Evolving sets, mixing and heat kernel bounds

BEN MORRIS* YUVAL PERES†

October 27, 2003

Abstract

We show that a new probabilistic technique, recently introduced by the first author, yields the sharpest bounds obtained to date on mixing times of Markov chains in terms of isoperimetric properties of the state space (also known as conductance bounds or Cheeger inequalities). We prove that the bounds for mixing time in total variation obtained by Lovász and Kannan, can be refined to apply to the maximum relative deviation $|p^n(x,y)/\pi(y) - 1|$ of the distribution at time n from the stationary distribution π. We then extend our results to Markov chains on infinite state spaces and to continuous-time chains. Our approach yields a direct link between isoperimetric inequalities and heat kernel bounds; previously, this link rested on analytic estimates known as Nash inequalities.

1 Introduction

It is well known that the absence of "bottlenecks" in the state space of a Markov chain implies rapid mixing. Precise formulations of this principle, related to Cheeger's inequality in differential geometry, have been proved by algebraic and combinatorial techniques [1, 14, 12, 17, 9, 15]. They have been used to approximate permanents, to sample from the lattice points in a convex set, to estimate volumes, and to analyze a random walk on a percolation cluster in a box.

In this paper, we show that a new probabilistic technique, introduced in [19], yields the sharpest bounds obtained to date on mixing times in terms of bottlenecks.

Let $\{p(x,y)\}$ be transition probabilities for an irreducible Markov chain on a countable state space V, with stationary distribution π (i.e., $\sum_{x\in V}\pi(x)p(x,y) = \pi(y)$ for all $x \in V$). For $x, y \in V$, let $Q(x,y) = \pi(x)p(x,y)$, and for $S, A \subset V$, define $Q(S,A) = \sum_{s\in S, a\in A} Q(s,a)$. For $S \subset V$, the "boundary size" of S is measured by $|\partial S| = Q(S, S^c)$. Following [12], we call $\Phi_S := \frac{|\partial S|}{\pi(S)}$ the *conductance* of S. Write $\pi_* := \min_{x\in V}\pi(x)$ and define $\Phi(r)$ for $r \in [\pi_*, 1/2]$ by

$$\Phi(r) = \inf\left\{\Phi_S : \pi(S) \le r\right\} . \tag{1}$$

For $r > 1/2$, let $\Phi(r) = \Phi_* = \Phi(1/2)$. Define the ϵ-*uniform mixing time* by

$$\tau(\epsilon) = \min\Big\{n : \Big|\frac{p^n(x,y) - \pi(y)}{\pi(y)}\Big| \le \epsilon \ \ \forall\, x, y \in V\Big\}.$$

*Department of Statistics, Evans Hall, University of California, Berkeley CA 94720.
Email: `morris@stat.berkeley.edu`. Supported by NSF post-doctoral fellowship.

†Departments of Statistics and Mathematics, University of California, Berkeley.
Email: `peres@stat.berkeley.edu` Research supported by the Miller Foundation and by NSF Grant DMS-0104073.

Jerrum and Sinclair [12] considered chains that are *reversible* (i.e., $Q(x, y) = Q(y, x)$ for all $x, y \in V$) and also satisfy

$$p(x, x) \geq 1/2 \text{ for all } x \in V\,. \tag{2}$$

They estimated the second eigenvalue of $p(\cdot, \cdot)$ in terms of conductance, and derived the bound

$$\tau(\epsilon) \leq 2\Phi_*^{-2}\left(\log\frac{1}{\pi_*} + \log\frac{1}{\epsilon}\right)\,. \tag{3}$$

Algorithmic applications of (3) are described in [22]. Extensions of (3) to nonreversible chains were obtained by Mihail [17] and Fill [9]. A striking new idea was introduced by Lovász and Kannan [15], who realized that in geometric examples, small sets often have larger conductance, and discovered a way to exploit this. Let $\|\mu - \nu\| = \frac{1}{2}\sum_{y \in V}|\mu(y) - \nu(y)|$ be the total variation distance, and denote by

$$\tau_V(\epsilon) := \min\Big\{n : \|p^n(x, \cdot) - \pi\| \leq \epsilon \text{ for all } x \in V\Big\} \tag{4}$$

the ϵ-mixing time in total variation. (This can be considerably smaller than the uniform mixing time $\tau(\epsilon)$, see the lamplighter walk discussed at the end of this section, or §8, Remark 1.) For reversible chains that satisfy (2), Lovász and Kannan proved that

$$\tau_V(1/4) \leq 2000\int_{\pi_*}^{3/4}\frac{du}{u\Phi^2(u)}, \tag{5}$$

This formula was the impetus for the present paper. Related formulae for infinite Markov chains were obtained earlier from Nash inequalities and are discussed below. (As noted in [18], there was a small error in [15]; the statement above is obtained from §3 in the survey by Kannan [13].)

Note that in general, $\tau_V(\epsilon) \leq \tau_V(1/4)\log_2(1/\epsilon)$. Therefore, ignoring constant factors, the bound in (5) is tighter than the bound of (3), but at the cost of employing a weaker notion of mixing.

Our main result sharpens (5) to a bound on the uniform mixing time. See Theorem 5 for a version that relaxes the assumption (2). We use the notation $\alpha \wedge \beta := \min\{\alpha, \beta\}$.

Theorem 1 *Assume (2). Then the ϵ-uniform mixing time satisfies*

$$\tau(\epsilon) \leq \int_{\pi_*}^{4/\epsilon}\frac{4du}{u\Phi^2(u)}\,. \tag{6}$$

More precisely, if

$$n \geq \int_{\pi(x)\wedge\pi(y)}^{4/\epsilon}\frac{4du}{u\Phi^2(u)}\,, \tag{7}$$

then

$$\left|\frac{p^n(x, y) - \pi(y)}{\pi(y)}\right| \leq \epsilon. \tag{8}$$

(Recall that $\Phi(r)$ is constant for $r \geq \frac{1}{2}$.) This result has several advantages over (5):

- The uniformity in (6).
- It yields a better bound when the approximation parameter ϵ is small.
- It applies to nonreversible chains.

- It yields an improvement of the upper bound on the time to achieve (8) when $\pi(x), \pi(y)$ are larger than π_*.

- The improved constant factors make the bound (6) potentially applicable as a stopping time in simulations. Under a convexity condition, these factors can be improved further; see §8, Remark 3.

Other ways to measure bottlenecks can yield sharper bounds. One approach, based on "blocking conductance functions" and restricted to the mixing time in total variation τ_V, is presented in [13, Theorem 3].

Another boundary gauge ψ is defined in §2 of the present paper. For the n-dimensional unit hypercube, this gauge (applied to the right class of sets, see §8) gives a bound of the right order $\tau(1/e) = O(n \log n)$ for the uniform mixing time. Previous methods of measuring bottlenecks did not yield the right order of magnitude for the uniform mixing time in this benchmark example.

Theorem 1 is related to another line of research, namely the derivation of heat kernel estimates for Markov chains using Nash and Sobolev inequalities. For finite Markov chains, such estimates were obtained by Chung and Yau [5], and by Diaconis and Saloff-Coste [8]. In particular, for the special case where Φ is a power law, the conclusion of Theorem 1 can be obtained by combining Theorems 2.3.1 and 3.3.11 of Saloff Coste [21]. For infinite Markov chains, Nash inequalities have been developed for general isoperimetric profiles; see Varopoulos [23], the survey by Pittet and Saloff Coste [20], the book [24], the papers [6, 7] and the references therein. Even in this highly developed subject, our probabilistic technique yields improved estimates when the stationary measure is not uniform; Suppose that π is an infinite stationary measure on V for the transition kernel p. As before, we define

$$Q(x,y) = \pi(x)p(x,y); \qquad |\partial S| = Q(S, S^c); \qquad \Phi_S := \frac{|\partial S|}{\pi(S)}.$$

Define $\Phi(r)$ for $r \in [\pi_*, \infty)$ by

$$\Phi(r) = \inf\left\{\Phi_S : \pi(S) \le r\right\}. \tag{9}$$

Theorem 2 (infinite stationary measure case)

Suppose that $0 < \gamma \le \frac{1}{2}$ *and* $p(x,x) \ge \gamma$ *for all* $x \in V$. *If*

$$n \ge \frac{(1-\gamma)^2}{\gamma^2} \int_{\pi(x)\wedge\pi(y)}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}, \tag{10}$$

then

$$\left|\frac{p^n(x,y)}{\pi(y)}\right| \le \epsilon. \tag{11}$$

This Theorem is proved in Section 6. For the rest of the introduction, we focus on the case of finite stationary measure.

Definition: Evolving sets. Given V, π and Q as above, consider the Markov chain $\{S_n\}$ on *subsets* of V with the following transition rule. If the current state S_n is $S \subset V$, choose U uniformly from $[0,1]$ and let the next state S_{n+1} be

$$\widetilde{S} = \left\{y : Q(S,y) \ge U\pi(y)\right\}.$$

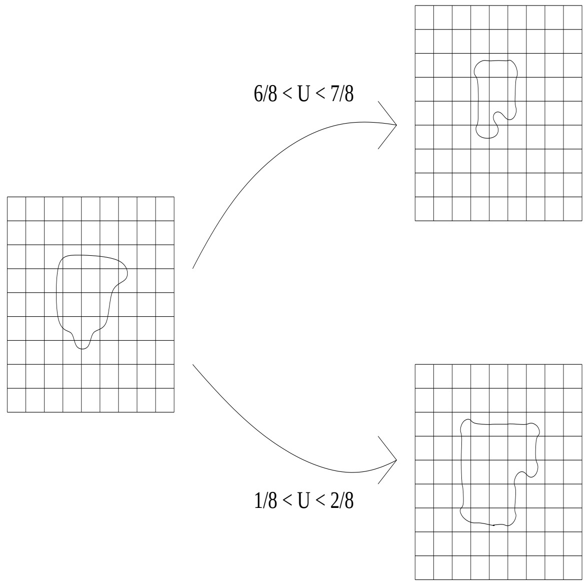

Figure 1: One step of the evolving set process.

Consequently,

$$\mathbf{P}(y \in \widetilde{S}) = \mathbf{P}\Big(Q(S,y) \geq U\pi(y)\Big) = \frac{Q(S,y)}{\pi(y)}\,. \tag{12}$$

Figure 1 illustrates one step of the evolving set process when the original Markov chain is a random walk in a box (with a holding probability of $\frac{1}{2}$). Since π is the stationary distribution, $\emptyset$ and V are absorbing states for the evolving set process.

Write $\mathbf{P}_S\Big(\,\cdot\,\Big) := \mathbf{P}\Big(\,\cdot\,\Big|\,S_0 = S\Big)$ and similarly for $\mathbf{E}_S\Big(\,\cdot\,\Big)$. The utility of evolving sets stems from the relation

$$p^n(x,y) = \frac{\pi(y)}{\pi(x)}\,\mathbf{P}_{\{x\}}(y \in S_n)$$

(see Proposition 7). Their connection to mixing is indicated by the inequality

$$\|\mu_n - \pi\| \leq \frac{1}{\pi(x)}\mathbf{E}_{\{x\}}\sqrt{\pi(S_n) \wedge \pi(S_n^c)}\,,$$

where $\mu_n := p^n(x,\cdot)$; see (20) for a sharper form of this. The connection of evolving sets to conductance can be seen in Lemma 3 below.

Example (Random Walk in a Box): Consider a simple random walk in an $n \times n$ box. To guarantee condition (2) we add a holding probability of $\frac{1}{2}$ to each state (i.e., with probability $\frac{1}{2}$ do nothing, else move as above). The conductance profile satisfies

$$\Phi(u) \geq \frac{a}{n\sqrt{u}}$$

for $1 \leq u \leq 1/2$, where a is a constant. Thus our bound implies that the ϵ uniform mixing time is at most

$$C_\epsilon + 4\int_{1/n^2}^{1/2} \frac{1}{u\left(\frac{a}{n\sqrt{u}}\right)^2}\,du = O(n^2),$$

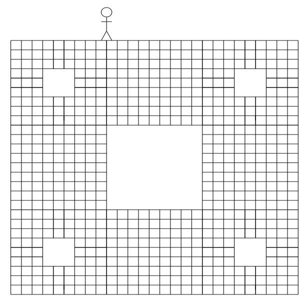

Figure 2: A box with holes.

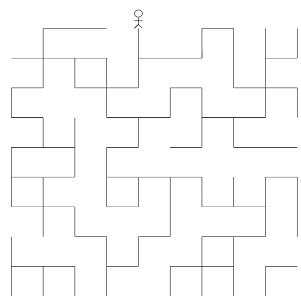

Figure 3: Random walk in a percolation cluster.

which is the correct order of magnitude. Of course, other techniques such as coupling or spectral methods would give the correct-order bound of $O(n^2)$ in this case. However, these techniques are not robust under small perturbations of the problem, whereas the conductance method is.

Example (Box with Holes): For a random walk in a box with holes (see Figure 1), it is considerably harder to apply coupling or spectral methods. However, it is clear that the conductance profile for the random walk is unchanged (up a constant factor), and hence the mixing time is still $O(n^2)$.

Example (Random Walk in a Percolation Cluster): In fact, the conductance method is robust enough to handle an even more extreme variant: Suppose that each edge in the box is deleted with probability $1-p$, where $p > \frac{1}{2}$. Then with high probability there is a connected component that contains a constant fraction of the original edges. Benjamini and Mossel [3] showed that for the random walk in the big component the conductance profile is sufficiently close (with high probability) to that of the box and deduced that the mixing time is still $O(n^2)$. (See [16] for analogous results in higher dimensions.) By our result, this also applies to the uniform mixing times.

Example (Random Walk on a Lamplighter Group): The following natural chain mixes more rapidly in the sense of total variation than in the uniform sense. A state of this chain consists of n lamps arrayed in a circle, each lamp either *on* (1) or *off* (0), and a *lamplighter* located next to one of the lamps. In one "active" step of the chain, the lamplighter either switches the current lamp or moves at random to one of the two adjacent lamps. We consider the lazy chain that stays put with probability 1/2 and makes and active step with probability 1/2. The path of the lamplighter is a delayed simple random walk on a cycle, and this implies that $\tau_V(1/4) = \Theta(n^2)$,

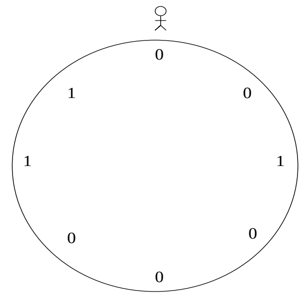

Figure 4: Random walk on a lamplighter group

see [10]. However, by considering the possibility that the lamplighter stays in one half of the cycle for a long time, one easily verifies that $\tau(1/4) \geq c_1 n^3$ for some constant $c_1 > 0$. Using the general estimate $\tau(\epsilon) = O(\tau_V(\epsilon)\log(1/\pi_*))$ gives a matching upper bound $\tau(1/4) = O(n^3)$.

2 Further results and proof of Theorem 1

We will actually prove a stronger form of Theorem 1, using the boundary gauge

$$\psi(S) := 1 - \mathbf{E}_S \sqrt{\frac{\pi(\widetilde{S})}{\pi(S)}}$$

instead of the conductance Φ_S. The next lemma relates these quantities.

Lemma 3 *Let $\emptyset \neq S \subset V$. If (2) holds, then $\psi(S) \geq \Phi_S^2/2$. More generally, if $0 < \gamma \leq \frac{1}{2}$ and $p(x,x) \geq \gamma$ for all $x \in V$, then $\psi(S) \geq \frac{\gamma^2}{2(1-\gamma)^2}\Phi_S^2$.*

See §4 for the proof. In fact, $\psi(S)$ is often much larger than Φ_S^2.

Define $\psi(r)$ for $r \in [\pi_*, 1/2]$ by

$$\psi(r) = \inf\{\psi(S) : \pi(S) \leq r\}, \tag{13}$$

and for $r > 1/2$, let $\psi(r) := \psi_* = \psi(\frac{1}{2})$. Observe that ψ is (weakly) decreasing on $[\pi_*, \infty)$.

For a measure μ on V, write

$$\chi^2(\mu, \pi) := \sum_{y \in V} \pi(y)\Big(\frac{\mu(y)}{\pi(y)} - 1\Big)^2 = \Big(\sum_{y \in V} \frac{\mu(y)^2}{\pi(y)}\Big) - 1\,. \tag{14}$$

By Cauchy-Schwarz,

$$2\|\mu - \pi\| = \Big\|\frac{\mu(\cdot)}{\pi(\cdot)} - 1\Big\|_{L^1(\pi)} \leq \Big\|\frac{\mu(\cdot)}{\pi(\cdot)} - 1\Big\|_{L^2(\pi)} = \chi(\mu, \pi)\,.$$

We can now state our key result relating evolving sets to mixing.

Theorem 4 *Denote $\mu_n = p^n(x, \cdot)$. Then $\chi^2(\mu_n, \pi) \leq \epsilon$ for all*

$$n \geq \int_{\pi(x)}^{4/\epsilon} \frac{du}{u\psi(u)}\,.$$

See §5 for the proof.

Derivation of Theorem 1 from Lemma 3 and Theorem 4:

The *time-reversal* of a Markov chain on V with stationary distribution π and transition matrix $p(x,y)$, is another Markov chain with stationary distribution π, and transition matrix $\overleftarrow{p}(\cdot,\cdot)$ that satisfies $\pi(y)p(y,z) = \pi(z)\overleftarrow{p}(z,y)$ for all $y,z \in V$. Summing over intermediate states gives $\pi(z)\overleftarrow{p}^{\,m}(z,y) = \pi(y)p^m(y,z)$ for all $z,y \in V$ and $m \geq 1$.

Since $p^{n+m}(x,z) = \sum_{y\in V} p^n(x,y)p^m(y,z)$, stationarity of π gives

$$p^{n+m}(x,z) - \pi(z) = \sum_{y\in V} \Big(p^n(x,y) - \pi(y)\Big)\Big(p^m(y,z) - \pi(z)\Big) \tag{15}$$

whence

$$\begin{aligned} & \Big|\frac{p^{n+m}(x,z) - \pi(z)}{\pi(z)}\Big| && (16)\\ = \;& \Big|\sum_{y\in V} \pi(y)\Big(\frac{p^n(x,y)}{\pi(y)} - 1\Big)\Big(\frac{\overleftarrow{p}^{\,m}(z,y)}{\pi(y)} - 1\Big)\Big| && (17)\\ \leq \;& \chi\Big(p^n(x,\cdot),\pi\Big)\chi\Big(\overleftarrow{p}^{\,m}(z,\cdot),\pi\Big) && (18)\end{aligned}$$

by Cauchy-Schwarz.

The quantity $Q(S,S^c)$ represents, for any $S \subset V$, the asymptotic frequency of transitions from S to S^c in the stationary Markov chain with transition matrix $p(\cdot,\cdot)$. In particular, stationarity of π implies that $Q(S,S^c) = Q(S^c,S)$; reversibility is not needed for this. It follows that the time-reversed chain has the same conductance profile $\Phi(\cdot)$ as the original Markov chain. Hence, Lemma 3 and Theorem 4 imply that if

$$m,n \geq \int_{\pi(x)\wedge\pi(y)}^{\frac{1}{2}} \frac{2du}{u\Phi^2(u)} + \frac{2}{\Phi_*^2}\log\frac{8}{\epsilon},$$

and (2) holds, then

$$\chi\Big(p^n(x,\cdot),\pi\Big) \leq \sqrt{\epsilon} \;\text{ and }\; \chi\Big(\overleftarrow{p}^{\,m}(z,\cdot),\pi\Big) \leq \sqrt{\epsilon}\,.$$

Thus by (18),

$$\Big|\frac{p^{n+m}(x,z) - \pi(z)}{\pi(z)}\Big| \leq \epsilon\,,$$

and Theorem 1 is established.

In fact, the argument above yields the following more general statement.

Theorem 5 *Suppose that $0 < \gamma \leq \frac{1}{2}$ and $p(x,x) \geq \gamma$ for all $x \in V$. If*

$$n \geq \frac{(1-\gamma)^2}{\gamma^2}\int_{\pi(x)\wedge\pi(y)}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}\,, \tag{19}$$

then (8) holds.

To complete the proof of Theorems 1 and 5, it suffices to prove Lemma 3 and Theorem 4. This is done in §4 and §5, respectively.

3 Properties of Evolving Sets

Lemma 6 *The sequence* $\{\pi(S_n)\}_{n\geq 0}$ *forms a martingale.*

Proof: By (12), we have

$$\begin{aligned}\mathbf{E}\left(\pi(S_{n+1})\Big|S_n\right) &= \sum_{y\in V}\pi(y)\,\mathbf{P}\left(y\in S_{n+1}\Big|S_n\right)\\ &= \sum_{y\in V} Q(S_n,y) \;=\; \pi(S_n)\,.\end{aligned}$$

□

The following proposition relates the nth order transition probabilities of the original chain to the evolving set process.

Proposition 7 *For all* $n\geq 0$ *and* $x,y\in V$ *we have*

$$p^n(x,y)=\frac{\pi(y)}{\pi(x)}\,\mathbf{P}_{\{x\}}(y\in S_n)\,.$$

Proof: The proof is by induction on n. The case $n=0$ is trivial. Fix $n>0$ and suppose that the result holds for $n-1$. Let U be the uniform random variable used to generate S_n from S_{n-1}. Then

$$\begin{aligned}p^n(x,y) &= \sum_{z\in V}p^{n-1}(x,z)p(z,y)\\ &= \sum_{z\in V}\mathbf{P}_{\{x\}}(z\in S_{n-1})\,\frac{\pi(z)}{\pi(x)}\,p(z,y)\\ &= \frac{\pi(y)}{\pi(x)}\mathbf{E}_{\{x\}}\left(\frac{1}{\pi(y)}Q(S_{n-1},y)\right)\\ &= \frac{\pi(y)}{\pi(x)}\,\mathbf{P}_{\{x\}}(y\in S_n).\end{aligned}$$

□

We will also use the following duality property of evolving sets.

Lemma 8 *Suppose that* $\{S_n\}_{n\geq 0}$ *is an evolving set process. Then the sequence of complements* $\{S_n^c\}_{n\geq 0}$ *is also an evolving set process, with the same transition probabilities.*

Proof: Fix n and let U be the uniform random variable used to generate S_{n+1} from S_n. Note that $Q(S_n,y)+Q(S_n^c,y)=Q(V,y)=\pi(y)$. Therefore, with probability 1,

$$\begin{aligned}S_{n+1}^c &= \Big\{y: Q(S_n,y)<U\pi(y)\Big\}\\ &= \Big\{y: Q(S_n^c,y)\geq(1-U)\pi(y)\Big\}.\end{aligned}$$

Thus, $\{S_n^c\}$ has the same transition probabilities as $\{S_n\}$, since $1-U$ is uniform. □

Next, we write the χ^2 distance between $\mu_n := p^n(x,\cdot)$ and π in terms of evolving sets. Let $\{S_n\}_{n\geq 0}$ and $\{\Lambda_n\}_{n\geq 0}$ be two independent replicas of the evolving set process, with $S_0 = \Lambda_0 = \{x\}$. Then by (14) and Proposition 7, $\chi^2(\mu_n,\pi)$ equals

$$\begin{aligned} & \sum_{y\in V} \pi(y)\frac{\mathbf{P}(y\in S_n)^2}{\pi(x)^2} - 1 \\ = \quad & \frac{1}{\pi(x)^2}\Big[\sum_{y\in V}\pi(y)\mathbf{P}\Big(\{y\in S_n\}\cap\{y\in\Lambda_n\}\Big) - \pi(x)^2\Big] \\ = \quad & \frac{1}{\pi(x)^2}\mathbf{E}\Big(\pi(S_n\cap\Lambda_n) - \pi(S_n)\pi(\Lambda_n)\Big)\,, \end{aligned}$$

where the last equation uses the relation $\pi(x) = \mathbf{E}\pi(S_n) = \mathbf{E}\pi(\Lambda_n)$. For any two sets $S,\Lambda\subset V$,

$$|\pi(S\cap\Lambda) - \pi(S)\pi(\Lambda)| = |\pi(S^c\cap\Lambda) - \pi(S^c)\pi(\Lambda)|.$$

Similarly, this expression doesn't change if we replace Λ by Λ^c. Thus, if we denote

$$S^\sharp := \begin{cases} S & \text{if } \pi(S)\leq \frac{1}{2}; \\ S^c & \text{otherwise,} \end{cases}$$

then

$$\begin{aligned} |\pi(S\cap\Lambda) - \pi(S)\pi(\Lambda)| &= |\pi(S^\sharp\cap\Lambda^\sharp) - \pi(S^\sharp)\pi(\Lambda^\sharp)| \\ &\leq |\pi(S^\sharp)\wedge\pi(\Lambda^\sharp)| \\ &\leq \sqrt{\pi(S^\sharp)\pi(\Lambda^\sharp)}\,. \end{aligned}$$

Inserting this into (20), we obtain

$$\chi^2(\mu_n,\pi) \leq \frac{1}{\pi(x)^2}\mathbf{E}\sqrt{\pi(S_n^\sharp)\pi(\Lambda_n^\sharp)}\,,$$

whence

$$2\|\mu_n - \pi\| \leq \chi(\mu_n,\pi) \leq \frac{1}{\pi(x)}\mathbf{E}\sqrt{\pi(S_n^\sharp)}\,. \tag{20}$$

4 Evolving sets and conductance profile: proof of Lemma 2

Lemma 9 *For every real number $u\leq\frac{1}{2}$, we have*

$$\frac{\sqrt{1+2u}+\sqrt{1-2u}}{2} \leq \sqrt{1-u^2} \leq 1 - u^2/2.$$

Proof: Squaring gives the second inequality and converts the first inequality into

$$1 + 2u + 1 - 2u + 2\sqrt{1-4u^2} \leq 4(1-u^2)$$

or equivalently, after halving both sides,

$$\sqrt{1-4u^2} \leq 1 - 2u^2\,,$$

which is verified by squaring again. $\square$

Lemma 10 *Let*

$$\varphi_S := \frac{1}{2\pi(S)} \sum_{y \in V} \Big(Q(S,y) \wedge Q(S^c,y) \Big) . \tag{21}$$

Then

$$1 - \psi(S) \le \frac{\sqrt{1+2\varphi_S} + \sqrt{1-2\varphi_S}}{2} \le 1 - \varphi_S^2/2 . \tag{22}$$

Proof: The second inequality in (22) follows immediately from Lemma 9. To see the first inequality, let U be the uniform random variable used to generate $\widetilde{S}$ from S. Then

$$\mathbf{P}_S\Big(y \in \widetilde{S} \,\Big|\, U < \tfrac{1}{2} \Big) = 1 \wedge \frac{2Q(S,y)}{\pi(y)} .$$

Consequently,

$$\pi(y) \mathbf{P}_S(y \in \widetilde{S} \,|\, U < \tfrac{1}{2}) = Q(S,y) + \Big(Q(S^c,y) \wedge Q(S,y) \Big) .$$

Summing over $y \in V$, we infer that

$$\mathbf{E}_S\Big(\pi(\widetilde{S}) \,\Big|\, U < \tfrac{1}{2} \Big) = \pi(S) + 2\pi(S)\varphi_S . \tag{23}$$

Therefore, $R := \pi(\widetilde{S})/\pi(S)$ satisfies $\mathbf{E}_S(R|U < \frac{1}{2}) = 1 + 2\varphi_S$. Since $\mathbf{E}_S R = 1$, it follows that

$$\mathbf{E}_S(R \,|\, U \ge \tfrac{1}{2}) = 1 - 2\varphi_S .$$

Thus

$$\begin{aligned} 1 - \psi(S) &= \mathbf{E}(\sqrt{R}) \\ &= \frac{\mathbf{E}(\sqrt{R}|U < \frac{1}{2}) + \mathbf{E}(\sqrt{R}|U \ge \frac{1}{2})}{2} \\ &\le \frac{\sqrt{\mathbf{E}(R|U < \frac{1}{2})} + \sqrt{\mathbf{E}(R|U \ge \frac{1}{2})}}{2} , \end{aligned}$$

by Jensen's inequality (or by Cauchy-Schwarz). This completes the proof. □

Proof of Lemma 3: If $p(y,y) \ge 1/2 \; \forall y \in V$, then it is easy to check directly that $\varphi_S = \Phi_S$ for all $S \subset V$.

If we are only given that $p(y,y) \ge \gamma \; \forall y \in V$, where $0 < \gamma \le \frac{1}{2}$, we can still conclude that for $y \in S$,

$$Q(S,y) \wedge Q(S^c,y) \ge \gamma\pi(y) \wedge Q(S^c,y) \ge \frac{\gamma}{1-\gamma} Q(S^c,y) .$$

Similarly, for $y \in S^c$ we have $Q(S,y) \wedge Q(S^c,y) \ge \frac{\gamma}{1-\gamma} Q(S,y)$. Therefore

$$\sum_{y \in V} [Q(S,y) \wedge Q(S^c,y)] \ge \frac{2\gamma}{1-\gamma} Q(S,S^c) ,$$

whence $\varphi_S \ge \frac{\gamma}{1-\gamma}\Phi_S$. This inequality, in conjunction with Lemma 10, yields Lemma 3.

5 Proof of Theorem 3

Denote by $K(S, A) = \mathbf{P}_S(\widetilde{S} = A)$ the transition kernel for the evolving set process. In this section we will use another Markov chain on sets with transition kernel

$$\widehat{K}(S, A) = \frac{\pi(A)}{\pi(S)} K(S, A). \tag{24}$$

The martingale property of the evolving set process, Lemma 6, implies that $\sum_A \widehat{K}(S, A) = 1$ for all $S \subset V$. The chain with kernel (24) represents the evolving set process conditioned to absorb in V; we will not use this fact explicitly.

Note that induction from equation (24) gives

$$\widehat{K}^n(S, A) = \frac{\pi(A)}{\pi(S)} K^n(S, A)$$

for every n, since

$$\begin{aligned} \widehat{K}^{n+1}(S, B) &= \sum_A \widehat{K}^n(S, A) \widehat{K}(A, B) \\ &= \sum_A \frac{\pi(B)}{\pi(S)} K^n(S, A) K(A, B) \\ &= \frac{\pi(B)}{\pi(S)} K^{n+1}(S, B) \end{aligned}$$

for every n and $B \subset V$. Therefore, for any function f

$$\widehat{\mathbf{E}}_S f(S_n) = \mathbf{E}_S \left[\frac{\pi(S_n)}{\pi(S)} f(S_n) \right], \tag{25}$$

for every function f, where we write $\widehat{\mathbf{E}}$ for the expectation when $\{S_n\}$ has transition kernel $\widehat{K}$. Define

$$Z_n = \frac{\sqrt{\pi(S_n^\sharp)}}{\pi(S_n)},$$

and note that $\pi(S_n) = Z_n^{-2}$ when $Z_n \geq \sqrt{2}$, that is, when $\pi(S_n) \leq \frac{1}{2}$. Then by equations (25) and (20), $\chi(\mu_n, \pi) \leq \widehat{\mathbf{E}}_{\{x\}}(Z_n)$ and

$$\begin{aligned} \widehat{\mathbf{E}} \left(\frac{Z_{n+1}}{Z_n} \Big| S_n \right) &= \mathbf{E} \left(\frac{\pi(S_{n+1})}{\pi(S_n)} \cdot \frac{Z_{n+1}}{Z_n} \Big| S_n \right) \\ &= \mathbf{E} \left(\frac{\sqrt{\pi(S_{n+1}^\sharp)}}{\sqrt{\pi(S_n^\sharp)}} \Big| S_n \right) & (26) \\ &\leq 1 - \psi(\pi(S_n)) = 1 - f_0(Z_n), & (27) \end{aligned}$$

where

$$f_0(z) := \begin{cases} \psi(1/z^2) & \text{if } z \geq \sqrt{2}; \\ \psi_* & \text{otherwise} \end{cases} \tag{28}$$

is increasing. Let $L_0 = Z_0 = \pi(x)^{-1/2}$. Next, observe that $\widehat{\mathbf{E}}(\cdot)$ is just the expectation operator with respect to a modified distribution, so we can apply Lemma 11 below, with $\widehat{\mathbf{E}}$ in place of $\mathbf{E}$. By part (iii) of that lemma, for all

$$n \geq \int_{\delta}^{L_0} \frac{2dz}{zf_0(z/2)} = \int_{\delta}^{L_0} \frac{2dz}{z\psi(4/z^2)}, \tag{29}$$

we have $\chi(\mu_n, \pi) \leq \widehat{\mathbf{E}}_{\{x\}}(Z_n) \leq \delta$. The change of variable $u = 4/z^2$ shows the integral (29) equals

$$\int_{4\pi(x)}^{4/\delta^2} \frac{du}{u\psi(u)} \leq \int_{\pi(x)}^{\frac{1}{2}} \frac{du}{u\psi(u)} + \frac{1}{\psi_*} \log \frac{8}{\delta^2}.$$

This establishes Theorem 4.

Lemma 11 *Let $f, f_0 : [0, \infty) \to [0, 1]$ be increasing functions. Suppose that $Z_n \geq 0$ for $n = 0, 1, 2, \ldots$ are random variables with $Z_0 = L_0$. Denote $L_n = \mathbf{E}(Z_n)$.*

(i) *If $L_n - L_{n+1} \geq L_n f(L_n)$ for all n, then for every $n \geq \int_{\delta}^{L_0} \frac{dz}{zf(z)}$, we have $L_n \leq \delta$.*

(ii) *If $\mathbf{E}(Z_{n+1}|Z_n) \leq Z_n(1 - f(Z_n))$ for all n and the function $u \mapsto uf(u)$ is convex on $(0, \infty)$, then the conclusion of (i) holds.*

(iii) *If $\mathbf{E}(Z_{n+1}|Z_n) \leq Z_n(1 - f_0(Z_n))$ for all n and $f(z) = f_0(z/2)/2$, then the conclusion of (i) holds.*

Proof: **(i)** It suffices to show that for every n we have

$$n \leq \int_{L_n}^{L_0} \frac{dz}{z\,f(z)}. \tag{30}$$

We verify this by induction. Clearly (30) holds for $n = 0$. Now fix $n \geq 0$ and suppose that (30) holds. Then

$$L_{n+1} \leq L_n\Big[1 - f(L_n)\Big] \leq L_n e^{-f(L_n)},$$

whence

$$\int_{L_{n+1}}^{L_n} \frac{dz}{zf(z)} \geq \frac{1}{f(L_n)} \int_{L_{n+1}}^{L_n} \frac{dz}{z} = \frac{1}{f(L_n)} \log \frac{L_n}{L_{n+1}} \geq 1.$$

Thus,

$$\int_{L_{n+1}}^{L_0} \frac{dz}{z\,f(z)} \geq \int_{L_n}^{L_0} \frac{dz}{z\,f(z)} + 1 \geq n + 1.$$

(ii) This is immediate from Jensen's inequality.
(iii) Fix $n \geq 0$. We have

$$\mathbf{E}\left(Z_n - Z_{n+1}\right) \geq \mathbf{E}\left[2Z_n f(2Z_n)\right] \geq L_n f(L_n), \tag{31}$$

by Lemma 12 below. This yields the hypothesis of (i). □

The following simple fact was used in the proof of Lemma 11.

Lemma 12 *Suppose that $Z \geq 0$ is a nonnegative random variable and f is a nonnegative increasing function. Then*

$$\mathbf{E}\Big(Zf(2Z)\Big) \geq \frac{\mathbf{E}Z}{2} \cdot f(\mathbf{E}Z).$$

Proof: Let A be the event $\{Z \geq \mathbf{E}Z/2\}$. Then $\mathbf{E}(Z\mathbf{1}_{A^c}) \leq \mathbf{E}Z/2$, so $\mathbf{E}(Z\mathbf{1}_A) \geq \mathbf{E}Z/2$. Therefore,

$$\mathbf{E}\Big(Zf(2Z)\Big) \geq \mathbf{E}\Big(Z\mathbf{1}_A \cdot f(\mathbf{E}Z)\Big) \geq \frac{\mathbf{E}Z}{2} f(\mathbf{E}Z)\,.$$

$\square$

6 Infinite stationary measures: proof of Theorem 2

Proof: For a probability measure μ on V, define $\chi^2(\mu, \pi)$ by

$$\chi^2(\mu, \pi) := \sum_{y \in V} \pi(y) \Big(\frac{\mu(y)}{\pi(y)}\Big)^2 = \sum_{y \in V} \frac{\mu(y)^2}{\pi(y)}\,. \tag{32}$$

We now write $\chi^2(\mu_n, \pi)$ in terms of evolving sets. Let $\{S_n\}_{n \geq 0}$ and $\{\Lambda_n\}_{n \geq 0}$ be two independent replicas of the evolving set process, with $S_0 = \Lambda_0 = \{x\}$. Then by (32) and Proposition 7,

$$\begin{aligned}
\chi^2(\mu_n, \pi) &= \sum_{y \in V} \pi(y) \frac{\mathbf{P}(y \in S_n)^2}{\pi(x)^2} && (33)\\
&= \frac{1}{\pi(x)^2} \Big[\sum_{y \in V} \pi(y) \mathbf{P}\Big(\{y \in S_n\} \cap \{y \in \Lambda_n\}\Big)\Big] && (34)\\
&= \frac{1}{\pi(x)^2} \mathbf{E}\Big(\pi(S_n \cap \Lambda_n)\Big) \leq \frac{1}{\pi(x)^2} \mathbf{E}\left(\sqrt{\pi(S_n)\pi(\Lambda_n)}\right)\,. && (35)
\end{aligned}$$

whence

$$\chi(\mu_n, \pi) \leq \frac{1}{\pi(x)} \mathbf{E}\sqrt{\pi(S_n)}\,. \tag{36}$$

As in the finite case, if $\widehat{K}$ is the Doob transform of K with respect to π, then

$$\widehat{\mathbf{E}}_S f(S_n) = \mathbf{E}_S \left[\frac{\pi(S_n)}{\pi(S)} f(S_n)\right]. \tag{37}$$

Define

$$Z_n = \frac{1}{\sqrt{\pi(S_n)}}.$$

Then by equations (37) and (36), $\chi(\mu_n, \pi) \le \widehat{\mathbf{E}}_{\{x\}}(Z_n)$ and

$$\begin{aligned}
\widehat{\mathbf{E}}\left(\frac{Z_{n+1}}{Z_n}\Big|\, S_n\right) &= \mathbf{E}\left(\frac{\pi(S_{n+1})}{\pi(S_n)}\cdot\frac{Z_{n+1}}{Z_n}\Big|\, S_n\right) \\
&= \mathbf{E}\left(\frac{\sqrt{\pi(S_{n+1})}}{\sqrt{\pi(S_n)}}\Big|\, S_n\right) \qquad (38)\\
&\le 1-\psi(\pi(S_n)) = 1 - f_0(Z_n),
\end{aligned}$$

where $f_0(z) = \psi(1/z^2)$ is increasing. Let $L_0 = Z_0 = \pi(x)^{-1/2}$. By Lemma 11(iii) above, for all

$$n \ge \int_{\sqrt{\epsilon}}^{L_0} \frac{2dz}{zf_0(z/2)} = \int_{\sqrt{\epsilon}}^{L_0} \frac{2dz}{z\psi(4/z^2)}\,, \qquad (39)$$

we have $\chi(\mu_n, \pi) \le \widehat{\mathbf{E}}_{\{x\}}(Z_n) \le \sqrt{\epsilon}$. The change of variable $u = 4/z^2$ shows the integral (39) equals

$$\int_{4\pi(x)}^{4/\epsilon} \frac{du}{u\psi(u)} \le \frac{(1-\gamma)^2}{\gamma^2}\int_{\pi(x)}^{4/\epsilon}\frac{2du}{u\Phi(u)}.$$

Let $\overleftarrow{p}$ denote the time-reversal of $p(\cdot,\cdot)$. Then for all

$$m, n \ge \frac{(1-\gamma)^2}{\gamma^2}\int_{\pi(x)\wedge\pi(y)}^{4/\epsilon}\frac{2du}{u\Phi^2(u)}$$

we have

$$\chi\Big(p^n(x,\cdot),\pi\Big) \le \sqrt{\epsilon} \text{ and } \chi\Big(\overleftarrow{p}^{\,m}(z,\cdot),\pi\Big) \le \sqrt{\epsilon}\,.$$

Thus

$$\begin{aligned}
\Big|\frac{p^{n+m}(x,z)}{\pi(z)}\Big| &= \Big|\frac{1}{\pi(z)}\sum_{y\in V} p^n(x,y)p^m(y,z)\Big| \qquad (40)\\
&= \Big|\sum_{y\in V}\pi(y)\Big(\frac{p^n(x,y)}{\pi(y)}\Big)\Big(\frac{\overleftarrow{p}^{\,m}(z,y)}{\pi(y)}\Big)\Big| \qquad (41)\\
&\le \chi\Big(p^n(x,\cdot),\pi\Big)\chi\Big(\overleftarrow{p}^{\,m}(z,\cdot),\pi\Big) \le \epsilon, \qquad (42)
\end{aligned}$$

where the first inequality is Cauchy-Schwarz. This establishes Theorem 2. □

7 Continuous Time

In this section we extend our results to continuous-time chains. We consider the chain $\{X_t, t \ge 0\}$ that moves at rate 1 according to P, where $P(x,y)$ is a transition kernel on V with stationary distribution π. We can assume that P satisfies the lazy condition (2). (If not, then increase the rate of transitions to 2 and replace P by $\frac{1}{2}(P+I)$.) Let Φ and ψ be defined with respect to P. Define $\tilde{p}^t(x,y) = \mathbf{P}_x(X_t = y)$.

Theorem 13 *The ϵ-uniform mixing time satisfies*

$$\tau(\epsilon) \le \int_{\pi_*}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}\,. \tag{43}$$

More precisely, if

$$t \ge \int_{\pi(x)\wedge\pi(y)}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}\,, \tag{44}$$

then

$$\Big|\frac{\tilde{p}^t(x,y)-\pi(y)}{\pi(y)}\Big| \le \epsilon. \tag{45}$$

Proof: As before, it is enough to show that if $\mu_t = \tilde{p}^t(x,\cdot)$, then $\chi^2(\mu_n,\pi) \le \epsilon$ for all

$$t \ge \int_{\pi(x)}^{4/\epsilon} \frac{du}{u\psi(u)}\,.$$

We accomplish this using the natural continuous-time evolving set process corresponding to X_t. Let $\{\tilde{S}_t : t \ge 0\}$ be the process which at rate 1 moves according to the evolving set process for P. Note that

$$\begin{aligned}
\tilde{p}^t(x,y) &= \sum_{j=0}^{\infty}\left[e^{-t}\,\frac{t^j}{j!}\right] p^n(x,y) && (46)\\
&= \sum_{j=0}^{\infty}\left[e^{-t}\,\frac{t^j}{j!}\right] \frac{\pi(y)}{\pi(x)}\,\mathbf{P}_{\{x\}}(y\in S_j) && (47)\\
&= \frac{\pi(y)}{\pi(x)}\,\mathbf{P}_{\{x\}}\Big(y\in\tilde{S}_t\Big)\,. && (48)
\end{aligned}$$

Our proof will parallel the proof of Theorem 4. Let μ_t be the distribution of X_t. Then one can argue as in Section 3 to obtain $\chi(\mu_t,\pi) \le \frac{1}{\pi(x)}\mathbf{E}\sqrt{\pi(S_t^\sharp)}$. Define

$$\tilde{Z}_t = \frac{\sqrt{\pi(\tilde{S}_t^\sharp)}}{\pi(\tilde{S}_t)}\,,$$

and let $\tilde{L}_t = \widehat{\mathbf{E}}(\tilde{Z}_t)$, so that $\chi(\mu_t,\pi) \le \tilde{L}_t$. Note that

$$\tilde{L}_t = \sum_{j=0}^{\infty}\left[e^{-t}\,\frac{t^j}{j!}\right] L_j$$

is differentiable. Equation (27) implies that

$$\widehat{\mathbf{E}}\left(\frac{\tilde{Z}_{t+\epsilon}}{\tilde{Z}_t}\Big|\,\tilde{S}_t, D\right) \le 1 - f_0(\tilde{Z}_t)$$

where D is the event that the evolving set process makes exactly one transition in $(t, t+\epsilon]$. It follows that for all $t \ge 0$ we have

$$\mathbf{E}(\tilde{Z}_t - \tilde{Z}_{t+\epsilon}|\tilde{S}_t) \ge \epsilon\tilde{Z}_t f_0(\tilde{Z}_t) + O(\epsilon^2).$$

Fix $t \geq 0$. Taking expectations above, we get

$$\mathbf{E}\left(\tilde{Z}_t - \tilde{Z}_{t+\epsilon}\right) \geq \epsilon \mathbf{E}\left[2\tilde{Z}_t f(2\tilde{Z}_t)\right] + O(\epsilon^2) \geq \epsilon \tilde{L}_t f(\tilde{L}_t) + O(\epsilon^2)\,, \tag{49}$$

where the last inequality holds by Lemma 12.

It follows that $\tilde{L}_{t+\epsilon} - \tilde{L}_t \leq -\epsilon \tilde{L}_t f(\tilde{L}_t) + O(\epsilon^2)$ and hence

$$\tilde{L}_t' \leq -\tilde{L}_t f(\tilde{L}_t). \tag{50}$$

The following Lemma is an analog of Lemma 11.

Lemma 14 *For every*

$$t \geq \int_\delta^{\tilde{L}_0} \frac{dz}{zf(z)},$$

we have $\tilde{L}_t \leq \delta$.

Proof: It's enough to show that for all $t \geq 0$ we have

$$t \leq \int_{\tilde{L}_t}^{\tilde{L}_0} \frac{dz}{zf(z)}. \tag{51}$$

It's true for $t = 0$, and differentiating both sides gives

$$1 \leq -\frac{1}{\tilde{L}_t f(\tilde{L}_t)} \tilde{L}_t',$$

which is true by equation (50). □

Lemma 14 implies that for all

$$t \geq \int_\delta^{\tilde{L}_0} \frac{2dz}{zf_0(z/2)} = \int_\delta^{\tilde{L}_0} \frac{2dz}{z\psi(4/z^2)} \tag{52}$$

we have $\chi(\mu_t, \pi) \leq \tilde{L}_t \leq \delta$. We calculated earlier that the integral on the RHS is at most

$$\int_{\pi(x)}^{\frac{1}{2}} \frac{du}{u\psi(u)} + \frac{1}{\psi_*} \log \frac{8}{\delta^2}.$$

This establishes Theorem 13. □

8 Concluding remarks

1. The example of the lamplighter group in the introduction shows that $\tau_V(1/4)$, the mixing time in total variation on the left-hand side of (5), can be considerably smaller than the corresponding uniform mixing time $\tau(1/4)$ (so an upper bound for $\tau(\cdot)$ is strictly stronger). We note that there are simpler examples of this phenomenon. For lazy random walk on a clique of n vertices, $\tau_V(1/4) = \Theta(1)$ while $\tau(1/4) = \Theta(\log n)$. To see a simple example with bounded degree, consider a graph consisting of two expanders of cardinality n and 2^n, respectively, joined by a single edge. In this case $\tau_V(1/4)$ is of order $\Theta(n)$, while $\tau(1/4) = \Theta(n^2)$.

2. Variants of conductance can give better bounds on Φ_S. For $S \subset V$, define

$$\theta_S := \frac{1}{\pi(S)} \sum_{y \in S} \sqrt{\pi(y) Q(S^c, y)}.$$

Note that for reversible chains we have

$$\theta_S = \frac{1}{\pi(S)} \sum_{y \in S} \pi(y) \sqrt{p(y, S^c)},$$

which is strictly greater than Φ_S for $S \notin \{\emptyset, V\}$.

Following Houdré and Tetali [11], denote

$$h_2^+ = \inf \left\{ \theta_S : \pi(S) \leq \tfrac{1}{2} \right\} . \tag{53}$$

Theorem 15 *Suppose that $p(x,x) \geq \frac{1}{2}$ for all x. Then*

$$\psi(S) \geq \frac{\theta_S^2}{8 \log(2/\theta_S^2)} .$$

Consequently, assuming reversibility, the spectral gap $1 - \lambda_2$ satisfies

$$1 - \lambda_2 \geq \frac{(h_2^+)^2}{8 \log \left(2/(h_2^+)^2 \right)} \geq c \, \frac{(h_2^+)^2}{\log(h_2^+)} . \tag{54}$$

Up to the logarithmic factor in the denominator, this type of inequality was conjectured by Houdré and Tetali [11, Remark 3.5].

Proof: For $u \in [0,1]$, let $A_u = \{y : Q(S,y) > u\pi(y)\}$. Note that $A_u \subset S$ for $u > \frac{1}{2}$. We have

$$\begin{aligned} \mathbf{E} \sqrt{\frac{\pi(\widetilde{S})}{\pi(S)}} &= \int_0^1 \sqrt{\frac{\pi(A_u)}{\pi(S)}} du \\ &= \int_0^1 \sqrt{1 + \frac{\pi(A_u) - \pi(S)}{\pi(S)}} du \\ &\leq 1 + \int_0^1 \frac{\pi(A_u) - \pi(S)}{\pi(S)} du \\ &\qquad - \tfrac{1}{8} \int_{\frac{1}{2}}^1 \frac{(\pi(A_u) - \pi(S))^2}{\pi(S)^2} du, \end{aligned}$$

by the inequality $\sqrt{1+t} \leq 1 + t/2 - \frac{1}{8} t^2 \mathbf{1}(t \leq 0)$, valid for $t \geq -1$. Define

$$\begin{aligned} B_t := S - A_{1-t} &= \{ y \in S : Q(S,y) \leq (1-t)\pi(y) \} \\ &= \{ y \in S : Q(S^c, y) \geq t\pi(y) \}. \end{aligned}$$

The middle term in (55) vanishes by the martingale property. Thus

$$\begin{aligned}\psi(S) &\geq \tfrac{1}{8}\int_{\frac{1}{2}}^{1} \frac{\pi(S-A_u)^2}{\pi(S)^2}du \\ &= \tfrac{1}{8}\int_{0}^{\frac{1}{2}} \frac{\pi(B_t)^2}{\pi(S)^2}\,dt,\end{aligned}$$

where we have made the substitution $t = 1-u$. Therefore, for any $\alpha > 0$ we have

$$\begin{aligned}8\pi(S)^2\psi(S)\cdot\log(1/2\alpha) &\geq \int_{\alpha}^{\frac{1}{2}} \pi(B_t)^2\,dt \cdot \int_{\alpha}^{\frac{1}{2}} \frac{1}{t}\,dt \\ &\geq \left[\int_{\alpha}^{\frac{1}{2}} \frac{\pi(B_t)}{\sqrt{t}}\,dt\right]^2,\end{aligned}$$

by Cauchy-Schwarz. But

$$\begin{aligned}\int_{\alpha}^{\frac{1}{2}} \frac{\pi(B_t)}{\sqrt{t}}\,dt &= \sum_{y\in S}\pi(y)\int_{\alpha}^{\frac{1}{2}} \mathbf{1}(y\in B_t)\frac{1}{\sqrt{t}}\,dt \\ &= \sum_{y\in S}\pi(y)\int_{\alpha}^{Q(S^c,y)/\pi(y)} \frac{1}{\sqrt{t}}\,dt \\ &= 2\left(\sum_{y\in S}\sqrt{Q(S^c,y)\pi(y)}\right) - 2\sqrt{\alpha}\cdot\pi(S) \\ &= 2\pi(S)\left(\theta_S - \sqrt{\alpha}\right).\end{aligned}$$

Setting $\alpha = \theta_S^2/4$ and using equation (55), we get

$$8\psi(S)\log(2/\theta_S^2) \geq \frac{1}{\pi(S)^2}\left[\int_{\theta_S^2/4}^{\frac{1}{2}} \frac{\pi(B_t)}{\sqrt{t}}\,dt\right]^2 \geq \theta_S^2,$$

and the theorem follows. $\square$

3. Theorems 4 and 5 can be improved under a convexity condition that holds in many examples. In the setting of Theorem 4, if $\psi(r) \geq \psi_c(r)$ for all $r \geq \pi_*$ where $u \mapsto u\psi_c(u^{-2})$ is a convex function of $u \in [\sqrt{2}, \infty)$, then $\chi^2(\mu_n, \pi) \leq \epsilon$ for all

$$n \geq \tfrac{1}{2}\int_{\pi(x)}^{\frac{1}{\epsilon}} \frac{du}{u\psi(u)}.$$

To prove this, follow the proof of Theorem 4 until (26), which implies that

$$\widehat{\mathbf{E}}_n\left(\frac{Z_{n+1}}{Z_n}\Big|S_n\right) \leq 1 - \psi_c(\pi(S_n)) = 1 - f(Z_n),$$

where $f(z) := \psi_c(z^{-2})$. Then invoke Lemma 11(ii) and apply the change of variable $u = 1/z^2$ to the integral there.

Consequently, the following variant of Theorem 5 holds. suppose that $p(x,x) \geq \gamma$ for all $x \in V$. If $\Phi(r) \geq \Phi_c(r)$ for all $r > 0$, where $u \mapsto u\Phi_c^2(u^{-2})$ is a convex function of $u \in [\sqrt{2}, \infty)$, then (8) holds for all

$$n \geq \frac{(1-\gamma)^2}{\gamma^2} \int_{\pi(x)\wedge\pi(y)}^{1/\epsilon} \frac{2du}{u\Phi_c^2(u)} \,. \tag{55}$$

4. Let $\mathcal{E}$ denote the support of the evolving set process. Theorem 4 can be improved by using $\psi_{\mathcal{E}}(r) = \inf\{\psi(S) : \pi(S) \leq r, S \in \mathcal{E}\}$, instead of $\psi(r)$. For random walk on the n dimensional hypercube, $\mathcal{E}$ consists of Hamming balls, $\psi_{\mathcal{E}}(r) \geq \frac{c}{n}\log(1/r)$ and this gives an upper bound of $O(n \log n)$ for the uniform mixing time $\tau(1/4)$.

   More generally, for any Markov chain $\{X_n\}$ on a poset with a monotone time-reversal, if X_0 is a maximal (or minimal) state, then $\mathcal{E}$ consists of increasing (respectively, decreasing) sets.

Acknowledgments. We are grateful to D. Aldous, L. Lovász , R. Montenegro, E. Mossel and A. Sinclair for useful discussions.

References

[1] Alon, N. (1986) Eigenvalues and expanders. *Combinatorica* **6**, 83–96.

[2] Alon, N. and Milman, V. D. (1985). λ_1, Isoperimetric inequalities for graphs and superconcentrators, *J. Combinatorial Theory Ser. B* **38**, 73–88.

[3] Benjamini, I. and Mossel, E. (2003). On the mixing time of a simple random walk on the super critical percolation cluster. *Probab. Th. Rel. Fields*, **125**, 408–420.

[4] Chung, F. R. K. (1996) Laplacians of graphs and Cheeger's inequalities. In *Combinatorics, Paul Erdös is eighty*, Vol. **2** 157–172, J. Bolyai Soc. Math. Stud., Budapest.

[5] Chung, F. R. K. and Yau, S. T. (1995) Eigenvalues of graphs and Sobolev inequalities, *Combinatorics, Probability and Computing* **4**, 11–26.

[6] Coulhon, T. (1996). Ultracontractivity and Nash type inequalities. *J. Funct. Anal.* **141**, 510–539.

[7] Coulhon, T., Grigorýan, A. and Pittet, C. (2001). A geometric approach to on-diagonal heat kernel lower bounds on groups. *Ann. Inst. Fourier (Grenoble)* **51**, 1763–1827.

[8] Diaconis, P. and Saloff-Coste, L. (1996). Nash inequalities for finite Markov chains. *J. Theoret. Probab.* **9**, 459–510.

[9] Fill, J. A. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.* **1**, 62–87.

[10] Häggström, O. and Jonasson, J. (1997). Rates of convergence for lamplighter processes. *Stochastic Process. Appl.* **67**, 227–249.

[11] Houdré, C. and Tetali, P. (2001). Isoperimetric Invariants for Product Markov Chains and Graph Products *Combinatorica*, to appear.

[12] Jerrum, M. R. and Sinclair, A. J. (1989). Approximating the permanent. *SIAM Journal on Computing* **18**, 1149–1178.

[13] Kannan, R. (2002). Rapid Mixing in Markov Chains *Proceedings of International Congress of Math.* 2002, Vol. III, 673–683.

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

[15] Lovász, L. and R. Kannan, R. (1999). Faster mixing via average conductance *Proceedings of the 27th Annual ACM Symposium on theory of computing.*

[16] Mathieu, P. and Remy, E. (2003) Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.*, to appear.

[17] Mihail, M. (1989). Conductance and convergence of Markov chains - A combinatorial treatment of expanders. *Proceedings of the 30th Annual Conference on Foundations of Computer Science*, 526–531.

[18] Montenegro, R. and Son, J.-B. (2001) Edge Isoperimetry and Rapid Mixing on Matroids and Geometric Markov Chains, *Proceedings of the 33rd Annual ACM Symposium on theory of computing.*

[19] Morris, B. (2002). A new, probabilistic approach to heat kernel bounds. Lecture at Sectional AMS meeting, Atlanta, GA, March 2002.

[20] Pittet, C. and Saloff-Coste, L. (2002) A survey on the relationships between volume growth, isoperimetry, and the behavior of simple random walk on Cayley graphs, with examples. Unpublished manuscript, available at `http://www.math.cornell.edu/~lsc/lau.html`

[21] Saloff-Coste, L. (1997). *Lectures on finite Markov chains.* Lecture Notes in Math. 1665, Springer, Berlin, 301–413.

[22] Sinclair, A. (1993) *Algorithms for Random Generation and Counting: A Markov Chain Approach*, Birkhäuser, Boston.

[23] Varopoulos, N. Th. (1985) Isoperimetric inequalities and Markov chains. *J. Funct. Anal.* **63**, 215–239.

[24] Woess, W. (2000). *Random walks on infinite graphs and groups.* Cambridge Tracts in Mathematics, 138. Cambridge University Press.