
&\leqslant \mathbf{Pr}\left[\mathcal{C}\cap\{X^\ell=r\}\right]\\
&\leqslant \mathbf{Pr}\left[\mathcal{C}\cap\{\tau\leqslant\ell^\delta\}\cap\{X^\ell=r\}\right]+\mathbf{Pr}\left[\tau>\ell^\delta\right]\\
&\leqslant \mathbf{Pr}\left[\mathcal{C}(\tau)\cap\{\tau\leqslant\ell^\delta\}\cap\{X^\ell=r\}\right]+\exp\left(-\ell^\delta/288\right)\\
&\leqslant \mathbf{Pr}\left[\tau\leqslant\ell^\delta\right]\cdot\mathbf{Pr}\left[\mathcal{C}(\tau)\mid\tau\leqslant\ell^\delta\right]\cdot\mathbf{Pr}\left[X^\ell=r\mid\mathcal{C}(\tau),\tau\leqslant\ell^\delta\right]+\exp\left(-\ell^\delta\right)\\
&\leqslant \mathbf{Pr}\left[X^\ell=r\mid\mathfrak{F}^{T(\tau)},T(\tau)\leqslant\ell\right]+\exp\left(-\ell^\delta/288\right)\\
&\leqslant \ell\cdot\max\left(2^{-D},2^{-\ell^\delta/12}\right)+\exp\left(-\ell^\delta/288\right)\\
&\leqslant 16\ell\cdot\max\left(\exp(-\ell^\delta/288),1/\beta\right)=:p_{\mathsf{bad}}.
\end{aligned}
$$

On the other hand, we will now lower bound the probability that a random walk starting at r never leaves the path of length ℓ^δ and is located at vertex r at step ℓ. By Lemma 4.4, we have that this probability is lower bounded by

$$
\mathbf{Pr}\left[\{\mathsf{supp}(\ell)\leqslant 2\ell^\delta\}\cap\{X^\ell=r\}\right]\geqslant\frac{1}{2}\cdot\ell^{-1/2+\varepsilon}\cdot 12^{-8\ell^{2\varepsilon}}\geqslant\ell^{-1/2}\cdot 12^{-8\ell^{2\varepsilon}}=:p_{\mathsf{good}}.
$$

Finally, we can now upper bound the expected size of the support of a closed random walk of length ℓ. Using the conditional probabilities and the definitions of p_{bad} and p_{good}, we have that

$$
\begin{aligned}
\frac{\mathbf{Pr}\left[\mathsf{supp}(\ell)\geqslant 2\ell^\delta+1\mid X^\ell=r\right]}{\mathbf{Pr}\left[\mathsf{supp}(\ell)\leqslant 2\ell^\delta+1\mid X^\ell=r\right]}&=\frac{\mathbf{Pr}\left[\{\mathsf{supp}(\ell)\geqslant 2\ell^\delta+1\}\cap\{X^\ell=r\}\right]}{\mathbf{Pr}\left[\{\mathsf{supp}(\ell)\leqslant 2\ell^\delta+1\}\cap\{X^\ell=r\}\right]}\leqslant\frac{p_{\mathsf{bad}}}{p_{\mathsf{good}}}\\
&\leqslant 16\ell^{1.5}\cdot\max\left(\exp(-\ell^\delta/288),1/\beta\right)\cdot 12^{8\ell^{2\varepsilon}},
\end{aligned}
$$

which implies that

$$
\mathbf{Pr}\left[\mathsf{supp}(\ell)\geqslant 2\ell^\delta+1\mid X^\ell=r\right]\leqslant 16\ell^{1.5}\cdot\max\left(\exp(-\ell^\delta/288),1/\beta\right)\cdot 12^{8\ell^{2\varepsilon}}.
$$

Therefore, it holds that

$$
\begin{aligned}
&\mathbf{E}\left[\mathsf{supp}(\ell)\mid X^\ell=r\right]\\
&\leqslant\mathbf{Pr}\left[\mathsf{supp}(\ell)<2\ell^\delta+1\mid X^\ell=r\right]\cdot 2\ell^\delta+\mathbf{Pr}\left[\mathsf{supp}(\ell)\geqslant 2\ell^\delta+1\mid X^\ell=r\right]\cdot\ell\\
&\leqslant 2\ell^\delta+1+16\ell^{2.5}\cdot\max\left(\exp(-\ell^\delta/288),1/\beta\right)\cdot 12^{8\ell^{2\varepsilon}}.
\end{aligned}
$$

Since $\varepsilon=\min\left(1/2\cdot\log(1/16\cdot\log_{12}\beta)/\log\ell,1/7\right)$ by definition, we have $\varepsilon\leqslant 1/7$. Together with $\delta=1/2-\varepsilon$, this implies $2\varepsilon\leqslant\delta-1/14$ and therefore

$$
12^{8\ell^{2\varepsilon}}\leqslant 12^{8\ell^{\delta-1/14}}\leqslant\exp\left(\ell^\delta/572\right),
$$

where the last inequality holds if ℓ is lower bounded by a sufficiently large constant $C>0$. Applying this gives us that

$$
16\ell^{2.5}\cdot\exp(-\ell^\delta/288)\cdot 12^{8\ell^{2\varepsilon}}\leqslant 16\ell^{2.5}\cdot\exp\left(-\ell^\delta/572\right).
$$

Similarly, we have $\varepsilon\leqslant 1/2\cdot\log(1/16\cdot\log_{12}\beta)/\log\ell$ by definition, and obtain

$$
16\ell^{2.5}\cdot\frac{1}{\beta}\cdot 12^{8\ell^{2\varepsilon}}\leqslant 16\ell^{2.5}\cdot\frac{1}{\beta}\cdot\sqrt{\beta}.
$$

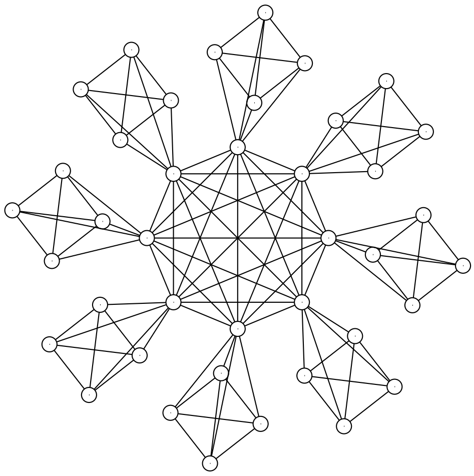

Figure 3 The construction of the graph G, where a large clique of $\beta/\lceil\log\log\beta\rceil$ vertices is connected to $\beta/\lceil\log\log\beta\rceil$ small cliques of size $\lceil\log\log\beta\rceil$ each; recall that our choice of β ensures that $\beta/\lceil\log\log\beta\rceil$ is an integer.

Combining the last two inequalities gives us that

$$\mathbf{E}\left[\mathsf{supp}(\ell) \mid X^s = r\right] \leqslant 2\ell^\delta + 1 + 16\ell^{2.5} \cdot \max\left(\exp\left(-\ell^\delta/572\right), \frac{1}{\sqrt{\beta}}\right) \leqslant 3\ell^\delta,$$

using that $\ell \geqslant C$ for some large constant $C > 0$ as well as $\ell \leqslant \beta^{1/5}$. This completes the proof. ◀

4.2 Proof of Theorem 1.3

In this subsection we present a more detailed formulation of Theorem 1.3, and prove the statement afterwards.

▶ **Theorem 4.5** (Formal version of Theorem 1.3). *Let $C \geqslant 1$ be a constant. Then, for every integer $\beta \geqslant C$ such that $\beta/\lceil\log\log\beta\rceil$ is an integer, there is a graph G with $n = \beta + \beta/\lceil\log\log\beta\rceil$ vertices such that a lazy random walk of length $\ell = \lfloor\log\beta\rfloor$ starting from some vertex chosen uniformly at random from $V(G)$ satisfies*

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 \sim \mathcal{U}\right] \leqslant 5\log\ell.$$

In particular, there is a start vertex $r \in V$ such that

$$\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = r\right] \leqslant 5\log\ell.$$

We first define the graphs used in proving Theorem 4.5. For any given parameter $\beta \in \mathbb{N}$, our graph G is defined as follows:

- Let $G_1 = (V_1, E_1)$ consist of $\beta/\lceil\log\log\beta\rceil$ disjoint, “small” cliques of size $\lceil\log\log\beta\rceil$ each.
- Let $G_2 = (V_2, E_2) = K_{\beta/\lceil\log\log\beta\rceil}$ be a “big” clique of size $\beta/\lceil\log\log\beta\rceil$.
- Our studied graph G is constructed by taking the union of G_1 and G_2, and additionally connecting each vertex of the smaller cliques to one distinct vertex in the big clique.

See Figure 3 for an illustration of our construction.

Proof of Theorem 4.5. We use the graph G defined above in the proof, and decompose the expected support based on the sampled start vertex according to the uniform distribution over $V(G)$:

$$\begin{aligned}
&\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 \sim \mathcal{U}\right] \\
&= \sum_{u \in V} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u\right] \cdot \mathbf{Pr}\left[X^0 = u \mid X^\ell = X^0\right] \\
&= \sum_{u \in V} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = X^0 = u\right]}{\mathbf{Pr}\left[X^\ell = X^0\right]} \\
&= \sum_{u \in V} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = u \mid X^0 = u\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[X^\ell = X^0\right]}
\end{aligned}$$

Splitting the above sum yields

$$\begin{aligned}
&\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 \sim \mathcal{U}\right] \\
&= \sum_{u \in V(G_1)} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = u \mid X^0 = u\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[X^\ell = X^0\right]} \\
&\qquad + \sum_{v \in V(G_2)} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = v\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = v \mid X^0 = v\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[X^\ell = X^0\right]} \\
&= |V(G_1)| \cdot \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = u \mid X^0 = u\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[X^\ell = X^0\right]} \\
&\qquad + |V(G_2)| \cdot \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = v\right] \cdot \frac{\mathbf{Pr}\left[X^\ell = v \mid X^0 = v\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[X^\ell = X^0\right]}, \qquad (6)
\end{aligned}$$

where u is an arbitrary vertex in G_1, v is an arbitrary vertex in G_2, and the last equation holds by symmetry.

Consider now a lazy random walk which starts from some vertex $u \in V_1$ in a small clique. Then the probability that this random walk never leaves the small clique and is at u in step $\ell \triangleq \lfloor \log \beta \rfloor$ is at least

$$\begin{aligned}
&\mathbf{Pr}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \cap \{X^\ell = u\} \mid X^0 = u\right] \\
&\geqslant \left(1 - \frac{1}{2\lceil \log\log\beta \rceil}\right)^{\lfloor \log\beta \rfloor - 1} \cdot \min\left(\frac{1}{2}, \frac{1}{2\lceil \log\log\beta \rceil}\right) \\
&\geqslant 8^{-\log\beta/\log\log\beta}. \qquad (7)
\end{aligned}$$

If the random walk leaves a small clique, then the probability of returning to the small clique within ℓ steps is at most $\ell \cdot (\lceil \log\log\beta \rceil)^2/n$; hence, it holds that

$$\begin{aligned}
\mathbf{Pr}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \cap \{X^\ell = u\} \mid X^0 = u\right] &\leqslant \ell \cdot \frac{\lceil \log\log\beta \rceil}{\beta/\lceil \log\log\beta \rceil - 1 + \lceil \log\log\beta \rceil} \\
&\leqslant \ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta}.
\end{aligned}$$

Therefore, as in the proof of Theorem 4.1,

$$\frac{\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \;\middle|\; X^\ell = X^0 = u\,\right]}{\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \leqslant \lceil \log\log\beta \rceil \;\middle|\; X^\ell = X^0 = u\,\right]}$$
$$= \frac{\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \cap \left\{X^\ell = u\right\} \;\middle|\; X^0 = u\,\right]}{\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \leqslant \lceil \log\log\beta \rceil \cap \left\{X^\ell = u\right\} \;\middle|\; X^0 = u\,\right]}$$
$$\leqslant \frac{\ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta}}{8^{-\log\beta/\log\log\beta}},$$

and upper bounding the denominator on the left hand side by 1 yields

$$\mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \;\middle|\; X^\ell = u\,\right] \leqslant \frac{\ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{n}}{8^{-\log\beta/\log\log\beta}}.$$

Therefore, it holds that

$$\begin{aligned}
&\mathbf{E}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^\ell = X^0 = u\,\right] \\
&\leqslant \mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \geqslant \lceil \log\log\beta \rceil \;\middle|\; X^0 = u\,\right] \cdot \ell \\
&\qquad + \mathbf{Pr}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \leqslant \lceil \log\log\beta \rceil \;\middle|\; X^0 = u\,\right] \cdot \lceil \log\log\beta \rceil \\
&\leqslant \ell^2 \cdot 8^{\log\beta/\log\log\beta} \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta} + 1 \cdot \lceil \log\log\beta \rceil \leqslant 2 \cdot \lceil \log\log\beta \rceil,
\end{aligned}$$

as $\ell = \lfloor \log\beta \rfloor$.

Now we return to (6). By using the three trivial estimates, (i) $|V_1| \leqslant n$, (ii) $|V_2| \leqslant n$ and (iii) $\mathbf{E}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = v\,\right] \leqslant \ell$, we have

$$\begin{aligned}
&\mathbf{E}\left[\,\mathsf{supp}_{\mathbf{P}}(\ell) \;\middle|\; X^\ell = X^0 \sim \mathcal{U}\,\right] \\
&\leqslant n \cdot 2 \cdot \lceil \log\log\beta \rceil \cdot \frac{\mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]} + n \cdot \ell \cdot \frac{\mathbf{Pr}\left[\,X^\ell = v \mid X^0 = v\,\right] \cdot \frac{1}{n}}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]} \\
&\leqslant 2 \cdot \lceil \log\log\beta \rceil \cdot \frac{\mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right]}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]} + \ell \cdot \frac{\mathbf{Pr}\left[\,X^\ell = v \mid X^0 = v\,\right]}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]}.
\end{aligned} \tag{8}$$

We now proceed to upper bound the expression in Equation (8), by considering the two addends separately. We first upper bound $\frac{\mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right]}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]}$. By decomposing and lower bounding the denominator, we have that

$$\begin{aligned}
\mathbf{Pr}\left[\,X^\ell = X^0\,\right] &\geqslant \sum_{u \in V_1} \mathbf{Pr}\left[\,X^0 = u\,\right] \cdot \mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right] \\
&\geqslant \frac{1}{2} \cdot \mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right]
\end{aligned} \tag{9}$$

since the probability $\mathbf{Pr}\left[\,X^\ell = X^0 \mid X^0 = u\,\right]$ is the same for all $u \in V_1$ by symmetry, and by construction of G, at least half of the vertices in G are in V_1. Therefore,

$$\frac{\mathbf{Pr}\left[\,X^\ell = u \mid X^0 = u\,\right]}{\mathbf{Pr}\left[\,X^\ell = X^0\,\right]} \leqslant 2. \tag{10}$$

We now turn to the second addend in (8). We first upper bound $\mathbf{Pr}\left[\,X^\ell = v \mid X^0 = v\,\right] = p^\ell_{v,v}$, where $v \in V_2$. To this end, note that the random walk can only be at v at step ℓ if at least one of the following three cases occurs: (i) the random walk always remains on v by taking ℓ self-loops, (ii) the random walk leaves v, and then returns to v from another vertex in the big clique, and (iii) the random walk leaves v, and then returns to v from a neighbor in the small clique. Regarding (i), the probability is $2^{-\ell}$. Regarding (ii), the probability of ever using an edge $\{x, v\} \in E$ with $x \in V_2$ during ℓ steps is at most

$$\ell \cdot \frac{1}{2\deg(x)} \leqslant \ell \cdot \frac{1}{2 \cdot \left(\frac{\beta}{\lceil \log\log\beta \rceil} - 1 + \lceil \log\log\beta \rceil\right)} \leqslant \ell \cdot \frac{\lceil \log\log\beta \rceil}{2\beta}.$$

Finally, regarding (iii), the probability that the random walk ever reaches any vertex in V_1 within ℓ steps is upper bounded by

$$\ell \cdot \max_{z \in V_2} \frac{\deg_{V_1}(z)}{2\deg(z)} \leqslant \ell \cdot \frac{\lceil \log\log\beta \rceil}{2 \cdot \left(\frac{\beta}{\lceil \log\log\beta \rceil} - 1 + \lceil \log\log\beta \rceil\right)} \leqslant \ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{2\beta}.$$

Combining these three cases, we have

$$\mathbf{Pr}\left[X^\ell = v \mid X^0 = v \right] \leqslant 2^{-\ell} + \ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta}. \tag{11}$$

To lower bound $\mathbf{Pr}\left[X^\ell = X^0 \right]$, we apply (9) and the estimate from (7) to obtain that

$$\mathbf{Pr}\left[X^\ell = X^0 \right] \geqslant \frac{1}{2} \cdot \mathbf{Pr}\left[X^\ell = u \mid X^0 = u \right] \geqslant \frac{1}{2} \cdot 8^{-\log\beta/\log\log\beta}. \tag{12}$$

Finally, combining (10), (11), (12) with (8) gives us that

$$\begin{aligned}
&\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 \sim \mathcal{U} \right] \\
&\leqslant 4 \cdot \lceil \log\log\beta \rceil + 2 \cdot \ell \cdot 8^{\log\beta/\log\log\beta} \cdot \left(2^{-\ell} + \ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta} \right) \\
&\leqslant 4 \cdot \lceil \log\log\beta \rceil + 2 \cdot \ell \cdot 8^{\log\beta/\log\log\beta} \cdot \left(2 \cdot \ell \cdot \frac{(\lceil \log\log\beta \rceil)^2}{\beta} \right) \\
&\leqslant 4 \cdot \lceil \log\log\beta \rceil + o(1) \\
&\leqslant 5 \cdot \log \ell,
\end{aligned}$$

where in the second inequality we used the definition $\ell = \lfloor \log\beta \rfloor \geqslant \log\beta - 1$, and similarly the third inequality holds since $\ell = \lfloor \log\beta \rfloor$, and thus the $1/\beta$ term dominates.

For the second statement, note that by conditioning on the vertex sampled for X^0,

$$\begin{aligned}
&\mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 \sim \mathcal{U} \right] \\
&= \sum_{u \in V} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u \right] \cdot \mathbf{Pr}\left[X^\ell = X^0 = u \mid X^\ell = X^0 \right] \\
&\geqslant \min_{u \in V} \mathbf{E}\left[\mathsf{supp}_{\mathbf{P}}(\ell) \mid X^\ell = X^0 = u \right],
\end{aligned}$$

and thereore the second statement follows immediately from the first statement. ◀

5 Results on Eigenvalue Multiplicity

This section presents some new eigenvalue multiplicity bounds on the transition matrices of random walks. For ease of presentation, we focus on lazy random walks in this section, but our presented method can be employed to analyze non-lazy random walks, too. Our first eigenvalue multiplicity bound is as follows:

▶ **Theorem 5.1** (Formal Statement of Theorem 1.4)**.** *Consider any connected, n-vertex graph $G = (V, E)$ with minimum degree δ and maximum degree Δ. Further assume $\lambda = \lambda_2 \geqslant \gamma$, where $\gamma \triangleq 1 - \frac{\delta}{32c\Delta} \cdot \frac{1}{\log^4 n}$ and $c \geqslant 1$ is the constant from Lemma 3.2. Then, it holds that*

$$M_{\mathbf{P}}\left[\left(1 - \frac{\delta}{32c\Delta \cdot \log^5 n}\right) \cdot \lambda, \lambda\right] = O\left(\frac{n}{\log n}\right). \tag{13}$$

It is shown in [16] that, for the normalized adjacency matrix of any n-vertex graph G with maximum degree Δ, the number of eigenvalues in the range $\left[\left(1 - \frac{\log\log_\Delta}{\log_\Delta n}\right) \cdot \lambda_2, \lambda_2\right]$ is

$$\widetilde{O}\left(n \cdot \frac{\Delta^{7/5}}{\log^{1/5} n}\right). \tag{14}$$

In comparison to their bound, our presented result only holds for graphs with poor expansion. However, Theorem 5.1 does show for such graphs that the number of eigenvalues in our studied range is $O(n/\log n)$, which is significantly smaller than the bound in (14).

The proof of Theorem 5.1 closely follows the approach in [16], by reducing the multiplicity analysis to the support of closed random walks. We consider a lazy random walk of length ℓ on G, and assume that the start vertex X^0 is sampled uniformly at random. We denote by W^ℓ the event $\{X^\ell = X^0\}$ and by $W^{\ell,s}$ the event $\{X^\ell = X^0, \mathsf{supp}(\ell) \leqslant s\}$; that is, the random walk is closed and has support at most s. Abusing notation a bit, let $W^{\ell,\geqslant s}$ be the event where the random walk is closed and has support at least s. We now state the following bound, which is based on the arguments of [16] and the probability bound in Theorem 3.1.

▶ **Lemma 5.2.** *Consider any connected, n-vertex graph, and a lazy random walk of length $\ell \leqslant 32c\mu \cdot \frac{\Delta}{\delta} n^2$ starting from a uniform random vertex. Then with $c > 0$ being the constant from Lemma 3.2, it holds for any $s \leqslant \left\lfloor \sqrt{\frac{1}{64c} \cdot \frac{\delta}{\Delta} \cdot \frac{\ell}{\mu}} \right\rfloor$ that*

$$\mathbf{Pr}\left[W^{\ell,s}\right] \leqslant \frac{\Delta}{\delta} \cdot n \cdot \left(\frac{5}{8}\right)^{\mu/2} \mathbf{Pr}\left[W^\ell\right].$$

Combining Lemma 5.2 with the techniques developed in [16] proves Theorem 5.1.

We further present a different and more elementary approach to bound the multiplicities of the eigenvalue λ_2, and our proof is based on the Random Target Lemma [14, (3.3)].

▶ **Theorem 5.3.** *Consider any connected, n-vertex graph $G = (V, E)$ with average degree d and minimum degree δ. Then there is some constant $C > 0$, such that it holds for any $\varepsilon > 0$ that*

$$M_{\mathbf{P}}[(1-\varepsilon)\lambda_2, \lambda_2] \leqslant C \cdot \frac{d}{\delta} \cdot n \cdot \frac{1 - (1-\varepsilon)\lambda_2}{\sqrt{1 - \lambda_2}}.$$

In particular, it holds with $\varepsilon = (1 - \lambda_2)/\lambda_2$ that

$$M_{\mathbf{P}}[(1-\varepsilon)\lambda_2, \lambda_2] \leqslant 2C \cdot \frac{d}{\delta} \cdot n \cdot \sqrt{1 - \lambda_2}.$$

6 Conclusions

In this work we analyze the support of closed random walks of length ℓ on different graph classes. Contrary to the well-understood worst-case support of *open* random walks, especially on regular and bounded-degree graphs, our results demonstrate that the (worst-case) support of *closed* random walks is much more complex, and undergoes a delicate phase transition as ℓ varies. While the support is $\Theta(\ell^{1/2}/\sqrt{\log n})$ for $\ell = \Omega((\log n)^{7/2})$, for smaller values of ℓ it is sandwiched between $\Omega(\ell^{1/5})$ and $O(\ell^{5/14})$. This proves a strong separation from the open random walk case, where the support is known to be $\Omega(\ell^{1/2})$ [3, 7], and provides a negative answer to [16, Open Problem 3].

For *(highly) irregular* graphs, we prove that even with a *randomly sampled* start vertex, the support may only be logarithmic in ℓ. This is once more in sharp contrast to open walks, where a lower bound of $\Omega(\ell^{1/3})$ holds for *any* start vertex and any $1 \leqslant \ell \leqslant n^3$ [3].

One interesting open problem is to derive refined bounds on the support of closed random walks. For instance, it is tempting to conjecture that on any bounded-degree expander graph, the lower bound on the support can be improved, possibly even to $\Omega(\ell)$, which would be tight and match the bound for open random walks. To the best of our knowledge, this is only known for the special case where ℓ is upper bounded by the girth of the expander [5].

References

1 David Aldous and James A. Fill. Reversible markov chains and random walks on graphs, 2002. Unpublished. `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

2 Noga Alon, Chen Avin, Michal Koucký, Gady Kozma, Zvi Lotker, and Mark R. Tuttle. Many random walks are faster than one. *Combinatorics, Probability and Computing*, 20(4):481–502, 2011. doi:10.1017/S0963548311000125.

3 Greg Barnes and Uriel Feige. Short random walks on graphs. *SIAM Journal on Discrete Mathematics*, 9(1):19–28, 1996. doi:10.1137/S0895480194264988.

4 Anna Ben-Hamou, Roberto I. Oliveira, and Yuval Peres. Estimating graph parameters via random walks with restarts. In *29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'18)*, pages 1702–1714, 2018.

5 Itai Benjamini, Roey Izkovsky, and Harry Kesten. On the Range of the Simple Random Walk Bridge on Groups. *Electronic Journal of Probability*, 12:591–612, 2007. doi:10.1214/EJP.v12-408.

6 Artur Czumaj and Christian Sohler. Testing expansion in bounded-degree graphs. *Combinatorics, Probability and Computing*, 19(5-6):693–709, 2010. doi:10.1017/S096354831000012X.

7 Uriel Feige. A spectrum of time–space trade-offs for undirected *s*-*t* connectivity. *Journal of Computer and System Sciences*, 54(2):305–316, 1997. doi:10.1006/jcss.1997.1471.

8 George Giakkoupis, Frederik Mallmann-Trenn, and Hayk Saribekyan. How to spread a rumor: Call your neighbors or take a walk? In *38th Annual ACM-SIGOPT Principles of Distributed Computing (PODC'19)*, pages 24–33, 2019.

9 Milan Haiman, Carl Schildkraut, Shengtong Zhang, and Yufei Zhao. Graphs with high second eigenvalue multiplicity. *Bulletin of the London Mathematical Society*, pages 1–23, 2022. doi:10.1112/blms.12647.

10 Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *BAMS: Bulletin of the American Mathematical Society*, 43, 2006.

11 Zilin Jiang, Jonathan Tidor, Yuan Yao, Shengtong Zhang, and Yufei Zhao. Equiangular lines with a fixed angle. *Annals of Mathematics*, 194(3):729–743, 2021.

12 James R. Lee, Shayan Oveis Gharan, and Luca Trevisan. Multiway spectral partitioning and higher-order cheeger inequalities. *Journal of the ACM*, 61(6):37:1–37:30, 2014.

13 David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2009. With a chapter by James G. Propp and David B. Wilson.

14 Lásló Lovász. Random walks on graphs: A survey. *Combinatorics, Paul Erdős is Eighty*, 2:1–46, 1993.

15 Theo McKenzie. *Random Walks and Delocalization through Graph Eigenvector Structure*. PhD thesis, UC Berkeley, 2022.

16 Theo McKenzie, Peter Michael Reichstein Rasmussen, and Nikhil Srivastava. Support of closed walks and second eigenvalue multiplicity of graphs. In *53rd Annual ACM Symposium on Theory of Computing (STOC'21)*, pages 396–407, 2021.

17 Roberto I. Oliveira. On the coalescence time of reversible random walks. *Transactions of the American Mathematical Society*, 364(4):2109–2128, 2012.

18 Roberto I. Oliveira and Yuval Peres. Random walks on graphs: new bounds on hitting, meeting, coalescing and returning. In *16th Workshop on Analytic Algorithmics and Combinatorics (ANALCO '19)*, pages 119–126, 2019.

19 Thomas Sauerwald and He Sun. Tight bounds for randomized load balancing on arbitrary network topologies. In *53th Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, pages 341–350, 2012. doi:10.1109/FOCS.2012.86.

20 Daniel A. Spielman and Shang-Hua Teng. A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning. *SIAM Journal on Computing*, 42(1):1–26, 2013.

A Auxiliary Tools

This section lists the auxiliary results used in the paper. Our first lemma is a simple upper bound on the lower tails of a sub-multiplicative random variable. We remark that this is a standard result, however we present the proof here for the sake of completeness.

▶ **Lemma A.1.** *Let X be a non-negative integer random variable such that $\mathbf{E}[X] \geqslant b$, and there exists integer $c \geqslant 1$ such that $\mathbf{Pr}[X > kc] \leqslant \mathbf{Pr}[X > c]^k$ for all integers $k \geqslant 0$. Then, it holds for any $a < c$ that*

$$\mathbf{Pr}[X > a] \geqslant \frac{b-a}{b+2c}.$$

Proof. Let $p \triangleq \mathbf{Pr}[X > a] \geqslant \mathbf{Pr}[X > c]$, since $a < c$ and $\mathbf{Pr}[X > x]$ decreasing in x. Hence, it holds that

$$\begin{aligned} b &\leqslant \mathbf{E}[X] \\ &= \sum_{i=0}^{\infty} \mathbf{Pr}[X > i] \\ &\leqslant a + \sum_{i=a}^{c-1} \mathbf{Pr}[X > i] + c\sum_{k=1}^{\infty} \mathbf{Pr}[X > kc] \\ &\leqslant a + \sum_{i=a}^{c-1} \mathbf{Pr}[X > i] + c\sum_{k=1}^{\infty} \mathbf{Pr}[X > c]^k \\ &\leqslant a + p(c-a) + cp/(1-p). \end{aligned}$$

This implies $(1-p)b \leqslant a + 2pc$, rearranging gives the result. ◀

▶ **Theorem A.2** (Cauchy's Interlacing Theorem). *Let $\mathbf{A}$ be a real symmetric $n \times n$ matrix, and $\mathbf{B}$ an $m \times m$ principal submatrix of $\mathbf{A}$ (that is, $\mathbf{B}$ is obtained by deleting both the i^{th} row and column for some values of i). Suppose A has eigenvalues $\lambda_1, \ldots, \lambda_n$, and B has eigenvalues $\beta_1, \ldots, \beta_m$. Then, it holds for $1 \leqslant k \leqslant m$ that $\lambda_k \leqslant \beta_k \leqslant \lambda_{k+n-m}$.*

▶ **Theorem A.3** (Hoeffding's Bound). *Let* $Y_1, \ldots, Y_\ell$ *be independent bounded random variables with* $Y_i \in [a, b]$ *for each* $i \leqslant \ell$, *and define* $Y = \sum_{i=1}^{\ell} Y_i$. *Then, the following hold for all* $\lambda \geqslant 0$*:*

$$\mathbf{Pr}\left[\, Y_i - \mathbf{E}\left[\, Y \,\right] \geqslant \lambda \,\right] \leqslant \exp\left(-\frac{2\lambda^2}{\ell(b-a)^2}\right)$$

and

$$\mathbf{Pr}\left[\, Y_i - \mathbf{E}\left[\, Y \,\right] \leqslant -\lambda \,\right] \leqslant \exp\left(-\frac{2\lambda^2}{\ell(b-a)^2}\right).$$