Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 116, 1–31.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1149

A local central limit theorem for random walks on expander graphs

Rafael Chiclana* Yuval Peres†

Abstract

There is a long history of establishing central limit theorems for Markov chains. Quantitative bounds for chains with a spectral gap were proved by Mann and refined later. Recently, rates of convergence for the total variation distance were obtained for random walks on expander graphs, which are often used to generate sequences satisfying desirable pseudorandom properties. We prove a local central limit theorem with an explicit rate of convergence for random walks on expander graphs, and derive an improved bound for the total variation distance.

Keywords: Markov chains; expander graphs; central limit theorem.
MSC2020 subject classifications: 05C81; 05C48; 60F05.
Submitted to EJP on August 21, 2023, final version accepted on May 23, 2024.

1 Introduction

Given $\lambda < 1$, a graph is considered to be a *λ-expander* when the absolute value of all the eigenvalues of its transition matrix except 1 are bounded above by λ. Expander graphs have a wide range of applications in areas such as derandomization, complexity theory, and coding theory (see [10]). In particular, random walks on expander graphs are typically used to generate sequences satisfying desirable pseudorandom properties. They serve as an efficient replacement of t independent sample vertices chosen uniformly at random. It is natural to study then how good of a replacement these sequences are, or equivalently, to measure the randomness of random walks on expander graphs. More precisely, consider a *balanced labelling* on a regular graph $G = (V, E)$, that is, a map $\mathrm{val}\colon V \longrightarrow \{0, 1\}$ with $\sum_{v \in V} \mathrm{val}(v) = |V|/2$. Given a test function $f\colon \{0, 1\}^t \longrightarrow \mathbb{R}$, we compare $f(\mathrm{val}(v_0), \dots, \mathrm{val}(v_{t-1}))$ when the vertices $v_0, \dots, v_{t-1}$ are sampled either from a random walk, or independently and uniformly at random. This problem was studied by Guruswami and Kumar in [8] for sticky random walks, and later on, by Cohen, Peri, and Ta-Shma in [3] for general expander graphs. Significant progress can be found also in Cohen et al. [2], Golowich-Vadhan [7], and Golowich [6]. In this paper, we focus on the

*Kent State University, United States. E-mail: chiclan1@msu.edu
†Beijing Institute of Mathematical Sciences and Applications, China. E-mail: yperes@gmail.com

asymptotic behavior of $f(\mathrm{val}(v_0), \dots, \mathrm{val}(v_{t-1}))$ as the size of the sample t grows. Our results answer Question 3 in [3] and Questions 2 and 3 in [2]. Moreover, we improve the bound of the main result of [6] for bounded degree graphs.

Throughout the paper, we assume that all graphs are finite and connected. We write $\mathcal{N}(\mu, \sigma^2)$ for a normal distribution with mean μ and variance σ^2, and ϕ for the density function of $\mathcal{N}(0,1)$. Local central limit theorems for Markov chains are known as early as the work of Kolmogorov [11], and the contributions due to Nagaev [14] and [15], who initiated the study of Markov chains by spectral methods. This topic has been widely studied over the last years. Our main result gives a local central limit theorem for the random walk on expander graphs with a uniform rate of convergence.

Theorem 1.1. *Let $G = (V, E)$ be a d-regular λ-expander graph with $\lambda < 1$. Fix a balanced labelling $\mathrm{val}\colon V \longrightarrow \{0,1\}$, let (X_i) be the simple random walk on G with uniform initial distribution, and let $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. There is a constant $C_1(\lambda, d)$ depending only on λ and d such that*

$$\left| \mathbb{P}\left\{Z_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k - t/2}{t^{1/2}\sigma}\right) \right| \le C_1(\lambda, d)\frac{1}{t} \quad \forall\, k \in \mathbb{Z} \quad \forall\, t \in \mathbb{N}.$$

We obtain this from a general local central limit theorem for Markov chains, given in Section 4.

Remark 1.2. Every d-regular connected graph that is not bipartite is a λ-expander for some $\lambda < 1$. It follows from Theorem 1.1 that the local central limit theorem holds for any of such graphs. However, the time t until Z_t and $\mathcal{N}(t/2, t\sigma^2)$ are close might depend on the number of vertices of G. The fundamental observation in Theorem 1.1 is that the dependence on the size of the graph disappears when G has an absolute spectral gap and bounded degree.

Write (U_i) for a sequence of independent vertices of G chosen uniformly at random. We are particularly interested in the total variation distance between the Hamming weights $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ and $B_t = \sum_{i=0}^{t-1} \mathrm{val}(U_i)$, denoted[1] $\|Z_t - B_t\|_{TV}$. This distance measures the best distinguishing probability a symmetric function can achieve on $(\mathrm{val}(X_i))_{i=0}^{t-1}$ and $(\mathrm{val}(U_i))_{i=0}^{t-1}$ (see Proposition 4.5 in [12]).

Mann showed in [13, Theorem 1] that Z_t converges in Kolmogorov distance to a normal distribution as t grows. One can show that a stronger convergence holds for Z_t as an application of Theorem 1.1. In fact, most of the mass of Z_t is concentrated in an interval of length $\sqrt{t \log t}$ around its mean (see Theorem 2.1 in [5]). Therefore, a local central limit theorem for Z_t implies convergence in total variation distance to a discretized normal distribution. Indeed, since Theorem 1.1 gives a convergence rate of $O(1/t)$, this simple argument gives a rate of convergence of $O(\sqrt{\log t}/\sqrt{t})$ for the total variation distance. A sharper analysis allows to improve this bound to $O(\log(t)^{1/4}/\sqrt{t})$. In an earlier version of this work [1], we demonstrated convergence in total variation distance without bounding the rate of convergence. Subsequently, Golowich presented in [6] a bound similar to (1.2), where $\log(t)^{1/4}$ is replaced by $\log(t)^{\eta_1 \log\log t + \eta_2}$ for some constants η_1, η_2 (see (2.2)).

There are different natural ways of discretizing the normal distribution. For convenience, we will consider $N_d(\mu, \sigma^2)$ with probability distribution $f_{N_d(\mu,\sigma^2)}$ given by

$$f_{N_d(\mu,\sigma^2)}(k) = \frac{1}{D(\mu, \sigma^2)}\sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) \quad \forall\, k \in \mathbb{Z}, \tag{1.1}$$

where $D(\mu, \sigma^2) = \sum_{k\in\mathbb{Z}} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right)$ is a normalizing constant.

[1]The total variation distance is defined between measures. We abuse notation and identify $\|Z_t - B_t\|_{TV}$ with $\|\mathcal{L}(Z_t) - \mathcal{L}(B_t)\|_{TV}$, where $\mathcal{L}(\cdot)$ stands for the law of a random variable.

Corollary 1.3. *Let $G = (V, E)$ be a d-regular λ-expander graph with $\lambda < 1$. Fix a balanced labelling $\mathrm{val}\colon V \longrightarrow \{0,1\}$, let (X_i) be the simple random walk on G with uniform initial distribution, and let $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. There is a constant $C_2(\lambda, d)$ depending only on λ and d such that*

$$\left\|Z_t - N_d(t/2, t\sigma^2)\right\|_{TV} \le C_2(\lambda, d)\frac{\log(t)^{1/4}}{\sqrt{t}} \quad \forall\, t \ge 2. \tag{1.2}$$

On the other hand, the asymptotic behavior of (B_t) is well known. Indeed, $\mathrm{val}(U_i)$ is a Bernoulli random variable with parameter $p = \frac{1}{2}$. Hence, B_t follows a binomial $\mathrm{Bin}(t, \frac{1}{2})$. Since binomial distributions are concentrated around their means (see Lemma 8.1 in [16]), the classic local central limit theorem (see Lemma 5 in [18]) implies that B_t converges in total variation distance to a discretized normal distribution. More precisely,

$$\lim_{t\to\infty} \|B_t - N_d(t/2, t/4)\|_{TV} = 0.$$

In view of Corollary 1.3, we deduce that

$$\lim_{t\to\infty} \|B_t - Z_t\|_{TV} = \|N_d(t/2, t/4) - N_d(t/2, t\sigma^2)\|_{TV}. \tag{1.3}$$

Even though both Z_t and B_t converge to discretized normal distributions with mean $\frac{t}{2}$, their variances may not be the same. Therefore, the ability of a random walk (X_i) on an expander graph to fool all symmetric functions as t grows is measured by the difference between the variances of Z_t and B_t. This difference can be bounded using the following formula for the variance of Z_t, which is written in terms of the eigenvalues λ_j, their normalized eigenvectors f_j (see discussion previous to (3.1)), and the labelling val. We give it also for unbalanced labellings since it will be useful when extending our main results. Balanced labellings correspond to taking $\alpha = 1/2$ below.

Proposition 1.4. *Let G be a d-regular graph with n vertices. Let (X_i) be the simple random walk on G with uniform initial distribution π, and consider $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$. For a labelling $\mathrm{val}\colon V \longrightarrow \{0,1\}$ with $\mathbb{E}_\pi(\mathrm{val}) = \alpha \in (0,1)$, we have*

$$\mathrm{Var}(Z_t) = \alpha(1-\alpha)t + 2\alpha^2 \sum_{k=1}^{t-1}(t-k)\sum_{j=2}^{n} \langle \pi_B, f_j\rangle^2 \lambda_j^k, \tag{1.4}$$

where $B = \{x \in V\colon\ \mathrm{val}(x) = 1\}$ and π_B is the uniform distribution on B. In particular, if G is a λ-expander then we have

$$|\mathrm{Var}(Z_t) - \alpha(1-\alpha)t| \le 2\alpha(1-\alpha)t\frac{\lambda}{1-\lambda}. \tag{1.5}$$

A sample obtained by running a random walk on a λ-expander graph, with small λ, behaves similar to a uniform sample. Indeed, taking $\alpha = 1/2$ in Proposition 1.4 shows that $\mathrm{Var}(Z_t)/t$ converges to $1/4$ as λ converges to zero. In view of (1.3), this implies that $\|B_t - Z_t\|$ goes to zero as $\lambda \to 0$. This provides an alternative proof for a fact first observed in [3], and subsequently optimized in [2], [7], and [6]. It is important to mention that Corollary 1.3 only provides convergence when the size of the sample t grows to infinity, since the constant $C_2(\lambda, d)$ does not converge to zero as λ does. Although the rate of convergence on t obtained in [6] is weaker than the one Corollary 1.3 provides, their bound shows simultaneous convergence as t grows to infinity or λ converges to zero. These contributions will be discussed in more detail in Section 2.

The *sticky random walk* with parameter $p \in (-1, 1)$ is a Markov chain (Q_i) on $\{0,1\}$ defined as follows. The initial state is chosen uniformly at random. At each step, the chain stays at the same state with probability $\frac{1+p}{2}$, and switches states with probability $\frac{1-p}{2}$. This simple chain can be seen as a simplified version of general random walks on expander graphs. Although a sequence of bits generated using the random walk may not fool all symmetric functions as t grows, it serves as a replacement for a sample of bits obtained from the sticky random walk. The following result is an immediate consequence of Corollary 1.3 and a similar result for the sticky random walk (see Lemma 7.3).

Theorem 1.5. *Let $G = (V, E)$ be a d-regular λ-expander graph with $\lambda < 1$. Fix a balanced labelling $\mathrm{val}\colon V \longrightarrow \{0, 1\}$, let (X_i) be the simple random walk on G with uniform initial distribution, $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$, and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. Let (Q_i) be the sticky random walk on $\{0, 1\}$ with parameter $p = \frac{4\sigma^2-1}{4\sigma^2+1}$ and $R_t = \sum_{i=0}^{t-1} Q_i$. There is a constant $C_3(\lambda, d)$ depending only on λ and d such that*

$$\|Z_t - R_t\|_{TV} \le C_3(\lambda, d)\frac{\sqrt{\log t}}{\sqrt{t}} \quad \forall\, t \ge 2. \tag{1.6}$$

We believe that power of the logarithm in (1.6) can be improved with additional work.

The rest of the paper is organized as follows. In Section 3, we introduce notation and definitions that will be used throughout the paper. In Section 2, we discuss significant previous work of several authors on this topic. In section 4, we present a general local central limit theorem for Markov chains. Our main result follows as a particular application of it. Section 5 is dedicated to prove Theorem 1.1. In Section 6, we prove Proposition 1.4 and Corollary 1.3. Section 7 is devoted to prove Theorem 1.5. In Section 8, we extend our main results to unbalanced labellings. Finally, we dedicate Section 9 to prove Example 2.1 and Example 2.2.

2 Previous work

In the recent paper [8], among other results, Guruswami and Kumar showed that the total variation distance between the Hamming weight of the sticky random walk with parameter p and the binomial distribution is $\Theta(p)$. As [3] states, "a major open problem they raise is whether the same is true for random walks on expander graphs". This problem has been studied very recently by several authors. We discuss here the most significant advances on the matter.

Cohen, Peri, and Ta-Shma present in [3] a Fourier-Analytic approach to study random walks on expander graphs. Their main result states that the SRW on λ-expander graphs "fools" symmetric functions for small values of λ. To be more precise, let G be a d-regular λ-expander graph, (X_i) the SRW on G with uniform initial distribution π, and (U_i) a sequence of independent vertices of G chosen uniformly at random. Theorem 1.1 in [3] states that for any $t \in \mathbb{N}$, any symmetric function $f\colon \{0, 1\}^t \longrightarrow \{0, 1\}$, and any balanced labelling $\mathrm{val}\colon V \longrightarrow \{0, 1\}$, we have

$$|\mathbb{E}_\pi(f(\mathrm{val}(X_0), \ldots, \mathrm{val}(X_{t-1}))) - \mathbb{E}_\pi(f(\mathrm{val}(U_0), \ldots, \mathrm{val}(U_{t-1})))| = O(\lambda \cdot \log^{3/2}(1/\lambda)).$$

This result can be rewritten in terms of the total variation distance between the Hamming weights $Z_t = \sum_{k=0}^{t-1} \mathrm{val}(X_k)$ and $B_t = \sum_{k=0}^{t-1} \mathrm{val}(U_k)$ as follows.

$$\|Z_t - B_t\|_{TV} = O(\lambda \cdot \log^{3/2}(1/\lambda)). \tag{2.1}$$

The authors then propose several open questions. First, they ask if (2.1) holds for unbalanced labellings. They also ask whether the above bound is sharp. These questions are addressed by Theorem 3 in [2]. It states that for any labelling $\mathrm{val}\colon V \longrightarrow \{1, -1\}$ with $\mathbb{E}_\pi(\mathrm{val}) = \alpha \in (-1, 1)$ and $0 < \lambda < \frac{1-|\alpha|}{128e}$ we have

$$\|Z_t - B_t\|_{TV} \le \frac{124}{\sqrt{1-|\alpha|}}\lambda.$$

An equivalent bound is achieved by Corollary 2 in [7], which also provides interesting bounds for the tails of the distributions. Moreover, Corollary 4 in [7] extends these bounds from binary to arbitrary labellings. Finally, the authors show that the dependence on λ in the above results is sharp up to a constant (see Theorem 5 in [7]).

Second, while (2.1) shows that the total variation distance between Z_t and B_t vanishes with λ, it leaves open the possibility that a better convergence exists, namely, for some fixed λ the total variation distance goes to 0 as the size of the sample t grows. This is the case for some well-known symmetric test functions, such as AND, OR, and PARITY, where the error decreases exponentially with t (see [3] for details), and MAJ, where the error goes down polynomially with t (see Theorem 4.6 in [3]). This question is addressed by Theorem 1 in [2], which shows that for every λ there is a λ-expander graph and a balanced labelling $\mathrm{val}\colon V \longrightarrow \{1,-1\}$ such that

$$\|Z_t - B_t\|_{TV} = \Theta(\lambda) \quad \forall\, t \in \mathbb{N}.$$

However, this estimate is obtained using Cayley graphs over Abelian groups, which cannot provide constant degree expanders. An open question that the authors in [2] propose is whether a similar bound holds for constant degree graphs. They also ask about the existence of a family of expander graphs that fools all symmetric functions with error going down to zero as the length of the walk t grows, independently of the chosen labelling. Corollary 1.3 answers both of these questions. It shows that the total variation distance between Z_t and B_t converges to zero as t grows if, and only if, the limits of the variances of $t^{-1/2}Z_t$ and $t^{-1/2}B_t$ are the same. The following examples, that we justify in Section 9, show that in general this is not the case. Recall that $\mathrm{Var}(B_t) = t/4$.

Example 2.1. Let K_4 be the complete graph with 4 nodes, let val be any balanced labelling on K_4, let (X_i) be the simple random walk on K_4 with uniform initial distribution, and consider $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$. Then K_4 is a 3-regular $\frac{1}{3}$-expander graph and

$$\mathrm{Var}(Z_t) = \frac{1}{8}t + O(1).$$

Example 2.2. Let G be a d-regular λ-expander graph with n vertices and let (X_i) be the simple random walk on G. There is a balanced labelling val on G for which $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ satisfies

$$\mathrm{Var}(Z_t) \geq \frac{t}{4} + \frac{1}{2}\left(\frac{1}{d} - \frac{3}{n-1}\right) t + O(1).$$

In the very recent paper [6], Golowich presents the first result with an explicit rate for the convergence in total variation distance of Z_t. For $\lambda \leq 1/100$, his main result gives universal constants η_1, η_2 such that

$$\|Z_t - \mathcal{N}^t_{\sigma^2}\|_{TV} \leq \frac{\lambda}{\sqrt{t}}(1+\log t)^{\eta_1 \log\log t + \eta_2} \quad \forall\, t \in \mathbb{N}, \tag{2.2}$$

where $\mathcal{N}^t_{\sigma^2}$ is a discretized normal distribution satisfying a number of axioms (see Definition 11 in [6]). It is asked in [6] whether the factor $(1+\log t)^{\eta_1 \log\log t + \eta_2}$ can be removed from the above bound. Although our approach does not allow us to completely get rid of it, Corollary 1.3 provides a better convergence with respect to t for bounded degree graphs. On the other hand, (2.2) gives a better convergence with respect to λ than Corollary 1.3.

3 Preliminaries

Given two probability measures μ, ν on a finite or countable set V, their *total variation distance* is

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{v \in V} |\mu(v) - \nu(v)|.$$

Let $G = (V, E)$ be a graph with n vertices, where V is the set of vertices and E the set of edges. We say that G is *d-regular* if every vertex $v \in V$ has degree d. Let (X_i) be the *simple random walk* (SRW for short) on G started at a vertex chosen uniformly at random from V, i.e., at every step the chain goes to an adjacent vertex chosen uniformly at random. The *transition matrix* of (X_i) is denoted by P, and its *stationary distribution* by π. Notice that π is the uniform distribution on V since G is regular. Denote $\langle \cdot, \cdot \rangle$ the usual inner product on $\mathbb{R}^V$ given by $\langle f, g \rangle = \sum_{x \in V} f(x)g(x)$. We will also consider the inner product $\langle \cdot, \cdot \rangle_\pi$ defined by

$$\langle f, g \rangle_\pi = \sum_{x \in V} f(x)g(x)\pi(x) \quad \forall\, f, g \colon V \longrightarrow \mathbb{R},$$

which induces a norm $\| \cdot \|_{2,\pi}$. It is well known that P is a self-adjoint stochastic matrix with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$, thus it has real eigenvalues $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n \geq -1$. Write $\lambda^* = \max\{|\lambda_j| \colon j \geq 2\}$. We say that G is a *λ-expander* graph if $\lambda^* \leq \lambda$. The *absolute spectral gap* of the chain is $1 - \lambda^*$. The spectral theorem applied to P gives an orthonormal basis of eigenvectors $(f_j)_{j=1}^n$ with respect to $\langle \cdot, \cdot \rangle_\pi$ corresponding to the eigenvalues $(\lambda_j)_{j=1}^n$. As a consequence, for any $f \colon V \longrightarrow \mathbb{R}$ we have

$$P^t f(x) = \sum_{j=1}^{n} \langle f, f_j \rangle_\pi f_j(x) \lambda_j^t \quad \forall\, x \in V \quad \forall\, t \in \mathbb{N}. \tag{3.1}$$

We refer to Lemma 12.2 in [12] for a more detailed explanation. Let W be a discrete random variable on $\mathbb{Z}$. The *characteristic function* (ch.f. for short) of W is $\varphi_W \colon [-\pi, \pi] \longrightarrow \mathbb{C}$ given by $\varphi_W(\theta) = \mathbb{E}(e^{i\theta W})$. If N is a continuous random variable on $\mathbb{R}$, then its ch.f. is $\varphi_N \colon \mathbb{R} \longrightarrow \mathbb{C}$ given by $\varphi_N(\theta) = \mathbb{E}(e^{i\theta N})$.

4 A general local central limit theorem for Markov chains

In this section, we prove a general central limit theorem for Markov chains with a spectral gap. Our main result Theorem 1.1 is obtained as an application of it.

Penrose and Peres proved in [16] the following useful principle: Let (Z_t) be a sequence of random variables that can be decomposed (with high probability) as $Z_t = S_t + Y_t$, where

- S_t and Y_t are independent;
- S_t is a sum of independent identically distributed random variables;
- (Y_t) satisfies the central limit theorem.

Then (Z_t) satisfies the local central limit theorem (with unspecified rate of convergence). See Theorem 2.1 in [16] for the precise statement.

The next result is of the same nature as this principle, but it provides an explicit rate of convergence. We need some preliminary notation. Let (X_i) be an irreducible and aperiodic Markov chain on a finite set V. Write $(\lambda_i)_{i=1}^{|V|}$ for its eigenvalues, where $\lambda_1 = 1$. We do not assume reversibility in the next result, so λ_i might be a complex number. We still write $\lambda^* = \max\{|\lambda_i| \colon i \geq 2\}$ and define the absolute spectral gap of the chain as $1 - \lambda^*$. Let W be a random variable taking integer values and pick $\eta > 0$ and $\theta_0 \in (0, \pi)$.

We say that W is an *η-nonlattice for θ_0* if $|\mathbb{E}(e^{i\theta W})| \leq 1-\eta$ for any $\theta \in \mathbb{R}$ satisfying $\theta_0 \leq |\theta| \leq \pi$. Clearly, $|\mathbb{E}(e^{i0W})| = 1$. If $|\mathbb{E}(e^{i\theta W})| = 1$ for some $\theta \neq 0$, then the distribution of $e^{i\theta W}$ must be concentrated at some point $e^{i\theta b}$, and $\mathbb{P}\{W \in b + (2\pi/\theta)\mathbb{Z}\} = 1$. The parameter η quantifies how far W is from behaving like this when $|\theta|$ is bigger than θ_0. Given two probability measures π, μ_0 on a finite or countable set V, let $\frac{\pi-\mu_0}{\pi}$ be the vector with entries $\frac{\pi-\mu_0}{\pi}(x) = \frac{\pi(x)-\mu_0(x)}{\pi(x)}$ for any $x \in V$.

Theorem 4.1. *Let (X_i) be an irreducible and aperiodic Markov chain on V with initial distribution μ_0, stationary distribution π, and absolute spectral gap $1-\lambda^* > 0$. Given $f\colon V \longrightarrow \mathbb{Z}$, write $Z_t = \sum_{i=0}^{t-1} f(X_i)$ and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. Let S_t and Y_t be independent random variables so that $\mathbb{E}_{\mu_0}(|Z_t - S_t - Y_t|) \leq \frac{M}{t}$ for some $M > 0$, and $S_t = \sum_{j=1}^{b_t} V_j$, where $b_t \in \mathbb{N}$ and V_j are independent η-nonlattice random variables for $\theta_0 > 0$ satisfying*

$$\theta_0 \leq \frac{(1-\lambda^*)^2\sigma^2}{2708\|f\|_\infty^3}.$$

Then there is a constant C_4 so that for any $k \in \mathbb{Z}$ and $t \in \mathbb{N}$,

$$\left|\mathbb{P}\{Z_t = k\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k - t\mathbb{E}_\pi(f)}{\sigma\sqrt{t}}\right)\right| \leq \left(\pi M + \frac{1}{\theta_0\sigma^2} + \frac{C_4\|f\|_\infty^3}{\sigma^4(1-\lambda^*)^2}\left(1 + \left\|\frac{\pi-\mu_0}{\pi}\right\|_{2,\pi}\right)\right)\frac{1}{t} + \frac{1}{e\eta}\frac{1}{b_t}.$$

Although the proof does not optimize the constant, it shows that we can take $C_4 = 6983$. Observe that the rate of convergence obtained in Theorem 4.1 depends on the absolute spectral gap $1-\lambda^*$ of the chain (X_i). It is worth to mention that (Y_t) satisfies the central limit theorem when λ^* is small. Indeed, recall that $Z_t = S_t + Y_t$, where (S_t) satisfies the central limit theorem, as it is a sum of independent identically distributed random variables. Furthermore, for small λ^* it is shown in [13, Theorem 1] that (Z_t) also satisfies the central limit theorem. Therefore, the claim follows from the characterization of convergence in distribution using pointwise convergence of the characteristic functions. We refer to the proof of Lemma 7.3 in [16] for more details. The strategy to prove Theorem 4.1 will be to analyze the difference between the characteristic functions of Z_t and the normal distribution. Then, we use Fourier Analysis to deduce the result. This analysis was previously done by Mann in [13] for low frequencies. The following proposition corresponds to the one-dimensional case of inequality (3.33) in [13], which will play a crucial role in the proof of Theorem 4.1.

Proposition 4.2 (Inequality (3.33) in [13])**.** *Let (X_i) be an irreducible and aperiodic Markov chain on V with initial distribution μ_0, stationary distribution π, and absolute spectral gap $1-\lambda^* > 0$. Given $f\colon V \longrightarrow \mathbb{R}$ with $\mathbb{E}_\pi(f) = 0$, write $Z_t = \sum_{i=0}^{t-1} f(X_i)$ and $\sigma^2 = \lim_{t\to\infty}\mathrm{Var}(Z_t)/t$. Let φ_Z be the ch.f. of Z_t and take*

$$\theta_0 = \frac{(1-\lambda^*)^2\sigma^2}{2708\|f\|_\infty^3}.$$

Then, for any $\theta \in [-\theta_0, \theta_0]$ we have

$$\left|\varphi_Z(\theta) - e^{\frac{-t\sigma^2\theta^2}{2}}\right| \leq (1-\lambda^*)^{-2}\|f\|_\infty^3\left(1 + \left\|\frac{\pi-\mu_0}{\pi}\right\|_{2,\pi}\right)e^{-t\sigma^2\theta^2/8}\left(683t|\theta|^3 + \frac{20}{\sigma^2}|\theta|\right).$$

We prove Theorem 4.1 first in the special case where $Z_t = S_t + Y_t$, where one can take $M = 0$, and then in full generality.

Lemma 4.3. *Theorem 4.1 holds in the case where $Z_t = S_t + Y_t$.*

Proof. First, assume that $\mathbb{E}_\pi(f) = 0$. Fix $t \in \mathbb{N}$ and let $\varphi_{t\sigma^2}$ be the ch.f. of a normal $\mathcal{N}(0, t\sigma^2)$, that is, $\varphi_{t\sigma^2}(\theta) = e^{-t\sigma^2\theta^2/2}$. Also, let φ_Z, φ_S, φ_Y, and φ_{V_j} be the ch.f.'s of Z_t,

S_t, Y_t, and V_j, respectively. The inversion formula (Theorem 3.3.14 in [4]) gives

$$\frac{1}{\sigma\sqrt{t}}\phi\left(\frac{y}{\sigma\sqrt{t}}\right) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\theta y}\varphi_{t\sigma^2}(\theta)\,d\theta. \quad \forall\, y \in \mathbb{R}.$$

Similarly, the inverse formula for discrete variables (Exercise 3.3.2 in [4]) gives

$$\mathbb{P}\left\{Z_t = k\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\theta k}\varphi_Z(\theta)\,d\theta \quad \forall\, k \in \mathbb{Z}.$$

Consequently, for any $k \in \mathbb{Z}$ we have

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k}{\sigma\sqrt{t}}\right)\right| \leq \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} e^{-i\theta k}(\varphi_Z(\theta) - \varphi_{t\sigma^2}(\theta))\,d\theta\right| + \frac{1}{2\pi}\left|\int_{|\theta|>\pi} e^{-i\theta k}\varphi_{t\sigma^2}(\theta)\,d\theta\right|. \tag{4.1}$$

It is not difficult to bound the second term in the left hand side of (4.1). Observe that

$$\begin{aligned}\left|\int_{|\theta|>\pi} e^{-i\theta k}\varphi_{t\sigma^2}(\theta)\,d\theta\right| &\leq \int_{|\theta|>\pi}|\varphi_{t\sigma^2}(\theta)|\,d\theta = 2\int_{\pi}^{\infty} e^{-t\sigma^2\theta^2/2}\,d\theta = 2\frac{1}{\sigma\sqrt{t}}\int_{\pi\sigma\sqrt{t}}^{\infty} e^{-x^2/2}\,dx \\ &\leq 2\frac{1}{\sigma\sqrt{t}}\int_{\pi\sigma\sqrt{t}}^{\infty}\frac{x}{\pi\sigma\sqrt{t}}e^{-x^2/2}\,dx = \frac{2}{\pi t\sigma^2}e^{-\pi^2 t\sigma^2} \leq \frac{2}{\pi\sigma^2}\frac{1}{t}.\end{aligned} \tag{4.2}$$

To bound the remaining term, we break it into two parts using θ_0. For convenience, write $c = (1-\lambda^*)^{-2}\|f\|_\infty^3(1 + \|\frac{\pi-\mu_0}{\pi}\|_{2,\pi})$. For any $\theta \in [-\theta_0, \theta_0]$, Proposition 4.2 gives

$$\left|\varphi_Z(\theta) - e^{\frac{-t\sigma^2\theta^2}{2}}\right| \leq ce^{-t\sigma^2\theta^2/8}\left(683t|\theta|^3 + \frac{20}{\sigma^2}|\theta|\right).$$

Therefore,

$$\begin{aligned}\left|\int_{-\theta_0}^{\theta_0} e^{-i\theta k}(\varphi_Z(\theta) - \varphi_{t\sigma^2}(\theta))\,d\theta\right| &\leq \int_{-\theta_0}^{\theta_0}|\varphi_Z(\theta) - \varphi_{t\sigma^2}(\theta)|\,d\theta \\ &\leq 2c\int_0^{\theta_0} e^{-t\sigma^2\theta^2/8}\left(683t\theta^3 + \frac{20}{\sigma^2}\theta\right)\,d\theta \\ &= 2c\frac{2}{\sigma\sqrt{t}}\int_0^{\theta_0\sigma\sqrt{t}/2} e^{-x^2/2}\left(\frac{683t\cdot 8}{\sigma^3 t^{3/2}}x^3 + \frac{20\cdot 2}{\sigma^3\sqrt{t}}x\right)\,dx \\ &\leq \frac{4c}{\sigma^4 t}\left(5464\int_0^{\infty} x^3 e^{-x^2/2}\,dx + 40\int_0^{\infty} xe^{-x^2/2}\,dx\right) \\ &= \frac{43872c}{\sigma^4}\frac{1}{t}.\end{aligned} \tag{4.3}$$

It remains to study the case $\theta_0 < |\theta| \leq \pi$. Observe that

$$\left|\int_{\theta_0\leq|\theta|\leq\pi} e^{-i\theta k}(\varphi_Z(\theta) - \varphi_{t\sigma^2}(\theta))\,d\theta\right| \leq \int_{\theta_0\leq|\theta|\leq\pi}|\varphi_Z(\theta)|\,d\theta + \int_{\theta_0\leq|\theta|\leq\pi}|\varphi_{t\sigma^2}(\theta)|\,d\theta.$$

We know that $\varphi_{t\sigma^2}(\theta) = e^{-t\sigma^2\theta^2/2}$, so we have

$$\int_{\theta_0\leq|\theta|\leq\pi}|\varphi_{t\sigma^2}(\theta)|\,d\theta = 2\int_{\theta_0}^{\pi} e^{-t\sigma^2\theta^2/2}\,d\theta = \frac{2}{\sigma\sqrt{t}}\int_{\theta_0\sigma\sqrt{t}}^{\pi\sigma\sqrt{t}} e^{-\theta^2/2}\,d\theta \leq \frac{2}{\sigma\sqrt{t}}\frac{1}{\theta_0\sigma\sqrt{t}} = \frac{2}{\theta_0\sigma^2}\frac{1}{t}. \tag{4.4}$$

Finally, recall that $Z_t = S_t + Y_t$ with $S_t = \sum_{j=1}^{b_t} V_j$, where $b_t \in \mathbb{N}$ and V_j are independent and η-nonlattice for θ_0. For θ with $\theta_0 \le |\theta| \le \pi$ we have

$$|\varphi_Z(\theta)| = |\varphi_S(\theta)\varphi_Y(\theta)| \le |\varphi_S(\theta)| = \prod_{j=1}^{b_t} |\varphi_{V_j}(\theta)| = \prod_{j=1}^{b_t} |\mathbb{E}(e^{i\theta V_j})| \le (1-\eta)^{b_t} \le e^{-\eta b_t} \le \frac{1}{e\eta b_t}, \tag{4.5}$$

where the last inequality holds since the function $f(x) = xe^{-\eta x}$ attains its maximum at η^{-1}. Therefore,

$$\int_{\theta_0 \le |\theta| \le \pi} |\varphi_Z(\theta)|\, d\theta \le 2\int_{\theta_0}^{\pi} \frac{1}{e\eta b_t}\, d\theta \le \frac{2\pi}{e\eta}\frac{1}{b_t}. \tag{4.6}$$

In view of the bounds obtained in (4.2), (4.3), (4.4), and (4.6), we conclude that

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k}{\sigma\sqrt{t}}\right)\right| \le \left(\frac{1}{\pi^2\sigma^2} + \frac{21936c}{\pi\sigma^4} + \frac{1}{\pi\theta_0\sigma^2}\right)\frac{1}{t} + \frac{1}{e\eta}\frac{1}{b_t} \quad \forall\, k \in \mathbb{Z}.$$

After substituting the value of c and some straightforward computations, we obtain

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k}{\sigma\sqrt{t}}\right)\right| \le \left(\frac{1}{\theta_0\sigma^2} + \frac{6983\|f\|_\infty^3}{\sigma^4(1-\lambda^*)^2}\left(1 + \left\|\frac{\pi-\mu_0}{\pi}\right\|_{2,\pi}\right)\right)\frac{1}{t} + \frac{1}{e\eta}\frac{1}{b_t}.$$

Consider now a general function $f\colon V \longrightarrow \mathbb{Z}$ and define $f_0 = f - \mathbb{E}_\pi(f)$ and $Z_t^0 = \sum_{i=0}^{t-1} f_0(X_i)$. Although f_0 might not be an integer-valued function, it takes values on $-\mathbb{E}_\pi(f) + \mathbb{Z}$. Also, Z_t^0 takes values on $-t\mathbb{E}_\pi(f) + \mathbb{Z}$ since

$$Z_t^0 = \sum_{i=0}^{t-1} f_0(X_i) = Z_t - t\mathbb{E}_\pi(f).$$

In the previous argument, we only used that Z_t takes integer values when applying the inverse formula, which is also valid for $-\mathbb{E}_\pi(f) - \mathbb{Z}$ (see Exercise 3.3.2 in [4]). Moreover, we can write $Z_t^0 = S_t + Y_t^0$, where $Y_t^0 = Y_t - t\mathbb{E}\pi(f)$. Therefore, the above argument applied to Z_t^0 gives

$$\left|\mathbb{P}\left\{Z_t^0 = k - t\mathbb{E}_\pi(f)\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k - t\mathbb{E}_\pi(f)}{\sigma\sqrt{t}}\right)\right| \le \left(\frac{1}{\theta_0\sigma^2} + \frac{6983\|f\|_\infty^3}{\sigma^4(1-\lambda^*)^2}\left(1 + \left\|\frac{\pi-\mu_0}{\pi}\right\|_{2,\pi}\right)\right)\frac{1}{t} + \frac{1}{e\eta b_t}.$$

The result follows from $\mathbb{P}\left\{Z_t^0 = k - t\mathbb{E}_\pi(f)\right\} = \mathbb{P}\left\{Z_t = k\right\}$. □

Proof of Theorem 4.1. Fix $t \in \mathbb{N}$, define $Z_t' = S_t + Y_t$, and denote its ch.f. by $\varphi_{Z'}$. Notice that the hypothesis $Z_t = S_t + Y_t$ in the proof of Lemma 4.3 was exclusively used to prove (4.5), which is clearly true for $\varphi_{Z'}$. Therefore, it is enough to bound $|\varphi_Z(\theta) - \varphi_{Z'}(\theta)|$ for $\theta \in [-\pi, \pi]$. First, recall that by the mean value theorem we have $|e^{ix} - e^{iy}| \le |x - y|$ for $x, y \in \mathbb{R}$. Consequently,

$$\begin{aligned}|\varphi_Z(\theta) - \varphi_{Z'}(\theta)| = |\mathbb{E}_{\mu_0}(e^{i\theta Z_t}) - \mathbb{E}_{\mu_0}(e^{i\theta Z_t'})| \le \mathbb{E}_{\mu_0}(|e^{i\theta Z_t} - e^{i\theta Z_t'}|) \le |\theta|\mathbb{E}_{\mu_0}(|Z_t - Z_t'|)\\ \le \pi\mathbb{E}_{\mu_0}(|Z_t - Z_t'|),\end{aligned}$$

for any $\theta \in [-\pi, \pi]$. Thus, the hypothesis on $\mathbb{E}_{\mu_0}(|Z_t - Z_t'|)$ and (4.5) give

$$|\varphi_Z(\theta)| \le |\varphi_{Z'}(\theta)| + |\varphi_Z(\theta) - \varphi_{Z'}(\theta)| \le \frac{1}{e\eta b_t} + \frac{\pi M}{t} \quad \forall\, \theta_0 \le |\theta| \le \pi. \tag{4.7}$$

To conclude the proof we just need to repeat the argument in the proof of Lemma 4.3, using (4.7) instead of (4.5). □

5 Proof of Theorem 1.1

Let G be a d-regular λ-expander graph. Let (X_i) be the SRW on G and write $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$, where val is a balanced labelling on G. We need some preliminary results to define the random variables S_t and Y_t that we use to decompose Z_t. The next lemma is a tighter version of the classic expander mixing lemma.

Lemma 5.1 (Lemma 4.15 in [17]). *Let G be a d-regular λ-expander graph with n vertices. For any subsets F_1, F_2 of V we have*

$$\left| |E(F_1, F_2)| - \frac{d}{n}|F_1||F_2| \right| \leq \lambda d \sqrt{\left(|F_1| - \frac{|F_1|^2}{n}\right)\left(|F_2| - \frac{|F_2|^2}{n}\right)}, \tag{5.1}$$

where $|E(F_1, F_2)| = \{(x, y) \in F_1 \times F_2 \colon \{x, y\} \in E\}|$ is the number of edges connecting F_1 and F_2 (counting edges contained in the intersection of F_1 and F_2 twice).

We will need lower bounds for $|E(F_1, F_2)|$ when F_1 is a subset of V and F_2 is its complement.

Corollary 5.2. *Let G be a d-regular λ-expander graph. For any $F_1 \subseteq V$ and $F_2 = {F_1}^{\mathsf{c}}$ we have*

$$|E(F_1, F_2)| \geq \frac{1}{2}(1-\lambda)d \min\{|F_1|, |F_2|\}.$$

Proof. Write $n = |V|$. Since $F_2 = {F_1}^{\mathsf{c}}$, we have $|F_1| - \frac{|F_1|^2}{n} = |F_2| - \frac{|F_2|^2}{n} = \frac{|F_1||F_2|}{n}$. Thus, the right hand side in (5.1) is $\lambda \frac{d}{n}|F_1||F_2|$. The result follows from the fact that $\frac{x(x-n)}{n} \geq \frac{x}{2}$ for every $x \in \left[0, \frac{n}{2}\right]$. ☐

Given a labelling $\mathrm{val} \colon V \longrightarrow \{0, 1\}$ on G, let $A = \{x \in V \colon \mathrm{val}(x) = 0\}$ and $B = \{x \in V \colon \mathrm{val}(x) = 1\}$. The labelling val is balanced when $|A| = |B| = \frac{|V|}{2}$. Given $x \in V$, write $q(x)$ for the number of neighbors y of x with $\mathrm{val}(y) = 0$. Define the sets

$$A_j = \{x \in A \colon q(x) = j\} \quad \text{and} \quad B_j = \{x \in B \colon q(x) = j\} \quad \forall\, j \in \{0, \dots, d\}.$$

The next lemma shows that for some $k^* \in \{1, \dots, d-1\}$ either the set A_{k^*} or B_{k^*} is relatively large.

Lemma 5.3. *Let G be a d-regular λ-expander graph and consider a balanced labelling val on G. Write $\delta = \frac{(1-\lambda)^2}{3}$. Then there is $k^* \in \{1, \dots, d-1\}$ such that either*

$$|A_{k^*}| \geq \frac{\delta|A|}{d-1} \quad \textit{or} \quad |B_{k^*}| \geq \frac{\delta|B|}{d-1}.$$

Proof. Assume that the statement is false. Then we must have

$$|A_0| + |A_d| > (1-\delta)|A| \quad \text{and} \quad |B_0| + |B_d| > (1-\delta)|B|. \tag{5.2}$$

Write $n = |V|$. Take $F_1 = A$ and $F_2 = B$. Corollary 5.2 gives $|E(A, B)| \geq (1-\lambda)d\frac{n}{4}$. Notice also that

$$|E(A, B)| \leq d|A \setminus A_d| \leq d(|A_0| + \delta|A|).$$

Therefore, we obtain

$$|A_0| \geq (1-\lambda)\frac{n}{4} - \delta\frac{n}{2} = (1-\lambda-2\delta)\frac{n}{4}. \tag{5.3}$$

A completely analogous argument replacing A_d with B_0 and A_0 with B_d gives

$$|B_d| \geq (1-\lambda-2\delta)\frac{n}{4}. \tag{5.4}$$

Next, take $F_1 = F_2 = A$. The expander mixing lemma gives $|E(A,A)| \geq (1-\lambda)d\frac{n}{4}$. Observe also that

$$|E(A,A)| \leq d|A \setminus A_0| \leq d(|A_d| + \delta|A|).$$

Thus, we get

$$|A_d| \geq (1-\lambda-2\delta)\frac{n}{4}. \tag{5.5}$$

Similarly, taking $F_1 = F_2 = B$ we obtain

$$|B_0| \geq (1-\lambda-2\delta)\frac{n}{4}. \tag{5.6}$$

Finally, take $F_1 = A_d \cup B_0$ and $F_2 = {F_1}^{\mathsf{c}}$. In view of (5.3), (5.4), (5.5), and (5.6) we have $\min\{|F_1|,|F_2|\} \geq (1-\lambda-2\delta)\frac{n}{2}$. Hence, Corollary 5.2 gives

$$|E(F_1,F_2)| \geq (1-\lambda)d(1-\lambda-2\delta)\frac{n}{4} = (1-\lambda)^2 d\left(1 - \frac{2(1-\lambda)}{3}\right)\frac{n}{4} \geq d\delta\frac{n}{4} = 2d\delta|A|.$$

Therefore, we must have either $|E(A_d,F_2)| \geq d\delta|A|$ or $|E(B_0,F_2)| \geq d\delta|B|$. In the first case, for any $e = \{x,y\} \in E$ with $x \in A_d$ and $y \in F_2$, we must have $\mathrm{val}(y) = 0$. Thus, $y \in A \setminus A_d$. Moreover, y cannot belong to A_0 since it is adjacent to x and $\mathrm{val}(x) = 0$. Therefore, $y \in A \setminus (A_0 \cup A_d)$. Consequently,

$$|A \setminus (A_0 \cup A_d)| \geq \frac{|E(A_0,F_2)|}{d} \geq \delta|A|,$$

which contradicts (5.2). If $|E(B_0,F_2)| \geq d\delta|B|$ we obtain $|B \setminus (B_0 \cup B_d)| \geq \delta|B|$, also a contradiction. □

Let (X_i) be the SRW on a d-regular λ-expander graph G with uniform initial distribution π. Recall that $Z_t = \sum_{i=0}^{t-1}\mathrm{val}(X_i)$ for some labelling val. The sets A_{k^*} and B_{k^*} provided by Lemma 5.3 will be used to decompose Z_t as a sum of two convenient independent random variables S_t and Y_t. The idea is that as we run the chain, we frequently see cycles of length 2. More precisely, $X_i = X_{i+2}$ with probability $\frac{1}{d}$. If we have $|A_{k^*}| \geq \frac{\delta|A|}{d-1}$, then many of these 2-cycles will start at a vertex of A_{k^*}. The contribution to Z_t of each one of these 2-cycles is either 0 with probability $\frac{k^*}{d}$, or 1 with probability $\frac{d-k^*}{d}$, and these contributions are independent of each other. If S_t represents the total contribution of the 2-cycles starting from A_{k^*}, and Y_t represents the contribution of the rest of the walk, then S_t is a sum of i.i.d. Bernoulli random variables and $Z_t = S_t + Y_t$. Although the idea is simple, making it rigorous requires a careful analysis.

Apply Lemma 5.3 to find $k^* \in \{1,\dots,d-1\}$ such that $\max\{|A_{k^*}|,|B_{k^*}|\} \geq \frac{\delta|A|}{d-1}$, where $\delta = \frac{(1-\lambda)^2}{3}$. By symmetry, we may assume that $|A_{k^*}| \geq \frac{\delta|A|}{d-1}$. Let (X_i^2) be the 2-steps SRW, that is, the Markov chain with transition matrix P^2. Let N_t be the number of times that $X_i^2 \in A_{k^*}$ within the first $\lfloor t/2 \rfloor - 1$ steps of the chain. Then

$$\begin{aligned}\mathbb{E}_\pi(N_t) &= \lfloor t/2\rfloor \pi(A_{k^*}) = \lfloor t/2\rfloor\frac{|A_{k^*}|}{|V|} = \lfloor t/2\rfloor\frac{|A_{k^*}|}{2|A|} \geq \lfloor t/2\rfloor\frac{\delta}{2(d-1)} \geq \frac{\delta(t-2)}{4(d-1)} \\ &= \frac{(1-\lambda)^2}{12(d-1)}(t-2).\end{aligned} \tag{5.7}$$

Recall that $\mathbb{P}\{X_i = X_{i+2}\} = \frac{1}{d}$ for any $i \geq 0$. Let $\widetilde{N}_t$ be a random variable that counts the number of times that one of these cycles of length 2 starting from a vertex of A_{k^*} appears at even time within the first t steps. To be more precise, let $i_1,\dots,i_{N_t}$ be the times for which $X_i^2 \in A_{k^*}$ and let U_j be the indicator that $X_{i_j} = X_{i_j+2}$, that is,

$$U_j = \begin{cases} 1 & \text{if } X_{i_j} = X_{i_j+2}; \\ 0 & \text{otherwise.}\end{cases}$$

Then (U_j) is a sequence of independent Bernoulli random variables with parameter $\frac{1}{d}$. We define

$$\widetilde{N}_t = \sum_{j=1}^{N_t} U_j \quad \text{and} \quad B_t = \sum_{j=1}^{\lceil \frac{\mathbb{E}_\pi(N_t)}{2} \rceil} U_j. \tag{5.8}$$

Let $b_t = \lfloor \mathbb{E}_\pi(\widetilde{N}_t)/4 \rfloor$. For every $i \in \{1, \dots, \widetilde{N}_t\}$, write (x_i, y_i) for the vertices appearing in the i-th 2-cycle, where $x_i \in A_{k^*}$ and y_i is some neighbor of x_i. Let V_i be the indicator of the event that $\mathrm{val}(y_i) = 1$ (which happens with probability $\frac{d-k^*}{d}$). Then $V_1, V_2, \dots$ are i.i.d. Bernoulli random variables. Moreover, the i-th 2-cycle adds V_i to the total sum of the labels. Consider as well $\widetilde{V}_1, \widetilde{V}_2, \dots$ independent from all previous random variables and identically distributed random variables given by

$$\widetilde{V}_i = \begin{cases} 1 & \text{with probability } \frac{d-k^*}{d}; \\ 0 & \text{otherwise.} \end{cases}$$

We can finally introduce the random variables used to decompose Z_t. Define

$$S'_t = \sum_{i=1}^{\min\{b_t, \widetilde{N}_t\}} V_i, \quad Y_t = Z_t - S'_t, \quad \text{and} \quad S_t = S'_t + \sum_{i=1}^{(b_t - \widetilde{N}_t)^+} \widetilde{V}_i, \tag{5.9}$$

where $(b_t - \widetilde{N}_t)^+ = \max\{b_t - \widetilde{N}_t, 0\}$. It turns out that $\widetilde{N}_t$ is concentrated around its mean, so one should expect $b_t \leq \widetilde{N}_t$. Hence, S'_t is equal to S_t with high probability, or equivalently, $Z_t = S_t + Y_t$ with high probability. To prove this fact, first we need the following Chernoff-type tail bound for the binomial distribution. For $a > 0$ set $\varphi(a) = 1 - a + a \log a$. Then $\varphi(a) > 0$ for $a \neq 1$ and $\varphi(1) = 0$.

Lemma 5.4 (Lemma 8.1 in [16]). *Let N be a binomial distributed random variable with $\mathbb{E}(N) = \mu > 0$. Then*

$$\mathbb{P}\{N \leq x\} \leq e^{-\mu\varphi\left(\frac{x}{\mu}\right)} \quad \forall\, 0 < x \leq \mu.$$

We also need the following consequence of Theorem 2.1 in [5]. It provides a Chernoff bound for random walks on expander graphs.

Lemma 5.5. *Let (X_i) be the random walk on a weighted graph $G = (V, E)$ starting from stationary distribution π, and let $1 - \lambda^*$ be its absolute spectral gap. Given $A \subseteq V$, let N_t be the number of visits to A in t steps. For any $0 < \gamma \leq t$,*

$$\mathbb{P}\{|N_t - \mathbb{E}_\pi(N_t)| \geq \gamma\} \leq 4e^{-\gamma^2(1-\lambda^*)/20t}.$$

Now we can show that $\widetilde{N}_t$ is expected to be bigger than b_t.

Lemma 5.6. *Let $\widetilde{N}_t$ be defined as in (5.8). Then*

$$\mathbb{P}\left\{\widetilde{N}_t \leq \mathbb{E}_\pi(\widetilde{N}_t)/4\right\} \leq 5\exp\left(-\frac{(1-\lambda)^5}{11520d^2}(t-4)\right).$$

Proof. Notice that

$$\begin{aligned} \mathbb{P}\left\{\widetilde{N}_t \leq \mathbb{E}_\pi(\tilde{N}_t)/4\right\} &\leq \mathbb{P}\left\{\widetilde{N}_t \leq \mathbb{E}_\pi(\tilde{N}_t)/4 | N_t > \mathbb{E}_\pi(N_t)/2\right\} + \mathbb{P}\{N_t \leq \mathbb{E}_\pi(N_t)/2\} \\ &\leq \mathbb{P}\left\{B_t \leq \mathbb{E}_\pi(\tilde{N}_t)/4\right\} + \mathbb{P}\{N_t \leq \mathbb{E}_\pi(N_t)/2\} \\ &\leq \mathbb{P}\{B_t \leq \mathbb{E}_\pi(B_t)/2\} + \mathbb{P}\{N_t \leq \mathbb{E}_\pi(N_t)/2\}, \end{aligned}$$

where the last inequality follows from the fact that $\frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} \leq \frac{\mathbb{E}_\pi(B_t)}{2}$. We can use Lemma 5.4 to get

$$\mathbb{P}\{B_t \leq \mathbb{E}_\pi(B_t)/2\} \leq e^{-\mathbb{E}_\pi(B_t)\varphi(1/2)} \leq \exp\left(-\frac{\delta\varphi(1/2)}{8d(d-1)}(t-2)\right) \leq \exp\left(-\frac{(1-\lambda)^2}{157d^2}(t-2)\right),$$

where $\delta = (1-\lambda)^2/3$. Finally, applying Lemma 5.5 to (X_i^2) and A_{k^*} with $\gamma = \mathbb{E}_\pi(N_t)/2$ gives

$$\mathbb{P}\{N_t \le \mathbb{E}_\pi(N_t)/2\} \le 4e^{-\mathbb{E}_\pi(N_t)^2(1-\lambda^2)/80t} \le 4\exp\left(-\frac{(1-\lambda^2)\delta^2}{16(d-1)^2}\frac{(t-2)^2}{80t}\right)$$
$$\le 4\exp\left(-\frac{(1-\lambda)^5}{11520d^2}(t-4)\right). \qquad \square$$

We can use Lemma 5.6 to obtain the bound for $\mathbb{E}_\pi(|Z_t - S_t - Y_t|)$ that we need to apply Theorem 4.1.

Lemma 5.7. *Let S_t, S_t', and Y_t defined as in (5.9). Then*

$$\mathbb{E}_\pi(|Z_t - S_t - Y_t|) \le \frac{10^{10}d^3}{(1-\lambda)^4}\frac{1}{t}.$$

Proof. Note that the function $f(t) = t^2e^{-\alpha t}$ attains its maximum at $t = 2\alpha^{-1}$. Write $\alpha = \frac{(1-\lambda)^2}{11520d^2}$. Then Lemma 5.6 gives

$$\mathbb{P}\left\{\widetilde{N}_t \le \mathbb{E}_\pi(\widetilde{N}_t)/4\right\} \le 5e^{-\alpha(t-4)} = \left(5e^4t^2e^{-\alpha t}\right)\frac{1}{t^2} \le \left(5e^4\frac{4e^{-2}}{\alpha^2}\right)\frac{1}{t^2} \le \frac{2\cdot 10^{10}d^4}{(1-\lambda)^4}\frac{1}{t^2}.$$

Observe that $|Z_t - S_t - Y_t| = |S_t' - S_t| \le b_t \le \frac{t}{4d}$. Since $Z_t = Y_t + S_t$ if $\widetilde{N}_t \ge b_t$, we have

$$\mathbb{E}_\pi(|Z_t - S_t - Y_t|) \le \frac{t}{4d}\mathbb{P}\{Z_t \ne S_t + Y_t\} \le \frac{10^{10}d^3}{(1-\lambda)^4}\frac{1}{t}. \qquad \square$$

Finally, the next simple result shows that a Bernoulli random variable is nonlattice. We include its proof for completeness.

Lemma 5.8. *Let V be a Bernoulli random variable with parameter $p \in (0,1)$ and take $\theta_0 \in (0,\pi)$. Then X is a η-nonlattice variable for θ_0 with*

$$\eta = p(1-p)(1-\cos(\theta_0)).$$

Proof. Write φ_V for the characteristic function of V, that is,

$$\varphi_V(\theta) = \mathbb{E}(e^{i\theta V}) = (1-p) + pe^{i\theta} \quad \forall\, \theta \in [-\pi,\pi],$$

which is a convex combination of the points 1 and $e^{i\theta}$. Therefore, when $\theta_0 \le |\theta| \le \pi$ we clearly have $|\varphi_V(\theta)| \le |\varphi_V(\theta_0)|$. A simple computation yields

$$|\varphi_V(\theta_0)|^2 = p^2\sin(\theta_0)^2 + ((1-p) + p\cos(\theta_0))^2 = p^2 + (1-p)^2 + 2p(1-p)\cos(\theta_0).$$

Using the equality $(a-b)(a+b) = a^2 - b^2$ and $|\varphi_V| \le 1$ we conclude that

$$1 - |\varphi_V(\theta_0)| \ge \frac{1}{2}(1 - |\varphi_V(\theta_0)|^2) = p(1-p)(1-\cos(\theta_0)). \qquad \square$$

We can now present the proof of our main result Theorem 1.1. Although it does not optimize the constant, it shows that we can take

$$C_1(\lambda, d) = \frac{2\cdot 10^{13}d^9}{(1-\lambda)^{10}}.$$

Proof of Theorem 1.1. Consider the random variables $\widetilde{N}_t$, V_i, $\widetilde{V}_i$, S_t', S_t, and Y_t appearing in (5.9). Let us check that the hypotheses of Theorem 4.1 are satisfied. First, (X_i) is the SRW on a finite λ-expander graph with $\lambda < 1$, whence it is irreducible and aperiodic.

The function considered is $\mathrm{val}\colon V \longrightarrow \mathbb{Z}$. Observe that S_t does not get affected by conditioning on Y_t. Indeed, regardless of the value of Y_t, S_t is a sum of b_t i.i.d. Bernouilli random variables that are independent of Y_t. Therefore, S_t and Y_t are independent. Moreover, Lemma 5.7 gives $\mathbb{E}_\pi(|Z_t - S_t - Y_t|) \le M/t$, with

$$M = \frac{10^{10} d^3}{(1-\lambda)^4}.$$

Finally, Lemma 5.8 shows that V_i and $\widetilde{V}_i$ are η-nonlattice for $\theta_0 = (1-\lambda)^2\sigma^2/2708$ with

$$\eta \ge \frac{1}{d}\left(1 - \frac{1}{d}\right)(1-\cos(\theta_0)) = \frac{d-1}{d^2}(1-\cos(\theta_0)).$$

Since in our case $\|\,\mathrm{val}\,\|_\infty = 1$ and $\mu_0 = \pi$, Theorem 4.1 gives

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k - t/2}{\sigma\sqrt{t}}\right)\right| \le \left(\pi M + \frac{1}{\theta_0\sigma^2} + \frac{C_4}{\sigma^4(1-\lambda)^2}\right)\frac{1}{t} + \frac{1}{e\eta}\frac{1}{b_t}. \tag{5.10}$$

Recall that $b_t = \lfloor \mathbb{E}_\pi(\widetilde{N}_t)/4 \rfloor$. Using the bound (5.7) for $\mathbb{E}_\pi(N_t)$ we get

$$\frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} = \frac{\mathbb{E}_\pi(N_t)}{4d} \ge \frac{(1-\lambda)^2}{48d(d-1)}(t-2) \ge \frac{(1-\lambda)^2 t - 2}{48d(d-1)}.$$

If $t \le 48(1-\lambda)^{-2}d^3$ there is nothing to prove. Otherwise, a simple computation shows that

$$b_t = \left\lfloor \frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} \right\rfloor \ge \frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} - 1 \ge \frac{(1-\lambda)^2 t - 48d^2}{48d(d-1)} \ge \frac{(1-\lambda)^2}{48d^2}t. \tag{5.11}$$

Next, we provide a bound for σ. Observe that

$$\begin{aligned}
\mathrm{Var}(Z_t) &\ge \mathbb{E}_\pi(\mathrm{Var}(Z_t|\widetilde{N}_t)) \ge \sum_{u=b_t}^{\infty} \mathrm{Var}(Z_t|\widetilde{N}_t = u)\mathbb{P}\left\{\widetilde{N}_t = u\right\} \\
&= \sum_{u=b_t}^{\infty} \mathrm{Var}(S_t + Y_t|\widetilde{N}_t = u)\mathbb{P}\left\{\widetilde{N}_t = u\right\} \ge \sum_{u=b_t}^{\infty} \mathrm{Var}(S_t|\widetilde{N}_t = u)\mathbb{P}\left\{\widetilde{N}_t = u\right\} \\
&= \mathrm{Var}(S_t)\sum_{u=b_t}^{\infty} \mathbb{P}\left\{\widetilde{N}_t = u\right\} \ge b_t\frac{d-1}{d^2}\mathbb{P}\left\{\widetilde{N}_t \ge b_t\right\}.
\end{aligned}$$

As t grows, Lemma 5.6 shows that $\mathbb{P}\left\{\widetilde{N}_t \ge b_t\right\}$ tends to 1. Thus, (5.11) gives

$$\sigma \ge \sqrt{\frac{(1-\lambda)^2(d-1)}{48d^4}} \ge \sqrt{\frac{(1-\lambda)^2}{96d^3}} \ge \frac{(1-\lambda)}{10d^{3/2}}. \tag{5.12}$$

Finally, we give a bound for η. It is straightforward to show that $\cos(x) \le 1 - x^2/5$ for any $x \in [-\pi, \pi]$. Consequently,

$$\eta \ge \frac{d-1}{d^2}(1-\cos(\theta_0)) \ge \frac{1}{2d}\frac{\theta_0^2}{5} = \frac{(1-\lambda)^4\sigma^4}{10 \cdot 2708^2 d} \ge \frac{(1-\lambda)^8}{8 \cdot 10^{11} d^7}. \tag{5.13}$$

Substituting the value of M and θ_0 and applying the bounds (5.11), (5.12), and (5.13) in (5.10) yields

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k-t/2}{\sigma\sqrt{t}}\right)\right| \le \left(\frac{\pi 10^{10} d^3}{(1-\lambda)^4} + \frac{10^8 d^6}{(1-\lambda)^6} + \frac{1.5 \cdot 10^{13} d^9}{(1-\lambda)^{10}}\right)\frac{1}{t} \le \frac{2 \cdot 10^{13} d^9}{(1-\lambda)^{10}}\frac{1}{t}.$$

□

Remark 5.9. It is possible to obtain a better bound in Theorem 1.1 if we assume that λ is small enough. If $\lambda \le \frac{1}{5}$, then Theorem 1.1 holds with $C_1(d) = 4 \cdot 10^{11} d^3$. This will be proved in Section 6.

6 Proofs of Proposition 1.4 and Corollary 1.3

Let $G = (V, E)$ be a d-regular λ-expander graph, let (X_i) be the SRW on G with uniform initial distribution π, and let $\mathrm{val}\colon V \longrightarrow \{0,1\}$ be a labelling with $\alpha = \mathbb{E}_\pi(\mathrm{val}) \in (0,1)$. For simplicity, write $Y_i = \mathrm{val}(X_i)$. Recall that $Z_t = \sum_{i=0}^{t-1} Y_i$, thus its variance can be expressed as

$$\mathrm{Var}(Z_t) = \sum_{j=0}^{t-1} \mathrm{Var}(Y_j) + 2\sum_{i<j} \mathrm{Cov}(Y_i, Y_j). \tag{6.1}$$

It is clear that $\mathrm{Var}(Y_j) = \mathbb{E}_\pi(Y_j^2) - \mathbb{E}_\pi(Y_j)^2 = \alpha - \alpha^2 = \alpha(1-\alpha)$ and $\mathrm{Cov}(Y_i, Y_j) = \mathbb{E}_\pi(Y_iY_j) - \alpha^2$. Recall that $A = \{x \in V\colon\ \mathrm{val}(x) = 0\}$ and $B = A^{\mathsf{c}}$. We have

$$\mathbb{E}_\pi(Y_iY_j) = \mathbb{P}\{Y_i = 1, Y_j = 1\} = \alpha\mathbb{P}\{Y_j = 1 | Y_i = 1\} = \alpha\mathbb{P}\{X_{j-i} \in B | X_0 \in B\}. \tag{6.2}$$

Thus, we want to find the probability that the chain is at a vertex of B after $j - i$ steps when the initial vertex is chosen uniformly at random from B.

Lemma 6.1. *Let $G = (V, E)$ be a d-regular graph with n vertices and let $B \subseteq V$. If (X_i) is the simple random walk on V starting uniformly at random from B, we have*

$$\mathbb{P}\{X_k \in B\} = \pi(B) + \pi(B)\sum_{j=2}^{n} \langle \pi_B, f_j\rangle^2 \lambda_j^k \quad \forall\, k \in \mathbb{N},$$

where $(f_j)_{j=1}^n$ is the orthonormal basis of eigenvectors corresponding to the eigenvalues $(\lambda_j)_{j=1}^n$ and π_B is the uniform distribution on B.

Proof. Let P denote the transition matrix of (X_i). Since P is symmetric, $P^k\pi_B$ is the vector of probabilities of the chain (X_i) after k steps. Therefore, $\mathbb{P}\{X_k \in B\} = \langle P^k\pi_B, \mathrm{val}\rangle$. We can use (3.1) to decompose P and obtain

$$P^k\pi_B = \sum_{j=1}^{n} \langle \pi_B, f_j\rangle_\pi f_j \lambda_j^k = \pi + \sum_{j=2}^{n} \langle \pi_B, f_j\rangle_\pi f_j \lambda_j^k.$$

Since $\pi_B = \frac{1}{|B|}\,\mathrm{val}$, we conclude

$$\langle P^k\pi_B, \mathrm{val}\rangle = \pi(B) + \sum_{j=2}^{n} \langle \pi_B, f_j\rangle_\pi \langle f_j, \mathrm{val}\rangle \lambda_j^k = \pi(B) + \pi(B)\sum_{j=2}^{n} \langle \pi_B, f_j\rangle^2 \lambda_j^k. \qquad \square$$

Proof of Proposition 1.4. Recall that $A = \{x \in V\colon\ \mathrm{val}(x) = 0\}$, $B = A^{\mathsf{c}}$, and $\alpha = \pi(B)$. Fix $0 \le i < j \le t-1$. Lemma 6.1 and (6.2) give

$$\mathrm{Cov}(Y_i, Y_j) = \mathbb{E}_\pi(Y_iY_j) - \alpha^2 = \alpha^2 \sum_{k=2}^{n} \langle \pi_B, f_j\rangle^2 \lambda_k^{j-i} \quad \forall\, i < j.$$

Adding all covariances yields

$$\sum_{i<j} \mathrm{Cov}(Y_i, Y_j) = \alpha^2 \sum_{k=1}^{t-1} (t-k) \sum_{j=2}^{n} \langle \pi_B, f_j\rangle^2 \lambda_j^k.$$

The formula (1.4) follows now from (6.1). Finally, if we asssume that G is a λ-expander we get

$$\begin{aligned} |\mathrm{Var}(Z_t) - \alpha(1-\alpha)t| &\le 2\alpha^2 \sum_{k=1}^{t-1} (t-k)\lambda^k \sum_{j=2}^{n} \langle \pi_B, f_j\rangle^2 \le 2\alpha^2 t \sum_{k=1}^{t-1} \lambda^k (n\|\pi_B\|_2^2 - \langle \pi_B, f_1\rangle^2) \\ &= 2\alpha^2 \left(\frac{n}{|B|} - 1\right) t \sum_{k=1}^{t-1} \lambda^k \le 2\alpha(1-\alpha)t\frac{\lambda}{1-\lambda}. \end{aligned} \qquad \square$$

As Remark 5.9 claims, for small values of λ we can use the bound for the variance provided by Proposition 1.4 to obtain a better bound in Theorem 1.1.

Proof of Remark 5.9. In this case, we have $\alpha = 1/2$. Dividing by t and sending t to infinity in (1.5) give

$$\left|\sigma^2 - \frac{1}{4}\right| \le \frac{1}{2}\frac{\lambda}{1-\lambda} \le \frac{1}{8},$$

since we assume $\lambda \le 1/5$. Therefore, $\sigma^2 \ge 1/8$. Using this bound instead of (5.12) also allow us to obtain better bounds for θ_0 and η. Indeed, we have

$$\theta_0 = \frac{(1-\lambda)^2\sigma^2}{2708} \ge \frac{1}{33850} \quad \text{and} \quad \eta \ge \frac{1}{2d}\frac{\theta_0^2}{5} \ge \frac{1}{1.2 \cdot 10^{10} d}.$$

Using these bounds and the bound (5.11) for b_t in (5.10) yields

$$\left|\mathbb{P}\{Z_t = k\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k - t/2}{\sigma\sqrt{t}}\right)\right| \le \left(8 \cdot 10^{10} d^3 + 10^6 + 3.4 \cdot 10^{11} d^3\right)\frac{1}{t} \le 4 \cdot 10^{11} d^3 \frac{1}{t}. \quad \square$$

Finally, we prove Corollary 1.3. First, we show that the normalizing constant $D(\mu, \sigma^2)$ appearing in (1.1) is close to 1 when the variance is large.

Lemma 6.2. *Let $D(\mu, \sigma^2)$ be the normalizing constant appearing in (1.1). For $\sigma^2 \ge 1$ we have*

$$|1 - D(\mu, \sigma^2)| \le \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}.$$

Proof. The density function of a normal distribution increases on $(-\infty, 0)$ and decreases on $(0, \infty)$. Therefore,

$$\phi(k) \ge \sigma \int_{k-\frac{1}{\sigma}}^{k} \phi(x)\,dx \quad \text{for } k < 0 \quad \text{and} \quad \phi(k) \ge \sigma \int_{k}^{k+\frac{1}{\sigma}} \phi(x)\,dx \quad \text{for } k \ge 0.$$

Let $I_1 = \{k \in \mathbb{Z} \colon k \le \mu\}$ and $I_2 = \mathbb{Z} \setminus I_1$. The previous observation gives

$$\sum_{k \in I_1} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) \ge \sum_{k \in I_1} \int_{\frac{k-\mu}{\sigma} - \frac{1}{\sigma}}^{\frac{k-\mu}{\sigma}} \phi(x)\,dx = \int_{-\infty}^{\frac{\lfloor \mu \rfloor - \mu}{\sigma}} \phi(x)\,dx.$$

Similarly,

$$\sum_{k \in I_2} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) \ge \sum_{k \in I_2} \int_{\frac{k-\mu}{\sigma}}^{\frac{k-\mu}{\sigma} + \frac{1}{\sigma}} \phi(x)\,dx = \int_{\frac{\lfloor \mu \rfloor - \mu}{\sigma} + \frac{1}{\sigma}}^{\infty} \phi(x)\,dx.$$

Since ϕ attains its maximum at $x = 0$, we deduce that

$$\sum_{k \in \mathbb{Z}} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) \ge 1 - \frac{1}{\sigma}\phi(0) = 1 - \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}.$$

Similarly, $\phi(k) \le \sigma \int_k^{k+1/\sigma} \phi(x)\,dx$ for any $k \le -\frac{1}{2\sigma}$ and $\phi(k) \le \sigma \int_{k-1/\sigma}^{k} \phi(x)\,dx$ for $k \ge \frac{1}{2\sigma}$. Notice that there is a unique integer $k^* \in \mathbb{Z}$ such that

$$-\frac{1}{2\sigma} \le \frac{k^* - \mu}{\sigma} < \frac{1}{2\sigma}.$$

Therefore,

$$\sum_{k \in \mathbb{Z}} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) \le \sum_{k \in \mathbb{Z}\setminus\{k^*\}} \sigma^{-1}\phi\left(\frac{k-\mu}{\sigma}\right) + \frac{1}{\sigma}\phi(0) \le 1 + \frac{1}{\sigma}\phi(0) = 1 + \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}. \quad \square$$

Let val be a balanced labelling on G and consider $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$. We write φ_Z for the ch.f. of Z_t and φ_t for the ch.f. of a normal distribution with mean $t/2$ and variance $t\sigma^2$, where $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. The next technical result bounds the L^2-distance between φ_Z and φ_t.

Lemma 6.3. *The characteristic function of Z_t satisfies*

$$\|\varphi_Z - \varphi_t\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta\right)^{1/2} \le \frac{3\cdot 10^{13} d^9}{(1-\lambda)^{10}} \frac{1}{t^{3/4}}.$$

Proof. We will break the integral into two parts as we did in the proof of Theorem 4.1. Set $\theta_0 = (1-\lambda)^2\sigma^2/2708$. For any $\theta \in [-\theta_0, \theta_0]$, Proposition 4.2 gives

$$|\varphi_Z(\theta) - \varphi_t(\theta)| \le ce^{-t\sigma^2\theta^2/8}\left(683t|\theta|^3 + \frac{20}{\sigma^2}|\theta|\right),$$

where $c = (1-\lambda)^{-2}$. Therefore,

$$\begin{aligned}\int_{-\theta_0}^{\theta_0} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta &\le 4c^2 \int_0^{\theta_0} e^{-t\sigma^2\theta^2/4}\left(683^2 t^2\theta^6 + \frac{400}{\sigma^4}\theta^2\right) d\theta \qquad (6.3)\\ &= 4c^2 \frac{2}{\sigma\sqrt{t}} \int_0^{\theta_0\sigma\sqrt{t}/2} e^{-x^2}\left(\frac{683^2\cdot 2^6 t^2}{\sigma^6 t^3}x^6 + \frac{400\cdot 2^2}{\sigma^4\cdot\sigma^2 t}x^2\right) dx\\ &\le \frac{2^9 c^2}{\sigma^7 t^{3/2}}\left(683^2\int_0^\infty x^6 e^{-x^2}\, dx + 25\int_0^\infty x^2 e^{-x^2}\, dx\right) \le \frac{5\cdot 10^8 c^2}{\sigma^7 t^{3/2}}.\end{aligned}$$

It remains to study the case $\theta_0 < |\theta| \le \pi$. Observe that

$$\int_{\theta_0\le|\theta|\le\pi} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta \le 2\int_{\theta_0\le|\theta|\le\pi} |\varphi_Z(\theta)|^2 \, d\theta + 2\int_{\theta_0\le|\theta|\le\pi} |\varphi_t(\theta)|^2 \, d\theta.$$

Consider $Z'_t = S_t + Y_t$, where S_t and Y_t are defined as in (5.9) and let $\varphi_{Z'}$ be its characteristic function. Recall that (4.5) gives $|\varphi_{Z'}(\theta)| \le \frac{1}{e\eta b_t}$ for $\theta_0 \le |\theta| \le \pi$. Moreover, the proof of Theorem 4.1 shows that

$$|\varphi_Z(\theta) - \varphi_{Z'}(\theta)| \le \pi\mathbb{E}_\pi(|Z_t - Z'_t|) \le \frac{\pi M}{t} \quad \forall\, \theta \in [-\pi, \pi],$$

where $M = \frac{10^{10} d^3}{(1-\lambda)^4}$ in view of Lemma 5.7. Consequently,

$$\int_{\theta_0\le|\theta|\le\pi} |\varphi_Z(\theta)|^2 \, d\theta \le \frac{4\pi}{e^2\eta^2 b_t^2} + \frac{2\pi^3 M^2}{t^2}.$$

On the other hand, a simple computation shows that the function $f(t) = \sqrt{t}e^{-\alpha t}$ attains its maximum at $t = 1/(2\alpha)$ and $f(1/(2\alpha)) \le 1/\sqrt{\alpha}$. Taking $\alpha = \sigma^2\theta_0^2$, this implies that $e^{-t\sigma^2\theta^2} \le \frac{1}{\sigma\theta_0\sqrt{t}}$. Therefore,

$$\begin{aligned}\int_{\theta_0\le|\theta|\le\pi} |\varphi_t(\theta)|^2 \, d\theta &= 2\int_{\theta_0}^{\pi} e^{-t\sigma^2\theta^2}\, d\theta = \frac{2}{\sigma\sqrt{t}}\int_{\theta_0\sigma\sqrt{t}}^{\pi\sigma\sqrt{t}} e^{-x^2}\, dx\\ &\le \frac{2}{\sigma\sqrt{t}}\frac{1}{2\theta_0\sigma\sqrt{t}}\int_{\theta_0\sigma\sqrt{t}}^{\infty} 2xe^{-x^2}\, dx = \frac{1}{\theta_0\sigma^2 t}e^{-\theta_0^2\sigma^2 t} \le \frac{1}{\theta_0\sigma^2 t}\frac{1}{\theta_0\sigma\sqrt{t}}\\ &= \frac{1}{\theta_0^2\sigma^3 t^{3/2}}.\end{aligned}$$

Putting everything together gives

$$\|\varphi_Z - \varphi_t\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta \le \frac{10^8 c^2}{\sigma^7 t^{3/2}} + \frac{4}{e^2\eta^2 b_t^2} + \frac{2\pi^2 M^2}{t^2} + \frac{1}{\pi\theta_0^2\sigma^3 t^{3/2}}.$$

Taking square root in both sides and using that $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ yields

$$\|\varphi_Z - \varphi_t\|_2 \le \frac{10^4 c}{\sigma^{7/2} t^{3/4}} + \frac{2}{e\eta b_t} + \frac{\sqrt{2}\pi M}{t} + \frac{1}{\pi^{1/2}\theta_0\sigma^{3/2}t^{3/4}}.$$

Recall that b_t, σ, and η, are bounded in (5.11), (5.12), and (5.13), respectively. The desired result follows from these bounds and the above inequality after straightforward computations. □

The proof of Corollary 1.3 does not optimize the constant, but it shows that we can take

$$C_2(\lambda, d) = \frac{10^{14} d^9}{(1-\lambda)^{41/4}}.$$

Proof of Corollary 1.3. Given $c \ge 1$, set $I_1 = \{k \in \mathbb{Z} \colon |k - t/2| \le c\sqrt{t} + 1\}$ and $I_2 = \mathbb{Z} \setminus I_1$. Recall that ϕ stands for the density function of a standard normal distribution and write ϕ_t for the density function of a normal distribution with mean $t/2$ and variance $t\sigma^2$, where $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. First, we bound

$$\sum_{k\in\mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)|$$

by breaking the sum using I_1 and I_2. To bound the sum on I_2 we study the tails of the distributions of Z_t and $\mathcal{N}(t/2, t\sigma^2)$. On one hand,

$$\sum_{k\in I_2} \phi_t(k) \le 2\int_{t/2+c\sqrt{t}}^{\infty} \phi_t(x)\,dx = 2\int_{c/\sigma}^{\infty} \phi(y)\,dy \le \frac{2}{\sqrt{2\pi}} \int_{c/\sigma}^{\infty} \frac{y}{c/\sigma} e^{-y^2/2}\,dy = \frac{2\sigma}{c\sqrt{2\pi}} e^{-\frac{c^2}{2\sigma^2}}.$$

On the other hand, Lemma 5.5 gives

$$\sum_{k\in I_2} \mathbb{P}\{Z_t = k\} = \mathbb{P}\left\{|Z_t - t/2| > c\sqrt{t} + 1\right\} \le 4e^{-\frac{(c\sqrt{t}+1)^2(1-\lambda)}{20t}} \le 4e^{-\frac{c^2(1-\lambda)}{20}}.$$

Consequently, we have

$$\sum_{k\in I_2} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| \le \frac{2\sigma}{c\sqrt{2\pi}} e^{-\frac{c^2}{2\sigma^2}} + 4e^{-\frac{c^2(1-\lambda)}{20}}. \tag{6.4}$$

We can use the Cauchy-Schwarz inequality to bound the sum on I_1 as follows.

$$\sum_{k\in I_1} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| \le |I_1|^{1/2} \left(\sum_{k\in I_1} |\mathbb{P}\{Z_t = k\} - \phi_t(k)|^2 \right)^{1/2}.$$

We clearly have $|I_1| \le 2c\sqrt{t} + 3 \le 5c\sqrt{t}$. To estimate the sum on the right hand side, we will use Parseval's identity (see [9, §1.4]). Define $F \colon [-\pi, \pi] \longrightarrow \mathbb{C}$ by

$$F(\theta) = \sum_{k\in\mathbb{Z}} \varphi_t(\theta + 2\pi k) \quad \forall\, \theta \in [-\pi, \pi],$$

where φ_t denotes the ch.f. of the normal distribution with mean $t/2$ and variance $t\sigma^2$. Then the Fourier coefficients of F, denoted by $a_k(F)$, satisfy

$$a_k(F) = \frac{1}{2\pi} \widehat{\varphi_t}(k) = \phi_t(k) \quad \forall\, k \in \mathbb{Z},$$

where $\widehat{\varphi_t}$ is the Fourier transform of φ_t as defined in [9, §1.2]. The first equality follows from (2.4.7) in [9], and the second one follows from the inversion formula (see Theorem 3.3.14 in [4]). Consequently, Parseval's identity gives

$$\sum_{k\in\mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)|^2 = \|\varphi_Z - F\|_2^2.$$

Recall that Lemma 6.3 gives a bound for $\|\varphi_Z - \varphi_t\|_2$, so by triangle inequality it suffices to estimate $\|\varphi_t - F\|_2$. For any $\theta \in [-\pi, \pi]$ we have

$$\begin{aligned} |\varphi_t(\theta) - F(\theta)| &\le \sum_{k\neq 0} |\varphi_t(\theta + 2\pi k)| = \sum_{k\neq 0} e^{-\frac{\sigma^2 t(\theta+2\pi k)^2}{2}} \le 2\sum_{k=1}^{\infty} e^{-\sigma^2 tk} = 2e^{-\sigma^2 t}\sum_{k=0}^{\infty} (e^{-\sigma^2 t})^k \\ &= \frac{2e^{-\sigma^2 t}}{1 - e^{-\sigma^2 t}}. \end{aligned}$$

It is easy to check that $e^x \ge 1 + x$ for any $x \ge 0$, or equivalently, $xe^{-x} \le 1 - e^{-x}$ for any $x \ge 0$. In view of the above bound, dividing by x and $1 - e^{-x}$ both sides and taking $x = \sigma^2 t$ shows that

$$|\varphi_t(\theta) - F(\theta)| \le \frac{2e^{-\sigma^2 t}}{1 - e^{-\sigma^2 t}} \le \frac{2}{\sigma^2 t} \le \frac{200d^3}{(1-\lambda)^2}\frac{1}{t} \quad \forall\, \theta \in [-\pi, \pi].$$

Therefore, we can apply Lemma 6.3 to get

$$\|\varphi_Z - F\|_2 \le \|\varphi_Z - \varphi_t\|_2 + \|\varphi_t - F\|_2 \le \frac{4\cdot 10^{13} d^9}{(1-\lambda)^{10}}\frac{1}{t^{3/4}}. \tag{6.5}$$

In view of (6.4) and (6.5), we conclude that

$$\sum_{k\in\mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| \le \frac{2\sigma}{c\sqrt{2\pi}} e^{-\frac{c^2}{2\sigma^2}} + 4e^{-\frac{c^2(1-\lambda)}{20}} + |5c\sqrt{t}|^{1/2}\frac{4\cdot 10^{13} d^9}{(1-\lambda)^{10}}\frac{1}{t^{3/4}}.$$

Notice that Proposition 1.4 implies $\sigma \le (1-\lambda)^{-1/2}$. Taking $c = \sqrt{10}(1-\lambda)^{-1/2}\sqrt{\log t}$ above gives

$$\begin{aligned} \sum_{k\in\mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| &\le \frac{1}{t} + \frac{4}{\sqrt{t}} + \frac{4}{(1-\lambda)^{1/4}} t^{1/4}\log(t)^{1/4}\frac{4\cdot 10^{13} d^9}{(1-\lambda)^{10}}\frac{1}{t^{3/4}} \\ &\le \frac{1.7\cdot 10^{14} d^9}{(1-\lambda)^{41/4}}\frac{\log(t)^{1/4}}{\sqrt{t}}. \end{aligned}$$

Finally, we proceed to study the total variation distance

$$\|Z_t - N_d(t/2, t\sigma^2)\|_{TV} = \frac{1}{2}\sum_{k\in\mathbb{Z}} \left|\mathbb{P}\{Z_t = k\} - f_{N_d(t/2,t\sigma^2)}(k)\right|.$$

Using the triangle inequality and (1.1) we can upper bound the previous sum by

$$\sum_{k\in\mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| + \left|1 - \frac{1}{D(\mu, t\sigma^2)}\right| \sum_{k\in\mathbb{Z}} \phi_t(k). \tag{6.6}$$

A bound for the first term in (6.6) is given above. For the second term, we can apply Lemma 6.2 to get

$$\left|1 - \frac{1}{D(\mu, t\sigma^2)}\right| \sum_{k\in\mathbb{Z}} \phi_t(k) = \left|1 - \frac{1}{D(\mu, t\sigma^2)}\right| D(\mu, t\sigma^2) = |D(\mu, t\sigma^2) - 1| \le \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma\sqrt{t}}. \quad \square$$

7 Proof of Theorem 1.5

This section is devoted to prove Theorem 1.5. Recall that Z_t and R_t denote the Hamming weights of the random walk on a expander graph and the sticky random walk, respectively. It is easy to show that R_t satisfies a local central limit theorem. Since R_t is concentrated around its mean, this implies convergence in total variation distance to a discretized normal distribution. In view of Corollary 1.3, we just need to match the means and variances of Z_t and R_t to obtain the result. The mean and variance of a sticky random walk on $\{0,1\}$ are easy to calculate. We do it in the next lemma.

Lemma 7.1. *Let (Q_i) be the sticky random walk on $\{0,1\}$ with parameter $p \in (-1,1)$ starting from stationary distribution π, and let $R_t = \sum_{i=0}^{t-1} Q_i$. Then $\mathbb{E}_\pi(R_t) = \frac{t}{2}$ and*

$$\mathrm{Var}(R_t) = \frac{p^{t+1} - p(t+1) + t}{2(1-p)^2} - \frac{t}{4}.$$

In particular, $\lim_{t\to\infty} \mathrm{Var}(R_t)/t = \frac{1}{4}\frac{1+p}{1-p}$.

Proof. Recall that Q_0 is chosen uniformly at random on $\{0,1\}$. Hence, $\mathbb{E}_\pi(R_t) = \sum_{i=0}^{t-1} \mathbb{E}_\pi(Q_i) = \frac{t}{2}$. To calculate the variance, consider P the transition matrix of (Q_i), that is,

$$P = \begin{pmatrix} \frac{1+p}{2} & \frac{1-p}{2} \\ \frac{1-p}{2} & \frac{1+p}{2} \end{pmatrix}.$$

For any $k \in \mathbb{N}$, we can multiply P by itself k times to obtain that $P^k(1,1) = \frac{1+p^k}{2}$. Consequently,

$$\mathbb{E}_\pi(Q_0 Q_k) = \mathbb{P}\{Q_0 = 1, Q_k = 1\} = \frac{1+p^k}{4}.$$

Using the Markov condition we have

$$\begin{aligned}
\mathbb{E}_\pi(R_t^2) &= \sum_{k=0}^{t-1} \mathbb{E}_\pi(Q_k^2) + 2\sum_{k=0}^{t-1}\sum_{j=k+1}^{t-1} \mathbb{E}_\pi(Q_k Q_j) = \sum_{k=0}^{t-1} \mathbb{E}_\pi(Q_0^2) + 2\sum_{k=1}^{t-1}(t-k)\mathbb{E}_\pi(Q_0 Q_k) \\
&= \frac{t}{2} + \frac{1}{2}\sum_{k=1}^{t-1}(t-k) + \frac{1}{2}\sum_{k=1}^{t-1}(t-k)p^k = \frac{t^2}{4} - \frac{t}{4} + \frac{1}{2}\sum_{k=0}^{t-1}(t-k)p^k \\
&= \frac{p^{t+1} - p(t+1) + t}{2(1-p)^2} - \frac{t}{4} + \frac{t^2}{4}.
\end{aligned}$$

□

The next result shows that the local central limit theorem holds for sticky random walks. We obtain it as an application of Theorem 4.1.

Lemma 7.2. *Let (Q_i) be the sticky random walk on $\{0,1\}$ with parameter $p \in (-1,1)$ starting from stationary distribution π, and let $R_t = \sum_{i=0}^{t-1} Q_i$. For $\sigma^2 = \frac{1}{4}\frac{1+p}{1-p}$ we have*

$$\left|\mathbb{P}\{R_t = k\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k - t/2}{\sigma\sqrt{t}}\right)\right| \le \frac{10^{11}}{(1-|p|)^7}\frac{1}{t} \quad \forall\, k \in \mathbb{Z} \quad \forall\, t \in \mathbb{N}.$$

Proof. We will decompose R_t into a sum of two independent random variables and use Theorem 4.1 to obtain the result. Let (Q_i^2) be the 2-steps Markov chain, that is, the markov chain with transition matrix P^2. Let N_t be a random variable that counts the number of times that $Q_i^2 \neq Q_{i+1}^2$ within $(Q_0, \ldots, Q_{t-1})$. Let $I = \{i\colon Q_i^2 \neq Q_{i+1}^2\}$ and denote its elements as $i_1, \ldots, i_{N_t}$. Since $\mathbb{P}\left\{Q_i^2 \neq Q_{i+1}^2\right\} = \frac{1-p^2}{2}$ independently of previous values of the chain, we deduce that N_t follows a binomial $\mathrm{Bin}(\lfloor\frac{t-1}{2}\rfloor, \frac{1-p^2}{2})$. For every $k \in \{1, \ldots, N_t\}$, let $V_k = Q_{2i_k+1}$ be the bit that we skip to go from $Q_{i_k}^2$ to $Q_{i_k+1}^2$. Then

(V_k) is a sequence of independent Bernoulli random variables with parameter $\frac{1}{2}$. Take $b_t = \left\lfloor \frac{\mathbb{E}_\pi(N_t)}{2} \right\rfloor$ and define the random variables

$$S_t' = \sum_{k=1}^{\min\{b_t, N_t\}} V_k, \quad Y_t = R_t - S_t', \quad S_t = S_t' + \sum_{k=1}^{(b_t - N_t)^+} V_k',$$

where (V_k') are Bernoulli random variables with parameter $\frac{1}{2}$ independent of everything else. We claim that (Y_t, S_t, R_t) satisfies the hypotheses of Theorem 4.1. First, notice that S_t and Y_t are independent. In fact, once we know that the $Q_{2i} \neq Q_{2i+2}$, the value of Q_{2i+1} does not affect the rest of the chain. Moreover, we can repeat the proof of Lemma 5.7 to obtain $\mathbb{E}_\pi(|R_t - S_t - Y_t|) \leq M/t$, where

$$M = \frac{779}{(1-|p|)^2}.$$

Indeed, since the function $f(t) = t^2 e^{-\alpha t}$ attains its maximum at $t = 2\alpha^{-1}$, Lemma 5.4 implies

$$\mathbb{P}\{N_t \leq b_t\} \leq e^{-\mathbb{E}_\pi(N_t)\varphi(1/2)} \leq e^{-\frac{(1-p^2)\varphi(1/2)}{4}(t-3)} \leq e^{3/4} \frac{4^3 e^{-2}}{(1-p^2)^2 \varphi(1/2)^2} \frac{1}{t^2} \leq \frac{779}{(1-|p|)^2} \frac{1}{t^2}.$$

Hence, the claim follows from the fact that $|R_t - S_t - Y_t| \leq t$ and that $R_t = S_t + Y_t$ if $b_t \leq N_t$. Finally, the eigenvalues of the sticky random walk with parameter p are 1 and p, so in this case we have $1 - \lambda = 1 - |p|$. Let $\theta_0 = (1-|p|)^2 \sigma^2 / 2708$. Since $\sigma^2 \geq (1-|p|)/8$, Lemma 5.8 shows that the random variables V_i and V_i' are η-nonlattice with

$$\eta \geq \frac{1}{4}(1 - \cos(\theta_0)) \geq \frac{1}{4} \frac{\theta_0^2}{5} \geq \frac{(1-|p|)^6}{20 \cdot 2708^2 \cdot 64} \geq \frac{(1-|p|)^6}{10^{10}}.$$

Therefore, Theorem 4.1 gives

$$\left| \mathbb{P}\{R_t = k\} - \frac{1}{\sigma\sqrt{t}} \phi\left(\frac{k - t/2}{\sigma\sqrt{t}}\right) \right| \leq \left(\pi M + \frac{1}{\theta_0 \sigma^2} + \frac{C_4}{\sigma^4 (1-|p|)^2} \right) \frac{1}{t} + \frac{1}{e\eta} \frac{1}{b_t}. \tag{7.1}$$

If $t \leq 22(1-p^2)^{-1}$, then there is nothing to prove. Otherwise, we have

$$\begin{aligned} b_t = \left\lfloor \frac{\mathbb{E}_\pi(N_t)}{2} \right\rfloor &\geq \frac{\left\lfloor \frac{t-1}{2} \right\rfloor \frac{1-p^2}{2}}{2} - 1 \geq \frac{(t-3)(1-p^2)}{8} - 1 \geq \frac{t(1-p^2) - 11}{8} \geq \frac{(1-p^2)}{16} t \\ &\geq \frac{(1-|p|)}{16} t. \end{aligned}$$

Substituting all previous bounds in (7.1) gives

$$\left| \mathbb{P}\{R_t = k\} - \frac{1}{\sigma\sqrt{t}} \phi\left(\frac{k - t/2}{\sigma\sqrt{t}}\right) \right| \leq \left(\frac{779\pi}{(1-|p|)^2} + \frac{10^6}{(1-|p|)^4} + \frac{8 \cdot 10^{10}}{(1-|p|)^7} \right) \frac{1}{t} \leq \frac{10^{11}}{(1-|p|)^7} \frac{1}{t}.$$

□

We can repeat word by word the proof of Corollary 1.3 using R_t instead of Z_t and Lemma 7.2 instead of Theorem 1.1 to produce the following result.

Lemma 7.3. *Let (Q_i) be the sticky random walk on $\{0,1\}$ with parameter $p \in (-1,1)$ starting from stationary distribution, and let $R_t = \sum_{i=0}^{t-1} Q_i$. For $\sigma^2 = \frac{1}{4}\frac{1+p}{1-p}$ we have*

$$\left\| R_t - N_d(t/2, t\sigma^2) \right\|_{TV} \leq \frac{10^{12}}{(1-|p|)^8} \frac{\sqrt{\log t}}{\sqrt{t}}.$$

Proof. As in the proof of Corollary 1.3, take $c \geq 1$ and let $I_1 = \{u \in \mathbb{Z} \colon |u - t/2| \leq c\sqrt{t} + 1\}$ and $I_2 = \mathbb{Z} \setminus I_1$. First, we bound

$$\sum_{k \in \mathbb{Z}} \left| \mathbb{P}\left\{R_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k - t/2}{t^{1/2}\sigma}\right)\right|$$

by breaking the sum using I_1 and I_2. Write $C(p) = 10^{11}(1 - |p|)^{-7}$. Lemma 7.2 gives

$$\sum_{k \in I_1} \left| \mathbb{P}\left\{R_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k - t/2}{t^{1/2}\sigma}\right)\right| \leq (2c\sqrt{t} + 3)\frac{C(p)}{t} \leq 5C(p)\frac{c}{\sqrt{t}}.$$

To bound the sum on I_2 we study the tails of the distributions of R_t and $\mathcal{N}(t/2, t\sigma^2)$. Since we can also apply Lemma 5.5 to study the tail of R_t, from the proof of Corollary 1.3 it follows that

$$\sum_{k \in I_2} \left| \mathbb{P}\left\{R_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k - t/2}{t^{1/2}\sigma}\right)\right| \leq \frac{2\sigma}{c\sqrt{2\pi}}e^{-\frac{c^2}{2\sigma^2}} + 4e^{-\frac{c^2(1-|p|)}{20}}.$$

Write $\alpha = \max\{\sigma, \sqrt{10}(1 - |p|)^{-1/2}\}$ and take $c = \alpha\sqrt{\log t}$. From the previous bounds we obtain

$$\sum_{k \in \mathbb{Z}} \left| \mathbb{P}\left\{R_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k - t/2}{t^{1/2}\sigma}\right)\right| \leq \left(5C(p)\alpha\sqrt{\log t} + \frac{\sqrt{2}}{\sqrt{\pi}} + 4\right)\frac{1}{\sqrt{t}}.$$

Recall that the total variation distance between R_r and $N_d(t/2, t\sigma^2)$ is given by

$$\|R_t - N_d(t/2, t\sigma^2)\|_{TV} = \frac{1}{2}\sum_{k \in \mathbb{Z}} \left| \mathbb{P}\left\{R_t = k\right\} - f_{N_d(t/2, t\sigma^2)}(k)\right|.$$

We can use the triangle inequality and repeat the argument in the proof of Corollary 1.3 to bound (6.6) to obtain that

$$\sum_{k \in \mathbb{Z}} \left| \mathbb{P}\left\{R_t = k\right\} - f_{N_d(t/2, t\sigma^2)}(k)\right| \leq \left(5C(p)\alpha\sqrt{\log t} + \frac{\sqrt{2}}{\sqrt{\pi}} + 4\right)\frac{1}{\sqrt{t}} + \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma\sqrt{t}}.$$

We clearly have $\sigma \leq (1 - |p|)^{-1}$, whence $\alpha \leq \sqrt{10}(1 - |p|)^{-1}$. We conclude that

$$\|R_t - N_d(t/2, t\sigma^2)\|_{TV} \leq \frac{1}{2}\left(5C_1(\lambda, d)\alpha\sqrt{\log t} + \frac{\sqrt{2}}{\sqrt{\pi}} + 4 + \frac{1}{\sigma\sqrt{2\pi}}\right)\frac{1}{\sqrt{t}} \leq \frac{3\sqrt{10}C(p)}{1 - |p|}\frac{\sqrt{\log t}}{\sqrt{t}}.$$

□

The proof of Theorem 1.5 is now immediate. Although we do not optimize the constant, we show that Theorem 1.5 holds with

$$C_3(\lambda, d) = \frac{2 \cdot 10^{23} d^{24}}{(1 - \lambda)^{16}}.$$

Proof of Theorem 1.5. We want to write the bound of Lemma 7.3 in terms of λ and d. Write $\sigma^2 = \lim_{t \to \infty} \operatorname{Var}(Z_t)/t$ and recall that p is chosen so that $\sigma^2 = \frac{1}{4}\frac{1+p}{1-p}$, that is,

$$p = \frac{4\sigma^2 - 1}{1 + 4\sigma^2}.$$

If $\sigma^2 \in (0, 1/4)$, then we have $|p| \leq 1 - 4\sigma^2$, and if $\sigma^2 \in (1/4, \infty)$, then $|p| \leq 1 - \frac{1}{4\sigma^2}$. Therefore,

$$1 - |p| \geq \min\left\{4\sigma^2, \frac{1}{4\sigma^2}\right\}.$$

Recall that the bound (5.12) gives $4\sigma^2 \geq \frac{(1-\lambda)^2}{25d^3}$. Moreover, Propositon 1.4 implies $4\sigma^2 \leq 2(1-\lambda)^{-1}$, so we have

$$1 - |p| \geq \frac{(1-\lambda)^2}{25d^3}.$$

Therefore, Lemma 7.3 gives

$$\left\| R_t - N_d(t/2, t\sigma^2) \right\|_{TV} \leq \frac{10^{12} \cdot 25^8 d^{24}}{(1-\lambda)^{16}} \frac{\sqrt{\log t}}{\sqrt{t}} \leq \frac{1.6 \cdot 10^{23} d^{24}}{(1-\lambda)^{16}} \frac{\sqrt{\log t}}{\sqrt{t}}.$$

The result now follows from Corollary 1.3 and the triangle inequality. □

8 Generalization to all labellings

In this section, we extend Theorem 1.1 and Corollary 1.3 to allow unbalanced labellings. First, we need a generalization of Lemma 5.3. Recall that $A = \{x \in V\colon \operatorname{val}(x) = 0\}$ and $B = A^{\mathsf{c}}$.

Lemma 8.1. *Let G be a d-regular λ-expander graph with n vertices. Fix a labelling $\operatorname{val}\colon V \longrightarrow \{0,1\}$ on G with $\mathbb{E}_\pi(\operatorname{val}) = \alpha \in [0,1]$. Write $\delta = \frac{1}{4}(1-\lambda)^2\alpha(1-\alpha)$. Then there is $k^* \in \{1, \dots, d-1\}$ such that either*

$$|A_{k^*}| \geq \frac{\delta(1-\alpha)n}{d-1} \quad \text{or} \quad |B_{k^*}| \geq \frac{\delta\alpha n}{d-1}.$$

Proof. We will follow the argument in the proof of Lemma 5.3. Suppose the statement is false. Then

$$|A_0| + |A_d| > (1-\delta)(1-\alpha)n \quad \text{and} \quad |B_0| + |B_d| > (1-\delta)\alpha n. \tag{8.1}$$

Take $F_1 = A$ and $F_2 = B$. Corollary 5.2 gives $|E(A,B)| \geq \frac{1}{2}(1-\lambda)d\alpha(1-\alpha)n$. Notice also that

$$|E(A,B)| \leq d|A \setminus A_d| \leq d(|A_0| + \delta|A|).$$

Therefore, we obtain

$$\begin{aligned} |A_0| &\geq \frac{1}{2}(1-\lambda)\alpha(1-\alpha)n - \delta(1-\alpha)n = \frac{1}{2}(1-\lambda)\alpha(1-\alpha)n\left(1 - \frac{1}{2}(1-\lambda)(1-\alpha)\right) \\ &\geq \frac{1}{4}(1-\lambda)\alpha(1-\alpha)n. \end{aligned} \tag{8.2}$$

A completely analogous argument replacing A_d with B_0 and A_0 with B_d gives

$$|B_d| \geq \frac{1}{2}(1-\lambda)\alpha(1-\alpha)n - \delta\alpha n \geq \frac{1}{4}(1-\lambda)\alpha(1-\alpha)n. \tag{8.3}$$

Next, consider $F_1 = F_2 = A$. Then the expander mixing lemma 5.1 gives

$$|E(A,A)| \geq (1-\alpha)^2 dn - \alpha(1-\alpha)\lambda dn = ((1-\alpha) - \alpha\lambda)(1-\alpha)dn.$$

Moreover, $|E(A,A)| \leq d|A \setminus A_0| \leq d(|A_d| + \delta|A|)$, from where we deduce that

$$|A_d| \geq ((1-\alpha) - \alpha\lambda - \delta)(1-\alpha)n.$$

We can take $F_1 = F_2 = B$ and repeat the previous argument, replacing A_d with B_0 and A_0 with B_d, to obtain that

$$|B_0| \geq (\alpha - (1-\alpha)\lambda - \delta)\alpha n.$$

Adding these bounds gives

$$\begin{aligned}|A_d| + |B_0| &\geq \left((1-\alpha)^2 - 2\alpha(1-\alpha)\lambda + \alpha^2\right) n - \delta n \\ &= ((1-\alpha)-\alpha)^2 n + 2\alpha(1-\alpha)(1-\lambda)n - \delta n \\ &\geq 2\alpha(1-\alpha)(1-\lambda)n - \delta n \geq \frac{1}{2}\alpha(1-\alpha)(1-\lambda)n.\end{aligned}$$

Finally, take $F_1 = A_0 \cup B_d$ and $F_2 = {F_1}^{\mathsf{c}}$. In view of the above bound, (8.2), and (8.3), we get $\min\{|F_1|, |F_2|\} \geq \frac{1}{2}(1-\lambda)\alpha(1-\alpha)n$. Hence, Corollary 5.2 gives

$$|E(F_1, F_2)| \geq \frac{1}{4} d(1-\lambda)^2\alpha(1-\alpha)n = d\delta n.$$

Therefore, we must have either $|E(A_0, F_2)| \geq d\delta\alpha n$ or $|E(B_d, F_2)| \geq d\delta(1-\alpha)n$. In the first case, for any $e = \{x, y\} \in E(A_0, F_2)$ with $x \in A_0$, we must have $\mathrm{val}(y) = 1$. Thus, $y \in B \setminus B_d$. Moreover, y cannot belong to B_0 since it is adjacent to x and $\mathrm{val}(x) = 0$. Therefore, $y \in B \setminus (B_0 \cup B_d)$. Consequently,

$$|B \setminus (B_0 \cup B_d)| \geq \frac{|E(A_0, F_2)|}{d} \geq \delta\alpha n,$$

which contradicts (8.1). Analogously, $|E(B_d, F_2)| \geq d\delta(1-\alpha)n$ also leads to a contradiction. □

To extend our main result for unbalanced labellings we just need to repeat its proof using Lemma 8.1 instead of Lemma 5.3.

Theorem 8.2. *Let G be a d-regular λ-expander graph with $\lambda < 1$. Let (X_i) be the simple random walk on G with uniform initial distribution π, fix a labelling $\mathrm{val}\colon V \longrightarrow \{0,1\}$ with $\alpha = \mathbb{E}_\pi(\mathrm{val}) \in (0,1)$, and let $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. There is a constant $C_5(\lambda, d, \alpha)$ depending on λ, d, and α such that*

$$\left|\mathbb{P}\left\{Z_t = k\right\} - t^{-1/2}\sigma^{-1}\phi\left(\frac{k-\alpha t}{t^{1/2}\sigma}\right)\right| \leq C_5(\lambda, d, \alpha)\frac{1}{t} \quad \forall\, k \in \mathbb{Z} \quad \forall\, t \in \mathbb{N}.$$

Although the proof of Theorem 8.2 does not optimize the constant, it shows that we can take

$$C_5(\lambda, d, \alpha) = \frac{4 \cdot 10^{12} d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{10}}.$$

Proof. Let $\delta = \frac{1}{4}(1-\lambda)^2\alpha(1-\alpha)$. Lemma 8.1 guarantees that there is $k^* \in \{1, \ldots, d-1\}$ such that either

$$|A_{k^*}| \geq \frac{\delta(1-\alpha)n}{d-1} \quad \text{or} \quad |B_{k^*}| \geq \frac{\delta\alpha n}{d-1}.$$

By symmetry, we may assume that $|A_{k^*}| \geq \frac{\delta(1-\alpha)n}{d-1}$. Fix $t \in \mathbb{N}$ and consider the random variables $\widetilde{N}_t$, V_i, $\widetilde{V}_i$, S_t', S_t, and Y_t appearing in (5.9). As the proof of Theorem 1.1 showed, the hypotheses of Theorem 4.1 are satisfied, so for any $k \in \mathbb{Z}$ we have

$$\left|\mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k-\alpha t}{\sigma\sqrt{t}}\right)\right| \leq \left(\pi M + \frac{1}{\theta_0\sigma^2} + \frac{C_4}{\sigma^4(1-\lambda)^2}\right)\frac{1}{t} + \frac{1}{e\eta}\frac{1}{b_t}. \tag{8.4}$$

The value of M is the same as in the proof of Theorem 1.1. In contrast, b_t now depends on α, so we need to find a new bound for it. Recall that $b_t = \lfloor \mathbb{E}_\pi(\widetilde{N}_t)/4 \rfloor$, where

$$\frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} = \frac{\mathbb{E}_\pi(N_t)}{4d} = \frac{\lfloor t/2 \rfloor \pi(A_{k^*})}{4d} \geq \frac{(t-2)\delta(1-\alpha)}{8d(d-1)} \geq \frac{t\delta(1-\alpha)-2}{8d(d-1)}.$$

If $t \le 8\delta^{-1}(1-\alpha)^{-1}d^3$, then there is nothing to prove. Otherwise, a simple computation shows that

$$b_t = \left\lfloor \frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} \right\rfloor \ge \frac{\mathbb{E}_\pi(\widetilde{N}_t)}{4} - 1 \ge \frac{t\delta(1-\alpha) - 48d^2}{48d(d-1)} \ge \frac{\delta(1-\alpha)}{8d^2} t = \frac{\alpha(1-\alpha)^2(1-\lambda)^2}{32d^2} t. \tag{8.5}$$

Thus, the argument used to get the bound (5.12) for σ shows that

$$\sigma^2 \ge \frac{b_t}{t} \frac{d-1}{d^2} \ge \frac{\alpha(1-\alpha)^2(1-\lambda)^2}{64d^3}. \tag{8.6}$$

Finally, to obtain a bound for η take

$$\theta_0 = \frac{(1-\lambda)^2\sigma^2}{2708}.$$

Then following the proof of Theorem 1.1 we get

$$\eta \ge \frac{d-1}{d^2}(1-\cos(\theta_0)) \ge \frac{1}{2d}\frac{\theta_0^2}{5} \ge \frac{1}{10\cdot 64^2 \cdot 2708^2}\frac{\alpha^2(1-\alpha)^4(1-\lambda)^8}{d^7} \ge \frac{\alpha^2(1-\alpha)^4(1-\lambda)^8}{3.1\cdot 10^{11} d^7}. \tag{8.7}$$

After substituting the previous bounds in (8.4), a simple computation gives

$$\left| \mathbb{P}\left\{Z_t = k\right\} - \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{k-\alpha t}{\sigma\sqrt{t}}\right) \right| \le \frac{4\cdot 10^{12} d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{10}} \frac{1}{t}.$$ □

Similarly, we can repeat word by word the proof of Corollary 1.3 to extend it for unbalanced labellings.

Corollary 8.3. *Let G be a d-regular λ-expander graph with $\lambda < 1$. Let (X_i) be the simple random walk on G with uniform initial distribution π, fix a labelling $\mathrm{val}\colon V \longrightarrow \{0,1\}$ with $\alpha = \mathbb{E}_\pi(\mathrm{val}) \in (0,1)$, and let $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$ and $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$. There is a constant $C_6(\lambda, d, \alpha)$ depending on λ, d, and α such that*

$$\left\| Z_t - N_d(t/2, t\sigma^2) \right\|_{TV} \le C_6(\lambda, d, \alpha) \frac{\log(t)^{1/4}}{\sqrt{t}} \quad \forall\, t \ge 2.$$

The proof of Corollary 8.3 does not optimize the constant, but it shows that we can take

$$C_6(\lambda, d, \alpha) = \frac{2\cdot 10^{13} d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{41/4}}.$$

We just need to verify that every result used for the proof of Corollary 1.3 is also valid for unbalanced labellings. First, Lemma 6.2 was proved for arbitrary normal distributions, so it can be used in the unbalanced setting. Next, we can repeat word by word the proof of Lemma 6.3 to extend it for unbalanced labellings. For convenience, we use φ_t for the characteristic function of a normal distribution with mean $t\alpha$ and variance $t\sigma^2$, where $\sigma^2 = \lim_{t\to\infty} \mathrm{Var}(Z_t)/t$.

Lemma 8.4. *The characteristic function of Z_t satisfies*

$$\|\varphi_Z - \varphi_t\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta\right)^{1/2} \le \frac{8\cdot 10^{12} d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{10}} \frac{1}{t^{3/4}}.$$

Proof. Set $\theta_0 = (1-\lambda)^2\sigma^2/2708$. For any $\theta \in [-\theta_0, \theta_0]$, Proposition 4.2 gives

$$|\varphi_Z(\theta) - \varphi_t(\theta)| \le c e^{-t\sigma^2\theta^2/8}\left(683t|\theta|^3 + \frac{20}{\sigma^2}|\theta|\right),$$

where $c = (1-\lambda)^{-2}$. By repeating the computations in (6.3) we obtain

$$\int_{-\theta_0}^{\theta_0} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta \le \frac{5 \cdot 10^8 c^2}{\sigma^7 t^{3/2}}. \tag{8.8}$$

It remains to study the case $\theta_0 < |\theta| \le \pi$. As we did for Lemma 6.3, consider $Z_t' = S_t + Y_t$, where S_t and Y_t are defined as in (5.9) and let $\varphi_{Z'}$ be its characteristic function. Recall that (4.5) gives $|\varphi_{Z'}(\theta)| \le \frac{1}{e\eta b_t}$ for $\theta_0 \le |\theta| \le \pi$. Moreover, the proof of Theorem 4.1 shows that

$$|\varphi_Z(\theta) - \varphi_{Z'}(\theta)| \le \pi \mathbb{E}_\pi(|Z_t - Z_t'|) \le \frac{\pi M}{t} \quad \forall\, \theta \in [-\pi, \pi],$$

where $M = \frac{10^{10} d^3}{(1-\lambda)^4}$ in view of Lemma 5.7. Consequently,

$$\int_{\theta_0 \le |\theta| \le \pi} |\varphi_Z(\theta)|^2 \, d\theta \le \frac{4\pi}{e^2 \eta^2 b_t^2} + \frac{2\pi^3 M^2}{t^2}.$$

On the other hand, we showed in the proof of Lemma 6.3 that

$$\int_{\theta_0 \le |\theta| \le \pi} |\varphi_t(\theta)|^2 \, d\theta \le \frac{1}{\theta_0^2 \sigma^3 t^{3/2}}.$$

Putting everything together gives

$$\|\varphi_Z - \varphi_t\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\varphi_Z(\theta) - \varphi_t(\theta)|^2 \, d\theta \le \frac{10^8 c^2}{\sigma^7 t^{3/2}} + \frac{4}{e^2 \eta^2 b_t^2} + \frac{2\pi^2 M^2}{t^2} + \frac{1}{\pi \theta_0^2 \sigma^3 t^{3/2}}.$$

Taking square root in both sides and using that $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ yields

$$\|\varphi_Z - \varphi_t\|_2 \le \frac{10^4 c}{\sigma^{7/2} t^{3/4}} + \frac{2}{e\eta b_t} + \frac{\sqrt{2}\pi M}{t} + \frac{1}{\pi^{1/2} \theta_0 \sigma^{3/2} t^{3/4}}.$$

Recall that b_t, σ, and η, are bounded in (8.5), (8.6), and (8.7), respectively. The claim follows from these bounds and the above inequality after straightforward computations. □

Proof of Corollary 8.3. The argument is analogous to the one in the proof of Corollary 1.3. Given $c \ge 1$, set $I_1 = \{k \in \mathbb{Z} \colon |k - t\alpha| \le c\sqrt{t} + 1\}$ and $I_2 = \mathbb{Z} \setminus I_1$. First, we bound

$$\sum_{k \in \mathbb{Z}} |\mathbb{P}\{Z_t = k\} - \phi_t(k)|$$

by breaking the sum using I_1 and I_2. To bound the sum on I_2 we study the tails of the distributions of Z_t and $\mathcal{N}(t\alpha, t\sigma^2)$. Let ϕ_t denote the density function of $\mathcal{N}(t\alpha, t\sigma^2)$. On one hand,

$$\sum_{k \in I_2} \phi_t(k) \le 2 \int_{t\alpha + c\sqrt{t}}^{\infty} \phi_t(x) \, dx = 2 \int_{c/\sigma}^{\infty} \phi(y) \, dy \le \frac{2}{\sqrt{2\pi}} \int_{c/\sigma}^{\infty} \frac{y}{c/\sigma} e^{-y^2/2} \, dy = \frac{2\sigma}{c\sqrt{2\pi}} e^{-\frac{c^2}{2\sigma^2}}.$$

On the other hand, Lemma 5.5 gives

$$\sum_{k \in I_2} \mathbb{P}\{Z_t = k\} = \mathbb{P}\left\{|Z_t - t\alpha| > c\sqrt{t} + 1\right\} \le 4e^{-\frac{(c\sqrt{t}+1)^2(1-\lambda)}{20t}} \le 4e^{-\frac{c^2(1-\lambda)}{20}}.$$

Consequently, we have

$$\sum_{k \in I_2} |\mathbb{P}\{Z_t = k\} - \phi_t(k)| \le \frac{2\sigma}{c\sqrt{2\pi}} e^{-\frac{c^2}{2\sigma^2}} + 4e^{-\frac{c^2(1-\lambda)}{20}}. \tag{8.9}$$

For the sum on I_1, the Cauchy-Schwarz inequality yields

$$\sum_{k\in I_1} |\mathbb{P}\{Z_t=k\}-\phi_t(k)| \le |I_1|^{1/2}\left(\sum_{k\in I_1} |\mathbb{P}\{Z_t=k\}-\phi_t(k)|^2\right)^{1/2}.$$

We have $|I_1|\le 2c\sqrt{t}+3\le 5c\sqrt{t}$. As in the proof of Corollary 1.3, to estimate the sum on the right hand side we will use Parseval's identity (see [9, §1.4]). Define $F\colon[-\pi,\pi]\longrightarrow\mathbb{C}$ by

$$F(\theta)=\sum_{k\in\mathbb{Z}}\varphi_t(\theta+2\pi k)\quad \forall\,\theta\in[-\pi,\pi],$$

where φ_t denotes the ch.f. of the normal distribution with mean $t\alpha$ and variance $t\sigma^2$. Then the Fourier coefficients of F, denoted by $a_k(F)$, satisfy

$$a_k(F)=\frac{1}{2\pi}\widehat{\varphi_t}(k)=\phi_t(k)\quad\forall\,k\in\mathbb{Z},$$

where $\widehat{\varphi_t}$ is the Fourier transform of φ_t as defined in [9, §1.2]. Consequently, Parseval's identity gives

$$\sum_{k\in\mathbb{Z}}|\mathbb{P}\{Z_t=k\}-\phi_t(k)|^2=\|\varphi_Z-F\|_2^2.$$

By repeating the computations in the proof of Corollary 1.3 we obtain

$$|\varphi_t(\theta)-F(\theta)|\le\frac{200d^3}{(1-\lambda)^2}\frac{1}{t}\quad\forall\,\theta\in[-\pi,\pi].$$

Therefore, from Lemma 8.4 and the triangle inequality we deduce that

$$\|\varphi_Z-F\|_2\le\|\varphi_Z-\varphi_t\|_2+\|\varphi_t-F\|_2\le\frac{9\cdot10^{12}d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{10}}\frac{1}{t^{3/4}}. \tag{8.10}$$

In view of (8.9) and (8.10), we conclude that

$$\sum_{k\in\mathbb{Z}}|\mathbb{P}\{Z_t=k\}-\phi_t(k)|\le\frac{2\sigma}{c\sqrt{2\pi}}e^{-\frac{c^2}{2\sigma^2}}+4e^{-\frac{c^2(1-\lambda)}{20}}+|5c\sqrt{t}|^{1/2}\frac{9\cdot10^{12}d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{10}}\frac{1}{t^{3/4}}.$$

Notice that Proposition 1.4 implies $\sigma\le(1-\lambda)^{-1/2}$. Taking $c=\sqrt{10}(1-\lambda)^{-1/2}\sqrt{\log t}$ above gives

$$\sum_{k\in\mathbb{Z}}|\mathbb{P}\{Z_t=k\}-\phi_t(k)|\le\frac{3.7\cdot10^{13}d^9}{\alpha^3(1-\alpha)^6(1-\lambda)^{41/4}}\frac{\log(t)^{1/4}}{\sqrt{t}}.$$

Finally, we proceed to study the total variation distance

$$\|Z_t-N_d(t/2,t\sigma^2)\|_{TV}=\frac{1}{2}\sum_{k\in\mathbb{Z}}\left|\mathbb{P}\{Z_t=k\}-f_{N_d(t/2,t\sigma^2)}(k)\right|.$$

Using the triangle inequality and (1.1) we can upper bound the previous sum by

$$\sum_{k\in\mathbb{Z}}|\mathbb{P}\{Z_t=k\}-\phi_t(k)|+\left|1-\frac{1}{D(\mu,t\sigma^2)}\right|\sum_{k\in\mathbb{Z}}\phi_t(k). \tag{8.11}$$

The first term in (8.11) was bounded above. For the second term, we can apply Lemma 6.2 to get

$$\left|1-\frac{1}{D(\mu,t\sigma^2)}\right|\sum_{k\in\mathbb{Z}}\phi_t(k)=\left|1-\frac{1}{D(\mu,t\sigma^2)}\right|D(\mu,t\sigma^2)=|D(\mu,t\sigma^2)-1|\le\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma\sqrt{t}}.\quad\square$$

9 Proofs of Example 2.1 and Example 2.2

Let (X_i) denote the SRW on a graph $G = (V, E)$ starting from stationary distribution π, and write $(Y_i) = (\mathrm{val}(X_i))$, where $\mathrm{val}\colon V \longrightarrow \{0,1\}$ is a labelling on G. First, we prove Example 2.1.

Proof of Example 2.1. We want to use the formula (6.1) for the variance of a sum of random variables. We claim that

$$\mathrm{Cov}(Y_0, Y_k) = \frac{1}{4(-3)^k} \quad \forall\, k \in \mathbb{N}.$$

Let $V = \{a, b, c, d\}$ be the set of vertices of K_4. Without loss of generality, we may assume that the balanced labelling that we consider satisfies $\mathrm{val}(a) = \mathrm{val}(b) = 1$ and $\mathrm{val}(c) = \mathrm{val}(d) = 0$. One can check that the transition matrix of the random walk on K_4 can be diagonalized as

$$\begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{-1}{3} & 0 & 0 & 0 \\ 0 & \frac{-1}{3} & 0 & 0 \\ 0 & 0 & \frac{-1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1/4 & -1/4 & -1/4 & 3/4 \\ -1/4 & -1/4 & 3/4 & -1/4 \\ -1/4 & 3/4 & -1/4 & -1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

This shows that K_4 is a $\frac{1}{3}$-expander graph. Moreover, using the above diagonalization one can compute powers of the transition matrix to get

$$P^k(a,a) = \frac{3^k + 3\cdot(-1)^k}{4 \cdot 3^k} \quad \text{and} \quad P^k(a,b) = \frac{3^k - (-1)^k}{4\cdot 3^k} \quad \forall\, k \in \mathbb{N}.$$

Therefore, for any $k \in \mathbb{N}$ we have

$$\begin{aligned} \mathbb{E}_\pi(Y_0 Y_k) &= \mathbb{P}\{Y_0 = 1, Y_k = 1\} = \frac{1}{2}\mathbb{P}\{Y_k = 1 | Y_0 = 1\} = \frac{1}{2}(P^k(a,a) + P^k(a,b)) \\ &= \frac{1}{4}\left(1 + \frac{1}{(-3)^k}\right). \end{aligned}$$

The claim follows from that fact that $\mathbb{E}_\pi(Y_0)\mathbb{E}_\pi(Y_k) = \frac{1}{4}$. In view of (6.1), adding the covariances gives the result since

$$2\sum_{i<j} \mathrm{Cov}(Y_i, Y_j) = 2\sum_{k=1}^{t-1}(t-k)\,\mathrm{Cov}(Y_0, Y_k) = \frac{1}{2}\sum_{k=1}^{t-1}(t-k)\frac{1}{(-3)^k} = -\frac{1}{8}t + O(1). \qquad \square$$

Example 2.2 also follows from computing the covariances and applying the formula (6.1) for the variance of a sum of random variables.

Proof of Example 2.2. Select a labelling val uniformly at random from all balanced labellings on G, and consider $Z_t = \sum_{i=0}^{t-1} \mathrm{val}(X_i)$. The variance of Z_t, which is taken with respect to the uniform distribution on the set of all balanced labellings on G, will be denoted by $\mathrm{Var}(Z_t)$. When the labelling val is fixed, we will write $\mathrm{Var}(Z_t | \,\mathrm{val})$ for its variance. We claim that

$$\mathrm{Var}(Z_t) \geq \frac{t}{4} + \frac{1}{2}\left(\frac{1}{d} - \frac{3}{n-1}\right)t + O(1).$$

Observe that the law of total variance tells us that

$$\mathrm{Var}(Z_t) = \mathbb{E}(\mathrm{Var}(Z_t | \,\mathrm{val})) + \mathrm{Var}(\mathbb{E}(Z_t | \,\mathrm{val})).$$

Since the initial distribution of (X_i) is the stationary one, we have that $\mathbb{E}(Z_t|\,\mathrm{val}) = t/2$ for any balanced labelling on G, whence $\mathrm{Var}(\mathbb{E}(Z_t|\,\mathrm{val})) = 0$. Therefore, the claim implies

$$\mathbb{E}(\mathrm{Var}(Z_t|\,\mathrm{val})) \geq \frac{t}{4} + \frac{1}{2}\left(\frac{1}{d} - \frac{3}{n-1}\right)t + O(1),$$

thus there must be a balanced labelling val for which $\mathrm{Var}(Z_t|\,\mathrm{val})$ is greater or equal than the right hand side. To prove the claim, we first calculate $\mathbb{E}(Y_0 Y_k)$, where the expectation is taken with respect to the uniform distribution on the set of all balanced labellings on G, and $Y_i = \mathrm{val}(X_i)$. Since the labeling is selected uniformly at random, observe that

$$\begin{aligned}\mathbb{E}(Y_0Y_k) &= \mathbb{P}\left\{Y_k = 1, Y_0 = 1 | X_0 = X_k\right\}\mathbb{P}\left\{X_k = X_0\right\} + \mathbb{P}\left\{Y_k = 1, Y_0 = 1 | X_0 \neq X_k\right\}\mathbb{P}\left\{X_k \neq X_0\right\} \\ &= \frac{1}{2}\mathbb{P}\left\{X_k = X_0\right\} + \frac{\frac{n}{2}-1}{2n-2}\mathbb{P}\left\{X_k \neq X_0\right\} = \frac{1}{2}\mathbb{P}\left\{X_k = X_0\right\} + \frac{1}{4}\left(1 - \frac{1}{n-1}\right)\mathbb{P}\left\{X_k \neq X_0\right\} \\ &= \frac{1}{4} + \frac{1}{4}\left(\mathbb{P}\left\{X_k = X_0\right\} - \frac{1}{n-1}\mathbb{P}\left\{X_k \neq X_0\right\}\right) = \frac{1}{4} + \frac{1}{4}\left(\frac{n}{n-1}\mathbb{P}\left\{X_k = X_0\right\} - \frac{1}{n-1}\right).\end{aligned}$$

Recall that $\mathbb{E}(Y_0) = \mathbb{E}(Y_k) = \frac{1}{2}$, which gives

$$\mathrm{Cov}(Y_0, Y_k) = \frac{1}{4}\left(\frac{n}{n-1}\mathbb{P}\left\{X_k = X_0\right\} - \frac{1}{n-1}\right).$$

Adding all covariances gives

$$\begin{aligned}\sum_{i<j}\mathrm{Cov}(Y_i, Y_j) &= \frac{1}{4}\frac{n}{n-1}\sum_{k=1}^{t-1}(t-k)\mathbb{P}\left\{X_k = X_0\right\} - \frac{1}{4}\sum_{k=1}^{t-1}(t-k)\frac{1}{n-1} \\ &= \frac{1}{4}\frac{n}{n-1}\sum_{k=1}^{t-1}(t-k)\mathbb{P}\left\{X_k = X_0\right\} - \frac{t(t-1)}{8(n-1)}.\end{aligned}$$

Write P for the transition matrix of (X_i). We can use the spectral expansion of P to calculate $\mathbb{P}\left\{X_k = X_0\right\}$. Let (λ_j) be the eigenvalues of P and let (f_j) be an orthonormal basis of $(\mathbb{R}^V, \langle\cdot,\cdot\rangle_\pi)$ corresponding to (λ_j). Lemma 12.2 in [12] gives for any $k \in \mathbb{N}$

$$P^k(x,y) = \sum_{j=1}^{n} f_j(x) f_j(y) \lambda_j^k \pi(y) \quad \forall\, x, y \in V,$$

from where we deduce that

$$\begin{aligned}\mathbb{P}\left\{X_k = X_0\right\} &= \frac{1}{n}\sum_{x\in V} P^k(x,x) = \frac{1}{n}\sum_{x\in V}\sum_{j=1}^{n} f_j(x)^2\lambda_j^k\pi(x) = \frac{1}{n}\sum_{j=1}^{n}\lambda_j^k\sum_{x\in V} f_j(x)^2\pi(x) \\ &= \frac{1}{n}\sum_{j=1}^{n}\lambda_j^k.\end{aligned}$$

Since $\lambda_1 = 1$, it is clear that $\mathbb{P}\left\{X_k = X_0\right\} \geq \frac{1}{n}$ for any even $k \in \mathbb{N}$. Moreover, for any even $k \in \mathbb{N}$ we have

$$\mathbb{P}\left\{X_k = X_0\right\} + \mathbb{P}\left\{X_{k+1} = X_0\right\} = \frac{1}{n}\sum_{j=1}^{n}\lambda_j^k + \lambda_j^{k+1} = \frac{1}{n}\sum_{j=1}^{n}\lambda_j^k(1+\lambda_j) \geq \frac{2}{n}.$$

Therefore,

$$\begin{aligned}\sum_{k=1}^{t-1}(t-k)\mathbb{P}\left\{X_k = X_0\right\} &\geq (t-2)\mathbb{P}\left\{X_2 = X_0\right\} + \frac{1}{n}\sum_{k=4}^{t-1}(t-k) \\ &= (t-2)\frac{1}{d} + \frac{1}{n}\left(\frac{t(t-1)}{2} - 3t + 6\right).\end{aligned}$$

Consequently,

$$\sum_{i<j} \mathrm{Cov}(Y_i, Y_j) \geq \frac{1}{4}\left((t-2)\frac{1}{d} + \frac{1}{n-1}\left(\frac{t(t-1)}{2} - 3t + 6\right)\right) - \frac{t(t-1)}{8(n-1)}$$
$$= \frac{1}{4}\left(\frac{1}{d} - \frac{3}{n-1}\right)t + O(1).$$

The claim now follows from the formula (6.1) for the variance of a sum of random variables. □

References

[1] Rafael Chiclana and Yuval Peres. A local central limit theorem for random walks on expander graphs, 2022. Preprint available on arXiv:2212.00958v1. MR4675822

[2] Gil Cohen, Dor Minzer, Shir Peleg, Aaron Potechin, and Amnon Ta-Shma. Expander random walks: the general case and limitations. In *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Paper No. 43, 18. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022. MR4473348

[3] Gil Cohen, Noam Peri, and Amnon Ta-Shma. Expander random walks: a Fourier-analytic approach. In *STOC'21—Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1643–1655. ACM, New York, [2021] ©2021. MR4398947

[4] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition of [MR1068527].

[5] David Gillman. A Chernoff bound for random walks on expander graphs. *SIAM J. Comput.*, 27(4):1203–1220, 1998. MR1621958

[6] Louis Golowich. A New Berry-Esseen Theorem for Expander Walks, 2022. MR4617364

[7] Louis Golowich and Salil Vadhan. Pseudorandomness of Expander Random Walks for Symmetric Functions and Permutation Branching Programs. In Shachar Lovett, editor, *37th Computational Complexity Conference (CCC 2022)*, volume 234 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:13, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. MR4481635

[8] Venkatesan Guruswami and Vinayak M. Kumar. Pseudobinomiality of the sticky random walk. In *12th Innovations in Theoretical Computer Science Conference*, volume 185 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 48, 19. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2021. MR4214292

[9] Henry Helson. *Harmonic analysis*, volume 7 of *Texts and Readings in Mathematics*. Hindustan Book Agency, New Delhi, second edition, 2010. MR2656971

[10] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561, 2006. MR2247919

[11] A. N. Kolmogorov. A local limit theorem for classical Markov chains. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 13:281–300, 1949. MR0031216

[12] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.

[13] Brad W. Mann. *Berry-Esseen central limit theorems for Markov chains*. ProQuest LLC, Ann Arbor, MI, 1996. Thesis (Ph.D.) – Harvard University. https://www.proquest.com/docview/304301471. MR2694356

[14] S. V. Nagaev. Some limit theorems for stationary Markov chains. *Teor. Veroyatnost. i Primenen.*, 2:389–416, 1957. MR0094846

[15] S. V. Nagaev. More exact statement of limit theorems for homogeneous Markov chains. *Theory Probab. Appl.*, 6:62–81, 1962. MR0131291

[16] Mathew D. Penrose and Yuval Peres. Local central limit theorems in stochastic geometry. *Electron. J. Probab.*, 16:no. 91, 2509–2544, 2011. MR2869414

[17] Salil Vadhan. Expander Graphs, 2009. Lecture notes, Chapter 4. https://people.seas.harvard.edu/~salil/cs225/spring09/lecnotes/Chap4.pdf.

[18] Anatolii Zolotukhin, Sergei Nagaev, and Vladimir Chebotarev. On a bound of the absolute constant in the Berry-Esseen inequality for i.i.d. Bernoulli random variables. *Mod. Stoch. Theory Appl.*, 5(3):385–410, 2018. MR3868547

Acknowledgments. We are grateful to Professor Fedor Nazarov for helpful suggestions that led to a sharper form of Theorem 4.1.