

JOHN'S WALK

ADAM GUSTAFSON,* *Microsoft Corporation*HARIHARAN NARAYANAN ,** *TIFR Mumbai*

Abstract

We present an affine-invariant random walk for drawing uniform random samples from a convex body $\mathcal{K} \subset \mathbb{R}^n$ that uses maximum-volume inscribed ellipsoids, known as John's ellipsoids, for the proposal distribution. Our algorithm makes steps using uniform sampling from the John's ellipsoid of the symmetrization of \mathcal{K} at the current point. We show that from a warm start, the random walk mixes in $\tilde{O}(n^7)$ steps, where the log factors hidden in the \tilde{O} depend only on constants associated with the warm start and desired total variation distance to uniformity. We also prove polynomial mixing bounds starting from any fixed point x such that for any chord pq of \mathcal{K} containing x , $\left| \log \frac{|p-x|}{|q-x|} \right|$ is bounded above by a polynomial in n .

Keywords: Random walk; convex body; John's ellipsoid

2020 Mathematics Subject Classification: Primary 60J05

Secondary 52-08

1. Introduction

Drawing random samples from a convex body in $\mathcal{K} \subset \mathbb{R}^n$ is an important problem for volume computation and optimization which has generated a large body of research. Usually \mathcal{K} is specified by a membership oracle which certifies whether or not a test point $x \in \mathbb{R}^n$ is contained in \mathcal{K} . Given such an oracle, geometric random walks are then used to explore \mathcal{K} in such a way that, after a sufficient number of steps, the walk has 'mixed', in the sense that the current point is suitably close to a point uniformly drawn from \mathcal{K} in terms of statistical distance. To use such walks, an assumption that $\mathcal{B}(r) \subset \mathcal{K} \subset \mathcal{B}(R)$ is often made, where $\mathcal{B}(r)$ represents the Euclidean ball of radius $r > 0$. One common example of such geometric walks is the ball walk, which generates the next point by uniformly randomly sampling from a ball of radius $\delta \leq r/\sqrt{n}$ centered at the current point, and mixes in $\tilde{O}(n(R^2/\delta^2))$ steps from an M -warm start (i.e., the starting distribution has a density bounded above by a constant M [5]). Here and elsewhere, the $\tilde{O}(\cdot)$ notation suppresses polylogarithmic factors as well as constants depending only on the error parameters. Another example is the hit-and-run walk, where the next point is chosen uniformly at random from a random chord in \mathcal{K} which intersects the current point. Hit-and-run mixes in $O(n^3(R/r)^2 \log(R/(d\epsilon)))$, where the starting point is at distance d from the boundary and ϵ is the desired distance to stationarity [11]. Affine-invariant walks (i.e., geometric walks whose mixing time is invariant to such affine transformations) are another class of random

Received 22 July 2020; revision received 8 June 2022.

* Email address: adam.marc.gustafson@gmail.com

** Email address: hariharan.narayanan@tifr.res.in

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

walks which avoid the problem of rounding. One such random walk is known as the Dikin walk [6], which uses uniform sampling from Dikin ellipsoids to make steps. Given a polytope with m inequality constraints, the Dikin walk mixes in $\tilde{O}(mn)$ steps from a warm start. This random walk was extended to general convex bodies equipped with a ν -self-concordant barrier in [13], and mixes in $\tilde{O}(n^3\nu^2)$ steps from a warm start. For the case of a polytope, this implies that the Dikin walk equipped with the Lee–Sidford (LS) barrier [8] mixes in $\tilde{O}(n^5)$ steps from a warm start, though at each step one must additionally compute the LS barrier, which requires $O(nnz(A) + n^2)$ arithmetic operations, where $nnz(A)$ is the number of non-zero entries in the matrix A which defines the polytope. A significantly improved analysis of this walk was performed by [2], whose algorithm reaches a total variation distance of ϵ from the uniform measure in $O\left(n^{2.5} \log^4\left(\frac{2m}{n}\right) \log\left(\frac{M}{\epsilon}\right)\right)$ steps from an M -warm start. Very recently, it was shown in [7] that for certain ‘strongly self-concordant’ barriers there is a Dikin walk that mixes in $\tilde{O}(n\tilde{\nu})$, where $\tilde{\nu}$ is related to the self-concordance parameter of the barrier.

In this paper we introduce another affine-invariant random walk akin to the Dikin walk which uses uniform sampling from John’s ellipsoids of a certain small radius of appropriately symmetrized convex sets to make steps. We show that this walk mixes to achieve a total variation distance ϵ in $O(n^7 \log \epsilon^{-1})$ steps from a warm start. The type of convex body \mathcal{K} is not specified (i.e., need not be a polytope) in our analysis of the mixing time, but one must have access to the John’s ellipsoid of the current symmetrization of the convex body. While this dependence on the dimension is admittedly steep, a significant feature of this walk is that its mixing time from a warm start, or alternatively a ‘central point’ such as the center of mass, can be bounded above by a quantity that has absolutely no dependence on any parameter associated with the body apart from its dimension.

Notation: We will denote a large universal constant by C and a small universal constant by c . Let π_0 be a measure that is absolutely continuous with respect to the uniform measure π on \mathcal{K} . Let π_t denote the measure after t steps of the Markov chain.

Our main theorems at the end of this paper are the following.

Theorem 1: Let $r = cn^{-\frac{5}{2}}$ be the radius of the John’s ellipsoids in the walk. Let $\epsilon > 0$ and $M = \sup \frac{\pi_0(A)}{\pi(A)}$. After $t(\epsilon) = Cn^7 \log(M/\epsilon)$ steps of John’s walk, we have $d_{TV}(\pi_{t(\epsilon)}, \pi) \leq \epsilon$.

Theorem 2: Let $r = cn^{-\frac{5}{2}}$ be the radius of the John’s ellipsoids in the walk. For all chords pq of \mathcal{K} containing x , assume $\frac{|p-x|}{|q-x|} \in (\eta, \eta^{-1})$ for some parameter $0 < \eta < 1$ that measures the centrality of x in \mathcal{K} . Then there is a random geometrically distributed time τ with mean bounded above by C such that for $\epsilon > 0$, after $t(\epsilon) + \tau = Cn^7 (n \log(\sqrt{n}/(r\eta)) + \log(1/\epsilon)) + \tau$ steps of John’s walk starting at x , we have $d_{TV}(\pi_{t(\epsilon)+\tau}, \pi) \leq \epsilon$.

It is known that for the center of mass, $\eta \geq \frac{c}{n}$.

2. John’s walk

In this section, we describe John’s maximum-volume ellipsoid for a convex body $\mathcal{K} \subset \mathbb{R}^n$, and describe a geometric random walk using such ellipsoids. We begin by reviewing John’s theorem and some implications of the theorem.

2.1. John’s theorem

Fritz John showed that any convex body contains a unique ellipsoid of maximal volume, and characterized the ellipsoid [4, 1]. Without loss of generality, we may assume that the ellipsoid

of maximal volume is the unit Euclidean ball $\mathcal{B} \subset \mathbb{R}^n$, since this is the case after an affine transformation. John's theorem as stated for the unit ball case is as follows.

Theorem 1. (John's theorem.) *Each convex body $\mathcal{K} \subset \mathbb{R}^n$ contains a unique ellipsoid of maximal volume. The ellipsoid is \mathcal{B} if and only if the following conditions are satisfied: $\mathcal{B} \subset \mathcal{K}$, and for some $m \geq n$ there are Euclidean unit vectors $\{u_i\}_{i=1}^m$ on the boundary of \mathcal{K} and positive constants $\{c_i\}_{i=1}^m$ satisfying*

$$\sum_{i=1}^m c_i u_i = 0, \tag{2.1}$$

$$\sum_{i=1}^m c_i u_i u_i^T = I_n, \tag{2.2}$$

where I_n denotes the identity matrix in $\mathbb{R}^{n \times n}$.

Note that the condition (2.2) is sometimes written equivalently as

$$\langle x, y \rangle = \sum_{i=1}^m c_i \langle u_i, x \rangle \langle u_i, y \rangle$$

for all $x, y \in \mathbb{R}^n$. Using the cyclic invariance of the trace and that the $\{u_i\}$ are unit vectors, (2.2) implies that

$$\sum_{i=1}^m c_i = n, \tag{2.3}$$

a property we employ in the subsequent analysis.

We now enumerate some properties from [1] which provide additional insight into the geometric properties of John's ellipsoids and are useful for the analysis in subsequent sections. Note that the condition (2.1) implies that all the contact points do not lie in one half-space of the unit ball, and this condition is redundant in the symmetric case, since for every contact point u_i , its reflection about the origin $-u_i$ is also a contact point. The condition (2.2) guarantees that such contact points do not lie close to a proper subspace. Furthermore, there are at most $n(n + 3)/2$ contact points for general \mathcal{K} , and $n(n + 1)/2$ non-redundant contact points if \mathcal{K} is origin-symmetric [3]. At each u_i , the supporting hyperplane to \mathcal{K} is unique and orthogonal to u_i , since this is the case for the unit ball. Thus, considering the polytope resulting from such supporting hyperplanes, $\mathcal{P} = \{x \in \mathbb{R}^n \mid \langle x, u_i \rangle \leq 1, i = 1, \dots, m\}$, the convex set \mathcal{K} obeys the sandwiching $\mathcal{B} \subset \mathcal{K} \subset \mathcal{P}$. By Cauchy-Schwarz, for any $x \in \mathcal{P}$, we have

$$-|x| \leq \langle u_i, x \rangle \leq 1.$$

Since the weights $\{c_i\}$ are positive, it follows from applying the conditions (2.1), (2.2), and (2.3) that

$$\begin{aligned} 0 &\leq \sum_{i=1}^m c_i (1 - \langle u_i, x \rangle)(|x| + \langle u_i, x \rangle) \\ &= |x| \sum_{i=1}^m c_i + (1 - |x|) \left\langle \sum_{i=1}^m c_i u_i, x \right\rangle - \sum_{i=1}^m c_i \langle u_i, x \rangle^2 \\ &= n|x| - |x|^2, \end{aligned}$$

Algorithm 2.1: John's walk step

Given $x \in \mathcal{K}_x^s$, $r > 0$, and \mathcal{E}_x , generate the next step y as follows:

1. Toss a fair coin. If the result is heads, let $y = x$.
2. If the result is tails:
 - (a) Draw a uniformly distributed random point z from $\mathcal{E}_x(r)$.
 - (b) Compute \mathcal{E}_z using \mathcal{K}_z^s .
 - (c) If $x \notin \mathcal{E}_z(r)$, let $y = x$. Otherwise, let

$$y = \begin{cases} z, & \text{with probability } \min\left(1, \frac{\lambda(\mathcal{E}_x(r))}{\lambda(\mathcal{E}_z(r))}\right) = \min\left(1, \frac{\det E_x}{\det E_z}\right), \\ x, & \text{else.} \end{cases}$$

from which it follows that $|x| \leq n$. If the convex body is origin-symmetric, then by substituting $-u_i$ for u_i , for any $x \in \mathcal{P}$, we have

$$|\langle u_i, x \rangle| \leq 1.$$

It follows that

$$|x|^2 = \sum_{i=1}^m c_i \langle u_i, x \rangle^2 \leq \sum_{i=1}^m c_i = n,$$

so $|x| \leq \sqrt{n}$. Therefore, if \mathcal{K} is origin-symmetric and the unit ball is the John's ellipsoid, then the containment is

$$\mathcal{B} \subset \mathcal{K} \subset \sqrt{n}\mathcal{B}. \quad (2.4)$$

2.2. The John's walk algorithm

We state the algorithm for a general convex body \mathcal{K} . At a given point $x \in \mathcal{K}$, let the symmetrization of \mathcal{K} at x be

$$\mathcal{K}_x^s \equiv \mathcal{K} \cap \{2x - y \mid y \in \mathcal{K}\},$$

and for some symmetric positive definite matrix E_x , let

$$\mathcal{E}_x = \{E_x u + x \mid |u| \leq 1\} \quad (2.5)$$

denote the John's ellipsoid of \mathcal{K}_x^s . Similarly, let the rescaled John's ellipsoid be $\mathcal{E}_x(r) = \{r(E_x u) + x \mid |u| \leq 1\}$, where the radius $r > 0$ will be specified in Section 3. Assume $0 = x_0 \in \text{int}(\mathcal{K})$, and we have computed \mathcal{E}_{x_0} . To generate a sample x_i given x_{i-1} , we use Algorithm 2.1, where $\lambda(\cdot)$ denotes the Lebesgue measure on \mathbb{R}^n .

Algorithm 2.1 is a Metropolis–Hastings geometric random walk which uses the uniform measure $Q_x(\cdot)$ on the dilated John's ellipsoid $\mathcal{E}_x(r)$ as the proposal distribution. Tossing a fair coin ensures that the transition probability kernel defined by the algorithm is positive definite, which is known as making the walk lazy. Lazy random walks have the same stationary distribution as the original walk, at the cost of a constant increase in mixing time (we will analyze the

non-lazy walk, noting that the mixing time is not affected in terms of complexity as a function of m and n). The rejection of any sample y such that $x \notin \mathcal{E}_y(r)$ is necessary to ensure that the random walk is reversible.

The uniform measure on the John's ellipsoid $\mathcal{E}_x(r)$ is absolutely continuous with respect to the Lebesgue measure λ , and thus the Radon–Nikodym derivative (i.e., density) for the proposal distribution is

$$\begin{aligned}
 q_x(y) &\equiv \frac{dQ_x}{d\lambda}(y) \\
 &= \left(\frac{1}{\lambda(\mathcal{E}_x(r))} \right) \cdot 1_{\{y \in \mathcal{E}_x(r)\}}.
 \end{aligned}
 \tag{2.6}$$

The acceptance probability corresponding to the uniform stationary measure in the Metropolis filter is

$$\alpha_x(y) = \min \left[1, \frac{\lambda(\mathcal{E}_x(r))}{\lambda(\mathcal{E}_y(r))} \right].$$

By the Lebesgue decomposition, the transition probability measure $P_x(\cdot)$ of the non-lazy version of Algorithm 2.1 is absolutely continuous with respect to the measure

$$\mu \equiv \lambda + \delta_x,
 \tag{2.7}$$

where $\delta_x(\cdot)$ is the Dirac measure at x corresponding to a rejected move. The transition density is thus

$$\begin{aligned}
 p_x(y) &\equiv \frac{dP_x}{d\mu}(y) \\
 &= \alpha_x(y)q_x(y)1_{\{y \neq x, x \in \mathcal{E}_y(r)\}} + \rho(x)1_{\{y=x\}} \\
 &= \min \left[\frac{1}{\lambda(\mathcal{E}_x(r))}, \frac{1}{\lambda(\mathcal{E}_y(r))} \right] 1_{\{y \neq x, x \in \mathcal{E}_y(r), y \in \mathcal{E}_x(r)\}} + \rho(x)1_{\{y=x\}},
 \end{aligned}
 \tag{2.8}$$

where $1_{\{\cdot\}}$ is the indicator function and the rejection probability is denoted by $\rho(x)$. Following the observation below, we analyze the mixing time of the walk.

2.2.1. *An observation.* As a consequence of John's theorem, $\mathcal{E}_x = \mathcal{B}$ for $x = 0 \in \mathcal{K}$ if and only if the following conditions are satisfied: $\mathcal{B} \subset \mathcal{K}$, and for some $m \geq n$ there are Euclidean unit vectors $\{u_i\}_{i=1}^m$ on the boundary of \mathcal{K} and positive constants $\{c_i\}_{i=1}^m$ satisfying

$$\sum_{i=1}^m c_i u_i u_i^T = I_n,
 \tag{2.9}$$

where I_n denotes the identity matrix in $\mathbb{R}^{n \times n}$. Note that we have done away with the first condition $\sum_i c_i u_i = 0$, because not all contact points with \mathcal{K}_x^s are contact points with \mathcal{K} . We are forced to place a zero weight on such vectors, and this is possible, because in such a case, $-u_i$ is a contact point with both \mathcal{K} and \mathcal{K}_x^s .

3. Analysis of mixing time

In what follows we let a discrete-time, homogeneous Markov chain be the triple $\{\mathcal{K}, \mathcal{A}, P_x(\cdot)\}$ along with a distribution P_0 for the starting point, where the sample space is the convex body $\mathcal{K} \subset \mathbb{R}^n$, the measurable sets on \mathcal{K} are denoted by \mathcal{A} , and $P_x(\cdot)$ denotes the

transition measure for any $x \in \mathcal{K}$. We say that π is a stationary distribution of $\{\mathcal{K}, \mathcal{A}, P_x(\cdot)\}$ if for all $A \in \mathcal{A}$,

$$\int_{\mathcal{K}} \int_A P_x(y)\pi(dx) = \pi(A).$$

We say that $\{\mathcal{K}, \mathcal{A}, P_x(\cdot)\}$ is reversible with respect to the stationary distribution π if for all $A, A' \in \mathcal{A}$,

$$\int_{A'} \int_A P_x(y)\pi(dx) = \int_A \int_{A'} P_y(x)\pi(dy).$$

3.1. Conductance and mixing times

We use the approach from [10] of lower-bounding the conductance of the chain to prove mixing times. The conductance is defined as follows.

Definition 1. (*Conductance.*) Let P be a discrete-time homogenous Markov chain with kernel $P_x(\cdot)$ that is reversible with respect to the stationary measure $\pi(\cdot)$. Given $A \in \mathcal{A}$ with $0 < \pi(A) < 1$, the conductance of A is defined as

$$\phi(A) \equiv \frac{\Phi(A)}{\min\{\pi(A), \pi(\mathcal{K} \setminus A)\}},$$

where $\Phi(A) \equiv \int_A P_u(\mathcal{K} \setminus A) d\pi(u)$. The conductance of the chain is defined as

$$\phi \equiv \inf\{\phi(A) \mid A \in \mathcal{A}, 0 < \pi(A) \leq 1/2\}.$$

Recall that the total variation distance between two measures P_1, P_2 on a measurable space $(\mathcal{K}, \mathcal{A})$ is

$$d_{TV}(P_1, P_2) = \sup_{A \in \mathcal{A}} |P_1(A) - P_2(A)|.$$

Note that $|P_1(A) - P_2(A)| = |P_1(\mathcal{K} \setminus A) - P_2(\mathcal{K} \setminus A)|$, so if the supremum is attained on any $A \in \mathcal{A}$, then it is attained on $\mathcal{K} \setminus A \in \mathcal{A}$ as well. If P_1 and P_2 are both absolutely continuous with respect to a dominating measure μ and thus have densities $p_1 \equiv \frac{dP_1}{d\mu}$ and $p_2 \equiv \frac{dP_2}{d\mu}$, respectively, the total variation distance may also be written as

$$\begin{aligned} d_{TV}(P_1, P_2) &= \frac{1}{2} \int |p_1 - p_2| d\mu \\ &= 1 - \int \min(p_1, p_2) d\mu \\ &= 1 - \int_{S_1} \left[\min\left(1, \frac{p_2}{p_1}\right) \right] p_1 d\mu \\ &= 1 - \mathbb{E}_{P_1} \left[\min\left(1, \frac{p_2}{p_1}\right) \right], \end{aligned} \tag{3.1}$$

where $S_1 = \{x \mid p_1(x) > 0\}$. Recall that (3.1) does not depend on the choice of dominating measure μ , but that, rather, the densities are correctly specified with respect to the dominating measure. Additionally, note that the equality is attained on $\{x \mid p_1(x) \geq p_2(x)\}$ almost everywhere with respect to μ (or alternatively on its complement). The following relationship

between conductance and the total variation distance to the stationary measure was proven in [10].

Theorem 2. (Lovász and Simonovits.) *Let π_0 be the initial distribution for a lazy, reversible Markov chain with conductance ϕ and stationary measure π , and let π_t denote the distribution after t steps. Let π_0 be an M -warm start for π ; i.e., we have $M = \sup_{A \in \mathcal{A}} \frac{\pi_0(A)}{\pi(A)}$. Then*

$$d_{TV}(\pi_t, \pi) \leq \sqrt{M} \left(1 - \frac{\phi^2}{2}\right)^t. \tag{3.2}$$

As a consequence, we have the following bound on the mixing time.

Corollary 1. *Given $\epsilon > 0$ and $M \equiv \sup \frac{\pi_0(A)}{\pi(A)}$, after $t(\epsilon) \equiv \lceil \frac{2}{\phi^2} \log(\sqrt{M}/\epsilon) \rceil$ steps of the chain, we have $d_{TV}(\pi_{t(\epsilon)}, \pi) \leq \epsilon$. Thus the Markov chain mixes in $\tilde{O}(\phi^{-2})$ steps from a warm start.*

To find mixing times, it then suffices to lower-bound the conductance ϕ .

3.2. Isoperimetry

The typical means by which one finds lower bounds on the conductance is via isoperimetric inequalities. We first restate the cross-ratio used in the isoperimetric inequality we will employ.

Definition 2. (Cross-ratio.) *Let $x, y \in \mathcal{K}$, and let p, q be the end points of a chord in \mathcal{K} passing through x, y , such that x lies in between p and y . The cross-ratio is defined to be*

$$\sigma(x, y) = \frac{|x - y||p - q|}{|p - x||y - q|},$$

where $|\cdot|$ denotes the Euclidean norm.

Additionally, for any $S_1, S_2 \subset \mathcal{K}$, let

$$\sigma(S_1, S_2) = \inf_{x \in S_1, y \in S_2} \sigma(x, y).$$

In [9], Lovász proved an isoperimetric inequality involving the cross-ratio, from which the conductance ϕ may be lower-bounded for the special case of the uniform distribution on a convex body $\mathcal{K} \subset \mathbb{R}^n$. This was extended to log-concave measures by Lovász and Vempala in [12], of which the uniform measure on a convex body is a special case. We state the latter result as follows.

Theorem 3. (Lovász and Vempala.) *For any log-concave probability measure $\pi(\cdot)$ supported on \mathcal{K} and a partition of \mathcal{K} into measurable subsets S_1, S_2 , and $S_3 = \mathcal{K} \setminus (S_1 \cup S_2)$, we have*

$$\pi(S_3) \geq \sigma(S_1, S_2)\pi(S_1)\pi(S_2). \tag{3.3}$$

3.3. Mixing of John's walk

The key step in proving conductance lower bounds is to show that if two points x and y are close in geometric distance, then the total variation distance between the corresponding marginals P_x and P_y is bounded away from 1 by a controlled quantity. Note that given John's ellipsoid $\mathcal{E}_x = \{E_x u + x \mid |u| \leq 1\}$, where E_x is symmetric and positive definite, a norm is induced via

$$\|y - x\|_x^2 = (y - x)^T E_x^{-2} (y - x).$$

We first relate this norm to the cross-ratio as follows.

Theorem 4. Let $\|\cdot\|_x$ denote the norm induced by the John’s ellipsoid of \mathcal{K}_x^s . Then

$$\sigma(x, y) \geq \frac{1}{\sqrt{n}} \|y - x\|_x.$$

Proof. Noting that the cross-ratio is invariant to affine transformations, without loss of generality we may assume by a suitable affine transformation that the John’s ellipsoid of \mathcal{K}_x^s is the unit ball, and thus $\|y - x\|_x = |y - x|$. Let p, x, y, q denote successive points on a chord through \mathcal{K}_x^s . Then

$$\begin{aligned} \sigma(x, y) &= \frac{|x - y||p - q|}{|p - x||y - q|} \\ &\geq \frac{|x - y| (|p - x| + |y - q|)}{|p - x||y - q|} \\ &\geq \max\left(\frac{|x - y|}{|y - q|}, \frac{|x - y|}{|p - x|}\right) \\ &= \frac{|x - y|}{\min(|y - q|, |p - x|)} \\ &\geq \frac{|x - y|}{\sqrt{n}}, \end{aligned}$$

where the last inequality follows from the containment

$$\mathcal{B} \subset \mathcal{K} \subset \sqrt{n}\mathcal{B}$$

(see (2.4)) and the observation that either $|x - p|$ or $|y - q|$ is contained inside \mathcal{K}_x^s . □

Before bounding the statistical distance between P_x and P_y given a bound on the geometric distance between x and y , we first state some useful lemmas regarding the ellipsoids \mathcal{E}_x and \mathcal{E}_y . The next lemma is a generalization of the Cauchy–Schwarz inequality to semidefinite matrices.

Lemma 1. (Semidefinite Cauchy–Schwarz.) Let $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ and let $A_1, \dots, A_m \in \mathbb{R}^{r \times n}$. Then

$$\left(\sum_{i=1}^m \alpha_i A_i\right) \left(\sum_{i=1}^m \alpha_i A_i\right)^T \preceq \left(\sum_{i=1}^m \alpha_i^2\right) \left(\sum_{i=1}^m A_i A_i^T\right), \tag{3.4}$$

where $A \preceq B$ signifies that $B - A$ is positive semidefinite.

Proof. The proof is as in Lemma 3.11 in [6]. For all i and j ,

$$(\alpha_j A_i - \alpha_i A_j)(\alpha_j A_i - \alpha_i A_j)^T \succeq 0.$$

Thus

$$\begin{aligned} 0 &\preceq \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\alpha_j A_i - \alpha_i A_j)(\alpha_j A_i - \alpha_i A_j)^T \\ &= \frac{1}{2} \sum_{i=1}^m \left[\left(\sum_{j=1}^m \alpha_j^2\right) A_i A_i^T - \alpha_i A_i \sum_{j=1}^m (\alpha_j A_j^T) - \left(\sum_{j=1}^m (\alpha_j A_j)\right) (\alpha_i A_i^T) + \alpha_i^2 \sum_{j=1}^m A_j A_j^T \right] \\ &= \left(\sum_{i=1}^m \alpha_i^2\right) \left(\sum_{i=1}^m A_i A_i^T\right) - \left(\sum_{i=1}^m \alpha_i A_i\right) \left(\sum_{i=1}^m \alpha_i A_i\right)^T. \end{aligned}$$

□

Now we study how the volume and aspect ratio of the John's ellipsoid changes in a move from x to y . If the John's ellipsoid centered at $x = 0$ is the unit ball (which we can assume after an affine transformation) and we make a move to y , the matrix E_y such that

$$\mathcal{E}_y = \left\{ z \mid (z - y)^T E_y^{-2} (z - y) \leq 1 \right\}$$

is the unique (from John's theorem) $\arg \max$ over positive definite matrices of $\log \det E$ under the constraint that

$$\mathcal{E} \subseteq \mathcal{K}_y^s. \tag{3.5}$$

Noting that the John's ellipsoid of \mathcal{K}_x^s is a unit ball at the origin, we see that the matrix E_x can be taken to be the identity matrix. Let us translate \mathcal{K} by $x - y$ (where $x = 0$). All the contact points with \mathcal{K} of \mathcal{E}_x must lie on the boundary of \mathcal{E}_y or outside \mathcal{E}_y . Note that if u_i is a contact point of \mathcal{E}_x with \mathcal{K} , then by definition, $u_i^T u_i = 1$. This implies that for any $\|u\| = 1$,

$$E_y u + y \in \mathcal{E}_y \subseteq \mathcal{K} \subseteq \left\{ x \mid u_i^T x \leq 1 \right\}.$$

The last containment above uses the fact that the John's ellipsoid \mathcal{E}_x to \mathcal{K}_x^s at $x = 0$ is the unit ball and that u_i is a contact point of \mathcal{E}_x with \mathcal{K} . This in turn implies that

$$u_i^T (E_y u + y) \leq 1.$$

Since u may be an arbitrary unit vector, and E_y is symmetric, this condition can be rewritten as

$$|E_y u_i| + \langle u_i, y \rangle \leq 1, \quad i = 1, \dots, m. \tag{3.6}$$

Note that we *do not* claim that E_y is the matrix with largest determinant that satisfies the above constraints. There may be other constraints corresponding to contact points for \mathcal{E}_y that are not contact points of \mathcal{E}_x .

Using Theorem 1 and Lemma 1, we deduce an upper bound on $\det E_y$ as follows.

Recall that we will denote a universal positive constant that is large by C and a universal positive constant that is small by c .

Lemma 2. *Let $r = cn^{-5/2}$, and assume y is chosen from a ball of radius r such that $\|y - x\|_x = |y - x| \leq r$. Then*

$$\det E_y \leq 1 + 2n^{-2}.$$

Proof. Note that by (3.6), E_y satisfies the constraints $u_i^T E_y u_i \leq 1 - u_i^T y$. Since the weights corresponding to the John's ellipsoid \mathcal{K}_x^s are positive, the constraint implies that

$$\sum_{i=1}^m c_i u_i^T E_y u_i \leq \sum_{i=1}^m c_i (1 - u_i^T y).$$

In the above equation, c_i is chosen so that it is non-zero for a contact point of the John's ellipsoid with \mathcal{K}_x^s only if it is also a contact point for \mathcal{K} itself. Such a choice of c_i can always be made, because, as explained in Subsection 2.2.1, if u_i is a contact point with \mathcal{K}_x^s that is not a contact point with \mathcal{K} , then $-u_i$ must be a contact point with both \mathcal{K}_x^s and \mathcal{K} , and so no weight

need fall on u_i ; we can compensate with a weight on $-u_i$. By (2.2) and (2.3), and using the linearity and cyclic invariance of the trace, we have

$$\begin{aligned} \text{tr}(E_y) &\leq \sum_{i=1}^m c_i - \sum_{i=1}^m c_i u_i^T y \\ &= n - \left(\sum_{i=1}^m c_i u_i \right)^T y \\ &= n - \left(\sum_{i=1}^m c_i u_i \right)^T y. \end{aligned}$$

Considering $\left| \sum_i c_i u_i \right|$ to bound the middle term, we may employ Lemma 1. Letting $\alpha_i = \sqrt{c_i}$ and $A_i = \sqrt{c_i} u_i$, we have

$$\left(\sum_{i=1}^m c_i u_i \right) \left(\sum_{i=1}^m c_i u_i \right)^T \preceq \left(\sum_{i=1}^m c_i \right) \left(\sum_{i=1}^m c_i u_i u_i^T \right).$$

Noting that the right side is equal to nI_n , and comparing the largest eigenvalue of the matrices on the left and the right side, it follows that

$$\left| \sum_{i=1}^m c_i u_i \right| \leq \sqrt{n}.$$

Therefore, if y is chosen from a ball of radius $cn^{-5/2}$, by Cauchy–Schwarz we conclude that

$$\text{tr}(E_y) \leq n + cn^{-2}. \tag{3.7}$$

Now, letting the eigenvalues of E_y be denoted by $d_i > 0$, we have by the arithmetic–geometric mean inequality

$$\begin{aligned} (\det E_y)^{1/n} &= \left(\prod_{i=1}^n d_i \right)^{1/n} \\ &\leq \frac{1}{n} \sum_{i=1}^n d_i \\ &\leq \frac{1}{n} (n + cn^{-2}). \end{aligned}$$

Thus,

$$\begin{aligned} \det E_y &\leq (1 + cn^{-3})^n \\ &\leq 1 + n^{-2}. \end{aligned}$$

The last inequality assumes that c is a sufficiently small positive constant. □

We deduce a lower bound on $\det E_y$ by considering a positive definite matrix of the form $E = \beta(I_n - \alpha y y^T)$ such that the corresponding ellipsoid $y + \mathcal{E}$ is contained in the unit ball \mathcal{E}_x .

Note that such a matrix has eigenvalue $\beta(1 - \alpha|y|^2)$ of multiplicity 1 corresponding to the unit eigenvector $y/|y|$, and eigenvalues β of multiplicity $n - 1$ corresponding to any unit vector z which is orthogonal to y . We choose $\beta = 1 - \frac{|y|}{\sqrt{n}}$, and $\alpha = 2\sqrt{n}|y|^{-1}$.

Lemma 3. *We have $y + \mathcal{E} \subseteq \mathcal{E}_x$, and hence $\text{vol}(\mathcal{E}_y) \geq \text{vol}(\mathcal{E})$.*

Proof. Assume after rescaling that \mathcal{E}_x is the unit ball \mathcal{B} centered at the origin. We divide the points u on the boundary of \mathcal{B} into two sets: $A = \left\{u \mid \langle u, y \rangle \leq \frac{|y|}{\sqrt{n}}\right\}$ and $B = \left\{u \mid \langle u, y \rangle > \frac{|y|}{\sqrt{n}}\right\}$.

If $u \in A$, we have

$$1 - \langle u, y \rangle \geq 1 - \frac{|y|}{\sqrt{n}},$$

and noting that $E > 0$ and $0 \prec (I_n - \alpha yy^T) \prec I_n$,

$$|Eu| \leq \beta \leq 1 - \langle u, y \rangle.$$

If $u \in B$, we have

$$\begin{aligned} |Eu|^2 &= \beta^2 u^T (I - \alpha yy^T)^2 u \\ &< \beta^2 (1 - \alpha \langle u, y \rangle)^2 \\ &= \beta^2 (1 - 2\sqrt{n}|y|^{-1} \langle u, y \rangle)^2 \\ &< (1 - \langle u, y \rangle)^2. \end{aligned}$$

Therefore, for any $u \in A \cup B$, we see that $|Eu| + \langle u, y \rangle \leq 1$. From this, it follows that for any $u, v \in \partial\mathcal{B}$,

$$u^T (Ev + y) \leq 1. \tag{3.8}$$

Therefore, for all $v \in \partial\mathcal{B}$, $|Ev + y| \leq 1$.

Thus, $\mathcal{E} + y \subseteq \mathcal{E}_x \subseteq \mathcal{K}$, and so $\mathcal{E} + y \subseteq \mathcal{K}_y^s$. Hence the volume of the John's ellipsoid of \mathcal{K}_y^s , namely \mathcal{E}_y , is at least the volume of \mathcal{E} . □

Lemma 4. *Let $r = cn^{-5/2}$, and assume y is chosen from a ball of radius r such that $\|y - x\|_x = |y - x| \leq r$. Then*

$$\det E_y \geq 1 - 3n^{-2}.$$

Proof. Considering the matrix E as provided by Lemma 3, E_y satisfies

$$\begin{aligned} \det E_y &\geq \det E \\ &= \beta^n (1 - \alpha|y|^2) \\ &= \left(1 - \frac{|y|}{\sqrt{n}}\right)^n (1 - 2\sqrt{n}|y|). \end{aligned}$$

Thus

$$\begin{aligned} \det E_y &\geq (1 - cn^{-2})(1 - 2cn^{-2}) \\ &\geq (1 - 3n^{-2}). \end{aligned}$$

□

Lemmas 2 and 4 establish that for some universal constant $c > 0$ and $|y - x| \leq cn^{-5/2}$, the volume ratio of \mathcal{E}_y and \mathcal{B}_x satisfies

$$1 - 3n^{-2} \leq \frac{\text{vol}(\mathcal{E}_y)}{\text{vol}(\mathcal{B}_x)} \leq 1 + 2n^{-2}. \tag{3.9}$$

This does not necessarily indicate that the shapes of \mathcal{E}_x and \mathcal{E}_y are close, a property we require so that rejection does not occur too frequently. The following lemma guarantees that this is indeed the case.

Lemma 5. *Let $d_1 \geq d_2 \geq \dots \geq d_n > 0$ denote the eigenvalues of E_y . Then the minimum eigenvalue d_n satisfies*

$$d_n \geq 1 - 4n^{-1}$$

for large enough n . Similarly, the maximum eigenvalue d_1 satisfies

$$d_1 \leq 1 + Cn^{-1}$$

for large enough n .

Proof. Assume the eigenvalues are ordered so that $d_1 \geq \dots \geq d_n$. If $d_n \geq 1$, then of course $d_n \geq 1 - 4n^{-1}$, so we may assume that $\zeta := 1 - d_n > 0$. By Lemma 4 we see that

$$\det E_y = \prod_{i=1}^n d_i \geq 1 - 3n^{-2}.$$

By the fact that $\text{tr}(E_y) \leq n + c/n^2$ (see (3.7)), making r smaller if necessary, we have that $\text{tr}(E_y) \leq n + n^{-2}$. By the arithmetic–geometric mean inequality, we thus have

$$\begin{aligned} (1 - 3n^{-2})/d_n &\leq \prod_{i < n} d_i \\ &\leq \left(\frac{\sum_{i < n} d_i}{n - 1} \right)^{n-1} \\ &\leq \left(\frac{n + n^{-2} - d_n}{n - 1} \right)^{n-1} \\ &\leq \left(1 + \frac{1 + n^{-2} - d_n}{n - 1} \right)^{n-1} \\ &\leq \exp(1 + n^{-2} - d_n). \end{aligned}$$

The last inequality holds because $(1 + z/m)^m \leq \exp(z)$ for $z \geq 0$ and integer $m \geq 1$. Thus, to summarize the above chain of inequalities,

$$(1 - 3/n^2) \leq \exp(1 + n^{-2} - d_n)d_n. \tag{3.10}$$

Then, (3.10) may be rephrased as

$$\exp(\zeta + n^{-2})(1 - \zeta) \geq \left(1 - \frac{3}{n^2}\right).$$

This implies that

$$\exp(\zeta)(1 - \zeta) \geq \left(1 - \frac{6}{n^2}\right). \tag{3.11}$$

We note that $\exp(\zeta)(1 - \zeta)$ is a monotonically decreasing function of ζ on $(0, 1)$, because its derivative is

$$-\exp(\zeta) + \exp(\zeta)(1 - \zeta) = -\zeta \exp(\zeta).$$

Recall that we say that a function τ of n satisfies $\tau = \Omega(1)$ if $\liminf_{n \rightarrow \infty} \tau > 0$. If $\zeta = \Omega(1)$, then $\exp(\zeta)(1 - \zeta) \leq 1 - \Omega(1)$, which would contradict (3.11). Therefore, we assume that $\zeta = o(1)$. But then

$$1 - \frac{\zeta^2(1 - \zeta)}{2} \geq \exp(\zeta)(1 - \zeta) \geq 1 - \frac{6}{n^2}$$

gives us $\zeta \leq \frac{4}{n}$, which implies the lower bound on d_n .

To obtain the desired upper bound on d_1 , we first note that since there is a uniform lower bound of $1 - \frac{4}{n}$ on all the d_i , and as the product lies within $(1 - \frac{3}{n^2}, 1 + \frac{2}{n^2})$, we obtain a uniform upper bound of $2e^4$ on the d_i and hence an upper bound of $2e^4$ on d_1 . Therefore, if $\|y - x\|_x \leq r$ then $\|x - y\|_y \leq 2e^4 r$. We may now reverse the roles of x and y , and obtain the following upper bound on d_1 :

$$d_1 \leq 1 + \frac{C}{n}. \quad \square$$

Now, to derive a lower bound on the conductance for John's walk, we first must bound the statistical distance between two points given a bound on their geometric distance with respect to the local norm. Again without loss of generality, in what follows we may assume that $x = 0$ and that the John's ellipsoid centered at x is the unit ball \mathcal{B}_x (otherwise, perform an affine transformation so that this is the case). Let $x, y \in \mathcal{K}$ represent any two points in the body such that $\|y - x\|_x = |y - x| \leq r$, where $r \in (0, 1)$ is a constant to be specified in terms of the dimension n . Let P_x and P_y denote the one-step transition probability measures defined at x and y , respectively. Let the uniform probability measures defined by the rescaled John's ellipsoids $\mathcal{E}_x(r)$ and $\mathcal{E}_y(r)$ be denoted by Q_x and Q_y , respectively. We seek to bound

$$d_{TV}(P_x, P_y) \leq d_{TV}(P_x, Q_x) + d_{TV}(Q_x, Q_y) + d_{TV}(Q_y, P_y) \tag{3.12}$$

by choosing r such that the right side of (3.12) is $1 - \Omega(1)$.

To bound $d_{TV}(Q_x, Q_y)$ in (3.12), letting \tilde{Q}_y denote the probability measure corresponding to the uniform distribution on a ball of radius r centered at y , we may alternatively bound

$$d_{TV}(Q_x, Q_y) \leq d_{TV}(Q_x, \tilde{Q}_y) + d_{TV}(\tilde{Q}_y, Q_y). \tag{3.13}$$

We bound each term in (3.13) separately. To bound $d_{TV}(Q_x, \tilde{Q}_y)$, note that by our assumption that $\mathcal{E}_x = \mathcal{B}_x$ (the unit ball at x), the corresponding densities with respect to the dominating Lebesgue measure λ are

$$q_x(z) = \left(\frac{1}{\lambda(\mathcal{B}_x(r))} \right) \cdot 1_{\{z \in \mathcal{B}_x(r)\}}$$

and

$$\tilde{q}_y(z) = \left(\frac{1}{\lambda(\mathcal{B}_y(r))} \right) \cdot 1_{\{z \in \mathcal{B}_y(r)\}}.$$

Thus, using (3.1) and noting that $\lambda(\mathcal{B}_x(r)) = \lambda(\mathcal{B}_y(r))$, we have

$$\begin{aligned} d_{TV}(Q_x, \tilde{Q}_y) &= 1 - \int_{\mathcal{B}_x(r) \cap \mathcal{B}_y(r)} q_x(z) d\lambda(z) \\ &= 1 - \frac{\lambda(\mathcal{B}_x(r) \cap \mathcal{B}_y(r))}{\lambda(\mathcal{B}_x(r))}. \end{aligned} \tag{3.14}$$

The Lebesgue measure of $\mathcal{B}_x \cap \mathcal{B}_y$ is equal to twice the volume of a spherical cap. The following lemma regarding the volume of a spherical cap (Lemma 0.1 from [10]) is useful.

Lemma 6. *Let $\mathcal{B}_x \subset \mathbb{R}^n$ be the Euclidean ball of unit radius centered at x . Let $\mathcal{H} \subset \mathbb{R}^n$ define a half-space at a distance of at least t from x (so x is not contained in the half-space). Then for $t \leq \frac{1}{\sqrt{n}}$, we have*

$$\lambda(\mathcal{H} \cap \mathcal{B}_x) \geq \frac{1}{2} (1 - t\sqrt{n}) \lambda(\mathcal{B}_x).$$

The following lemma results trivially from Lemma 6 and (3.14).

Lemma 7. *Let $t \leq 1$. If $\|y - x\|_x = |y - x| \leq \frac{t}{\sqrt{n}}$, then*

$$d_{TV}(Q_x, \tilde{Q}_y) \leq t.$$

To bound $d_{TV}(\tilde{Q}_y, Q_y)$, note that we are bounding the total variation distance between a density supported on a ball and a density supported on an ellipsoid with the same center. The following lemma provides the bound.

Lemma 8. *If $\|y - x\|_x \leq r = cn^{-5/2}$, the total variation distance between \tilde{Q}_y and Q_y satisfies*

$$d_{TV}(\tilde{Q}_y, Q_y) \leq 1/4.$$

Proof. Note that by (3.1), we have

$$\begin{aligned} d_{TV}(\tilde{Q}_y, Q_y) &= 1 - \mathbb{E}_{\tilde{Q}_y} \left[\min \left(1, \frac{\lambda(\mathcal{B}_y(r))}{\lambda(\mathcal{E}_y(r))} \right) \mathbf{1}_{z \in \mathcal{E}_y(r)} \right] \\ &= 1 - \min \left[1, \frac{\lambda(\mathcal{B}_y)}{\lambda(\mathcal{E}_y)} \right] \mathbb{P}_{\tilde{Q}_y} (Z \in \mathcal{E}_y(r)) \\ &\leq 1 - \min \left[1, \frac{\lambda(\mathcal{B}_y)}{\lambda(\mathcal{E}_y)} \right] \mathbb{P}_{\tilde{Q}_y} (Z \in \mathcal{E}_y(r) | B) \mathbb{P}_{\tilde{Q}_y} (Z \in B), \end{aligned}$$

where B denotes the event in which $Z \in (1 - \frac{c}{n}) \cdot \mathcal{B}_y(r)$. By Lemma 5, it follows that $\mathbb{P}_{\tilde{Q}_y} (Z \in \mathcal{E}_y(r) | B) = 1$, since the smallest eigenvalue of E_y is at least $1 - Cn^{-1}$. Additionally, by (3.9),

$$\begin{aligned} \min \left[1, \frac{\lambda(\mathcal{B}_y)}{\lambda(\mathcal{E}_y)} \right] &\geq \frac{1}{1 + Cn^{-2}} \\ &\geq \exp(-Cn^{-2}). \end{aligned}$$

Now, noting that $(1 - \frac{x}{2}) \geq e^{-x}$ for $x \in [0, 1]$, we have $\mathbb{P}_{\tilde{Q}_y} (Z \in B) = \left(1 - \frac{c}{n}\right)^n \geq e^{-2c}$, and

$$d_{TV}(\tilde{Q}_y, Q_y) \leq 1 - \exp(-C(2 + n^{-2})).$$

□

To bound $d_{TV}(P_x, Q_x)$, we provide the following lemma.

Lemma 9. *If $\|y - x\|_x \leq r = cn^{-5/2}$, the total variation distance between P_x and Q_x satisfies*

$$d_{TV}(P_x, Q_x) \leq 1/4.$$

Proof. With some abuse of notation with regard to (2.6), temporarily let the density of Q_x with respect to the dominating measure μ as defined by (2.7) be

$$q_x(y) \equiv \frac{dQ}{d\mu}(y) = \left(\frac{1}{\lambda(\mathcal{E}_x(r))} \right) \cdot 1_{\{y \in \mathcal{E}_x(r), y \neq x\}}.$$

Then, since $q_x(x) = 0$ and $p_x(y) \leq q_x(y)$ for $y \neq x$, by (3.1) we have

$$\begin{aligned} d_{TV}(P_x, Q_x) &= 1 - \int_{\{y \mid q_x(y) \geq p_x(y)\}} \min[q_x(y), p_x(y)] d\mu(y) \\ &= 1 - \int_{\mathcal{K} \setminus \{x\}} \min[q_x(y), p_x(y)] d\lambda(y) \\ &= 1 - \int_{\mathcal{E}_x(r) \setminus \{x\}} \min\left[\frac{1}{\lambda(\mathcal{E}_x(r))}, \frac{1}{\lambda(\mathcal{E}_y(r))}\right] \cdot 1_{\{x \in \mathcal{E}_y(r)\}} d\lambda(y) \\ &= 1 - \int_{\mathcal{E}_x(r)} \left[\min\left(1, \frac{\lambda(\mathcal{E}_x(r))}{\lambda(\mathcal{E}_y(r))}\right) \right] \cdot 1_{\{x \in \mathcal{E}_y(r)\}} \cdot \left(\frac{1}{\lambda(\mathcal{E}_x(r))}\right) d\lambda(y) \\ &= 1 - \mathbb{E}_{Q_x} \left[\min\left(1, \frac{\lambda(\mathcal{E}_x)}{\lambda(\mathcal{E}_Y)}\right) \cdot 1_{\{Y \in A\}} \right], \end{aligned}$$

where we let A denote the ‘accept’ event in which $x \in \mathcal{E}_Y(r)$. Since $\mathcal{E}_x = \mathcal{B}_x$, as in the proof of Lemma 8, we have for all $Y \in A$

$$\min\left[1, \frac{\lambda(\mathcal{E}_x)}{\lambda(\mathcal{E}_Y)}\right] \geq \frac{1}{1 + Cn^{-2}}.$$

Therefore,

$$\begin{aligned} d_{TV}(P_x, Q_x) &\leq 1 - \left(\frac{1}{1 + Cn^{-2}}\right) \mathbb{P}_{Q_x}(Y \in A) \\ &\leq 1 - \left(\frac{1}{1 + Cn^{-2}}\right) \mathbb{P}_{Q_x}(Y \in A \mid Y \in B) \mathbb{P}_{Q_x}(Y \in B), \end{aligned}$$

where B is the event in which $Y \in \left(1 - \frac{c}{n}\right) \cdot \mathcal{E}_x(r) = r\left(1 - \frac{c}{n}\right) \cdot \mathcal{B}_x$. Again by Lemma 5, $\mathbb{P}_{Q_x}(Y \in A \mid Y \in B) = 1$. The remainder of the proof is as in Lemma 8. \square

Note that by a similar argument, $d_{TV}(Q_y, P_y) \leq 1/4$ for some universal $c > 0$ as well. Combining this with Lemmas 7, 8, and 9, we obtain the following theorem.

Theorem 5. *If $\|y - x\|_x \leq \frac{rt}{\sqrt{n}} = ct n^{-3}$ for some universal constant $c > 0$ and some $t \leq 1$, the total variation distance between P_x and P_y satisfies*

$$d_{TV}(P_x, P_y) \leq 3/4 + t = 1 - \epsilon.$$

In particular, we may choose $t = 1/8$, so that $\epsilon = 1/8$.

We finally arrive at a lower bound on the conductance for John's walk using Theorems 3, 4, and 5. The proof of the next result is similar to that of Corollary 10 and Theorem 11 in [9].

Theorem 6. (Conductance lower bound.) *Consider the partition $\mathcal{K} = S_1 \cup S_2$ where $S_1, S_2 \in \mathcal{A}$, and let π be the uniform measure on \mathcal{K} , i.e.,*

$$\pi(A) = \frac{\lambda(A)}{\lambda(\mathcal{K})} \quad \text{for all } A \in \mathcal{A}.$$

Then for large enough n and $t = 1/8$, we have

$$\int_{S_1} P_x(S_2)d\pi(x) \geq \left(\frac{c}{512n^{7/2}}\right) \min(\pi(S_1), \pi(S_2)),$$

so $\phi = \Omega(n^{-7/2})$.

Proof. Note that the Radon–Nikodym derivative of P_x with respect to the Lebesgue measure λ is well-defined for all $y \in \mathcal{K} \setminus \{x\}$, and is given as

$$\frac{dP_x}{d\lambda}(y) = \min\left[\frac{1}{\lambda(\mathcal{E}_x(r))}, \frac{1}{\lambda(\mathcal{E}_y(r))}\right] 1_{\{x \in \mathcal{E}_y(r), y \in \mathcal{E}_x(r)\}}.$$

Let

$$\rho(x) \equiv \frac{d\pi}{d\lambda}(x) = \frac{1}{\lambda(\mathcal{K})} \cdot 1_{\{x \in \mathcal{K}\}}$$

be the density for π . Then for any $x, y \in \mathcal{K}$ such that $y \neq x$, we have

$$\rho(x) \frac{dP_x}{d\lambda}(y) = \rho(y) \frac{dP_y}{d\lambda}(x),$$

from which it follows that π is the stationary measure for the chain.

Now consider points far inside S_1 that are unlikely to cross over to S_2 . Letting $t = 1/8$ so $\epsilon = 1/8$ as in Theorem 5, we define

$$S'_1 \equiv S_1 \cap \left\{x \mid \rho(x)P_x(S_2) < \frac{\epsilon}{2\lambda(\mathcal{K})}\right\}.$$

Similarly, let

$$S'_2 \equiv S_2 \cap \left\{y \mid \rho(y)P_y(S_1) < \frac{\epsilon}{2\lambda(\mathcal{K})}\right\}.$$

Since $\rho(x)P_x(S_2) \geq \epsilon/(2\lambda(\mathcal{K}))$ for $x \in S_1 \setminus S'_1$, we have

$$\begin{aligned} \int_{S_1} P_x(S_2)d\pi(x) &\geq \int_{S_1 \setminus S'_1} \rho(x)P_x(S_2)d\lambda(x) \\ &\geq \frac{\epsilon\lambda(S_1 \setminus S'_1)}{2\lambda(\mathcal{K})} \\ &= (\epsilon/2)\pi(S_1 \setminus S'_1). \end{aligned}$$

Similarly, for $y \in S_2 \setminus S'_2$ we have

$$\int_{S_2} P_y(S_1)d\pi(y) \geq (\epsilon/2)\pi(S_2 \setminus S'_2).$$

By the reversibility of the chain, we have

$$\int_{S_1} P_x(S_2)d\pi(x) = \int_{S_2} P_y(S_1)d\pi(y), \tag{3.15}$$

so it follows that

$$\begin{aligned} \int_{S_1} P_x(S_2)d\pi(x) &= \frac{1}{2} \int_{S_1} P_x(S_2)d\pi(x) + \frac{1}{2} \int_{S_2} P_y(S_1)d\pi(y) \\ &\geq \frac{\epsilon}{4} (\pi(S_1 \setminus S'_1) + \pi(S_2 \setminus S'_2)) \\ &= (\epsilon/4)\pi(\mathcal{K} \setminus (S'_1 \cup S'_2)). \end{aligned}$$

Now let $\delta = ctn^{-3}$. Assuming that $\pi(S'_1) \leq (1 - \delta)\pi(S_1)$, we have $\pi(S_1 \setminus S'_1) = \pi(S_1) - \pi(S'_1) \geq \delta\pi(S_1)$, and thus

$$\begin{aligned} \int_{S_1} P_x(S_2)d\pi(x) &\geq (\epsilon\delta/2)\pi(S_1) \geq (\epsilon\delta/2) \min(\pi(S_1), \pi(S_2)) \\ &= \left(\frac{c}{128n^3}\right) \min(\pi(S_1), \pi(S_2)), \end{aligned}$$

which proves the claim. Similarly, if $\pi(S'_2) \leq (1 - \delta)\pi(S_2)$, the claim is proved again using (3.15). Thus, assume that $\pi(S'_1) > (1 - \delta)\pi(S_1)$ and $\pi(S'_2) > (1 - \delta)\pi(S_2)$. By Theorem 3, we have

$$\pi(\mathcal{K} \setminus (S'_1 \cup S'_2)) \geq \sigma(S'_1, S'_2)\pi(S'_1)\pi(S'_2).$$

Now, given $x \in S'_1$ and $y \in S'_2$, the total variation between P_x and P_y satisfies

$$\begin{aligned} d_{TV}(P_x, P_y) &\geq P_x(S_1) - P_y(S_1) \\ &= 1 - P_x(S_2) - P_y(S_1) \\ &> 1 - \epsilon. \end{aligned}$$

By Theorem 5, it follows that $\|y - x\|_x > \delta$. Then by Theorem 4 it follows that

$$\sigma(S'_1, S'_2) \geq n^{-1/2}\|y - x\|_x > \delta n^{-1/2}.$$

Finally, we deduce that

$$\begin{aligned} \int_{S_1} P_x(S_2)d\pi(x) &\geq (\epsilon\delta n^{-1/2}/4)\pi(S'_1)\pi(S'_2) \\ &\geq \left(\frac{\epsilon\delta(1 - \delta)^2}{4\sqrt{n}}\right) \pi(S_1)\pi(S_2) \\ &\geq \left(\frac{\epsilon\delta(1 - \delta)^2}{8\sqrt{n}}\right) \min(\pi(S_1), \pi(S_2)) \\ &= \left(\frac{c(1 - \delta)^2}{512n^{7/2}}\right) \min(\pi(S_1), \pi(S_2)). \end{aligned}$$

The claim follows by absorbing terms into the constant for large enough n . □

Now applying Corollary 1, we obtain our main theorem.

Theorem 7. For $\epsilon > 0$ and $M \geq \sup \frac{\pi_0(A)}{\pi(A)}$, after $t(\epsilon) = Cn^7 \log(\sqrt{M}/\epsilon)$ steps of John's walk, we have $d_{TV}(\pi_{t(\epsilon)}, \pi) \leq \epsilon$.

As a matter of fact, this theorem allows us to find mixing time bounds starting from any point that is not on the boundary of \mathcal{K} . Suppose we know that x belongs to the interior of

\mathcal{K} and satisfies the following chord condition. For all chords pq of \mathcal{K} containing x , assume $\frac{|p-x|}{|q-x|} \in (\eta, \eta^{-1})$ for some parameter $0 < \eta < 1$ that measures the centrality of x in \mathcal{K} . Then we see that

$$\lambda(\mathcal{B}_x(r)) \geq \left(\frac{r\eta}{\sqrt{n}} \right)^n \lambda(\mathcal{K}).$$

After a random geometrically distributed time τ with mean bounded above by an absolute constant, the first nontrivial move occurs. Then the distribution of x_τ has a density bounded above by

$$M = \left(\frac{r\eta}{\sqrt{n}} \right)^{-n}.$$

We thus have the following theorem.

Theorem 8. *Let $r = cn^{-\frac{5}{2}}$ be the radius of the John's ellipsoids in the walk. For all chords pq of \mathcal{K} containing x , assume $\frac{|p-x|}{|q-x|} \in (\eta, \eta^{-1})$ for some parameter $0 < \eta < 1$ that measures the centrality of x in \mathcal{K} . Then there is a random geometrically distributed time τ with mean bounded above by C such that for $\epsilon > 0$, after $t(\epsilon) + \tau = Cn^7 (n \log(\sqrt{n}/(r\eta)) + \log(1/\epsilon)) + \tau$ steps of John's walk, we have $d_{TV}(\pi_{t(\epsilon)+\tau}, \pi) \leq \epsilon$.*

4. Concluding remarks

We introduced an affine-invariant random walk akin to the Dikin walk which uses uniform sampling from John's ellipsoids of a certain small radius of appropriately symmetrized convex sets to make steps, and showed that this walk mixes to within a total variation distance ϵ in $O(n^7 \log \epsilon^{-1})$ steps from a warm start. The type of convex body \mathcal{K} is not specified (i.e., it need not be a polytope) in our analysis of the mixing time, but one must have access to the John's ellipsoid of the current symmetrization of the convex body. A significant feature of this walk is that its mixing time from a warm start, or alternatively a 'central point' such as the center of mass, can be bounded above by a quantity that has absolutely no dependence on any parameter associated with the body apart from its dimension.

One potential avenue for obtaining an improved mixing time is to increase r and control the relative change of shape of the John's ellipsoids \mathcal{E}_x and \mathcal{E}_y when $\|x - y\|_x < r/\sqrt{n}$ in a more direct manner. We suspect that the method of controlling the change in shape using the arithmetic-geometric mean inequality and the ratio of the volumes of \mathcal{E}_x and \mathcal{E}_y that appears in Lemma 3.3 might be suboptimal.

Acknowledgements

We are deeply grateful to the anonymous reviewers for their role in improving this paper by pointing out numerous places where the arguments needed greater clarity.

Funding information

H. N. was partially supported by a Ramanujan Fellowship and a Swarna Jayanti Fellowship instituted by the Government of India. H. N. also acknowledges the support of DAE project no. 12-R&D-TFR-5.01-0500.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] BALL, K. (1992). Ellipsoids of maximal volume in convex bodies. *Geom. Dedicata* **41**, 241–250.
- [2] CHEN, Y., DWIVEDI, R., WAINWRIGHT, M. J. AND YU, B. (2017). Fast MCMC sampling algorithms on polytopes. Preprint. Available at <https://arxiv.org/abs/1710.08165>.
- [3] GRUBER, P. M. (1988). Minimal ellipsoids and their duals. *Rend. Circ. Mat. Palermo* **37**, 35–64.
- [4] JOHN, F. (1948). Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays: Courant Anniversary Volume*, eds K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker, Wiley-Interscience, New York, pp. 187–204.
- [5] KANNAN, R., LOVÁSZ, L. AND SIMONOVITS, M. (1997). Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures Algorithms* **11**, 1–50.
- [6] KANNAN, R. AND NARAYANAN, H. (2012). Random walks on polytopes and an affine interior point method for linear programming. *Math. Operat. Res.* **37**, 1–20.
- [7] LADDHA, A., LEE, Y. T. AND VEMPALA, S. S. (2020). Strong self-concordance and sampling. In *Proc. 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC 2020)*, eds K. Makarychev, Y. Makarychev, M. Tulsiani, G. Kamath and J. Chuzhoy, Association for Computing Machinery, New York, pp. 1212–1222.
- [8] LEE, Y. T. AND SIDFORD, A. (2013). Path finding I: solving linear programs with $\tilde{O}(\sqrt{\text{rank}})$ linear system solves. Preprint. Available at <https://arxiv.org/abs/1312.6677>.
- [9] LOVÁSZ, L. (1999). Hit-and-run mixes fast. *Math. Program.* **86**, 443–461.
- [10] LOVÁSZ, L. AND SIMONOVITS, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms* **4**, 359–412.
- [11] LOVÁSZ, L. AND VEMPALA, S. (2006). Hit-and-run from a corner. *SIAM J. Computing* **35**, 985–1005.
- [12] LOVÁSZ, L. AND VEMPALA, S. (2007). The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms* **30**, 307–358.
- [13] NARAYANAN, H. (2016). Randomized interior point methods for sampling and optimization. *Ann. Appl. Prob.* **26**, 597–641.