Electronic Journal of Probability

Vol. 11 (2006), Paper no. 1, pages 1-26.

Journal URL
http://www.math.washington.edu/∼ejpecp/

MIXING TIME BOUNDS VIA THE SPECTRAL PROFILE

SHARAD GOEL, RAVI MONTENEGRO AND PRASAD TETALI

ABSTRACT. On complete, non-compact manifolds and infinite graphs, Faber-Krahn inequalities have been used to estimate the rate of decay of the heat kernel. We develop this technique in the setting of finite Markov chains, proving upper and lower L^∞ mixing time bounds via the spectral profile. This approach lets us recover and refine previous conductance-based bounds of mixing time (including the Morris-Peres result), and in general leads to sharper estimates of convergence rates. We apply this method to several models including groups with moderate growth, the fractal-like Viscek graphs, and the product group $\mathbb{Z}_a \times \mathbb{Z}_b$, to obtain tight bounds on the corresponding mixing times.

Keywords and phrases: finite Markov chains, mixing time, spectral profile, conductance, Faber-Krahn inequalities, log-Sobolev inequalities, Nash inequalities

AMS subject classification (2000): Primary 60,68.

Submitted to EJP on May 26, 2005. Final version accepted onOctober 11, 2005.

Research supported in part by NSF grants DMS-0306194, DMS-0401239.

1. Introduction

It is well known that the spectral gap of a Markov chain can be estimated in terms of conductance, facilitating isoperimetric bounds on mixing time (see [SJ89, LS88]). Observing that small sets often have large conductance, Lovász and Kannan ([LK99]) strengthened this result by bounding total variation mixing time for reversible chains in terms of the "average conductance" taken over sets of various sizes. Morris and Peres ([MP]) introduced the idea of evolving sets to analyze reversible and non-reversible chains, and found bounds on the larger L^∞ mixing time.

To sidestep conductance, we introduce "spectral profile" and develop Faber-Krahn inequalities in the context of finite Markov chains, bounding mixing time directly in terms of the spectral profile. FK-inequalities were introduced by Grigor'yan and developed together with Coulhon and Pittet ([Gri94, Cou96, CGP01, BCG01]) to estimate the rate of decay of the heat kernel on manifolds and infinite graphs. Their techniques build on functional analytic methods presented, for example, in [Dav90]. We adapt this approach to the setting of finite Markov chains and derive L^∞ mixing time estimates for both reversible and non-reversible walks.

These bounds let us recover the previous conductance-based results, and in general lead to sharper estimates on rates of convergence to stationarity. We also show that the spectral profile can be bounded in terms of both log-Sobolev and Nash inequalities, leading to new and elementary proofs for previous mixing time results – for example, we re-derive Theorem 3.7 of Diaconis–Saloff-Coste [DSC96a] and Theorem 42 (Chapter 8) of Aldous-Fill [AF].

In terms of applications, we first observe that for simple examples such as the random walk on a complete graph and the n-cycle, the spectral profile gives the correct bounds. As more interesting examples, we also analyze walks on graphs with moderate growth, the fractal-like Viscek graphs, product groups like $Z_a \times Z_b$, and show optimal bounds. In the case of the graphs with moderate growth, we show that the mixing time is of the order of the square of the diameter, a result originally due to Diaconis and Saloff-Coste (see [DSC94, DSC96b]). In the case of the Viscek graphs, we show that the spectral profile provides tight upper and lower bounds on mixing time, and observe that the conductance-based bounds give much weaker upper bounds.

In Section 1 we introduce notation, review preliminary ideas and state our main results. Section 2 presents the proofs of both the continuous and discrete time versions of the spectral profile upper bound on mixing time. In Section 3 we recall a complementary lower bound shown in [CGP01]. Section 4 discusses applications, including the relationship between the spectral profile, log-Sobolev and Nash inequalities. Section 4.3 discusses the more elaborate example of the Viscek graphs. Section 4.4 discusses the spectral profile of the random walk on $Z_a \times Z_b$, which turns out to be a bit subtle.

1.1. Preliminaries.

A Markov chain on a finite state space $\mathcal{X}$ can be identified with a kernel K satisfying

$$K(x,y) \geq 0 \qquad \sum_{y\in\mathcal{X}} K(x,y) = 1.$$

The kernel of K^n is then given iteratively by $K_n(x,y) = \sum_{z \in \mathcal{X}} K_{n-1}(x,z)K(z,y)$ and can be interpreted as the probability of moving from state x to y in exactly n steps.

We say that a probability measure π on $\mathcal{X}$ is invariant with respect to K if $\sum_{x \in \mathcal{X}} \pi(x)K(x,y) = \pi(y)$. That is, starting with distribution π and moving according to the kernel K leaves the distribution of the chain unchanged. Throughout, we assume that K is irreducible: For each $x, y \in \mathcal{X}$ there is an n such that $K_n(x,y) > 0$. Under this assumption K has a unique invariant measure π and $\pi_* = \min_x \pi(x) > 0$.

The chain (K, π) is reversible if, $K = K^*$ is a self-adjoint operator on the Hilbert space $L^2(\pi)$. In general, $K^*(x,y) = \pi(y)K(y,x)/\pi(x)$, and so reversibility is equivalent to requiring that K satisfy the detailed balance equation: for all $x, y \in \mathcal{X}$, we have $\pi(x)K(x,y) = \pi(y)K(y,x)$.

The kernel K describes a discrete-time chain which at each step moves with distribution according to K. Alternatively, we can consider the continuous-time chain H_t, which waits an exponential time before moving. More precisely, as operators

$$H_t = e^{-t\Delta} \qquad \Delta = I - K.$$

The kernel of H_t is then given explicitly by

$$H_t(x,y) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} K^n(x,y).$$

Let $h(x,y,t) = H_t(x,y)/\pi(y)$ denote the density of $H_t(x,\cdot)$ with respect to its stationary measure π.

To measure the rate of convergence to equilibrium, we first need to decide on a distance.

Definition 1.1. For two measures μ and ν with densities $f(x) = \mu(x)/\pi(x)$ and $g(x) = \nu(x)/\pi(x)$ with respect to the positive measure π, their $L^p(\pi)$ distance is

$$d_{p,\pi}(\mu,\nu) = \|f - g\|_{L^p(\pi)} \qquad \text{for } 1 \le p \le \infty.$$

For $p = 1$, this is twice the usual total variation distance. Furthermore, by Jensen's inequality, the function $p \mapsto d_{p,\pi}(\mu,\nu)$ is non-decreasing.

Definition 1.2. The L^p mixing time $\tau_p(\epsilon)$ for the continuous time chain with kernel $H_t(x,y)$ and stationary distribution π is given by

$$\tau_p(\epsilon) = \inf \left\{ t > 0 : \sup_{x \in \mathcal{X}} d_{p,\pi}(H_t(x,\cdot), \pi) \le \epsilon \right\}.$$

Our main result bounds the L^∞ mixing time $\tau_\infty(\epsilon)$, also known as the ϵ-uniform mixing time. Explicitly,

$$\tau_\infty(\epsilon) = \inf \left\{ t > 0 : \sup_{x,y \in \mathcal{X}} \left| \frac{H_t(x,y) - \pi(y)}{\pi(y)} \right| \le \epsilon \right\}.$$

To estimate mixing time, we prove lower bounds on the Dirichlet form associated to the walk.

Definition 1.3. The Dirichlet form associated to K is

$$\mathcal{E}_K(f,g) = \sum_{x \in \mathcal{X}} \Delta f(x) \cdot g(x)\, \pi(x) = \langle \Delta\, f, g \rangle_\pi$$

where $\Delta = I - K$ and $\langle \cdot, \cdot \rangle_\pi$ is the standard inner product for $L^2(\pi)$.

In particular,

$$\mathcal{E}_K(f,f) = \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2 K(x,y)\pi(x). \tag{1.1}$$

Furthermore, $\mathcal{E}_K(f,f) = \mathcal{E}_{K^*}(f,f) = \mathcal{E}_{\frac{K+K^*}{2}}(f,f)$, which follows from equation (1.1) and the identity $K^*(x,y)\pi(x) = K(y,x)\pi(y)$. Fix $x \in \mathcal{X}$ and set $u_{x,t}(y) = h(x,y,t)$. Then recall that

$$\frac{d}{dt}\mathrm{Var}(u_{x,t}) = \sum_y \frac{d}{dt}u_{x,t}^2 d\pi = -2\sum_y u_{x,t}\Delta^* u_{x,t} d\pi = -2\mathcal{E}(u_{x,t}, u_{x,t}). \tag{1.2}$$

This argument motivates the standard definition of the spectral gap

$$\lambda_1 = \inf_f \frac{\mathcal{E}(f,f)}{\mathrm{Var}(f)}$$

and the well known mixing time bounds using the spectral gap:

$$\tau_2(\epsilon) \le \frac{1}{\lambda_1}\log\frac{1}{\epsilon\sqrt{\pi_*}} \quad \text{and} \quad \tau_\infty(1/e) \le \frac{1}{\lambda_1}\left(1+\log\frac{1}{\pi_*}\right). \tag{1.3}$$

Note that by the Courant-Fischer minmax characterization of eigenvalues, λ_1 is the second smallest eigenvalue of the symmetric operator $(\Delta+\Delta^*)/2$.

1.2. **Statement of the Main Result.** Our main result bounds the L^∞ mixing time of a chain through eigenvalues of restricted Laplace operators.

Definition 1.4. For a non-empty subset $S \subset \mathcal{X}$, define

$$\lambda(S) = \inf_{f\in c_0^+(S)} \frac{\mathcal{E}(f,f)}{\mathrm{Var}(f)}$$

where $c_0^+(S) = \{f : \mathrm{supp}(f) \subset S,\ f \ge 0,\ f \ne \mathit{constant}\}$.

In the reversible case,

$$\lambda_0(S) \le \lambda(S) \le \frac{1}{1-\pi(S)}\,\lambda_0(S) \tag{1.4}$$

where λ_0 is the smallest eigenvalue of the restricted Laplacian $\Delta_S : c_0(S) \to c_0(S)$ with $c_0(S) = \{f : \mathrm{supp}(f) \subset S\}$ and

$$\Delta_S f(x) = \begin{cases} \Delta f(x) & x \in S \\ 0 & x \notin S \end{cases}$$

The kernel of $\Delta_S = I - K_S$ is given explicitly by

$$K_S(x,y) = \begin{cases} K(x,y) & x,y \in S \\ 0 & \text{otherwise} \end{cases}$$

By the Courant-Fischer minmax characterization of eigenvalues, (1.4) is equivalent to the statement:

$$\inf_{f\in c_0(S)} \frac{\mathcal{E}_{K_S}(f,f)}{\|f\|_2^2} \le \inf_{f\in c_0^+(S)} \frac{\mathcal{E}_K(f,f)}{\mathrm{Var}(f)} \le \frac{1}{1-\pi(S)} \inf_{f\in c_0(S)} \frac{\mathcal{E}_{K_S}(f,f)}{\|f\|_2^2}$$

The lower bound is due to the identity $\mathcal{E}_K(f,f) = \mathcal{E}_{K_S}(f,f)$ when $f \in c_0(S)$, which follows from $\Delta f(x) = \Delta_S f(x)$ when $f \in c_0(S)$ and $x \in S$. The upper bound also requires the inequality $(x-y)^2 \ge (|x|-|y|)^2$ to show that $\mathcal{E}(f,f) \ge \mathcal{E}(|f|,|f|)$, while Cauchy-Schwartz gives $\|f\|_1 \le \|f\|_2\sqrt{\pi(S)}$ which implies $\mathrm{Var}(f) \ge (1-\pi(S))\|f\|_2^2$.

In general, when $\pi(S) \leq 1/2$ then $\lambda(S)$ is within a factor two of the smallest eigenvalue of the symmetric operator $(\Delta_S + \Delta_S^*)/2$.

We are interested in how $\lambda(S)$ decays as the size of S increases.

Definition 1.5. Define the *spectral profile* $\Lambda : [\pi_*, \infty) \to \mathsf{R}$ by

$$\Lambda(r) = \inf_{\pi_* \leq \pi(S) \leq r} \lambda(S)\,.$$

Observe that $\Lambda(r)$ is non-increasing, and $\Lambda(r) \geq \lambda_1$. For $r \geq 1/2$, Lemma 2.2 shows that $\Lambda(r)$ is within a factor two of the spectral gap λ_1. Furthermore, by construction the walk (K, π) satisfies the *Faber-Krahn* inequality

$$\lambda(S) \geq \Lambda(\pi(S)) \qquad \forall S \subset \mathcal{X}\,.$$

Theorem 1.1 is our main result:

Theorem 1.1. *For $\epsilon > 0$, the L^∞ mixing time $\tau_\infty(\epsilon)$ for a chain $H_t(x,y)$ satisfies*

$$\tau_\infty(\epsilon) \leq \int_{4\pi_*}^{4/\epsilon} \frac{2dv}{v\Lambda(v)}\,.$$

In Section 2.3, we prove an analogous result for discrete-time walks. Since $\Lambda(r) \geq \lambda_1$, Theorem 1.1 shows that

$$\tau_\infty(1/e) \leq \int_{4\pi_*}^{4e} \frac{2\,dv}{v\Lambda(v)} \leq \frac{2}{\lambda_1}\left(1 + \log\frac{1}{\pi_*}\right).$$

But since we can expect $\Lambda(r) \gg \lambda_1$ for small r, Theorem 1.1 offers an improvement over the standard spectral gap mixing time bound (1.3). In particular, by a discrete version of the Cheeger inequality of differential geometry,

$$\Phi_*^2(r)/2 \leq \Lambda(r) \leq 2\Phi_*(r)$$

where $\Phi_*(r)$ is the (truncated) conductance profile (see Section 2.2). Consequently, by Theorem 1.1:

Corollary 1.1. *For $\epsilon > 0$, the L^∞ mixing time $\tau_\infty(\epsilon)$ for a chain $H_t(x,y)$ satisfies*

$$\tau_\infty(\epsilon) \leq \int_{4\pi_*}^{4/\epsilon} \frac{4dv}{v\Phi_*^2(v)}.$$

Theorem 13 of Morris and Peres [MP] is a factor two weaker than this.

Although Theorem 1.1 implies mixing time estimates in terms of conductance, it is reasonable to expect that for many models $\Lambda(r) \gg \Phi_*^2(r)$. In these cases, compared to Corollary 1.1, presently the best known conductance bound, our spectral approach leads to sharper mixing time results. We provide below examples of such cases (see Sections 4.1 and 4.3).

2. Upper Bounds on Mixing Time

2.1. **Spectral Profile Bounds.** In this section, we prove one of the main results, Theorem 1.1. The proof uses the techniques of [Gri94] for estimating heat kernel decay on non-compact manifolds. The first Dirichlet eigenvalue $\lambda_0(S)$ for small sets S captures the convergence behavior at the start of the walk, when the fact that the state space is finite has minimal influence. The spectral gap λ_1 governs the long-term convergence. The spectral profile $\Lambda(r)$ takes into account these two effects, since $\lambda(S) \approx \lambda_0(S)$ for $\pi(S) \leq 1/2$, and $\Lambda(r) \approx \lambda_1$ for $r \geq 1/2$.

To bound mixing times, we first lower bound $\mathcal{E}(f,f)$ in terms of the spectral profile Λ, and as such Lemma 2.1 is the crucial step in the proof of Theorem 1.1.

We regularly use the notation that, given a function f, $f_+ = f \vee 0$ denotes its positive part, and $f_- = -(f \wedge 0)$ its negative part.

Lemma 2.1. *For every non-constant function $u : \mathcal{X} \mapsto \mathsf{R}_+$,*

$$\frac{\mathcal{E}(u,u)}{\operatorname{Var} u} \geq \frac{1}{2}\,\Lambda\Big(\frac{4(\mathbb{E}u)^2}{\operatorname{Var} u}\Big).$$

Proof. For c constant, $\mathcal{E}(u,u) = \mathcal{E}(u-c, u-c)$. Also, $\forall a,b \in \mathsf{R} : (a-b)^2 \geq (a_+ - b_+)^2$ so $\mathcal{E}(f,f) \geq \mathcal{E}(f_+, f_+)$. It follows that when $0 \leq c < \max u$ then

$$\begin{aligned}
\mathcal{E}(u,u) &\geq \mathcal{E}((u-c)_+, (u-c)_+) \\
&\geq \operatorname{Var}((u-c)_+) \inf_{f \in c_0^+(u>c)} \frac{\mathcal{E}(f,f)}{\operatorname{Var}(f)} \\
&\geq \operatorname{Var}((u-c)_+)\,\Lambda(\pi(u>c))\,.
\end{aligned}$$

Now, $\forall a,b \geq 0 :\ (a-b)_+^2 \geq a^2 - 2ba$ and $(a-b)_+ \leq a$ so

$$\begin{aligned}
\operatorname{Var}((u-c)_+) &= \mathbb{E}(u-c)_+^2 - (\mathbb{E}(u-c)_+)^2 \geq \mathbb{E}u^2 - 2c\,\mathbb{E}u - (\mathbb{E}u)^2 \\
&= \operatorname{Var}(u) - 2c\mathbb{E}u.
\end{aligned}$$

Let $c = \operatorname{Var}(u)/4\mathbb{E}u$ and apply Markov's inequality $\pi(u > c) < (\mathbb{E}u)/c$,

$$\mathcal{E}(u,u) \geq (\operatorname{Var}(u) - 2c\,\mathbb{E}u)\,\Lambda(\mathbb{E}u/c) = \frac{1}{2}\operatorname{Var}(u)\,\Lambda\left(\frac{4(\mathbb{E}u)^2}{\operatorname{Var} u}\right).$$

□

Now we bound the L^2 distance of a chain from equilibrium in terms of the function $V(t) : [0,\infty) \to \mathsf{R}$ given by

$$t = \int_{4\pi_*}^{V(t)} \frac{dv}{v\Lambda(v)}.$$

Since the integral diverges, $V(t)$ is well-defined for $t \geq 0$.

The L^2 bound of Theorem 2.1 implies the L^∞ bound that is our main result. To prove the L^2 bound, we simply apply Lemma 2.1 to the heat kernel $h(x,y,t)$.

Theorem 2.1. *For the chain (K,π), we have*

$$\sup_{x\in\mathcal{X}} d_{2,\pi}^2(H_t(x,\cdot),\pi) \leq \frac{4}{V(t)}.$$

Proof. Given $x \in \mathcal{X}$ a value where the supremum occurs, define $u_{x,t}(y) = h(x,y,t)$ and $I_x(t) = \operatorname{Var}(u_{x,t})$. If $u_{x,t} = 1$ then the theorem follows trivially. Otherwise, $u_{x,t}$ is non-constant and since $\mathbb{E}u_{x,t} = 1$, then by (1.2) and Lemma 2.1

$$I_x'(t) = -2\mathcal{E}(u_{x,t}, u_{x,t}) \leq -I_x\Lambda(4/I_x). \tag{2.1}$$

Integrating over $[0,t]$ we have

$$\int_{I_x(0)}^{I_x(t)} \frac{dI_x}{I_x\Lambda(4/I_x)} \leq -t.$$

With the change of variable $v = 4/I_x$,

$$t \leq \int_{4/I_x(0)}^{4/I_x(t)} \frac{dv}{v\Lambda(v)}.$$

Since $I_x(0) = 1/\pi(x) - 1 < 1/\pi_*$

$$V(t) \le \frac{4}{I_x(t)} = \frac{4}{\|h(x,\cdot,t)-1\|_2^2}$$

and the result follows. □

Now we show how to transfer the L^2 bounds of Theorem 2.1 to the L^∞ bounds of our main result.

Proof of Theorem 1.1. Observe that

$$\begin{aligned}
\left|\frac{H_t(x,y)-\pi(y)}{\pi(y)}\right| &= \left|\frac{\sum_z \left(H_{t/2}(x,z)-\pi(z)\right)\left(H_{t/2}(z,y)-\pi(y)\right)}{\pi(y)}\right| \\
&= \left|\sum_z \pi(z)\left(\frac{H_{t/2}(x,z)}{\pi(z)}-1\right)\left(\frac{H^*_{t/2}(y,z)}{\pi(z)}-1\right)\right| \\
&\le d_{2,\pi}(H_{t/2}(x,\cdot),\pi)\, d_{2,\pi}(H^*_{t/2}(y,\cdot),\pi) \qquad (2.2)
\end{aligned}$$

where the inequality follows from Cauchy-Schwartz. Since we can apply Theorem 2.1 to either H_t or H_t^*, we have

$$\sup_{x,y\in\mathcal{X}} |h(x,y,t)-1| \le \frac{4}{V(t/2)}.$$

So $|h(x,y,t)-1| \le \epsilon$ for $V(t/2) \ge 4/\epsilon$, that is, for t such that

$$t/2 \ge \int_{4\pi_*}^{4/\epsilon} \frac{dv}{v\Lambda(v)}$$

proving the result. □

The next result shows that any improvement in using the spectral profile $\Lambda(r)$ instead of the spectral gap λ_1 comes from looking at small sets since for $r = 1/2$, already $\Lambda(r) \approx \lambda_1$.

Lemma 2.2. *The spectral gap λ_1 and the spectral profile $\Lambda(r)$ satisfy*

$$\lambda_1 \le \Lambda(1/2) \le 2\lambda_1.$$

Proof. The lower bound follows immediately from the definition of the spectral gap. For the upper bound, let m be a median of f. Then using Lemma 2.3,

$$\begin{aligned}
\mathcal{E}(f,f) &= \mathcal{E}(f-m,f-m) \\
&\ge \mathcal{E}((f-m)_+,(f-m)_+) + \mathcal{E}((f-m)_-,(f-m)_-).
\end{aligned}$$

Since $\pi(\{f>m\}) = \pi(\{f<m\}) \le 1/2$, we have

$$\mathcal{E}((f-m)_+,(f-m)_+) \ge \|(f-m)_+\|_2^2\, \lambda_0(\{f>m\})$$

and

$$\mathcal{E}((f-m)_-,(f-m)_-) \ge \|(f-m)_-\|_2^2\, \lambda_0(\{f<m\})$$

Consequently,

$$\begin{aligned}
\mathcal{E}(f,f) &\ge \|f-m\|_2^2 \inf_{\pi(S)\le 1/2} \lambda_0(S) \\
&\ge \operatorname{Var}(f)\,\frac{\Lambda(1/2)}{2}\,.
\end{aligned}$$

The upper bound follows by minimizing over f. □

The proof required the following lemma.

Lemma 2.3. *Given a function* $f : \mathcal{X} \mapsto \mathsf{R}$ *then*

$$\mathcal{E}(f,f) \geq \mathcal{E}(f_+,f_+) + \mathcal{E}(f_-,f_-) \geq \mathcal{E}(|f|,|f|).$$

Proof. Given $g, h : \mathcal{X} \mapsto \mathsf{R}$ with $g, h \geq 0$ and $(\operatorname{supp} g) \cap (\operatorname{supp} h) = \emptyset$ then

$$\mathcal{E}(g,h) = \sum_x g(x)h(x)\pi(x) - \sum_{x,y} g(y)h(x)K(x,y)\pi(x) \leq 0$$

because the first sum is zero and every term in the second is non-negative. In particular, $f_+, f_- \geq 0$ with $(\operatorname{supp} f_+) \cap (\operatorname{supp} f_-) = \emptyset$, and so by linearity

$$\begin{aligned}
\mathcal{E}(f,f) &= \mathcal{E}(f_+ - f_-, f_+ - f_-) \\
&= \mathcal{E}(f_+,f_+) + \mathcal{E}(f_-,f_-) - \mathcal{E}(f_+,f_-) - \mathcal{E}(f_-,f_+) \\
&\geq \mathcal{E}(f_+,f_+) + \mathcal{E}(f_-,f_-) \\
&\geq \mathcal{E}(f_+,f_+) + \mathcal{E}(f_-,f_-) + \mathcal{E}(f_+,f_-) + \mathcal{E}(f_-,f_+) \\
&= \mathcal{E}(|f|,|f|).
\end{aligned}$$

□

2.2. **Conductance Bounds.** In this section, we show how to use Theorem 1.1 to recover previous bounds on mixing time in terms of the conductance profile.

Definition 2.1. For non-empty $A, B \subset \mathcal{X}$, the flow is given by

$$Q(A,B) = \sum_{x \in A,\, y \in B} Q(x,y)$$

where $Q(x,y) = \pi(x)K(x,y)$ can be viewed as a probability measure on $\mathcal{X} \times \mathcal{X}$. The boundary of a subset is defined by

$$\partial S = \{x \in S : \exists y \notin S,\ K(x,y) > 0\}$$

and $|\partial S| = Q(S, S^c)$.

Observe that

$$\pi(S) = Q(S, \mathcal{X}) = Q(S,S) + Q(S,S^c)$$

and also

$$\pi(S) = Q(\mathcal{X}, S) = Q(S,S) + Q(S^c,S).$$

It follows that $Q(S,S^c) = Q(S^c,S)$.

Like the spectral profile $\Lambda(r)$, the conductance profile $\Phi(r)$ measures how conductance changes with the size of the set S.

Definition 2.2. Define the conductance profile $\Phi : [\pi_*, 1) \to \mathsf{R}$ by

$$\Phi(r) = \inf_{\pi_* \leq \pi(S) \leq r} \frac{|\partial S|}{\pi(S)}$$

and the truncated conductance profile $\Phi_* : [\pi_*, 1) \to \mathsf{R}$ by

$$\Phi_*(r) = \begin{cases} \Phi(r) & r < 1/2 \\ \Phi(1/2) & r \geq 1/2 \end{cases}$$

The value $\Phi(1/2)$ is often referred to as the conductance, or the isoperimetric constant, of the chain.

The next lemma is a discrete version of the "Cheeger inequality" of differential geometry, and will let us apply Theorem 1.1 to recover the conductance profile bound of Corollary 1.1. The proof of the lemma is similar to the proof given in [SC96] of the fact that

$$\frac{\Phi^2(1/2)}{8} \le \lambda_1 \le 2\Phi(1/2).$$

Lemma 2.4. *For $r \in [\pi_*, 1)$, the spectral profile Λ and the conductance profile Φ satisfy*

$$\frac{\Phi^2(r)}{2} \le \Lambda(r) \le \frac{\Phi(r)}{1-r}.$$

Proof. It suffices to show that $\frac{1}{2}\,\Phi^2(\pi(A)) \le \lambda_0(A) \le \frac{|\partial A|}{\pi(A)}$ for every $A \subset \mathcal{X}$. The bound then follows from (1.4) by minimizing over sets with $\pi(A) \le r$.

For the upper bound,

$$\lambda_0(A) \le \frac{\mathcal{E}(1_A, 1_A)}{\|1_A\|_2^2} = \frac{|\partial A|}{\pi(A)}.$$

To show the lower bound, for a non-negative function f, define the level sets $F_t = \{x \in \mathcal{X} : f(x) \ge t\}$ and the indicator functions $f_t = 1_{F_t}$. Then

$$\begin{aligned} \pi(f) &= \sum_{x\in\mathcal{X}} \left(\int_0^\infty f_t(x)dt\right)\pi(x) \\ &= \int_0^\infty \pi(F_t)\,dt. \end{aligned} \tag{2.3}$$

Furthermore,

$$\begin{aligned} \sum_{x,y} |f(x)-f(y)|\,Q(x,y) &= \sum_{f(x)>f(y)} [f(x)-f(y)]\,[Q(x,y)+Q(y,x)] \\ &= \sum_{f(x)>f(y)} \int_0^\infty 1_{\{f(y)<t\le f(x)\}}\,[Q(x,y)+Q(y,x)]\,dt \\ &= \int_0^\infty |\partial F_t|\,dt + \int_0^\infty |\partial F_t^c|\,dt \\ &= 2\int_0^\infty |\partial F_t|\,dt. \end{aligned} \tag{2.4}$$

Observe that (2.4) is a discrete analog of the co-area formula. For non-negative $f \in c_0(A)$, $F_t \subset A$ for $t > 0$, and so

$$\begin{aligned} \sum_{x,y} |f(x)-f(y)|Q(x,y) &= 2\int_0^\infty |\partial F_t|\,dt \qquad \text{by (2.4)} \\ &\ge 2\Phi(\pi(A))\int_0^\infty \pi(F_t)\,dt \\ &= 2\Phi(\pi(A))\pi(f) \qquad \text{by (2.3).} \end{aligned}$$

Consequently, for any non-negative $f \in c_0(A)$, by the above

$$\begin{aligned}
2\Phi(\pi(A))\pi(f^2) &\le \sum_{x,y} |f^2(x) - f^2(y)|\, Q(x,y) \\
&= \sum_{x,y} |f(x) - f(y)| \cdot (f(x) + f(y))\, Q(x,y) \\
&\le \left(\sum_{x,y} (f(x) - f(y))^2\, Q(x,y)\right)^{1/2} \\
&\qquad \times \left(\sum_{x,y} (f(x) + f(y))^2\, Q(x,y)\right)^{1/2} \\
&\le (2\mathcal{E}(f,f))^{1/2}\,(4\pi(f^2))^{1/2}.
\end{aligned}$$

Then

$$\lambda_0(A) = \inf_{f \in c_0^+(A)} \frac{\mathcal{E}(f,f)}{\pi(f^2)} \ge \frac{\Phi^2(\pi(A))}{2}\,.$$

The infimum for $\lambda_0(A)$ occured at $f \in c_0^+(A)$ because for general $f \in c_0(A)$, $\mathcal{E}(f,f) \ge \mathcal{E}(|f|,|f|)$ and $\pi(f^2) = \pi(|f|^2)$. $\square$

Remark 2.1. ¿From the proofs of Lemma 2.4 and Lemma 2.2, we have that

$$\frac{\Phi^2(1/2)}{2} \le \inf_{\pi(A)\le 1/2} \lambda_0(A) \le \lambda_1.$$

Consequently, when $r > 1/2$ then $\Phi_*^2(r)/2 \le \lambda_1 \le \Lambda(r)$, proving the conductance profile bound of Corollary 1.1.

2.3. **Discrete-Time Walks.** In this section we consider discrete-time chains, deriving spectral profile bounds on mixing time similar to those for continuous-time walks. For $u_{x,t}(y) = h(x,y,t)$, the rate of decay of the heat operator in the continuous setting is given by

$$\frac{d}{dt}\mathrm{Var}(u_{x,t}) = -2\mathcal{E}(u_{x,t}, u_{x,t}). \tag{2.5}$$

In the discrete-time setting, set $u_{x,n}(y) = k(x,y,n) = \frac{K_n(x,y)}{\pi(y)}$. Then, since $K^* u_{x,n} = u_{x,n+1}$ and $\mathbb{E}(u_{x,n}) = 1$,

$$\begin{aligned}
\mathrm{Var}(u_{x,n+1}) - \mathrm{Var}(u_{x,n}) &= \langle u_{x,n+1}, u_{x,n+1}\rangle - \langle u_{x,n}, u_{x,n}\rangle \\
&= -\langle (I - KK^*)u_{x,n},\, u_{x,n}\rangle \\
&= -\mathcal{E}_{KK^*}(u_{x,n}, u_{x,n})
\end{aligned} \tag{2.6}$$

and so it is natural to consider the multiplicative symmeterizations KK^* and K^*K. In order to relate mixing time directly to the kernel K of the original walk, we use the assumption that for $\alpha > 0$

$$K(x,x) \ge \alpha \qquad \forall x \in \mathcal{X}.$$

Define Λ_{KK^*} and V_{KK^*} to be the analogs of Λ and V where $\mathcal{E}_K(f,f)$ is replaced by $\mathcal{E}_{KK^*}(f,f)$. If KK^* is reducible, then $\lambda_1^{KK^*} = 0$, and so we restrict ourselves to the irreducible case. We define Λ_{K^*K} and V_{K^*K} similarly and also assume irreducibility. The following result is a discrete-time version of Theorem 2.1, and its proof is analogous.

Theorem 2.2. *For a discrete-time chain (K,π) with K^*K and KK^* irreducible*

$$\sup_{x\in\mathcal{X}} d_{2,\pi}^2(K_n(x,\cdot),\pi) \le \frac{4}{V_{KK^*}(n/2)} \qquad \textit{and} \qquad \sup_{x\in\mathcal{X}} d_{2,\pi}^2(K_n^*(x,\cdot),\pi) \le \frac{4}{V_{K^*K}(n/2)}.$$

Proof. The second statement follows from the first by replacing K by K^*. For fixed $x \in \mathcal{X}$, define $u_{x,n}(y) = k(x,y,n)$ and $I_x(n) = \mathrm{Var}(u_{x,n})$. By (2.6) and Lemma 2.1

$$\begin{aligned} I_x(n+1) - I_x(n) &= -\mathcal{E}_{KK^*}(u_{x,n},u_{x,n}) \\ &\le -\frac{1}{2} I_x(n)\Lambda_{KK^*}(4/I_x(n)). \end{aligned}$$

Since both $I_x(n)$ and $\Lambda_{KK^*}(r)$ are non-increasing, the piecewise linear extension of $I_x(n)$ to $\mathbb{R}_+$ satisfies

$$I_x'(t) \le -\frac{1}{2} I_x(t)\Lambda_{KK^*}(4/I_x(t)).$$

At integer t, we can take either the derivative from the right or the left. Solving this differential equation as in Theorem 2.1, we have

$$V_{KK^*}(t/2) \le \frac{4}{I_x(t)}$$

and the result follows. □

Corollary 2.1. *Assume that $K(x,x) \ge \alpha > 0$ for all $x \in \mathcal{X}$. Then for $\epsilon > 0$, the L^∞ mixing time for the discrete-time chain K_n satisfies*

$$\tau_\infty(\epsilon) \le 2\left\lceil \int_{4\pi_*}^{4/\epsilon} \frac{dv}{\alpha v \Lambda(v)} \right\rceil.$$

Proof. Since $K^*(x,x) = K(x,x) \ge \alpha$, observe that

$$\begin{aligned} KK^*(x,y)\pi(x) &\ge K^*(x,x)K(x,y)\pi(x) + K^*(x,y)K(y,y)\pi(x) \\ &\ge \alpha K(x,y)\pi(x) + \alpha K(y,x)\pi(y) \end{aligned}$$

and so,

$$\mathcal{E}_{KK^*}(f,f) \ge 2\alpha\mathcal{E}_K(f,f).$$

Consequently, $\Lambda_{KK^*} \ge 2\alpha\Lambda$, from which it follows that

$$\begin{aligned} \alpha t &= \int_{4\pi_*}^{V(\alpha t)} \frac{dv}{v\Lambda(v)} \\ &\ge 2\alpha \int_{4\pi_*}^{V(\alpha t)} \frac{dv}{v\Lambda_{KK^*}(v)}. \end{aligned}$$

Accordingly, $V_{KK^*}(t/2) \ge V(\alpha t)$, and similarly $V_{K^*K}(t/2) \ge V(\alpha t)$. As in Theorem 1.1,

$$\begin{aligned} |k(x,y,2n) - 1| &\le d_{2,\pi}(K_n(x,\cdot),\pi)\, d_{2,\pi}(K_n^*(y,\cdot),\pi) \\ &\le \frac{4}{V(\alpha n)}. \end{aligned}$$

And so, $|k(x,y,2n) - 1| \le \epsilon$ for

$$n \ge \int_{4\pi_*}^{4/\epsilon} \frac{dv}{\alpha v \Lambda(v)}.$$

□

Improvement for discrete-time using rescaling. Given a Markov kernel K let $\Lambda^K(r)$ and $\Phi^K(r)$ denote the spectral and conductance profiles, respectively. Then

$$\begin{aligned}\Lambda^K(r) = (1-\alpha)\Lambda^{\frac{K-\alpha I}{1-\alpha}}(r) &\geq \frac{1-\alpha}{2}\,\Phi^{\frac{K-\alpha I}{1-\alpha}}(r)^2 \\ = \quad \frac{1-\alpha}{2}\left(\frac{\Phi^K(r)}{1-\alpha}\right)^2 &= \frac{\Phi^K(r)^2}{2(1-\alpha)}\end{aligned}$$

The appropriate discrete time version of Corollary 1.1 is then

Corollary 2.2. *For $\epsilon > 0$, the L^∞ mixing time $\tau_\infty(\epsilon)$ for the chain K_n satisfies*

$$\tau_\infty(\epsilon) \leq 2\left\lceil \int_{4\pi_*}^{4/\epsilon} \frac{2\,dv}{\frac{\alpha}{1-\alpha}\,v\Phi_*^2(v)} \right\rceil.$$

In contrast, the bound of Morris and Peres [MP] is

$$\tau_\infty(\epsilon) \leq 2\left\lceil \int_{4\pi_*}^{4/\epsilon} \frac{2\,dv}{\min\left\{\frac{\alpha^2}{(1-\alpha)^2}, 1\right\}\,v\Phi_*^2(v)} \right\rceil$$

which is similar, but slightly weaker when $\alpha \neq 1/2$.

3. Lower Bounds on Mixing Time

In this section, we recall a result of [CGP01] to show that for reversible chains the spectral profile describes well the decay behavior of the heat kernel $h_t(x,y) = H_t(x,y)/\pi(y)$. These results are based on the idea of *anti-Faber-Krahn inequalities.*

For reversible chains, (2.2) implies

$$\sup_{x,y} \frac{H_t(x,y)}{\pi(y)} - 1 \leq \sup_x \sum_z \pi(z)\left(\frac{H_{t/2}(x,z)}{\pi(z)} - 1\right)^2 = \sup_x \frac{H_t(x,x)}{\pi(x)} - 1$$

and so

$$\sup_{x,y\in\mathcal{X}} h_t(x,y) = \sup_{x\in\mathcal{X}} h_t(x,x)\,.$$

Lemma 3.1 gives a simple lower bound on the heat kernel.

Lemma 3.1 ([CGP01])**.** *For a reversible chain (K,π) and non-empty $S \subset \mathcal{X}$,*

$$\sup_{x\in\mathcal{X}} h_t(x,x) \geq \frac{\exp(-t\lambda_0(S))}{2\pi(S)}.$$

Proof. Let $\lambda_0(S) \leq \lambda_1(S) \leq \cdots \leq \lambda_{|S|-1}(S)$ be the eigenvalues of $I - K_S$. Then K_S has eigenvalues $\{1-\lambda_i(S)\}$. Since $\mathrm{tr}(K_S^k)$ can be written as either the sum of eigenvalues, or the sum of diagonal entries, we have

$$\begin{aligned}\sum_{i=0}^{|S|-1} (1-\lambda_i(S))^k &= \sum_{x\in S} K_S^k(x,x) \\ &\leq \sum_{x\in S} K_k(x,x).\end{aligned}$$

For k even, all the terms in the first sum are non-negative, and consequently

$$(1-\lambda_0(S))^k \leq \sum_{x\in S} K_k(x,x).$$

Finally, to bound the continuous-time kernel, note that

$$\begin{aligned}\pi(S)\sup_{x\in\mathcal{X}} h_t(x,x) &\geq \sum_{x\in S} h_t(x,x)\pi(x)\\ &\geq \sum_{x\in S} e^{-t}\sum_{k=0}^{\infty}\frac{t^{2k}}{(2k)!}K_{2k}(x,x)\\ &\geq e^{-t}\sum_{k=0}^{\infty}\frac{t^{2k}(1-\lambda_0(S))^{2k}}{(2k)!}\\ &= e^{-t}\frac{\exp[t(1-\lambda_0(S))]+\exp[-t(1-\lambda_0(S))]}{2}\end{aligned}$$

from which the result follows. □

Theorem 3.1 is a partial converse of the upper bound given in Theorem 2.1 under the restriction of δ-regularity.

Definition 3.1. A positive, increasing function $f \in C^1(0,T)$ is δ-regular if for all $0 < t < s \leq 2t < T$

$$\frac{f'(s)}{f(s)} \geq \delta\frac{f'(t)}{f(t)}.$$

Definition 3.2. The walk (K,π) satisfies the anti-Faber-Krahn inequality with function $L : [\pi_*,\infty) \to \mathsf{R}$ if for all $r \in [\pi_*,\infty)$,

$$\inf_{\pi_*\leq\pi(S)\leq r}\lambda_0(S) \leq L(r).$$

Remark 3.1. Observe that (K,π) satisfies the anti-Faber-Krahn inequality with $L(r) = \Lambda(r)$, in light of (1.4).

Theorem 3.1 ([CGP01]). *Let (K,π) be a reversible Markov chain that satisfies the anti-Faber-Krahn inequality with $L : (\pi_*,\infty) \to \mathsf{R}$, and that $\gamma(t)$, defined implicitly by*

$$t = \int_{\pi_*}^{\gamma(t)}\frac{dv}{vL(v)},$$

is δ-regular on $(0,T)$. Then for $t \in (0,\delta T/2)$

$$\sup_{x\in\mathcal{X}} h_t(x,x) \geq \frac{1}{2\gamma(2t/\delta)}.$$

Proof. Fix $t \in (0,\delta T/2)$ and set $r = \gamma(t/\delta)$. By the anti-Faber-Krahn inequality, there exists $S \subset \mathcal{X}$ with $\pi(S) \leq r$ and $\lambda_0(S) \leq L(r)$. Consequently, by Lemma 3.1,

$$\sup_{x\in\mathcal{X}} h_t(x,x) \geq \frac{\exp(-t\lambda_0(S))}{2\pi(S)} \geq \frac{\exp(-tL(r))}{2r}.$$

So, $\sup_x h_t(x,x) \geq \exp(-C_t)$ for $C_t = \log 2r + tL(r)$. Since $L(\gamma(s)) = (\log\gamma)'(s)$

$$C_t = \log 2\gamma(t/\delta) + t(\log\gamma)'(t/\delta).$$

By the mean value theorem, there exists $\theta \in (t/\delta, 2t/\delta)$ such that

$$(\log\gamma)'(\theta) = \frac{\log\gamma(2t/\delta) - \log\gamma(t/\delta)}{t/\delta}.$$

By δ-regularity

$$(\log\gamma)'(\theta) \geq \delta(\log\gamma)'(t/\delta)$$

and so $C_t \leq \log[2\gamma(2t/\delta)]$, showing the result. $\square$

4. Applications

The following lemma, while hardly surprising, is often effective in reducing computation in specific examples. In particular, it is used in computing the spectral profile of the random walk on the n-cycle in the present section.

Lemma 4.1. *Let $S = S_1 \cup \cdots \cup S_k$ be a decomposition of S into connected components. Then*

$$\lambda(S) = \min_{S_i}\{\lambda(S_i)\}.$$

Proof. Clearly $\lambda(S) \leq \min_{S_i}\{\lambda(S_i)\}$, and we need only show the reverse inequality. For a function $f \geq 0$, define $f_{S_i} = 1_{S_i} f$. Then

$$\mathrm{Var}(f) = \mathrm{Var}\left(\sum_{S_i} f_{S_i}\right) = \sum_{S_i} \mathbb{E} f_{S_i}^2 - \left(\sum_{S_i} \mathbb{E} f_{S_i}\right)^2 \leq \sum_{S_i} \mathrm{Var}(f_{S_i}).$$

Consequently,

$$\begin{aligned}
\lambda(S) &= \inf_{f \in c_0^+(S)} \frac{\mathcal{E}(f,f)}{\mathrm{Var}(f)} \\
&= \inf_{f \in c_0^+(S)} \frac{\sum_{S_i} \mathcal{E}(f_{S_i}, f_{S_i})}{\mathrm{Var}(f)} \\
&\geq \inf_{f \in c_0^+(S)} \frac{\sum_{S_i} \lambda(S_i)\mathrm{Var}(f_{S_i})}{\mathrm{Var}(f)}
\end{aligned}$$

and the result follows. $\square$

4.1. First Examples.

4.1.1. *The Complete Graph.* Consider the continuous-time walk on the complete graph in the n-point space $\Omega = \{x_1, \ldots, x_n\}$ with kernel $K(x_i, x_j) = 1/n \ \forall i, j$. To find the eigenvalues of the restricted operator $K_S : c_0(S) \mapsto c_0(S)$, we consider functions $f : \{x_1, \ldots, x_{|S|}\} \mapsto \mathsf{R}$. Since

$$K_S f(x_j) = \frac{1}{n} \sum_{i=1}^{|S|} f(x_i) = \bar{f} \qquad 1 \leq j \leq |S|$$

f is an eigenfunction of K_S with corresponding eigenvalue λ if and only if $\lambda f(x_j) = \bar{f}$ for $1 \leq j \leq |S|$. If $\lambda \neq 0$, then this implies that f is constant with eigenvalue $\lambda = |S|/n$. So, the smallest eigenvalue of $I - K_S$ satisfies $\lambda_0(S) = 1 - |S|/n$, and the second smallest eigenvalue of $I - K$ satisfies $\lambda_1 = 1$. Since

$$\lambda_1 \leq \lambda(S) \leq \frac{\lambda_0(S)}{1 - \pi(S)}$$

$\lambda(S) = 1$ and accordingly $\Lambda(r) \equiv 1$.

Theorem 1.1 then shows that for the complete graph $\tau_\infty(\epsilon) \leq 2\log(n/\epsilon)$. Since the distribution of the chain at any time $t \geq 0$ is given explicitly by

$$H_t(x_i, x_j) = e^{-t}\delta_{x_i}(x_j) + \frac{(1 - e^{-t})}{n}$$

we see that $\tau_\infty(\epsilon) = \log[(n-1)/\epsilon]$, and so our estimate is off by a factor of 2.

4.1.2. *The n-Cycle.* To bound the spectral profile $\Lambda(r)$ for simple random walk on the n-cycle, recall (see Lemma 4.1) that it is sufficient to restrict our attention to connected subsets. Now consider simple random walk on the n-cycle $\Omega = \{x_0, \dots x_{n-1}\}$ given by kernel $K(x_i, x_j) = 1/2$ if $j = i \pm 1 \pmod{n}$ and zero otherwise. By Lemma 4.1, to find $\lambda_0(S)$ we need only consider connected subsets $S \subset \Omega$. For S such that $\pi(S) < 1$, $I - K_S$ corresponds to the tridiagonal Toeplitz matrix with 1's along the diagonal and -1/2's along the upper and lower off-diagonals (and 0's everywhere else). In this case, the least eigenvalue is given explicitly by

$$\lambda_0(S) = 1 - \cos\left(\frac{\pi}{|S|+1}\right).$$

Since the spectral gap satisfies $\lambda_1 = 1 - \cos(2\pi/n)$, we have $\Lambda(r) \approx 1/(rn)^2$ for $1/n \le r \le 1$. Theorem 1.1 then shows the correct $O(n^2)$ mixing time bound.

4.2. **Log-Sobolev and Nash Inequalities.** Logarithmic Sobolev and Nash inequalities are among the strongest tools available to study L^2 convergence rates of finite Markov chains. Log-Sobolev inequalities were introduced by Gross [Gro75, Gro93] to study Markov semigroups in infinite dimensional settings, and developed in the discrete setting by Diaconis and Saloff-Coste [DSC96a]. Nash inequalities were originally formulated to study the decay of the heat kernel in certain parabolic equations (see [Nas58]). Building on ideas in [CKS87, CSC90b, CSC90a], Diaconis and Saloff-Coste [DSC96b] show how to apply Nash's argument to finite Markov chains. In this section we show that both log-Sobolev and Nash inequalities yield bounds on the spectral profile $\Lambda(r)$, leading to new proofs of previous mixing time estimates in terms of these inequalities.

Definition 4.1. The log-Sobolev constant ρ is given by

$$\rho = \inf_{\mathrm{Ent}_\pi f^2 \neq 0} \frac{\mathcal{E}(f,f)}{\mathrm{Ent}_\pi f^2}$$

where the entropy $\mathrm{Ent}_\pi(f^2) = \sum_{x \in \mathcal{X}} f^2(x) \log\left(f^2(x)/\|f\|_2^2\right) \pi(x)$.

Lemma 4.2 (Log-Sobolev)**.** *The spectral profile* $\Lambda(r)$ *and log-Sobolev constant* ρ *satisfy*

$$\Lambda(r) \ge \rho \, \frac{\log(1/r)}{1-r}\,.$$

Proof. By definition

$$\Lambda(r) = \inf_{\pi(S) \le r} \inf_{f \in c_0^+(S)} \frac{\mathcal{E}(f,f)}{\mathrm{Var}_\pi(f)} \ge \rho \inf_{\pi(S) \le r} \inf_{f \in c_0^+(S)} \frac{\mathrm{Ent}_\pi(f^2)}{\mathrm{Var}_\pi(f)}$$

The lemma will follow if for every set $S \subset \mathcal{X}$

$$\inf_{f \in c_0^+(S)} \frac{\mathrm{Ent}_\pi(f^2)}{\mathrm{Var}_\pi(f)} \ge \frac{\log \frac{1}{\pi(S)}}{1 - \pi(S)}\,.$$

Define a probability measure $\pi'(x) = \frac{\pi(x)}{\pi(S)}$ if $x \in S$ and $\pi'(x) = 0$ otherwise. Then

$$\inf_{f \in c_0^+(S)} \frac{\mathrm{Ent}_\pi(f^2)}{\mathrm{Var}_\pi(f)} = \inf_{f \in c_0^+(S)} \frac{\mathrm{Ent}_{\pi'}(f^2) + \log \frac{1}{\pi(S)}\, E_{\pi'} f^2}{E_{\pi'} f^2 - \pi(S)\,(E_{\pi'} f)^2}$$

Rearranging the terms, it suffices to show that

$$\inf_{f\in c_0^+(S)} \frac{\mathrm{Ent}_{\pi'}(f^2)}{\mathrm{Var}_{\pi'}(f)} \geq \frac{\pi(S)\log\frac{1}{\pi(S)}}{1-\pi(S)}\,.$$

However, since $\pi(S)\in(0,1)$ then $\pi(S)\frac{\log(1/\pi(S))}{1-\pi(S)}\leq 1$ and so it suffices that for every probability measure and $f\geq 0$ that $\mathrm{Ent}(f^2)/\mathrm{Var}(f)\geq 1$. This is true, as observed in [LO00] and recalled in Remark 6.7 of [BT03]. $\square$

The bound $\lambda_0(A)\geq\rho\log(1/\pi(A))$ can be shown similarly, but without need for the result of [LO00]. Like log-Sobolev inequalities, Nash inequalities also yield bounds on the spectral profile:

Lemma 4.3 (Nash Inequality)**.** *Given a Nash inequality*

$$\|f\|_2^{2+1/D} \leq C\left[\mathcal{E}(f,f)+\frac{1}{T}\|f\|_2^2\right]\|f\|_1^{1/D}$$

which holds for every function $f:\mathcal{X}\mapsto\mathsf{R}$ *and some constants* C, D, $T\in\mathsf{R}_+$, *then*

$$\Lambda(r)\geq\frac{1}{C\,r^{1/2D}}-\frac{1}{T}.$$

Proof. The Nash inequality can be rewritten as

$$\frac{\mathcal{E}(f,f)}{\|f\|_2^2}\geq\frac{1}{C}\left(\frac{\|f\|_2}{\|f\|_1}\right)^{1/D}-\frac{1}{T}$$

Then,

$$\begin{aligned}\lambda_0(A)=\inf_{f\in c_0(A)}\frac{\mathcal{E}(f,f)}{\|f\|_2^2} &\geq \inf_{f\in c_0(A)}\frac{1}{C}\left(\frac{\|f\|_2}{\|f\|_1}\right)^{1/D}-\frac{1}{T}\\ &\geq \frac{1}{C\,\pi(A)^{1/2D}}-\frac{1}{T}\,.\end{aligned}$$

The final inequality was due to Cauchy-Schwartz: $\|f\|_1\leq\|f\|_2\sqrt{\pi(\mathrm{supp}\,f)}$. The lemma follows by minimizing over $\pi(A)\leq r$. $\square$

Although the spectral profile $\Lambda(r)$ is controlled by the spectral gap λ_1 for $r\geq 1/2$, Nash inequalities tend to be better for r close to 0, and log-Sobolev inequalities for intermediate r. Combining Lemmas 4.2 and 4.3, we get the following bounds on mixing time:

Corollary 4.1. *Given the spectral gap* λ_1 *and the log-Sobolev constant* ρ *and/or a Nash inequality with* $DC\geq T$, $D\geq 1$ *and* $\pi_*\leq 1/4e$, *the* L^∞ *mixing time for the continuous-time Markov chain with* $\epsilon\leq 8$ *satisfies*

$$\begin{aligned}\tau_\infty(\epsilon) &\leq \frac{2}{\rho}\log\log\frac{1}{4\pi_*}+\frac{2}{\lambda_1}\log\frac{8}{\epsilon}\\ \tau_\infty(\epsilon) &\leq 4T+\frac{2}{\lambda_1}\left(2D\log\frac{2DC}{T}+\log\frac{4}{\epsilon}\right)\\ \tau_\infty(\epsilon) &\leq 4T+\frac{2}{\rho}\log\log\left(\frac{2DC}{T}\right)^{2D}+\frac{2}{\lambda_1}\log\frac{8}{\epsilon}\end{aligned}$$

Proof. For the first upper bound use the log-Sobolev bound $\Lambda(r) \geq \rho \log(1/r)$ when $r < 1/2$ and the spectral gap bound when $r \geq 1/2$. Simple integration gives the result.

For the second upper bound use the Nash bound when $r \leq (T/2DC)^{2D}$ and spectral gap bound for the remainder. Then

$$\begin{aligned}
\tau_\infty(\epsilon) &\leq \int_{4\pi_*}^{(T/2DC)^{2D}} \frac{2dr}{r\,\frac{1}{C\,r^{1/2D}}\left(1-\frac{C\,r^{1/2D}}{T}\right)} + \int_{(T/2DC)^{2D}}^{4/\epsilon} \frac{2dr}{r\,\lambda_1} \\
&\leq 4T + \frac{2}{\lambda_1}\,\log\frac{4/\epsilon}{(T/2DC)^{2D}}
\end{aligned}$$

where the second inequality used the bound $1-\frac{C\,r^{1/2D}}{T} \geq 1-\frac{1}{2D} \geq 1/2$ before integrating. Simplification gives the result.

For the mixed bound use the Nash bound when $r \leq (T/2DC)^{2D}$, the log-Sobolev bound for $(T/2DC)^{2D} \leq r < 1/2$ and the spectral gap bound when $r \geq 1/2$. □

Similar discrete time bounds follow from Corollary 2.1. When $\forall x:\ K(x,x) \geq \alpha$ then these bounds are roughly a factor α^{-1} larger than the continuous time case.

These bounds compare well with previous results shown through different methods. For instance, Aldous and Fill [AF] combine results of Diaconis and Saloff-Coste [DSC96a, DSC96b] to show a continuous time bound on reversible chains of

$$\tau_\infty(\epsilon) \leq 2T + \frac{1}{2\rho}\,\log\log\left(\frac{DC}{T}\right)^D + \frac{1}{\lambda_1}\,(4+\log(1/\epsilon))$$

whenever $DC \geq T$.

4.2.1. *Walks with Moderate Growth.* In this section, we describe how estimates on the volume growth of a walk give estimates on the spectral profile $\Lambda(v)$. The treatment given here is analogous to the method of Nash inequalities described in [DSC96b].

Define the Cayley graph of (K,π) to be the undirected graph on the state space $\mathcal{X}$ with edge set $E = \{(x,y) : \pi(x)K(x,y) + \pi(y)K(y,x) > 0\}$. Let $d(x,y)$ be the usual graph distance, and denote the closed ball of radius r around x by $B(x,r) = \{z : d(x,z) \leq r\}$. The volume of $B(x,r)$ is given by $V(x,r) = \sum_{z\in B(x,r)} \pi(z)$.

Definition 4.2. For $A, d \geq 1$, the finite Markov chain (K,π) has (A,d)-moderate growth if

$$V(x,r) \geq \frac{1}{A}\left(\frac{r+1}{\gamma}\right)^d \qquad \forall x\in\mathcal{X},\ 0\leq r\leq\gamma \tag{4.1}$$

where γ is the diameter of the graph.

For any f and $r \geq 0$, set

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

Definition 4.3. The finite Markov chain (K,π) satisfies a local Poincaré inequality with constant a if for all f and $r \geq 0$

$$\|f-f_r\|_2^2 \leq ar^2\mathcal{E}(f,f). \tag{4.2}$$

Under assumptions (4.1) and (4.2), Diaconis and Saloff-Coste [DSC96b] derive the Nash inequality

$$\|f\|_2^{2+4/d} \leq C\left[\mathcal{E}(f,f) + \frac{1}{a\gamma^2}\|f\|_2^2\right]\|f\|_1^{4/d} \tag{4.3}$$

where $C = (1+1/d)^2(1+d)^{2/d}A^{2/d}a\gamma^2$. By Lemma 4.3, this immediately implies the lower bound on the spectral profile

$$\Lambda(v) \geq \left(\frac{d^2}{(d+1)^{2+2/d}A^{2/d}v^{2/d}} - 1\right)\frac{1}{a\gamma^2}.$$

Theorem 4.1 below shows how to bound $\Lambda(r)$ in terms of a local Poincaré inequality and the volume growth function

$$V_*(r) = \inf_x V(x,r).$$

The proof is similar to the derivation of Nash inequalities for walks with moderate growth shown in [DSC96b].

Theorem 4.1. *Let (K,π) be a finite Markov chain that satisfies the local Poincaré inequality with constant a. For $v \leq 1/2$, the spectral profile satisfies*

$$\Lambda(v) \geq \frac{1}{4aW^2(2v)}$$

where $W(v) = \inf\{r : V_*(r) \geq v\}$.

Proof. Fix $S \subset \mathcal{X}$ with $\pi(S) \leq 1/2$ and $f \in c_0(S)$. It is sufficient to show that

$$\frac{\mathcal{E}(f,f)}{\|f\|_2^2} \geq \frac{1}{4aW^2(2\pi(S))}.$$

First observe that

$$\begin{aligned}\|f\|_2^2 &= \langle f - f_r, f\rangle + \langle f_r, f\rangle \\ &\leq \|f - f_r\|_2 \cdot \|f\|_2 + \langle f_r, f\rangle.\end{aligned}$$

Now,

$$\begin{aligned}\langle f_r, f\rangle &= \sum_x \left(\frac{1}{V(x,r)}\sum_{y\in B(x,r)} f(y)\pi(y)\right) f(x)\pi(x) \\ &\leq \frac{1}{V_*(r)}\|f\|_1^2 \\ &\leq \frac{\pi(S)}{V_*(r)}\|f\|_2^2.\end{aligned}$$

Consequently, by the local Poincaré inequality,

$$\|f\|_2^2 \leq \sqrt{a}r\mathcal{E}(f,f)^{1/2}\|f\|_2 + \frac{\pi(S)}{V_*(r)}\|f\|_2^2.$$

Dividing by $\|f\|_2^2$ and choosing $r = W(2\pi(S))$ we have

$$1 \leq \sqrt{a}r\frac{\mathcal{E}(f,f)^{1/2}}{\|f\|_2} + 1/2$$

and the result follows. ∎

Corollary 4.2. *Let (K, π) be a finite Markov chain that satisfies (A, d)-moderate growth and the local Poincaré inequality with constant a. Then the L^∞ mixing time satisfies*

$$\tau_\infty(\epsilon) \le C(a, A, d, \epsilon)\gamma^2$$

where γ is the diameter of the graph and $C(a, A, d, \epsilon)$ is a constant depending only on a, A, d and ϵ.

Proof. By the moderate growth assumption, $W(v) \le \gamma(Av)^{1/d}$. And so for $v \le 1/2$

$$\Lambda(v) \ge \frac{1}{4aW^2(2v)} \ge \frac{1}{8aA^{1/d}\gamma^2 v^{2/d}}.$$

For $v \ge 1/2$, note that $\Lambda(v) \ge \lambda_1 \ge \Lambda(1/2)/2$. The result now follows immediately from Theorem 1.1. □

In Theorem 3.1 of [DSC94] Diaconis and Saloff-Coste show that for walks on groups with (A, d)-moderate growth, local Poincaré inequality with constant a, and $\gamma \ge A4^{d+1}$

$$\tau_\infty(1/e) \ge \frac{\gamma^2}{4^{2d+1}A^2}\,.$$

It follows that $\tau_\infty(1/e) = \Theta(\gamma^2)$, and Corollary 4.2 was of the correct order γ^2.

For instance, consider the example of simple random walk on the n-cycle discussed in Section 4.1.2. For this walk $V(x_i, r) = (1 + 2\lfloor r\rfloor)/n$, and so it satisfies the moderate growth criterion (4.1) with $A = 6$, $d = 1$ and diameter $\gamma = \lfloor n/2\rfloor$. Moreover, it is shown in [DSC96b] that every group walk satisfies the local Poincaré inequality

$$\|f - f_r\|_2^2 \le 2|S|r^2\mathcal{E}(f, f)$$

where S is a symmetric generating set for the walk. Consequently, Corollary 4.2 shows that the walk on the n-cycle mixes in $O(n^2)$ time. For several additional examples of walks with moderate growth, see [DSC94, DSC96b].

4.3. **The Viscek Graphs.** For a random walk (K, π) consider its Cayley graph defined in Section 4.2.1. Define the minimum volume of a disk of radius r by $V_*(r) = \inf_x\{V(x, r)\}$. Here we first use a result of [BCG01] that shows that the spectral profile $\Lambda(r)$ can be bounded in terms of the volume growth $V_*(r)$ alone (see Lemma 4.4). We then apply this technique to analyze walks on the fractal-like Viscek family of finite graphs.

Lemma 4.4 ([BCG01])**.** *Let $Q_* = \inf_{x\sim y}[\pi(x)K(x, y) + \pi(y)K(y, x)]$ and*

$$w(r) = \inf\{k : V_*(k) > r\}.$$

Then

$$\lambda(A) \ge \frac{Q_*}{4\pi(A)w(\pi(A))}.$$

Proof. Fix $f \in c_0(A)$ normalized so that $\|f\|_\infty = 1$. Then,

$$\|f\|_2^2 = \sum_x |f(x)|^2\pi(x) \le \pi(A).$$

Let x_0 be a point such that $|f(x_0)| = 1$ and let $k = \max\{l \in \mathbb{N} : B(x_0, l) \subset A\}$. Then there is a sequence of points $x_0, x_1, \ldots, x_{k+1}$ with $x_i \sim x_{i+1}$, $x_0, \ldots, x_k \in A$ and $x_{k+1} \notin A$. So,

$$\begin{aligned}
\mathcal{E}(f,f) &= \frac{1}{2}\sum_{x,y} |f(x) - f(y)|^2 \pi(x) K(x,y) \\
&\geq \frac{1}{2}\sum_{i=0}^{k} |f(x_{i+1}) - f(x_i)|^2 [\pi(x_i)K(x_i,x_{i+1}) + \pi(x_{i+1})K(x_{i+1},x_i)] \\
&\geq \frac{Q_*}{2(k+2)} \left(\sum_{i=0}^{k} |f(x_{i+1}) - f(x_i)| \right)^2 \\
&= \frac{Q_*}{2(k+2)} |f(x_{k+1}) - f(x_0)|^2 \\
&= \frac{Q_*}{2(k+2)}.
\end{aligned}$$

Consequently,

$$\lambda_0(A) = \inf_{f \in c_0(A)} \frac{\mathcal{E}(f,f)}{\|f\|_2^2} \geq \frac{Q_*}{2(k+2)\pi(A)}.$$

To finish the proof, observe that $\pi(A) \geq V(x_0, k) \geq V_*(k)$, and so $w(\pi(A)) \geq k + 1 \geq (k+2)/2$. $\square$

The Viscek graphs are a two parameter family of finite trees that are inductively defined as follows. Fix the parameter $N \geq 2$, and define $\mathcal{V}_N(0)$ to be the star graph on $N+1$ vertices (i.e. a central vertex surrounded by N vertices). Given $\mathcal{V}_N(n-1)$ choose N vertices $x_1, \ldots, x_N$ such that $d(x_i, x_j) = \mathrm{diam}(\mathcal{V}_N(n-1))$ for $i \neq j$. Construct $\mathcal{V}_N(n)$ by taking $N+1$ copies $\{\mathcal{V}_N^i(n-1)\}_{i=0}^N$ of the $(n-1)^{th}$ generation graph, and for $1 \leq i \leq N$ identifying $x_i^0 \in \mathcal{V}_N^0(n-1)$ with $x_i^i \in \mathcal{V}_N^i(n-1)$. Observe that a different choice of vertices $x_1, \ldots, x_N$ leads to an isomorphic construction. For $N = 2$, $\mathcal{V}_2(n)$ is a path for each n. Figure 1 illustrates the first three generations of a Viscek graph for $N = 4$.

The following lemma bounds the spectral profile and mixing time for simple random walk on $\mathcal{V}_N(n)$. The proof is analogous to the volume growth computation for the infinite Viscek graph $\mathcal{V}_N(\infty) = \lim_{n\to\infty} \mathcal{V}_N(n)$ given in [BCG01] and recalled in [PSC].

Lemma 4.5. *For $N \geq 2$, $r \leq 1$ the spectral profile $\Lambda(r)$ for simple random walk on $\mathcal{V}_N(n)$ satisfies*

$$\frac{a(N)}{\gamma^{d+1} r^{1+1/d}} \leq \Lambda(r) \leq \frac{A(N)}{\gamma^{d+1} r^{1+1/d}} \qquad d = \log_3(N+1)$$

where $\gamma = \mathrm{diam}(\mathcal{V}_N(n)) = 2 \cdot 3^n$ and the constants $a, A > 0$ depend only on N.

In particular, there exist constants $b, B > 0$ depending only on N such that the mixing time for the continuous-time walk satisfies

$$b(N)\gamma^{d+1} \leq \tau_1(1/e) \leq \tau_\infty(1/e) \leq B(N)\gamma^{d+1}.$$

Observe that since the conductance profile for $\mathcal{V}_N(n)$ satisfies

$$\Phi(r) \approx \frac{1}{|E_N(n)| r} \approx \frac{1}{\gamma^d r},$$

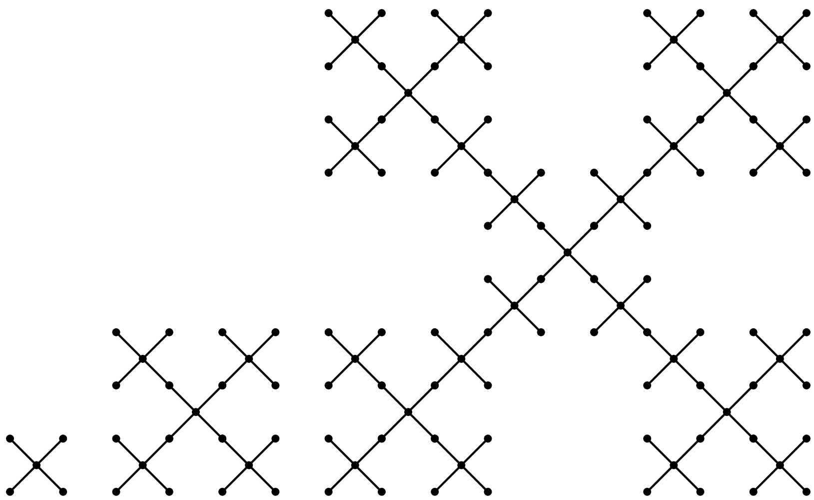

FIGURE 1. The first three generations $\mathcal{V}_4(0)$, $\mathcal{V}_4(1)$ and $\mathcal{V}_4(2)$ of a Viscek graph with $N = 4$.

using the conductance profile bound of Corollary 1.1 results in the upper bound $\tau_\infty(1/e) \preceq \gamma^{2d}$ which overestimates the mixing time for $N \geq 3$.

Proof. We first show that the mixing time bound follows from the spectral profile estimate. The upper bound is a direct consequence of Theorem 1.1. Recall that for an ergodic chain, the spectral gap λ_1 and L^1 mixing time are related by $1/\lambda_1 \leq \tau_1(1/e)$ (see e.g. [SC96]). Since $\Lambda(r) \geq \lambda_1$, the lower bound is immediate.

To estimate the spectral profile, first note that the number of edges $|E_N(n)| = N(N+1)^n$. Since $\mathcal{V}_N(n)$ is a tree, $|\mathcal{V}_N(n)| = N(N+1)^n + 1$. Furthermore, $\mathrm{diam}(\mathcal{V}_N(n)) = 2 \cdot 3^n$.

For $0 \leq k \leq n$, define a k-block to be a subgraph of $\mathcal{V}_N(n)$ isomorphic to the k^{th} generation graph $\mathcal{V}_N(k)$. Fix $x \in \mathcal{V}_N(n)$ and $3 \leq r \leq \mathrm{diam}(\mathcal{V}_N(n))$. Then there is a unique integer m such that $3^{m+1} \leq r < 3^{m+2}$. Moreover, the vertex x is contained in some m block B. Since $\mathrm{diam}(B) = 2 \cdot 3^m$, $B(x,r) \supseteq B$. Consequently,

$$|B(x,r)| \geq |B| = N(N+1)^m + 1$$

and since $\pi_* = 1/(2|E_N(n)|)$

$$V_*(r) \geq \frac{N(N+1)^m + 1}{2N(N+1)^n} \succeq \left(\frac{r}{\gamma}\right)^d$$

where $d = \log_3(N+1)$ and the notation $a \preceq b$ indicates that there is some constant $c(N) > 0$ depending only on N such that $a \leq c(N)b$. Thus, using the notation of Lemma 4.4, $w(s) \preceq \gamma s^{1/d}$. Since $Q_* = 1/|E_N(n)| \succeq 1/\gamma^d$, Lemma 4.4 gives the lower bound on the spectral profile.

For the upper bound we construct test functions f_m supported on m-blocks. Given an m-block $A \subset \mathcal{V}_N(n)$, choose vertices $x_1, \ldots, x_N$ such that $d(x_i, x_j) = \mathrm{diam}(A)$ for $i \neq j$, and call the shortest paths between these vertices diagonals. These diagonals meet in a unique point o at the center of the m-block. Define the

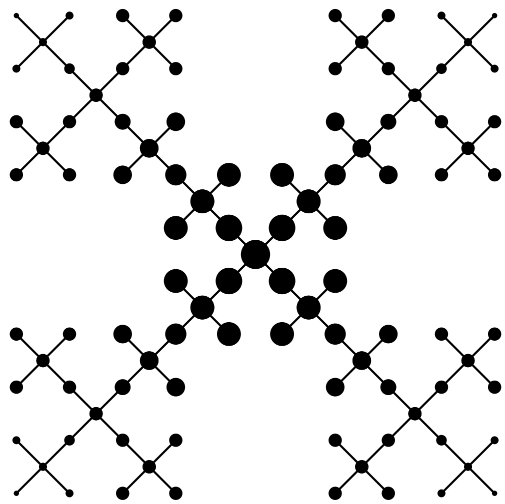

FIGURE 2. A graphical representation of the test function f_2 supported on a 2-block.

function $f_m \in c_0(A)$ as follows: Along diagonals, f_m varies linearly with $f_m(o) = 1$ and $f_m(x_i) = 0$. Since $d(o, x_i) = \mathrm{diam}(A)/2 = 3^m$, along diagonals the function is given explicitly by $f_m(x) = 1 - d(o,x)/3^m$. For a point x off of the diagonals, let $f_m(x) = f_m(x')$ where x' is the closest point to x that lies on a diagonal. (See Figure 2 for a graphical representation of f_m). Now, since $K(x,y)\pi(x) = 1/(2|E_N(n)|)$ for $x \sim y$

$$\begin{aligned}
\mathcal{E}(f_m, f_m) &= \frac{1}{2}\sum_{x,y} |f_m(x) - f_m(y)|^2 K(x,y)\pi(x) \\
&= 3^{-2m} \cdot \frac{N3^m}{2|E_N(n)|} \\
&\approx \frac{1}{\gamma^d 3^m}.
\end{aligned}$$

Define the central $m-1$ block of A to be $A' = \{x \in A : d(o,x) \leq 3^{m-1}\}$. Since $f_m(x) \geq 2/3$ on A',

$$\|f_m\|_2^2 \geq \frac{4}{9}\pi(A') \approx \frac{(N+1)^m}{\gamma^d}.$$

It is sufficient to prove the upper bound on $\Lambda(r)$ for $1/(N+1)^{n-2} < r \leq 1/2$. For these r take

$$m(r) = \left\lfloor \frac{\log r(N+1)^{n-2}}{\log N + 1} \right\rfloor \leq n.$$

Then $(N+1)^{m(r)} \leq r(N+1)^{n-2}$ and so for an $m(r)$-block K, $\pi(K) \leq r$. Consequently, for r in this range

$$\begin{aligned}\frac{\mathcal{E}(f_m, f_m)}{\mathrm{Var}(f_m)} &\leq \frac{2\mathcal{E}(f_m, f_m)}{\|f_m\|_2^2} \\ &\preceq \frac{1}{[3(N+1)]^m}\end{aligned}$$

Finally, since $(N+1)^m \succeq r\gamma^d$

$$\begin{aligned}\frac{\mathcal{E}(f_m, f_m)}{\mathrm{Var}(f_m)} &\preceq (N+1)^{-m\frac{\log 3(N+1)}{\log N+1}} \\ &= (N+1)^{-m(1+1/d)} \\ &\preceq \frac{1}{\gamma^{d+1} r^{1+1/d}}\end{aligned}$$

and the upper bound on $\Lambda(r)$ follows. □

4.4. **A delicate example.** Consider simple random walk on the product group $\mathbb{Z}_n \times \mathbb{Z}_{n^2}$. For this model, it is not hard to see that $\gamma = \Theta(n^2)$ and that the volume satisfies

$$V_*(r) \asymp \begin{cases} (r+1)^2/n^3 & 0 \leq r \leq n \\ r/n^2 & n \leq r \leq n^2 \end{cases}.$$

Taking $r = 0$ in (4.1) shows that walks with moderate growth must have

$$\frac{1}{n^3} \geq \frac{1}{A}\left(\frac{1}{n^2}\right)^d.$$

Consequently, $\mathbb{Z}_n \times \mathbb{Z}_{n^2}$ is of moderate growth with $d = 3/2$ and furthermore, this is the optimal choice of d (assuming A and d are constant). Corollary 4.2 gives the correct $\gamma^2 = n^4$ mixing time, but gives the underestimate

$$\Lambda(v) \geq \frac{C(a, A)}{\gamma^2 v^{4/3}}$$

for the spectral profile. The problem is that the moderate growth criterion alone is not sufficient to identify the two different scales of volume growth present in this example: For $r \ll 1/n$ the space appears 2-dimensional, while for $r \gg 1/n$ it looks 1-dimensional. However, we can apply Theorem 4.1 to directly take into account volume estimates, leading to sharp bounds on both the spectral profile and the rate of decay of $d_{\infty,\pi}(H_t, \pi)$.

Lemma 4.6. *For $2 \leq a \leq b$, the walk on $G = \mathbb{Z}_a \times \mathbb{Z}_b$ with generating set $\{(\pm 1, 0), (0, \pm 1)\}$ has spectral profile satisfying*

$$\Lambda(v) \asymp \begin{cases} 1/vab & 1/ab \leq v \leq a/b \\ 1/v^2 b^2 & a/b \leq v \leq 1 \\ 1/b^2 & 1 \leq v \end{cases}.$$

In particular,

$$d_{\infty,\pi}(H_t, \pi) \asymp \begin{cases} ab/(t+1) & 0 \leq t \leq a^2 \\ b/t^{1/2} & a^2 \leq t \leq b^2 \end{cases}$$

and there are constants $c_1, c_2 > 0$ such that for $t \geq b^2$

$$e^{-c_1 t/b^2} \preceq d_{\infty,\pi}(H_t, \pi) \preceq e^{-c_2 t/b^2}.$$

Proof. Observe that $\mathrm{diam}(G) = \Theta(b)$, and the volume function is given by

$$V_*(r) \asymp \begin{cases} (r+1)^2/(ab) & 0 \le r \le a \\ r/b & a \le r \le \mathrm{diam}(G) \end{cases}.$$

Consequently, $W(v) = \inf\{r : V_*(r) \ge v\}$ satisfies

$$W(v) \asymp \begin{cases} v^{1/2}(ab)^{1/2} & 1/ab \le v \le a/b \\ vb & a/b \le v \le 1 \end{cases}.$$

Since this group walk is driven by a constant number of generators, the chain satisifes a local Poincaré inequality with constant independent of a and b. For $v \le 1/2$, the lower bound on the $\Lambda(v)$ now follows from Theorem 4.1. By Lemma 2.2, $\lambda_1 \ge \Lambda(1/2)/2$. For $v \ge 1/2$, the lower bound on $\Lambda(r)$ then follows from the fact $\Lambda(r) \ge \lambda_1$.

For the upper bound, for $m \ge 1$ define linear functions $f_m \in c_0(B(0,m))$ by

$$f_m(x) = 1 - \frac{d_G(0,x)}{m}.$$

Then,

$$\begin{aligned} \mathcal{E}(f_m, f_m) &= \frac{1}{2|G|}\sum_{x,y}[f_m(x) - f_m(y)]^2 K(x,y) \\ &\le \frac{V_*(m)}{2m^2}. \end{aligned}$$

Since G is volume doubling (i.e. $V_*(2m) \approx V_*(m)$)

$$\|f_m\|_2^2 \ge \left(\frac{1}{2}\right)^2 V_*(m/2) \succeq V_*(m)$$

and consequently, for $V_*(m) \le 1/2$

$$\frac{\mathcal{E}(f_m, f_m)}{\mathrm{Var}(f_m)} \le 2\frac{\mathcal{E}(f_m, f_m)}{\|f_m\|_2^2} \preceq \frac{1}{m^2}.$$

Since it is sufficient to consider only $v \le 1/2$, the upper bound on $\Lambda(v)$ follows by taking $m = W(v) - 1$.

Following the notation of Theorem 3.1, define $\gamma(t)$ implicitly through

$$L(r) = \begin{cases} C/rab & 1/ab \le r \le a/b \\ C/r^2b^2 & a/b \le r \le 1 \end{cases}$$

and observe that for C sufficiently large, the walk satisfies the anti-Faber-Krahn inequality with $L(r)$ by the upper bound on $\Lambda(r)$. Then $\gamma(t)$ is given explicitly as

$$\gamma(t) = \begin{cases} (Ct+1)/ab & 0 \le t \le (a^2-1)/C \\ (2Ct - a^2 + 2)^{1/2}/b & (a^2-1)/C \le t \le (b^2 + a^2 - 2)/(2C) \end{cases}.$$

Furthermore,

$$\frac{\gamma'(t)}{\gamma(t)} = \begin{cases} C/(Ct+1) & 0 < t \le (a^2-1)/C \\ C/(2Ct - a^2 + 2) & (a^2-1)/C \le t < (b^2 + a^2 - 2)/(2C) \end{cases}.$$

So, $\gamma(t)$ is δ-regular on $(0,T)$ with $\delta = 1/6$ and $T = (b^2 + a^2 - 2)/(2C)$. For $t \le cb^2$ and c sufficiently small, the lower bound on convergence now follows from Theorem 3.1.

Let $\{\lambda_i\}$ be the eigenvalues of $I-K$ with corresponding real orthonormal eigenfunctions $\{\psi_i\}$. Since $h_t(x,y) = H_t 1_y(x)$, writing $1_y(\cdot)$ in this L^2 basis we have

$$h_t(x,y) = \sum_{i=0}^{n-1} e^{-t\lambda_i}\psi_i(x)\psi_i(y) = 1 + \sum_{i=1}^{n-1} e^{-t\lambda_i}\psi_i(x)\psi_i(y).$$

In particular,

$$\sup_x h_t(x,x) - 1 \geq \sup_x e^{-t\lambda_1}\psi_i^2(x) \geq e^{-t\lambda_1}$$

since the eigenfunctions are normalized in L^2. Since $\lambda_1 \approx \Lambda(1/2) \approx 1/b^2$, the lower bound on convergence rate follows.

Now we show the upper bound. By the lower bound on $\Lambda(r)$, using the notation of Theorem 2.1,

$$V(t) \succeq \begin{cases} t/ab & 0 \leq t \leq a^2 \\ t^{1/2}/b & a^2 \leq t \leq b^2 \\ e^{ct/b^2} & b^2 \leq t \end{cases}.$$

Consequently,

$$\sup_{x\in\mathcal{X}} d^2_{2,\pi}(H_t(x,\cdot),\pi) \preceq \begin{cases} ab/t & 0 \leq t \leq a^2 \\ b/t^{1/2} & a^2 \leq t \leq b^2 \\ e^{-ct/b^2} & b^2 \leq t \end{cases}.$$

Since the walk is reversible, the upper bound on $d_{\infty,\pi}$ then follows from the argument of Theorem 1.1 $\square$

For $a \ll b$, Lemma 4.6 captures the fact that decay is fast at the start of the walk and then slows down. More specifically,

$$d_{\infty,\pi}(H_t,\pi) \asymp 1/V_*(t^{1/2}) \qquad t \leq \mathrm{diam}^2(G).$$

As shown in [DSC95], this relationship holds in general for random walks on groups with volume doubling. While we considered only the simple case of $\mathbb{Z}_a \times \mathbb{Z}_b$, the technique applies well to more general k-fold products.

Acknowledgments

We thank Laurent Saloff-Coste for initiating this project by pointing out that the method of Faber-Krahn inequalities developed in [Gri94] also applies in the setting of finite Markov chains. His PhD student Sharad Goel thanks Prof. Saloff-Coste for his plentiful guidance, advice and encouragement during this work.

References

[AF] D. Aldous and J. Fill, *Reversible markov chains and random walks on graphs*, Monograph in preparation available on the web at http://stat-www.berkeley.edu/users/aldous/RWG/book.html.

[BCG01] Martin Barlow, Thierry Coulhon, and Alexander Grigor'yan, *Manifolds and graphs with slow heat kernel decay*, Invent. Math. **144** (2001), 609–649.

[BT03] Sergey G. Bobkov and Prasad Tetali, *Modified log-sobolev inequalities in discrete settings*, Proc. of the ACM STOC 2003, 2003, pp. 287–296.

[CGP01] T. Coulhon, A. Grigor'yan, and C. Pittet, *A geometric approach to on-diagonal heat kernel lower bound on groups*, Ann. Inst. Fourier **51** (2001), 1763–1827.

[CKS87] E. Carlen, S. Kusuoka, and D. Stroock, *Upper bounds for symmetric markov transition functions*, Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), 245–287.

[Cou96] T. Coulhon, *ultracontractivity and nash type inequalities*, J. Funct. Anal. **141** (1996), 510–539.

[CSC90a] Thierry Coulhon and Laurent Saloff-Coste, *Marches aléatoires non symétriques sur les groupes unimodulaires*, C. R. Acad. Sci. Paris Sér. I Math. **310** (1990), no. 8, 627–630.

[CSC90b] ———, *Puissances d'un op'erateur régularisant*, Ann. Inst. H. Poincaré Probab. Statist. **26** (1990), no. 3, 419–436.

[Dav90] E.B. Davies, *Heat kernels and spectral theory*, Cambridge University Press, 1990.

[DSC94] P. Diaconis and L. Saloff-Coste, *Moderate growth and random walk on finite groups*, Geom. and Funct. Anal. **4** (1994), 1–34.

[DSC95] Persi Diaconis and Laurent Saloff-Coste, *An application of harnack inequalities to random walk on nilpotent quotients*, Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993), J. Fourier Anal. Appl, Special Issue, 1995, pp. 189–207.

[DSC96a] P. Diaconis and L. Saloff-Coste, *Logarithmic sobolev inequalities for finite markov chains*, Ann. Appl. Prob. **6** (1996), 695–750.

[DSC96b] ———, *Nash inequalities for finite markov chains*, J. Theoret. Probab. **9** (1996), no. 2, 459–510.

[Gri94] A. Grigor'yan, *Heat kernel upper bounds on a complete non-compact manifold*, Revista Matemática Iberoamericana **10** (1994), 395–452.

[Gro75] L. Gross, *Logarithmic sobolev inequalities*, Amer. J. Math. **97** (1975), 1061–1083.

[Gro93] ———, *Logarithmic sobolev inequalities and contractivity properties of semigroups*, Lecture Notes in Mathematics, vol. 1563, Springer, 1993.

[LK99] L. Lovász and R. Kannan, *Faster mixing via average conductance*, Annual ACM Symposium on theory of computing (Atlanta, GA, 1999), ACM, 1999, pp. 282–287.

[LO00] R. Latala and K. Oleszkiewicz, *Between sobolev and poincaré*, Geometric Aspects of Functional Analysis, Lect. Notes in Math., vol. 1745, 2000, pp. 147–168.

[LS88] G.F. Lawler and A.D. Sokal, *Bounds on the L^2 spectrum for markov chains and markov processes: a generalization of cheeger's inequality*, Trans. Amer. Math. Soc. **309** (1988), 557–580.

[MP] Ben Morris and Yuval Peres, *Evolving sets, mixing and heat kernel bounds*, Preprint.

[Nas58] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math **80** (1958), 931–954.

[PSC] Ch. Pittet and L. Saloff-Coste, *Isoperimetry, volume growth and random walks*, Monograph in preparation.

[SC96] Laurent Saloff-Coste, *Lectures on finite markov chains*, Lectures on Probability Theory and Statistics (P. Bernard, ed.), Lecture Notes in Mathematics, vol. 1665, Ecole d'Eté de Probabiltés de Saint-Flour XXVI, Springer, 1996.

[SJ89] A.J. Sinclair and M. R. Jerrum, *Approximate counting, uniform generation and rapidly mixing markov chains*, Information and Computation **82** (1989), 93–133.