Random walks on finite nilpotent groups driven by long-jump measures[*]

Laurent Saloff-Coste[†]   Yuwen Wang[‡]

Abstract

We consider a variant of simple random walk on a finite group. At each step, we choose an element, s, from a set of generators ("directions") uniformly, and an integer, j, from a power law distribution ("speed") associated with the chosen direction, and move from the current position, g, to gs^j. We show that if the finite group is nilpotent, the time it takes this walk to reach its uniform equilibrium is of the same order of magnitude as the diameter of a suitable pseudo-metric on the group, which is attached to the generators and speeds. Additionally, we give sharp bounds on the ℓ^2-distance between the distribution of the position of the walker and the stationary distribution, and compute the relevant diameter for some examples.

Keywords: random walk; mixing time; group.
MSC2020 subject classifications: 60B15.
Submitted to EJP on July 22, 2021, final version accepted on January 25, 2022.

1 Introduction

A probability measure μ on a finite group G with identity e induces a Markov kernel

$$K(x,y) = K(e, x^{-1}y) = \mu(x^{-1}y), \ x, y \in G.$$

The associated iterated kernel K^n is then given by

$$K^n(x,y) = \mu^{(n)}(x^{-1}y),$$

where $\mu^{(n)}$ is the iterated convolution of μ with itself n times. In this paper, we consider only probability measures μ that are symmetric, i.e. $\mu(g) = \mu(g^{-1})$, irreducible and aperiodic. Consequently, $K^n(x,y)$ converges to $\pi(y)$ as n goes to infinity, where π is the

uniform distribution on G. A *random walk on G driven by μ* is a sequence of G-valued random variables $\{X_n\}_{n \geq 0}$ of the form

$$X_n = \xi_0 \xi_1 \cdots \xi_n,$$

where ξ_0 is the initial, possibly random, position and $(\xi_i)_{i \geq 1}$ is an i.i.d. sequence of random variables with common distribution μ. When ξ_0 is e, the distribution X_n is $K^n(e, \cdot)$. The mixing time of such a random walk is

$$t_{mix} = \min\{n : ||K^n(e, \cdot) - \pi||_{TV} \leq 1/4\}.$$

In a simple random walk with respect to a set $S \subseteq G$, at each time step, the walker chooses s uniformly from S, and steps from her current location, g, to gs. In this paper, we consider a variant of this walk, where the walker may "jump" further away in the direction of s, not just take a single step. Specifically, in a finite group G, let $S = (s_1, \cdots, s_k)$ be a k-tuple of elements that generate G. For each i, let μ_i be a symmetric distribution supported on the cyclic subgroup $\langle s_i \rangle$ and set

$$\mu(g) = \frac{1}{k} \sum_{i=1}^{k} \mu_i(g). \tag{1.1}$$

One natural choice is to set μ_i to be the uniform measure on $\langle s_i \rangle$, for each i. In this case, the probability given to s_i by μ_i varies drastically depending on the order of s_i. In this paper, we consider a model that is more regular, which we call *long-jump random walk*, and is inspired by classical stable processes, see [5, Ch. 6], and approximation algorithms of convex bodies [7]. Before defining the μ_i's that we will use for the rest of the paper, we describe a more intuitive wrap-around model that is comparable. We associate with each i in $\{1, \cdots, k\}$ a number $\alpha_i \in (0, 2)$ and the probability distribution $q_i(x) = c_{\alpha_i}(1 + |x|)^{-(1+\alpha_i)}$, $x \in \mathbb{Z}$. After choosing i uniformly in $\{1, \cdots, k\}$, j is chosen from the probability distribution q_i, and, in this time step, the walker jumps from the current location g to gs_i^j. Thus, for a fixed i, the smaller the α_i the larger the probability that a high power of s_i is chosen. We will actually work with the following variant.

Definition 1.1. *Let G be a finite nilpotent group of nilpotency class ℓ, S be a k-tuple of elements that generate G and $\mathbf{a} = (\alpha_1, \ldots, \alpha_k) \in (0, 2)^k$. A long-jump measure on G is*

$$\mu_{S,\mathbf{a}}(g) = \frac{1}{k} \sum_{i=1}^{k} \sum_{j \in \mathbb{Z}/N_i\mathbb{Z}} \mathbb{1}_{s_i^j}(g) p_i(j), \tag{1.2}$$

where N_i is order of s_i in G and $p_i : \mathbb{Z}/N_i\mathbb{Z} \to \mathbb{R}$ is

$$p_i(j) = \frac{c_i}{(1 + |\,j\,|)^{1+\alpha_i}}, \qquad \text{where } 0 \leq j < N_i, \tag{1.3}$$

$$|\,j\,| = \min(j, N_i - j), c_i^{-1} = \sum_{j \in \mathbb{Z}/N_i\mathbb{Z}} \frac{1}{(1 + |\,j\,|)^{1+\alpha_i}}.$$

In addition we will use

$$\mu_{S,\mathbf{a}}(g) = \frac{1}{k} \sum_{i=1}^{k} \mu_i(g), \qquad \text{where } \mu_i(g) = \sum_{j \in \mathbb{Z}/N_i\mathbb{Z}} \mathbb{1}_{s_i^j}(g) p_i(j). \tag{1.4}$$

An $(S, \mathbf{a})$-long-jump random walk on G is a random walk driven by a long-jump measure $\mu_{S,\mathbf{a}}$.

The relationship between this definition and wrap-around model is discussed in Appendix A.

Fix a $C_0 > 0$. When G is a finite nilpotent group of class $\ell < C_0$ and generated by a symmetric set S of size less than C_0, it is shown in [3] that the mixing time of the simple random walk associated with the set S is of the same order as the square of the diameter of G with respect to S. In this paper, we answer the following question:

Given the modification of long-jumps, what is the mixing time?

This a challenging question even in the case when $G = \mathbb{Z}/N\mathbb{Z}$ as illustrated by the following example.

Example 1.2. For any positive integer t, consider the cyclic group $\mathbb{Z}/N\mathbb{Z}$ with $N = t^5$. Let S be the generating 2-tuple $(1, s)$, where $s = t^4$. The simple random walk driven by the uniform measure on $\{\pm 1, \pm s\}$ mixes in order D^2 where D is the diameter of the Cayley graph $(\mathbb{Z}/N\mathbb{Z}, \{\pm 1, \pm s\})$. The diameter D is of order $N^{4/5}$ and the mixing time of the simple random walk is of order $N^{8/5}$.

Now, consider the random walk driven by $\mu_{S, \mathbf{a}}$ where $\mathbf{a} = (\alpha, 1)$ and $\alpha \in (0, 2)$. Intuitively, as α decreases, since the variance of the length of the steps increases, one expects that the mixing time decreases. Indeed, following from the results of this paper, the mixing time is of order

$$
\begin{cases}
N^\alpha & \text{for } 0 < \alpha \leq 1/5, \\
N^{1/5} & \text{for } 1/5 \leq \alpha \leq 1/4, \\
N^{4\alpha/5} & \text{for } 1/4 \leq \alpha < 2.
\end{cases}
\tag{1.5}
$$

Let us summarize what is behind this mixing time estimate. For a general finite nilpotent group, G, and a probability measure $\mu_{S, \mathbf{a}}$, we introduce a quasinorm $||\cdot||_{S, \mathbf{a}}$ and show that the associated random walk mixes in time of order $D_{S, \mathbf{a}}$, the diameter of G with respect to $||\cdot||_{S, \mathbf{a}}$. Even when $G = \mathbb{Z}/N\mathbb{Z}$, as in Example 1.2, computing the diameter $D_{S, \mathbf{a}}$ is a non-trivial task, see Appendix C. The estimate (1.5) is obtained using this method.

We now explain how to associate with $(S, \mathbf{a})$ a quasi-norm on G. Referring to the notation in Definition 1.1, from $S = (s_1, s_2, \ldots, s_k)$, create a formal alphabet $\mathcal{S} = \{\mathbf{s}_1^{\pm 1}, \mathbf{s}_2^{\pm 1}, \ldots, \mathbf{s}_k^{\pm 1}\}$. Let $\mathcal{W}$ be the set of finite words generated by $\mathcal{S}$ and $\deg_{\mathbf{s}_i}(w)$ be the number of times either $\mathbf{s}_i^{+1}$ or $\mathbf{s}_i^{-1}$ appears in the word w. There is natural projection map ρ from $\mathcal{W} \to G$, mapping $\mathbf{s}_i^{\pm 1}$ to $s_i^{\pm 1}$, $i = 1, \ldots, k$. For example, $w_1 = \mathbf{s}_1^{+1}$ and $w_2 = \mathbf{s}_1^{+1}\mathbf{s}_1^{-1}\mathbf{s}_1^{-1}$ both map to s_1 under ρ. However, their degrees with respect to $\mathbf{s}_1$ differs:

$$
\deg_{\mathbf{s}_1}(w_1) = 1 \qquad \deg_{\mathbf{s}_1}(w_2) = 3.
$$

For $\mathbf{a} \in (0, 2)^k$, define the *cost* of g to be

$$
||g||_{S, \mathbf{a}} = \min_{w \in \mathcal{W}: g = \rho(w)} \left\{ \max_i \left\{ (\deg_{\mathbf{s}_i}(w))^{\alpha_i} \right\} \right\}.
\tag{1.6}
$$

The function $||\cdot||_{S, \mathbf{a}} : G \to \mathbb{R}_+$ is a *quasi-norm* on G: it only satisfies the triangle inequality up to a multiplicative constant of 2 because we assume $\alpha_i \in (0, 2)$. We define $D_{S, \mathbf{a}}$ to be the diameter of the quasi-norm $||\cdot||_{S, \mathbf{a}}$, that is, the largest $||g||_{S, \mathbf{a}}$ can be when g varies over G.

We are ready to state the main results of the article, which relate the spectral gap and mixing time of $(S, \mathbf{a})$-long jump random walks to $D_{S, \mathbf{a}}$. Because the long-jump random walk is symmetric, its spectrum has the form

$$
-1 \leq \beta_{min} \leq \cdots \leq \beta_1 < \beta_0 = 1.
$$

The eigenvalue $\beta_{\min}$ is bounded away from -1 by a constant, see (2.2) below. Concerning β_1 and mixing time, we prove the following theorems.

Theorem 1.3. *Fix $C_0 > 0$ and $0 < \varepsilon < 1$. There exist $c_1, c_2 > 0$ (depending on C_0 and ε) such that for any $\ell, k < C_0$, $\mathbf{a} \in (\varepsilon, 2 - \varepsilon)^k$, and any finite nilpotent group G of nilpotency class ℓ generated by a k-tuple S, the $(S, \mathbf{a})$-long jump random walk satisfies*

$$c_1/D_{S, \mathbf{a}} \leq 1 - \beta_1 \leq c_2/D_{S, \mathbf{a}}.$$

Theorem 1.4. *Fix $C_0 > 0$ and $0 < \varepsilon < 1$. There exist $c_1, c_2 > 0$ (depending on C_0 and ε) such that for any $\ell, k < C_0$, $\mathbf{a} \in (\varepsilon, 2 - \varepsilon)^k$, and any finite nilpotent group G of nilpotency class ℓ generated by a k-tuple S, the $(S, \mathbf{a})$-long jump random walk satisfies*

$$e^{-c_1 n/D_{S, \mathbf{a}}} \leq ||K^n(e, \cdot) - \pi||_{T.V.} \leq e^{-c_2 n/D_{S, \mathbf{a}}}$$

for all $n > 0$. In particular, t_{mix} is of the same order as $D_{S, \mathbf{a}}$.

Although the result stated here is with respect to the total variation norm, throughout the paper we will work with ℓ^2, which will give more quantitative information; see Theorem 4.1.

The main results above rely on volume growth properties of $||\cdot||_{S, \mathbf{a}}$, which we describe now. For a given quasi-norm $|| \cdot ||$, define

$$B(x, r) = \{y \in G : ||x^{-1}y|| \leq r\} \quad \text{and} \quad V(x, r) = \sum_{y \in B(x,r)} \pi(y),$$

where π is the uniform measure on G.

Definition 1.5. *A finite group G equipped with a quasi-norm $|| \cdot ||$ is* doubling *if there exists $A > 1$ so that*

$$V(2r) \leq AV(r), \qquad \text{for all } r \geq 0.$$

We will call A a doubling constant *for the pair G and $|| \cdot ||$.*

Theorem 1.6. *Fix $C_0 > 0$ and $0 < \varepsilon < 1$. There exists $A > 0$ (depending on C_0 and ε) such that for any $\ell, k < C_0$, $\mathbf{a} \in (\varepsilon, 2)^k$, and any finite nilpotent group G of nilpotency class ℓ generated by a k-tuple S, the group G equipped with $|| \cdot ||_{S, \mathbf{a}}$ is doubling with constant at most A.*

Organization This paper is an extension of work done in [12, 1]; these papers are concerned with infinite groups, whereas this paper studies finite groups. Many of the techniques used in here take inspiration from proofs from those two papers; we will give specific citations as we use them. In Section 2, we start by proving Theorem 1.3. We prove the upper bound by using a pseudo-Poincaré inequality, where we rely heavily on results developed in [12]. For the lower bound, we use the Courant-Fischer characterization of β_1 with a test function and bounds that are similar to those in [1]. In Section 3, we prove the doubling property, i.e., Theorem 1.6 using growth results from [12] for free nilpotent groups and a lemma of [6] to transport the result to finite nilpotent groups. In Section 4, we give precise mixing ℓ^2-estimates, which gives a proof of Theorem 1.4 as a corollary. For these results, we use now standard techniques of Nash inequalities developed in [3, 4]. The ℓ^2-mixing upper bound for time less than $D_{S, \mathbf{a}}$ uses intermediate Nash inequalities. The matching lower bound uses the spectral lower bound on balls from Section 2 and relate it to $\mu_{S, \mathbf{a}}^{(n)}$ by using an argument inspired by proofs from [2, 14].

In Section 5, we discuss some of the challenges in computing $D_{S, \mathbf{a}}$ in general by providing some illustrative examples. Up until this point in the paper, we have assumed

that $\mathbf{a} \in (0,2)^k$. In Section 6, we explain how to generalize the main results when $\mathbf{a} \in (0,\infty)^k$ by changing the definition of $||\cdot||_{S,\mathbf{a}}$

In Appendix A, we discuss the relationship between the intuitive model presented in paragraph 2 of the introduction and the model that we presented in Definition 1.1 with which we work throughout the paper. In Appendix B, we outline properties of the Dirichlet form and present explicit computations for the more laborious bounds used in Section 2. In Appendix C, we present an algorithm for computing $D_{S,\mathbf{a}}$ when $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1,s)$ and $\mathbf{a} = (\alpha_1, \alpha_2)$, and a proof of its correctness. We use this algorithm for many of the examples in Section 5.

Notation We conclude the introduction with some notation that we will use in the paper.

When G is the cyclic group $\mathbb{Z}/N\mathbb{Z}$, and $g \in \mathbb{Z}/N\mathbb{Z}$ is represented as a number in $[0, N-1]$, it will be convenient to define $\mathbf{|}\, g\, \mathbf{|} = \min(|g|, |N-g|)$.

All Markov kernels $K : G \times G \to \mathbb{R}$ considered in this paper are symmetric and irreducible, and their stationary distributions π are uniform on G. Note that for random walks on groups driven by μ, $K(x,y) = K(e, x^{-1}y) = \mu(x^{-1}y)$. Define $Kf(x) = \sum_{y \in G} K(x,y)f(y)$. The corresponding continuous-time Markov chain has kernel $H_t = e^{-t(I-K)} = e^{-t}\sum_{n=0}^{\infty} \frac{t^n}{n!}K^n$. Let $k_e^n(x) = K^n(e,x)/\pi(x)$ and $h_t^e(x) = H_t(e,x)/\pi(x)$ be the densities with respect to π of the discrete- and continuous-time kernels.

The space $\ell^p(\pi)$ is the set of functions from G to $\mathbb{R}$ under the norm

$$||f||_p = \left(\sum_{x \in G} |f(x)|^p \pi(x) \right)^{1/p}$$

if $p \geq 1$ and $||f||_\infty = \sup_{x \in G} |f(x)|$. Given $p, q \in [1,\infty]$ and $K : \ell^p(\pi) \to \ell^q(\pi)$, define

$$||K||_{p \to q} = \sup_{f \in \ell^p(\pi)} \left\{ \frac{||Kf||_q}{||f||_p} \right\}.$$

The inner product on $\ell^2(\pi)$ we will use is $\langle f, g \rangle_\pi = \sum_x f(x)g(x)\pi(x)$. The Dirichlet form associated with μ on $\ell^2(\pi)$ is

$$\mathcal{E}_\mu(f,g) = \langle (I-K)f, g \rangle = \frac{1}{2}\sum_{x,y}(f(x)-f(xy))(g(x)-g(xy))\mu(y)\pi(x).$$

The relation $f \asymp g$, where f and g are positive functions, means that there exist constants $c_1, c_2 > 0$ so that $c_1 f \leq g \leq c_2 f$.

2 Spectral gap estimates

The main tool we use to study the spectral gap is the Dirichlet form. It is related to the spectral gap by

$$1 - \beta_1 = \min_{\substack{\langle f, \mathbf{1} \rangle = 0 \\ f \neq 0}} \left\{ \frac{\mathcal{E}_\mu(f,f)}{||f||_2^2} \right\} = \min_{f \neq 0} \left\{ \frac{\mathcal{E}_\mu(f,f)}{\mathrm{Var}_\pi(f)} \right\}, \tag{2.1}$$

which is explained in [10, Section 2]. Moreover, the form is linear in μ, so bounds for $\mathcal{E}_{p_i}$ can be aggregated to a bound for $\mathcal{E}_{\mu_{S,\mathbf{a}}}$. The details of these computations are included in Appendix B.

Define $\alpha_* = \min_{\alpha \in \mathbf{a}} \frac{\alpha}{2(1+\alpha)}$. It follows from Lemma A.1 that

$$\alpha_* \leq \min_{1 \leq i \leq k} c_i \leq \min_{1 \leq i \leq k} p_i(e) \leq \mu_{S,\mathbf{a}}(e).$$

By a standard bound, see e.g. Theorem 6.6 of [11],

$$\beta_{min} \geq 2\mu_{S,\mathbf{a}}(e) - 1 \geq 2\alpha_* - 1. \tag{2.2}$$

2.1 Spectral gap lower bound

A probability distribution on G, μ, satisfies *the pseudo-Poincaré inequality* if, for any $r > 0$, there exists $a(r) > 0$ such that, for all $f \in \ell^2(\pi)$,

$$||f - f_r||_2^2 \leq a(r)\,\mathcal{E}_\mu(f,f),$$

where

$$f_r(x) = \frac{1}{V(x,r)} \sum_{y \in B(x,r)} f(y)\pi(y) V(x,r) = \sum_{y \in B(x,r)} \pi(y). \tag{2.3}$$

We need the following result regarding nilpotent groups.

Theorem 2.1. *[12, Theorem 2.10] Let k and ℓ be positive integers and $\mathbf{a} \in (0,2)^k$. There exist $C = C(\ell, k, \mathbf{a})$, $p = p(\ell, k, \mathbf{a})$, and $(i_1, \ldots, i_p) \in \{1, \ldots, k\}^p$ so that for any finite nilpotent group G of class ℓ and generating k-tuple S, any $g \in G$ with $||g||_{S,\mathbf{a}} \leq r$ can be written as*

$$g = \prod_{j=1}^{p} s_{i_j}^{m_j} \quad \text{with } |m_j| \leq C r^{1/\alpha_{i_j}}.$$

Proof. Let $\hat{G} = N(\ell, k)$ be the free nilpotent group of nilpotency class ℓ and generated by S. Theorem 2.10 of [12] states that there exist an integer $p = p(\ell, k, \mathbf{a})$, a constant $C = C(\ell, k, \mathbf{a})$, and $(i_1, \ldots, i_p) \in [k]^p$, such that for all $\hat{g} \in \hat{G}$ that can be expressed a word $\hat{w}$ where $\deg_{\mathbf{s}_i} \hat{w} \leq r^{1/\alpha_i}$, $\hat{g}$ can be rewritten as

$$\hat{g} = \prod_{j=1}^{p} s_{i_j}^{m_j} \quad \text{with } |m_j| \leq C r^{\alpha_{i_j}}. \tag{2.4}$$

Define ρ and $\hat{\rho}$ be the projections maps from $\mathcal{W}$ to G and $\hat{G}$ respectively, mapping $\mathbf{s}_i \to s_i$. There exists a group homomorphism φ so that the following diagram commutes, i.e., such that $\varphi(s_i) = s_i$ for all i.

$$
\begin{array}{ccc}
 & & \hat{G} = N(\ell, k) \\
 & \stackrel{\hat{\rho}}{\nearrow} & \downarrow \varphi \\
S \stackrel{i}{\hookrightarrow} \mathcal{W} & \stackrel{\rho}{\longrightarrow} & G
\end{array}
$$

Let $g \in G$ be an element satisfying the conditions in the theorem. Let w_0 be a word that realizes $||g||_{S,\mathbf{a}}$ in the sense of (1.6), and $\hat{g} = \hat{\rho}(w_0)$. Since for all i, $\deg_{\mathbf{s}_i} w_0 \leq r^{1/\alpha_i}$, there exist $p(\ell, k, \mathbf{a})$, $C(\ell, k, \mathbf{a})$ and $(i_1, \ldots, i_p)$ so that (2.4) is satisfied. After applying φ to both sides, we get the desired result. $\square$

Theorem 2.2. *[12, Theorem 4.3] Let k, ℓ, and $\mathbf{a} \in (0,2)^k$ be fixed. There exists a constant $a = a(k, \ell, \mathbf{a})$ such that for any long-jump random measure $\mu_{S,\mathbf{a}}$ on a finite nilpotent group G of class ℓ with $|S| = k$ and $f : G \to \mathbb{R}$,*

$$||f - f_r||_2^2 \leq ar\,\mathcal{E}_{\mu_{S,\mathbf{a}}}(f,f).$$

Proof. Fix $r > 0$, $y_0 \in B(e, r)$ and $w_0 \in \mathcal{W}$ so that w_0 realizes $||y_0||_{S, \mathbf{a}}$. By Theorem 2.1, there exists $C_0 = C_0(\ell, k, \mathbf{a})$, $p = p(\ell, k, \mathbf{a})$ and $(i_1, \ldots, i_p) \in [k]^p$ so that y_0 can be written as $y_0 = s_{i_1}^{m_1} \cdots s_{i_p}^{m_p}$ where $|m_j| \leq C_0 r^{1/\alpha_{i_j}}$. For each j, Theorem B.5 gives the existence of a constant $C_1(\alpha_{i_j}) > 0$ so that, for all $f : G \to \mathbb{R}$ and $m \in \mathbb{Z}/N_{i_j}\mathbb{Z}$ where $|m|^{\alpha_{i_j}} \leq C_0^{\alpha_{i_j}} r$,

$$\frac{1}{|G|} \sum_{x \in G} |f(x) - f(xs^m)|^2 \leq C_1(\alpha_{i_j}) C_0^{\alpha_{i_j}} r \, \mathcal{E}_{\mu_{i_j}}(f, f). \tag{2.5}$$

By Theorem B.1 (2), for all $f \in \ell^2(\pi)$,

$$\begin{aligned}
\frac{1}{|G|} \sum_{x \in G} |f(x) - f(xy_0)|^2 &= \frac{1}{|G|} \sum_{x \in G} |f(x) - f(xs_{i_1}^{m_1} \cdots s_{i_p}^{m_p})|^2 \\
&\leq \frac{p}{|G|} \sum_{j=1}^{p} \sum_{x \in G} |f(x) - f(xs_{i_j}^{m_j})|^2 \quad \text{(telescoping sum and C-S)} \\
&\leq C_1(\alpha_{i_j}) p \sum_{j=1}^{p} C_0^{\alpha_{i_j}} r \, \mathcal{E}_{\mu_{i_j}}(f, f) \quad \text{(by (2.5))} \\
&\leq p^2 k \max_{1 \leq i \leq k} \{C_1(\alpha_i) C_0^{\alpha_i}\} r \, \mathcal{E}_{\mu_{S, \mathbf{a}}}(f, f) \quad \text{(since } \mathcal{E}_{\mu_i} \leq k \, \mathcal{E}_{\mu_{S, \mathbf{a}}}).
\end{aligned}$$

Let $a = kp^2 \max_i \{C_1(\alpha_i) C_0^{\alpha_i}\}$. The theorem now follows immediately from Proposition B.1 (1). $\qquad \square$

Proof of lower bound of Theorem 1.3. Let $r = D_{S, \mathbf{a}}$. In this case,

$$f_r = \mathbb{E}_\pi[f] \quad ||f - f_r||_2^2 = \mathrm{Var}_\pi(f).$$

Using the setting of Theorem 2.2, we have that for all $f \in \ell^2(\pi)$, $\mathrm{Var}_\pi(f) \leq aD_{S, \mathbf{a}} \mathcal{E}_{\mu_{S, \mathbf{a}}}(f, f)$. From the spectral gap characterization (2.1), we obtain $1 - \beta_1 \geq 1/aD_{S, \mathbf{a}}$. $\qquad \square$

2.2 Spectral gap upper bound

In this section, we will prove the upper bound of Theorem 1.3 as a consequence of the following result, which is similar to [1, Lemma 4.2].

Theorem 2.3. *Let $\mu_{S, \mathbf{a}}$ be a long-jump measure on a finite group G that is doubling with constant A with respect to $|| \cdot ||_{S, \mathbf{a}}$. There exists $\zeta \in \ell^2(\pi)$ and $a(A, \mathbf{a}) > 0$ such that*

$$\frac{\mathcal{E}_{\mu_{S, \mathbf{a}}}(\zeta, \zeta)}{||\zeta||_2^2} \leq \frac{a(A, \mathbf{a})}{D_{S, \mathbf{a}}}.$$

Proof of the upper bound of Theorem 1.3. Let G be a finite nilpotent group with nilpotency class ℓ, S be a list of k generating elements, and $\mathbf{a} \in (0, 2)^k$. This gives a long-jump measure $\mu_{S, \mathbf{a}}$ and cost function $|| \cdot ||_{S, \mathbf{a}}$. By Theorem 1.6, G is doubling with respect to $|| \cdot ||_{S, \mathbf{a}}$, with doubling constant $A(\ell, k, \mathbf{a})$. From Theorem 2.3, there exists a constant $a(\ell, k, \mathbf{a}) > 0$ and function ζ so that

$$\frac{\mathcal{E}_{\mu_{S, \mathbf{a}}}(\zeta, \zeta)}{||\zeta||_2^2} \leq \frac{a(\ell, k, \mathbf{a})}{D_{S, \mathbf{a}}}.$$

From the spectral gap characterization (2.1), $1 - \beta_1 \leq a(\ell, k, \mathbf{a})/D_{S, \mathbf{a}}$. $\qquad \square$

We are left with the task of proving Theorem 2.3.

Lemma 2.4. *Let $o \in G$ where $||o||_{S,\mathbf{a}} = D_{S,\mathbf{a}}$. In addition, define*

$$\Omega_+ = \left\{ x \in G : ||x||_{S,\mathbf{a}} \le \frac{1}{12}D_{S,\mathbf{a}} \right\} \Omega_- = \left\{ x \in G : ||o^{-1}x||_{S,\mathbf{a}} \le \frac{1}{12}D_{S,\mathbf{a}} \right\}.$$

If $g \in \Omega_+$ and $gh \in \Omega_-$, then $||h||_{S,\mathbf{a}} \ge \frac{1}{8}D_{S,\mathbf{a}}$. Thus Ω_+ and Ω_- are disjoint.

Proof. We know that

$$||o||_{S,\mathbf{a}} \le 2(||o^{-1}g||_{S,\mathbf{a}} + ||g||_{S,\mathbf{a}})||o^{-1}g||_{S,\mathbf{a}} \le 2(||o^{-1}gh||_{S,\mathbf{a}} + ||h||_{S,\mathbf{a}}).$$

It follows that

$$D_{S,\mathbf{a}} = ||o||_{S,\mathbf{a}} \le 2\left(2\left(\frac{1}{12}D_{S,\mathbf{a}} + ||h||_{S,\mathbf{a}}\right) + \frac{1}{12}D_{S,\mathbf{a}} \right) = \frac{1}{2}D_{S,\mathbf{a}} + 4||h||_{S,\mathbf{a}}.$$

Thus, $||h||_{S,\mathbf{a}} \ge D_{S,\mathbf{a}}/8$. $\qquad\square$

We now define the test function for Theorem 2.3. For $R = D_{S,\mathbf{a}}/16$, let $\zeta : G \to \mathbb{R}$ be

$$\zeta(g) = \zeta_+(g) - \zeta_-(g),$$

where $\alpha_\star = \min(\mathbf{a})$,

$$\zeta_+(g) = (R^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star})_+ \zeta_-(g) = (R^{1/\alpha_\star} - ||o^{-1}g||_{S,\mathbf{a}}^{1/\alpha_\star})_+. \tag{2.6}$$

Because $R = D_{S,\mathbf{a}}/16$, by Lemma 2.4, the supports of ζ_+ and ζ_- are disjoint and

$$||\zeta||_2^2 = ||\zeta_+||^2 + ||\zeta_-||^2 = 2||\zeta_+||^2 = 2||\zeta_+||^2.$$

Let A be the doubling constant of G with respect to $||\cdot||_{S,\mathbf{a}}$ and $B(e,R)$ is a ball with respect to the quasi-metric $||\cdot||_{S,\mathbf{a}}$. Because $R^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star} \ge (1 - 2^{-1/\alpha_\star})R^{2/\alpha_\star}$ when $g \in B(e, R/2)$, it follows that

$$||\zeta_+||_2^2 = \frac{1}{|G|}\sum_{g \in G}(R^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star})_+^2 \ge \frac{1}{|G|}\sum_{g \in B(e,R/2)}\left(1 - \frac{1}{2^{1/\alpha_\star}}\right)^2 R^{2/\alpha_\star}$$

$$= \frac{1}{|G|}\left(1 - \frac{1}{2^{1/\alpha_\star}}\right)^2 R^{2/\alpha_\star}\#B(e,R/2) \ge \frac{1}{|G|}\left(1 - \frac{1}{2^{1/\alpha_\star}}\right)^2 \frac{1}{A}R^{2/\alpha_\star}\#B(e,R).$$

Thus,

$$||\zeta||_2^2 \ge C_0 R^{2/\alpha_\star}\frac{\#B(e,R)}{|G|}, \qquad \text{where } C_0 = \frac{2}{A}(1 - 2^{-1/\alpha_\star})^2. \tag{2.7}$$

Now Theorem 2.3 follows from (2.7) and the following lemma.

Lemma 2.5. *Let ζ be defined as above. Then there exists $C(k, \ell, \mathbf{a}) > 0$ such that*

$$\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta, \zeta) \le \frac{C(k, \ell, \mathbf{a})\#B(e,R)}{|G|}R^{-1+2/\alpha_\star}.$$

Because $\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta, \zeta) = 2\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_+) - 2\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_-)$, Lemma 2.5 reduces to the following statement.

Lemma 2.6. *Let ζ be defined as above. Then there exists $C(k, \ell, \mathbf{a}) > 0$ such that*

$$\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_+) \le \frac{C(k, \ell, \mathbf{a})\#B(e,R)}{|G|}R^{-1+2/\alpha_\star} \tag{2.8}$$

and

$$-\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_-) \le \frac{C(k, \ell, \mathbf{a})\#B(e,R)}{|G|}R^{-1+2/\alpha_\star}. \tag{2.9}$$

Proof of (2.8). For this bound, we will use the notation from Definition 1.1 to describe $\mu_{S,\mathbf{a}}$. Fix $i_0 \in [1,k]$, and let $s_0 = S(i_0)$, $\alpha_0 = \mathbf{a}(i_0)$, $\mu_0 = \mu_{i_0}$ and $p_0 = p_{i_0}$. We will first prove the inequality from the theorem for each i_0 and then take the average of both sides for the final result.

Keeping this notation in mind, we begin by giving an upper bound for

$$\mathcal{E}_{\mu_0}(\zeta_+, \zeta_+) = \frac{1}{2|G|} \sum_{g,h \in G} |\zeta_+(gh) - \zeta_+(g)|^2 \mu_0(h).$$

Let $\Omega = \{(g,h) \in G \times \langle s_0 \rangle : \zeta_+(gh) + \zeta_+(g) > 0\}$. So we can restrict the sum above to just Ω. For a fixed h, we have that

$$\#\{g \in G : (g,h) \in \Omega\} \le 2\#B(e, R).$$

Note that $\mu_0(h)$ is only non-zero when $h \in \langle s_0 \rangle$, so we can write $h = s_0^t$. Thus, we can further break the sum into two parts: (1) when $|t| > \rho$ and (2) when $|t| \le \rho$, where $\rho = (12R)^{1/\alpha_0}$. For the first sum, by Lemma 2.4, we have

$$\sum_{(g,h):\Omega:|t|\ge\rho} |\zeta_+(gh) - \zeta_+(g)|^2 \mu_0(h) \le 2(R^{1/\alpha_\star})^2 \#B(e, R) \sum_{|t|\ge\rho} p_0(t)$$

$$\le \frac{4R^{2/\alpha_\star} \#B(e, R)}{\alpha_0 \rho^{\alpha_0}}.$$

For sum (2), fix $g \in G$ and $h \in \langle s_0 \rangle$, and choose the smallest t in absolute value so that $s_0^t = h$. It will be convenient, for all $g \in G$, to set w_g to be a word that realizes the cost of g, and set $x = \deg_{\mathbf{s_0}} w_g$, the number of times either $\mathbf{s_0}$ or $\mathbf{s_0^{-1}}$ appears in w_g. To start, we would like to bound the term $|\zeta_+(gh) - \zeta_+(g)|$. Note that we can assume that $||gh||_{S,\mathbf{a}} \ge ||g||_{S,\mathbf{a}}$; otherwise we can set $g_0 = gh$ and $g_0 h^{-1} = g$, and the bound would proceed the same since we do not make assumptions about g and $||h||_{S,\mathbf{a}} = ||h^{-1}||_{S,\mathbf{a}}$. This implies that

$$|\zeta_+(gh) - \zeta_+(g)| = (R^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star})_+ - (R^{1/\alpha_\star} - ||gh||_{S,\mathbf{a}}^{1/\alpha_\star})_+.$$

We will now show that this expression is less than or equal to $||gh||_{S,\mathbf{a}}^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star}$. Since we assume that $||gh||_{S,\mathbf{a}} \ge ||g||_{S,\mathbf{a}}$, if the first term is zero, then so is the second term. Therefore, three cases remain: (1) if both term are zero, then the inequality holds trivially, (2) if both terms are non-zero, then the two lines are equal, and (3) if the second term is zero, but the first is not, then, $||gh||_{S,\mathbf{a}} \ge R$, and

$$|\zeta_+(gh) - \zeta_+(g)| = R^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star} \le ||gh||_{S,\mathbf{a}}^{1/\alpha_\star} - ||g||_{S,\mathbf{a}}^{1/\alpha_\star}.$$

We are ready to evaluate

$$|\zeta_+(gh) - \zeta_+(g)| \le \max_{1\le i\le k}\{(\deg_{\mathbf{s}_i} w_g)^{\alpha_i/\alpha_\star}, (x+|t|)^{\alpha_0/\alpha_\star}\} - \max_{1\le i\le k}\{(\deg_{\mathbf{s}_i} w_g)^{\alpha_i/\alpha_\star}, x^{\alpha_0/\alpha_\star}\}$$

$$\le (x+|t|)^{\alpha_0/\alpha_\star} - x^{\alpha_0/\alpha_\star}.$$

By the fundamental theorem of calculus and $x \le (12R)^{1/\alpha_0}$, we have

$$|\zeta_+(gh) - \zeta_+(g)| \le \int_x^{x+|t|} \frac{\alpha_0}{\alpha_\star} s^{\frac{\alpha_0}{\alpha_\star}-1} \, ds \le \frac{\alpha_0}{\alpha_\star}\left((12R)^{1/\alpha_0} + \rho\right)^{\frac{\alpha_0}{\alpha_\star}-1} |t|$$

$$\le \frac{\alpha_0}{\alpha_\star}(2^{1/\alpha_0}12R)^{\frac{1}{\alpha_\star}-\frac{1}{\alpha_0}} |t|.$$

Summing over $h = s^t$, where $|t| \leq \rho$,

$$\frac{1}{2} \sum_{|t| \leq \rho} |\zeta_+(gs^t) - \zeta_+(g)|^2 \mu_0(t) \leq \frac{\alpha_0^2}{\alpha_\star^2} \left(2^{1/\alpha_0} 12R\right)^{2/\alpha_\star - 2/\alpha_0} \#B(e,R) \sum_{|t| \leq \rho} |t|^2 \mu_0(t)$$

$$\leq C_1 R^{2/\alpha_\star - 2/\alpha_0} \#B(e,R)\rho^{2-\alpha_0} \qquad \text{(by Lemma A.2)},$$

where $C_1 = C_1(\alpha_0) = C_0(\alpha_0)\alpha_0^2 12^{2/\alpha_\star - 2/\alpha_0} \frac{3^{2-\alpha_0}}{2-\alpha_0}$ and $C_0(\alpha_0)$ is the constant that appears in Lemma A.2. Putting the two parts together, we have

$$|G|\,\mathcal{E}_{\mu_0}(\zeta_+, \zeta_+) \leq \#B(e,R)R^{2/\alpha_\star} \left(\frac{4}{\alpha_0}\rho^{-\alpha_0} + C_1 R^{-2/\alpha_0}\rho^{2-\alpha_0}\right)$$

$$\leq \#B(e,R)R^{2/\alpha_\star} \left(\frac{4}{\alpha_0 12^{\alpha_0}}R^{-1} + C_1 12^{2-\alpha_0}R^{-2/\alpha_0}R^{2/\alpha_0 - 1}\right)$$

$$\leq C_2 \#B(e,R)R^{-1+2/\alpha_\star},$$

where $C_2 = C_2(\alpha_0) = \frac{4}{\alpha_0 12^{\alpha_0}} + C_1(\alpha_0)12^{2-\alpha_0}$. $\qquad\qquad\square$

Proof of (2.9). We use the notation s_0, α_0, μ_0, and p_0 as above, and give a lower bound for

$$\mathcal{E}_{\mu_0}(\zeta_+, \zeta_-) = \frac{1}{2|G|} \sum_{g,h \in G} (\zeta_+(gh) - \zeta_+(g))(\zeta_-(gh) - \zeta_-(g))\mu_0(h).$$

Let Ω_+ be the support of ζ_+ and Ω_- be the support of ζ_-. As we chose $R = D_{S,\mathbf{a}}/12$, Lemma 2.4 implies that Ω_+ and Ω_- are disjoint. We see that the only non-zero summands are those where $g \in \Omega_+$ and $gh \in \Omega_0$ or vice versa, in which case $||h||_{S,\mathbf{a}} > R$. By symmetry, we have

$$-\mathcal{E}_{\mu_0}(\zeta_+, \zeta_-) = \frac{1}{|G|} \sum_{\substack{g \in \Omega_+ \\ gh \in \Omega_-}} \zeta_+(g)\zeta_-(gh)\mu_0(h)$$

Because $|\zeta_+|, |\zeta_+| \leq R^{1/\alpha_\star}$ and ζ_+ has support in $B(e,R)$,

$$-|G|\,\mathcal{E}_{\mu_0}(\zeta_+, \zeta_-) \leq R^{2/\alpha_\star} \sum_{\substack{g \in \Omega_+ \\ gh \in \Omega_-}} \mu_0(h) \leq R^{2/\alpha_\star} \#B(e,R) \sum_{||h||_{S,\mathbf{a}} > R} \mu_0(h)$$

$$\leq R^{2/\alpha_\star} \#B(e,R) \sum_{|t|^{\alpha_0} > R} p_0(t) \leq C_3(\alpha_0) \#B(e,R)R^{-1+2/\alpha_\star},$$

where $C_3(\alpha_0)$ is as in Lemma A.2. Averaging over all components μ_i of $\mu_{S,\mathbf{a}}$, we get that the inequality also holds for $\mu_{S,\mathbf{a}}$. $\qquad\qquad\square$

3 Volume estimates

For proving the doubling statement of Theorem 1.6, there are two main ingredients: (1) [12, Example 1.5] that shows doubling with respect to $||\cdot||_{S,\mathbf{a}}$ for free nilpotent groups and (2) the finite version of [6, Lemma 1.1] stated below, which allows us to translate doubling from the free nilpotent group to the finite nilpotent group.

Lemma 3.1 ([6, Lemma 1.1]). *Let G be a finitely-generated countable group acting on a set X, which we will write on the right. Let A and B be finite subsets of G, and Y a subset of X. Then,*

$$\#A\#(YB) \leq \#(AB)\#(YA^{-1}).$$

Proof of Theorem 1.6. Let $\hat{G} = N(\ell, k)$ denote the free nilpotent group of class ℓ generated by S of class ℓ. Let $\mathcal{W}$ be the set words generated by entries of S and $\hat{\rho}$ and ρ be the natural projection maps from $\hat{G}$ and G, respectively, to $\mathcal{W}$. Using this lifting, we can define an $(S, \mathbf{a})$-cost function on $\hat{G}$, which we will call $||| \cdot |||_{S, \mathbf{a}}$ to differentiate. We will also use $\hat{B}$ and B to denote balls with respect to $||| \cdot |||_{S, \mathbf{a}}$ and $|| \cdot ||_{S, \mathbf{a}}$ respectively. Further, there exists $\varphi : \mathcal{W} \to \hat{G}$ so that the following diagram commutes.

$$
\begin{array}{ccc}
 & & \hat{G} = N(\ell, k) \\
 & \overset{\hat{\rho}}{\nearrow} & \downarrow \varphi \\
S \overset{i}{\hookrightarrow} \mathcal{W} & \overset{\rho}{\longrightarrow} & G
\end{array}
$$

With this, by the way that the cost function is defined, for all $\hat{g} \in N(\ell, k)$, $|||\hat{g}|||_{S, \mathbf{a}} \geq ||\varphi(\hat{g})||_{S, \mathbf{a}}$.

By Example 1.5 from [12], there exist constants $c_1, c_2 > 0$

$$c_1 r^{d(\ell,k)} \leq \#\hat{B}(e, r) \leq c_2 r^{d(\ell,k)} \quad \text{where } d(\ell, k) = \sum_{m=1}^{\ell} \sum_{d|m} \boldsymbol{\mu}(d) k^{m/d} \tag{3.1}$$

and $\boldsymbol{\mu}$ is the classical Möbius function. Thus, $N(\ell, k)$ has polynomial growth with respect to $||| \cdot |||_{S, \mathbf{a}}$.

Next, we apply Lemma 3.1 with the group action of $\hat{g} \in \hat{G}$ on $x \in G$ via $x \cdot \hat{g} = x\varphi(\hat{g})$, and the sets

$$
\begin{aligned}
Y &= \{e_G\}, \\
A &= \hat{B}(e, r) = \{x \in N(\ell, k) : |||x|||_{S, \mathbf{a}} \leq r\} \text{ and} \\
B &= \hat{B}(e, 2r) = \{x \in N(\ell, k) : |||x|||_{S, \mathbf{a}} \leq 2r\}.
\end{aligned}
$$

First, notice that since $||x||_{S, \mathbf{a}} = ||x^{-1}||_{S, \mathbf{a}}$, $YA^{-1} = YA$. We then show $YA^{-1} = YA = B(e, r)$ by showing inclusion both ways. Let $x \in B(e, r)$ and $w \in \mathcal{W}$ that realizes the cost of x, i.e., $||x||_{S, \mathbf{a}} = \max_i (\deg_{\mathbf{s}_i} w)^{\alpha_i}$. Consider $\hat{x} = \hat{\rho}(w) \in N(k, \ell)$. By construction, $\varphi(\hat{x}) = x$, $|||\hat{x}|||_{S, \mathbf{a}} \leq r$, and thus $\hat{x} \in \hat{B}(e_G, r)$. It follows that $YA \ni e_G \cdot \hat{x} = e_G\varphi(\hat{x}) = x$, and $B(e_G, r) \subseteq AY$.

Now let $y \cdot a \in YA$. Let w be a word that realizes cost of a, i.e., $|||a|||_{S, \mathbf{a}} = \max_i (\deg_{\mathbf{s}_i} w)^{\alpha_i}$. As in the previous case $\rho(w)$, which is also equal to $\varphi(a) = y \cdot a$, must have cost less than or equal to r. Therefore, $YA \subseteq B(e_G, r)$.

Next, we want to show that $AB \subseteq \hat{B}(e, 6r)$. Let $a \in A$ and $b \in B$. Let $w_0, w_1,$ and w_2 be words that realizes the costs of ab, a and b respectively. By the triangle inequality for our quasi-norm, $||ab||_{S, \mathbf{a}} \leq 2(||a||_{S, \mathbf{a}} + ||b||_{S, \mathbf{a}})$. So $A \subseteq \hat{B}(e, 6r)$.

Finally, we can show doubling

$$
\begin{aligned}
\frac{\#B(e, 2r)}{\#B(e, r)} &= \frac{\#YB}{\#YA^{-1}} \leq \frac{\#AB}{\#A} \quad \text{(by Lemma 3.1)} \\
&\leq \frac{\#\hat{B}(e, 6r)}{\#\hat{B}(e, r)} \leq \frac{c_2 (6r)^{d(\ell,k)}}{c_1 r^{d(\ell,k)}} \leq 6^{d(\ell,k)}(c_2/c_1). \qquad \square
\end{aligned}
$$

Doubling imples the following property, which we use in Section 4.

Corollary 3.2. *For all $0 \leq r \leq R \leq D_{S, \mathbf{a}}$,*

$$V(e, r) \geq A^{-2} V(e, R) \left(\frac{r+1}{R+1}\right)^d, \qquad \text{where } d = \log_2 A.$$

Proof. We have that for $\frac{R+1}{2^k} \leq r+1$,

$$V(e, R) \leq V(e, R+1) \leq 2V(e, \frac{R+1}{2}) \leq \cdots \leq A^k V(e, \frac{R+1}{2^k}) \leq A^k V(e, r+1) \leq A^{k+1} V(e, r),$$

After taking $\log_2(\frac{R+1}{r+1}) \leq k \leq \log_2(\frac{R+1}{r+1}) + 1$, we can deduce

$$\frac{V(e, R)}{V(e, r)} \leq A^{\log_2(\frac{R+1}{r+1})+2} = A^2 \left(\frac{R+1}{r+1}\right)^d. \qquad \square$$

To conclude this section, we show that $|| \cdot ||_{S, \mathbf{a}}$ also satisfies a "reverse doubling" property, i.e. a lower bound of $V(e, R)/V(e, r)$ by a quantity that is a polynomial of R/r.

Lemma 3.3. *Let G be a finite group, $|| \cdot ||_{S, \mathbf{a}}$ be the cost function of an $(S, \mathbf{a})$-long-jump random walk. Let $1 \leq R < D_{S, \mathbf{a}}$, there exists $g \in G$ such that $R/4 \leq ||g||_{S, \mathbf{a}} \leq R$.*

Proof. In the case that $1 \leq R \leq 4$, fix s in S that is not the identity. Then, $s \in B(e, R)$, and $R/4 \leq ||s||_{S, \mathbf{a}} \leq R$. Now consider R such that $4 \leq R < D_{S, \mathbf{a}}$. Since R is strictly smaller than $D_{S, \mathbf{a}}$, $G \setminus B(e, R)$ is non-empty and there exists $g \in B(e, R)$ such that $gs \in G \setminus B(e, R)$ where s is an entry in S. Therefore, $R < ||gs||_{S, \mathbf{a}} \leq 2(||g||_{S, \mathbf{a}} + ||s||_{S, \mathbf{a}}) = 2(||g||_{S, \mathbf{a}} + 1)$, and $||g||_{S, \mathbf{a}} > R/4$. $\qquad \square$

Proposition 3.4. *For all $1 \leq R \leq D_{S, \mathbf{a}}$ and $r = 2^{-7}R$,*

$$\frac{V(e, R)}{V(e, r)} \geq 2.$$

Consequently, for all $1 \leq r \leq R \leq D_{S, \mathbf{a}}$,

$$\frac{V(e, R)}{V(e, r)} \geq \frac{1}{2} (R/r)^{1/7}. \tag{3.2}$$

Proof. There are no elements with cost in $(0, 1)$. So if $1 \leq R \leq 4$, $\#B(e, r) = 1$ and $\#B(e, R) \geq 2$. Now, we assume that $4 \leq R \leq D_{S, \mathbf{a}}$. By Lemma 3.3, there exists $o \in B(e, R/4)$, such that $2^3 r = R/2^4 \leq ||o||_{S, \mathbf{a}} \leq R/2^2$. We will show that

1. $B(e, r) \cap B(o, r) = \emptyset$

2. $B(e, r) \cup B(o, r) \subseteq B(e, R)$.

This immediately implies that there are two disjoint balls of radius r in $B(e, R)$, which is our desired result. To show (1), suppose there exists $g \in B(o, r) \cap B(e, r)$. By definition, we know that $||g||_{S, \mathbf{a}} \leq r$ and $||o^{-1}g||_{S, \mathbf{a}} \leq r$. This implies that $2^3 r \leq ||o||_{S, \mathbf{a}} \leq 2(||o^{-1}g||_{S, \mathbf{a}} + ||g||_{S, \mathbf{a}}) \leq 2^2 r$, which is a contradition. For (2), the fact that $B(e, r) \subseteq B(e, R)$ is clear. If $g \in B(o, r)$, then $||g||_{S, \mathbf{a}} \leq 2(||o||_{S, \mathbf{a}} + ||o^{-1}g||_{S, \mathbf{a}}) \leq 2(R/4 + R/2^7) \leq R$.

To show (3.2), observe that for $R \geq 2^{7k}r$,

$$V(e, R) \geq 2V(e, 2^{-7}R) \geq \cdots \geq 2^k V(e, 2^{-7k}R) \geq 2^k V(e, r).$$

By choosing k so that $\frac{1}{7}\log_2(R/r) - 1 \leq k \leq \frac{1}{7}\log_2(R/r)$, we get the desired result. $\qquad \square$

4 Estimates on mixing and proof of Theorem 1.4

Theorem 4.1. *Let K be the Markov kernel of an $(S, \mathbf{a})$-long-jump random walk on a finite group G with nilpotency ℓ, and π be the uniform distribution. There exists $b_1, b_2, c_1, c_2 > 0$ such that for all $n > 0$*

$$\frac{c_1}{V(e, n)^{1/2}} \exp\left(-n/b_1 D_{S, \mathbf{a}}\right) \leq ||k_e^n - 1||_2 \leq \frac{c_2}{V(e, n)^{1/2}} \exp\left(-n/b_2 D_{S, \mathbf{a}}\right).$$

Proof of the upper bound of Theorem 4.1. We have shown that there exists positive real numbers $d = d(\ell, k, \mathbf{a}) \geq 1$ and $a = a(\ell, k, \mathbf{a}) \geq 1$, for all $0 \leq r \leq R \leq D_{S,\mathbf{a}}$ and $f \in \ell^2(\pi)$,

$$V(e,r) \geq A^{-2} V(e,R) \left(\frac{r+1}{R+1} \right)^d \quad ||f - f_r||_2^2 \leq ar\, \mathcal{E}_{\mu_{S,\mathbf{a}}}(f,f). \tag{4.1}$$

It is straighforward to check that the proof of [4, Theorem 5.2] works for quasi-norms. Using [4, Remark 5.4 (2)] with $\alpha = 1$ and $M = \frac{A^2(R+1)^d}{V(e,R)}$, (4.1) implies that

$$\forall f \in \ell^2(\pi), \qquad ||f||_2^{2+2/d} \leq C \left(\mathcal{E}_{\mu_{S,\mathbf{a}}}(f,f) + \frac{1}{aR^2}||f||_2^2 \right) ||f||_1^{2/d},$$

where $C = (1 + 1/(2d))^2 (1 + 2d)^{1/d} A^{2/d} (R+1)(V(e,R))^{-1/d} a$. This is called a *Nash inequality* [4]. By [4, Corollary 3.1] with $R = n$, we obtain that,

$$\forall n \leq D_{S,\mathbf{a}}, \qquad ||K^n||_{2\to\infty} \leq \frac{c_3}{V(e,n)^{1/2}}, \tag{4.2}$$

where $c_3 = 2\sqrt{2}(2^{d/2})(1 + \lceil 2d \rceil)^{2d}(1 + 1/(2d))^d (1 + 2d)^{1/2} a^{d/2} A$.

Now, fix $n > 0$ and write $n = n_1 + n_2$ with $n_1 = \min(\lfloor D_{S,\mathbf{a}} \rfloor, n)$. We have

$$||k_e^n - 1||_2 = ||K^n - 1||_{2\to\infty} \leq ||K^{n_1}||_{2\to\infty}||K^{n_2} - \pi||_{2\to2},$$

e.g. [10, Section 1.2.4]. Inequality (4.2) implies that $||K^{n_1}||_{2\to\infty} \leq \frac{c_3}{V(e,n)^{1/2}}$, and Theorem 1.3 and (2.2) give

$$||K^{n_2} - \pi||_{2\to2} \leq (1 - \min(2\alpha_*, 1/aD_{S,\mathbf{a}}))^{n_2} \qquad \left(\text{where } \alpha_* = \min_{\alpha \in \mathbf{a}} \frac{\alpha}{2(1+\alpha)}\right)$$

$$\leq (1 - 2\alpha_*/aD_{S,\mathbf{a}})^{n_2} \qquad (\text{since } 2\alpha_* \leq 1 \text{ and } aD_{S,\mathbf{a}} \geq 1)$$

$$\leq \exp(-2\alpha_* n_2/aD_{S,\mathbf{a}}).$$

It follows that

$$||k_e^n - 1||_2 \leq ||K^{n_1}||_{2\to\infty}||K^{n_2} - \pi||_{2\to2} \leq \frac{c_3}{V(e,n_1)^{1/2}} \exp(-2\alpha_* n_2/aD_{S,\mathbf{a}})$$

$$\leq \frac{c_2}{V(e,n_1)^{1/2}} \exp(-n/b_2 D_{S,\mathbf{a}}) \qquad (\text{where } c_2 = e^{1/a}c_3 \text{ and } b_2 = a/(2\alpha_*)). \quad \square$$

For the lower bound, we will use this following lemma which orginates from [2, Proposition 2.3] and [14, Lemma 3.1].

Lemma 4.2. *For all $n \geq 0$, there exists a constant $C \geq 0$ such that*

$$||k_e^n||_2 \geq \frac{e^{-2C}}{V(e,n)^{1/2}}.$$

Proof. Let $U = B(e,n)$ and $K_U(x,y) = K(x,y)$ when x or y are in U and $K_U(x,y) = 0$ otherwise. Let β_U is the largest eigenvalue of K_U and φ_U be its associated eigenvector. The argument is based on the fact that as a consequence of Cauchy-Schwarz inequality, the function $n \mapsto \frac{||K^n f||_2}{||K^{(n-1)} f||_2}$ is decreasing. We have

$$||k_e^n||_2 = ||K^n||_{1\to2} = \max_{f \neq 0} \left\{ \frac{||K^n f||_2}{||f||_1} \right\}$$

$$= \max_{f \neq 0} \left\{ \frac{||K^n f||_2}{||K^{n-1} f||_2} \cdots \frac{||K f||_2}{||f||_2} \frac{||f||_2}{||f||_1} \right\}$$

$$\geq \max_{f \neq 0} \left\{ \left(\frac{||K f||_2}{||f||_2} \right)^{n-1} \frac{||f||_2}{||f||_1} \right\}$$

$$\geq \max_{\substack{f \neq 0 \\ \operatorname{supp}(f) \subseteq U}} \left\{ \left(\frac{||Kf||_2}{||f||_2} \right)^{n-1} \frac{1}{V(e,n)^{1/2}} \right\} \qquad \text{(by C-S and } U = B(e,n))$$

$$\geq \left(\frac{||K\varphi_U||_2}{||\varphi_U||_2} \right)^{n-1} \frac{1}{V(e,n)^{1/2}}$$

$$\geq \left(\frac{||K_U\varphi_U||_2}{||\varphi_U||_2} \right)^{n-1} \frac{1}{V(e,n)^{1/2}} \qquad (\varphi_U \text{ is positive})$$

$$= \beta_U^{n-1} \frac{1}{V(e,n)^{1/2}}.$$

Consider the value

$$\beta_U = \max_{\substack{f \neq 0, \operatorname{supp}(f) \subseteq U \\ ||f||_2 = 1}} ||K_U f||_2 = \max_{\substack{f \neq 0, \operatorname{supp}(f) \subseteq U \\ ||f||_2 = 1}} \langle Kf, f \rangle_\pi = 1 - \min_{\substack{f \neq 0, \operatorname{supp}(f) \subseteq U \\ ||f||_2 = 1}} \mathcal{E}_{\mu_{S,\mathbf{a}}}(f,f).$$

Consider the test functions $\mathbb{1}_e$ and ζ_+, the function from (2.6) with $R = n$. From Lemma 2.6 and (2.7), there exist a constant $C = C(k, \ell, \mathbf{a})$ such that $\frac{\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_+)}{||\zeta_+||^2} \leq C/n$. Thus, we have

$$\beta_U \geq 1 - \min \left\{ \frac{\mathcal{E}_{\mu_{S,\mathbf{a}}}(\mathbb{1}_e, \mathbb{1}_e)}{||\mathbb{1}_e||^2}, \frac{\mathcal{E}_{\mu_{S,\mathbf{a}}}(\zeta_+, \zeta_+)}{||\zeta_+||^2} \right\}$$

$$\geq 1 - \min \left\{ \frac{\alpha_*}{8}, \frac{C}{n} \right\}.$$

Collecting our lower bound on $||k_e^n||_2$ and β_U, we derive

$$||k_e^n||_2 \geq \left(1 - \min \left\{ \frac{\alpha_*}{8}, \frac{C}{n} \right\} \right)^n \frac{1}{V(e,n)^{1/2}}$$

$$\geq \exp \left(-\min \left\{ \frac{\alpha_*}{4}, \frac{2C}{n} \right\} n \right) \frac{1}{V(e,n)^{1/2}}$$

$$\geq e^{-2C} \frac{1}{V(e,n)^{1/2}}. \qquad \square$$

Proof of lower bound of Theorem 4.1. We have that $||k_e^n - 1||_2 \geq 2||\mu_{S,\mathbf{a}}^{(n)} - \pi||_{TV} \geq \beta_1^n$. From Theorem 1.3, we know that there exists $a > 0$ such that $\beta_1 \geq 1 - a/D_{S,\mathbf{a}}$. We also have the bound $\beta_{min} \geq -1 + \frac{1}{8}\alpha_*$ by using test function $\mathbb{1}_e$ in (2.1). Let $c = \min(a/D_{S,\mathbf{a}}, \alpha_*/8)$, and we compute further

$$||k_e^n - 1||_2 \geq (1 - c)^n \geq e^{-2an/D_{S,\mathbf{a}}}, \qquad \text{(since } 0 \leq c \leq 1/2.) \qquad (4.3)$$

Let $C > 0$ be the constant from Lemma 4.2. In the case that $V(e,n) \leq e^{-4C}/4$, we have $n \leq D_{S,\mathbf{a}}$, and the term $\exp(-2an/D_{S,\mathbf{a}})$ is roughly constant, namely,

$$e^{-2a} \leq \exp(-2an/D_{S,\mathbf{a}}) \leq 1.$$

Hence, it follows from Lemma 4.2 that

$$||k_e^n - 1||_2 \geq ||k_e^n||_2 - 1 \geq \frac{e^{-2C}}{V(e,n)^{1/2}} - 1 \geq \frac{e^{-2C}}{2V(e,n)^{1/2}} \geq \frac{e^{-2C}}{2V(e,n)^{1/2}} \exp(-2an/D_{S,\mathbf{a}}).$$

When $V(e,n) \geq e^{-4C}/4$, by (4.3), we have

$$||k_e^n - 1||_2 \geq \exp(-2an/D_{S,\mathbf{a}}) \geq \frac{e^{-2C}}{2V(e,n)^{1/2}} \exp(-2an/D_{S,\mathbf{a}}).$$

Thus, the lower bound is true for $c_1 = \exp(-2C)/2$ and $b_1 = 1/2a$. $\qquad \square$

The proof for continuous time is similar and we have the following result. For the definition of H_t and h_t^e, see the Notation section at the end of Section 1.

Theorem 4.3. *Let H_t, $t > 0$, be the continuous time Markov kernel of an $(S, \mathbf{a})$-long-jump random walk on a finite group G with nilpotency ℓ, and π be the uniform distribution. Then, there exists $a_1, a_2, c_1, c_2 > 0$ such that for all $t > 0$*

$$\frac{c_1}{V(e,t)^{1/2}} \exp\left(-t/a_1 D_{S,\mathbf{a}}\right) \leq \|h_t^e - 1\|_2 \leq \frac{c_2}{V(e,t)^{1/2}} \exp\left(-t/a_2 D_{S,\mathbf{a}}\right).$$

We now have the ingredients to prove Theorem 1.4, the mixing time result.

Proof of Theorem 1.4. We want to show that there exists $a_1, a_2, b_1, b_2 > 0$ such that for all $n > 0$,

$$a_1 \exp(-n/b_1 D_{S,\mathbf{a}}) \leq \|K^n(e,\cdot) - \pi\|_{TV} \leq a_1 \exp(-n/b_2 D_{S,\mathbf{a}}).$$

The lower bound follows directly from the fact that $\|K^n(x,\cdot) - \pi\|_{TV} \geq \beta_1^n$, see [11, Proposition 5.5]. For the upper bound, we know from Theorem 4.1,

$$\|K^n(e,\cdot) - \pi\|_{TV} \leq \frac{1}{2}\|k_e^n - 1\|_2 \leq \frac{c_1}{V(e,n)^{1/2}} \exp\left(-n/c_2 D_{S,\mathbf{a}}\right), \qquad (4.4)$$

for some $c_1, c_2 > 0$. When $n \geq D_{S,\mathbf{a}}$, $V(e,n)$ is equal to one and only the exponential term remains. The total variation is always bounded above by 2, and when $n \leq D_{S,\mathbf{a}}$ the exponential term bounded above by a constant, which gives us our upper bound. It thus follows that

$$a_1(\log 2)D_{S,\mathbf{a}} \leq t_{mix} \leq a_2(\log 4)D_{S,\mathbf{a}}. \qquad \square$$

5 On computing the diameter

As one would expect, computing $D_{S,\mathbf{a}}$ for arbitrary groups, S, and $\mathbf{a}$ is a difficult problem in general. More surprisingly, even just on the cyclic group computing $D_{S,\mathbf{a}}$ is still quite nuanced. In Section C, we give an exact formula for $D_{S,\mathbf{a}}$ when the $S = (1, s)$ and arbitrary $\mathbf{a}$ and use those results in our examples. We start with a remark about the relationship between $D_{S,\mathbf{a}}$ and the diameter of the Cayley graph:

Remark 5.1. Let G be a finite group and $S = (s_1, s_2, \ldots, s_k)$ be a k-tuple whose elements generate G. Recall from the introduction, $\mathcal{W}$ is the set of words generated by an alphabet $\mathcal{S} = \{\mathbf{s}_1^{\pm 1}, \mathbf{s}_2^{\pm 1}, \ldots, \mathbf{s}_k^{\pm 1}\}$ generated from S. We define the following quantity, which is comparable to the diameter of the Cayley graph.

$$D_{\mathcal{S}} = \max_{g \in G} \left\{ \min_{w \in \mathcal{W}: g = \rho(w)} \max_{1 \leq i \leq k} |\deg_{\mathbf{s}_i}(w)| \right\}.$$

Notice the following facts:

1. if $\mathbf{a} = (\alpha, \ldots, \alpha)$ for some $\alpha \in (0, 2)$, then $D_{S,\mathbf{a}} = D_{\mathcal{S}}^{\alpha}$, and

2. if $c > 0$ and $\mathbf{a}$ such that $\mathbf{a}$ and $c\mathbf{a} \in (0, 2)^k$, then $D_{S,\mathbf{a}}^c = D_{S,c\mathbf{a}}$.

Now we are ready to present three examples in the vein of the example in the introduction, Example 1.2. We will compute $D_{S,\mathbf{a}}$ for $G = \mathbb{Z}/N\mathbb{Z}$ where $\mathbf{a}$ and N are fixed and S is set to $(1, s)$ for various s of the same order. See Appendix C for detailed computations.

Example 5.2 (Simple variation of Example 1.2). Let t be a postive integer larger than 5 and

$$N = t(t^2 + 1)(t^2 + 2).$$

We want to find $D_{S,\mathbf{a}}$ for $G = \mathbb{Z}/N\mathbb{Z}$, $\mathbf{a} = (\alpha, 1)$, $\alpha \in (0, 2)$, and $S = (1, s)$, where $s = (t^2 + 1)(t^2 + 2)$. By Theorem C.1,

$$D_{S,\mathbf{a}} \asymp \min\{N^\alpha, \max\{N^{4\alpha/5}, N^{1/5}\}\}.$$

Breaking this into cases, we have

$$D_{S,\mathbf{a}} \asymp \left\{ \begin{array}{ll} N^\alpha & \text{if } \alpha \in (0, 1/5), \\ N^{1/5} & \text{if } \alpha \in [1/5, 1/4), \\ N^{4\alpha/5} & \text{if } \alpha \in [1/4, 2). \end{array} \right.$$

For this example, s is of order $N^{4/5}$ where s divides N, which is the same set up in Example 1.2 from the introduction.

Example 5.3. Next, we still have $t > 5$, $N = t(t^2 + 1)(t^2 + 2)$, $G = \mathbb{Z}/N\mathbb{Z}$, and $\mathbf{a} = (\alpha, 1)$ for some $\alpha \in (0, 2)$. For this example, we pick $S' = (1, s')$, with $s' = t^2(t^2 + 2)$ and we will compute $D_{S',\mathbf{a}}$. As in the previous example, $s' \asymp N^{4/5}$, but s' doesn't quite divide N. Dividing N by s' using the Euclidean algorithm, we get

$$N = ts' + r \qquad \text{(where } r = t^3 + 2t)$$
$$s' = tr.$$

Applying Theorem C.1, we have

$$D_{S',\mathbf{a}} \asymp \min\{N^\alpha, \max\{N^{4\alpha/5}, N^{1/5}\}, \max\{N^{3\alpha/5}, N^{2/5}\}\},$$

which gives us what we got in Example 5.2 with two more cases

$$D_{S',\mathbf{a}} \asymp \left\{ \begin{array}{ll} N^\alpha & \text{if } \alpha \in (0, 1/5), \\ N^{1/5} & \text{if } \alpha \in [1/5, 1/4), \\ N^{4\alpha/5} & \text{if } \alpha \in [1/4, 1/2), \\ N^{2/5} & \text{if } \alpha \in [1/2, 2/3), \\ N^{3\alpha/5} & \text{if } \alpha \in [2/3, 2). \end{array} \right.$$

Example 5.4. Again, we let $t > 5$, $N = t(t^2 + 1)(t^2 + 2)$, $G = \mathbb{Z}/N\mathbb{Z}$, $\mathbf{a} = (\alpha, 1)$, $\alpha \in (0, 2)$. We choose $S'' = (1, s'')$, with $s'' = (t^2 + 1)^2$, which does not divide N "even more" than in the previous example. Specifically, dividing N by s'' using the Euclidean algorithm terminates in three steps instead of two:

$$N = ts' + t^3 + t \qquad \text{(where } r_1 = t(t^2 + 1))$$
$$s'' = tr_1 + r_2 \qquad \text{(where } r_2 = t^2 + 1)$$
$$r_1 = tr_2.$$

Applying Theorem C.1, we get

$$D_{S'',\mathbf{a}} \asymp \min\{N^\alpha, \max\{N^{4\alpha/5}, N^{1/5}\}, \max\{N^{3\alpha/5}, N^{2/5}\}, \max\{N^{2\alpha/5}, N^{3/5}\}\},$$

and $D_{S'',\mathbf{a}}$

$$D_{S'',\mathbf{a}} \asymp \left\{ \begin{array}{ll} N^\alpha & \text{if } \alpha \in (0, 1/5), \\ N^{1/5} & \text{if } \alpha \in [1/5, 1/4), \\ N^{4\alpha/5} & \text{if } \alpha \in [1/4, 1/2), \\ N^{2/5} & \text{if } \alpha \in [1/2, 2/3), \\ N^{3\alpha/5} & \text{if } \alpha \in [2/3, 1), \\ N^{3/5} & \text{if } \alpha \in [1, 3/2), \\ N^{2\alpha/5} & \text{if } \alpha \in [3/2, 2). \end{array} \right.$$

Remark 5.5. In the three examples above, N is the same and s, s' and s'' are comparable in size ($N^{4/5}$), but the resulting diameters $D_{S,\mathbf{a}}$ change according to the length of the Euclidean division of N by s, s' or s''.

Next we give $D_{S,\mathbf{a}}$ for a non-abelian group for different sets of generators.

Example 5.6. Let $G = H_3(\mathbb{Z}/N\mathbb{Z})$ be the group of upper triangular matrices in $M_{3\times 3}$ $(\mathbb{Z}/N\mathbb{Z})$ with 1's on the diagonal. Let g be a element of $H_3(\mathbb{Z}/N\mathbb{Z})$, which we will write of the form

$$\begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.1}$$

Let

$$s_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad s_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad s_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Consider $D_{S,\mathbf{a}}$ with $S = (s_1, s_2, s_3)$ and $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3)$, where each $\alpha_i \in (0,2)$. Then,

$$\|g\|_{S,\mathbf{a}} \asymp \max\left\{ \mid x \mid^{\alpha_1}, \mid y \mid^{\alpha_2}, \min\left\{ \mid z \mid^{\alpha_3}, |z|^{\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}} \right\} \right\}. \tag{5.2}$$

Therefore, $D_{S,\mathbf{a}} \asymp N^{\max\{\alpha_1, \alpha_2, \alpha_3\}}$.

If we also include s_1^t in S, this decreases the cost of elements in both the s_1 and s_3 direction.

Example 5.7. Fix $t > 0$, and $N = t^2$. Let $G = H_3(\mathbb{Z}/N\mathbb{Z})$, $S = (s_1, s_1', s_2, s_3)$, $s_1' = s_1^t$, and $\mathbf{a} = (\alpha_1, \alpha_1, \alpha_2, \alpha_3)$. Let $g = s_3^{m_3} s_2^{m_2} s_1^{m_1}$. Define $x(m)$ and $y(m)$ so that $m = y(m)t + x(m)$ where $|x(m)| \le t/2$, and therefore, $|y(m)| \le t$, Then

$$\|g\|_{S,\mathbf{a}} \asymp \max\Big\{ \max\{|x(m_1)|, |y(m_1)|\}^{\alpha_1}, |m_2|^{\alpha_2},$$
$$\min\left\{ |m_3|^{\alpha_3}, \max\{|x(m_3)|, |y(m_3)|\}^{\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}} \right\} \Big\}.$$

Therefore, $D_{S,\mathbf{a}} \asymp N^{\max\left\{ \frac{\alpha_1}{2}, \alpha_2, \alpha_3, \frac{\alpha_1\alpha_2}{2(\alpha_1+\alpha_2)} \right\}}$.

6 Generalizing results to $\mathbf{a} \in (0,\infty)^k$.

In this section, we discuss how to generalize the main results of the paper (Theorem 1.3, Theorem 1.4, and Theorem 4.1) when $\mathbf{a} \in (0,\infty)^k$.

Definition 6.1. *For any $\alpha > 0$, define a function $\Phi_\alpha : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ as follows.*

$$\Phi_\alpha(x) = \begin{cases} \mid x \mid^\alpha & \text{if } \alpha \in (0,2) \\ \mid x \mid^2 / \log \mid x \mid & \text{if } \alpha = 2 \\ \mid x \mid^2 & \text{if } \alpha > 2 \end{cases}. \tag{6.1}$$

We redefine the cost function from (1.6) as follows.

Definition 6.2. *For $g \in G$,*

$$\|g\|_{S,\mathbf{a}} = \min_{\substack{w \in \mathcal{W}: \\ \rho(w)=g}} \left\{ \max_{1 \le i \le k} \left\{ \Phi_{\alpha_i}(\deg_{\mathbf{s}_i}(w)) \right\} \right\}, \tag{6.2}$$

where $\mathcal{W}$ is is the set of words generated by the alphabet $\mathcal{S} = \{\mathbf{s}_1^\pm, \dots, \mathbf{s}_k^\pm\}$, and ρ is the canonical projection from $\mathcal{W}$ to G. We redefine the $(S, \mathbf{a})$-diameter with respect to the new cost function

$$D_{S,\mathbf{a}} = \max_{g \in G} \|g\|_{S,\mathbf{a}}.$$

Theorem 6.3. *Fix $C_0 > 0$ and $0 < \varepsilon < 1$. There exist $c_1, c_2 > 0$ (depending on C_0 and ε) such that for any $\ell, k < C_0$, $\mathbf{a} \in (\varepsilon, 1/\varepsilon)^k$, and any finite nilpotent group G of nilpotency class ℓ generated by a k-tuple S, the $(S, \mathbf{a})$-long jump random walk satifies*

$$c_1/D_{S,\mathbf{a}} \le 1 - \beta_1 \le c_2/D_{S,\mathbf{a}}.$$

To prove this we need the following lemma:

Lemma 6.4. *Fix $N > 0$, $\alpha > 0$, and $G = \mathbb{Z}/N\mathbb{Z}$. Then, there exists $C(\alpha) > 0$ so that for all $r > 0$, $\Phi_\alpha(y) \le r$ and $f \in \ell^2(\pi)$,*

$$\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(x+y)|^2 \le C(\alpha)\Phi_\alpha(y)\, \mathcal{E}_{p_{N,\alpha}}(f, f), \tag{6.3}$$

where

$$p_{N,\alpha}(x) = \frac{c}{(1 + \mid x \mid)^{1+\alpha}} c^{-1} = \sum_{j \in \mathbb{Z}/N\mathbb{Z}} \frac{1}{(1 + \mid j \mid)^{1+\alpha}}.$$

Proof. When $\alpha > 2$, both $\mathcal{E}_p$ and $\mathcal{E}_{p_{N,\alpha}}$ have finite second moment, where p is the measure that drives lazy simple random walks on $\mathbb{Z}/N\mathbb{Z}$. Therefore, the two forms are comparable up to a constant, see [9, Corollary 1.5]. The case when $\alpha = 2$ is treated in [14, Proposition A.4]. $\square$

Proof of Theorem 6.3. For the lower bound, we repeat the argument from Section 2. The appropriate version of Theorem 2.1 comes from [12, Theorem 2.10]. Lemma 6.4 is the corresponding version of (2.5). From there, the proof follows the same line of reasoning.

For the lower bound, adapting details of Section 2.2 is a straightfoward calculus exercise. The main details of the computation is also covered in [1, Lemma 4.2]. $\square$

Still following the same reasoning as for $\mathbf{a} \in (0, 2)^k$, we arrive to the following theorem.

Theorem 6.5. *Let K be the Markov kernel of an $(S, \mathbf{a})$-long-jump random walk on a finite group G with nilpotency ℓ, and π be the uniform distribution. There exists $b_1, b_2, c_1, c_2 > 0$ such that for all $n > 0$*

$$\frac{c_1}{V(e, n)^{1/2}} \exp\left(-n/b_1 D_{S,\mathbf{a}}\right) \le ||k_e^n - 1||_2 \le \frac{c_2}{V(e, n)^{1/2}} \exp\left(-n/b_2 D_{S,\mathbf{a}}\right),$$

where

$$V(x, r) = \sum_{y \in G : ||x^{-1}y||_{S,\mathbf{a}} \le r} \pi(y).$$

Example 6.6. Fix $t > 0$, and let $N = t^2$, $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1, t)$, and $\mathbf{a} = (1, 2)$. For each $g \in \mathbb{Z}/N\mathbb{Z}$, we can write $g = x_1 + x_2 t$ so that $|x_1|$ and $|x_2|$ are strictly less than t. Then,

$$||g||_{S,\mathbf{a}} = ||x_1 + x_2 t||_{S,\mathbf{a}} \asymp \max\left\{|x_1|, \frac{|x_2|^2}{\log |x_2|}\right\}.$$

A A note on properties of $\mu_{S,\mathbf{a}}$

In this section, we prove some useful lemmas about $\mu_{S,\mathbf{a}}$ that are used frequently throughout the paper, such as bounds for the normalization constants and its moments. We will also discuss how $\mu_{S,\mathbf{a}}$ compares to the probability measure that drives the wrap-around model.

Lemma A.1. *Let N be a positive integer, α be a positive real number, and*

$$c = \left(\sum_{j \in \mathbb{Z}/N\mathbb{Z}} \frac{1}{(1 + \mid j \mid)^{1+\alpha}} \right)^{-1},$$

the normalization constant on individual cycles from Definition 1.1. Then,

$$\frac{\alpha}{2(1+\alpha)} \leq c \leq 1.$$

Proof. Since the summand corresponding to $j = 0$ is 1, c is less than or equal to 1. For the lower bound, we have that for all $N \geq 1$,

$$c^{-1} \leq 2\left(1 + \sum_{k=1}^{N/2} \frac{1}{(1+k)^{1+\alpha}} \right) \leq 2\left(1 + \int_0^{N/2} \frac{1}{(1+s)^{1+\alpha}} \, ds \right) \leq 2\left(1 + \frac{1}{\alpha} \right). \qquad \square$$

Lemma A.2. *Let $p : \mathbb{Z}/N\mathbb{Z} \to [0,1]$, $p(j) = c/(1 + \mid j \mid)^{1+\alpha}$ with $\sum_{j \in \mathbb{Z}/N\mathbb{Z}} p(j) = 1$. There exists a constant $C(\alpha) > 0$, so that*

$$\sum_{|t|>a} p(t) \leq \frac{C(\alpha)}{a^\alpha} \sum_{|t|<a} |t|^2 p(t) \leq C(\alpha) a^{2-\alpha}.$$

Proof. If $a = 1$, then both sums are less than or equal to $1 = 1/a^{\alpha_i} = a^{2-\alpha}$. If $a > 1$, we can compute

$$\sum_{|t|>a} p(t) \leq 2 \int_{\max(2,a)-1}^\infty p(t) \, dt \leq 2c \int_{a/2}^\infty \frac{dt}{(1+t)^{1+\alpha}}$$

$$= \frac{2c}{\alpha(1+a/2)^\alpha} \leq \frac{2c(2^\alpha)}{\alpha} \frac{1}{a^\alpha}.$$

Moreover,

$$\sum_{|t|<a} |t|^2 p(t) \leq 2 \int_0^{4a} t^2 p(t) \, dt \leq 2c \int_0^{4a} (1+t)^{1-\alpha} \, dt$$

$$\leq \frac{2^5 c}{2-\alpha} a^{2-\alpha}$$

Thus, the statement of the lemma is true for $C(\alpha) = \max\{1, 2c(2^\alpha)/\alpha, 2^5 c/(2-\alpha)\}$. $\square$

Next, we prove that the wrap-around definition described in the introduction and one given in Definition 1.1 are comparable. Specifically, we will show that on cycles, the probability measures are comparable up to multiplicative constants depending only on α. Therefore, the probability measures on the full group are comparable up to constants depending on **a**. For fixed $N > 0$ and $\alpha \in (0,2)$, the measure driving our long-jump random walks on $\mathbb{Z}/N\mathbb{Z}$ is

$$p(g) = \frac{c}{(1 + \mid g \mid)^{1+\alpha}}, \qquad \text{where } c^{-1} = \sum_{j \in \mathbb{Z}/N\mathbb{Z}} \frac{1}{(1 + \mid j \mid)^{1+\alpha}},$$

and the measure driving the wrap-around model is

$$\tilde{p}(g) = \sum_{j \in \mathbb{Z}} \frac{\tilde{c}}{(1 + |g + Nj|)^{1+\alpha}}, \qquad \text{where } \tilde{c}^{-1} = \sum_{g \in \mathbb{Z}/N\mathbb{Z}} \left(\sum_{j \in \mathbb{Z}} \frac{1}{(1 + |g + Nj|)^{1+\alpha}} \right).$$

Lemma A.3. *For all $\alpha > 0$, there exist constants $c_1, c_2 > 0$ depending on α, such that for all positive integer N and $k \in \mathbb{Z}/N\mathbb{Z}$,*

$$c_1 p(k) \leq \tilde{p}(k) \leq c_2 p(k).$$

Proof. Fix an integer $k \in [0, N/2]$, and consider

$$\frac{\tilde{p}(k)}{p(k)} = \frac{\tilde{c}}{c} \sum_{j \in \mathbb{Z}} \frac{(1+k)^{1+\alpha}}{(1 + |k + Nj|)^{1+\alpha}} = \frac{\tilde{c}}{c} \left(1 + \sum_{j \neq 0} \frac{(1+k)^{1+\alpha}}{(1 + |k + Nj|)^{1+\alpha}} \right) \tag{A.1}$$

It will be convenient to define the constant $A = \sum_{j=1}^{\infty} \frac{1}{(1+j)^{1+\alpha}}$. For the lower bound of (A.1), notice that the term in the parenthesis is bounded below by 1. Moreover, since rearranging the summand gives that $\tilde{c}^{-1} \leq 2A$, combined with Lemma A.1, we see that we can set $c_1 = \frac{2(1+\alpha)}{\alpha A}$.

For the upper bound of (A.1), first when $k = 0$, for all $k \in [0, N/2]$ and $j > 0$,

$$\frac{1+k}{1+k+jN} = \frac{1}{1 + j\frac{N}{k+1}} \leq \frac{2}{1+j},$$

where the last inequality is because $k + 1 \leq 2N$. When $j < 0$, we have

$$\frac{1+k}{1 - jN + k} \leq \frac{1}{1 + |j|\frac{N}{k+1}} \leq \frac{2}{1 + |j|}.$$

Then,

$$\sum_{j \neq 0} \frac{(1+k)^{1+\alpha}}{(1 + |k + Nj|)^{1+\alpha}} \leq \sum_{j \neq 0} \left(\frac{2}{1 + |j|} \right)^{1+\alpha} \leq 2^{2+\alpha} A.$$

Thus, we can set $c_2 = (1 + 2^{2+\alpha})A/c$. $\qquad \square$

B Dirichlet form estimates

In this section, we establish various estimates on the Dirichlet form. We will also let π always be the uniform distribution for simplicity of proof, but all theorems can be made to work for arbitrary distributions. The techniques in this section were inspired by techniques developed in [12], [13], and [14]. In particular, see Section 4 of [13].

Proposition B.1. *Let G be a finite group, $||\cdot||$ a quasi-norm on G, μ a probability measure on G, and π be the uniform distribution.*

1. *Suppose that there exists a function $a(r) \geq 0$, such that for all $r \geq 0$, $f \in \ell^2(\pi)$, $y \in B(0, r)$,*

$$\sum_{x \in G} |f(x) - f(xy)|^2 \pi(x) \leq a(r)\, \mathcal{E}_\mu(f, f).$$

   Then for all $r \geq 0$, $f \in \ell^2(\pi)$,

$$||f - f_r||_2^2 \leq a(r)\, \mathcal{E}_\mu(f, f).$$

2. *Fix $s \in G$ and n to be the order of s in G. Let μ is a probability distribution on G of the form*

$$\mu(g) = \sum_{j \in \mathbb{Z}/N\mathbb{Z}} \mathbb{1}_{s^j}(g) p(j),$$

   where p is a probability distribution on $\mathbb{Z}/N\mathbb{Z}$. Let also that $||\cdot||$ be a quasi-norm of the form

$$||g|| = \begin{cases} ||m||_0 & \text{if } g = s^m \\ \infty & \text{otherwise} \end{cases},$$

where $||\cdot||_0$ is a quasi-norm on $\mathbb{Z}/N\mathbb{Z}$.

Suppose that there exists a real-valued function $a(r) \geq 0$ such that for all $r \geq 0$, $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$, $y \in \mathbb{Z}/N\mathbb{Z}$ where $||y||_0 \leq r$,

$$\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(x+y)|^2 \leq a(r)\, \mathcal{E}_p(f,f).$$

Then we have that for all $r \geq 0$, $f : G \to \mathbb{R}$, $y \in G$ where $||y|| \leq r$ then

$$\frac{1}{|G|} \sum_{x \in G} |f(x) - f(xy)|^2 \leq a(r)\, \mathcal{E}_\mu(f,f).$$

3. Let $\mu : G \to \mathbb{R}$ be a convex combination of probability measures μ_i: $\mu = \sum_{i=1}^k c_i \mu_i$. Then for any $f \in \ell^2(\pi)$, then

$$\mathcal{E}_\mu(f,f) = c_1\, \mathcal{E}_{\mu_1}(f,f) + \cdots + c_k\, \mathcal{E}_{\mu_k}(f,f).$$

Proof. 1. For all $r \geq 0$ and $f \in \ell^2(\pi)$, we have

$$||f - f_r||_2^2 \leq \sum_{x \in G} |f(x) - f_r(x)|^2 \pi(x)$$

$$= \sum_{x \in G} \left| \frac{1}{V(e,r)} \sum_{y \in B(e,r)} (f(x) - f(xy))\pi(xy) \right|^2 \pi(x)$$

$$= \sum_{x \in G} \left| \sum_{y \in B(e,r)} (f(x) - f(xy)) \frac{1}{\#B(e,r)} \right|^2 \pi(x)$$

$$\leq \sum_{x \in G} \frac{1}{V(e,r)} \sum_{y \in B(e,r)} |(f(x) - f(xy))|^2 \pi(x) \qquad \text{(by Jensen's inequality)}$$

$$\leq \frac{1}{\#B(e,r)} \sum_{y \in B(e,r)} a(r)\, \mathcal{E}_\mu(f,f) \qquad \text{(by assumption)}$$

$$= a(r)\, \mathcal{E}_\mu(f,f).$$

2. Fix $r > 0$, $f : G \to \mathbb{R}$ and $y_0 \in G$ such that $||y_0|| \leq r$. By the definition of $||\cdot||$, y_0 is of the form s^m where $||m||_0 \leq r$. We denote the cosets of $\langle s \rangle \leq G$ as $[x_j \langle s \rangle]$ where the x_j's are fixed representatives of the cosets. Then we have

$$\mathcal{E}_\mu(f,f) = \frac{1}{|G|} \sum_{x,y \in G} |f(x) - f(xy)|^2 \mu(y)$$

$$= \frac{1}{|G|} \sum_{x,y \in G} |f(x) - f(xy)|^2 \sum_{j \in \mathbb{Z}/N\mathbb{Z}} \mathbb{1}_{s^j}(y) p(j)$$

$$= \frac{1}{|G|} \sum_{x \in G} \sum_{j \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(xs^j)|^2 p(j).$$

For each x there is an unique representation as a product of one of the x_j's and an element in $\langle s \rangle$. So we have

$$\mathcal{E}_\mu(f,f) = \frac{N}{|G|} \sum_{j=1}^{|G|/N} \sum_{\ell,\ell' \in \mathbb{Z}/N\mathbb{Z}} |f(x_j s^{\ell'}) - f(x_j s^{\ell+\ell'})|^2 p(\ell) \frac{1}{N}.$$

Then define $f_j : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ to map $\ell \mapsto f(x_j s^\ell)$, and we have

$$\mathcal{E}_\mu(f,f) = \frac{N}{|G|} \sum_{j=1}^{|G|/N} \mathcal{E}_p(f_j, f_j). \tag{B.1}$$

Then

$$\begin{aligned}
\frac{1}{|G|} \sum_{x \in G} |f(x) - f(xs^m)|^2 &= \frac{1}{|G|} \sum_{j=1}^{|G|/N} \sum_{\ell \in \mathbb{Z}/N\mathbb{Z}} |f(x_j s^\ell) - f(x_j s^{\ell+m})|^2 \\
&= \frac{1}{|G|} \sum_{j=1}^{|G|/N} \sum_{\ell \in \mathbb{Z}/N\mathbb{Z}} |f_j(\ell) - f_j(\ell + m)|^2 \\
&\leq \frac{N}{|G|} a(r) \sum_{j=1}^{|G|/N} \mathcal{E}_\mu(f_j, f_j) \qquad \text{(by assumption)} \\
&= a(r) \mathcal{E}_\mu(f,f) \qquad \text{(by (B.1)).}
\end{aligned}$$

3.

$$\begin{aligned}
\mathcal{E}_\mu(f,f) &= \sum_{x,y} |f(x) - f(xy)|^2 \mu(y)\pi(x) \\
&= \sum_{x,y} |f(x) - f(xy)|^2 (c_1\mu_1(y) + \cdots + c_k\mu_k(y))\pi(x) \\
&= c_1 \sum_{x,y} |f(x) - f(xy)|^2 \mu_1(y)\pi(x) + \cdots + c_k \sum_{x,y} |f(x) - f(xy)|^2 \mu_k(y)\pi(x) \\
&= c_1 \mathcal{E}_{\mu_1}(f,f) + \cdots + c_k \mathcal{E}_{\mu_k}(f,f) \qquad\qquad\qquad \square
\end{aligned}$$

Definition B.2. *Define p a symmetric distribution on $\mathbb{Z}/N\mathbb{Z}$ to satisfy* regularity condition (A) *if there exists a constant $C_p > 0$ such that for all $k \in [0, N/2]$*

$$\min_{I_k} p \geq C_p \max_{I_k} p,$$

where $I_k = [\lfloor k/9 \rfloor, k]$. Since p is symmetric the inequality remains true for $k \in [-N/2, 0]$ with $I_k = [k, \lceil k/9 \rceil]$.

Lemma B.3. *Let $N \geq 0$ and $\alpha > 0$. The probability distribution $p_{N,\alpha} : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$, where*

$$p_{N,\alpha}(x) = \frac{c_{N,\alpha}}{(1+|x|)^{1+\alpha}}, \qquad \text{and} \quad c_{N,\alpha}^{-1} = \sum_{j \in \mathbb{Z}/N\mathbb{Z}} \frac{1}{(1+|j|)^{1+\alpha}}. \tag{B.2}$$

satisfies regularity condition (A) where the constant $C_{p_{N,\alpha}}$ depends only on α, and not N.

Proof. By Lemma A.2, we know that

$$\frac{\alpha}{2(1+\alpha)} \leq c_{N,\alpha} \leq 1.$$

Let $k \in [0, N/2]$. In the trivial case, when $k \in [0, 8]$, I_k only includes 0, so we have $\min_{I_k} p_{N,\alpha} = \max_{I_k} p_{N,\alpha} = c_{N,\alpha}$.

In the nontrivial case, when $k \geq 9$, we have $\lfloor k/9 \rfloor \geq \max\{1, k/9 - 1\}$.

Then we have

$$
\begin{aligned}
\max_{I_k} p_{N,\alpha} = c(1 + \lfloor k/9 \rfloor)^{-(1+\alpha)} &\leq (1 + \max\{1, k/9 - 1\})^{-(1+\alpha)} \\
&\leq 2^{1+\alpha}(2 + 1 + k/9 - 1)^{-(1+\alpha)} \leq 18^{1+\alpha}(18 + k)^{-(1+\alpha)} \\
&\leq 18^{1+\alpha}(1 + k)^{-(1+\alpha)} \leq 18^{1+\alpha}\min_{I_k} p_{N,\alpha}.
\end{aligned}
$$

Thus, $C_{p_{N,\alpha}}$ can be set to $18^{-1-\alpha}$. $\qquad\square$

Next we show that that the pseudo-Poincaré inequality holds for $p_{N,\alpha}$, as defined in (B.2), on the cyclic group.

Theorem B.4. *Fix $\alpha \in (0,2)$ and $n > 0$. Then there exists $C(\alpha)$ so that for all $r > 0$, $|y|^\alpha < r$ and $f \in \ell^2(\pi)$*

$$
\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(x+y)|^2 \leq C(\alpha)|y|^\alpha \, \mathcal{E}_{p_{N,\alpha}}(f, f). \tag{B.3}
$$

Proof. The statement is trivially true when $y = 0$. And for $y \neq 0$, we first define

$$
I_0 = \begin{cases} [\lfloor y/4 \rfloor, y/2] & \text{if } y \geq 0 \\ [y/2, \lceil y/4 \rceil] & \text{if } y \leq 0 \end{cases}.
$$

Note that I_0 is always non-empty, since if $|y| \in \{1, 2, 3\}$, then $0 \in I_0$. For all other y's, $\lfloor |y|/4 \rfloor$ and $\lfloor |y|/2 \rfloor$ are at least one apart. Then, first multiplying the left hand side of (B.3) by $p_{N,\alpha}(y)$, we create two sums A and B:

$$
\sum_{x \in G} |f(x) - f(x+y)|^2 p_{N,\alpha}(y) \leq \frac{1}{|I_0|}\Bigg(\underbrace{\sum_{z \in I_0}\sum_{x \in G} |f(x) - f(x+z)|^2 p_{N,\alpha}(y)}_{A}
$$

$$
+ \underbrace{\sum_{z \in I_0}\sum_{x \in G} |f(x+z) - f(x+y)|^2 p_{N,\alpha}(y)}_{B} \Bigg).
$$

Define

$$
J_y = \begin{cases} [\lfloor y/9 \rfloor, y] & \text{if } y \geq 0 \\ [y, \lceil y/9 \rceil] & \text{if } y < 0 \end{cases}.
$$

By the regularity property (Definition B.2) of $p_{N,\alpha}$ and the fact that $p_{N,\alpha}$ is symmetric and $I_0 \subseteq J_y$, we have

$$
p_{N,\alpha}(y) \leq \frac{1}{C_{p_{N,\alpha}}} p_{N,\alpha}(z) \qquad\qquad \text{for all } y \in I_0
$$

Thus,

$$
A \leq \frac{1}{C_{p_{N,\alpha}}} \sum_{x \in \mathbb{Z}/N\mathbb{Z}}\sum_{z \in I_0} |f(x) - f(x+z)|^2 p_{N,\alpha}(z) \leq \frac{N}{C_{p_{N,\alpha}}} \mathcal{E}_{p_{N,\alpha}}(f, f).
$$

Moreover, since $z \in I_0$, we have

$$
y - z \in \begin{cases} [\lfloor y/4 \rfloor, y/2] & \text{if } y \geq 0 \\ [y/2, \lceil y/4 \rceil] & \text{if } y < 0 \end{cases}.
$$

So

$$p_{N,\alpha}(y) \leq \frac{1}{C_{p_{N,\alpha}}} p_{N,\alpha}(y - z),$$

and

$$B \leq \frac{1}{C_{p_{N,\alpha}}} \sum_{z \in I_0} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x+z) - f(x+y)|^2 p_{N,\alpha}(y-z) \leq \frac{N}{C_{p_{N,\alpha}}} \mathcal{E}_{p_{N,\alpha}}(f, f).$$

Then combining what we computed, we have

$$\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(x+y)|^2 \leq \frac{2}{C_{p_{N,\alpha}} p_{N,\alpha}(y) \# I_0} \mathcal{E}_{p_{N,\alpha}}(f, f) \leq \frac{4|y|^{1+\alpha}}{\# I_0 c_{N,\alpha} C_{p_{N,\alpha}}} \mathcal{E}_{p_{N,\alpha}}(f, f).$$

where the last inequality is by:

$$\begin{cases} (1 + |y|)^{1+\alpha} \leq 2^{1-\alpha} |y|^{1+\alpha} & \text{if } y \neq 0 \\ (1 + |y|)^{1+\alpha} \leq |y|^{1+\alpha} & \text{if } y = 0 \end{cases}.$$

Then to count the number of elements in I_0, we see that when $|y| = 1$, I_0 has one element; when $|y| = 2, 3, 4$, I_0 has 2 elements, and for $|y| \geq 8$, we have

$$\# I_0 \geq \left\lfloor \frac{|y|}{2} - \lfloor |y|/4 \rfloor \right\rfloor \geq \lfloor |y|/4 \rfloor \geq |y|/4 - 1 \geq |y|/8.$$

And for $4 < |y| < 8$, we have that $|y|/4$ is one, and $|y|/2 > 2$, so $\# I_0 \geq 2$. In all cases, the $\# I_0 \geq |y|/8$.

Therefore if we set, using previous bounds for $c_{N,\alpha}$ and $C_{p_{N,\alpha}}$, $C(\alpha)$ to $2^{9+\alpha} 3^{2+\alpha} (\alpha + 1)/\alpha$, then

$$\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f(x) - f(x+y)|^2 \leq C(\alpha) |y|^\alpha \mathcal{E}_\mu(f, f),$$

which is what we were looking for in (B.3). $\square$

By Theorem B.4, Proposition B.1 (2), and the definition of $|| \cdot ||_{s,\alpha}$, we obtain the following theorem.

Theorem B.5. *Let G be finite group, $s \in G$, and $\alpha \in (0, 2)$. Then as defined in the end of Section 1,*

$$\mu_{s,\alpha}(g) = \sum_{\ell \in \mathbb{Z}/N_i\mathbb{Z}} \mathbb{1}_{s_i^\ell}(g) p_i(\ell). \tag{B.4}$$

There exists a constant $C(\alpha) > 0$ such that for all $r \geq 0$, $f \in \ell^2(\pi)$, and $y \in G$ where $||y||_{s,\alpha} \leq r$,

$$\sum_{x \in G} |f(x) - f(xy)|^2 \pi(y) \leq C(\alpha) r \mathcal{E}_{\mu_{s,\alpha}}(f, f),$$

where $C(\alpha)$ can be defined as $2^{9+\alpha} 3^{2\alpha} (1 + \alpha)/\alpha$.

At this point, it may be illustrative to use the above theorem to prove a pseudo-Poincaré inequality for finite abelian groups. Let G be a finite abelian group, S be a k-tuple of generating elements of G, and $\mathbf{a} \in (0, 2)^k$. Fix $r > 0$, $f : G \to \mathbb{R}$, $y \in G$ where $||y||_{S,\mathbf{a}} \leq r$. Then y can be written as $y = y_1 y_2 \cdots y_k$ so that for all i, $||y_i||_{s_i,\alpha_i} \leq r$. So by Theorem B.5, for all $1 \leq i \leq k$,

$$\sum_{x \in G} |f(x) - f(xy_i)|^2 \pi(x) \leq C(\alpha_i) r \mathcal{E}_{\mu_{s_i,\alpha_i}}(f, f),$$

where $C(\alpha_i)$ is the number defined in Theorem B.5. Using these inequalities we have

$$\frac{1}{|G|} \sum_{x \in G} |f(x) - f(xy)|^2 = \frac{1}{|G|} \sum_{x \in G} |f(x) - f(xy_1 \cdots y_k)|^2$$

$$\leq k \sum_{i=1}^{k} \frac{1}{|G|} \sum_{x \in G} |f(x) - f(xy_i)|^2 \quad \text{(by Cauchy-Schwarz inequality)}$$

$$\leq C(\mathbf{a})rk \sum_{i=1}^{k} \mathcal{E}_{\mu_{s_i,\alpha_i}}(f, f) \quad \text{(using Proposition B.1)}$$

$$= C(\mathbf{a})rk^2 \mathcal{E}_{\mu_{S,\mathbf{a}}}(f, f),$$

where $C(\mathbf{a}) = \max_{1 \leq i \leq k} C(\alpha_i)$.

C   Algorithm for computing $D_{S,\mathbf{a}}$ for cyclic groups

The goal of this section is to give an algorithmic way to compute $D_{S,\mathbf{a}}$ on the cyclic group $\mathbb{Z}/N\mathbb{Z}$ with $S = (1, s)$ and $\mathbf{a} = (\alpha_1, \alpha_2)$. We use this process to arrive at the examples outlined in Section 5. For convenience, we will think of elements of $\mathbb{Z}/N\mathbb{Z}$ as integers in $\{0, \ldots, n-1\}$, and fix $1 \leq s \leq N/2$. We know that for all positive integers $0 < a \leq b/2$, there exists positive integers q and r such that

$$b = qa - \varepsilon r,$$

such that $\varepsilon \in \{\pm 1\}$ and $0 \leq r \leq a/2$.

Using this fact to modify the Euclidean algorithm, we can expand N as follows:

$$r_{-1} = N = q_1 s - \varepsilon_1 r_1$$
$$r_0 = s = q_2 r_1 - \varepsilon_2 r_2$$
$$r_1 = q_3 r_2 - \varepsilon_3 r_3$$
$$\vdots$$
$$r_{i-1} = q_{i+1} r_i - \varepsilon_{i+1} r_{i+1}$$
$$\vdots$$
$$r_{K-1} = q_{K+1} r_K - r_{K+1} \tag{C.1}$$

where r_{K+1} is the first r_i that's equal to zero, so r_K is equal to the greatest common divisor of N and s. The connection between this algorithm and continued fractions is well studied, see Section 4.5.3 [8].

For $1 \leq i \leq K$, we choose r_i and ε_i so that

$$r_{i-1} = q_{i+1} r_i - \varepsilon_{i+1} r_{i+1} \tag{C.2}$$
$$r_{i+1} \leq r_i/2 \qquad \text{and} \tag{C.3}$$
$$\text{if } r_{i+1} = r_i/2, \text{ then } \varepsilon_{i+1} = -1. \tag{C.4}$$

For each i, we can write $\varepsilon_i r_i$ in terms of N and s:

$$m_i' N + \bar{\varepsilon}_i r_i = m_i s,$$

for some $m_i, m_i' > 0$, where $\bar{\varepsilon}_i = \varepsilon_1 \cdots \varepsilon_i$. One should interpret this as "using m_i s-steps (positive ones only), one can reach $\bar{\varepsilon}_i r_i$ by going around the circle m_i' times."

Using the expansion in (C.1), we get

$$CN + \bar{\varepsilon}_{i+1} r_{i+1} = (q_{i+1} m_i - \varepsilon_i m_{i-1}) s.$$

Immediately, we see that m_i satisfy the recurrence relation

$$m_{i+1} = q_{i+1} m_i - \varepsilon_i m_{i-1}$$

with base cases $m_{-1} = 0$ and $m_0 = 1$. The next few elements in the series are

$$m_1 = q_1 \qquad m_2 = q_1 q_2 - \varepsilon_1 \qquad m_3 = q_1 q_2 q_3 - \varepsilon_1 q_3 - \varepsilon_2 q_1$$

As we will show, the sequence of m_i's for $-1 \le m_i \le K$ is non-negative and strictly increasing. In addition, m_i is the smallest positive integer, ℓ such that $\ell s = \bar{\varepsilon}_i r_i \mod N$.

Theorem C.1.

$$D_{S,\mathbf{a}} \asymp \min_{-1 \le i \le K} \{\max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}\}.$$

In other words, there exist constants $c_1, c_2 > 0$ such that

$$c_1 \min_i \{\max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}\} \le D_{S,\mathbf{a}} \le c_2 \min_i \{\max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}\}.$$

In particular we can set $c_1 = 1/2^{5(\alpha_1 + \alpha_2)}$ and $c_2 = 1$.

First, we present some simple corollaries.

Corollary C.2.   1. *Let $N = st$, where $s, t > 0$, $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1, s)$, and $\mathbf{a} = (\alpha_1, \alpha_2)$. Then,*

$$D_{S,\mathbf{a}} \asymp \min\{N^{\alpha_1}, \max\{s^{\alpha_1}, t^{\alpha_2}\}\}.$$

2. *Suppose $s^{\alpha_1} \le (N/s)^{\alpha_2}$. Let $N = st$, where $s, t > 0$, $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1, s)$, and $\mathbf{a} = (\alpha_1, \alpha_2)$. Then,*

$$D_{S,\mathbf{a}} \asymp \min\{N^{\alpha_1}, (N/s)^{\alpha_2}\}.$$

3. *Let $N = st_1 + s_2$, where $0 \le s_2 \le s/2$ and $s = s_2 t_2$, $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1, s)$, and $\mathbf{a} = (\alpha_1, \alpha_2)$. Then,*

$$D_{S,\mathbf{a}} \asymp \min\{N^{\alpha_1}, \max\{s^{\alpha_1}, t_1^{\alpha_2}\}, \max\{(s_2^{\alpha_1}, (t_1 t_2)^{\alpha_2}\}\}.$$

4. *Let*

$$N = st_1 + s_2$$
$$s = s_2 t_2 + s_3$$
$$s_2 = s_3 t_3,$$

where $0 < s$, $0 < s_2 \le s/2$, and $0 < s_3 \le s_2/2$, $G = \mathbb{Z}/N\mathbb{Z}$, $S = (1, s)$, and $\mathbf{a} = (\alpha_1, \alpha_2)$. Then

$$D_{S,\mathbf{a}} \asymp \min\{N^{\alpha_1}, \max\{s^{\alpha_1}, t_1^{\alpha_2}\}, \max\{(s_2^{\alpha_1}, (t_1 t_2)^{\alpha_2}\}, \max\{s_3^{\alpha_1}, (t_1 t_2 t_3)^{\alpha_2}\}\}.$$

Proof of the upper bound in Theorem C.1. It suffices to show that for all $x \in \mathbb{Z}/N\mathbb{Z}$,

$$\|x\|_{S,\mathbf{a}} \le \min_i \{\max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}\}.$$

Fix i. For all $x \in \mathbb{Z}/N\mathbb{Z}$, $|ks - x| < r_i$ for some $0 \le k \le m_{i+1}$. Then $x = ks + r$ for some $|r| \le r_i$ and $|k| \le m_{i+1}$. Therefore, $\|x\|_{S,\mathbf{a}} \le \max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}$, for all i. Taking the minimum over all i, we achieve the desired result. $\qquad\square$

Figure 1: We visualize $\mathbb{Z}/47\mathbb{Z}$, whose elements are labeled inside the circles. With $s = 17$, we label $[\cdot]_s$ for each element on the outside of the circle. Following the positive s-steps, we can see that r_1 is the first time that the path has visited $(-s/2, s/2)$, at which time the path turn from yellow to red. Similarly, at the 11th s-step, we visit -1 and it is the first time we visit the interval $(-r_1/2, r_1/2)$.

Our proof of the lower bound is much more involved, and will use the following definition and proposition

Definition C.3. *Fix positive integers N and s with $0 < s \leq N/2$, define*

$$[x]_s = \operatorname*{argmin}_{\ell \in \mathbb{Z}} \{|\ell| : \ell s \equiv x \mod N\},$$

and if there are two options, choose the positive one.

Proposition C.4. *For all $0 \leq i \leq K$, $|[r_i]_s| = m_i$.*

First we give some properties of $[\cdot]_s$ in the following lemma:

Lemma C.5.   *1. For any $x \in \mathbb{Z}/N\mathbb{Z}$, represented as $0 < x \leq N/2$, if $[x]_s = -[x]_s$, then x divides N.*

    2. For all $i < K$, $[-r_i]_s = -[r_i]_s$.

    3. Let $x, y \in \mathbb{Z}/N\mathbb{Z}$, where $[x]_s$ and $[y]_s$ are positive. If $[x+y]_s \geq \min([x]_s, [y]_s)$, then

$$[x+y]_s = [x]_s + [y]_s.$$

    4. Let $x, y \in \mathbb{Z}/N\mathbb{Z}$, where $[x]_s$ and $[y]_s$ are positive. If $[x]_s > [y]_s > 0$, then $[x-y]_s = [x]_s - [y]_s$.

Proof. 1. By definition, $[-x]_s$ is either $-[x]_s$ or $[x]_s$. If the former is true, then we are done. If the latter, then let $\ell = [x]_s = [-x]_s$, and then we have

$$\ell s \equiv -x \mod N$$
$$\ell s \equiv x \mod N.$$

Thus, $2x \equiv 0 \mod N$.

2. Follows directly from (1).

3. Let $a = [x]_s$ and $b = [y]_s$. Without loss of generality, we can assume that $a > b$. We know that $[x+y]_s \leq a+b$ by definition. Thus suppose that $0 < [x+y]_s < a+b$ would imply that $0 < [x+y]_s - b < a$, which contradicts the assumption that $[x]_s = a$.

4. We know that $[x-y]_s \leq a-$ by definition. If $[x-y]_s < a-b$, then it would be true that $0 < [x-y]_s + b < a$, which contradicts the assumption that $[x]_s = a$. $\qquad\square$

Proof of Proposition C.4. We prove by induction with the following induction hypotheses:

(a) $[\bar{\varepsilon}_i r_i]_s = m_i$ (note in particular, this means that $[\bar{\varepsilon}_i r_i]_s$ is positive), where $\bar{\varepsilon}_i = \varepsilon_1 \cdots \varepsilon_i$.

(b) For all $x \in S_i = (-2r_i - r_{i-1}, 2r_i + r_{i-1}) \setminus \{\pm r_{i-1}, 0\}$, $|[x]_s| > m_i$.

The base cases are for $i = 0$ and 1, i.e. r_i being s and r_1, which are trivial for both hypotheses.

For the induction step, we first consider when $i < K - 1$, and case-split based on the signs of $-r_{i-1}, -r_i, r_i$, and r_{i-1}:

If the signs are $(-, -, +, +)$, then we know that $\bar{\varepsilon}_{i-1} = \varepsilon_i = 1$. Then by Lemma C.5 (3) and the induction hypothesis, we can deduce that $[r_i]_s = m_i$, $[2r_i]_s = 2m_i$, ..., and $[q_{i+1} r_i]_s = q_{i+1} m_i$. Thus,

$$\begin{aligned}
[\varepsilon_{i+1} r_{i+1}]_s = [\bar{\varepsilon}_{i+1} r_{i+1}]_s &= [q_{i+1} r_i - r_{i-1}]_s \\
&= [q_{i+1} r_i]_s - [r_{i-1}]_s \qquad \text{(Lemma C.5 (4))} \\
&= q_{i+1} m_i - m_{i-1} = m_{i+1}.
\end{aligned}$$

For induction hypothesis (b), note that it suffices to show this for points with positive $[\cdot]_s$, by Lemma C.5 (1). Starting from $-r_{i-1}$, consider the path along positive s-steps. By the induction hypothesis, the next visit to the set S_i is at $r_i - r_{i-1}$, and the next at $2r_i - r_{i-1}$, and so on. Therefore, the first visit to S_{i+1} is at the point $q_{i+1} r_i - r_{i-1}$, by the definition of q_{i+1}. This point is $\varepsilon_{i+1} r_{i+1}$. Therefore, for all other points in S_i, specifically the ones in S_{i-1}, have $[\cdot]_s$ greater than m_{i+1}, if it is realized by a positive s-path.

If the signs are $(-, +, -, +)$, then we know that $\bar{\varepsilon}_{i-1} = 1$ and $\varepsilon_i = -1$. As in the previous case, by Lemma C.5 (2) and the induction hypothesis, $[q_{i+1} r_i]_s = -q_{i+1} m_i$, and

$$[\bar{\varepsilon}_{i+1} r_{i+1}]_s = -[\varepsilon_{i+1} r_{i+1}]_s = -[q_{i+1} r_i - r_{i-1}]_s = q_{i+1} m_i + m_{i-1} = m_{i+1}.$$

We follow the format of the previous case, instead starting at r_{i-1}, considering again positive s-steps. In this case, the next visits to S_i are $r_{i-1} - r_i$, $r_{i-1} - 2r_i$, ..., and $r_{i-1} - q_{i+1} r_i$, which is $\bar{\varepsilon}_{i+1} r_{i+1}$. The last two cases are the same as the cases above with the signs switched.

Remaining are the cases when $i = K$. The possible sign combinations for $-r_{K-1}, -r_K$, r_K, and r_{K-1} are $(-, +, +, +)$ and $(+, +, +, -)$. Then the argument proceeds exactly the same as above. $\qquad\square$

Lemma C.6. *1. Let $r > 0$ and $x \geq n$, where n is a positive integer. Then*

$$\lfloor rx \rfloor \geq \frac{\lfloor rn \rfloor - 1}{\lfloor rn \rfloor} rx \quad \text{and} \quad \lceil rx \rceil \leq \frac{\lceil rn \rceil}{\lceil rn \rceil - 1} rx.$$

Note that the first inequality is only of interest of when $rn > 1$ and the second inequality when $rn > 2$.

2. For all $1 \leq i \leq K + 1$, $q_i \geq 2$.

3. The m_i's are strictly increasing.

4. Let $2 \leq i \leq K$. Then

$$q_{i+1} m_i \geq \frac{m_{i+1}}{4}.$$

5. Let $i \geq 1$. Then

$$q_{i+1} r_i \geq \frac{3 r_{i-1}}{4}.$$

Proof. 1. The proof is simple and we omit it here.

2. Fix $1 \leq i \leq K + 1$.
$$r_{i-2} + \varepsilon_i r_i = q_i r_{i-1}.$$

The algorithm requires that for each $i \geq 0$, $r_{i+1} \leq r_i/2$. For each $i \geq 1$ we know that $r_{i-2} \geq 2r_{i-1}$ and $r_{i-1} \leq r_i/2$, so

$$q_i r_{i-1} = r_{i-2} + \varepsilon_i r_i \geq 2r_{i-1} - r_{i-1}/2 = (3/2)r_{i-1}$$

So $q_i \geq 3/2$, and since the q_i's must be positive integers, $q_i \geq 2$.

3. From our inductive definition we have for the base case

$$m_2 = q_1 q_2 - \varepsilon_1 \geq q_1(2 - 1/2) \geq (3/2)q_1 > m_1.$$

and for the inductive case,

$$\begin{aligned} m_{i+1} &= q_{i+1} m_i - \varepsilon_i m_{i-1} \\ &\geq 2m_i - m_{i-1} \quad \text{(Lemma C.6 (2))} \\ &> m_i \quad \text{(induction hypothesis)}. \end{aligned}$$

4. If $i = 1$, we have

$$q_2 m_1 = q_2 q_1 \geq \frac{3}{4}(q_2 q_1 - 1) \geq \frac{3}{4}(q_2 q_1 - \varepsilon_1) = \frac{3}{4} m_2.$$

If $i > 1$, then $q_{i+1} m_i = m_{i+1} - \varepsilon_i m_{i-1}$. If $m_{i+1}/m_{i-1} \leq 2$, then since m_i's are increasing by Lemma C.6 (3), we know $m_{i+1} \leq 2m_i$, as well. So,

$$q_{i+1} m_i \geq q_{i+1} \frac{m_{i+1}}{2} \geq m_{i+1},$$

with the last inequality following from $q_{i+1} \geq 2$, Lemma C.6 (2). If $m_{i+1}/m_{i-1} \geq 2$,

$$q_{i+1} m_i \geq m_{i+1} - m_{i-1} \geq \frac{1}{2} m_{i+1}.$$

5. $r_{i-1} = q_{i+1} r_i - \varepsilon_i r_{i+1} \leq q_{i+1} r_i + r_{i+1} \leq q_{i+1} r_i + \frac{r_{i-1}}{4}$ $\qquad \square$

Proposition C.7. *Let* $0 \leq i \leq K$. *Define*

$$x_i = \begin{cases} \lfloor q_{i+1}/2 \rfloor \, r_i & \text{if } q_{i+1} \geq 8 \\ r_i & \text{if } q_{i+1} < 8 \end{cases}. \tag{C.5}$$

Then,

$$||x_i||_{S,\mathbf{a}} \geq \frac{1}{2^{5(\alpha_1+\alpha_2)}} \min\{r_{i-1}^{\alpha_1}, m_{i+1}^{\alpha_2}\}.$$

Proof. First we consider the case when $q_{i+1} \geq 8$, and thus $x_i = \lfloor q_{i+1}/2 \rfloor \, r_i$. Let $n_i = \lfloor q_{i+1}/4 \rfloor$. Consider the interval $[-r_{i-1}, r_{i-1}]$ with the points reachable using at most $n_i m_i$ large steps, and the two colors signify the large steps that were used with generators of the opposite sign. The particular picture uses $q_{i+1} = 10$:
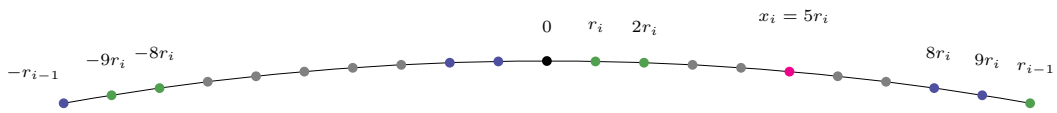
Consider a path that w that maps to x_i under the standard projection. If $\deg_{\mathbf{s}}(w) \geq n_i m_i$ or more large steps,

$$
\begin{aligned}
n_i m_i &= \left\lfloor \frac{q_{i+1}}{4} \right\rfloor m_i \\
&\geq \frac{q_{i+1} m_i}{2^3} \quad \text{(Lemma C.6 (1))} \\
&\geq \frac{1}{2^5} m_{i+1} \quad \text{(Lemma C.6 (4))}
\end{aligned}
$$

If $\deg_{\mathbf{s}}(w) \leq n_i m_i$, then

$$
\begin{aligned}
\deg_1(w) &\geq x_i - n_i r_i = \left(\left\lfloor \frac{q_{i+1}}{2} \right\rfloor - \left\lfloor \frac{q_{i+1}}{4} \right\rfloor \right) r_i \\
&\geq \left(\frac{3}{4} \frac{q_{i+1}}{2} - \frac{q_{i+1}}{4} \right) r_i \quad \text{(Lemma C.6 (1))} \\
&= \frac{q_{i+1}}{2^3} r_i \\
&\geq \frac{r_{i-1}}{2^5} \quad \text{(Lemma C.6 (5).)} \qquad \square
\end{aligned}
$$

Now consider consider the case when $q_{i+1} < 8$, and thus $x_i = r_i$. If we use fewer than m_i large steps, then the number of small steps required is

$$r_i \geq \frac{q_{i+1} r_i}{2^3} \geq \frac{r_{i-1}}{2^5}.$$

If we can use m_i large steps, then we can reach r_i. But

$$m_i \geq \frac{q_{i+1} m_i}{2^3} \geq \frac{m_{i+1}}{2^5}.$$

Proof of the lower bound of Theorem C.1. For the lower bound, let

$$L = \underset{i}{\operatorname{argmin}}\{\max\{r_i^{\alpha_1}, m_{i+1}^{\alpha_2}\}\}.$$

Suppose $L = -1$. Then $N^{\alpha_1} < q_1^{\alpha_2}$ and

$$D_{S,\mathbf{a}} \geq ||x_0||_{S,\mathbf{a}} \geq \frac{1}{2^{5(\alpha_1+\alpha_2)}} N^{\alpha_1}.$$

Now we can assume that $L \geq 0$,

Case 1 ($r_L^{\alpha_1} \leq m_{L+1}^{\alpha_2}$) Then $m_{L+1}^{\alpha_2} = \max\{r_L^{\alpha_1}, m_{L+1}^{\alpha_2}\} \leq \max\{r_{L-1}^{\alpha_1}, m_L^{\alpha_2}\} = r_{L-1}^{\alpha_1}$, since the m_i's are increasing. By substituting $i = L$ into Lemma C.7, we have that

$$D_{S,\mathbf{a}} \geq \frac{1}{2^{5(\alpha_1+\alpha_2)}} \min\{r_{L-1}^{\alpha_1}, m_{L+1}^{\alpha_2}\} \geq \frac{m_{L+1}^{\alpha_2}}{2^{5(\alpha_1+\alpha_2)}}.$$

Case 2 ($r_L^{\alpha_1} \geq m_{L+1}^{\alpha_2}$) Then $r_L^{\alpha_1} = \max\{r_L^{\alpha_1}, m_{L+1}^{\alpha_2}\} \leq \max\{r_{L+1}^{\alpha_1}, m_{L+2}^{\alpha_2}\} = m_{L+2}^{\alpha_2}$, since the r_i's are decreasing. By substituting $i = L+1$ into Lemma C.7, we have that

$$D_{S,\mathbf{a}} \geq \frac{1}{2^{5(\alpha_1+\alpha_2)}} \min\{r_L^{\alpha_1}, m_{L+2}^{\alpha_2}\} \geq \frac{r_L^{\alpha_1}}{2^{5(\alpha_1+\alpha_2)}}. \qquad \square$$

References

[1] Zhen-Qing Chen, Takashi Kumagai, Laurent Saloff-Coste, Jian Wang, and Tianyi Zheng, *Long range random walks and associated geometries on groups of polynomial growth*, arXiv preprint arXiv:1807.00354 (2018).

[2] Thierry Coulhon and Alexander Grigor'yan, *On-diagonal lower bounds for heat kernels and Markov chains*, Duke Math. J. **89** (1997), no. 1, 133–199. MR1458975

[3] Persi Diaconis and Laurent Saloff-Coste, *Moderate growth and random walk on finite groups*, Geom. Funct. Anal. **4** (1994), no. 1, 1–36. MR1254308

[4] Persi Diaconis and Laurent Saloff-Coste, *Nash inequalities for finite Markov chains*, J. Theoret. Probab. **9** (1996), no. 2, 459–510. MR1385408

[5] William Feller, *An introduction to probability theory and its applications. Vol. II*, Second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971. MR0270403

[6] Yves Guivarc'h, *Croissance polynomiale et périodes des fonctions harmoniques*, Bull. Soc. Math. France **101** (1973), 333–379. MR0369608

[7] Ravi Kannan, László Lovász, and Miklós Simonovits, *Random walks and an $O^*(n^5)$ volume algorithm for convex bodies*, Random Structures Algorithms **11** (1997), no. 1, 1–50. MR1608200

[8] Donald E. Knuth, *The art of computer programming. Vol. 2*, Addison-Wesley, Reading, MA, 1998, Seminumerical algorithms, Third edition. MR3077153

[9] Ch. Pittet and L. Saloff-Coste, *On the stability of the behavior of random walks on groups*, J. Geom. Anal. **10** (2000), no. 4, 713–737. MR1817783

[10] Laurent Saloff-Coste, *Lectures on finite Markov chains*, Lecture Notes in Math., vol. 1665, pp. 301–413, Springer, Berlin, 1997. MR1490046

[11] Laurent Saloff-Coste, *Random walks on finite groups*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 263–346. MR2023654

[12] Laurent Saloff-Coste and Tianyi Zheng, *Random walks on nilpotent groups driven by measures supported on powers of generators*, Groups Geom. Dyn. **9** (2015), no. 4, 1047–1129. MR3428409

[13] Laurent Saloff-Coste and Tianyi Zheng, *Random walks under slowly varying moment conditions on groups of polynomial volume growth*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 4, 837–855. MR3434259

[14] Laurent Saloff-Coste and Tianyi Zheng, *Random walks and isoperimetric profiles under moment conditions*, Ann. Probab. **44** (2016), no. 6, 4133–4183. MR3572333

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm