Authors:	Richard F. Bass, Krzysztof Burdzy, Zhen-Qing Chen and Martin Hairer
Title:	Stationary distributions for diffusions with inert drift
Year of publication:	2009
Link to published version:	http://dx.doi.org/10.1007/s00440-008-0182-6
Publisher statement:	The original publication is available at www.springerlink.com

Stationary distributions for diffusions with inert drift *

Richard F. Bass, **Krzysztof Burdzy**, **Zhen-Qing Chen** and **Martin Hairer**

(October 31, 2007)

Abstract

Consider a process reflected inside a domain in $\mathbb{R}^d$ which acquires drift in proportion to the amount of local time spent on the boundary of the domain. We show that the stationary distribution for the vector of two processes, the position of the reflected process and the value of the drift vector, has a product form. Moreover, the first component is uniform measure on the domain, and the second component is a Gaussian distribution. We also consider processes where the drift is given in terms of the gradient of a potential.

1 Introduction

Our paper is concerned with the stationary distribution for a process originally introduced by Knight [18]. Computer simulations presented in [9] led to the conjecture that the stationary distribution for the process has a certain interesting structure. We prove this conjecture, and moreover answer questions about the stationary distribution left open in [9].

The model has two mathematically equivalent representations in one dimension. We will first informally describe the original one-dimensional model introduced in [18]. Consider two stochastic processes Z_t and Y_t with values in $\mathbb{R}$. The dynamics of the system (Z_t, Y_t) are the following. Assume that $Z_0 > Y_0$; we will always have $Z_t \geq Y_t$, by construction. As long as the two processes stay away from each other, Z_t moves at a constant speed, and the second process, Y_t, moves as a Brownian motion B_t. Assume that the initial speed of Z is 0. When the particles meet, they reflect from each other instantaneously. Let L_t be the local time spent by $Z_t - Y_t$ at 0. The velocity V_t of Z_t increases proportionally to L_t, that is, $V_t = c_1 L_t$, and hence $Z_t = Z_0 + \int_0^t c_1 L_s ds$. The effect of the reflection on the second particle is that of a shift, that is $Y_t = B_t - c_2 L_t$. After a finite random time T, the velocity of Z_t will increase to the point that Z_t will escape to infinity with a constant speed and Y_t will never hit Z_t again after time T. The construction of the process, proof of its existence, and the finiteness of the time T can be found in [18, 29, 30].

We will transform this one-dimensional model to obtain a representation readily amenable to multi-dimensional generalizations. Recall that $Z_t = \int_0^t c_1 L_s ds$. We subtract $\int_0^t c_1 L_s ds$ from

*Research supported in part by NSF Grants DMS-0601783 and DMS-0600206 and by EPSRC Grant EP/D071593/1. The second and fourth authors gratefully acknowledge hospitality and support of Institut Mittag-Leffler where part of this research was done.

both components of (Z_t, Y_t) so that we obtain $(0, X_t) := (0, Y_t - Z_t)$. The process X_t has the representation $X_t = B_t - c_2 L_t - \int_0^t c_1 L_s ds$. The process X_t is confined to $(-\infty, 0]$. Now L_t represents the local time spent by X_t at 0, the boundary of the fixed domain $(-\infty, 0]$.

Next, we present a multidimensional generalization of the model, still in an informal way. We consider a bounded multidimensional domain D with a smooth boundary ∂D. The process X_t moves like a Brownian motion B_t with a drift K_t when it is away from ∂D. Let $\mathbf{n}(x)$ be the unit inward normal vector to ∂D at $x \in \partial D$ and let L_t be the local time of X on ∂D. When X hits ∂D at time s, two things happen. First, the process instantly reflects along the normal inner vector, that is,

$$dX_s = c_2 \mathbf{n}(X_s) dL_s,$$

where $c_2 > 0$ is a constant chosen so that X always stays within the domain D. Second, the drift K_t is modified as follows,

$$dK_s = c_1 \mathbf{n}(X_s) dL_s,$$

where $c_1 > 0$ is an arbitrary positive constant. Overall, the process X_t has the following representation,

$$X_t = X_0 + B_t + \int_0^t c_2 \mathbf{n}(X_s) dL_s + \int_0^t K_s ds,$$

where

$$K_t = K_0 + \int_0^t c_1 \mathbf{n}(X_s) dL_s.$$

We call the process K "inert drift" because it plays a role analogous to the inert particle in Knight's original model [18].

The main simulation result of [9] suggests that a stationary distribution for (X, K) exists and has a product form, i.e., that X_t and K_t are independent for each time t under the stationary distribution. Moreover, the first component of the stationary distribution is the uniform probability measure on D. We will show that the second component of the stationary distribution is Gaussian.

The product form of the stationary distribution was initially a mystery to us, especially since the components X and K of the vector (X, K) are *not* Markov processes. There are models known in mathematical physics when the stationary distribution of a Markov process has a product form although each component of the Markov process is not a Markov process itself. Examples may be found in Chapter VIII of [20], in particular, Theorem 2.1 on page 380. As we will see in Section 4, the explanation of the product form of the stationary distribution in our model is related to considerations of a suitable energy function.

The Gaussian nature of the stationary distribution for K has been known in the one-dimensional case, see [29, 10].

The main goal of this paper is to give an explicit formula for the stationary distribution of normally reflecting Brownian motion with inert drift—see Theorem 6.2. Our argument proving this result consists of a number of steps, some of which may have independent interest. We start by showing in Section 2 weak existence and weak uniqueness of solutions to an SDE representing a large family of diffusions with reflection and inert drift. The main observation of this section is that the reflected diffusion with inert drift can be obtained from reflected diffusion without inert drift by a suitable Girsanov transform, and vice versa. The questions of strong existence and strong uniqueness for the same SDE are discussed in Section 3. They are resolved in a positive way under assumptions stronger than those in Section 2. This section uses some ideas and results from [21] but the main idea of our argument is different from the one in that paper, and we believe ours is somewhat simpler. In Section 4 we consider Brownian motion with drift given as the gradient of a smooth potential. This is because the analysis of the stationary distribution is much easier in the case of a smooth potential than the "singular" potential representing reflection on the boundary of a domain. In Section 5, we prove weak convergence of a sequence of Brownian motions with inert drifts corresponding to smooth potentials to normally reflecting Brownian motion with inert drift. This implies existence of a stationary distribution with product form for reflecting Brownian motion with inert drift. Finally, Section 6 completes the program by showing irreducibility in the sense of Harris for reflecting Brownian motion with inert drift. Uniqueness of the stationary distribution follows from irreducibility.

The level of generality of our results decreases throughout the paper, for technical reasons. We leave it as an open problem to prove a statement analogous to Theorem 6.2 in the general setting of Theorem 2.1.

Our model belongs to a family of processes with "reinforcement" surveyed by Pemantle in [25]; see especially Section 6 of that survey and references therein. Papers [5, 6, 7] study a process with a drift defined in terms of a "potential" and the "normalized" occupation measure. While there is no direct relationship to our results, there are clear similarities between that model and ours.

We are grateful to David White for very useful advice.

2 Weak existence and uniqueness

This section is devoted to weak existence and uniqueness of solutions to an SDE representing a family of reflecting diffusions with inert drift.

Let D be a bounded C^2-smooth domain in $\mathbb{R}^d$, $d \geq 2$, and denote by $\mathbf{n}$ the inward unit normal vector field on ∂D. Throughout this paper, all vectors are column vectors. Let $A(x) = (a_{ij}(x))_{n\times n}$ be a matrix-valued function on $\mathbb{R}^d$ that is symmetric, uniformly positive definite, and each a_{ij} is bounded and C^2 on $\overline{D}$. The vector $\mathbf{u}(x) := \frac{1}{2}A(x)\mathbf{n}(x)$ is called the conormal vector at $x \in \partial D$. Clearly there exists $c_1 > 0$ such that $\mathbf{u}(x) \cdot \mathbf{n}(x) > c_1$ for all $x \in \partial D$. Let $\sigma(x) = (\sigma_{ij}(x))$ be the positive symmetric square root of $A(x)$. For notational convenience, we use ∂_i to denote $\frac{\partial}{\partial x_i}$. For $\varphi \in C^2(\mathbb{R}^d)$, let

$$\mathcal{L}\varphi(x) := \frac{1}{2}\sum_{i,j=1}^{d} \partial_i \left(a_{ij}(x)\partial_j\varphi(x)\right).$$

Let $\mathbf{b}(x)$ be the vector whose k^{th} component is

$$b_k(x) = \frac{1}{2}\sum_{i=1}^{d} \partial_i(a_{ik}(x)).$$

Thus b_k is the same as $\mathcal{L}$ operating on the function $f_k(x) = x_k$.

Let B be standard d-dimensional Brownian motion and $\mathbf{v}$ a bounded measurable vector field on ∂D. Consider the following system of stochastic differential equations, with the extra condition that $X_t \in \overline{D}$ for all $t \geq 0$:

$$\begin{cases} dX_t = \sigma(X_t)dB_t + \mathbf{b}(X_t)dt + \mathbf{u}(X_t)dL_t + K_t dt, \\ t \mapsto L_t \text{ is continuous and non-decreasing with } L_t = \int_0^t \mathbf{1}_{\partial D}(X_s)dL_s, \\ dK_t = \mathbf{v}(X_t)dL_t. \end{cases} \tag{2.1}$$

Theorem 2.1 *For every $x \in \overline{D}$ and $y \in \mathbb{R}^d$ there exists a unique weak solution $\{(X_t, K_t), t \in [0,\infty)\}$ to (2.1) with $(X_0, K_0) = (x, y)$.*

Proof. Consider the following SDE,

$$\begin{cases} dX_t = \sigma(X_t)dB_t + \mathbf{b}(X_t)dt + \mathbf{u}(X_t)dL_t, \\ t \mapsto L_t \text{ is continuous and non-decreasing with } L_t = \int_0^t \mathbf{1}_{\partial D}(X_s)dL_s, \end{cases} \tag{2.2}$$

with $X_t \in \overline{D}$ for every $t \geq 0$. Weak existence and uniqueness of solutions to (2.2) follow from [11] or [13] (while we have been informed by the authors that there is a gap in the proof of Case 2 in [13], we need only Case 1). The distribution of the solution to (2.2) with $X_0 = x \in D$ will be denoted by $\mathbb{P}_x$.

Note that the remaining proof uses only the C^1-smoothness of the domain and Lipschitz continuity of a_{ij}. Let $K_t := y + \int_0^s \mathbf{v}(X_s)dL_s$ and $\sigma^{-1}(x)$ be the inverse matrix of $\sigma(x)$. Define for $t \geq 0$,

$$M_t = \exp\left(\int_0^t \sigma^{-1}(X_s)K_s dB_s - \frac{1}{2}\int_0^t |\sigma^{-1}(X_s)K_s|^2 ds\right).$$

It is clear that M is a continuous positive local martingale with respect to the minimal augmented filtration $\{\mathcal{F}_t, t \geq 0\}$ of X. Let $T_n = \inf\{t > 0 : |K_t| \geq 2^n\}$. Since $L_t < \infty$ for every $t < \infty$, $\mathbb{P}_x$-a.s., and $|\mathbf{v}|$ is uniformly bounded, we see that $K_t < \infty$ for every $t < \infty$, $\mathbb{P}_x$-a.s. Hence, $T_\infty := \lim_{n\to\infty} T_n = \infty$, $\mathbb{P}_x$-a.s. For every $n \geq 1$, $\{M_{T_n \wedge t}, \mathcal{F}_t, t \geq 0\}$ is a martingale.

For $x \in \overline{D}$, $y \in \mathbb{R}^d$ and $n \geq 1$, define a new probability measure $\mathbb{Q}_{x,y}$ by

$$d\mathbb{Q}_{x,y} = M_{T_n}\, d\mathbb{P}_x \qquad \text{on } \mathcal{F}_{T_n} \text{ for every } n \geq 1.$$

It is routine to check this defines a probability measure $\mathbb{Q}_{x,y}$ on $\mathcal{F}_\infty$. By the Girsanov theorem (cf. [26]), the process

$$W_t := B_t - \int_0^t \sigma^{-1}(X_s)K_s ds,$$

is a Brownian motion up to time T_n for every $n \geq 1$, under the measure $\mathbb{Q}_{x,y}$. Thus we have from (2.2) that under $\mathbb{Q}_{x,y}$, up to time T_n for every $n \geq 1$,

$$dX_t = \sigma(X_t)dW_t + \mathbf{b}(X_t)dt + \mathbf{u}(X_t)dL_t + K_t dt.$$

In other words, $\{(X_t, K_t), 0 \leq t < T_\infty\}$ under the measure $\mathbb{Q}_{x,y}$ is a weak solution of (2.1).

We make a digression on the use of the strong Markov property. At this point in the proof, we cannot claim that (X, K) is a strong Markov process. Note however that if T is a finite stopping time, then (X_{t+T}, K_{t+T}) is again a solution of (2.1) with initial values (X_T, K_T). We can therefore use regular conditional probabilities as a technical substitute for the strong Markov property; this technique has been described in great detail in Remark 2.1 of [3], where we used the name "pseudo-strong Markov property." Throughout the remainder of this proof we will use the pseudo-strong Markov property in place of the traditional strong Markov property and refer the reader to [3] for details.

We will next show that $T_\infty = \infty$, $\mathbb{Q}_{x,y}$-a.s., i.e., the process is conservative. This is the same as saying $|K|$ does not "explode" in finite time under $\mathbb{Q}_{x,y}$. The intuitive reason why this should be true is the following. Consider a one-dimensional Brownian motion starting at 1 with a very large constant negative drift of size c and reflect it at the origin. Then, a simple calculation shows that the local time accumulated at the origin by this process up to time 1 is also of order c. This suggests that if K is very large, the local time accumulated in a time interval of order 1 by X on the boundary ∂D will be approximately proportional to K. Since this feeds back into the right hand side of the definition of K in (2.1), one would expect K to grow at most exponentially fast in time.

Consider $\varepsilon = 2^{-j} > 0$ such that $j \geq 1$ is an integer. Our argument applies only to small $\varepsilon > 0$ so we will now impose some assumptions on ε. Consider $x_0 \in \partial D$ and let CS_{x_0} be an orthonormal coordinate system such that $x_0 = 0$ in CS_{x_0} and the positive part of the d-th axis contains $\mathbf{n}(x_0)$. Let $\mathbf{n}_0 = \mathbf{n}(x_0)$. Recall that D has a C^1-smooth boundary and that there exists $c_1 > 0$ such that $\mathbf{u}(x) \cdot \mathbf{n}(x) > c_1$ for all $x \in \partial D$. Hence, there exist $\varepsilon_0 > 0$ and $c_2 > 0$, such that for all $\varepsilon \in (0, \varepsilon_0)$, every $x_0 \in \partial D$ and all points $x = (x_1, \dots, x_d) \in \partial D \cap B(x_0, (6/c_2 + 5)\varepsilon)$, we have $|x_d| < \varepsilon/2$ and $\mathbf{u}(x) \cdot \mathbf{n}_0 > c_2$, in CS_{x_0}. Since $|\mathbf{v}(x)| \leq c_3 < \infty$ for all $x \in \partial D$, we can make $\varepsilon_0 > 0$ smaller, if necessary, so that $(2c_2)/(c_3\varepsilon) - 5\varepsilon \geq \varepsilon$ for all $\varepsilon \in (0, \varepsilon_0]$.

Let

$$S_0 = \inf\{t > 0 : |K_t| \geq 1/\varepsilon\},$$
$$S_{n+1} = \inf\{t > S_n : | |K_t| - |K_{S_n}| | \geq (6c_3/c_2)\varepsilon\}, \qquad n \geq 0.$$

We will estimate $\mathbb{Q}_{x,y}(S_{n+1} - S_n > \varepsilon^2 \mid \mathcal{F}_{S_n})$ for $n = 0, \dots, [\varepsilon^{-2}c_2/(6c_3)]$. Note that $X_{S_n} \in \partial D$ for every n because K does not change when X is in the interior of the domain. Note also that $|K_{S_n}| \leq 2/\varepsilon$ for every $n \leq \varepsilon^{-2}c_2/(6c_3)$ by construction.

For $n \geq 0$, let

$$Y_t^{(n)} = \int_{S_n}^{S_n+t} \sigma(X_s)dW_s + \int_{S_n}^{S_n+t} \mathbf{b}(X_s)ds,$$
$$F_n = \left\{ \sup_{t\in[0,\varepsilon^2]} |Y_t^{(n)}| < \varepsilon \right\}.$$

It is standard to show that there exist $\varepsilon_0, p_0 > 0$, not depending on n, such that if $\varepsilon < \varepsilon_0$, then

$$\mathbb{Q}_{x,y}(F_n \mid \mathcal{F}_{S_n}) \geq p_0.$$

Let

$$R_n = (S_n + \varepsilon^2) \wedge \inf\{t \geq S_n : |K_t| \geq 4/\varepsilon\}.$$

Suppose that the event F_n holds. We will analyze the path of the process $\{(X_t, K_t), S_n \leq t \leq R_n\}$. First, we will argue that $X_t \in B(X_{S_n}, (6/c_2 + 5)\varepsilon)$ for $S_n \leq t \leq R_n$. Suppose otherwise. Let $U_1 = \inf\{t > S_n : X_t \notin B(X_{S_n}, (6/c_2 + 5)\varepsilon)\}$ and $U_2 = \sup\{t < U_1 : X_t \in \partial D\}$. We have

$$(6/c_2 + 5)\varepsilon = |X_{U_1} - X_{S_n}| = \left| Y_{U_1 - S_n} + \int_{S_n}^{U_1} K_t dt + \int_{S_n}^{U_1} \mathbf{u}(X_t)dL_t \right|.$$

So on $F_n \cap \{U_1 \leq R_n\}$,

$$\begin{aligned}
\left| \int_{S_n}^{U_2} \mathbf{u}(X_t)dL_t \right| &= \left| \int_{S_n}^{U_1} \mathbf{u}(X_t)dL_t \right| \\
&\geq (6/c_2 + 5)\varepsilon - \left| Y_{U_1 - S_n} + \int_{S_n}^{U_1} K_t\, dt \right| \\
&\geq (6/c_2 + 5)\varepsilon - \varepsilon - \int_{S_n}^{S_n+\varepsilon^2} 4/\varepsilon\, dt \\
&\geq (6/c_2 + 5)\varepsilon - \varepsilon - 4\varepsilon \\
&= (6/c_2)\varepsilon.
\end{aligned}$$

We will use the coordinate system $CS_{X_{S_n}}$ to make the following observations. The last formula implies that the d-th coordinate of $\int_{S_n}^{U_2} \mathbf{u}(X_t)dL_t$ is not less than 6ε. Hence, the d-th coordinate of X_{U_2} must be greater than or equal to

$$6\varepsilon - \left| Y_{U_1 - S_n} + \int_{S_n}^{U_1} K_t dt \right| \geq 6\varepsilon - \varepsilon - \int_{S_n}^{S_n+\varepsilon^2} 4/\varepsilon dt \geq 6\varepsilon - \varepsilon - 4\varepsilon = \varepsilon.$$

This is a contradiction because $X_{U_2} \in \partial D$ and the d-th coordinate for all $x \in \partial D \cap B(X_{S_n}, (6/c_2 + 5)\varepsilon)$ is bounded by $\varepsilon/2$. So on F_n, we have $X_t \in B(X_{S_n}, (6/c_2 + 5)\varepsilon)$ for $t \in [S_n, R_n]$.

We will use a similar argument to show that $R_n = S_n + \varepsilon^2$ on F_n. Suppose that $R_n < S_n + \varepsilon^2$. Then $S_n \leq R_n$ and $|K_{R_n}| = 4/\varepsilon$. Let $U_3 = \sup\{t < R_n : X_t \in \partial D\}$. The definition of S_n implies that $K_{S_n} \leq 2/\varepsilon$ for $n \leq [\varepsilon^{-2} c_2/(6c_3)]$. We have

$$\left| \int_{S_n}^{R_n} \mathbf{v}(X_t)dL_t \right| = |K_{R_n} - K_{S_n}| \geq 4/\varepsilon - 2/\varepsilon = 2/\varepsilon.$$

Since $|\mathbf{v}(x)| \leq c_3$, we have $L_{R_n} - L_{S_n} \geq 2/(c_3\varepsilon)$, so the d-th coordinate of $\int_{S_n}^{R_n} \mathbf{u}(X_t)dL_t$, which is the same as $\int_{S_n}^{U_3} \mathbf{u}(X_t)dL_t$, is bounded below by $(2c_2)/(c_3\varepsilon)$. But then on F_n, the d-th coordinate of X_{U_3} must be greater than or equal to

$$(2c_2)/(c_3\varepsilon) - \left|Y_{R_n-S_n} + \int_{S_n}^{R_n} K_t dt\right| \geq (2c_2)/(c_3\varepsilon) - \varepsilon - \int_{S_n}^{S_n+\varepsilon^2} 4/\varepsilon dt = (2c_2)/(c_3\varepsilon) - 5\varepsilon \geq \varepsilon.$$

This is a contradiction because $X_{U_3} \in \partial D$ and the d-th coordinate for all $x \in \partial D \cap B(x_0, (6/c_2+5)\varepsilon)$ is bounded by $\varepsilon/2$. So on F_n we have $R_n = S_n + \varepsilon^2$.

We will now use the same idea to show that on F_n, $L_{S_n+\varepsilon^2} - L_{S_n} < (6/c_2)\varepsilon$. Assume that $L_{S_n+\varepsilon^2} - L_{S_n} \geq (6/c_2)\varepsilon$. Let $U_4 = \sup\{t \leq S_n + \varepsilon^2 : X_t \in \partial D\}$. The d-th coordinate of $\int_{S_n}^{U_4} \mathbf{u}(X_t)dL_t$ is bounded below by 6ε. But then on F_n, the d-th coordinate of X_{U_3} must be greater than or equal to

$$6\varepsilon - \left|Y_{R_n-S_n} + \int_{S_n}^{R_n} K_t dt\right| \geq 6\varepsilon - \varepsilon - \int_{S_n}^{S_n+\varepsilon^2} 4/\varepsilon dt \geq \varepsilon.$$

This is a contradiction because $X_{U_3} \in \partial D$ and the d-th coordinate for all $x \in \partial D \cap B(x_0, (6/c_2+5)\varepsilon)$ is bounded by $\varepsilon/2$. We see that if the event F_n holds, then

$$\left|\,|K_{S_n+\varepsilon^2}| - |K_{S_n}|\,\right| \leq \int_{S_n}^{S_n+\varepsilon^2} |\mathbf{v}(X_t)|dL_t \leq c_3(L_{S_n+\varepsilon^2} - L_{S_n}) \leq c_3(6/c_2)\varepsilon.$$

Hence, if the event F_n holds, then $S_{n+1} > S_n + \varepsilon^2$. We see that

$$\mathbb{Q}_{x,y}(S_{n+1} > S_n + \varepsilon^2 \mid \mathcal{F}_{S_n}) \geq p_0.$$

Recall that we took $\varepsilon = 2^{-j}$ so that $S_0 = T_j$. The last estimate and the pseudo-strong Markov property applied at stopping times S_n allow us to apply some estimates known for a Bernoulli sequence with success probability p_0 to the sequence of events $\{S_{n+1} > S_n + \varepsilon^2\}$. Specifically, there is some $p_1 > 0$ so that for all sufficiently small $\varepsilon = 2^{-j} > 0$,

$$\begin{aligned}
&\mathbb{Q}_{x,y}(T_{j+1} - T_j \geq (c_2/(6c_3))p_0/2 \,\big|\, \mathcal{F}_{T_j}) \\
&\qquad = \mathbb{Q}_{x,y}\left(\sum_{n=0}^{[\varepsilon^{-2}c_2/(6c_3)]} (S_{n+1} - S_n) \geq (c_2/(6c_3))p_0/2 \,\Big|\, \mathcal{F}_{T_j}\right) \\
&\qquad \geq \mathbb{Q}_{x,y}\left(\frac{\varepsilon^2}{c_2/(6c_3)} \sum_{n=0}^{[\varepsilon^{-2}c_2/(6c_3)]} \mathbf{1}_{\{S_{n+1} > S_n + \varepsilon^2\}} \geq p_0/2 \,\Big|\, \mathcal{F}_{T_j}\right) \\
&\qquad \geq p_1.
\end{aligned}$$

Once again, we use an argument based on comparison with a Bernoulli sequence, this time with success probability p_1. We conclude that there are infinitely many n such that $T_{j+1} - T_j \geq$

$(c_2/(6c_3))p_0/2$, $\mathbb{Q}_{x,y}$-a.s. We conclude that $T_\infty = \infty$, $\mathbb{Q}_{x,y}$-a.s., so our process (X_t, K_t) is defined for all $t \in [0, \infty)$.

Next we will prove weak uniqueness. Suppose that $\mathbb{Q}'_{x,y}$ is the distribution of any weak solution to (2.1) and define $\mathbb{P}'_x$ by

$$d\mathbb{P}'_x = \frac{1}{M_{T_n}} d\mathbb{Q}'_{x,y} \qquad \text{on } \mathcal{F}_{T_n} \text{ for every } n \geq 1.$$

Reversing the argument in the first part of the proof, we conclude that X under $\mathbb{P}'_x$ solves (2.2). It follows from the strong uniqueness for (2.2) that $\mathbb{P}'_x = \mathbb{P}_x$, and, therefore, $\mathbb{Q}'_{x,y} = \mathbb{Q}_{x,y}$ on $\mathcal{F}_{T_n}$. Since this holds for all n, we see see that $\mathbb{Q}'_{x,y} = \mathbb{Q}_{x,y}$ on $\mathcal{F}_\infty$. □

Remark 2.2 The assumption that D is a bounded C^2-smooth domain and that a_{ij}'s are C^2 on $\overline{D}$ are only used in the weak existence and uniqueness of solution X to (2.2). Remaining proof only requires that D to be bounded C^1-smooth domain and that a_{ij}'s are Lipschitz on $\overline{D}$. In fact, when D to be bounded C^1-smooth domain and that a_{ij}'s are Lipschitz on $\overline{D}$, the weak existence of solutions to (2.2) follow from [11] and so we have the weak existence to SDE (2.1). We believe the weak uniqueness for solution to (2.2) and consequently to (2.1) also holds on this weakened assumption by an argument analogous to that in [2, Section 4]. However to give the full details of the proof will take a significant number of pages so we leave the details to the reader. □

We remarked earlier that if T is a finite stopping time, then (X_{t+T}, K_{t+T}) is again a solution to (2.1) with starting point (X_T, K_T). This observation together with the weak uniqueness of the solution to (2.1) implies that (X_t, K_t) is a strong Markov process in the usual sense; cf. [1, Section I.5].

3 Pathwise uniqueness

We will prove strong existence and strong uniqueness for solutions to (2.1) under assumptions stronger than those in Section 2, namely, we will assume that D is a bounded C^2-smooth domain in $\mathbb{R}^d$, each a_{ij} is $C^{1,1}$-smooth on $\overline{D}$ and so the vector $\mathbf{b}$ in (2.1) is Lipschitz continuous, and that the vector field $\mathbf{v}$ is a fixed multiple of $\mathbf{u}$. Our approach to the strong existence and uniqueness for solutions to (2.1) uses some ideas and results from [21] but the main idea of our argument is different from the one in that paper, and we believe ours is somewhat simpler. It might be possible to produce a proof along the lines of [21] but a detailed version of that argument adapted to our setting would be at least as long as the one we give here.

Theorem 3.1 *Suppose that $\mathbf{v} \equiv a_0 \mathbf{u}$ for some constant $a_0 \in \mathbb{R}$. For each $(x, y) \in \overline{D} \times \mathbb{R}^d$, there exists a unique strong solution $\{(X_t, K_t), t \in [0, \infty)\}$ to (2.1) with $(X_0, K_0) = (x, y)$.*

Proof. When $a_0 = 0$, $K_t \equiv K_0$. In this case, the result follows from [13] (as mentioned above there is a gap in the proof of Case 2 in [13]; we need only Case 1). So without loss of generality, we assume $a_0 \neq 0$.

First we will prove pathwise uniqueness. Suppose that there exist two solutions (X_t, K_t) and (X'_t, K'_t), driven by the same Brownian motion B, starting with the same initial values $(X_0, K_0) = (X'_0, K'_0) = (x, y)$, and such that $(X_t, K_t) \neq (X'_t, K'_t)$ for some t with positive probability.

According to Lemma 4.1 of [21], there exists a matrix-valued function $\Lambda(x) = \{\lambda_{ij}(x)\}_{1\le i,j\le d}$, $x \in \mathbb{R}^d$ such that $x \to \Lambda(x)$ is uniformly elliptic and in C_b^2 and such that the following two formulas hold. We have

$$\mathbf{u}(x)^T \Lambda(x) = \mathbf{n}(x)^T \qquad \text{for } x \in \partial D. \tag{3.1}$$

Moreover, there exists $c_1 < \infty$ such that

$$c_1|x - x'|^2 + \mathbf{u}(x)^T \Lambda(x)(x - x') \ge 0 \qquad \text{for } x \in \partial D \text{ and } x' \in \overline{D}. \tag{3.2}$$

In our setting, we can take $\Lambda = 2A^{-1}$. Let $c_2 > 1$ be such that

$$c_2^{-1} I_{d\times d} \le \Lambda(x) \le c_2 I_{d\times d} \qquad \text{for every } x \in \mathbb{R}^d, \tag{3.3}$$

where $I_{d\times d}$ is the $d \times d$-dimensional identity matrix. Define $\lambda := c_2^3$.

Let

$$\begin{aligned} U_t &= \int_0^t \sigma(X_t) dB_t + \int_0^t \mathbf{b}(X_t) dt, \\ V_t &= \int_0^t K_t dt, \end{aligned}$$

and define U'_t and V'_t in a similar way relative to X'.

We fix an arbitrary $p_1 < 1$ and an integer k_0 such that in k_0 Bernoulli trials with success probability 1/2, at least $k_0/4$ of them will occur with probability p_1 or greater. Let $t_1 = 1/k_0$.

Consider some $\varepsilon \ge \varepsilon_0 > 0$, $c_0 > 0$, and let

$$\begin{aligned} T_1 &= \inf\{t > 0 : |X_t - X'_t| \ge \varepsilon_0 \text{ or } |U_t - U'_t| \ge \varepsilon_0 \text{ or } |K_t - K'_t| \ge \varepsilon_0\}, \\ T_k &= (T_{k-1} + t_1) \wedge \inf\{t > T_{k-1} : |X_t - X'_t| \vee |U_t - U'_t| \ge \lambda^{k-1}\varepsilon\} \\ &\qquad \wedge\ \inf\{t > T_{k-1} : |L_t - L_{T_{k-1}}| \vee |L'_t - L'_{T_{k-1}}| > c_0\}, \qquad k \ge 2. \end{aligned}$$

We will specify the values of ε_0, ε and c_0 later in the proof.

It is easy to see from (2.1) that it is impossible to have X_t equal to X'_t for all $t \in [0, T_1]$, and at the same time $U_{T_1} \neq U'_{T_1}$ or $K_{T_1} \neq K'_{T_1}$. Hence, $X_{T_1} \neq X'_{T_1}$, a.s. Note that T_1 is a function of ε_0 and set $T_1^* = \lim_{\varepsilon\downarrow 0} T_1(\varepsilon_0)$. Then with probability 1, for every $\delta > 0$ there exists $t_* \in [T_1^*, T_1^* + \delta]$ such that $X_{t_*} \neq X'_{t_*}$.

The idea of our pathwise uniqueness proof is as follows. Define

$$S = (T_1^* + 1/4) \wedge \inf\{t \ge T_1^* : (L_t - L_{T_1^*}) \vee (L'_t - L'_{T_1^*}) \ge c_0/(4t_1)\}.$$

We will show that for any $p_1 < 1$ and integer $k_0 = k_0(p_1)$ as defined above and every $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,S]} |X_t - X'_t| \le \varepsilon\lambda^{k_0}\right) \ge p_1.$$

As $\varepsilon > 0$ is arbitrary, we have

$$\mathbb{P}\left(X_t = X'_t \text{ for every } t \in [0,S]\right) \ge p_1 > 0,$$

which contradicts the definition of T_1^*. This contradiction proves the pathwise uniqueness.

Gronwall's inequality says that if $g(t)$ is nonnegative and locally bounded and $g(t) \le a + b\int_0^t g(s)\,ds$, then $g(t) \le ae^{bt}$. Suppose now that f is a nonnegative nondecreasing function and $g(t) \le f(t) + b\int_0^t g(s)\,ds$. Applying Gronwall's inequality for $t \le t_1$ with $a = f(t_1)$, we have

$$g(t_1) \le f(t_1)e^{bt_1}. \tag{3.4}$$

We apply the inequality with $g(t) = |K_t - K'_t|$ and

$$f(t) = |K_{T_1} - K'_{T_1}| + \sup_{T_1\le s\le t}\left(|X_s - X'_s| + |U_s - U'_s|\right).$$

Since $\mathbf{v} \equiv a_0\mathbf{u}$, we have $\int_0^t \mathbf{u}(X_s)dL_s = \frac{1}{a_0}(K_t - K_0)$ and so

$$K_t - K_0 = a_0(X_t - x_0 - U_t - V_t),$$

and similarly for K'. Thus

$$K_t - K'_t = a_0(X_t - X'_t - U_t + U'_t - V_t + V'_t),$$

and hence for $t \ge T_1$,

$$|K_t - K'_t| \le |K_{T_1} - K'_{T_1}| + 2|a_0|\left(\sup_{T_1\le s\le t}(|X_s - X'_s| + |U_s - U'_s|) + \int_{T_1}^t |K_s - K'_s|\,ds\right).$$

By (3.4), for $t \ge T_1$,

$$|K_t - K'_t| \le e^{2|a_0|(t-T_1)}\left(|K_{T_1} - K'_{T_1}| + 2|a_0|\sup_{T_1\le s\le t}(|X_s - X'_s| + |U_s - U'_s|)\right).$$

Recall that $t_1 = 1/k_0$. It follows that $T_{k_0} - T_1 \le 1$ and $e^{2|a_0|(t-T_1)} \le e^{2|a_0|}$ for $t \le T_{k_0}$. The definition of the T_k's implies that

$$\sup_{0\le s\le T_k}(|X_s - X'_s| + |U_s - U'_s|) \le 2\lambda^{k-1}\varepsilon.$$

We obtain for $t \le T_k$ with $k \le k_0$,

$$|K_t - K'_t| \le e^{2|a_0|}(1 + 4|a_0|)\lambda^{k-1}\varepsilon. \tag{3.5}$$

We have

$$
\begin{aligned}
\Xi &:= (X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k} - X'_{T_k}) - (X_{T_{k-1}} - X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}} - X'_{T_{k-1}}) \\
&= \int_{T_{k-1}}^{T_k} d\left((X_t - X'_t)^T \Lambda(X_t)(X_t - X'_t)\right) \\
&\le c_3 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)((\sigma(X_t) - \sigma(X'_t))dB_t \\
&+ \ c_4 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)(\mathbf{b}(X_t) - \mathbf{b}(X'_t))dt \\
&+ \ c_5 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)(K_t - K'_t)dt \\
&+ \ c_6 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)\mathbf{u}(X_t)dL_t \\
&- \ c_7 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)\mathbf{u}(X'_t)dL'_t \\
&+ \ c_8 \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2(\sigma(X_t)dB_t + \mathbf{b}(X_t)dt + K_t dt + \mathbf{u}(X_t)dL_t).
\end{aligned}
$$

Rewrite the second to the last term as

$$
-c_7 \int_{T_{k-1}}^{T_k} \Lambda(X'_t)(X_t - X'_t)\mathbf{u}(X'_t)dL'_t - c_7 \int_{T_{k-1}}^{T_k} (\Lambda(X_t) - \Lambda(X'_t))(X_t - X'_t)\mathbf{u}(X'_t)dL'_t.
$$

We apply (3.5) to $\int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)(K_t - K'_t)\, dt$ and use the Lipschitz property of Λ to obtain

$$
\begin{aligned}
\Xi \le\ & c_3 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)((\sigma(X_t) - \sigma(X'_t))dB_t \\
&+ \ c_4 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)(\mathbf{b}(X_t) - \mathbf{b}(X'_t))dt \\
&+ \ c_9 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)\lambda^{k-1}\varepsilon dt \\
&+ \ c_{10} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL_t \\
&+ \ c_{11} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL'_t + \ c_{12} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL'_t \\
&+ \ c_8 \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2(\sigma(X_t)dB_t + \mathbf{b}(X_t)dt + K_t dt + \mathbf{u}(X_t)dL_t).
\end{aligned} \tag{3.6}
$$

It follows from (3.3) that there exists $c_{13} > 0$ such that if

$$
(X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k} - X'_{T_k}) - (X_{T_{k-1}} - X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}} - X'_{T_{k-1}}) \le c_{13}\lambda^{2(k-1)}\varepsilon^2 \tag{3.7}
$$

and

$$|X_{T_{k-1}} - X'_{T_{k-1}}| \le \lambda^{k-2}\varepsilon, \tag{3.8}$$

then

$$|X_{T_k} - X'_{T_k}| \le \lambda^{-1/3}\,\lambda^{k-1}\varepsilon < \lambda^{k-1}\varepsilon. \tag{3.9}$$

In view of our assumptions on Λ and σ,

$$\begin{aligned}
&\mathrm{Var}\Big(c_3 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)((\sigma(X_t) - \sigma(X'_t))dB_t \;\Big|\; \mathcal{F}_{T_{k-1}}\Big)\\
&\qquad \le \mathbb{E}\left(c_{14} \int_{T_{k-1}}^{T_k} \sup_{T_{k-1}\le s\le T_k} |X_s - X'_s|^4 dt \;\Big|\; \mathcal{F}_{T_{k-1}}\right)\\
&\qquad \le c_{14}\lambda^{4(k-1)}\varepsilon^4 t_1.
\end{aligned}$$

We make $t_1 > 0$ smaller (and therefore k_0 larger), if necessary, so that by Doob's and Chebyshev's inequalities,

$$\mathbb{P}\left(\left|c_3 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)((\sigma(X_t) - \sigma(X'_t))dB_t\right| \ge (1/10)c_{13}\lambda^{2(k-1)}\varepsilon^2 \;\Big|\; \mathcal{F}_{T_{k-1}}\right) \le 1/8. \tag{3.10}$$

We make $t_1 > 0$ smaller, if necessary, so that

$$\begin{aligned}
\left|c_4 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)(\mathbf{b}(X_t) - \mathbf{b}(X'_t))dt\right| &\le c_{15} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dt \\
&\le c_{15} \int_{T_{k-1}}^{T_k} \sup_{T_{k-1}\le s\le T_k} |X_s - X'_s|^2 dt \\
&\le c_{15} t_1 \lambda^{2(k-1)}\varepsilon^2 \le (1/10)c_{13}\lambda^{2(k-1)}\varepsilon^2.
\end{aligned} \tag{3.11}$$

If $t_1 > 0$ is sufficiently small then

$$\begin{aligned}
\left|c_9 \int_{T_{k-1}}^{T_k} \Lambda(X_t)(X_t - X'_t)\lambda^{k-1}\varepsilon dt\right| &\le c_{17} \int_{T_{k-1}}^{T_k} \sup_{T_{k-1}\le s\le T_k} |X_s - X'_s|\lambda^{k-1}\varepsilon dt \\
&\le (1/10)c_{13}\lambda^{2(k-1)}\varepsilon^2.
\end{aligned} \tag{3.12}$$

We make $c_0 > 0$ in the definition of T_k so small that

$$c_{10} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL_t + c_{11} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL'_t + c_{12} \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 dL'_t \le (1/10)c_{13}\lambda^{2(k-1)}\varepsilon^2. \tag{3.13}$$

Completely analogous estimates show that for sufficiently small $t_1, c_0 > 0$,

$$\mathbb{P}\Big(\Big|c_8 \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2(\sigma(X_t)dB_t + \mathbf{b}(X_t)dt + K_t dt + \mathbf{u}(X_t)dL_t)\Big|$$

$$\geq (1/10) c_{13} \lambda^{2(k-1)} \varepsilon^2 \;\Big|\; \mathcal{F}_{T_{k-1}} \Big) \leq 1/8. \tag{3.14}$$

Combining (3.6) with (3.10)-(3.14), we see that conditioned on $\mathcal{F}_{T_{k-1}}$, with probability 3/4 or greater the following event holds:

$$(X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k} - X'_{T_k}) - (X_{T_{k-1}} - X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}} - X'_{T_{k-1}}) \leq c_{13} \lambda^{2(k-1)} \varepsilon^2.$$

In view of (3.7)-(3.9), this implies that if (3.8) holds, then

$$|X_{T_k} - X'_{T_k}| \leq \lambda^{-1/3} \, \lambda^{k-1} \varepsilon. \tag{3.15}$$

Let

$$T'_k = (T_{k-1} + t_1) \wedge \inf\{t > T_{k-1} : |X_t - X'_t| \geq \lambda^{k-1} \varepsilon\}.$$

By Doob's inequality, if $t_1 > 0$ is sufficiently small,

$$\mathbb{P}\left(\sup_{T_{k-1} \leq s \leq T'_k} |U_s - U'_s| \geq (1/2) \lambda^{k-1} \varepsilon \;\Big|\; \mathcal{F}_{T_{k-1}} \right) \leq 1/4. \tag{3.16}$$

Let

$$F_k = \{T_k = T_{k-1} + t_1\} \cup \{|L_{T_k} - L_{T_{k-1}}| \vee |L'_{T_k} - L'_{T_{k-1}}| > c_0\}.$$

By (3.15) and (3.16),

$$\mathbb{P}(F_k \mid \mathcal{F}_{T_{k-1}}) \geq 1/2. \tag{3.17}$$

Recall the definition of k_0 relative to $p_1 \in (0,1)$. Repeated application of the strong Markov property at the stopping times T_k and comparison with a Bernoulli sequence prove that at least $k_0/4$ of the events F_k will occur with probability p_1 or greater. This implies that with probability p_1 or greater,

$$\text{either} \quad T_{k_0} - T_1 \geq 1/4 \quad \text{or} \quad L_{T_{k_0}} - L_{T_1} \geq c_0/(4t_1) \quad \text{or} \quad L'_{T_{k_0}} - L'_{T_1} \geq c_0/(4t_1).$$

From the definitions of T_1^* and S,

$$\mathbb{P}\left(\sup_{t \in [0,S]} |X_t - X'_t| \leq \varepsilon \lambda^{k_0} \right) \geq p_1.$$

As $\varepsilon > 0$ is arbitrary, we have

$$\mathbb{P}\left(X_t = X'_t \text{ for every } t \in [0,S] \right) \geq p_1 > 0,$$

which contradicts to the definition of T_1^*. This completes the proof of pathwise uniqueness.

Strong existence follows from weak existence and pathwise uniqueness using a standard argument; see [26, Section IX.1]. □

4 Diffusions with smooth drifts

This section is devoted to analysis of a diffusion with inert drift given as the gradient of a smooth potential. We will use such diffusions to approximate diffusions with reflection. The analysis of the stationary measure is easier in the case when the drift is smooth. Let V be a C^2-smooth function on D that goes to $+\infty$ sufficiently fast as x approaches the boundary of D. We will consider the diffusion process on D associated with generator $\mathcal{L} = \frac{1}{2} e^V \nabla (e^{-V} A \nabla)$ but with an additional "inert" drift. More precisely, let Γ be a non-degenerate constant $d \times d$-matrix. We consider the SDE

$$\begin{cases} dX_t &= \sigma(X_t) dB_t + \mathbf{b}(X_t) dt - (A\nabla V)(X_t)\, dt + K_t\, dt\,, \\ dK_t &= -\Gamma\, \nabla V(X_t)\, dt\,, \end{cases} \tag{4.1}$$

where $A = \sigma^T \sigma$ is uniformly elliptic and bounded having C^2-smooth entries a_{ij} on $\overline{D}$, $\mathbf{b} = (b_1, \cdots, b_d)$ with $b_k(x) := \frac{1}{2} \sum_{i=1}^d \partial_i (a_{ik}(x))$.

To find the candidate for the stationary distribution for (X, K), we will do some computations with processes of the form $f(X_t, K_t)$, where $f \in C^2(\mathbb{R}^{2d})$. We will use the following notation,

$$\begin{aligned}
\nabla_x f(x,y) &= \left(\frac{\partial}{\partial x_1} f(x_1, \ldots, x_d, y_1, \ldots, y_d), \ldots, \frac{\partial}{\partial x_d} f(x_1, \ldots, x_d, y_1, \ldots, y_d) \right)^T, \\
\nabla_y f(x,y) &= \left(\frac{\partial}{\partial y_1} f(x_1, \ldots, x_d, y_1, \ldots, y_d), \ldots, \frac{\partial}{\partial y_n} f(x_1, \ldots, x_d, y_1, \ldots, y_d) \right)^T, \\
\mathcal{L}_x f(x,y) &= \frac{1}{2} e^{V(x)} \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(e^{-V(x)} a_{ij}(x) \frac{\partial}{\partial x_j} f(x,y) \right), \\
\mathcal{L}_x^* f(x,y) &= \frac{1}{2} \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(e^{-V(x)} a_{ij}(x) \frac{\partial}{\partial x_j} (e^{V(x)} f(x,y) \right),
\end{aligned}$$

For $f \in C^2(\mathbb{R}^{2d})$, by Ito's formula, we have

$$\begin{aligned}
df(X_t, K_t) &= \nabla_x f dX_t + \nabla_y f dK_t + \frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j f d\langle X^i, X^j \rangle_t \\
&= \text{local martingale} + (\mathcal{L}_x f + \nabla_x f \cdot K_t - \nabla_y f \cdot \Gamma \nabla_x V)\, dt.
\end{aligned}$$

So the process (X, K) has the generator

$$\mathcal{G} f(x,y) := \mathcal{L}_x f(x,y) + y \cdot \nabla_x f(x,y) - \Gamma\, \nabla_x V(x) \cdot \nabla_y f(x,y) \tag{4.2}$$

for $x \in D$ and $y \in \mathbb{R}^d$.

We will now assume that (X, K) has a stationary measure of a special form. Then we will do some calculations to find an explicit formula for the stationary measure, and finally we will show that the calculations can be traced back to complete the proof that the measure we started with is indeed the stationary distribution.

Suppose that (X,K) has a stationary distribution π of the form $\pi(dx,dy)=\rho_1(x)\rho_2(y)dx\,dy$. Then $\mathcal{G}^*\pi=0$, in the sense that for every $f\in \mathrm{Dom}(\mathcal{G})$,

$$\int_{\mathbb{R}^{2d}} \mathcal{G}f(x,y)\pi(dx,dy)=0.$$

Then we have for every $f\in C_c^2(D\times\mathbb{R}^d)$, by integration by parts formula,

$$\begin{aligned}
0 &= \int_{\mathbb{R}^d}\left(\int_D \left(\rho_1(x)y\cdot\nabla_x f(x,y)+\rho_1(x)\mathcal{L}_x f(x,y)\right)dx\right)\rho_2(y)dy \\
&\quad -\int_D\left(\int_{\mathbb{R}^d}\rho_2(y)\big(\Gamma\,\nabla_x V(x)\cdot\nabla_y f(x,y)dy\right)\rho_1(x)dx \\
&= \int_{\mathbb{R}^d}\left(\int_D \left(-\nabla_x\rho_1(x)\cdot yf(x,y)-\mathcal{L}_x^*\rho_1(x)f(x,y)\right)dx\right)\rho_2(y)dy \\
&\quad +\int_D\left(\int_{\mathbb{R}^d}\Gamma\,\nabla_x V(x)\cdot\nabla_y\rho_2(y)f(x,y)dy\right)\rho_1(x)dx
\end{aligned}$$

.

This implies that

$$\rho_2(y)\left(\nabla_x\rho_1(x)\cdot y+\mathcal{L}_x\rho_1(x)\right)+\rho_1(x)\Gamma\,\nabla_x V(x)\cdot\nabla_y\rho_2(y)=0 \qquad \text{for } (x,y)\in D\times\mathbb{R}^d. \tag{4.3}$$

We now make an extra assumption that $\rho_1(x)=ce^{-V(x)}$. Then we have

$$\rho_2(y)\nabla_x V(x)\cdot y+\Gamma\,\nabla_x V(x)\cdot\nabla_y\rho_2(y)=0 \qquad \text{for } (x,y)\in D\times\mathbb{R}^d.$$

Since $V(x)$ blows up as x approaches the boundary ∂D, $\{\nabla_x V(x), x\in D\}$ spans the whole $\mathbb{R}^d$. So we must have

$$y+\Gamma^T\nabla_y\log\rho_2(y)=0 \qquad \text{for every } y\in\mathbb{R}^d.$$

This implies that Γ is symmetric and that

$$\nabla_y\log\rho_2(y)=-\Gamma^{-1}y \qquad \text{for every } y\in\mathbb{R}^d.$$

Hence we have

$$\log\rho_2(y)=-\frac{1}{2}(\Gamma^{-1}y,y)+c_1,$$

or

$$\rho_2(y)=c_2\exp\Big(-\frac{1}{2}(\Gamma^{-1}y,y)\Big).$$

The above calculations suggests that when Γ is a symmetric positive definite matrix, the stationary distribution for the process (X,K) in (2.1) has the form

$$c_3\,\mathbf{1}_D(x)\,\exp\Big(-V(x)-\tfrac{1}{2}(\Gamma^{-1}y,y)\Big)\,dx\,dy,$$

where $c_3>0$ is the normalizing constant.

Theorem 4.1 *Suppose that Γ is a symmetric positive definite matrix, each a_{ij} is C^2-smooth on $\overline{D}$ and $V \geq 0$ is a C^2-smooth potential on D that goes to infinity when x approaches to the boundary ∂D at a sufficient rate so that the diffusion having infinitesimal generator $\frac{1}{2}e^V \sum_{i,j=1}^{\infty} \partial_i (e^{-v} a_{ij} \partial_j)$ is conservative. Then the process (X, K) of (4.1) has a unique stationary distribution*

$$\pi(dx, dy) := c_0 \mathbf{1}_D(x) \exp\Big(- V(x) - \tfrac{1}{2}(\Gamma^{-1}y, y)\Big)\, dx\, dy.$$

Proof. (**Existence: not rigorous**): Clearly we have

$$\int_{D\times\mathbb{R}^d} \mathcal{G} f(x,y)\pi(dx,dy) = 0 \qquad \text{for every } f \in C_c^2(D\times\mathbb{R}^d), \tag{4.4}$$

where $\mathcal{G}$ is given by (4.2). For $f \in C_c^2(D\times\mathbb{R}^d)$, let $u(t,x,y) := \mathbb{E}_{(x,y)}[f(X_t, K_t)]$. **If we can show that** $u(t,x,y) \in C_\infty^2(D, \mathbb{R}^d)$ for each fixed $t > 0$ and C^1 in time $t > 0$, then we have by Ito's formula that

$$\frac{\partial u(t,x,y)}{\partial t} = \mathcal{G} u(t,x,y).$$

It then follows from (4.4)

$$\begin{aligned} \frac{d}{dt}\int_{D\times\mathbb{R}^d} u(t,x,y)\pi(dx,dy) &= \int_{D\times\mathbb{R}^d} \frac{\partial}{\partial t} u(t,x,y)\pi(dx,dy) \\ &= \int_{D\times\mathbb{R}^d} \mathcal{G} u(t,x,y)\pi(dx,dy) \\ &= 0. \end{aligned}$$

The above is the argument from Ikeda-Watanabe book that Martin quoted. However in order to get the last identity from (4.4)**, we need to establish semigroup regularity at least** $u \in C_\infty^2(D\times\mathbb{R}^d)$ **(it is not true that** $u \in C_c^2(D\times\mathbb{R}^d)$**). Or we can go the other way around. Let** $u(t,x,y)$ **be the PDE solution for** $\frac{\partial u(t,x,y)}{\partial t} = \mathcal{G}u(t,x,y)$ **with initial value** $u(0,x,y) = f(x,y) \in C_c^2(D\times\mathbb{R}^d)$**. If we can show that** $u \in C^{1,2}$ **and that for each fixed** $t > 0$**,** $u(t,x,y) \in C_\infty^2(D\times\mathbb{R}^d$**, then by Ito's formula we have** $u(t,x,y) = \mathbb{E}_{(x,y)}[f(X_t, K_t)]$ **and then apply the above reasoning. Can Martin supply the regularity for** u**?** Thus we have

$$\int_{D\times\mathbb{R}^d} u(t,x,y)\pi(dx,dy) = \int_{D\times\mathbb{R}^d} f(x,y)\pi(dx,dy) \qquad \text{for every } t > 0.$$

This proves that π is a stationary distribution for the process (X, K).

From ZC: Can we modify the following argument to establish the uniqueness of the above theorem in this generality? The following (4.6) is a sufficient condition to guarantee a global solution. (i) Can we give some concrete examples for the existence of such V for bounded C^2-domain D? (ii) The V_n in Section 5 may not satisfy condition (4.6) but we still have a conservative process. This is one of the reasons that I formulated above Theorem 4.1 in the current form.

This section is devoted to analysis of a diffusion with inert drift given as the gradient of a smooth potential. We will use such diffusions to approximate diffusions with reflection. The analysis of the stationary measure is easier in the case when the drift is smooth. We will limit ourselves to the case when σ is the identity matrix, i.e., the diffusion part of the process is Brownian motion.

We consider the SDE

$$\begin{aligned} dX_t &= -\nabla V(X_t)\,dt + K_t\,dt + dB_t\;, \\ dK_t &= -\nabla V(X_t)\,dt\;, \end{aligned} \tag{4.5}$$

where $V\colon D \to \mathbb{R}^+$ is some 'potential'. For $D \neq \mathbb{R}^d$, we think of V as going to infinity at the boundary of D. In order to analyze the behavior of (4.5), it is natural to introduce the 'energy'

$$H(x,y) = \frac{y^2}{2} + V(x), \qquad x, y \in \mathbb{R}^d\;.$$

Applying Itô's formula to $H_t := H(X_t, K_t)$, we obtain

$$dH_t = -|\nabla V(X_t)|^2\,dt + \frac{1}{2}\Delta V(X_t)\,dt + \nabla V(X_t)\,dB_t\;.$$

In order to ensure non-explosion of H_t, it is therefore natural to introduce the assumption

Assumption 4.2 *There exists a constant $C \in \mathbb{R}$ such that the relation*

$$\Delta V(x) \leq 2|\nabla V(x)|^2 + C\bigl(1 + V(x)\bigr) \tag{4.6}$$

holds for every $x \in D$. Furthermore, one has $\lim_{x\to\partial D} V(x) = +\infty$ and $\int_D \exp(-2V(x))\,dx < \infty$. If D is unbounded, one should interpret ∂D as the boundary of D in the Alexandrov compactification of $\mathbb{R}^d$.

Proposition 4.3 *Assume that $V \in \mathcal{C}^\infty(D)$ satisfies Assumption 4.2. Then there exists a unique global strong solution to (4.5) and $\mathbb{P}\bigl(\sup_{t\in[0,T]} V(X_t) = +\infty\bigr) = 0$ for every initial condition $X_0 \in D$ and every $T > 0$. Furthermore, the measure $\mu(dx, dy) := \exp\bigl(-2H(x,y)\bigr)\,dx\,dy$ is an invariant measure for this solution.*

Remark 4.4 *The assumption (4.6) is close to being optimal for ensuring that the solution to (4.5) never hits the boundary of the domain D. Consider the case $n = 1$, $D = \mathbb{R}_+$, $V\colon \mathbb{R}_+ \to \mathbb{R}_+$, and remove inertia so that $dX = -\nabla V(X)\,dt + dB$. Define a function $\Psi\colon \mathbb{R}_+ \to \mathbb{R}$ by*

$$\Psi(x) = \int_1^x \exp\bigl(2V(y)\bigr)\,dy\;.$$

Then one can check that $\Psi(X_t)$ is a local martingale with quadratic variation greater or equal to 1. In particular, it has a positive probability of taking arbitrarily large values during any finite time interval. If V now diverges at 0 sufficiently slowly so that $\exp(2V)$ is still integrable, this immediately implies that this diffusion has a positive probability of reaching 0 in any finite time.

Let us now consider the case where $\exp\big(2V(x)\big) \approx 1/x^\alpha$ *near* $x = 0$.

From ZC: the following does not follow from the above unless you have a specific V *in mind*

In this case, one has $\Delta V = \alpha/(2x^2)$ *and* $2|\nabla V|^2 = \alpha^2/(2x^2)$*. Therefore, (4.6) is satisfied if and only if* $\alpha \ge 1$*, which is precisely the range of parameters for which* $\exp(2V)$ *is no longer integrable at* 0.

Proof of Proposition 4.3. Since V is smooth, (4.5) has a unique solution as long as it stays in the interior of D. Fix now an initial condition $(x_0, K_0) \in D \times \mathbb{R}$, a (non-random) time $T > 0$ and consider the stopping time $\tau_N = T \wedge \inf\{t \ge 0 \,:\, V(X_t) \ge N\}$. Setting

$$J(t) = \mathbb{E}\, H(X_{t\wedge\tau_N}, K_{t\wedge\tau_N}),$$

we have

$$\begin{aligned} J(t) - J(0) &= \mathbb{E} \int_0^{t\wedge\tau_N} \Big(-|\nabla V(X_s)|^2 + \frac{1}{2}\Delta V(X_s)\Big)\, ds \\ &\le C\mathbb{E} \int_0^t \big(1 + V(X_{s\wedge\tau_N})\big)\, ds \le C \int_0^t \big(1 + J(s)\big)\, ds \end{aligned}$$

It follows from Gronwall's inequality that $J(t) \le C_1 \exp(C_2 t)$ for some constants C_1 and C_2. Since $\mathbb{P}\big(\sup_{t\in[0,T]} V(X_t) \ge N\big) \le J(t)/N$, it follows that (4.5) has a unique global solution and that $\mathbb{P}\big(\sup_{t\in[0,T]} V(X_t) = +\infty\big) = 0$, for every $T > 0$.

Direct calculations show that the generator $\mathcal{L}$ of the process (X, K) is given by

$$\mathcal{L}f(x, y) = \frac{1}{2}\Delta_x f(x, y) - \nabla V(x)\nabla_x f(x, y) - \nabla V(x)\nabla_y f(x, y) + y \cdot \nabla_x f(x, y),$$

and $\mathcal{L}^* \exp\big(-2H(x, y)\big) = 0$, where $\mathcal{L}^*$ is the formal adjoint of $\mathcal{L}$.

Assume now for the moment that V is defined on all of $\mathbb{R}^d$, grows linearly at infinity and is such that $\nabla V \in \mathcal{C}_b^\infty(\mathbb{R}^d)$. In this case, for any initial distribution on $\mathbb{R}^{2d}$, we can interpret the law of the corresponding solution to (4.5) as a Schwartz distribution on $\mathbb{R}^{2d+1}$ via the formula

$$p(\phi) = \int_0^\infty \mathbb{E}\left(\phi(t, X_t, K_t)\right) dt\ .$$

Choosing ϕ such that $\phi(t, x, K) = 0$ for $t \le 0$ and applying Itô's formula to $\phi(t, X_t, K_t)$, we see that p satisfies the forward Kolmogorov equation $\big(\partial_t - \mathcal{L}^*\big)p = 0$. Since $\mathcal{L}^* \exp\big(-2H(x, y)\big) = 0$, this shows that

$$\mu(dx\, dy) = \exp\big(-2H(x, y)\big)\, dx\, dy \tag{4.7}$$

is an invariant measure for (4.5). The fact that this is still true for the class of potentials considered in Assumption 4.2 follows from the following approximation argument. Let $V_N \colon \mathbb{R}^d \to \mathbb{R}_+$ be a sequence of smooth function such that

(i) One has $V_N(x) = V(x)$ for every $x \in D$ such that $V(x) \leq N$.

(ii) For every $N > 0$, ∇V_N belongs to $\mathcal{C}_b^\infty(\mathbb{R}^d)$.

(iii) The sequence of measures $\mu_N(dx\, dK) = \exp(-2H_N(x, K))\, dx\, dK$ converges to μ in total variation.

By (ii) and the above consideration, this yields a sequence $(X^{(N)}, K^{(N)})$ of stationary processes with fixed-time distribution μ_N. By (i) and (iii), this sequence converges (in the total variation topology over any fixed time interval) to a stationary process with fixed-time distribution μ that solves (4.5). □

It is possible to show under very weak additional assumptions that μ is the only invariant measure for (4.5). This is not obvious *a priori* since we would like to consider the case where $V(x) = f(d(x, \partial D))$ for some function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ which has a singularity at 0 and is such that $f(x) = 0$ for x larger than some (small) constant ε. On the set $\{x : d(x, \partial D) > \varepsilon\} \times \mathbb{R}^d$, the diffusion (4.5) has then a deterministic component $dK = 0$. In particular, this shows that (4.5) is not hypoelliptic and that its transition probabilities are not absolutely continuous with respect to Lebesgue measure.

From ZC: Are we assuming D is C^∞ here? If D is C^2, then the distance function $d(x, \partial D)$ may only be C^2 so V can not be C^∞.

For any multi-index $\alpha = \{\alpha_1, \dots, \alpha_k\}$, we define the vector field $\nabla_\alpha V$ by

$$\big(\nabla_\alpha V(x)\big)_k = \frac{\partial^{k+1} V(x)}{\partial x_k \partial x_{\alpha_1} \dots \partial x_{\alpha_k}} \;.$$

Denoting by $|\alpha|$ the size of the multi-index, we furthermore assume that

Assumption 4.5 *There exists $x_* \in D$ such that the collection $\{\nabla_\alpha V(x_*)\}_{|\alpha|>0}$ spans all of $\mathbb{R}^d$.*

Remark 4.6 *In the case where V is of the form $V(x) = f(d(x, \partial D))$ for some smooth function f diverging at 0, Assumption 4.5 is satisfied as soon as there exists a point on the boundary such that its curvature has full rank.*

We then have

Proposition 4.7 *If both Assumptions 4.2 and 4.5 hold, then μ is the unique invariant measure for (4.5).*

Proof. Let us first introduce the following concept. Given a Markov operator $\mathcal{P}$ over a Polish space $\mathcal{X}$, we say that $\mathcal{P}$ is *strong Feller at x* if $\mathcal{P}\phi$ is continuous at x for every bounded measurable function $\phi\colon \mathcal{X} \to \mathbb{R}$. The proof of Proposition 4.7 is then based on the following fact, which can be found for example in [12]: If $\{\mathcal{P}_t\}_{t\geq 0}$ is a Markov semigroup over $\mathcal{X}$ such that

(i) There exists $t > 0$ and $x \in \mathcal{X}$ such that $\mathcal{P}_t$ is strong Feller at x,

(ii) There exists $s > 0$ such that $x \in \mathit{supp}\, \mathcal{P}_s(y,\cdot\,)$ for every $y \in \mathcal{X}$,

then the semigroup $\{\mathcal{P}_t\}$ can have at most one invariant probability measure.

Denote now by $\mathcal{P}_t$ the Markov semigroup generated by solutions to (4.5). It is easy to check that Assumption 4.5 means precisely that the operator $\partial_t - \mathcal{L}$, where $\mathcal{L}$ is the generator of (4.5), satisfies Hörmander's condition [16, 23] in some neighborhood of (x_*, K, t) for every $K \in \mathbb{R}^d$ and every $t > 0$. Since for any $\phi \in \mathcal{B}_b(\mathbb{R}^d)$, the map $\Phi(t, y, K) = \big(\mathcal{P}_t\phi\big)(y, K)$ is a solution (in the sense of distributions) of the equation $\big(\partial_t - \mathcal{L}\big)\Phi = 0$, this implies that $\mathcal{P}_t\phi$ is $\mathcal{C}^\infty$ in a neighborhood of $(x_*, 0)$ for every $t > 0$. In particular, $\mathcal{P}_t$ is strong Feller at $(x_*, 0)$ for every $t > 0$.

It now remains to show that the point $(x_*, 0)$ belongs to the support of every $\mathcal{P}_t(x,\cdot)$. For this, it suffices for example to check that for every $(x, K) \in \mathbb{R}^{2d}$, there exists $T > 0$ such that $\mathcal{P}_T(x, K; A) > 0$ for every neighborhood A of the point $(x_*, 0)$. In order to show this, we are going to apply the Stroock-Varadhan support theorem [28], so we consider the control system

$$\dot{x} = -\nabla V(x) + K + u(t)\;, \qquad \dot{K} = -\nabla V(x)\;, \tag{4.8}$$

where $u\colon [0, T] \to \mathbb{R}^d$ is a smooth control. The claim is proved if we can show that for every (x_0, K_0), there exists $T > 0$ such that, for every $\varepsilon > 0$ there exists a control such that the solution to (4.8) at time T is located in an ε-neighborhood of the point $(x_*, 0)$. Our proof is based on the fact that, since we assumed that $V(x)$ grows to $+\infty$ as x approaches ∂D, there exists a collection of ℓ points $x_1, \ldots, x_\ell$ ($\ell > n$) such that the positive cone generated by $\nabla V(x_1), \ldots, \nabla V(x_\ell)$ is all of $\mathbb{R}^d$. (Assume otherwise, so that there exists $v \in \mathbb{R}^d$ such that $\langle v, \nabla V(x)\rangle \leq 0$ for every $x \in \mathbb{R}^d$. One then immediately gets a contradiction to the fact that there is $t_0 > 0$ (possibly $t_0 = +\infty$) such that $\lim_{t\to t_0} V(x + vt) = \infty$ for every $x \in D$.)

Fix now an initial condition (x_0, K_0). From the previous argument, there exist positive constants $\alpha_1, \ldots, \alpha_\ell$ such that $\sum_{i=1}^\ell \alpha_i \nabla V(x_i) = -K_0$. Fix now $T = \sum_{i=1}^\ell \alpha_i$ and consider a family $X_\varepsilon\colon [0, T] \to \mathbb{R}^d$ of smooth trajectories such that:

(i) there are time intervals I_i of lengths greater than or equal to $\alpha_i - \varepsilon$ such that $X_\varepsilon(t) = x_i$ for $t \in I_i$,

(ii) $X_\varepsilon(0) = x_0$ and $X_\varepsilon(T) = x_*$,

(iii) one has $\int_{[0,T]\setminus\bigcup I_i} |\nabla V(X_\varepsilon(t))|\, dt \leq \varepsilon$.

Such a family of trajectories can be easily constructed (take a single trajectory going through all the x_i and run through it at the appropriate speed). It now suffices to take as control

$$u(t) = \dot{X}_\varepsilon(t) + \nabla V(X_\varepsilon(t)) + \int_0^t \nabla V(X_\varepsilon(s)) \ ,$$

and to note that the solution to (4.8) is given by $\big(X_\varepsilon(t), K_0 - \int_0^t \nabla V(X_\varepsilon(s))\big)$. This solution has the desired properties by construction. □

5 Weak convergence

In this section, we will prove that reflecting Brownian motion with inert drift can be approximated by diffusions with smooth inert drifts. As in the previous section, we assume that the diffusion part of each of our processes is a Brownian motion.

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$ and use $\delta_D(x)$ to denote the Euclidean distance between x and D^c. A continuous function $x \mapsto \delta(x)$ is called a regularized distance function to D^c if there are constants $c_2 > c_1 > 0$ such that

(i) $c_1 \delta_D(x) \leq \delta(x) \leq c_2 \delta_D(x)$ for every $x \in \mathbb{R}^d$;

(ii) $\delta \in C^\infty(D)$ and for any multi-index $\alpha = (\alpha_1, \cdots, \alpha_d)$ with $|\alpha| := \sum_{k=1}^d \alpha_k$, there is a constant c_α such that

$$\left| \frac{\partial^{|\alpha|} \delta(x)}{(\partial x_1)^{\alpha_1} \cdots (\partial x_d)^{\alpha_d}} \right| \leq c_\alpha \delta_D(x)^{1-|\alpha|} \qquad \text{for every } x \in D.$$

The existence of such a regularized distance function δ is given by [31, Lemma 2.1]. For $n \geq 1$, define for $x \in D$,

$$V_n(x) := \exp\left(n\delta(x)^{-1}\right). \tag{5.1}$$

Let $X^{(n)}$ be the diffusion given by

$$dX_t^{(n)} = -\frac{1}{2}\nabla V_n(X_t^{(n)})dt + dB_t,$$

with initial probability distribution $m_n(dx) := c_n \exp(-V_n(x))1_D(x)dx$, where B is a d-dimensional Brownian motion and $c_n > 0$ is a normalizing constant so that $m_n(\mathbb{R}^d) = 1$. It is known (see [31]) that $X^{(n)}$ is conservative and never reaches ∂D. In addition, $X^{(n)}$ is a symmetric diffusion with respect to the measure m_n and its Dirichlet form $(\mathcal{E}^{(n)}, \mathcal{F}^{(n)})$ in $L^2(\mathbb{R}^d, m_n(dx))$ is given by (see [31, Lemma 3.5])

$$\begin{aligned} \mathcal{E}^{(n)}(f,f) &:= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla f(x)|^2 m_n(dx); \\ \mathcal{F}^{(n)} &:= \left\{ f \in L^2(\mathbb{R}^d, m_n) : \mathcal{E}^{(n)}(f,f) < \infty \right\}. \end{aligned}$$

Without loss of generality, we assume that $X^{(n)}$ is defined on the canonical path space $\Omega := C([0,\infty),\mathbb{R}^d)$ with $X_t^{(n)}(\omega) = \omega(t)$. On Ω, for every $t > 0$, there is a time-reversal operator r_t, defined for $\omega \in \Omega$ by

$$r_t(\omega)(s) := \begin{cases} \omega(t-s), & \text{if } 0 \le s \le t, \\ \omega(0), & \text{if } s \ge t. \end{cases} \tag{5.2}$$

Let $\mathbb{P}_n$ denote the law of $X^{(n)}$ with initial distribution m_n and let $\{\mathcal{F}_t^{(n)}, t \ge 0\}$ denote the minimal augmented filtration generated by $X^{(n)}$. Then it follows from the reversibility of $X^{(n)}$ that for every $t > 0$, the measure $\mathbb{P}_n$ on $\mathcal{F}_t^{(n)}$ is invariant under the time-reversal operator r_t (cf. [15, Lemma 5.7.1]). We know that

$$X_t^{(n)} - X_0^{(n)} = B_t^{(n)} - \int_0^t \frac{1}{2}\nabla V_n(X_s^{(n)})ds, \qquad t \ge 0, \tag{5.3}$$

where $B^{(n)}$ is a d-dimensional Brownian motion. On the other hand, by Lyons-Zheng's forward and backward martingale decomposition (see [15, Theorem 5.7.1]), we have for every $T > 0$,

$$X_t^{(n)} - X_0^{(n)} = \frac{1}{2}B_t^{(n)} - \frac{1}{2}(B_T^{(n)} - B_{T-t}^{(n)}) \circ r_T \qquad \text{for } t \in [0,T]. \tag{5.4}$$

Clearly, for every $T > 0$, the law of $\{B_t^{(n)}, t \ge 0\}$ under $\mathbb{P}_n$ is tight in the space $C([0,T],\mathbb{R}^d)$ and so is $\frac{1}{2}(B_T^{(n)} - B_{T-t}^{(n)}) \circ r_T$. It follows that the laws of $(X^{(n)}, B^{(n)}, (B_T^{(n)} - B_{T-\cdot}^{(n)}) \circ r_T)$ under $\mathbb{P}_n$ are tight in the space $C([0,T],\mathbb{R}^{3d})$ for every $T > 0$. On the other hand, by (5.3)-(5.4),

$$K^{(n)}(t) := -\frac{1}{2}\int_0^t \nabla V_n(X_s^{(n)})ds = -\frac{1}{2}B_t^{(n)} - \frac{1}{2}(B_T^{(n)} - B_{T-t}^{(n)}) \circ r_T.$$

Thus the laws of $(X^{(n)}, B^{(n)}, K^{(n)})$ under $\mathbb{P}_n$ are also tight in the space $C([0,T],\mathbb{R}^{3d})$ for every $T > 0$. By Theorem 3.1 and Theorem 3.2 of [31], $(X^{(n)}, \mathbb{P}_n)$ converges weakly in $C([0,T],\mathbb{R}^d)$ to stationary reflecting Brownian motion X in D with uniform initial distribution on D. Passing to a subsequence, if necessary, we conclude that $((X^{(n)}, B^{(n)}, K^{(n)}), \mathbb{P}_n)$ converges weakly in $C([0,T],\mathbb{R}^{3d})$ to a process (X, B, K).

Now we apply the Skorokhod lemma (see Theorem 3.1.8 in [14]) to construct the processes $(X^{(n)}, B^{(n)}, K^{(n)})$ and (X, B, K) on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in such a way that $(X^{(n)}, B^{(n)}, K^{(n)})$ converge to (X, B, K) a.s., on the time interval $[0,1]$ in the supremum norm. Clearly, $X = X_0 + B + K$. By the proof of [31, Lemma 4.1], K is a continuous process of locally finite variation. On the other hand, since X is reflecting Brownian motion in a Lipschitz domain, it admits a Skorokhod decomposition. So by the uniqueness of Doob-Meyer decomposition, we conclude that

$$X_t = X_0 + B_t + \int_0^t \mathbf{n}(X_s)dL_s, \qquad t \ge 0,$$

where $\mathbf{n}$ is the unit inward normal vector field on ∂D, which is well defined a.e. with respect to the surface measure σ, and L is a positive continuous additive functional of X with Revuz measure σ.

In other words, we have

$$K_t = \int_0^t \mathbf{n}(X_s)dL_s, \qquad \text{for } t \geq 0.$$

Set

$$M_t = \exp\left(\int_0^t K_s dB_s - \frac{1}{2}\int_0^t |K_s|^2 ds\right), \quad t \geq 0$$

and

$$M_t^{(n)} = \exp\left(\int_0^t K_s^{(n)} dB_s^{(n)} - \frac{1}{2}\int_0^t |K_s^{(n)}|^2 ds\right), \qquad t \geq 0.$$

Let $d\mathbb{Q} = M_1 d\mathbb{P}$ and $d\mathbb{Q}_n = M_1^{(n)} d\mathbb{P}$. By the Girsanov theorem, under $\mathbb{Q}_n$, $(X^{(n)}, K^{(n)})$ satisfies the following equation

$$\begin{aligned} dX_t^{(n)} &= dB_t^{(n)} - \frac{1}{2}V_n(X_t^{(n)})dt + K_t^{(n)}dt, \\ dK_t &= -\frac{1}{2}V_n(X_t^{(n)})dt, \end{aligned}$$

where $W^{(n)}$ is a d-dimensional Brownian motion. On the other hand, by the proof of Theorem 2.1, (X, K) under $\mathbb{Q}$ is Brownian motion with inert drift, and satisfies

$$\begin{aligned} dX_t &= dB_t + \mathbf{n}(X_t)dL_t + K_t dt, \\ dK_t &= \mathbf{n}(X_t)dL_t, \end{aligned}$$

where W is a d-dimensional Brownian motion.

In the following, the initial distribution of $(X_0^{(n)}, K_0^{(n)})$ and (X_0, K_0) under $\mathbb{P}$ are taken to be $c_n 1_D(x)e^{-2V_n(x)-y^2}dxdy$ and $c1_D(x)e^{-|y|^2}$, respectively, where $c_n > 0$ and $c > 0$ are normalizing constant to make the corresponding measures to be a probability measures.

Theorem 5.1 *For every $T > 0$, the law of $(X^{(n)}, K^{(n)})$ under $\mathbb{Q}_n$ converges weakly in the space $C([0,T], \mathbb{R}^{2d})$ to that of the law of (X, K) under $\mathbb{Q}$.*

Proof. Without loss of generality, we take $T = 1$.

We have $\int_0^1 |K_s^{(n)}|^2 ds \to \int_0^1 |K_s|^2 ds$ as $n \to \infty$, $\mathbb{P}$-a.s., because $K^{(n)} \to K$ uniformly on $[0,1]$. Let $|K^{(n)}|_t$ denote the total variation process of $K_s^{(n)}$ over the interval $[0,t]$. Then we know from [31, Lemma 2.2 and p. 483] that

$$\sup_{n\geq 1} \mathbb{E}_{\mathbb{P}_n}\left[|K^{(n)}|_t\right] < \infty \qquad \text{for every } t \geq 0.$$

Hence by Theorem 2.2 of [19], $\int_0^1 K_s^{(n)} dB_s^{(n)} \to \int_0^1 K_s dB_s$ as $n \to \infty$, in probability. By passing to a subsequence, we may assume that the convergence is $\mathbb{P}$-almost sure. We conclude that $\lim_{n\to\infty} M_1^{(n)} = M_1$ $\mathbb{P}$-a.s.

Let Φ be a continuous function on $(C[0,1])^2$ with $0 \le \Phi \le 1$. Since $\Phi(X^{(n)}, K^{(n)}) \to \Phi(X,K)$, $\mathbb{P}$-a.s., and $M_1^{(n)} \to M_1$, $\mathbb{P}$-a.s., by Fatou's lemma,

$$\mathbb{E}_{\mathbb{P}}[\Phi(X,K)M_1] \le \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)}, K^{(n)})M_1^{(n)}] \le \limsup_{n\to\infty} \mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)}, K^{(n)})M_1^{(n)}] \tag{5.5}$$

and

$$\mathbb{E}_{\mathbb{P}}[(1-\Phi)(X,K)M_1] \le \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[(1-\Phi)(X^{(n)}, K^{(n)})M_1^{(n)}]. \tag{5.6}$$

Summing (5.5) and (5.6) we obtain $\mathbb{E}_{\mathbb{P}}[M_1] \le \mathbb{E}_{\mathbb{P}}[M_1^{(n)}]$. Note that $M^{(n)}$ is a continuous non-negative local martingale and hence a supermartingale, while by the proof of Theorem 2.1, M is a continuous martingale. Hence, $\mathbb{E}_{\mathbb{P}}[M_1] = 1 \ge \mathbb{E}_{\mathbb{P}}[M_1^{(n)}]$ and, therefore, the inequalities in (5.5) and (5.6) are in fact equalities. It follows that

$$\lim_{n\to\infty} \mathbb{Q}_n[\Phi(X^{(n)}, K^{(n)})] = \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)}, K^{(n)})M_1^{(n)}] = \mathbb{E}_{\mathbb{P}}[\Phi(X,K)M_1] = \mathbb{Q}[\Phi(X,K)].$$

This proves the weak convergence of $(X^{(n)}, K^{(n)})$ under $\mathbb{Q}_n$ to (X,K) under $\mathbb{Q}$. □

The following proposition, which gives the uniform integrability of $\{M_t^{(n)}, \mathbb{P}_n, n \ge 1\}$ for small times $t > 0$, can be used to give another proof of weak convergence, provided one starts with the stationary distributions, but the estimate is also of independent interest. Recall that $(X^{(n)}, K^{(n)})$ is Brownian motion with drift given by (5.1).

Proposition 5.2 *Let μ_n be the stationary distribution for $(X^{(n)}, K^{(n)})$ with inert drift given by (4.7) and suppose $(X_0^{(n)}, K_0^{(n)})$ has distribution μ_n. Then there exist t_0 and c_1 independent of n such that*

$$\mathbb{E}_{\mathbb{Q}_n}\left[e^{\int_0^{t_0} |K_s^{(n)}|^2\, ds}\right] \le c_1 \tag{5.7}$$

and

$$\mathbb{E}_{\mathbb{P}}\left[e^{\int_0^{t_0} |K_s^{(n)}|^2\, ds}\right] \le c_1. \tag{5.8}$$

Proof. By (4.7), under $\mathbb{Q}_n$, K_n is Gaussian, and there exist c_2 and c_3 not depending on n such that if $r \le c_2$, then

$$\mathbb{E}_{\mathbb{Q}_n} e^{r|K_n(s)|^2} \le c_3.$$

Then by Jensen's inequality applied with the measure $\frac{1}{t_0}1_{[0,t_0]}(s)\,ds$ and the function e^x we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}_n} e^{\int_0^{t_0} |K_n(s)|^2\, ds} &= \mathbb{E}_{\mathbb{Q}_n} e^{\frac{1}{t_0}\int_0^{t_0} (t_0|K_n(s)|^2)\, ds} \\
&\le \mathbb{E}_{\mathbb{Q}_n} \tfrac{1}{t_0} \int_0^{t_0} e^{t_0|K_n(s)|^2}\, ds \\
&\le \tfrac{1}{t_0} \int_0^{t_0} c_3\, ds = c_3,
\end{aligned}$$

provided we take $t_0 \le c_2$.

To prove the second inequality, let $N_t^{(n)} = \int_0^t K_s^{(n)}\, dB_s^{(n)}$. This is a martingale with respect to $\mathbb{P}$ and its quadratic variation $\langle N^{(n)} \rangle_t$ is equal to $\int_0^t |K_s^{(n)}|^2\, ds$ under both $\mathbb{P}$ and $\mathbb{Q}_n$. Note that $\exp(-N_t^{(n)} - \frac{1}{2}\langle N^{(n)} \rangle_t)$ is a martingale with respect to $\mathbb{P}$, hence

$$\mathbb{E}_{\mathbb{Q}_n}\left[e^{-2N_t^{(n)}}\right] = \mathbb{E}_{\mathbb{P}}\left[e^{-N_t^{(n)} - \frac{1}{2}\langle N^{(n)} \rangle_t}\right] = 1.$$

Using Cauchy-Schwarz,

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[e^{\int_0^t |K_s^{(n)}|^2\, ds}\right] &= \mathbb{E}_{\mathbb{P}}\left[e^{\langle N^{(n)} \rangle_t}\right] \\
&= \mathbb{E}_{\mathbb{Q}_n}\left[e^{\langle N^{(n)} \rangle_t} e^{-N_t^{(n)} - \frac{1}{2}\langle N \rangle_t}\right] \\
&\le \left(\mathbb{E}_{\mathbb{Q}_n}\left[e^{-2N_t^{(n)}}\right]\right)^{1/2} \left(\mathbb{E}_{\mathbb{Q}_n}\left[e^{\langle N^{(n)} \rangle_t}\right]\right)^{1/2} \\
&= \left(\mathbb{E}_{\mathbb{Q}_n}\left[e^{\langle N^{(n)} \rangle_t}\right]\right)^{1/2}.
\end{aligned}$$

The last term is bounded using (5.7) if t is small enough. □

Theorem 5.3 *Consider the SDE* (2.1) *with σ being the identity matrix, $\mathbf{b} \equiv 0$ and $\mathbf{u} \equiv \mathbf{v} \equiv \mathbf{n}$. A stationary distribution for the solution to* (2.1) *is*

$$\mu(A) = \mathbb{Q}((X_t, K_t) \in A) = \int_A c_1 1_D(x) e^{-|y|^2}\, dx\, dy, \qquad A \subset \overline{D} \times \mathbb{R}^d,$$

where c_1 is the normalizing constant given by $c_1^{-1} = |D|(2\pi)^{d/2}$ and $|D|$ is the Lebesgue measure of D.

Proof. Recall the notation from previous proofs in this section. Let us start $(X^{(n)}, K^{(n)})$ with the stationary distribution μ_n of Section 4, namely,

$$\mathbb{Q}_n(X_0^{(n)}, K_0^{(n)}) \in A) = \mu_n(A) = c_n \int_A e^{-V^{(n)}(x)} e^{-|y|^2}\, dx\, dy, \qquad A \subset \overline{D} \times \mathbb{R}^d,$$

where c_n is a normalizing constant. Clearly μ_n converge weakly on $\overline{D} \times \mathbb{R}^d$ to the probability μ. Also, the $\mathbb{Q}_n$ laws of $(X^{(n)}, B^{(n)}, K^{(n)})$ converge weakly to the $\mathbb{Q}$ law of (X, B, K). If f is any continuous and bounded function on $\mathbb{R}^{2d}$, then for $t_0 \le 1$,

$$\mathbb{E}_{\mathbb{Q}} f(X_{t_0}, K_{t_0}) = \lim_{n \to \infty} \mathbb{E}_{\mathbb{Q}_n} f(X_{t_0}^{(n)}, K_{t_0}^{(n)}) = \lim_{n \to \infty} \int_{\overline{D} \times \mathbb{R}^d} f(x, y) \mu_n(dx, dy) = \int_{\overline{D} \times \mathbb{R}^d} f(x, y) \mu(dx, dy).$$

This shows that the μ is a stationary distribution for the solution to (2.1). □

6 Irreducibility

In this section we again assume that we are in the case of normally reflecting Brownian motion. Let $\mathbb{Q}_{x,y}$ denote the distribution of (X, K) with inert drift starting from (x, y). We say that (X, K) is irreducible in the sense of Harris if there exists a positive measure ψ on $\overline{D} \times \mathbb{R}^d$ and $t_0 > 0$ such that if $\psi(A) > 0$, then for all $(x, y) \in \overline{D} \times \mathbb{R}^d$, $\mathbb{Q}_{x,y}((X_{t_0}, K_{t_0}) \in A) > 0$.

Theorem 6.1 *Assume that σ is the identity matrix and $\mathbf{u} \equiv \mathbf{v} \equiv \mathbf{n}$. Then the solution (X, K) to* (2.1) *is irreducible in the sense of Harris.*

Proof. Let X denote a solution to (2.2), i.e., the usual normally reflecting Brownian motion in D, let L be the local time of X on ∂D, and $K_t = y_0 + \int_0^s \mathbf{n}(X_s) dL_s$. The distribution of (X, K) starting from $(X_0, K_0) = (x, y)$ will be denoted $\mathbb{P}_{x,y}$. First, we will prove Harris irreducibility for (X, K) under $\mathbb{P}_{x,y}$.

Step 1. Fix any $t_0, r > 0$ and $z_1 \in D$. In this step, we will show that for any $(x_0, y_0) \in \overline{D} \times \mathbb{R}^d$ there exists $p_1 > 0$ such that $\mathbb{P}_{x_0,y_0}((X_{t_0/2}, K_{t_0/2}) \in \mathcal{B}(z_1, r) \times \mathcal{B}(0, r)) \geq p_1$.

We recall the deterministic Skorokhod problem in D with normal vector of reflection. Suppose a continuous function $f : [0, T] \to \mathbb{R}^d$ is such that $f(0) \in \overline{D}$. Then the Skorokhod problem is to find a continuous function $g : [0, T] \to \overline{D}$ and a non-decreasing function $\ell : [0, T] \to [0, \infty)$, such that $\ell(0) = 0$, $g(0) = f(0)$, $\int_0^T 1_D(g(s)) d\ell_s = 0$, and $g(t) = f(t) + \int_0^t \mathbf{n}(g(s)) d\ell_s$. It has been proved in [21] that the Skorokhod problem has a unique solution (g, ℓ) in every C^2 domain.

Since D is a bounded smooth domain, the set $\{\mathbf{n}(x)/|\mathbf{n}(x)|, x \in \partial D\}$ is the whole unit sphere in $\mathbb{R}^d$. Find $x_1 \in \partial D$ such that $\mathbf{n}(x_1) = -c_0 y_0$ for some $c_0 > 0$. It is elementary to construct a continuous function $f : [0, t_0/2] \to \mathbb{R}^d$ such that $f(0) = x_0$, $f(t) \in D$ for $t \in (0, t_0/4)$, $f(t_0/4) = x_1$, f is linear on $[t_0/4, 3t_0/8]$, $f(3t_0/8) = x_1 + y_0$, $f(t) - y_0 \in D$ for $t \in (3t_0/8, t_0/2)$, and $f(t_0/2) - y_0 = z_1$. It is straightforward to check that the pair (g, ℓ) that solves the Skorokhod problem for f has the following properties: $g(t) = f(t)$ for $t \in (0, t_0/4)$, $g(t) = x_1$ for $t \in [t_0/4, 3t_0/8]$, $g(t) = f(t) - y_0$ for $t \in (3t_0/8, t_0/2)$, and $\int_0^{t_0/2} \mathbf{n}(g(s)) d\ell_s = -y_0$.

For a function $f^1 : [0, t_0/2] \to \mathbb{R}^d$ with $f^1(0) \in \overline{D}$, let (g^1, ℓ^1) denote the solution of the Skorokhod problem for f^1. Let $\mathcal{B}_C(f, \delta)$ be the ball in $C([0, t_0/2], \mathbb{R}^d)$ centered at f, with radius δ, in the supremum norm. By Theorem 2.1 and Remark 2.1 of [21], for any $r > 0$ there exists $\delta \in (0, r/2)$, such that if $f^1 \in \mathcal{B}_C(f, \delta)$, then $g^1 \in \mathcal{B}_C(g, r/2)$. This implies that

$$\sup_{t \in [0, t_0/2]} \left| \int_0^{t_0/2} \mathbf{n}(g^1(s)) d\ell^1_s - \int_0^{t_0/2} \mathbf{n}(g(s)) d\ell_s \right| \leq \sup_{t \in [0, t_0/2]} (|f^1(t) - f(t)| + |g^1(t) - g(t)|) \leq \delta + r/2 \leq r.$$

Thus, if $f^1 \in \mathcal{B}_C(f, \delta)$, then $\left(g^1(t_0/2), \int_0^{t_0/2} \mathbf{n}(g^1(s)) d\ell^1_s\right) \in \mathcal{B}(z_1, r) \times \mathcal{B}(-y_0, r)$. Let $\widetilde{\mathbb{P}}_x$ denote the distribution of standard Brownian motion. By the support theorem for Brownian motion, $\widetilde{\mathbb{P}}_{x_0}(\mathcal{B}_C(f, \delta)) > p_1$, for some $p_1 > 0$. Hence, $\mathbb{P}_{x_0,y_0}((X_{t_0/2}, K_{t_0/2}) \in \mathcal{B}(z_1, r) \times \mathcal{B}(0, r)) > p_1 > 0$.

Step 2. This step is mostly a review of the excursion theory needed in the rest of the argument. See, e.g., [22] for the foundations of excursion theory in abstract settings and [8] for the special case of excursions of Brownian motion. See also [17] for excursions of reflected Brownian motion on C^3-smooth domains. Although [8] does not discuss reflecting Brownian motion, all the results we need from that book readily apply in the present context. We will use two different but closely related exit systems. The first one represents excursions of reflecting Brownian motion from ∂D.

We consider X under a probability measure $\mathbb{P}_x$, i.e., X denotes reflecting Brownian motion without inert drift.

An exit system for excursions of reflecting Brownian motion X from ∂D is a pair (L^*_t, H_x) consisting of a positive continuous additive functional L^*_t and a family of excursion laws $\{H_x\}_{x\in\partial D}$. We will soon show that $L^*_t = L_t$. Let Δ denote a "cemetery" point outside $\mathbb{R}^d$ and let $\mathcal{C}$ be the space of all functions $f : [0,\infty) \to \mathbb{R}^d \cup \{\Delta\}$ which are continuous and take values in $\mathbb{R}^d$ on some interval $[0,\zeta)$, and are equal to Δ on $[\zeta,\infty)$. For $x \in \partial D$, the excursion law H_x is a σ-finite (positive) measure on $\mathcal{C}$, such that the canonical process is strong Markov on (t_0,∞) for every $t_0 > 0$, with the transition probabilities of Brownian motion killed upon hitting ∂D. Moreover, H_x gives zero mass to paths which do not start from x. We will be concerned only with "standard" excursion laws; see Definition 3.2 of [8]. For every $x \in \partial D$ there exists a standard excursion law H_x in D, unique up to a multiplicative constant.

Excursions of X from ∂D will be denoted e or e_s, i.e., if $s < u$, $X_s, X_u \in \partial D$, and $X_t \notin \partial D$ for $t \in (s,u)$, then $e_s = \{e_s(t) = X_{t+s},\ t \in [0, u-s)\}$ and $\zeta(e_s) = u - s$. By convention, $e_s(t) = \Delta$ for $t \ge \zeta$, so $e_t \equiv \Delta$ if $\inf\{s > t : X_s \in \partial D\} = t$. Let $\mathcal{E}_u = \{e_s : s \le u\}$.

Let $\sigma_t = \inf\{s \ge 0 : L^*_s \ge t\}$ and let I be the set of left endpoints of all connected components of $(0,\infty) \setminus \{t \ge 0 : X_t \in \partial D\}$. The following is a special case of the exit system formula of [22]. For every $x \in \overline{D}$,

$$\mathbb{E}_x \left[\sum_{t\in I} Z_t \cdot f(e_t)\right] = \mathbb{E}_x \int_0^\infty Z_{\sigma_s} H_{X(\sigma_s)}(f) ds = \mathbb{E}_x \int_0^\infty Z_t H_{X_t}(f) dL^*_t, \tag{6.1}$$

where Z_t is a predictable process and $f : \mathcal{C} \to [0,\infty)$ is a universally measurable function which vanishes on those excursions e_t identically equal to Δ. Here and elsewhere $H_x(f) = \int_{\mathcal{C}} f dH_x$.

The normalization of the exit system is somewhat arbitrary, for example, if (L^*_t, H_x) is an exit system and $c \in (0,\infty)$ is a constant then $(cL^*_t, (1/c)H_x)$ is also an exit system. Let $\mathbb{P}^D_y$ denote the distribution of Brownian motion starting from y and killed upon exiting D. Theorem 7.2 of [8] shows how to choose a "canonical" exit system; that theorem is stated for the usual planar Brownian motion but it is easy to check that both the statement and the proof apply to reflecting Brownian motion in $D \subset \mathbb{R}^d$. According to that result, we can take L^*_t to be the continuous additive functional whose Revuz measure is a constant multiple of the surface area measure dx on ∂D and the H_x's to be standard excursion laws normalized so that for some constant $c_1 \in (0,\infty)$,

$$H_x(A) = c_1 \lim_{\delta\downarrow 0} \frac{1}{\delta} \mathbb{P}^D_{x+\delta\mathbf{n}(x)}(A) \tag{6.2}$$

for any event A in a σ-field generated by the process on an interval $[t_0, \infty)$ for any $t_0 > 0$. The Revuz measure of L is $c_2 dx$ on ∂D. We choose c_1 so that (L_t, H_x) is an exit system.

We will now discuss another exit system, for a different process X'. Let $U \subset D$ be a fixed closed ball with non-zero radius, and let X' be the process X conditioned by the event $\{T_U^X > \sigma_1\}$, where $T_U^X = \inf\{t > 0 : X \in U\}$. One can show using Theorem 2.1 and Remark 2.1 of [21] that for any starting point in $D \setminus U$, the probability of $\{T_U^X > \sigma_1\}$ is greater than 0. It is easy to see that (X'_t, L_t) is a time-homogeneous Markov process. For notational consistency, we will write (X'_t, L'_t) instead of (X'_t, L_t).

We will now describe an exit system $(L'_t, H'_{x,\ell})$ for (X'_t, L'_t). We will construct this exit system on the basis of (L_t, H_x) because of the way that X' has been defined in relation to X. It is clear that L' does not change within any excursion interval of X' away from ∂D, so we will assume that $H'_{x,\ell}$ is a measure on paths representing X' only. The local time L' is the continuous additive functional with Revuz measure $c_2 dx$ on ∂D. For $\ell \geq 1$ we let $H'_{x,\ell} = H_x$. Let $\widehat{\mathbb{P}}_y^D$ denote the distribution of Brownian motion starting from $y \in D \setminus U$, conditioned to hit ∂D before hitting U, and killed upon exiting D. For $\ell < 1$, we have

$$H'_{x,\ell}(A) = c_1 \lim_{\delta \downarrow 0} \frac{1}{\delta} \widehat{\mathbb{P}}^D_{x+\delta \mathbf{n}(x)}(A). \tag{6.3}$$

Let $A_* \subset \mathcal{C}$ be the event that the path hits U. It follows from (6.2) and (6.3) that for $\ell < 1$,

$$H'_{x,\ell}(A) = H_x(A \setminus A_*). \tag{6.4}$$

One can deduce easily from (6.2) and standard estimates for Brownian motion that for some $c_3, c_4 \in (0, \infty)$ and all $x \in \partial D$,

$$c_3 < H_x(A_*) < c_4. \tag{6.5}$$

Let $\sigma'_t = \inf\{s \geq 0 : L'_s \geq t\}$. The exit system formula (6.1) and (6.4) imply that we can construct X (on a random interval, to be specified below) using X' as a building block, in the following way. Suppose that X' is given. We enlarge the probability space, if necessary, and construct a Poisson point process $\mathcal{E}$ with state space $[0, \infty) \times \mathcal{C}$ whose intensity measure conditional on the whole trajectory $\{X'_t, t \geq 0\}$ is given by

$$\mu([s_1, s_2] \times F) = \int_{1 \wedge s_1}^{1 \wedge s_2} H_{X'_{\sigma'_t}}(F \cap A_*) dt.$$

Since $\mu([0, \infty) \times \mathcal{C}) < \infty$, the Poisson point process $\mathcal{E}$ may be empty. Consider the case when it is not empty and let S_1 be the minimum of the first coordinates of points in $\mathcal{E}$. Note that there can be only one point $(S_1, e_{S_1}) \in \mathcal{E}$ with first coordinate S_1, because of (6.5). By convention, let $S_1 = \infty$ if $\mathcal{E} = \emptyset$. Recall that $T_U^X = \inf\{t > 0 : X_t \in U\}$ and let

$$\begin{aligned} T_U^{X'} &= \inf\{t > 0 : X'_t \in U\}, \\ T_* &= \sigma'_{S_1} + \inf\{t > 0 : e_{S_1}(t) \in U\}. \end{aligned}$$

It follows from the exit system formula (6.1) that the distribution of the process

$$\widehat{X}_t = \begin{cases} X'_t & \text{if } 0 \le t \le T_U^{X'} \wedge \sigma'_{S_1}, \\ e_{S_1}(t - \sigma'_{S_1}) & \text{if } \mathcal{E} \neq \emptyset \text{ and } \sigma'_{S_1} < t \le T_*, \end{cases}$$

is the same as the distribution of $\{X_t, 0 \le t \le T_U^X\}$.

Step 3. We will now construct reflecting Brownian motion in D from several trajectories, including a family of independent paths.

Recall z_1 from Step 1. Let $U_j = \overline{\mathcal{B}(z_j, r)}$ for $j = 1, \dots, d+1$, where $z_j \in D$ and $r > 0$ are chosen so that $U_j \cap U_k = \emptyset$ for $j \neq k$, and $\bigcup_{1 \le j \le d+1} U_j \subset D$.

Recall from the last step how the process X was constructed from a process X'. Fix some $x_1 \in U_1$ and let X^1 be a process starting from $X_0^1 = x_1$, with the same transition probabilities as X', relative to U_2. We then construct Y^1 based on X^1, by adding an excursion that hits U_2, in the same way as X was constructed from X'. We thus obtain a process $\{Y_t^1, 0 \le t \le T_1\}$, where $T_1 = \inf\{t > 0 : Y_t^1 \in U_2\}$, whose distribution is that of reflecting Brownian motion in D, observed until the first hit of U_2.

We next construct a family of independent reflecting Brownian motions $\{Y^j\}_{1 \le j \le d}$. For a fixed $j = 2, \dots, d$, we let X^j be a process with the same transition probabilities as X', relative to U_{j+1}, and initial distribution uniform in U_j. We then construct Y^j based on X^j, by adding an excursion that hits U_{j+1}, in the same way as X was constructed from X'. We thus obtain a process $\{Y_t^j, 0 \le t \le T_j\}$, where $T_j = \inf\{t > 0 : Y_t^1 \in U_{j+1}\}$, whose distribution is that of reflecting Brownian motion in D, observed until the first hit of U_{j+1}.

Note that for some $c_5 > 0$ and all $x, y \in U_{j+1}$, $j = 1, \dots, d$,

$$\mathbb{P}_x(X_1 \in dy, X_t \notin \partial D \text{ for } t \in [0,1]) \ge c_5 dy.$$

We can assume that all X^j's and Y^j's are defined on the same probability space. The last formula and standard coupling techniques show that on an enlarged probability space, there exist reflecting Brownian motions Z^j, $j = 1, \dots, d$, with the following properties. For $1 \le j \le d-1$, $Z_0^j = Y_{T_j}^j$, and for some $c_6 > 0$,

$$\mathbb{P}\left(Z_1^j = Y_0^{j+1}, Z_t^j \notin \partial D \text{ for } t \in [0,1] \mid \{Y^k\}_{1 \le k \le j}, \{Z^k\}_{1 \le k \le j-1}\right) \ge c_6. \tag{6.6}$$

The process Z^j does not depend otherwise on $\{Y^k\}_{1 \le k \le d}$ and $\{Z^k\}_{k \neq j}$. We define Z^d as a reflecting Brownian motion in D with $Z_0^d = Y_{T_d}^d$ but otherwise independent of $\{Y^k\}_{1 \le k \le d}$ and $\{Z^k\}_{1 \le k \le d-1}$.

Let $F_j = \{Z_1^j = Y_0^{j+1}, Z_t^j \notin \partial D \text{ for } t \in [0,1]\}$. We define a process X^* as follows. We let $X_t^* = Y_t^1$ for $0 \le t \le T_1$. If F_1^c holds, then we let $X_t^* = Z_{t-T_1}^1$ for $t \ge T_1$. If F_1 holds, then we let $X_t^* = Z_{t-T_1}^1$ for $t \in [T_1, T_1 + 1]$ and $X_t^* = Y_{t-T_1-1}^2$ for $t \in [T_1 + 1, T_1 + 1 + T_2]$. We proceed by induction. Suppose that X_t^* has been defined so far only for

$$t \in [0, T_1 + 1 + T_2 + 1 + \cdots + T_k],$$

for some $k < d$. If F_k^c holds, then we let

$$X_t^* = Z^k_{t-T_1-1-T_2-1-\cdots-T_k}$$

for $t \geq T_1 + 1 + T_2 + 1 + \cdots + T_k$. If F_k holds, then we let

$$X_t^* = Z^k_{t-T_1-1-T_2-1-\cdots-T_k}$$

for $t \in [T_1 + 1 + T_2 + 1 + \cdots + T_k, T_1 + 1 + T_2 + 1 + \cdots + T_k + 1]$ and

$$X_t^* = Y^{k+1}_{t-T_1-1-T_2-1-\cdots-T_k-1}$$

for $t \in [T_1 + 1 + T_2 + 1 + \cdots + T_k + 1, T_1 + 1 + T_2 + 1 + \cdots + T_k + 1 + T_{k+1}]$. We let

$$X_t^* = Z^d_{t-T_1-1-T_2-1-\cdots-T_d}$$

for $t \geq T_1 + 1 + T_2 + 1 + \cdots + T_k$.

By construction, X^* is a reflecting Brownian motion in D starting from x_1. Note that in view of (6.6), conditional on $\{X_t^j, t \geq 0\}$, $j = 1, \dots, d$, there is at least probability c_6^d that X^* is a time-shifted path of X_t^j on an appropriate interval, for all $j = 1, \dots, d$.

Step 4. In this step, we will show that with a positive probability, the process K can have "almost" independent and "almost" perpendicular increments over disjoint time intervals. Moreover, the distributions of the increments have densities in an appropriate sense.

We find d points $y_1, \dots, y_d \in \partial D$ such that the $\mathbf{n}(y_j)$'s point in d orthogonal directions. Let $C_j = \{\mathbf{z} \in \mathbb{R}^d : \angle(\mathbf{n}(y_j), \mathbf{z}) \leq \delta_0\}$, for some $\delta_0 > 0$ so small that for any $\mathbf{z}_j \in C_j$, $j = 1, \dots, d$, the vectors $\{\mathbf{z}_j\}$ are linearly independent. Let $\delta_1 > 0$ be so small that for every $j = 1, \dots, d$, and any $y \in \partial D \cap \mathcal{B}(y_j, \delta_1)$, we have $\mathbf{n}(y) \in C_j$.

Let L^j be the local time of X^j on ∂D and $\sigma_t^j = \inf\{s \geq 0 : L_s^j \geq t\}$. It is easy to see that there exists $p_2 > 0$ such that with probability greater than p_2, for every $j = 1, \dots, d$, we have $X_t^j \notin \partial D \setminus \mathcal{B}(y_j, \delta_1)$, for $t \in [0, \sigma_1^j]$. Let $R_j = \sup\{t < T_j : Y_t^j \in \partial D\}$ and $S_j = L^j_{R_j}$. Let F_* be the event that for every $j = 1, \dots, d$, $X_t^j \notin \partial D \setminus \mathcal{B}(y_j, \delta_1)$ for $t \in [0, \sigma_1^j]$ and $S_j < \sigma_1^j$. Then (6.5) shows that $\mathbb{P}_{x_1}(F_*) \geq p_2(1 - e^{-c_3})^d$.

Let $K_t^j = \int_0^t \mathbf{n}(X^j_{\sigma_s^j}) ds$ and note that if F_* holds, then $K_t^j \in C_j$ for all $j = 1, \dots, d$ and $t \in [0, 1]$.

Let

$$\Lambda([a_1, b_1], [a_2, b_2], \dots, [a_d, b_d]) = \{K^1_{t_1} + K^2_{t_2} + \cdots + K^d_{t_d} : t_k \in [a_k, b_k] \text{ for } 1 \leq k \leq d\}.$$

It is easy to show using the definition of C_j's that the d-dimensional volume of $\Lambda([a_1, b_1], \dots, [a_d, b_d])$ is bounded below by $c_7 \prod_{1 \leq k \leq d}(b_k - a_k)$, and bounded above by $c_8 \prod_{1 \leq k \leq d}(b_k - a_k)$.

Let us consider the processes defined above, conditioned on the σ-field

$$\mathcal{G} = \sigma\left(\{K_t^j, t \in [0, 1]\}_{1 \leq j \leq d}\right).$$

By (6.5), conditional on $\mathcal{G}$, the random variables S_j have distributions whose densities on $[0,1]$ are bounded below. In view of our remarks on the volume of Λ, it follows that conditional on $\mathcal{G}$, the vector $K^1_{S_1} + \cdots + K^d_{S_d}$ has a density with respect to d-dimensional Lebesgue measure that is bounded below by $c_9 > 0$ on a ball U_* with positive radius. We now remove the conditioning to conclude that $K^1_{S_1} + \cdots + K^d_{S_d}$ has a density with respect to d-dimensional Lebesgue measure that is bounded below on U_*.

Let $K^*_t = \int_0^t \mathbf{n}(X^*_{\sigma^j_s})ds$ and $T_* = T_d$. Using conditioning on F_*, we see that the distribution of $K^*_{T_*}$ has a component with density greater than c_9 on U_*. Since K^* does not change when X^* is inside the domain and X^* is a reflecting Brownian motion, we conclude that $(X^*_{T_*+1}, K^*_{T_*+1})$ has a density with respect to $2d$-dimensional Lebesgue measure on a non-empty open set. It follows that for some fixed $t_* > 0$, $(X^*_{t_*}, K^*_{t_*})$ has a density with respect to $2d$-dimensional Lebesgue measure on a non-empty open set.

We leave it to the reader to check that the argument can be easily modified so that we can show that for any fixed $t_0 > 0$, $(X^*_{t_0/2}, K^*_{t_0/2})$ has a density with respect to $2d$-dimensional Lebesgue measure on a non-empty open set. We can now combine this with the result of Step 1 using the Markov property to see that for some non-empty open set $\widetilde{U}$ and any starting point $(X_0, K_0) = (x_0, y_0)$, the process (X_{t_0}, K_{t_0}) has a positive density with respect to $2d$-dimensional Lebesgue measure on $\widetilde{U}$ under $\mathbb{P}_{x_0,y_0}$.

By the Girsanov theorem, the same conclusion holds for (X, K) under the measure $\mathbb{Q}_{x_0,y_0}$. □

Theorem 6.2 *Consider the SDE* (2.1) *with σ being the identity matrix,* $\mathbf{b} \equiv 0$ *and* $\mathbf{u} \equiv \mathbf{v} \equiv \mathbf{n}$. *The probability distribution $\mu(dx, dy)$ defined by*

$$\mu(A) = \int_A c_1 1_D(x) e^{-|y|^2}\, dx\, dy, \qquad A \subset \overline{D} \times \mathbb{R}^d,$$

is the only invariant measure for the solution (X, K) to (2.1).

Proof. In view of Theorem 5.3, all we have to show is uniqueness. If there were more than one invariant measure, at least two of them (say μ and ν) would be mutually singular by Birkhoff's ergodic theorem [27]. However, we have just shown that there exists a strictly positive measure ψ which is absolutely continuous with respect to any transition probability, so that in particular, $\psi \ll \mu$ and $\psi \ll \nu$. Since $\mu \perp \nu$, there exists a set A such that $\mu(A) = 0$ and $\nu(A^c) = 0$. Therefore, one must have $\psi(A) = \psi(A^c) = 0$ which contradicts the fact that ψ is strictly positive. □

References

[1] R. Bass, *Diffusions and Elliptic Operators.* Berlin, Springer, 1997.

[2] R. Bass, K Burdzy and Z.-Q. Chen, Uniqueness for reflecting Brownian motion in lip domains. *Ann. Inst. Henri Poincare Probab. Statist.* **41** (2005), 197-235.

[3] R. F. Bass, K Burdzy and Z.-Q. Chen, Pathwise uniqueness for a degenerate stochastic differential equation. *Ann. Probab.* **35** (2007) 2385–2418.

[4] R. F. Bass and P. Hsu, Some potential theory for reflecting Brownian motion in Hölder and Lipschitz domains. *Ann. Probab.* **19** (1991), 486–508.

[5] M. Benaïm, M. Ledoux and O. Raimond, Self-interacting diffusions. *Probab. Theory Related Fields* **122**, (2002), 1–41.

[6] M. Benaïm and O. Raimond, Self-interacting diffusions. II. Convergence in law. *Ann. Inst. H. Poincaré Probab. Statist.* **39**, (2003), 1043–1055.

[7] M. Benaïm and O. Raimond, Self-interacting diffusions. III. Symmetric interactions. *Ann. Probab.* **33**, (2005), 1717–1759.

[8] K. Burdzy, *Multidimensional Brownian Excursions and Potential Theory*, Longman Sci. Tech., Harlow, 1987.

[9] K. Burdzy, R. Hołyst and Ł. Pruski, Brownian motion with inert drift, but without flux: a model. *Physica A* **384** (2007) 278–284.

[10] K. Burdzy and D. White, A Gaussian oscillator. *Electron. Comm. Probab.* **9** (2004) paper 10, pp. 92–95.

[11] Z.-Q. Chen, On reflecting diffusion processes and Skorokhod decompositions. *Probab. Theory Related Fields* **94**, (1993), 281–315.

[12] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite-Dimensional Systems*, Cambridge University Press 1996.

[13] P. Dupuis and H. Ishii, SDEs with oblique reflection on nonsmooth domains. *Ann. Probab.* **21**, (1993), 554–580.

[14] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 1986.

[15] M. Fukushima, Y. Oshima and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*. de Gruyter, 1994.

[16] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, Springer 1985.

[17] P. Hsu, On excursions of reflecting Brownian motion. *Trans. Amer. Math. Soc.* **296** (1986), 239-264.

[18] F. Knight, On the path of an inert object impinged on one side by a Brownian particle. *Probab. Theory Related Fields* **121**, (2001) 577–598.

[19] T. Kurtz and P. Protter, Weak Limit Theorems for Stochastic Integrals and Stochastic Differential Equations. *Ann. Probab.* **19**, (1991), 1035–1070.

[20] T.M. Liggett, *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[21] P. L. Lions and A. S. Sznitman, Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37** (1984), 511-537.

[22] B. Maisonneuve, Exit systems, *Ann. Probability* **3** (1975), no. 3, 399–411.

[23] D. Nualart, *The Malliavin Calculus and Related Topics*. Springer 1995.

[24] E. Pardoux and R. J. Williams, Symmetric reflected diffusions. *Ann. Inst. H. Poincar Probab. Statist.* **30** (1994), 13-62.

[25] R. Pemantle, A survey of random processes with reinforcement. *Probability Surveys* **4**, (2007) 1–79.

[26] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion, third edition*. Springer 1999.

[27] Y. Sinai, *Topics in Ergodic Theory*, Princeton University Press 1994,

[28] D.W. Stroock and S.R.S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability theory*, 333–359. Univ. California Press, Berkeley, Calif., 1972.

[29] D. White, Processes with inert drift. Ph.D. Thesis, University of Washington (2005)

[30] D. White, Processes with inert drift. Mathematics ArXiv, preprint math.PR/0604052 (2006)

[31] R.J. Williams and W. Zheng, On reflecting Brownian motion—a weak convergence approach. *Ann. Inst. H. Poincaré Probab. Statist.* **26** (1990), no. 3, 461–488.

[32] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** (1971) 155–167.

Richard F. Bass:
Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009, USA.
Email: `bass@math.uconn.edu`

Krzysztof Burdzy:
Department of Mathematics, University of Washington, Seattle, WA 98195, USA.
Email: `burdzy@math.washington.edu`

Zhen-Qing Chen:
Department of Mathematics, University of Washington, Seattle, WA 98195, USA.
Email: `zchen@math.washington.edu`

Martin Hairer:
Mathematics Institute, The University of Warwick, Coventry CV4 7AL, UK.
Email: `M.Hairer@Warwick.ac.uk`