The Annals of Applied Probability
1994, Vol. 4, No. 3, 812–837

SAMPLING FROM LOG-CONCAVE DISTRIBUTIONS

By Alan Frieze, Ravi Kannan and Nick Polson

Carnegie Mellon University, BellCore and Carnegie Mellon University, and University of Chicago

We consider the problem of sampling according to a distribution with log-concave density F over a convex body $K \subseteq \mathbb{R}^n$. The sampling is done using a biased random walk, and we prove polynomial upper bounds on the time to get a sample point with distribution close to F.

1. Introduction. This paper is concerned with the efficient sampling of random points from $\mathbb{R}^n$, where the underlying density F is log-concave (i.e., $\log F$ is concave). This is a natural restriction which is satisfied by many common distributions, for example, the multivariate normal. The algorithm we use generates a sample path from a Markov chain whose stationary distribution is (close to) F. The algorithm falls into the class of Metropolis algorithms. It has applications to the problem of computing the volume of convex bodies and in statistics. Recent statistics literature has focused on the many applications of Markov chain Monte Carlo algorithms (see [11]). However, theoretical bounds on the convergence rate of such algorithms have been limited. Using recent developments in the theory of rapidly mixing Markov chains, in particular the notion of conductance [6, 10], Applegate and Kannan [2] proved a bound on the convergence rate of the chain considered in this paper. These theoretical bounds are too large to be useful in practice. In this paper, we prove tighter bounds using an approach related to the classical Poincaré inequalities instead of conductance. These bounds are small enough to be useful in practice and have already found application in the generation of random contingency tables [4].

Instead of sampling from the continuum of points in $\mathbb{R}^n$, we discretize the problem by assuming that $\mathbb{R}^n$ is divided into a set of hypercubes $\mathscr{C}_{\mathbb{R}}$ of side δ (δ is a given small positive real number) and the problem is to choose one of these cubes each with probability proportional to the integral of F over the cube. (If necessary, a sample from the continuum can then be picked by standard rejection sampling techniques from the cube chosen; we omit details of this.) Second, we assume that we have a compact convex set K and we wish to choose points only from K (not all of $\mathbb{R}^n$). This is justified because clearly for any positive real number ε, we can find a compact convex set (for example, a ball) such that the integral of F over the set is at least $(1 - \varepsilon)$

Received May 1993; revised November 1993.
[1] Supported in part by NSF Grants CCR 90-24935 and CCR 92-08597.
AMS 1991 *subject classifications*. 60J15, 68Q25.
Key words and phrases. Sampling, random walk, log-concave functions.

times the integral over $\mathbb{R}^n$. While it would suffice to consider the case when K is a ball, the generality of convex sets is useful in many contexts.

In what follows $\mathscr{C}$ denotes a subset of the cubes $\mathscr{C}_{\mathbb{R}}$ whose union contains K. For example, $\mathscr{C}$ could be the set of cubes $\mathscr{C}_I(K)$, which have nonempty intersection with K. Let C denote the set of centres of these cubes. For $x \in \mathbb{R}^n$, we denote the cube of side δ and centre x by $C(x)$. [Thus $C(x) \in \mathscr{C}$ if and only if $x \in C$.] We choose our sample point X by performing a random walk over C. The walk is biased so that its steady state is (close to) what we want, and we run the walk until it is close enough to the steady state. The main results of this paper concern the rate of convergence of the walk to its steady state.

We may not be able to compute F exactly and so we assume we have good approximations $\overline{F}(x)$, $x \in C$. Further, we assume that $\overline{F}(x)$ is strictly positive for all $x \in C$.

We can only take advantage of the log-concavity of F if our grid is sufficiently fine and our approximations $\overline{F}(x)$ are sufficiently good. In this context, we will assume that for some small $\alpha > 0$,

$$(1) \qquad (1+\alpha)^{-1}\overline{F}(x) \le \overline{F}(y) \le (1+\alpha)\overline{F}(x)$$

and

$$(2) \qquad (1+\alpha)^{-1}\delta^{n-1}\overline{F}(x) \le \int_{C(x)\cap C(y)} F(\xi)\,d\xi \le (1+\alpha)\delta^{n-1}\overline{F}(x)$$

whenever $C(x), C(y)$ are cubes of $\mathscr{C}$ sharing a face $C(x) \cap C(y)$ of dimension $n-1$. Furthermore, we assume

$$(3) \quad (1+\alpha)^{-1}\delta^n\overline{F}(x) \le \int_{C(x)} F(\xi)\,d\xi \le (1+\alpha)\delta^n\overline{F}(x), \qquad \forall\, x \in C.$$

When we have $\overline{F} = F$, it is easy to check that all three conditions are satisfied if we choose α to be $e^{M\delta} - 1$, where M is the Lipschitz constant of $\ln F$ with respect to the infinity norm [i.e., M satisfies $|\ln F(x) - \ln F(y)| \le M|x-y|_\infty$, $\forall\, x, y \in K$]. However, an α smaller than $e^{M\delta} - 1$ may satisfy (1)–(3); this is, in fact, the case for important functions like $F(x) = e^{-c|x|}$ and $F(x) = e^{-c|x|^2}$, as tedious, but simple, calculations show. This is the reason for stating the cumbersome conditions (1)–(3). As we will see (Theorems 1 and 2), the rate of convergence to the steady state depends upon $(1+\alpha)$. In typical applications, one would make $1+\alpha$ a constant. [For example, this can be ensured by choosing $\delta = O(1/M)$.]

The walk we consider fits into the scheme of Metropolis algorithms introduced in [8]. It was used by Applegate and Kannan [2] in their paper on volume computation.

In the following text, for any natural number m, we let $[m] = \{1, 2, \dots, m\}$ and $e_1, e_2, \dots, e_m$ are the standard basis vectors of $\mathbb{R}^m$.

The random walk. This generates a random sequence X_0, X_1, $X_2, \dots, X_t, \dots \in C$, where X_0 is picked according to some initial distribution $p_0(x)$ and X_{t+1} is obtained from X_t as follows:

STEP 1. Choose j randomly from $[n]$. Choose σ randomly from $\{\pm 1\}$.

STEP 2. Let $y = X_t + \sigma e_j$.

STEP 3. If $y \notin C$, then $X_{t+1} = X_t$. Otherwise, put $X_{t+1} = y$ with probability $\theta = \min\{1, \bar{F}(y)/\bar{F}(X_t)\}$ and $X_{t+1} = X_t$ with probability $1 - \theta$.

Formally, the transition probabilities $P(x, y) = \Pr(X_{t+1} = y | X_t = x)$ are given by

$$P(x,y) = \frac{1}{2n} \min\{1, \bar{F}(y)/\bar{F}(x)\} \quad \text{for } x \neq y,\ x, y \text{ adjacent},$$

and

$$P(x,x) = 1 - \sum_{y \neq x} P(x,y).$$

We refer to this as "the random walk" in this paper. It will be useful also to consider a modified random walk. The "modified random walk" has

$$P(x,y) = \frac{\nu}{2n} \min\{1, \bar{F}(y)/\bar{F}(x)\} \quad \text{for } x \neq y,\ x, y \text{ adjacent}$$

and

$$P(x,x) = 1 - \sum_{y \neq x} P(x,y),$$

where $\nu = (1 + 2\delta^2/((1+\alpha)^3 nd(d + 2\sqrt{n}\,\delta)))^{-1}$. In most applications, ν will be very close to 1. So the modification can be thought of as follows: With a small probability, the walk stays put. Otherwise, it does what "the (unmodified) random walk" did.

When $\mathscr{C} = \mathscr{C}_I(K)$ we will refer to the walk as the "body intersecting" walk. In this case, testing whether $y \in C$ can be a significant computational problem, but nevertheless polynomially solvable, in general. In specific cases, for example, when $K = B(0, R)$ (the ball of radius R centred at the origin), the problem is rather trivial.

We will also consider the computationally simpler random walk over those cubes $\mathscr{C}_C(K)$ in $\mathscr{C}_{\mathbb{R}}$ whose centres are in K. We call this the "centre point" random walk.

It is easy to see that the chain is ergodic and thus, there are steady state probabilities $\pi(x)$ with $\lim_{t\to\infty} \Pr(X_t = x) = \pi(x)$ for all x independent of the distribution of X_0. It is easy to verify that

$$\pi(x) = \bar{F}(x)/\Delta,$$

where $\Delta = \sum_{x \in C} \bar{F}(x)$. We assume that the $\bar{F}(x)$ are sufficiently good approximations so that sampling according to π can be considered to be our objective.

Note that this chain is *time reversible*, that is,

$$\pi(x)P(x, y) = \pi(y)P(y, x) \quad \text{for all } x, y \in C.$$

The main theorem describing the rate of convergence is stated in the following text as Theorem 1. We use the so-called chi-squared distance of the distribution after t steps to the steady state as a measure of closeness (see [9] and [5]). Let $p_t(x) = \Pr(X_t = x)$ be the distribution after t steps. Then this distance is given by

$$\sum_{x \in C} \left(\frac{p_t(x)}{\pi(x)} - 1 \right)^2 \pi(x).$$

There are other measures one could use, including the more traditional variational distance given by

$$\sum_{x \in \mathscr{C}} |p_t(x) - \pi(x)|.$$

The chi-squared measure turns out to yield stronger theorems and, in fact (as Fill points out), one can derive a bound on the variational distance using the bound on the chi-squared distance (see Corollary 1).

In Theorems 1 and 2, which will be proved later, we have a random walk over a set of cubes $\mathscr{C}$:

$$T = \bigcup_{D \in \mathscr{C}} D$$

and

$$d = \operatorname{diam}(T)$$

(the maximum distance between two points of T).

THEOREM 1. *Let K be a convex set in $\mathbb{R}^n$. Let $\mathscr{C}$ be a subset of $\mathscr{C}_{\mathbb{R}}$ (with $\delta \le d/10$) such that $K \subseteq T$. Let C be the set of centres of these cubes. Let S be the set of centres of those cubes in $\mathscr{C}$ which are not wholly contained in K and let $\bar{S} = C \setminus S$. Let F be a log-concave positive real-valued function on K with α satisfying* (1), (2) *and* (3). *Consider the random walk $X_0, X_1, \dots X_t, \dots$ described previously. Then, with $p_t(x) = \Pr(X_t = x)$ [and $\pi(x)$, the steady state probability of being at x], we have*

$$\left(\sum_{x \in C} \left(\frac{p_t(x)}{\pi(x)} - 1 \right)^2 \pi(x) \right)^{1/2} \le (1 - \lambda_1)^t \left(\sum_{x \in C} \left(\frac{p_0(x)}{\pi(x)} - 1 \right)^2 \pi(x) \right)^{1/2} + \left(\frac{3}{1 - \lambda_1} \max_{x \in C} \left| \frac{p_0(x)}{\pi(x)} - 1 \right| \right) \left(\frac{\pi(S)}{\pi(\bar{S})} \right)^{1/2},$$

where

$$\kappa_0 = \frac{d}{4\delta}(d + 2\sqrt{n}\,\delta),$$

$$\lambda_1^{-1} = \max\left\{\frac{(1+\alpha)^3\kappa_0 n}{\delta}, \frac{d^2}{6\delta^2}\right\} \tag{4}$$

$$\approx \frac{nd^2}{4\delta^2}, \tag{5}$$

and the final approximate equation holds under the assumptions $\alpha \approx 0$ *and* $2\sqrt{n}\,\delta/d \approx 0$.

Further, the modified random walk satisfies the inequality with λ_1 *replaced by* $\hat{\lambda}_1$, *where*

$$\hat{\lambda}_1^{-1} = \nu^{-1}\frac{(1+\alpha)^3\kappa_0 n}{\delta}$$

and where $\nu = (1 + \delta/(2(1+\alpha)^3\kappa_0 n))^{-1}$. (*Also*, $\nu \approx 1$ *under the same assumptions as before.*)

REMARK 1. When n is large, the term $\tau_2 = d^2/(6\delta^2)$ in (4) is dominated by the other term. However, when n is small, this may not be the case and the modified random walk will be preferable.

REMARK 2. The quantity $1 - \lambda_1$ estimates the second largest eigenvalue of the transition matrix of the walk. The constant 4 in (5) is reasonably close to optimal. For example, when $n = 1$, the second largest eigenvalue is $1 - 2\sin(\pi\delta/d)^2 \approx 1 - \pi^2\delta^2/(2d^2)$; see [1].

COROLLARY 1. *Under the same conditions as in Theorem* 1, *we have the following bound on the variational distance*:

$$\sum_{x\in C}|p_t(x) - \pi(x)| \le (1-\lambda_1)^t\left(\sum_{x\in C}\left(\frac{p_0(x)}{\pi(x)} - 1\right)^2\pi(x)\right)^{1/2} + \left(\frac{3}{1-\lambda_1}\max_{x\in C}\left|\frac{p_0(x)}{\pi(x)} - 1\right|\right)\left(\frac{\pi(S)}{\pi(\bar{S})}\right)^{1/2}.$$

PROOF. Using the Cauchy–Schwarz inequality, we have

$$\sum_{x\in C}|p_t(x) - \pi(x)| \le \left(\sum_{x\in C}\frac{(p_t(x) - \pi(x))^2}{\pi(x)}\right)^{1/2}\left(\sum_{x\in C}\pi(x)\right)^{1/2}.$$

The latter quantity equals the left-hand side of the inequality in Theorem 1 and so we have the corollary. □

We first consider the body intersecting walk. Here we can apply Theorem 1 and its corollary directly. The first term of the bound falls off exponentially to zero with t. Theorem 1 is only of use if $\pi(S)$ is small. For any log-concave function F, we can choose a sufficiently large compact set such that the integral of F over the set is close to the integral of F over $\mathbb{R}^n$. If now we choose K to be somewhat larger than this set, then clearly, $\pi(S)$ will be small. We do not quantify these observations here, because that would depend upon the properties of the specific F's.

Let us now consider the centre point walk. We define the convex set $K' = \{x \in K: C(x) \subseteq K\}$, run the walk with $\mathscr{C} = \mathscr{C}_C(K)$ and apply Theorem 1 to $\mathscr{C}$ and K'. This is valid as $K' \subseteq \bigcup_{d \in \mathscr{C}} D$ and when δ is small, K and K' will be "close" to each other. The comments in the previous paragraph are valid in this case too.

It would be preferable to remove the term involving $\pi(S)$ in Theorem 1, since it does not tend to zero at $t \to \infty$. With extra assumptions, we can manage this.

Assume without loss that $o \in K$. For $u \in \mathbb{R}^n$, with $|u| = 1$, let $L_u = \{ru: r \in \mathbb{R}^+\}$ be the ray in direction u emanating from the origin. Let $h(r) = h_u(r) = r^{n-1}F(ru)$. Theorem 2, (which follows) is proved under the assumption that the following conditions hold for all unit vectors u:

$$L_u \cap T \text{ is an interval.} \tag{6}$$

Let $R = R(u) = |L_u \cap K|$ and $R_1 = R_1(u) = |L_u \cap T|$. Then

$$R_1 < 2R. \tag{7}$$

Let $\varsigma = R_1 - R$. Then

$$h(r') \le \kappa_1 h(r), \quad \text{for } R - \varsigma \le r \le r' \le r + \varsigma, \tag{8}$$

where $\kappa_1 \ge 1$ is defined to be the smallest value satisfying (8). Let

$$\kappa_2 = 2\kappa_1\left(\frac{\varsigma}{\delta} + \sqrt{n}\right).$$

THEOREM 2. *Let F be a log-concave positive real-valued function on $\mathbb{R}^n$ with α satisfying* (1)–(3). *Assume* (6)–(8) *hold. Consider the random walk $X_0, X_1, \dots X_t, \dots$ described previously, with $\mathscr{C} = \mathscr{C}_I(K)$ and $5\delta \le R$. Then we have*

$$\left(\sum_{x \in C}\left(\frac{p_t(x)}{\pi(x)} - 1\right)^2 \pi(x)\right)^{1/2} \le (1 - \lambda_2)^t \left(\sum_{x \in C}\left(\frac{p_0(x)}{\pi(x)} - 1\right)^2 \pi(x)\right)^{1/2},$$

where

$$\lambda_2^{-1} = \max\left\{(1 + \alpha)^3\left(\kappa_0\kappa_1 + \kappa_0 + \kappa_2 + 2\sqrt{\kappa_0\kappa_1\kappa_2}\right)\frac{n}{\delta}, \frac{d^2}{6\delta^2}\right\}$$
$$\approx \frac{n}{2}\left(\frac{d}{\delta}\right)^2$$

and the final approximate equation holds under the assumptions $\alpha \approx 0$, $\kappa_1 \approx 1$, $\kappa_2 \ll \kappa_0$ and $18\delta\sqrt{n}/d \approx 0$.

COROLLARY 2. *Under the same conditions as in Theorem 2, we have the following bound on the variational distance*:

$$\sum_{x \in C} |p_t(x) - \pi(x)| \le (1-\lambda_2)^t \left(\sum_{x \in C} \left(\frac{p_0(x)}{\pi(x)} - 1 \right)^2 \pi(x) \right)^{1/2}.$$

In the next section, we prove a general result about the rate of convergence to the steady state for arbitrary time-reversible chains. It is then a matter of estimating quantities associated with this result.

2. Convergence rates of time-reversible chains: Dealing with small sets. In the analysis of geometric random walks, it turns out that certain sets of states with relatively small steady state probabilities are not easy to analyse. In the case of random walks on convex sets, where the problem is usually discretized by using a set of cubes as the states, this problem arises for the set of S of *boundary cubes*, which are those cubes in $\mathscr{C}$ which are not wholly contained in the convex set. This irksome problem has cost volume computation algorithms significant added complexity [3, 6] and has also led to new ideas like the use of log-concave damping functions as in [2]. In this section, we propose a general way of tackling the problem of small sets. In some sense, this serves a purpose analogous to that of μ-conductance proposed by Lovász and Simonovits [6].

Suppose P is the $N \times N$ transition probability matrix of an ergodic Markov chain with N states and steady state probabilities $\pi(\cdot)$. We assume throughout that the chain is time-reversible, that is, that $\pi(x)P(x,y) = \pi(y)P(y,x)$ for all x, y. Suppose we start the chain with initial distribution $p_0(\cdot)$ [i.e., $\Pr(X_0 = x) = p_0(x)$ for all x]. Let $p_t(\cdot)$ be the probability distribution of X_t. We define the quantity

$$\phi_t(x) = \frac{p_t(x)}{\pi(x)} - 1.$$

This leads to what Fill [5] calls the chi-squared distance between distributions.

To define this distance, we first define an inner product on $\mathbb{R}^n$ by

$$\langle \phi, \psi \rangle = \sum_{x \in [N]} \phi(x)\psi(x)\pi(x)$$

and denote its associated norm by

$$|\phi|_\pi = \langle \phi, \phi \rangle^{1/2}.$$

We also retain the more familiar notation that $|\ |$ denotes Euclidean length and that $|\phi|_\infty = \max\{|\phi(x)|: x \in [N]\}$.

Suppose

$$\sigma_0 = \sup\left\{ \frac{|\langle \phi, P\phi \rangle|}{\langle \phi, \phi \rangle} : \phi \ne 0,\ \pi^T\phi = 0 \right\}.$$

(From linear algebra and time-reversibility, it can be seen that the largest eigenvalue of P is 1 and that σ_0 is the largest absolute value of an eigenvalue of P other than 1. We will only use the first of these facts.)

Using arguments similar to the ones used by Fill [5] to get his inequality (2.11), we can get the inequality

$$|\phi_t|_\pi \le \sigma_0^t |\phi_0|_\pi .$$

So it would suffice then to prove an upper bound on σ_0. We are not always able to do this for the random walk described in Section 1. We are able to prove a bound on a quantity similar to σ_0 which ignores a small set S. Our aim in this section is to prove an inequality for general time-reversible chains, similar to the foregoing one, but with the new quantity σ replacing σ_0.

Let S be an arbitrary set of states which will remain fixed for the rest of this section. Let

$$\Omega = \{\phi \in \mathbb{R}^N : \pi^T\phi = 0 \text{ and } \phi(x) = 0,\ x \in S\}$$

and

$$\sigma = \sup\left\{ \frac{|\langle \phi, P\phi\rangle|}{\langle \phi, \phi\rangle} : \phi \in \Omega \setminus \{0\},\ \pi^T\phi = 0 \right\}.$$

The main theorem of this section follows.

THEOREM 3.

$$|\phi_t|_\pi \le \sigma^t |\phi_0|_\pi + \frac{3|\phi_0|_\infty}{(1-\sigma)} \sqrt{\frac{\pi(S)}{\pi(\bar{S})}} . \tag{9}$$

We need some preliminary linear algebra. Let D denote the $N \times N$ diagonal matrix whose (x, x)th entry is $\sqrt{\pi(x)}$. Then time-reversibility is equivalent to

$$D^2 P = P^T D^2. \tag{10}$$

Let

$$Q = DPD^{-1}.$$

Then Q has the same eigenvalues as P and (10) implies that it is symmetric. The matrix D "converts" $|\ |_\pi$ to $|\ |$ in a natural way. This follows from

$$\langle \phi, \psi \rangle = (D\phi)^T (D\psi)$$

for $\phi, \psi \in \mathbb{R}^N$. Thus

$$|\phi|_\pi = |D\phi| \quad \text{and} \quad \langle \phi, P\phi\rangle = (D\phi)^T Q (D\phi).$$

So using $\psi = D\phi$, we obtain

$$\sigma = \sup_{\substack{\psi \in \tilde{\Omega} \\ \psi \ne 0}} \left\{ \left| \frac{\psi^T Q \psi}{\psi^T \psi} \right| \right\}, \tag{11}$$

where

$$\tilde{\Omega} = \{\psi \colon \zeta^T\psi = 0 \text{ and } \psi(x) = 0,\ x \in S\}$$

and $\zeta(x) = \sqrt{\pi(x)}$ for $x \in [N]$. Let A denote the orthogonal projection of $\mathbb{R}^N$ to $\tilde{\Omega}$. It is easy to see that A is defined by

$$A\psi(x) = \begin{cases} \psi(x) - \dfrac{\sqrt{\pi(x)}}{\pi(S)}\left(\sum_{y \in \bar{S}} \psi(y)\sqrt{\pi(y)}\right), & x \in \bar{S}, \\ 0, & x \in S. \end{cases}$$

LEMMA 1. *Let $Q, \tilde{\Omega}, A$, and σ be as above. Then*

$$|AQA\phi| \le \sigma|\phi| \quad \textit{for all } \phi \in \mathbb{R}^N.$$

PROOF. For $\phi \in \tilde{\Omega}$, we have $\phi^T Q\phi = \phi^T AQA\phi$. So,

$$\sigma = \sup_{\substack{\phi \in \tilde{\Omega} \\ \phi \ne 0}} \left\{\left|\frac{\phi^T AQA\phi}{\phi^T\phi}\right|\right\} \le \sup_{\substack{\phi \in \mathbb{R}^N \\ \phi \ne 0}} \left\{\left|\frac{\phi^T AQA\phi}{\phi^T\phi}\right|\right\} = \sigma_1,$$

where σ_1 is the largest eigenvalue of the symmetric matrix $B = AQA$. [From standard linear algebra, if $\tilde{\Omega} = \{x\colon x = Cy\}$, where C is a matrix with independent columns, then $A = C(C^TC)^{-1}C^T$ and is, therefore, symmetric.] However, if ϕ is a corresponding eigenvector of B, then $B\phi = \sigma_1\phi$ implies that $\phi \in \tilde{\Omega}$. Hence $\sigma = \sigma_1$ and the lemma follows. □

LEMMA 2. $\zeta^T\psi = 0$ *implies*

$$|\psi - A\psi| \le |D^{-1}\psi|_\infty \sqrt{\frac{\pi(S)}{\pi(\bar{S})}}\,.$$

PROOF. Assume $\zeta^T\psi = 0$. Let $\psi' = \psi - A\psi$, $\phi = D^{-1}\psi$ and $M = |\phi|_\infty$. Then

$$\begin{aligned} \sum_{x \in [N]} \psi'(x)^2 &= \sum_{x \in S} \psi(x)^2 + \frac{1}{\pi(\bar{S})^2} \sum_{x \in \bar{S}} \left(\sum_{y \in \bar{S}} \psi(y)\sqrt{\pi(y)}\right)^2 \pi(x) \\ &= \sum_{x \in S} \phi(x)^2\pi(x) + \frac{1}{\pi(\bar{S})^2} \sum_{x \in \bar{S}} \pi(x)\left(\sum_{y \in \bar{S}} \phi(y)\pi(y)\right)^2 \\ &= \sum_{x \in S} \phi(x)^2\pi(x) + \frac{1}{\pi(\bar{S})^2} \sum_{x \in \bar{S}} \pi(x)\left(\sum_{y \in S} \phi(y)\pi(y)\right)^2 \\ &\le M^2\pi(S) + \frac{M^2\pi(S)^2}{\pi(\bar{S})} \\ &= \frac{M^2\pi(S)}{\pi(\bar{S})}. \end{aligned}$$

□

We now go back to the proof of the main theorem of this section—Theorem 3.

PROOF OF THEOREM 3. Let $\psi_1 = D\phi_t$. Observe that $|\phi|_\pi = |\psi|$. We will prove that

$$(12) \qquad |\psi_{t+1}| \le \sigma|\psi_t| + 3|\phi_0|_\infty \sqrt{\frac{\pi(S)}{\pi(\bar{S})}},$$

from which we get, by induction on t,

$$\begin{aligned} |\psi_t| &\le \sigma^t|\psi_0| + 3|\phi_0|_\infty\left(\frac{1-\sigma^{t+1}}{1-\sigma}\right)\sqrt{\frac{\pi(S)}{\pi(\bar{S})}} \\ &\le \sigma^t|\psi_0| + 3\frac{|\phi_0|_\infty}{1-\sigma}\sqrt{\frac{\pi(S)}{\pi(\bar{S})}}, \end{aligned}$$

as desired. Now

$$\zeta^T\psi_t = \sum_{x\in[N]}(p_t(x) - \pi(x)) = 0.$$

Also, using time-reversibility, we get

$$\begin{aligned} (P\phi_t)(x) &= \sum_{y\in[N]}\frac{P(x,y)p_t(y)}{\pi(y)} - 1 \\ &= \sum_{y\in[N]}\frac{P(y,x)p_t(y)}{\pi(x)} - 1 = \phi_{t+1}(x). \end{aligned}$$

Thus $P\phi_t = \phi_{t+1}$, and premultiplying by D, we get

$$\psi_{t+1} = Q\psi_t \quad \text{for } t \ge 0.$$

To prove (12), we proceed as follows:

$$|\psi_{t+1}| = |Q\psi_t| \le |QA\psi_t| + |Q(\psi_t - A\psi_t)|.$$

Using the fact that the eigenvalues of Q have absolute value at most 1 and Lemma 2, along with the fact that $|\phi_{t+1}|_\infty \le |\phi_t|_\infty$ (since P has row sums equal to 1), we have

$$|Q(\psi_t - A\psi_t)| \le |\psi_t - A\psi_t| \le |\phi_0|_\infty\sqrt{\frac{\pi(S)}{\pi(\bar{S})}}.$$

Now we need to bound $|QA\psi_t|$. However,

$$|QA\psi_t| = |AQA\psi_t| + |QA\psi_t - AQA\psi_t|.$$

We have by Lemma 1 that $|AQA\psi_t| \le \sigma|\psi_t|$. So it now suffices to prove that

$$|QA\psi_t - AQA\psi_t| \le 2|\phi_0|_\infty \sqrt{\frac{\pi(S)}{\pi(\bar{S})}}\,.$$

Now

$$\begin{aligned}|QA\psi_t - A(QA\psi_t)| &= |(I-A)QA\psi_t| \\ &\le |(I-A)Q\psi_t| + |(I-A)Q(A\psi_t - \psi_t)| \\ &\le |(I-A)\psi_{t+1}| + |A\psi_t - \psi_t|,\end{aligned}$$

since Q has all eigenvalues of absolute value at most 1 and also for all vectors v, $|(I-A)v| \le |v|$, because A is a projection. Each of the last two quantities is at most $|\phi_0|_\infty\sqrt{\pi(S)/\pi(\bar{S})}$ by Lemma 2. This finishes the proof of Theorem 3. □

Our aim is to prove an upper bound on

$$\sigma = \sup_{\substack{\phi \in \Omega\setminus\{0\} \\ \pi^T\phi = 0}} \left\{ \frac{|\langle \phi, P\phi\rangle|}{\langle \phi, \phi\rangle} \right\}$$

for the case of the walk of Section 1. We will split the task into two parts. In the next section, we prove that for any ϕ in $\Omega \setminus \{0\}$ with $\pi^T\phi = 0$, we have

$$\frac{\langle \phi, P\phi\rangle}{\langle \phi, \phi\rangle} \ge -\left(1 - \frac{\delta^2}{6d^2}\right). \tag{13}$$

Clearly, for this it suffices to consider the ϕ satisfying $\langle \phi, P\phi\rangle < 0$. Then in the ensuing section, we prove the more difficult bounds:

$$\frac{\langle \phi, P\phi\rangle}{\langle \phi, \phi\rangle} \ge 1 - \lambda,$$

where $\lambda = \lambda_1$ (Theorem 1) or $\lambda = \lambda_2$ (Theorem 2).

3. Proof of (13): $\langle \phi, P\phi\rangle < 0$. We now return to the specific chain corresponding to our random walk. The following lemma considers the case where $\langle \phi, P\phi\rangle$ is negative

LEMMA 3. *$\phi \in \Omega$ implies*

$$\langle \phi, P\phi\rangle \ge -(1-\tau)\langle \phi, \phi\rangle,$$

where $\tau = \delta^2/6d^2$.

PROOF. Consider

$$\begin{aligned}
&\langle \phi, P\phi\rangle + (1-\tau)\langle \phi, \phi\rangle \\
&\quad = \sum_{x\in C}\Big(\sum_{y\in C} P(x,y)\phi(y)\Big)\phi(x)\pi(x) + \sum_{x\in C}(1-\tau)\phi(x)^2\pi(x) \\
&\quad = \sum_{x\in C}\sum_{y\neq x} P(x,y)\phi(y)\phi(x)\pi(x) + \sum_{x\in C}(1-P(x,x))\phi(x)^2\pi(x) \\
&\qquad + \sum_{x\in C}(2P(x,x)-\tau)\phi(x)^2\pi(x) \\
&\quad = \sum_{x\in C}\sum_{y\neq x}\big(P(x,y)\phi(y)\phi(x)\pi(x) + P(x,y)\phi(x)^2\pi(x)\big) \qquad (14)\\
&\qquad + \sum_{x\in C}(2P(x,x)-\tau)\phi(x)^2\pi(x) \\
&\quad = \sum_{\{x,y\}\in E}\pi(x)P(x,y)(\phi(x)+\phi(y))^2 \\
&\qquad + \sum_{x\in C}(2P(x,x)-\tau)\phi(x)^2\pi(x),
\end{aligned}$$

where $E = \{\{x,y\}: x \neq y$ and $P(x,y) > 0\}$ and the last equation follows by time-reversibility.

Suppose next that the walk is at $x \in C$. Let $P_i(x,x)$ denote the probability that the ith direction is chosen and that no movement is made at the current iteration. Thus

$$P_i(x,x) = \frac{1}{2n}\left(1 - \min\left\{1, \frac{\pi(x+e_i)}{\pi(x)}\right\} + 1 - \min\left\{1, \frac{\pi(x-e_i)}{\pi(x)}\right\}\right)$$

and, of course,

$$P(x,x) = \sum_{i=1}^{n} P_i(x,x).$$

We will show for each i, $i = 1,2,\ldots,n$, that with $E_i = \{(x,y)\in E: x-y = \pm\delta e_i\}$,

$$\begin{aligned}
&\sum_{\{x,y\}\in E_i}\pi(x)P(x,y)(\phi(x)+\phi(y))^2 \\
&\qquad + \sum_{x\in C}\left(2P_i(x,x) - \frac{\tau}{n}\right)\phi(x)^2\pi(x) \geq 0.
\end{aligned} \tag{15}$$

Fix i and consider the lines in $\mathbb{R}^n$ parallel to e_i which go through the centres C of the cubes $\mathscr{C}$. These induce a natural partition of C into $\mathscr{L}_i$, where each $L \in \mathscr{L}_i$ is the set of cube centres lying on some line. Now fix $L \in \mathscr{L}_i$ and suppose that

$$\begin{aligned}
L &= \{x, x+e_i, x+2e_i, \ldots, x+se_i\} \\
&= \{x^{(0)}, x^{(1)}, \ldots, x^{(s)}\}.
\end{aligned}$$

We break L into *maximal* contiguous segments $L_1, L_2, \dots$ of the form

$$\{x^{(k)}, x^{(k+1)}, \dots, x^{(l)}\}$$

such that

$$\pi(x^{(r+1)}) \geq \pi(x^{(r)})\left(1 - \frac{\delta}{d}\right)$$

for $r = k,\ k+1, \dots, l-1$. [A segment ends when $\pi(x^{(l+1)}) < \pi(x^{(l)})(1 - \delta/d)$, which includes $l = s$, as the "next" cube centre has $\pi(\cdot) = 0$.] Thus

$$(16) \qquad P(x^{(r)}, x^{(r+1)}) \geq \frac{1}{2n}\left(1 - \frac{\delta}{d}\right), \qquad r = k, k+1, \dots, l-1,$$

and

$$(17) \qquad \pi(x^{(r)}) \geq \theta\pi(x^{(r')}), \qquad k \leq r' \leq r \leq l,$$

where $\theta = (1 - \delta/d)^{(d/\delta)-1}$. Finally,

$$(18) \qquad P_i(x^{(l)}, x^{(l)}) \geq \frac{1}{2n}\frac{\delta}{d}.$$

We prove (15) by showing that

$$(19) \qquad \begin{aligned} &\sum_{r=k}^{l-1} \pi(x^{(r)})P(x^{(r)}, x^{(r+1)})\big(\phi(x^{(r)}) + \phi(x^{(r+1)})\big)^2 \\ &\quad + \sum_{r=k}^{l}\left(2P_i(x^{(r)}, x^{(r)}) - \frac{\tau}{n}\right)\phi(x^{(r)})^2\pi(x^{(r)}) \geq 0. \end{aligned}$$

Now for $k \leq r < l$,

$$\begin{aligned} \phi(x^{(r)}) = {} & \big(\phi(x^{(r)}) + \phi(x^{(r+1)})\big) - \big(\phi(x^{(r+1)}) + \phi(x^{(r+2)})\big) + \cdots \\ & \pm\big(\phi(x^{(l-1)}) + \phi(x^{(l)})\big) \pm \phi(x^{(l)}). \end{aligned}$$

Applying the Cauchy–Schwarz inequality twice, we get

$$\begin{aligned} &\phi(x^{(r)})^2\pi(x^{(r)}) \\ &\quad \leq 2\pi(x^{(r)})\left(\sum_{r'=r}^{l-1}|\phi(x^{(r')}) + \phi(x^{(r'+1)})|\right)^2 + 2\phi(x^{(l)})^2\pi(x^{(r)}) \\ &\quad \leq 2(l-r)\pi(x^{(r)})\sum_{r'=r}^{l-1}\big(\phi(x^{(r')}) + \phi(x^{(r'+1)})\big)^2 + 2\phi(x^{(l)})^2\pi(x^{(r)}) \\ &\quad \leq 4\theta^{-1}\left(1 - \frac{\delta}{d}\right)^{-1} n(l-r)\sum_{r'=r}^{l-1}\pi(x^{(r')})P(x^{(r')}, x^{(r'+1)}) \\ &\qquad\qquad \times\big(\phi(x^{(r')}) + \phi(x^{(r'+1)})\big)^2 + 2\theta^{-1}\phi(x^{(l)})^2\pi(x^{(l)}). \end{aligned}$$

Summing over $r = k, k+1, \dots, l-1$ and adding $\phi(x^{(l)})^2\pi(x^{(l)})$, we obtain

$$\sum_{r=k}^{l} \phi(x^{(r)})^2 \pi(x^{(r)}) \le 2\theta^{-1}\left(1 - \frac{\delta}{d}\right)^{-1} n\left(\frac{d}{\delta}\right)^2$$
$$\times \sum_{r=k}^{l-1} \pi(x^{(r)}) P(x^{(r)}, x^{(r+1)})\left(\phi(x^{(r)}) + \phi(x^{(r+1)})\right)^2$$
$$+ \left(1 + 2\theta^{-1}\frac{d}{\delta}\right)\phi(x^{(l)})^2 \pi(x^{(l)}).$$

Comparing this with what we want, that is, (19), we see that the following equations suffice:

$$2\theta^{-1}\left(1 - \frac{\delta}{d}\right)^{-1} \frac{\tau}{n}\left(\frac{d}{\delta}\right)^2 n \le 1$$

and

$$\frac{\tau}{n}\left(1 + 2\theta^{-1}\frac{d}{\delta}\right) \le 2P_i(x^{(l)}, x^{(l)}).$$

Under the assumption that $10\delta \le d$, we get that $\tau = \delta^2/6d^2$ satisfies the inequalities since we have (18). □

REMARK 3. The modified walk has $P(x, x) \ge 1 - \nu$ and the argument can be stopped at (14) provided $\tau \le 2(1 - \nu)$. We thus need $\nu \ge 1 - \hat{\lambda}_1/2$ in the modified walk. Note that we have $\hat{\lambda}_1$ rather than λ_1 because the modification introduces a factor of ν into our eigenvalue estimate; see Remark 4.

4. Reduction to a continuous problem.

We first introduce the quantity

$$E(\phi, \phi) = \langle \phi, \phi \rangle - \langle \phi, P\phi \rangle$$

for all $\phi \in \mathbb{R}^C$. The time-reversibility of P allows us to conclude

$$2E(\phi, \phi) = \sum_{x \in C} \sum_{y \in \Gamma(x)} (\phi(x) - \phi(y))^2 \pi(x) P(x, y), \tag{20}$$

where $\Gamma(x) = \{y \in C\colon P(x, y) < 0\}$.

To prove (20), write

$$\sum_{x \in C} \sum_{y \in C} \pi(x) P(x, y)(\phi(x) - \phi(y))^2$$
$$= \sum_{x \in C} \pi(x)\phi(x)^2 \sum_{y \in C} P(x, y) + \sum_{y \in C} \pi(y)\phi(y)^2 \sum_{x \in C} P(y, x)$$
$$- 2\sum_{x \in C} \pi(x)\phi(x) \sum_{y \in C} P(x, y)\phi(y).$$

In this notation we must now prove that

$$\inf_{\phi \in \Omega} \frac{E(\phi, \phi)}{\langle \phi, \phi \rangle} \geq \lambda.$$

Let 1_C denote the C-vector of 1's, $\bar{\phi} = \Sigma \pi(x)\phi(x)$ and let $\bar{\Omega} = \{\phi\colon \phi(x) = \bar{\phi}$ for $x \in S\}$. If $\phi \in \bar{\Omega}$, then $\phi - \bar{\phi}1_C \in \Omega$. Also, for all $\phi \in \mathbb{R}^N$ and for all $z \in \mathbb{R}$, we have

$$E(\phi - z1_N, \phi - z1_N) = E(\phi, \phi),$$

$$\langle \phi - \bar{\phi}1_C, \phi - \bar{\phi}1_C \rangle = \langle \phi, \phi \rangle - \bar{\phi}^2$$

and so we can redefine our objective as proving

$$\inf_{\substack{\phi \in \bar{\Omega} \\ \phi \text{ nonconstant}}} \frac{E(\phi, \phi)}{\langle \phi, \phi \rangle - \bar{\phi}^2} \geq \lambda. \tag{21}$$

For the rest of this section, we fix a particular nonconstant ϕ in $\tilde{\Omega}$. Let $T = \bigcup_{x \in C} C(x)$, which is "slightly" larger than K. Define $\tilde{F}\colon T \to \mathbb{R}$ as follows: Suppose $\xi \in C(x)$ for some $x \in C$. Then $\tilde{F}(\xi) = \bar{F}(x)$ if $\xi \notin \partial C(x)$ and $\tilde{F}(\xi) = 0$ otherwise.

Given $\phi \in \mathbb{R}^N$ and a small $\varepsilon > 0$, we define $\Phi_\varepsilon\colon T \to \mathbb{R}$ as follows. Suppose $z \in C(x)$ for some $x \in C$. Let $C(x, \varepsilon)$ denote the cube centred at x with side $\delta - 2\varepsilon$. If $z \in C(x, \varepsilon)$, we let $\Phi_\varepsilon(z) = \phi(x)$. If $z \notin C(x, \varepsilon)$, let D be a face of $C(x)$ which is closest to z. (If there is a tie for D, the value of Φ_ε does not matter, as we will see.) Suppose first that $D = C(x) \cap C(y)$ for some $y \in C$ and that $\text{dist}(z, D) = \eta\varepsilon$, where $0 \leq \eta < 1$. In this case, we let $\Phi_\varepsilon(z) = ((1 + \eta)\phi(x) + (1 - \eta)\phi(y))/2$. In this way, if we start at a point on a face of $C(x, \varepsilon)$ parallel to D and move toward D, then Φ_ε changes linearly from $\phi(x)$ to $\phi(y)$ over a distance 2ε. Finally, if the hypercube on the other side of D to $C(x)$ is not in $\mathscr{C}$, then we keep $\Phi_\varepsilon(z) = \phi(x)$.

Let

$$I_\varepsilon = \int_T |\nabla \Phi_\varepsilon(z)|^2 \tilde{F}(z)\, dz.$$

The function Φ_ε is not differentiable on a set Z of measure zero (consisting of points for which there is a tie for D). We can, however, easily "smooth out" Φ_ε close to Z so that (20) and (1) imply

$$\begin{aligned} \varepsilon I_\varepsilon &= \delta^{n-1} \sum_{x \in C} \sum_{y \in \Gamma(x)} \left(\frac{\phi(x) - \phi(y)}{2} \right)^2 \bar{F}(x) + O(\varepsilon) \\ &\leq \left(\frac{1 + \alpha}{4} \right) \delta^{n-1} \sum_{x \in C} \sum_{y \in \Gamma(x)} (\phi(x) - \phi(y))^2 \min\{\bar{F}(x), \bar{F}(y)\} + O(\varepsilon) \\ &= (1 + \alpha)\Delta n \delta^{n-1} E(\phi, \phi) + O(\varepsilon), \end{aligned}$$

where the hidden constant in $O(\varepsilon)$ may depend on n, F, ϕ.

REMARK 4. For the modified walk we need a factor of ν^{-1} in order to get the final equation.

On the other hand, the concavity of $\ln F$ implies that it is continuous, so for small enough ε, we have $\int_S F(\zeta)\,d\zeta \le (1+f(\varepsilon))\int_{C(x)\cap C(y)} F(\zeta)\,d\zeta$, where S is obtained by translating $C(x)\cap C(y)$ toward x by some $\varepsilon' \in [0,\varepsilon]$ and f is some function with $\limsup_{\varepsilon\to 0} f(\varepsilon) = 0$. This along with (2) implies that

$$\varepsilon I_\varepsilon \ge (1+f(\varepsilon))^{-1}(1+\alpha)^{-1}\varepsilon\int_T |\nabla\Phi_\varepsilon(z)|^2 F(z)\,dz - O(\varepsilon).$$

So

$$E(\phi,\phi) \ge \frac{\varepsilon\int_T |\nabla\Phi_\varepsilon(z)|^2 F(z)\,dz}{(1+\alpha)^2(1+f(\varepsilon))n\delta^{n-1}\Delta} - O(\varepsilon). \tag{22}$$

This deals with the numerator in (21). For the denominator,

$$\begin{aligned}\langle\phi,\phi\rangle - \bar{\phi}^2 &= \Delta^{-1}\sum_{x\in C}\phi(x)^2\bar{F}(x) - \bar{\phi}^2\\ &= \Delta^{-1}\sum_{x\in C}(\phi(x)-\zeta)^2\bar{F}(x) - (\bar{\phi}-\zeta)^2 \quad \text{for any } \zeta\in\mathbb{R}.\end{aligned} \tag{23}$$

We will use the foregoing equation with $\zeta = \mu$, where

$$\mu = \frac{\int_K \Phi_\varepsilon(z)F(z)\,dz}{\int_K F(z)\,dz}$$

and then let $\psi(z) = \Phi_\varepsilon(z) - \mu$. We will show in Section 5 that

$$\int_K \psi(z)^2 F(z)\,dz \le \varepsilon\kappa_0\int_K |\nabla\psi|^2 F(z)\,dz \tag{24}$$

and in Section 7 (for use in Theorem 2) that

$$\int_{T\setminus K}\psi(z)^2F(z)\,dz \le a\kappa_1\int_K\psi(z)^2F(z)\,dz + b\varepsilon\kappa_2\int_T|\nabla\psi|^2F(z)\,dz, \tag{25}$$

for any $a, b \ge 1$ satisfying $(a-1)(b-1)\ge 1$.

Consider first Theorem 1. It follows from (23) that

$$\begin{aligned}\langle\phi,\phi\rangle - \bar{\phi}^2 &= \Delta^{-1}\sum_{x\in C\setminus S}(\phi(x)-\mu)^2\bar{F}(x)\\ &\quad + \Delta^{-1}\sum_{x\in S}(\phi(x)-\mu)^2\bar{F}(x) - (\bar{\phi}-\mu)^2\\ &\le \frac{1+\alpha}{\delta^n\Delta}\int_K\psi(z)^2F(z)\,dz + O(\varepsilon)\end{aligned} \tag{26}$$

since $\phi = \overline{\phi}$ for $x \in S$. Thus, by (24),

$$\begin{aligned}\langle \phi, \phi \rangle - \overline{\phi}^2 &\le \frac{1+\alpha}{\delta^n \Delta} \varepsilon \kappa_0 \int_K |\nabla \psi(z)|^2 F(z)\, dz + O(\varepsilon) \\ &\le \frac{1+\alpha}{\delta^n \Delta} \kappa_0 (1+\alpha)^2 (1 + f(\varepsilon)) n \delta^{n-1} \Delta E(\phi, \phi) + O(\varepsilon).\end{aligned} \tag{27}$$

The second inequality comes from (22) and $K \subseteq T$. Now, (27) is true for all $\varepsilon > 0$ and so (21) follows in the case of Theorem 1.

Now consider Theorem 2. It follows from (23) that

$$\begin{aligned}\langle \phi, \phi \rangle - \overline{\phi}^2 &\le \frac{1+\alpha}{\delta^n \Delta} \int_T (\Phi_\varepsilon(z) - \mu)^2 F(z)\, dz + O(\varepsilon) \\ &= \frac{1+\alpha}{\delta^n \Delta} \left(\int_K \psi^2 F(x)\, dx + \int_{T \setminus K} \psi^2 F(z)\, dz \right) + O(\varepsilon) \\ &\le \frac{1+\alpha}{\delta^n \Delta} \Bigg((a\kappa_1 + 1) \int_K \psi^2 F(z)\, dz \\ &\qquad + b \varepsilon \kappa_2 \int_T |\nabla \Phi_\varepsilon(x)|^2 F(z)\, dz \Bigg) + O(\varepsilon) \\ &\le \frac{1+\alpha}{\delta^n \Delta} (\kappa_0(a\kappa_1 + 1) + b\kappa_2) \varepsilon \int_T |\nabla \Phi_\varepsilon(z)|^2 F(z)\, dz + O(\varepsilon) \\ &\le \frac{1+\alpha}{\delta^n \Delta} (\kappa_0(a\kappa_1 + 1) + b\kappa_2)(1+\alpha)^2 (1 + f(\varepsilon)) \\ &\qquad \times n \delta^{n-1} \Delta E(\phi, \phi) + O(\varepsilon).\end{aligned}$$

Substituting $a = 1\sqrt{\kappa_2/(\kappa_0 \kappa_1)}$ and $b = 1 + (a-1)^{-1}$, we obtain

$$\begin{aligned}\langle \phi, \phi \rangle - \overline{\phi}^2 &\le \frac{1+\alpha}{\delta^n \Delta} \left(\kappa_0 \kappa_1 + \kappa_0 + \kappa_2 + 2\sqrt{\kappa_0 \kappa_1 \kappa_2} \right) \\ &\qquad \times (1+\alpha)^2 (1 + f(\varepsilon)) n \delta^{n-1} \Delta E(\phi, \phi) + O(\varepsilon).\end{aligned}$$

This is true for all $\varepsilon > 0$ and so (21) follows for the case of Theorem 2.

5. Proof of (24). We will reduce the geometry to one dimension by applying the following *localisation lemma* of Lovász and Simonovits [7]:

LEMMA 4. *Let f_1, f_2 be upper semicontinuous functions defined on $\mathbb{R}^n$ such that*

$$\int_{\mathbb{R}^n} f_i(z)\, dz > 0, \qquad i = 1, 2.$$

Then there exist a, $b \in \mathbb{R}^n$ and a linear function l: $[0,1] \to \mathbb{R}_+$ such that

$$\int_{t=0}^1 f(ta + (1-t)b) l(t)^{n-1}\, dt > 0, \qquad i = 1, 2.$$

To apply this we replace (24) by the equivalent

$$\int_K \Phi_\varepsilon^2(z)F(z)\,dz - \frac{\left(\int_K \Phi_\varepsilon(z)F(z)\,dz\right)^2}{\int_K F(z)\,dz} \le \varepsilon\kappa_0 \int_K |\nabla\Phi_\varepsilon(z)|^2 F(z)\,dz. \tag{28}$$

Adding a constant to Φ_ε does not change either side of this inequality and so we can assume that $\Phi_\varepsilon(z)$ is positive on K. We then use the fact that under this assumption, (28) *fails* to hold if and only if there exists $\alpha > 0$ such that

$$\int_K \Phi_\varepsilon^2(z)F(z)\,dz - \alpha\int_K \Phi_\varepsilon(z)F(z)\,dz > \varepsilon\kappa_0\int_K |\nabla\Phi_\varepsilon(z)|^2 F(z)\,dz$$

and

$$-\int_K \Phi_\varepsilon(z)F(z)\,dz + \alpha\int_K F(z)\,dz > 0.$$

We apply Lemma 4 with

$$f_1 = \left(\Phi_\varepsilon^2 F - \alpha\Phi_\varepsilon F - \varepsilon\kappa_0|\nabla\Phi_\varepsilon(x)|^2 F\right)\chi_K$$

and

$$f_2 = (-\Phi_\varepsilon F + \alpha F)\chi_K.$$

(Here χ_K is the indicator function of the body K.)

Let a, b, l be as in Lemma 4. We observe that we can take $a,\ b \in K$ because of the factor χ_K. Let $g(t) = \Phi_\varepsilon((1-t)a + tb)$, $h(t) = F((1-t)a + tb)l(t)^{n-1}$, $\pi(t) = h(t)/\int_{\xi=0}^1 h(\xi)\,d\xi$ and $\tilde{g}(t) = |\nabla\Phi_\varepsilon((1-t)a + tb)|$. Note that $\pi(t)$ is log-concave.

We then see that if (28) fails to hold, then

$$\int_{t=0}^1 g(t)^2\pi(t)\,dt - \left(\int_{t=0}^1 g(t)\pi(t)\,dt\right)^2 \le \varepsilon\kappa_0\int_{t=0}^1 \tilde{g}(t)^2\pi(t)\,dt \tag{29}$$

fails to hold.

We can thus prove (24) by proving (29). We replace the LHS of (29) using the identity

$$\begin{aligned}\int_{t=0}^1 g(t)^2\pi(t)\,dt - \left(\int_{t=0}^1 g(t)\pi(t)\,dt\right)^2 \\ = \frac{1}{2}\int_{s=0}^1\int_{t=0}^1 (g(t) - g(s))^2\pi(s)\pi(t)\,ds\,dt.\end{aligned} \tag{30}$$

Let $\hat{d} = |b - a|$ and $u = (b-a)/\hat{d}$. Then, for $0 \le s \le t \le 1$,

$$\begin{aligned}(g(t) - g(s))^2 &= \left(\int_{\xi=s}^t g'(\xi)\,d\xi\right)^2 \\ &= \hat{d}^2\left(\int_{\xi=s}^t \sum_{j=1}^n u_j\chi_j(\xi)\frac{\partial\Phi_\varepsilon}{\partial x_j}((1-\xi)a + \xi b)\,d\xi\right)^2,\end{aligned}$$

where $\chi_j(\cdot)$ is defined by

$$\chi_j(\xi) = \begin{cases} 1, & \text{if } \dfrac{\partial \Phi_\varepsilon}{\partial x_j}((1-\xi)a + \xi b) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

By the Cauchy–Schwarz inequality, we have

$$\sum_{j=1}^{n} \left| u_j \chi_j(\xi) \frac{\partial \Phi_\varepsilon}{\partial x_j}((1-\xi)a + \xi b) \right| \leq \left(\sum_{j=1}^{n} u_j^2 \chi_j(\xi) \right)^{1/2} \tilde{g}(\xi).$$

So we get with another application of Cauchy–Schwarz,

$$\begin{aligned}(g(t) - g(s))^2 &\leq \hat{d}^2 \left(\int_{\xi=s}^{t} \left(\sum_{j=1}^{n} u_j^2 \chi_j(\xi) \right)^{1/2} \tilde{g}(\xi)\, d\xi \right)^2 \\ &\leq \hat{d}^2 \left(\int_{\xi=s}^{t} \sum_{j=1}^{n} u_j^2 \chi_j(\xi)\, d\xi \right) \left(\int_{\xi=s}^{t} \tilde{g}(\xi)^2\, d\xi \right).\end{aligned}$$

Now each time the line from s to t crosses a hyperplane of the form $x_j = m\delta$, m an integer, we get a contribution of $2\varepsilon/(\hat{d}u_j)$ to $\int \chi_j(\xi)\, d\xi$. Furthermore, the number of such crossings is at most

$$\frac{\hat{d}u_j(t-s)}{\delta} + 1. \tag{31}$$

So we get (using the facts that $\sum_{j=1}^{n} u_j^2 = 1$ and $\sum |u_j| \sqrt{n}$)

$$(g(t) - g(s))^2 \leq 2\hat{d}^2 \varepsilon \left(\frac{|t-s|}{\delta} + \frac{\sqrt{n}}{\hat{d}} \right) \int_{\xi=s}^{t} \tilde{g}(\xi)^2\, d\xi.$$

Thus if RHS (30) denotes the right hand side of (30), then

$$\text{RHS}(30) \leq \hat{d}^2 \varepsilon \int_{s=0}^{1} \int_{t=0}^{1} \left(\frac{|t-s|}{\delta} + \frac{\sqrt{n}}{\hat{d}} \right) \int_{\xi=s}^{t} \tilde{g}(\xi)^2\, d\xi\, \pi(t)\pi(s)\, dt\, ds$$

$$\begin{aligned} &= 2\hat{d}^2 \varepsilon \int_{\xi=0}^{1} \tilde{g}(\xi)^2 \pi(\xi) \\ &\quad \times \left(\int_{s=0}^{\xi} \int_{t=\xi}^{1} \left(\frac{t-s}{\delta} + \frac{\sqrt{n}}{\hat{d}} \right) \frac{\pi(s)\pi(t)}{\pi(\xi)}\, dt\, ds \right) d\xi, \end{aligned} \tag{32}$$

where the last equation is obtained by interchanging the order of integration, the factor of 2 coming from the fact that s, t are interchangable in the previous expression. Now let

$$\Pi = \left\{ \pi \in [0,1] \to \mathbb{R}_+ : \pi \text{ is log-concave and } \int_{t=0}^{1} \pi(t)\, dt = 1 \right\}.$$

Let

$$M_1 = \sup_{\substack{\pi \in \Pi \\ \xi \in [0,1]}} \int_{s=0}^{\xi} \int_{t=\xi}^{1} (t-s) \frac{\pi(s)\pi(t)}{\pi(\xi)}\, dt\, ds$$

and

$$M_2 = \sup_{\substack{\pi \in \Pi \\ \xi \in [0,1]}} \int_{s=0}^{\xi} \int_{t=\xi}^{1} \frac{\pi(s)\pi(t)}{\pi(\xi)}\, dt\, ds.$$

Then (32) implies

$$\text{RHS}(30) \le 2\hat{d}^2 \varepsilon \left(\frac{M_1}{\delta} + \frac{M_2 \sqrt{n}}{\hat{d}} \right) \int_{\xi=0}^{1} \tilde{g}(\xi)^2 \pi(\xi)\, d\xi. \tag{33}$$

We will prove in succeeding text that

$$M_1 = \tfrac{1}{8} \quad \text{and} \quad M_2 = \tfrac{1}{4}$$

and (24) follows immediately from (33).

6. Computation of M_1 and M_2. We begin with M_1. Rather than restrict ourselves to $\pi \in \Pi$, we prove that if $h\colon [0,1] \to \mathbb{R}_+$ is log-concave, then

$$\int_{s=0}^{\xi} \int_{t=\xi}^{1} (t-s) h(s) h(t)\, dt\, ds \le \frac{h(\xi)}{8} h(s)\, ds. \tag{34}$$

This will prove $M_1 \le 1/8$, which is what we want. We then note that $h = 1$ satisfies (34) with equality.

Since $\log h$ is concave, there exists $\alpha \in \mathbb{R}$ such that

$$h(s) \le h(\xi) e^{\alpha(s-\xi)}, \qquad s \in [0,1]. \tag{35}$$

Our aim first is to show that the extremal h satisfies (35) with equality for all $s \in [0,1]$, for some α. Let

$$I_0 = \int_{s=0}^{\xi} h(s)\, ds \quad \text{and} \quad I_1 = \int_{s=\xi}^{1} h(s)\, ds,$$

$$J_0 = \int_{s=0}^{\xi} (\xi - s) h(s)\, ds \quad \text{and} \quad J_1 = \int_{t=\xi}^{1} (t-\xi) h(t)\, dt.$$

Then (34) is equivalent to

$$\frac{I_0 J_1 + I_1 J_0}{I_0 + I_1} \le \frac{h(\xi)}{8}$$

or

$$\left(\frac{J_1}{I_1} + \frac{J_0}{I_0} \right) \frac{1}{I_0^{-1} + I_1^{-1}} \le \frac{h(\xi)}{8}.$$

Suppose now that we fix α and also $h(\xi) = \beta > 0$. Let $I_0(\alpha, \beta) = \int_{s=0}^{\xi} \beta e^{\alpha(s-\xi)}\, ds$ and let $I_1(\alpha, \beta)$, $J_0(\alpha, \beta)$ and $J_1(\alpha, \beta)$ be defined analogously. Clearly, under these circumstances $I_k \le I_k(\alpha,\beta)$ for $k = 0,1$ and so

$$\frac{1}{I_0^{-1} + I_1^{-1}} \le \frac{1}{I_0(\alpha,\beta)^{-1} + I_1(\alpha,\beta)^{-1}}. \tag{36}$$

We now prove that

$$\frac{J_k}{I_k} \le \frac{J_k(\alpha,\beta)}{I_k(\alpha,\beta)}, \qquad k = 0,1. \tag{37}$$

We prove (37) for $k = 1$, the proof for $k = 0$ being essentially the same. Let M be a positive integer and $\theta = (1-\xi)/M$. Let $g = \log h$ and let $g_l = g(\xi + l\theta)$, $l = 0,1,\dots,M$. The concavity of g implies that

$$g_{l+2} - g_{l+1} \le g_{l+1} - g_l, \qquad 0 \le l < M-2, \tag{38}$$

and (35) implies

$$g_l \le g_0 + al\theta, \qquad 0 \le l \le M. \tag{39}$$

Now choose $\eta > 0$ small. By choosing M sufficiently large, we can ensure that

$$\frac{J_1}{I_1} \le \frac{\sum_{l=0}^{M} l e^{g_l}}{\sum_{l=0}^{M} e^{g_l}} + \eta. \tag{40}$$

Denote the RHS of (4) by $\rho(g_0, g_1, \dots, g_M)$. Since η is arbitrary, we need only show that, subject to (38) and (39), ρ is maximised when $g_l = g_0 + \alpha l\theta$, that is,

$$\rho(g) \le \frac{\sum_{l=0}^{M} l e^{g_0 + \alpha l\theta}}{\sum_{l=0}^{M} e^{g_0 + \alpha l\theta}}. \tag{41}$$

This is easy to do by backward induction on

$$k = k(g) = \max\{m: g_l = g_0 + \alpha l\theta, 0 \le l \le m\}.$$

The base case $k = M$ is trivial and so assume that (41) holds for $k = \kappa + 1$ and assume that $k = \kappa$. Let

$$\nu = g_0 + \alpha\kappa\theta - g_\kappa > 0.$$

Define $\bar{g}$ by

$$\bar{g}_l = \begin{cases} g_l, & 0 \le l < \kappa, \\ g_l + \nu, & \kappa \le l \le M. \end{cases}$$

Note that $\bar{g}$ satisfies (38) and (39) and that $k(\bar{g}) = \kappa + 1$. Hence our inductive assumption implies that $\bar{g}$ satisfies (41). On the other hand, if $e^{\nu\theta} = 1 + \varepsilon$ and $h_l = e^{gl}$, then

$$\begin{aligned} \rho(\bar{g}) - \rho(g) &= \frac{\sum_{l=0}^{M} l h_l + \varepsilon \sum_{l=\kappa}^{M} l h_l}{\sum_{l=0}^{M} h_l + \varepsilon \sum_{l=\kappa}^{M} h_l} - \frac{\sum_{l=0}^{M} l h_l}{\sum_{l=0}^{M} h_l} \\ &= \frac{\varepsilon\left(\sum_{l=0}^{\kappa-1} h_l \sum_{l=\kappa}^{M} l h_l - \sum_{l=0}^{\kappa-1} l h_l \sum_{l=\kappa}^{M} hl\right)}{\left(\sum_{l=0}^{M} h_l + \varepsilon \sum_{l=\kappa}^{M} h_l\right)\sum_{l=0}^{M} h_l} \\ &\ge 0, \end{aligned}$$

since clearly

$$\frac{\sum_{l=\kappa}^{M} l h_l}{\sum_{l=\kappa}^{M} h_l} \geq \kappa > \frac{\sum_{l=0}^{\kappa-1} l h_l}{\sum_{l=0}^{\kappa-1} h_l}.$$

We can, therefore, assume from now on that $h(s) = \beta e^{\alpha(s-\xi)}$ for some α, β. We can, in fact, assume that $\beta = 1$ as β contributes a factor of β^2 to both sides of (34). So now let

$$\gamma(\alpha, \xi) = \frac{\int_{s=0}^{\xi} \int_{t=\xi}^{1} (t-s) e^{\alpha(s+t)}\, dt\, ds}{e^{\alpha\xi} \int_{s=0}^{1} e^{\alpha s}\, ds}$$

$$= \frac{e^{\alpha} + \xi - \xi e^{\alpha} - e^{\alpha(1-\xi)}}{\alpha(e^{\alpha} - 1)}.$$

We must show that $\gamma(\alpha, \xi) \leq 1/8$ for $\alpha \in \mathbb{R}$ and $\xi \in [0, 1]$. Now,

$$\frac{\partial \gamma}{\partial \xi} = \frac{1}{\alpha(e^{\alpha} - 1)} (1 - e^{\alpha} + \alpha e^{\alpha(1-\xi)})$$

and

$$\frac{\partial^2 \gamma}{\partial \xi^2} = \frac{1}{\alpha(e^{\alpha} - 1)} (-\alpha^2 e^{\alpha(1-\xi)}) < 0.$$

It follows that for fixed α, γ is maximised when

$$\xi = \xi^* = \frac{1}{\alpha} \log\left(\frac{\alpha}{1 - e^{-\alpha}} \right).$$

Now let

$$f(\alpha) = \gamma(\alpha, \xi^*(\alpha))$$

$$= \frac{1}{\alpha^2} \log\left(\frac{e^{\alpha} - 1}{\alpha} \right) - \frac{1}{\alpha^2} + \frac{1}{\alpha(e^{\alpha} - 1)}.$$

A simple computation yields

$$\lim_{\alpha \to 0} f(\alpha) = \frac{1}{8}$$

and that f is even, that is, $f(-\alpha) = f(\alpha)$ for all $\alpha \in \mathbb{R}$. Next let

$$g(\alpha) = \alpha^2 f(\alpha) - \frac{\alpha^2}{8}.$$

Then

$$g'(\alpha) = -\frac{e^{2\alpha}(\alpha - 2)^2 + e^{\alpha}(2\alpha^2 - 8) + (\alpha + 2)^2}{4\alpha(e^{\alpha} - 1)^2}.$$

By checking that the series expansion of the preceding numerator contains only nonnegative coefficients, we see that $g'(\alpha) \leq 0$ for $\alpha \geq 0$. Thus $g(\alpha) \leq g(0) = 0$ for $\alpha \geq 0$, and since f is even, this completes the proof that $M_1 = 1/8$.

For the computation of M_2 we need to show

$$\int_{s=0}^{\xi}\int_{t=\xi}^{1} h(s)h(t)\,dt\,ds \le \frac{h(\xi)}{4}\int_{s=0}^{1} h(s)\,ds.$$

Fixing $h(\xi) = \beta$ and assuming (35), this amounts to

$$\frac{I_0 I_1}{I_0 + I_1} \le \frac{\beta}{4}.$$

Thus we can once again assume that the extremal h satisfies (35) and that $\beta = 1$.

So now let

$$\begin{aligned}
\zeta(\alpha, \xi) &= \frac{\int_{s=0}^{\xi}\int_{t=\xi}^{1} e^{\alpha(s+t)}\,dt\,ds}{e^{\alpha\xi}\int_{s=0}^{1} e^{\alpha s}\,ds}\\
&= \frac{\int_{s=0}^{\xi} e^{\alpha s}\,ds \int_{t=\xi}^{1} e^{\alpha t}\,dt}{e^{\alpha\xi}\int_{s=0}^{1} e^{\alpha s}\,ds}\\
&= \frac{(e^{\alpha\xi} - 1)(e^{\alpha} - e^{\alpha\xi})}{\alpha e^{\alpha\xi}(e^{\alpha} - 1)}\\
&= \frac{e^{\alpha} - e^{\alpha\xi} - e^{\alpha(1-\xi)} + 1}{\alpha(e^{\alpha} - 1)}.
\end{aligned}$$

We need to show that $\xi(\alpha, \xi) \le 1/4$ for $\xi \in [0, 1]$. Now

$$\frac{\partial\zeta}{\partial\xi} = \frac{1}{(e^{\alpha} - 1)}\left(e^{\alpha(1-\xi)} - e^{\alpha\xi}\right)$$

and

$$\frac{\partial^2\zeta}{\partial\xi^2} = -\frac{\alpha}{(e^{\alpha} - 1)}\left(e^{\alpha(1-\xi)} + e^{\alpha\xi}\right) < 0.$$

Thus $\zeta(\alpha, \xi)$ is maximised at $\xi = 1/2$, independent of α. Now let

$$f(\alpha) = \zeta\left(\alpha, \frac{1}{2}\right) = \frac{(e^{\alpha/2} - 1)^2}{\alpha(e^{\alpha} - 1)}.$$

Then

$$\lim_{\alpha\to 0} f(\alpha) = 1/4$$

and if $g(\alpha) = \alpha f(\alpha) - \alpha/4$, then

$$g'(\alpha) = -\frac{(e^{\alpha/2} - 1)^2}{4(e^{\alpha/2} + 1)^2} \le 0.$$

This shows that $g(\alpha) \le g(0) = 0$ for $\alpha \ge 0$ and $g(\alpha) \ge g(0)$ for $\alpha \le 0$ and completes our proof that $M_2 = 1/4$.

7. Proof of (25). Recall assumptions (6)–(8). We introduce polar coordinates $(r, \theta) = (\theta_1, \theta_2, \dots, \theta_{n-1}))$ so that

$$x_j = r\sin(\theta_j)\prod_{i=1}^{j-1}\cos(\theta_i), \qquad j \in [n].$$

Then

$$(42) \quad \int_{T\setminus K}\psi(z)^2F(z)\,dz = \int_\theta\left(\int_{r=R}^{R_1} r^{n-1}\psi(r,\theta)^2F(r,\theta)\,dr\right)J(\theta)\,d\theta,$$

where $r^{n-1}J(\theta)$ is the Jacobian of the transformation. Similarly,

$$(43) \quad \int_B\psi(z)^2F(z)\,dz = \int_\theta\left(\int_{r=0}^{R} r^{n-1}\psi(r,\theta)^2F(r,\theta)\,dr\right)J(\theta)\,d\theta$$

and

$$(44) \quad \int_T|\nabla\psi|^2F(z)\,dz = \int_\theta\left(\int_{r=R}^{R_1(\theta)} r^{n-1}|\nabla\psi|^2F(r,\theta)\,dr\right)J(\theta)\,d\theta.$$

Now consider a fixed θ. Let $h(r) = r^{n-1}F(r,\theta)$ and $g(r) = \psi(r,\theta)$. Note that h is log-concave. Let u denote the vector of Euclidean length 1 in the direction θ. Then we can write

$$g'(r) = \sum_{j=1}^{n} u_j\chi_j\frac{\partial\psi}{\partial z_j},$$

where $\chi_j = \chi_j(r)$ is the indicator for $\partial\psi/\partial z_j \neq 0$.

We prove the following lemma, where $R' = R - \varsigma$.

LEMMA 5. *For any real $a, b \geq 1$ such that $(a-1)(b-1) \geq 1$,*

$$(45) \quad \int_{r=R}^{R_1} g(r)^2h(r)\,dr \leq a\kappa_1\int_{r=R'}^{R} g(r)^2h(r)\,dr + \varepsilon b\kappa_2\int_{r=R'}^{R_1}|\nabla\psi|^2h(r)\,dr.$$

Inequality (25) follows from (42)–(45).

PROOF OF LEMMA 5. Let $r \geq R$. We have, using the inequalities $g(r)^2 \leq a(g(r-\varsigma))^2 + b(g(r) - g(r-\varsigma))^2$ and $h(r) \leq \kappa_1 h(r-\varsigma)$,

$$(46) \quad \begin{aligned}\int_{r=R}^{R_1} g(r)^2h(r)\,dr &\leq a\kappa_1\int_{R'}^{R} g(r)^2h(r)\,dr\\ &\quad + b\int_{R'}^{R}(g(r+\varsigma) - g(r))^2h(r+\varsigma)\,dr.\end{aligned}$$

We will bound the second term. To do so, note that for any ζ and α,

$$(47) \quad \begin{aligned}|g(\zeta) - g(\alpha)| &= \left|\int_{\lambda=\alpha}^{\zeta} g'(\lambda)\,d\lambda\right|\\ &= \left|\int_{\lambda=\alpha}^{\zeta}\sum_{j=1}^{n} u_j\chi_j\frac{\partial\psi}{\partial z_j}\sqrt{h(\lambda)/h(\lambda)}\,d\lambda\right|.\end{aligned}$$

Now using Cauchy–Schwarz twice, we get

$$
\begin{aligned}
|g(\zeta)-g(\alpha)| &\le \int_{\lambda=\alpha}^{\zeta}\left(\sum_{j=1}^{n} u_j^2\chi_j^2 h(\lambda)^{-1}\right)^{1/2}\left(\sum_{j=1}^{n}\left(\frac{\partial\psi}{\partial z_j}\right)^2 h(\lambda)\right)^{1/2} d\lambda \\
&\le \left(\int_{\lambda=\alpha}^{\zeta}\sum_{j=1}^{n} u_j^2\chi_j^2 h(\lambda)^{-1}\,d\lambda\right)^{1/2}\left(\int_{\lambda=\alpha}^{\zeta}|\nabla\psi|^2 h(\lambda)\,d\lambda\right)^{1/2}.
\end{aligned} \tag{48}
$$

We wish to apply this to (46) with $\zeta = r + \varsigma$ and $\alpha = r$, where $r \in [R', R]$. Then as λ runs from r to $r + \varsigma$, we have $\kappa_1 h(\lambda) \ge h(r + \varsigma)$, so we have

$$
(g(r+\varsigma)-g(r))^2 h(r+\varsigma) \le \kappa_1\left(\int_{\lambda=r}^{r+\varsigma}\sum_{j=1}^{n} u_j^2\chi_j^2\,d\lambda\right)\left(\int_{\lambda=r}^{r+\varsigma}|\nabla\psi|^2 h(\lambda)\,d\lambda\right).
$$

We bound the first factor on the right hand side as in (31) to obtain

$$
\int_{\lambda=r}^{r+\varsigma}\sum_{j=1}^{n} u_j^2\chi_j^2\,d\lambda \le 2\varepsilon\left(\frac{\varsigma}{\delta}+\sqrt{n}\right).
$$

So we get

$$
\int_{R'}^{R}(g(r+\varsigma)-g(r))^2 h(r+\varsigma)\,dr \le 2\varepsilon\kappa_1\left(\frac{l_0}{\delta}+\sqrt{n}\right)\int_{R'}^{R_1}|\nabla\psi|^2 h(\lambda)\,d\lambda.
$$

Plugging this into (46), we get the lemma. □

8. Conclusion. We have given good estimates of the second eigenvalue of the transition matrices of our chains. It would be of great interest if we could replace the l_2 diameter d by the smaller l_∞ diameter.

REFERENCES

[1] ANDERSON, W. N. and MORLEY, T. D. (1985). Eigenvalues of the Laplacian of a graph. *Linear and Multilinear Algebra* **18** 141–145.

[2] APPLEGATE, D. and KANNAN, R. (1991). Sampling and integration of log-concave functions. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing* 156–163. Assoc. Comput. Mach., New York.

[3] DYER, M. E., FRIEZE, A. M. and KANNAN, R. (1991). A random polynomial time algorithm for approximating the volume of convexbodies, *J. Assoc. Comput. Mach.* **38** 1–17.

[4] DYER, M. E., KANNAN, R. and MOUNT J. (1993). Random sampling of contingency tables. Unpublished manuscript.

[5] 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.

[6] LOVÁSZ, L. and SIMONOVITS, M. (1990). Mixing rate of Markov chains, an isoperimetric inequality and computing the volume. In *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science* 346–355. IEEE, New York.

[7] LOVÁSZ, L. and SIMONOVITS, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures and Algorithms.* To appear.

[8] METROPOLIS, N., ROSENBERG, A., ROSENBLUTH, R., TELLER, A. and TELLER, E. (1953). Equation of state calculation by fast computing machines. *J. Chem. Phys.* **21** 1087–1092.

[9] MIHAIL, M. (1989). Combinatorial aspects of expanders. Ph.D. dissertation, Dept. Computer Science, Harvard Univ.

[10] SINCLAIR, A. J. and JERRUM, M. R. (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. Comput.* **82** 93–133.

[11] TIERNEY, L. (1994). Exploring posterior distributions using Markov chains. *Ann. Statist.* To appear.

ALAN FRIEZE
DEPARTMENT OF MATHEMATICS
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213

RAVI KANNAN
BELLCORE
MORRISTOWN, NEW JERSEY 07962
AND
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213

NICK POLSON
GRADUATE SCHOOL OF BUSINESS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS 60637