Journal of Theoretical Probability (2020) 33:22–35
https://doi.org/10.1007/s10959-018-0848-7

Uniform Ergodicity for Brownian Motion in a Bounded Convex Set

Jackson Loper[1]

Received: 17 January 2018 / Revised: 10 May 2018 / Published online: 31 July 2018
© The Author(s) 2018

Abstract
We consider an n-dimensional Brownian motion trapped inside a bounded convex set by normally reflecting boundaries. It is well known that this process is uniformly ergodic. However, the rates of this ergodicity are not well understood, especially in the regime of very high-dimensional sets. Here we present new bounds on these rates for convex sets with a given diameter. Our bounds do not depend upon the smoothness of the boundary nor the value of the ambient dimension, n.

Keywords Reflected Brownian motion · Uniform ergodicity · Convex sets

Mathematics Subject Classification (2010) 60J65 · 37A25

1 Introduction

Let Ω be an open, convex set with diameter $d < \infty$. We will consider a Brownian motion, $\{X_t\}_{t\geq 0}$, trapped inside Ω by normally reflecting boundaries, initialized at some point $X_0 = x$. We will give a more precise definition of this process momentarily; intuitively, X is a homogeneous Markov process which behaves like a Brownian motion on Ω, spends essentially no time at the boundary (which we will denote $\partial\Omega$), and is always contained in the closure (which we will denote $\bar{\Omega}$). Let $\{p(t, x, \mathrm{d}y)\}_{t\geq 0, x\in\bar{\Omega}}$ denote the transition measures of the process, i.e., $\mathbb{P}(X_t \in A | X_0 = x) = \int_A p(t, x, \mathrm{d}y)$. It is well known that the process X is ergodic, with stationary distribution σ, where $\sigma(\mathrm{d}y) \triangleq \mathrm{d}y/\mathrm{Vol}(\Omega)$ (cf. [4]). Thus, for every x we must have that

$$p(t, x, \cdot) \to \sigma \quad \text{as} \quad t \to \infty$$

✉ Jackson Loper
jl5116@columbia.edu

1 Columbia University, New York City, NY 10027, USA

in various modes of convergence. The rate and exact nature of this convergence has been investigated in a number of ways. For example, consider the case that Ω is a convex polytope which can be contained inside a cube of diameter d, such as $[0, d/\sqrt{n}]^n$. In this special case, it has long been known that

$$\sup_x \|p(t, x, \cdot) - \sigma\|_{\mathrm{TV}} \leq \sqrt{\frac{d^2}{2\pi t}} \tag{1}$$

where $\|\mu\|_{\mathrm{TV}} \triangleq \sup_A |\mu(A)|$ (cf. [16]). Another important result follows from a certain Poincare constant on arbitrary bounded convex domains, first rigorously shown by Bebendorf in [3]. Using this result one may readily show that

$$\begin{aligned} &\left| \int g(x) \left(\int f(y) p(t, x, \mathrm{d}y) \right) \mathrm{d}y - \int f(y) \sigma(\mathrm{d}y) \right| \\ &\leq \exp\left(-\frac{1}{2} \left(\frac{\pi}{d} \right)^2 t \right) \left\| g - \frac{1}{\mathrm{Vol}(\Omega)} \right\|_{\mathscr{L}^2} \|f\|_{\mathscr{L}^2} \end{aligned} \tag{2}$$

where we define $\|f\|^2_{\mathscr{L}^2} \triangleq \int f^2 \mathrm{d}x$ and take g to be any density (i.e., $g \geq 0$ and $\int g = 1$). One way to see the significance of this formula is to take Y to be some variable with $Y \sim \sigma$, and $\rho(\mathrm{d}x) = g(x)\mathrm{d}x$ to be some initial distribution. Equation (2) can then be understood as a bound on the rate at which $\mathbb{E}[f(X_t)|X_0 \sim \rho] \to \mathbb{E}[f(Y)]$. Comparing Bebendorf's result with Eq. (1), we see that Bebendorf's result has several advantages:

1. Equation (1) becomes less powerful as $n \to \infty$, but Eq. (2) does not suffer from this deficit. To see how Eq. (1) fails in high dimensions, consider what it says about a convex polytope of diameter d which is roughly spherical. Such a polytope generally cannot fit inside a cube of diameter d. Indeed, to enclose an n-dimensional ball of diameter d, one needs a cube with diameter $d\sqrt{n}$. Thus, as $n \to \infty$, Eq. (1) becomes quite weak for certain kinds of diameter-d sets. Equation (2) does not suffer from this problem.
2. Equation (2) has exponential decay instead of polynomial decay.
3. Equation (2) does not require Ω to be a convex polytope.

However, Bebendorf's result also has some problems: it is not directly applicable when the initial distribution on X_0 isn't absolutely continuous with respect to Lebesgue measure. For example, let δ_x denote the degenerate initial distribution defined by $\delta_x(A) = \mathbb{I}_{x \in A}$. If we try applying Eq. (2) to the limiting case as $g(x)\mathrm{d}x \to \delta_x(\mathrm{d}x)$, the right-hand side of the bound must become infinite (and thus quite useless).

More generally, a rich understanding of the rate of ergodicity remains elusive, even in this simple convex case. How does it depend upon the dimension? How does it depend upon the initial condition?

It turns out that a simple one-dimensional diffusion can shed some light on these questions. Let $\{W_t\}_{t \geq 0}$ denote a one-dimensional Brownian motion. Let $\tilde{\tau}_d = \inf\{t : W_t \notin (-d, d)\}$. Note that the distribution of this object is straightforward to

calculate and analyze. For example, in [11] it is shown that the survival function of $\tilde{\tau}_d$ is given by

$$\mathbb{P}(t \leq \tilde{\tau}_d | W_t = k) = F_d(t, k) \triangleq \sum_{n=0}^{\infty} \mathrm{e}^{-\frac{\pi^2}{8d^2}(2n+1)^2 t} \frac{4(-1)^n}{\pi(2n+1)} \cos\left(\frac{2n+1}{2} \times \frac{\pi k}{d}\right)$$

Numerical estimation of this sum is straightforward and effective in practice. Indeed, $F_d(t, k)$ is simply the solution to the heat equation on $[-d, d]$ with homogeneous Dirichlet boundary conditions and initial condition $F_d(0, k) = \mathbb{I}_{k \in (-d,d)}$; this partial differential equation is well understood (cf. [5]). It is also easy to bound F_d using the moment-generating function of $\tilde{\tau}$ and a Chernoff bound; the moment-generating function can be deduced by an application of the Kac moment formula, yielding $\mathbb{E}\left[\mathrm{e}^{\gamma \tilde{\tau}_d} | W_t = k\right] = \cos(\sqrt{2\gamma} k) / \cos(\sqrt{2\gamma} d)$ for any $\gamma \leq \pi^2/8d^2$ (cf. [8,12]).

We can use the distribution of $\tilde{\tau}_d$ to help us understand the rate of ergodicity for convex domains:

Theorem 1 *Let $\Omega \subset \mathbb{R}^n$ bounded, open, convex, with diameter d. Let $\{p(t, x, dy)\}_{t \geq 0, x \in \Omega}$ denote the transition measures of Brownian motion trapped inside Ω by normally reflecting barriers. Then*

$$\|p(t, x, \cdot) - p(t, y, \cdot)\|_{\mathrm{TV}} \leq F_d(4t, d - |x - y|)$$
$$\|p(t, x, \cdot) - \sigma\|_{\mathrm{TV}} \leq \int F_d(4t, d - |x - y|)\, \sigma(\mathrm{d}y) \leq F_d(4t, 0)$$

This very last bound cannot be improved by more than a factor of 2. In particular, taking the special case that $\Omega = [-d/2, d/2] \times [0, \epsilon]^{n-1} \subset \mathbb{R}^n$ for any $\epsilon > 0$, we have that $\|p(t, d/2, \cdot) - \sigma\|_{\mathrm{TV}} \geq F_d(4t, 0)/2$.

We will defer the proof to Sect. 3.

Notice that the leading $\exp\left(-\pi^2 t / 2d^2\right)$ rate in $F_d(4t, \cdot)$ is the same as the rate given by Bebendorf in Eq. (2). This is no accident. Both quantities reflect the spectral gap for the Neumann Laplacian on the interval $[0, d]$, namely π/d.

The author's particular interest in this problem arose from a question about hitting probabilities. Let A, B denote two open disjoint subsets of Ω. Let $T = \inf\{t : X_t \in A \cup B\}$, and consider the problem of estimating $u(x) = \mathbb{P}(X_T \in \partial A | X_0 = x)$. In general it can be quite tricky to analyze u. However, there are some circumstances in which it simplifies considerably. Let $x \in \Omega$ denote some point such that $\mathbb{P}(T > t | X_0 = x) \approx 1$ and $\|p(t, x, \cdot) - \sigma\|_{\mathrm{TV}} \approx 0$. Then it is easy to show that $u(x) \approx \int u(y) \sigma(\mathrm{d}y)$. This quantity is of course independent of x and can be estimated efficiently. Unfortunately, it is not immediately obvious how to tell whether a given point x satisfies these criteria. In particular, exact computation of $\|p(t, x, \cdot) - \sigma\|_{\mathrm{TV}}$ is nontrivial, and so we turned to finding accurate rates of the uniform ergodicity as a way to bound this quantity.

The remainder of this article is divided into three sections:

1. *Known results* We give a rigorous definition for reflecting Brownian motion in a convex set and formalize some aspects of our introductory exposition. We summa-

rize known results, look at the equations governing $p(t, x, \mathrm{d}y)$, and see how the work by Bebendorf yields the rate of convergence found in Eq. (2). We will examine a coupling idea whose first rigorous construction is due to Atar and Burdzy (cf. [1]). Finally, we will give some remarks on the function $F_d(t, k)$ and begin to see why it will appear in our main theorem.

2. *Application of known results* Here we prove our main theorem, using the coupling construction of Atar and Burdzy.
3. *Conclusions* We will consider possible directions for future research.

2 Known Results

The theory of reflected Brownian motion in convex sets is fairly well developed. To get a sense of this history, we will here recall Tanaka's early work on the subject:

Definition 1 Let $\Omega \in \mathbb{R}^n$ any set. A plane $\{x : \langle x, y\rangle = c\}$ is said to be a **supporting hyperplane** of Ω if $\Omega \subset \{x : \langle x, y\rangle \geq c\}$. The vector y is said to be the **inward-facing normal vector**.

Definition 2 Let $\Omega \subset \mathbb{R}^n$ denote an open set. Let $\{W_t\}_{t\geq 0}$ denote an n-dimensional Brownian motion adapted to $\{\mathcal{F}_t\}_{t\geq 0}$. Let us say there exists

- a continuous process $\{X_t\}_{t\geq 0} \subset \bar{\Omega}$ which is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$
- a positive locally finite random measure μ on $[0, \infty)$
- a random function $\boldsymbol{n} : \mathbb{R}^+ \to \mathbb{R}^n$

such that

$$X_t = W_t + \int_0^t \boldsymbol{n}(s)\mu(\mathrm{d}s)$$
$$\mu(\{t : X_t \notin \partial\Omega\}) = 0$$

for every t and $\boldsymbol{n}(t)$ is the inward-facing normal vector of a supporting hyperplane of Ω at the point X_t for μ almost every value of t. Then we will call X a **reflecting Brownian motion in Ω driven by W**.

It is worth taking a moment to consider the case $n = 1$, in which case much of the complexity of the definition above evaporates. To make it as simple as possible, say Ω is simply the interval $[0, \infty)$. Then it is easy to see that

$$X_t \triangleq W_t - 0 \wedge \inf_{s\leq t} W_s$$

is a reflecting Brownian motion on Ω. We refer the reader to [7] for a useful exposition on this point. The corresponding measure μ is simply the local time of X on the boundary, and the vector $\boldsymbol{n}$ is simply the number 1. That is, $-0 \wedge \inf_{s\leq t} W_s = \int_0^t 1\mu(\mathrm{d}s)$. This formulation will be useful to us later, as we will see that the general problem for $n > 1$ can be approximated by looking at a much simpler problem with $n = 1$.

For the general case $n > 1$, Tanaka showed that as long as Ω is convex and bounded, we can always find a unique reflecting Brownian motion:

Theorem 2 (Tanaka [18]) *Let W denote any n-dimensional Brownian motion with $W_0 \in \Omega$. If $\Omega \subset \mathbb{R}^n$ is bounded and convex then there is a pathwise-unique reflecting Brownian motion in Ω driven by W.*

Proof Cf. [18]. □

Here we summarize some well-known facts about the process X.

Lemma 1 (Properties of reflected Brownian motion in a convex set Ω) *Let W denote an $\{\mathcal{F}_t\}_{t\geq 0}$-adapted Brownian motion and let X denote a reflecting Brownian motion in Ω driven by W. If Ω is convex, then*

1. *$\{X_t\}_{t\geq 0}$ is a strongly $\{\mathcal{F}_t\}_{t\geq 0}$-adapted homogeneous Markov process. Let $\{p(t,x,\mathrm{d}y)\}_{t\geq 0, x\in\bar{\Omega}}$ denote the transition measures of this process. The process X is reversible with respect to σ, i.e., $\int_{x\in A}\int_{y\in B}\sigma(\mathrm{d}x)p(t,x,\mathrm{d}y) = \int_{x\in B}\int_{y\in A}\sigma(\mathrm{d}x)p(t,x,\mathrm{d}y)$. In particular, σ is a stationary distribution of X.*
2. *X is uniformly ergodic with stationary distribution σ, i.e.,*

$$\lim_{t\to\infty}\sup_x \|p(t,x,\cdot)-\sigma\|_{\mathrm{TV}} = 0$$

3. *Let $\lambda \geq 0$ denote any constant such that*

$$\int f\,\mathrm{d}x = 0 \implies \lambda\|f\|_{\mathscr{L}^2} \leq \sqrt{\int |\nabla f(x)|^2\,\mathrm{d}x}$$

for weakly differentiable functions $f : \Omega \to \mathbb{R}$. Then for any density $g \in \mathscr{L}^2$ (i.e., $g \geq 0$ and $\int g(x)\mathrm{d}x = 1$), we have that

$$|\mathbb{E}[f(X_t)|X_0\sim\rho] - \mathbb{E}[f(Y)]| \leq \exp\left(-\frac{1}{2}\lambda^2 t\right)\left\|g - \frac{1}{\mathrm{Vol}(\Omega)}\right\|_{\mathscr{L}^2}\|f\|_{\mathscr{L}^2}$$

where $Y\sim\sigma$ and $\rho(\mathrm{d}x) = g(x)\mathrm{d}x$.

Proof These results are well known. We relate them at a high level here for the convenience of the reader.

The key is to grasp the connection between X and a certain so-called Dirichlet form,

$$\mathscr{E}(f,g) = \frac{1}{2}\int_\Omega \langle\nabla f, \nabla g\rangle\,\mathrm{d}x$$

which is understood as a bilinear form on the Sobolev space H^1 of weakly differentiable functions on Ω. The arc of this connection is the content of the treatise [9]. We will only sketch it briefly in this paragraph. Let $\mathscr{L}^2(\Omega)$ denote the space of square-integrable measurable functions on Ω, equipped with the inner product $\langle f,g\rangle_{\mathscr{L}^2} = \int f(x)g(x)\mathrm{d}x$. One can find a unique nonnegative definite operator

$A : H^1 \to \mathscr{L}^2$ such that $\mathscr{E}(f, g) = \int (Af)(Ag)\,\mathrm{d}x$. It turns out that A is self-adjoint. One can thus obtain a family of operators of the form $T_t : \mathscr{L}^2 \to \mathscr{L}^2$, uniquely defined as

$$T_t = \mathrm{e}^{-A^2 t}$$

Note that $\mathrm{e}^{-A^2 t}$ is defined on all of $\mathscr{L}^2$ even though A is only defined on H^1; we refer the reader to [17] for a very clear introduction to these considerations. If Ω has Lipschitz boundary (i.e., the boundary can locally be represented as the epigraph of a Lipschitz function), one can then (not necessarily uniquely) define a strong Markov process $\{Y_t\}_{t\geq 0}$ such that

$$\mathbb{E}\left[f(Y_{t+s})|Y_t = y\right] = (T_t f)(y)$$

almost surely with respect to Lebesgue measure, for every $f \in \mathscr{L}^2$. In this case we may say that Y is "weakly determined" by $\mathscr{E}$.

It is shown in [2] that if Ω is bounded with Lipschitz boundary, then *any* process which is weakly determined by $\mathscr{E}$ will be a reflecting Brownian motion driven by some Brownian motion W. It is also shown that at least one such process exists. Since a bounded convex set automatically has a Lipschitz boundary (cf. Corollary 1.2.2.3 of [10]) and Tanaka showed that reflecting Brownian motion on a convex set is uniquely defined, it follows that the process X must be weakly determined by $\mathscr{E}$.

This allows us to prove all three of our claims:

1. Since [9] shows that every process with Lipschitz boundary that is weakly determined by $\mathscr{E}$ is a strong Markov process, it follows that X is a strong Markov process. The reversibility of X with respect to σ then follows from the fact that T_t is self-adjoint as an operator on $\mathscr{L}^2(\Omega, \sigma)$ (this, in turns follows from the fact that A is self-adjoint).
2. In [4] it is shown that if a process $\{Y_t\}_{t\geq 0}$ is weakly determined by $\mathscr{E}$ and Ω is convex, then

$$\lim_{t\to\infty} \sup_y \sup_A |\mathbb{P}(Y_t \in A|Y_0 = y) - \sigma(A)| = 0$$

   Thus, the same follows for our process, X.
3. Let λ denote any constant so that $\int f\,\mathrm{d}x = 0 \implies \lambda^2 \|f\|^2_{\mathscr{L}^2} \leq \mathscr{E}(f, f)$. Recall that we have said there is a unique operator $A : H^1 \to \mathscr{L}^2$ such that $\mathscr{E}(f, g) = \int (Af)(Ag)\,\mathrm{d}x$. In particular, $\mathscr{E}(f, f) = \|Af\|^2_{\mathscr{L}^2}$. We may thus rephrase our understanding of λ by saying that $\int f\,\mathrm{d}x = 0 \implies \lambda \|f\|_{\mathscr{L}^2} \leq \|Af\|_{\mathscr{L}^2}$. Using spectral methods it is thus straightforward to show that

$$\int f\,\mathrm{d}x = 0 \implies \left\|\mathrm{e}^{-A^2 t} f\right\|^2_{\mathscr{L}^2} \leq \mathrm{e}^{-\lambda^2 t} \|f\|^2_{\mathscr{L}^2} \tag{3}$$

Using this and Cauchy–Schwarz, one can readily show that

$$|\mathbb{E}[f(X_t)|X_0 \sim \rho] - \mathbb{E}[f(Y)]| \leq \mathrm{e}^{-\frac{1}{2}\lambda^2} \left\| g - \frac{1}{\mathrm{Vol}(\Omega)} \right\|_{\mathscr{L}^2} \|f\|_{\mathscr{L}^2}$$

□

Corollary 1 *Let $Y \sim \sigma$, $\rho(\mathrm{d}x) = g(x)\mathrm{d}x$, $g \geq 0$ and $\rho(\Omega) = 1$. Then*

$$|\mathbb{E}[f(X_t)|X_0 \sim \rho] - \mathbb{E}[f(Y)]| \leq \exp\left(-\frac{1}{2}\left(\frac{\pi}{d}\right)^2 t\right) \|g - \sigma\|_{\mathscr{L}^2} \|f\|_{\mathscr{L}^2}$$

Proof The work in [3] shows that $\lambda = \pi/d$ fills the required role for statement 3 of Lemma 1. □

This last corollary gives a satisfying grip on the $\mathscr{L}^2$ ergodic convergence for Brownian motion in convex domains. Our endeavor here is to complement this with a comparable analysis of the total variation convergence.

Toward this end, we will employ a coupling construction. That is, we will construct a joint process $\{X_t, Y_t\}_{t\geq 0}$ so that X and Y both carry the law of reflecting Brownian motion, but each has a different initial condition. We will construct them in such a way that $\tau = \inf\{t : X_t = Y_t\}$ is almost surely finite:

Theorem 3 (Atar and Burdzy [1]) *Fix any bounded convex set Ω. Let $\{W_t\}_{t\geq 0}$ denote a Brownian motion. Then there exists a pathwise-unique solution to the equations*

$$\begin{aligned} X_t &= x + W_t + L_t \in \bar{\Omega} & |L|_t &= \int_0^t \mathbb{I}_{X_s \in \partial\Omega} \mathrm{d}\,|L|_s < \infty \\ Y_t &= y + Z_t + M_t \in \bar{\Omega} & L_t &= -\int_0^t \boldsymbol{n}_L(s)\mathrm{d}\,|L|_s \\ Z_t &= W_t - \int_0^t 2\eta_s \langle \eta_s, \mathrm{d}W_s \rangle & |M|_t &= \int_0^t \mathbb{I}_{Y_s \in \partial\Omega} \mathrm{d}\,|M|_s < \infty \\ \eta_t &= \tfrac{X_t - Y_t}{|X_t - Y_t|} & M_t &= -\int_0^t \boldsymbol{n}_M(s)\mathrm{d}\,|M|_s \end{aligned} \tag{4}$$

Here $\mathrm{d}\,|L|_s$ plays the role of the measure μ in Definition 2; we require that $\boldsymbol{n}_L(s)$ is a normal vector of supporting hyperplanes of Ω at X_s, $\mathrm{d}\,|L|_s$-almost surely. Likewise for $\boldsymbol{n}_M$, Y_s, $\mathrm{d}\,|M|_s$. Let $\tau = \inf\{t : X_t = Y_t\}$ and

$$\tilde{Y}_t = \begin{cases} Y_t & t \leq \tau \\ X_t & t \geq \tau \end{cases}$$

Then $\left\{\left(X_t, \tilde{Y}_t\right)\right\}_{t\geq 0}$ constitute a strongly Markovian process, and both X and Y are reflecting Brownian motions.

Proof We refer the reader to the work of Atar and Burdzy in [1]. Note that although this article focuses on the case that $\partial\Omega$ is smooth, it also mentions that all of the reasoning goes through for any set which is "admissible" according to the lights of work by Lions and Sznitman [15]. Convex sets are indeed "admissible" according to Remark 3.1 of the work by Lions and Sznitman. □

We emphasize that even though X and Y in this theorem are profoundly coupled, individually they both behave like Brownian motions trapped inside Ω by reflecting boundaries. It is also worth emphasizing that there are two completely different conceptual "reflections" at play here:

1. The normally reflecting boundaries keep X, Y inside Ω
2. The mirror coupling causes Y to generally behave like the mirror image of X, reflected over a plane halfway between X and Y.

These two kinds of reflections may interact when X or Y hits a point in $\partial\Omega$. In this case the direction of reflection (which is mathematically expressed as η_t) may rotate.

Finally, we turn to the function $F_d(t,k)$. This function comes up several times for different reasons; here we hope to clarify the relationships between these different forms.

Lemma 2 (Connections between $F_d(t,k)$ and one-dimensional Brownian motion) *Let W_t denote a one-dimensional reflected Brownian motion trapped inside $[-d,d]$. Let $\tau_d = \inf\{t : W_t \notin (-d,d)$ and $\tau_0 = \inf\{t : W_t = 0\}$. Let*

$$F_d(t,k) \triangleq \sum_{n=0}^{\infty} \mathrm{e}^{-\frac{\pi^2}{8d^2}(2n+1)^2 t} \frac{4(-1)^n}{\pi(2n+1)} \cos\left(\frac{2n+1}{2} \times \frac{\pi k}{d}\right)$$

Then

1. $\mathbb{P}(t \le \tau_d | W_0 = w) = F_d(t,w)$.
2. $\mathbb{P}(t \le \tau_0 | W_0 = w \ge 0) = F_d(t, d-w)$.
3. $\mathbb{P}(W_t \le 0 | W_0 = w \ge 0) = \frac{1}{2} + \frac{1}{2} F_d(t, d-w)$.

Proof The foundations of these results may be found in [14]; the particular function $F_d(t,k)$ can be found in [11], Equations (3.8) and (3.9). Here we give a sketch.

1. Let X_t denote the process W_t killed at $-d$ and d. That is, $X_t = W_t$ for $t \le \tau_d$ but $X_t = \perp$ for $t \ge \tau_d$, where $\perp$ designates a graveyard state. Then for any function f on $(-d,d)$, we may define $f(\perp) = f(-d) = f(d) = 0$ and consider

$$\phi(w,t) = \mathbb{E}[f(X_t) | W_0 = w]$$

We can make a slight extension to the Dirichlet form arguments used in Lemma 1 to account for this kind of killed Brownian motion. In brief, we replace the Sobolev space H^1 with the Sobolev space H_0^1 of functions which vanish at the boundary. The operator A of Lemma 1 turns out to be the derivative, so the equation $\phi(w,t) = (\mathrm{e}^{-A^2 t} f)(w)$ from that Lemma actually corresponds to the differential equation

$$\phi_t = -\frac{1}{2}\phi_{ww} \qquad \phi(w,0) = f(w) \qquad f(-d,t) = f(d,t) = 0$$

(Here we have adopted the notation that subscripts indicate derivatives.) We can easily solve this function for special trigonometric functions of the form $f(x) = \cos(n\pi x/2d)$. Indeed, in this case we have $\phi(w,t) = \cos(n\pi w/2d)\mathrm{e}^{-tn^2\pi^2/8d^2}$. Now our primary interest is in $\mathbb{P}(t \le \tau_d | W_0 = w)$, which can be articulated as $\mathbb{E}[f(X_t)|W_0 = w]$ in the case $f(x) = 1$ for all $x \in (-d,d)$ and $f(x) = 0$ otherwise. Thus, we can obtain our object of interest if we can write this f in a spectral representation, i.e., in terms of the nice trigonometric functions. This is easily done:

$$f(x) = \sum_{n=0}^{\infty} \frac{4(-1)^n}{\pi(2n+1)} \cos\left(\frac{2n+1}{2} \times \frac{\pi x}{d}\right) = \begin{cases} 1 & x \in (-d,d) \\ 0 & x \in \{-d,d\} \end{cases}$$

Using linearity (and checking to make sure nothing goes wrong in taking the limits (cf. [9,14]), we conclude that $\mathbb{E}[f(X_t)|W_0 = w]$ is given by $F_d(t,w)$, so the same must be true of $\mathbb{P}(t \le \tau_d | W_0 = w)$.

2. This is nearly the same problem. Only the boundary conditions are slightly different. In particular, we have a graveyard state at $W = 0$ and a *reflection* at $W = d$. These translate into the boundary conditions $\phi(0,t) = 0$ and $\phi_w(d,t) = 0$. However, applying the symmetry of the problem, we can readily show that this is equivalent to the boundary conditions $\phi(0,t) = 0$ and $\phi(2d,t) = 0$, which reduces to the same problem we just handled, only offset by d and reflected. Equivalently, one may consider that the time it takes W to hit $-d$ or d starting from w must be the same as the time it takes a simple (unreflected) Brownian motion to hit 0 or $2d$ starting at $d - w$, and by symmetry this must be the same as the time it takes the reflected Brownian motion to hit 0 starting from $d - w$.
3. This problem carries the reflecting boundary conditions $\phi_w(-d,t) = \phi_w(d,t) = 0$ and the initial condition $\phi(w,t) = 1$ for $w < 0$ and $\phi(w,t) = 0$ for $w > 0$. This is just our first problem, vertically scaled by $1/2$, vertically shifted by $+1/2$ and a horizontally shifted by d.

□

3 Application of Known Results

To show our main theorem, we must understand the distribution of the coupling time in the mirror construction from Theorem 3. Exact calculation of this distribution may be impossible, but some bounds are straightforward to obtain:

Lemma 3 *Let Ω, X, Y, $\tilde{Y}$, τ be as in Theorem 3. Let d denote the diameter of Ω. Then*

$$\mathbb{P}_x(t \le \tau) \le F_d(4t, d - |x - y|)$$

Proof Let us consider

$$R_t = |X_t - Y_t|$$

The key question is how long it takes before $R_t = 0$. Ito's Lemma yields that

$$R_t = |x - y| + \int_0^t \left\langle \eta_s, \boldsymbol{n}_M(s)\mathrm{d}\,|M|_s - \boldsymbol{n}_L(s)\mathrm{d}\,|L|_s \right\rangle + 2\left\langle \eta_s, \mathrm{d}W_s \right\rangle$$

Dambis–Dubins–Schwarz then yields that we can find some one-dimensional Brownian motion B such that

$$R_t = |x - y| + \int_0^t \left\langle \eta_s, \boldsymbol{n}_M(s)\mathrm{d}\,|M|_s - \boldsymbol{n}_L(s)\mathrm{d}\,|L|_s \right\rangle + B_{4t}$$

Let us now inspect

$$\Phi_t = \int_0^t \left\langle \eta_s, \boldsymbol{n}_M(s)\mathrm{d}\,|M|_s - \boldsymbol{n}_L(s)\mathrm{d}\,|L|_s \right\rangle$$

Let us focus on three properties of this object:

1. $t \mapsto \Phi_t$ is monotone decreasing. Indeed, recall that $\boldsymbol{n}_M(s)$ is a supporting hyperplane of Ω at X_s, $\mathrm{d}\,|M|_s$ almost surely. That is,

   $$\Omega \subset \{y : \langle \boldsymbol{n}_M(s), y - X_t \rangle \geq 0\}$$

   So certainly

   $$\langle \eta_s, \boldsymbol{n}_M(s) \rangle = \left\langle \frac{X_s - Y_s}{|X_s - Y_s|}, \boldsymbol{n}_M(s) \right\rangle \leq 0$$

   The same arguments apply to $-\langle \eta_s, \boldsymbol{n}_L(s) \rangle$.
2. Since the diameter of Ω is d, we have that $R_t \leq d$ for all t. Put another way, $\Phi_t \leq d - B_{4t} - |x - y|$.
3. $\Phi_0 = 0$.

Putting these facts together, we obtain the overall bound of

$$\Phi_t \leq \inf_{s \leq t} (d - B_{4t} - |x - y|) \wedge 0$$

And thus,

$$R_t \leq |x - y| + B_{4t} + \inf_{s \leq t} (d - B_{4t} - |x - y|) \wedge 0 \triangleq \tilde{R}_{4t}$$

This is useful because the law of $\tilde{R}_t$ is well understood. It is that of a one-dimensional Brownian motion with reflection at the point d, initialized at $\tilde{R}_0 = |x - y|$. (Indeed, it is easy to see that it satisfies Definition 2.) Intuitively, this signifies that the amount of time it takes R to go to zero is *shorter* than the amount of time it takes a (time-rescaled) reflected one-dimensional Brownian motion to hit zero.

In particular, let $\tilde{\tau} = \inf\left\{t : \tilde{R}_t = 0\right\}$. Since $R_t \leq \tilde{R}_{4t}$ for every t, it follows that $\tilde{\tau} \geq 4\tau$. On the other hand, Lemma 2 yields

$$\mathbb{P}_x\left(t \leq \tilde{\tau}\right) = F_d\left(t, d - |x - y|\right)$$

Our result follows immediately. □

This leads to our main theorem, which we here restate here the convenience of the reader:

Theorem 4 *Let $\Omega \subset \mathbb{R}^n$ bounded, open, convex, with diameter d. Let $\{p\left(t, x, \mathrm{d}y\right)\}_{t \geq 0, x \in \Omega}$ denote the transition measures of Brownian motion trapped inside Ω by normally reflecting barriers. Then*

$$\begin{aligned}\|p\left(t, x, \cdot\right) - p\left(t, y, \cdot\right)\|_{\mathrm{TV}} &\leq F_d\left(4t, d - |x - y|\right)\\ \|p\left(t, x, \cdot\right) - \sigma\|_{\mathrm{TV}} &\leq \int F_d\left(4t, d - |x - y|\right)\sigma\left(\mathrm{d}y\right) \leq F_d\left(4t, 0\right)\end{aligned}$$

This very last bound cannot be improved by more than a factor of 2. In particular, taking the special case that $\Omega = [-d/2, d/2] \times [0, \epsilon]^{n-1} \subset \mathbb{R}^n$ for any $\epsilon > 0$, we have that $\|p\left(t, d/2, \cdot\right) - \sigma\|_{\mathrm{TV}} \geq F_d\left(4t, 0\right)/2$.

Proof The total variation distance between $p\left(t, x, \cdot\right) - p\left(t, y, \cdot\right)$ is easy to bound using Lemma 3. Recall that $X_t, \tilde{Y}_t$ were both reflecting Brownian motions, with $X_0 = x$, $\tilde{Y}_0 = y$, and $X_t = \tilde{Y}_t$ for $t \geq \tau$. The lemma then shows that $\mathbb{P}\left(t < \tau\right) \leq F_d\left(4t, d - |x - y|\right)$. We thus obtain our first total variation bound:

$$\begin{aligned}\|p\left(t, x, \cdot\right) - p\left(t, y, \cdot\right)\|_{\mathrm{TV}} &= \sup_A \left|\mathbb{P}\left(X_t \in A\right) - \mathbb{P}\left(\tilde{Y}_t \in A\right)\right|\\ &\leq \mathbb{P}\left(X_t \neq \tilde{Y}_t\right) \leq \mathbb{P}\left(t < \tau\right) \leq F_d\left(4t, d - |x - y|\right)\end{aligned}$$

The second total variation bound then follows immediately from the fact that σ is the stationary distribution of the process, i.e., $\int p\left(t, x, A\right)\sigma\left(\mathrm{d}x\right) = \sigma\left(A\right)$. The second total variation bound is thus a kind of "average" of the first variation bound:

$$\begin{aligned}\|p\left(t, x, \cdot\right) - \sigma\|_{\mathrm{TV}} &= \sup_A \left|p\left(t, x, A\right) - \int p\left(t, y, A\right)\sigma\left(\mathrm{d}y\right)\right|\\ &\leq \int \|p\left(t, x, \cdot\right) - p\left(t, y, \cdot\right)\|_{\mathrm{TV}}\,\sigma\left(\mathrm{d}y\right)\\ &\leq \int F_d\left(t, d - |x - y|\right)\sigma\left(\mathrm{d}y\right)\end{aligned}$$

Finally, it is well known that $\sup_k F_d\left(t, k\right) = F_d\left(t, 0\right)$, so we obtain the overall bound of $\|p\left(t, x, \cdot\right) - \sigma\|_{\mathrm{TV}} \leq F_d\left(t, 0\right)$.

Finally we turn to how much this bound could be further improved: any upper bound on the rate of uniform ergodicity which relies only on diameter and dimension

cannot be smaller than the bound given here by more than a factor of two. Indeed, consider any pipe $\Omega = [-d/2, d/2] \times [0, \epsilon]^{n-1} \subset \mathbb{R}^n$ for $\epsilon > 0$. Notice that the law of X can be decomposed into n independent one-dimensional reflecting Brownian motions, and so the total variation can likewise be decomposed. The total variation distance may therefore be lower bounded by the total variation distance along any dimension. Thus, since we seek a lower bound, without loss of generality, let us assume $\Omega = [-d/2, d/2]$. Then $\|p(t, x, \cdot) - \sigma\|_{\mathrm{TV}}$ may further be lower bounded by $|\mathbb{P}(X_t \leq 0 | X_0 = d/2) - 1/2|$. The relevant probability is given by the third result of Lemma 2:

$$\mathbb{P}(X_t \leq 0 | X_0 = d/2) = \frac{1}{2} + \frac{1}{2} F_{d/2}(t, d/2 - d/2) = \frac{1}{2} + \frac{1}{2} F_d(4t, 0)$$

which yields the desired result. □

4 Conclusions

Among all convex sets with diameter d, this work begins to suggest that the one-dimensional interval (e.g., $\Omega = [0, d]$) may provide a kind of worst-case scenario for mixing rates. This is helpful because the one-dimensional interval is easy to analyze.

However, "typical" high-dimensional sets of diameter d may mix much faster than our bounds would suggest. Thus, Eq. (1) from [16] seemed to imply that mixing might get slower in high dimensions, our theorem suggests that the mixing does not depend upon the dimension, but the reality may be that mixing typically gets faster as in higher dimensions. For example, preliminary analysis suggests that n-dimensional Brownian motion in the unit n-dimensional ball mixes quite a bit faster than one-dimensional Brownian motion on $[0, d]$, especially as $n \to \infty$. On the other hand, we are simply not sure about the total variation mixing rate for Brownian motion in a high-dimensional cube. Are there simple ways to improve our bounds when we know more about Ω? This is a possible direction of future research.

In another direction, it should be possible to extend this basic method of proof to accommodate a wide variety of Ito diffusions, beyond Brownian motion. Mirror couplings are available for many such diffusions; we refer the reader to [13] and the many papers which have cited it. For every such mirror coupling process, $\{X_t, Y_t\}_{t \geq 0}$, one can analyze the one-dimensional process $R_t = |X_t - Y_t|$. By applying Ito's lemma, Dambis–Dubins–Schwarz, and taking bounds, one can often obtain a stochastic differential inequality of the form $\mathrm{d}R_t \leq \mathrm{d}\tilde{R}_t$, where $\tilde{R}_t$ is some semimartingale that is better understood. Applying stochastic differential inequality results such as those found in [6], one can then obtain bounds for the high-dimensional process with some simple one-dimensional process.

For example, consider the stochastic differential equation $\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + W_t$ where μ is some Lipschitz vector field. Similar to the technique we used, let $\mathrm{d}Y_t = \mu(Y_t)\mathrm{d}t - \mathrm{d}W_t - 2\eta_t \langle \eta_t, \mathrm{d}W_t \rangle$, where η is a normalized version of $X - Y$. As in our case, if we define $B_{4t} = \int_0^t 2 \langle \eta_s, \mathrm{d}W_s \rangle$ then we can show that B is a one-dimensional Brownian motion. Finally, let Γ denote some Lipschitz function satisfying $\Gamma(r) \geq$

$\sup_{|x-y|=r} \langle x-y, \mu(x)-\mu(y)\rangle$, and consider the simple one-dimensional stochastic differential equation

$$Z_t = |x-y| + \int_0^t \frac{\Gamma(Z_s)}{Z_s} \mathrm{d}s + B_{4t}$$

Using the results from [6], one can readily show that $R_t \leq Z_t$. Thus,

$$\inf\{t : X_t = Y_t\} \leq \inf\{t : Z_t = 0\}$$

In short, by analyzing the simple one-dimensional diffusion Z, one can estimate coupling times for extremely complex and high-dimensional processes. These coupling times can then be used to estimate rates of ergodicity. Of course, these kinds of estimates may be quite poor in some cases (e.g., consider the catastrophic case that $\sup_{|x-y|=r} \langle x-y, \mu(x)-\mu(y)\rangle = \infty$). This difficulty is related to the problem with which we began these conclusions: in some cases it may be very difficult to get high-quality bounds using only a one-dimensional diffusion.

We have seen that one-dimensional diffusions can be used to analyze very high-dimensional diffusions, although the bounds may not be ideal in certain cases. Might it be possible to improve this basic technique to allow two or three-dimensional diffusions to give rigorous bounds on high-dimensional processes? This is an intriguing question for future research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Atar, R., Burdzy, K.: On neumann eigenfunctions in lip domains. J. Am. Math. Soc. **17**(2), 243–265 (2004)
2. Bass, R.F., Hsu, P.: The semimartingale structure of reflecting Brownian motion. In: Proceedings of the American Mathematical Society, pp. 1007–1010 (1990)
3. Bebendorf, M.: A note on the poincaré inequality for convex domains. J. Anal. Appl. **22**, 751–756 (2003)
4. Burdzy, K., Chen, Z.Q., Marshall, D.E.: Traps for reflected brownian motion. Math. Z. **252**(1), 103–132 (2006)
5. Cannon, J.: The One-Dimensional Heat Equation. Encyclopedia of Mathematics, vol. 23. Cambridge University Press, Cambridge (1984)
6. Ding, X., Wu, R.: A new proof for comparison theorems for stochastic differential inequalities with respect to semimartingales. Stoch. Process. Appl. **78**(2), 155–171 (1998)
7. Dupuis, P., Ishii, H.: On lipschitz continuity of the solution mapping to the skorokhod problem, with applications. Stoch. Int. J. Probab. Stoch. Process. **35**(1), 31–62 (1991)
8. Fitzsimmons, P., Pitman, J.: Kac's moment formula and the feynman-kac formula for additive functionals of a markov process. Stoch. Process. Appl. **79**(1), 117–134 (1999)
9. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes, vol. 19. Walter de Gruyter, Berlin (2010)
10. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. SIAM (2011)

11. Hu, Q., Wang, Y., Yang, X.: The hitting time density for a reflected brownian motion. Comput. Econ. **40**(1), 1–18 (2012)
12. Khas' minskii, R.: On positive solutions of the equation au + vu = 0. Theory Probab. Appl. **4**(3), 309–318 (1959)
13. Lindvall, T., Rogers, L.C.G.: Coupling of multidimensional diffusions by reflection. Ann. Probab. **14**, 860–872 (1986)
14. Linetsky, V.: Lookback options and diffusion hitting times: a spectral expansion approach. Finance Stoch. **8**(3), 373–398 (2004)
15. Lions, P.L., Sznitman, A.S.: Stochastic differential equations with reflecting boundary conditions. Commun. Pure Appl. Math. **37**(4), 511–537 (1984)
16. Matthews, P.: Mixing rates for brownian motion in a convex polyhedron. J. Appl. Probab. **27**(02), 259–268 (1990)
17. Schmüdgen, K.: Unbounded Self-adjoint Operators on Hilbert Space, vol. 265. Springer', Berlin (2012)
18. Tanaka, H.: Stochastic differential equations with reflecting. Stoch. Process. Sel. Pap. Hiroshi Tanaka **9**, 157 (1979)