Probab. Theory Relat. Fields (2010) 146:1–47
DOI 10.1007/s00440-008-0182-6

Stationary distributions for diffusions with inert drift

Richard F. Bass · Krzysztof Burdzy · Zhen-Qing Chen · Martin Hairer

Received: 25 April 2008 / Revised: 5 October 2008 / Published online: 4 November 2008

Abstract Consider reflecting Brownian motion in a bounded domain in $\mathbb{R}^d$ that 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 joint law of the position of the reflecting Brownian motion and the value of the drift vector has a product form. Moreover, the first component is uniformly distributed on the domain, and the second component has a Gaussian distribution. We also consider more general reflecting diffusions with inert drift as well as processes where the drift is given in terms of the gradient of a potential.

Mathematics Subject Classification (2000) Primary: 60H10; Secondary: 60J55 · 60J60

Research supported in part by NSF Grants DMS-0601783 and DMS-0600206 and by EPSRC Grant EP/D071593/.

K. Burdzy and M. Hairer gratefully acknowledge the hospitality and support of the Institut Mittag-Leffler, where part of this research was done.

R. F. Bass
Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009, USA
e-mail: bass@math.uconn.edu

K. Burdzy (✉) · Z.-Q. Chen
Department of Mathematics, University of Washington, Seattle, WA 98195, USA
e-mail: burdzy@math.washington.edu

Z.-Q. Chen
e-mail: zchen@math.washington.edu

M. Hairer
Mathematics Institute, The University of Warwick, Coventry CV4 7AL, UK
e-mail: M.Hairer@Warwick.ac.uk

1 Introduction

This article is concerned with the higher dimensional version of a one-dimensional model originally introduced by Knight [22] and studied in more detail in [34,35]. Computer simulations presented in [9] led to the conjecture that the stationary distribution for this higher dimensional process has a certain interesting structure. We prove this conjecture, and moreover answer questions about the stationary distribution left open in [9].

We start with a presentation of the model in a simple case, that of [9]. We consider a bounded smooth domain $D \subset \mathbb{R}^d$ and reflecting Brownian motion X_t in $\overline{D}$ with drift K_t. Let B_t be d-dimensional Brownian motion, $\mathbf{n}(x)$ be the unit inward normal vector of D at $x \in \partial D$ and let L_t be the local time of X on ∂D, that is, a nondecreasing one-dimensional process with continuous paths that increases only when $X_t \in \partial D$. The pair of processes (X, K) has the following representation:

$$\begin{cases} X_t = X_0 + B_t + \int_0^t \mathbf{n}(X_s)\, dL_s + \int_0^t K_s\, ds, \\ K_t = K_0 + \int_0^t \mathbf{n}(X_s) dL_s, \end{cases}$$

with $(X_t, K_t) \in \overline{D} \times \mathbb{R}^d$ for all $t \geq 0$. Note that X_t is reflecting Brownian motion in $\overline{D}$ with normal reflection at the boundary and with drift K_t, K_t is an $\mathbb{R}^d$-valued process that represents the accumulated local time on the boundary in the direction normal to the boundary, and the drift K_t does not change when X_t is in the interior of D. We call the process K "inert drift" because it plays a role analogous to the inert particle in Knight's original model [22] in one dimension. The main simulation result of [9], which is done for a flat two-dimensional torus with a closed subset A removed from its center, where A is either a disk, an ellipse or a square, 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 prove rigorously in this paper that this indeed holds, and we further 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 where 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 [24], in particular, Theorem 2.1 on page 380. As we will see at the beginning of Sect. 4, the product form of the stationary distribution in our model comes naturally from a computation with infinitesimal generators.

When $d = 1$ and D is a finite interval in $\mathbb{R}$, it is shown in [10,34] that the process K, when time-changed by the local time L, has a Gaussian stationary distribution. The results of the present paper yield as a special case that the process K under its original time clock also has a Gaussian stationary distribution. Moreover, we show that if the inward normal vector field $\mathbf{n}$ in the equation for K is replaced by $\Gamma\mathbf{n}$ for some constant symmetric positive definite matrix Γ, (X, K) continues to have a product form for the stationary distribution, but this time the component K has a Gaussian distribution with covariance matrix Γ.

The main goal of this paper is to address the existence and uniqueness of the stationary distribution of normally reflecting Brownian motion with inert drift and to give an explicit formula for the stationary distribution. We also consider a larger class of reflecting diffusions, including what is sometimes known as distorted reflecting Brownian motion—see Theorem 5.2 below. Distorted reflected Brownian motion is the reflecting diffusion with generator $\frac{1}{2\rho}\nabla(\rho\nabla)$ for a suitable function ρ.

We start by showing in Sect. 2 the weak existence and weak uniqueness of solutions to an SDE representing a large family of diffusions with reflection and inert drift. More specifically, let ρ be a C^2 function on $\overline{D}$ that is bounded between two positive constants and $A(x) = (a_{ij}(x))_{n\times n}$ 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 $\mathbf{u}(x) := A(x)\mathbf{n}(x)$ is called the conormal vector at $x \in \partial D$. Let $\sigma(x) = (\sigma_{ij}(x))$ be the positive symmetric square root of $A(x)$. For notational convenience, we sometimes use ∂_i to denote $\frac{\partial}{\partial x_i}$. For $\varphi \in C^2(\mathbb{R}^d)$, let

$$\mathcal{L}\varphi(x) := \frac{1}{2\rho(x)} \sum_{i,j=1}^{d} \partial_i \left(\rho(x) a_{ij}(x) \partial_j \varphi(x)\right). \tag{1.1}$$

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

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

Let B be standard d-dimensional Brownian motion and $\mathbf{v}$ a bounded measurable vector field on ∂D. Consider the following diffusion process X taking values in $\overline{D}$ such that for all $t \geq 0$,

$$\begin{cases} dX_t = \sigma(X_t)\,dB_t + \mathbf{b}(X_t)\,dt + \frac{1}{2}(A\nabla\log\rho)(X_t)dt + \mathbf{u}(X_t)\,dL_t + K_t\,dt, \\ t \mapsto L_t \text{ is continuous and non-decreasing with } L_t = \int_0^t \mathbf{1}_{\partial D}(X_s)\,dL_s, \\ dK_t = \mathbf{v}(X_t)\,dL_t. \end{cases} \tag{1.2}$$

In Theorem 2.1, we show that the above stochastic differential equation (SDE) has a unique weak solution (X, K) for every starting point $(x_0, k_0) \in \overline{D} \times \mathbb{R}^d$. The solution (X, K) of (1.2) is called a (symmetric) reflecting diffusion on D with inert drift. Let X^0 be symmetric reflecting diffusion on D with infinitesimal generator $\mathcal{L}$ in (1.1); that is, X^0 is a continuous process taking values in $\overline{D}$ such that for every $t \geq 0$,

$$\begin{cases} dX_t^0 = \sigma(X_t^0)\,dB_t + \mathbf{b}(X_t^0)\,dt + \frac{1}{2}(A\nabla\log\rho)(X_t)dt + \mathbf{u}(X_t^0)\,dL_t^0, \\ t \mapsto L_t^0 \text{ is continuous and non-decreasing with } L_t^0 = \int_0^t \mathbf{1}_{\partial D}(X_s^0)\,dL_s^0. \end{cases} \tag{1.3}$$

The continuous non-decreasing process L^0 is called the boundary local time of X^0. When σ is the identity matrix, X^0 is distorted reflecting Brownian motion on $\overline{D}$. The main observation of Sect. 2 is that the reflecting diffusion (X, K) with inert drift can be obtained from the reflecting diffusion $\left(X^0, K_0 + \int_0^\cdot \mathbf{n}(X_s^0)\, dL_s^0\right)$ without inert drift by a suitable Girsanov transform, and vice versa. A further change of measure (see Theorem 2.3) shows that, if $\mathbf{c}$ is a bounded $\mathbb{R}^d$-valued function on $\overline{D}$, the following SDE with "generalized inert drift":

$$\begin{cases} dX_t = \sigma(X_t)\, dB_t + \mathbf{b}(X_t)\, dt + \frac{1}{2}(A\nabla \log \rho)(X_t)dt + \mathbf{u}(X_t)\, dL_t + K_t\, dt, \\ t \mapsto L_t \text{ is continuous and non-decreasing with } L_t = \int_0^t \mathbf{1}_{\partial D}(X_s)\, dL_s, \\ dK_t = \mathbf{v}(X_t)\, dL_t + \mathbf{c}(X_t)dt, \end{cases} \tag{1.4}$$

has a unique weak solution (X, K) taking values in $\overline{D} \times \mathbb{R}^d$ for every starting point $(x_0, k_0) \in \overline{D} \times \mathbb{R}^d$.

The questions of strong existence and strong uniqueness for solutions to (1.2) are discussed in Sect. 3. They are resolved positively under the additional assumption that $\mathbf{v} = a_0\mathbf{u}$ for some constant $a_0 \in \mathbb{R}$. This section uses some ideas and results from [25], but the main idea of our argument is different from the one in that paper, and we believe ours is somewhat simpler.

In Sect. 4 we consider symmetric diffusions with drift given as the gradient of a potential. We do this 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. More specifically, let Γ be a symmetric positive definite constant $d \times d$-matrix and $V \in C^1(D)$ tending to infinity in a suitable way as x approaches the boundary ∂D. Consider the following diffusion process X on D associated with generator $\mathcal{L} = \frac{1}{2}e^V \nabla \left(e^{-V} A \nabla\right)$ but with an additional "inert" drift K_t:

$$\begin{cases} dX_t = \sigma(X_t)\, dB_t + \mathbf{b}(X_t)\, dt - \frac{1}{2}(A\nabla V)(X_t)\, dt + K_t\, dt, \\ dK_t = -\frac{1}{2}\Gamma\, \nabla V(X_t)\, dt. \end{cases} \tag{1.5}$$

Here $A = A(x) = \sigma^T \sigma$ is a $d \times d$ matrix-valued function that is uniformly elliptic and bounded, with $\sigma_{ij} \in C^1(\overline{D})$. Note that Eq. (1.5) is missing the term $\mathbf{u}(X_t)\, dL_t$ which is present in (1.2). This is because the process defined in (1.5) never hits the boundary of the domain. We show in Theorem 4.3 that if $e^{-V/2} \in W_0^{1,2}(D)$, then the SDE (1.5) has a unique conservative solution (X, K) which has a stationary probability distribution

$$\pi(dx, dy) = c_1 \mathbf{1}_D(x) e^{-V(x)}\, e^{-(\Gamma^{-1} y, y)}\, dx\, dy.$$

In other words, (X, K) has a product form stationary distribution, with $c\, e^{-V}\, dx$ in the X component, and a Gaussian distribution with covariance matrix Γ in the K component. Observe that $c\, e^{-V}\, dx$ with normalizing constant $c > 0$ is the stationary distribution for the conservative symmetric diffusion X^0 in D with generator

$\frac{1}{2}e^V\nabla(e^{-V}A\nabla)$. The uniqueness of the stationary measure for solutions of (1.5) is addressed in Proposition 4.8 under much stronger conditions.

In Sect. 5, we prove weak convergence of a sequence of solutions of (1.5) to symmetric reflecting diffusions with inert drift given by (1.2) where $\rho \equiv 1$ and $\mathbf{v} = \Gamma\,\mathbf{n}$ for some symmetric positive definite constant matrix Γ. This implies

$$\pi(dx, dy) = c_2 \mathbf{1}_D(x)\, e^{-(\Gamma^{-1}y, y)}\, dx\, dy$$

is a stationary distribution for solutions (X, K) of (1.2) with $\rho \equiv 1$ and $\mathbf{v} = \Gamma\mathbf{n}$; see Theorem 5.2. Observe that $c_3\,\mathbf{1}_D(x)\,dx$ with $c_3 > 0$ being a normalizing constant is the stationary distribution for the symmetric reflecting diffusion X^0 on $\overline{D}$ with generator $\frac{1}{2}\nabla(A\nabla)$ in (1.3). If one prefers to have a stationary measure of the form $\pi(dx, dy) = c_3\mathbf{1}_D(x)\rho(x)\, e^{-(\Gamma^{-1}y, y)}\, dx\, dy$, where $c_3 > 0$ is a normalizing constant, then one needs to consider the SDE (X, K) with generalized inert drift (1.4) by taking $\mathbf{c}(x)$ there to be $\Gamma\nabla \log \rho(x)$; see Theorem 5.2.

Finally, Sect. 6 completes our program by showing irreducibility in the sense of Harris for reflecting diffusions with inert drift given by (1.2) under the assumption that $\rho \equiv 1$, $\mathbf{v} = \Gamma\,\mathbf{n}$ for some symmetric positive definite constant matrix Γ, A is the identity matrix, $\mathbf{u} = \mathbf{n}$, and $\mathbf{b} = 0$. Uniqueness of the stationary distribution follows from the irreducibility.

The level of generality of our results varies 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 2.3. The calculation at the beginning of Sect. 4 indicates that in order for the solution (X, K) of (1.2) to have a product form stationary distribution, the inert drift vector field $\mathbf{v}$ has to be of the form $\Gamma\,\mathbf{n}$ for some symmetric positive definite constant matrix Γ. We leave the verification of this conjecture as another open problem.

Our model belongs to a family of processes with "reinforcement" surveyed by Pemantle in [30]; see especially Sect. 6 of that survey and references therein. Papers [4–6] 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 Boris Rozovskii and David White for very useful advice. We also thank the referee for suggestions that significantly improved the substance and presentation of this paper.

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 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 ρ be a C^2 function on $\overline{D}$ that is bounded between two positive constants and $A(x) = (a_{ij}(x))_{n\times n}$ 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 $\mathbf{u}(x) := A(x)\mathbf{n}(x)$ is called the conormal vector at $x \in \partial D$. Clearly there exists $c_1 > 0$ such that $\mathbf{u}(x) \cdot \mathbf{n}(x) \geq c_1$ for all $x \in \partial D$. Let $\sigma(x) = (\sigma_{ij}(x))$ be the positive symmetric square root of $A(x)$. For $\varphi \in C^2(\mathbb{R}^d)$, let

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

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

$$\widehat{b}_k(x) = \frac{1}{2\rho(x)} \sum_{i=1}^{d} \partial_i(\rho(x) a_{ik}(x)).$$

Thus $\widehat{b}_k$ is the same as $\mathcal{L}$ operating on the function $f_k(x) = x_k$. Note that $\widehat{\mathbf{b}} = \mathbf{b} + \frac{1}{2} A \nabla \log \rho$.

Let B be a 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 $t \geq 0$:

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

The proof of the next theorem says that reflecting diffusions with inert drift can be obtained from the corresponding symmetric reflecting diffusions without inert drift by suitable Girsanov transforms, and vice versa.

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,

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

with $X_t \in \overline{D}$ for every $t \geq 0$. Weak existence and uniqueness of solutions to (2.2) follow from [11] or Corollary 5.2 in [14], where the existence and uniqueness of a strong solution are given (see also [15]). 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 part of the proof uses only the C^1-smoothness of the domain and Lipschitz continuity of a_{ij} and ρ. 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

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

It is routine to check this defines a probability measure $\mathbb{Q}_{x,y}$ on $\mathcal{F}_\infty$. By the Girsanov theorem (cf. [31]), 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 $n \geq 1$,

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

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 $\varepsilon = 2^{-j} > 0$ where $j \geq 1$ 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 dth axis contains $\mathbf{n}(x_0)$. Let $\mathbf{n}_0 = \mathbf{n}(x_0)$. Recall that D has a C^1 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

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

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 $F_n \cap \{U_1 \le R_n\}$,

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

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

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

This is a contradiction because $X_{U_2} \in \partial D$ and the dth 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 $S_n \le R_n$ 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 $K_{S_n} \le 2/\varepsilon$ for $n \le [\varepsilon^{-2} c_2/(6c_3)]$. We have

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

Since $|\mathbf{v}(x)| \le c_3$, we have $L_{R_n} - L_{S_n} \ge 2/(c_3\varepsilon)$, so the dth 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 dth coordinate of X_{U_3} must be greater than or equal to

$$\begin{aligned}(2c_2)/(c_3\varepsilon) - \left|Y_{R_n-S_n} + \int_{S_n}^{R_n} K_t\,dt\right| &\ge (2c_2)/(c_3\varepsilon) - \varepsilon - \int_{S_n}^{S_n+\varepsilon^2} 4/\varepsilon\,dt \\ &= (2c_2)/(c_3\varepsilon) - 5\varepsilon \ge \varepsilon.\end{aligned}$$

This is a contradiction because $X_{U_3} \in \partial D$ and the dth 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 dth coordinate of $\int_{S_n}^{U_4} \mathbf{u}(X_t)\, dL_t$ is bounded below by 6ε. But then on F_n, the dth 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 dth 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 $\varepsilon = 2^{-j} > 0$,

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

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

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

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 that $\mathbb{Q}'_{x,y} = \mathbb{Q}_{x,y}$ on $\mathcal{F}_\infty$. □

Remark 2.2 The assumptions that D is a bounded C^2 domain and that the a_{ij}'s and ρ are C^2 on $\overline{D}$ are only used in the weak existence and uniqueness of the solution X to (2.2). The remaining proof only requires that D be a bounded C^1 domain and that the a_{ij}'s and ρ are Lipschitz in $\overline{D}$. In fact, when D is a bounded C^1 domain and the a_{ij}'s and ρ are Lipschitz on $\overline{D}$, the weak existence of solutions to (2.2) follows from [11] and so we have the weak existence to the SDE (2.1). We believe the weak uniqueness for solutions to (2.2) and consequently to (2.1) also holds under this weaker assumption by an argument analogous to that in [2, Section 4]. However to give the full details of the proof would 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].

For later use, we present a result on diffusions with "generalized inert drift." Let $\mathbf{c}$ be a bounded $\mathbb{R}^d$-valued function on $\overline{D}$. Consider the following system of stochastic differential equations with the extra condition that $X_t \in \overline{D}$ for all $t \geq 0$:

$$\begin{cases} dX_t = \sigma(X_t)\, dB_t + \widehat{\mathbf{b}}(X_t)\, dt + \mathbf{u}(X_t)\, dL_t + K_t\, dt, \\ t \mapsto L_t \text{ is continuous and non-decreasing with } L_t = \int_0^t \mathbf{1}_{\partial D}(X_s)\, dL_s, \\ dK_t = \mathbf{v}(X_t)\, dL_t + \mathbf{c}(X_t)dt. \end{cases} \tag{2.3}$$

Theorem 2.3 *Under the conditions of this section and the assumption that* $\mathbf{c}$ *is a bounded* $\mathbb{R}^d$*-valued function on* $\overline{D}$*, 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* (2.3) *with* $(X_0, K_0) = (x, y)$.

Proof Note that the weak existence and uniqueness for solution of (2.3) with $(X_0, K_0) = (x, y)$ is equivalent to that for the solution of

$$X_t = x + \int_0^t \sigma(X_s)\, dB_s + \int_0^t \widehat{\mathbf{b}}(X_s)\, ds + \int_0^t \mathbf{u}(X_s)\, dL_s + \int_0^t \left(y + \int_0^s \mathbf{v}(X_r) dL_r \right) ds + \int_0^t \left(\int_0^s \mathbf{c}(X_r) dr \right) ds, \tag{2.4}$$

with $X_t \in \overline{D}$. Similarly, the weak existence and uniqueness for the solution of (2.1) with $(X_0, K_0) = (x, y)$ is equivalent to that for the solution of

$$X_t = x + \int_0^t \sigma(X_s)\, dB_s + \int_0^t \widehat{\mathbf{b}}(X_s)\, ds + \int_0^t \mathbf{u}(X_s)\, dL_s + \int_0^t \left(y + \int_0^s \mathbf{v}(X_r) dL_r \right) ds, \tag{2.5}$$

with $X_t \in \overline{D}$. The solutions to (2.4) and (2.5) are related by a Girsanov transform. Since $\mathbf{c}$ is bounded, the result of this theorem follows from Theorem 2.1. □

3 Pathwise uniqueness

We will prove strong existence and strong uniqueness for solutions to (2.1) under assumptions stronger than those in Sect. 2, namely, we will assume that the vector field $\mathbf{v}$ is a fixed constant multiple of $\mathbf{u}$. Throughout this section, D is a bounded C^2 domain in $\mathbb{R}^d$, each a_{ij} and ρ are $C^{1,1}$ on $\overline{D}$ and so the vector $\widehat{\mathbf{b}}$ in (2.1) is Lipschitz continuous. Our approach to the strong existence and uniqueness for solutions to (2.1) uses some ideas and results from [25], but the main idea of our argument is different from the one in that paper, and we believe ours is somewhat simpler. It is probably possible to produce a proof along the lines of [25], 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 $a_0 = 0$, $K_t = K_0$ for every $t \geq 0$. In this case, the result follows from [14] (as mentioned above there is a gap in the proof of Case 2 in [14]; 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.

We let $\Lambda = 2A^{-1}$. The matrix-valued function $\Lambda(x) = \{\lambda_{ij}(x)\}_{1 \leq i,j \leq d}$, $x \in \mathbb{R}^d$, is symmetric and such that $x \to \Lambda(x)$ is uniformly elliptic and in C_b^2. We have

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

Moreover, since the vector $\mathbf{n}(x)$ points inwards, there exists $c_1 < \infty$ such that

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

Let $c_2 > 1$ be such that

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

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

Let

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

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 $\varepsilon, c_0 > 0$, and let

$$\begin{aligned} T_1 &= \inf\{t > 0 : |X_t - X_t'| > 0\},\\ T_k &= (T_{k-1} + t_1) \wedge \inf\{t > T_{k-1} : |X_t - X_t'| \vee |U_t - U_t'| \ge \lambda^{k-1}\varepsilon\}\\ &\quad \wedge \inf\{t > T_{k-1} : |L_t - L_{T_{k-1}}| \vee |L_t' - L_{T_{k-1}}'| > c_0\}, \quad k \ge 2. \end{aligned}$$

We will specify the values of 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' on some random interval $[0, T^*]$, and at the same time $U_t \ne U_t'$ or $K_t \ne K_t'$ for some $t \in [0, T^*]$. Hence, with probability 1, $U_t = U_t'$ and $K_t = K_t'$ for $t \in [0, T_1]$.

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 $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,S]} |X_t - X_t'| \le \varepsilon\lambda^{k_0} \right) \ge p_1. \tag{3.4}$$

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, in view of the remarks following the definition. 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

$$g(t_1) \le f(t_1)e^{bt_1}. \tag{3.5}$$

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 $t \ge T_1$,

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

By (3.5), for $t \ge T_1$,

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

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

$$\sup_{0 \le s \le T_k} |X_s - X'_s| \le \lambda^{k-1}\varepsilon \quad \text{and} \quad \sup_{0 \le s \le T_k} |U_s - U'_s| \le \lambda^{k-1}\varepsilon. \tag{3.6}$$

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

$$|K_t - K'_t| \le 4e^{2|a_0|}|a_0|\lambda^{k-1}\varepsilon. \tag{3.7}$$

By Ito's formula we have

$$
\begin{aligned}
&(X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k}-X'_{T_k})-(X_{T_{k-1}}-X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}}-X'_{T_{k-1}}) \quad (3.8)\\
&= \int_{T_{k-1}}^{T_k} d\left((X_t - X'_t)^T \Lambda(X_t)(X_t - X'_t)\right)\\
&= 2\int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(\sigma(X_t) - \sigma(X'_t))\, dB_t \quad (3.9)\\
&\quad +2\int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(\widehat{\mathbf{b}}(X_t) - \widehat{\mathbf{b}}(X'_t))\, dt \quad (3.10)\\
&\quad +2\int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(K_t - K'_t)\, dt \quad (3.11)\\
&\quad +2\int_{T_{k-1}}^{T_k} \mathbf{u}(X_t)^T \Lambda(X_t)(X_t - X'_t)\, dL_t \quad (3.12)\\
&\quad -2\int_{T_{k-1}}^{T_k} \mathbf{u}(X'_t)^T \Lambda(X_t)(X_t - X'_t)\, dL'_t \quad (3.13)\\
&\quad +\int_{T_{k-1}}^{T_k} \sum_{i=1}^{d}(X_t - X'_t)^T \frac{\partial}{\partial x_i}\Lambda(X_t)(X_t - X'_t)\, dX^i_t \quad (3.14)\\
&\quad +\frac{1}{2}\int_{T_{k-1}}^{T_k} \sum_{i,j=1}^{d}(X_t - X'_t)^T \left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\Lambda(X_t)\right)(X_t - X'_t)(\sigma^2(X_t))_{ij}\, dt \quad (3.15)\\
&\quad +2\int_{T_{k-1}}^{T_k} \sum_{i,j,n,m=1}^{d}(X^i_t-(X')^i_t)\frac{\partial}{\partial x_n}\lambda_{ij}(X_t)\sigma_{nm}(X_t)(\sigma_{jm}(X_t)-\sigma_{jm}(X'_t))dt \quad (3.16)\\
&\quad +\int_{T_{k-1}}^{T_k} \sum_{i,j,m=1}^{d}\lambda_{ij}(X_t)(\sigma_{im}(X_t) - \sigma_{im}(X'_t))(\sigma_{jm}(X_t) - \sigma_{jm}(X'_t))\, dt. \quad (3.17)
\end{aligned}
$$

We are now going to bound each of these terms individually in order to show that (3.4) holds. Recall that $\lambda = c_2^3 > 1$. Taking $c_3 > 0$ so that $\lambda^{1/3}c_3 + \lambda^{-4/3} < 1$, it follows from (3.3) that if

$$(X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k} - X'_{T_k}) - (X_{T_{k-1}} - X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}} - X'_{T_{k-1}}) \\ \leq c_3 \lambda^{2(k-1)} \varepsilon^2 \tag{3.18}$$

and

$$|X_{T_{k-1}} - X'_{T_{k-1}}| \leq \lambda^{k-2} \varepsilon, \tag{3.19}$$

then

$$|X_{T_k} - X'_{T_k}| < \lambda^{k-1} \varepsilon. \tag{3.20}$$

In view of our assumptions on Λ and σ,

$$\mathrm{Var}\left(2 \int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(\sigma(X_t) - \sigma(X'_t))\, dB_t \;\Big|\; \mathcal{F}_{T_{k-1}} \right) \\ \leq \mathbb{E}\left(c_4 \int_{T_{k-1}}^{T_k} \sup_{T_{k-1} \leq s \leq T_k} |X_s - X'_s|^4 \, dt \;\Big|\; \mathcal{F}_{T_{k-1}} \right) \leq c_4 \lambda^{4(k-1)} \varepsilon^4 t_1.$$

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

$$\mathbb{P}\left(\left| 2 \int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(\sigma(X_t) - \sigma(X'_t))\, dB_t \right| \geq (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2 \;\Big|\; \mathcal{F}_{T_{k-1}} \right) \\ \leq \frac{1}{100}. \tag{3.21}$$

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

$$\left| 2 \int_{T_{k-1}}^{T_k} (X_t - X'_t)^T \Lambda(X_t)(\widehat{\mathbf{b}}(X_t) - \widehat{\mathbf{b}}(X'_t))\, dt \right| \leq c_5 \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2 \, dt \tag{3.22}$$
$$\leq c_5 t_1 \lambda^{2(k-1)} \varepsilon^2 \leq (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2.$$

We use (3.7) and make $t_1 > 0$ smaller, if necessary, to obtain

$$\begin{aligned}\left|2\int_{T_{k-1}}^{T_k}(X_t-X_t')^T\Lambda(X_t)(K_t-K_t')\,dt\right| &\le \left|c_6\int_{T_{k-1}}^{T_k}|(X_t-X_t')^T\Lambda(X_t)|\lambda^{k-1}\varepsilon\,dt\right|\\ &\le c_7\int_{T_{k-1}}^{T_k}|X_t-X_t'|\lambda^{k-1}\varepsilon\,dt\\ &\le (1/100)c_3\lambda^{2(k-1)}\varepsilon^2. \end{aligned} \tag{3.23}$$

We now apply (3.2) and make $c_0 > 0$ in the definition of T_k sufficiently small so that

$$\begin{aligned}2\int_{T_{k-1}}^{T_k}\mathbf{u}(X_t)^T\Lambda(X_t)(X_t-X_t')\,dL_t &\le 2c_1\int_{T_{k-1}}^{T_k}|X_t-X_t'|^2\,dL_t\\ &\le (1/100)c_3\lambda^{2(k-1)}\varepsilon^2. \end{aligned} \tag{3.24}$$

We have

$$\begin{aligned}&-2\int_{T_{k-1}}^{T_k}\mathbf{u}(X_t')^T\Lambda(X_t)(X_t-X_t')\,dL_t'\\ &\quad=2\int_{T_{k-1}}^{T_k}\mathbf{u}(X_t')^T\Lambda(X_t')(X_t'-X_t)\,dL_t'\\ &\qquad-2\int_{T_{k-1}}^{T_k}\mathbf{u}(X_t')^T(\Lambda(X_t)-\Lambda(X_t'))(X_t-X_t')\,dL_t', \end{aligned}$$

so, by using (3.2) as before and making $c_0 > 0$ even smaller, we obtain the following estimate,

$$\begin{aligned}&-2\int_{T_{k-1}}^{T_k}\mathbf{u}(X_t')^T\Lambda(X_t)(X_t-X_t')\,dL_t'\\ &\quad\le 2c_1\int_{T_{k-1}}^{T_k}|X_t-X_t'|^2\,dL_t'+c_8\int_{T_{k-1}}^{T_k}|X_t-X_t'|^2\,dL_t'\\ &\quad\le (1/100)c_3\lambda^{2(k-1)}\varepsilon^2. \end{aligned} \tag{3.25}$$

We have

$$\begin{aligned}
&\int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t)\, dX_t^i \\
&= \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t) \sum_{j=1}^{d} \sigma_{ij}(X_t)\, dB_t^j \\
&\quad + \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t)(\widehat{b}_i(X_t) + K_t^i)\, dt \\
&\quad + \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t)\mathbf{u}^i(X_t)\, dL_t. \qquad (3.26)
\end{aligned}$$

The three terms on the right hand side of the above formula can be estimated in the same way as in (3.21), (3.22) and (3.25). Hence,

$$\begin{aligned}
&\mathbb{P}\left(\left| \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t) \sum_{j=1}^{d} \sigma_{ij}(X_t)\, dB_t^j \right| \right. \\
&\qquad \left. \geq (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2 \;\middle|\; \mathcal{F}_{T_{k-1}} \right) \\
&\leq \frac{1}{100}, \qquad (3.27)
\end{aligned}$$

$$\left| \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t)(\widehat{b}_i(X_t) + K_t^i)\, dt \right| \leq (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2, \qquad (3.28)$$

$$\left| \int_{T_{k-1}}^{T_k} \sum_{i=1}^{d} (X_t - X'_t)^T \frac{\partial}{\partial x_i} \Lambda(X_t)(X_t - X'_t)\mathbf{u}^i(X_t)\, dL_t \right| \leq (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2. \qquad (3.29)$$

The following estimate is completely analogous to (3.22), and may require that we make $t_0 > 0$ smaller,

$$\begin{aligned}
&\left| \frac{1}{2} \int_{T_{k-1}}^{T_k} \sum_{i,j=1}^{d} (X_t - X'_t)^T \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \Lambda(X_t) \right) (X_t - X'_t)(\sigma^2(X_t))_{ij}\, dt \right. \\
&\quad + 2 \int_{T_{k-1}}^{T_k} \sum_{i,j,n,m=1}^{d} (X^i_t - (X')^i_t) \frac{\partial}{\partial x_n} \lambda_{ij}(X_t) \sigma_{nm}(X_t)(\sigma_{jm}(X_t) - \sigma_{jm}(X'_t))\, dt \\
&\quad \left. + \int_{T_{k-1}}^{T_k} \sum_{i,j,m=1}^{d} \lambda_{ij}(X_t)(\sigma_{im}(X_t) - \sigma_{im}(X'_t))(\sigma_{jm}(X_t) - \sigma_{jm}(X'_t))\, dt \right| \\
&\quad \le c_9 \int_{T_{k-1}}^{T_k} |X_t - X'_t|^2\, dt \\
&\quad \le c_9 t_1 \lambda^{2(k-1)} \varepsilon^2 \le (1/100) c_3 \lambda^{2(k-1)} \varepsilon^2. \qquad (3.30)
\end{aligned}$$

Combining (3.8)–(3.30), we see that conditioned on $\mathcal{F}_{T_{k-1}}$, with probability 3/4 or greater the following event holds:

$$\begin{aligned}
&(X_{T_k} - X'_{T_k})^T \Lambda(X_{T_k})(X_{T_k} - X'_{T_k}) - (X_{T_{k-1}} - X'_{T_{k-1}})^T \Lambda(X_{T_{k-1}})(X_{T_{k-1}} - X'_{T_{k-1}}) \\
&\quad \le c_3 \lambda^{2(k-1)} \varepsilon^2.
\end{aligned}$$

In view of (3.18)–(3.20), this implies that if (3.19) and the above display hold, then

$$|X_{T_k} - X'_{T_k}| < \lambda^{k-1} \varepsilon. \qquad (3.31)$$

Let

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

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

$$\mathbb{P}\left(\sup_{T_{k-1} \le s \le T'_k} |U_s - U'_s| \ge (1/2) \lambda^{k-1} \varepsilon \;\middle|\; \mathcal{F}_{T_{k-1}} \right) \le 1/4. \qquad (3.32)$$

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 (3.31) and (3.32),

$$\mathbb{P}(F_k \mid \mathcal{F}_{T_{k-1}}) \geq 1/2. \tag{3.33}$$

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 (3.6) and the definitions of T_1 and S,

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

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 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 [31, Section IX.1]. □

4 Diffusions with gradient drifts

This section is devoted to analysis of a diffusion with inert drift given as the gradient of a potential. We will use such diffusions to approximate diffusions with reflection, but the analysis of the stationary measure is easier in the case when the drift is smooth. Let V be a C^1 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

$$\begin{cases} dX_t &= \sigma(X_t)\, dB_t + \mathbf{b}(X_t)\, dt - \frac{1}{2}(A \nabla V)(X_t)\, dt + K_t\, dt, \\ dK_t &= -\frac{1}{2} \Gamma\, \nabla V(X_t)\, dt, \end{cases} \tag{4.1}$$

where $A = A(x) = (a_{ij}) = \sigma^T \sigma$ is uniformly elliptic and bounded. We assume that $\sigma \in C^1(\overline{D})$ and $\mathbf{b} = (b_1, \ldots, b_d)$ with $b_k(x) := \frac{1}{2} \sum_{i=1}^d \partial_i a_{ik}(x)$. Note that, since $\sigma \in C^1(\overline{D})$, $V \in C^1(D)$ and $\mathbf{b}$ is bounded, (4.1) has a unique weak solution (X, K) up to the time $\inf\{t > 0 : X_t \notin D \text{ or } K_t = \infty\}$.

To find a 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,

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

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

$$\begin{aligned} df(X_t, K_t) &= \nabla_x f dX_t \\ &\quad + \nabla_y f\, dK_t + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x_1, \dots, x_d, y_1, \dots, y_d) d\langle X^i, X^j \rangle_t \\ &= \text{local martingale} + \left(\mathcal{L}_x f + \nabla_x f \cdot K_t - \frac{1}{2} \nabla_y f \cdot \Gamma \nabla_x V\right) dt. \end{aligned}$$

So the process (X, K) has the generator

$$\mathcal{G} f(x, y) := \mathcal{L}_x f(x, y) + y \cdot \nabla_x f(x, y) - \frac{1}{2} \Gamma\, \nabla_x V(x) \cdot \nabla_y f(x, y) \tag{4.2}$$

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 $\pi(dx, dy) = \rho_1(x)\rho_2(y)\, dx\, dy$. Let $\mathcal{G}$ be defined by (4.2) and set $\mathcal{D}(\mathcal{G})$, the domain of $\mathcal{G}$, to be $C_c^2(D \times \mathbb{R}^d)$. It follows that $\mathcal{G}^* \pi = 0$, in the sense that for every $f \in \mathcal{D}(\mathcal{G})$,

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

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

$$\begin{aligned}
0 &= \int_{\mathbb{R}^d} \left(\int_D (\rho_1(x) y \cdot \nabla_x f(x,y) + \rho_1(x) \mathcal{L}_x f(x,y))\, dx \right) \rho_2(y)\, dy \\
&\quad - \frac{1}{2} \int_D \left(\int_{\mathbb{R}^d} \rho_2(y) \Gamma\, \nabla_x V(x) \cdot \nabla_y f(x,y)\, dy \right) \rho_1(x)\, dx \\
&= \int_{\mathbb{R}^d} \left(\int_D \big(-\nabla_x \rho_1(x) \cdot y f(x,y) + \mathcal{L}_x^* \rho_1(x) f(x,y)\big)\, dx \right) \rho_2(y)\, dy \\
&\quad + \frac{1}{2} \int_D \left(\int_{\mathbb{R}^d} \Gamma\, \nabla_x V(x) \cdot \nabla_y \rho_2(y) f(x,y)\, dy \right) \rho_1(x)\, dx.
\end{aligned}$$

This implies that

$$\rho_2(y) \big(\nabla_x \rho_1(x) \cdot y - \mathcal{L}_x^* \rho_1(x)\big) - \frac{1}{2} \rho_1(x) \Gamma\, \nabla_x V(x) \cdot \nabla_y \rho_2(y) = 0$$
$$\text{for } (x,y) \in D \times \mathbb{R}^d. \tag{4.3}$$

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

$$\rho_2(y) \nabla_x V(x) \cdot y + \frac{1}{2} \Gamma\, \nabla_x V(x) \cdot \nabla_y \rho_2(y) = 0 \quad \text{for } (x,y) \in D \times \mathbb{R}^d.$$

Since $V(x)$ blows up as x approaches the boundary ∂D, $\{\nabla_x V(x), x \in D\}$ spans the whole of $\mathbb{R}^d$. (If not, there exists $v \in \mathbb{R}^d$ such that $\langle v, \nabla V(x) \rangle \le 0$ for every $x \in \mathbb{R}^d$. One then gets a contradiction to the fact that there exists $r_0 > 0$ (possibly $r_0 = +\infty$) such that $\lim_{r \to r_0} V(x + vr) = \infty$ for every $x \in D$.) So we must have

$$y + \frac{1}{2} \Gamma^T \nabla_y \log \rho_2(y) = 0 \quad \text{for every } y \in \mathbb{R}^d. \tag{4.4}$$

This implies that Γ is symmetric. To see this, denote the entries of Γ by (γ_{ij}). Then for every $y = (y_1, \ldots, y_d)^T \in \mathbb{R}^d$ and $1 \le i \le d$, we have by (4.4)

$$y_i + \frac{1}{2} \sum_{j=1}^d \gamma_{ij} \frac{\partial}{\partial y_j} \log \rho_2(y) = 0.$$

Taking the partial derivative with respect to y_k yields

$$\delta_{ik} + \frac{1}{2}\sum_{j=1}^{d} \gamma_{ij} \frac{\partial^2}{\partial y_j \partial y_k} \log \rho_2(y) = 0 \quad \text{for every } 1 \le i, k \le d.$$

This says that $-\frac{1}{2}\Gamma$ is the inverse matrix of the Hessian of $\log \rho_2$, which is symmetric, and so is Γ. We obtain from (4.4) that

$$\nabla_y \log \rho_2(y) = -2\Gamma^{-1} y \quad \text{for every } y \in \mathbb{R}^d.$$

Hence we have

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

or

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

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

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

where $c_3 > 0$ is the normalizing constant. This is made rigorous in the following result. Recall that a process is said to be conservative if it does not explode in finite time. In the present setting, since we consider processes taking values on $D \times \mathbb{R}^d$, this means that the second component does not explode and that the first component does not reach the boundary of D.

Theorem 4.1 *Suppose that Γ is a symmetric positive definite matrix, each σ_{ij} is C^1 on $\overline{D}$ so that $A = \sigma^T \sigma$ is uniformly elliptic and bounded on D, and V is a C^1 potential on D. Suppose that*

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

is a probability measure on $D \times \mathbb{R}^d$ such that the diffusion process of (4.1) with initial distribution π is conservative. Then the process of (4.1) has π as a (possibly not unique) stationary distribution.

Proof Let $\mathcal{G}$ be the operator defined by (4.2) with domain $\mathcal{D}(\mathcal{G}) = C_c^2(D \times \mathbb{R}^d)$. The above calculation shows that

$$\int_{D \times \mathbb{R}^d} \mathcal{G} f(x, y) \pi(dx, dy) = 0 \quad \text{for every } f \in \mathcal{D}(\mathcal{G}). \tag{4.5}$$

Let $E := D \times \mathbb{R}^d$ and $E_\Delta = E \cup \{\Delta\}$ be the one-point compactification of E. Recall $\mathcal{G}$ is defined by (4.2) with $\mathcal{D}(\mathcal{G}) = C_c^2(E)$. Since (4.5) holds, we have by Theorem 9.17 in Chapter 4 of [16] that π is a stationary measure for some solution $\mathbb{P}$ to the martingale problem for $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$. In [16, Theorem 4.9.17], the measure $\mathbb{P}$ is a probability measure on the product space $E^{\mathbb{R}_+}$. However by [16, Corollary 4.3.7], any solution of the martingale problem for $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$ has a modification with sample paths in the Skorokhod space $\mathbb{D}(\mathbb{R}_+, E_\Delta)$ of right continuous paths on $\mathbb{E}_\Delta$ having left limits. Thus we can assume that $\mathbb{P}$ is a $\mathbb{D}(\mathbb{R}_+, E_\Delta)$-solution to the martingale problem $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$ with stationary distribution π. Let (X, K) denote the coordinate maps on E_Δ and set

$$\zeta := \inf\{t > 0 : (X_t, K_t) = \Delta\}.$$

We show next that $\{(X_t, K_t), t < \zeta\}$ under $\mathbb{P}$ is a solution to SDE (4.1) up to time ζ.

Let $f(x, y) = g(x)h(y)$ with $g \in C_c^2(D)$ and $h \in C_c^2(\mathbb{R}^d)$. Then under $\mathbb{P}$,

$$\begin{aligned} g(X_t)h(K_t) = g(X_0)h(K_0) &+ \text{martingale} \\ &+ \int_0^t \Big(h(K_s)\mathcal{L}_x f(X_s) + h(K_s)K_s \cdot \nabla_x g(X_s) \\ &- \frac{1}{2} g(X_s)\Gamma \nabla_x V(X_s) \cdot \nabla_y h(K_s) \Big)\, ds. \end{aligned} \tag{4.6}$$

Here and below, we use the convention that for any function f on E, $f(\Delta) := 0$. Let D_n be subdomains of D increasing to D, let $E_n = D_n \times B(0, n)$, and let $\tau_n = \inf\{t : (X_t, K_t) \notin E_n\}$. Suppose $g \in C^2(\overline{D})$. Let g_n be a function in $C_c^2(D)$ that equals g on D_n and h_n a function in $C_c^2(\mathbb{R}^d)$ that equals 1 on $B(0, n)$. Applying (4.6) to g_n and h_n, we have, under $\mathbb{P}$, that

$$g(X_t) = g(X_0) + \text{local martingale} + \int_0^t (\mathcal{L}_x g(X_s) + K_s \cdot \nabla_x g(X_s))\, ds \tag{4.7}$$

for $0 \le t \le \tau_n$. Letting $n \to \infty$, (4.7) holds for $0 \le t < \zeta$. Let $g(x) = x_i$. There are local martingales $M = (M^1, \dots, M^d)$ such that

$$dX_t = dM_t + \mathbf{b}(X_t)\, dt - \frac{1}{2}(A\nabla V)(X_t)\, dt + K_t\, dt, \quad t < \zeta.$$

Applying Ito's formula to $X_t^i X_t^j$ yields $\langle M^i, M^j \rangle_t = \langle X^i, X^j \rangle_t = \int_0^t a_{ij}(X_s)\, ds$ for $t < \zeta$. Define $dB_t := \sigma^{-1}(X_t)\, dM_t$. Then $\{B_t, t < \zeta\}$ is a Brownian motion up to time ζ and $M_t = \int_0^t \sigma(X_s)\, dB_s$ for $t < \zeta$. Hence we have

$$dX_t = \sigma(X_t)\, dB_t + \mathbf{b}(X_t)\, dt - \frac{1}{2}(A\nabla V)(X_t)\, dt + K_t\, dt, \quad t < \zeta. \tag{4.8}$$

Similarly to the derivation of (4.7), from (4.6), we have for every $h \in C^2(\mathbb{R}^d)$ that has bounded derivatives,

$$h(K_t) = h(K_0) + \text{local martingale} - \frac{1}{2}\int_0^t \Gamma\nabla_x V(X_s)\cdot\nabla_y h(K_s)\,ds, \quad t < \zeta.$$

In particular, taking $h(y) = y_i$, $1 \le i \le d$, there are local martingales $N = (N^1, \dots, N^d)$ such that

$$dK_t = dN_t - \frac{1}{2}\Gamma\nabla_x V(X_s)\,ds, \quad t < \zeta.$$

Applying Ito's formula to $K^i K^j$ yields $\langle N^i, N^j\rangle_t = \langle K^i, K^j\rangle_t = 0$ for $t < \zeta$. Hence

$$dK_t = -\frac{1}{2}\Gamma\nabla_x V(X_s)\,ds, \quad t < \zeta.$$

This together with (4.5) implies that $\{(X_t, K_t), t < \zeta\}$ under $\mathbb{P}$ is a (continuous) weak solution to (4.1) with initial distribution π up to time ζ. Since $a_{ij}, V \in C^1(\overline{D})$, weak uniqueness holds for solutions of (4.1) (see [33]). So under our conservativeness assumption, (4.1) has a conservative weak solution with initial distribution π which is unique in distribution. By standard techniques (cf. the proof of [1, Proposition I.2.1]), any weak solution to (4.1) gives rise to a solution to the martingale problem for $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$. This implies that $\zeta = \infty$ because (X, K) with the initial distribution π is a conservative solution to (4.1), by assumption. We conclude that π is a stationary distribution for (4.1). □

We will present an easily verifiable condition on V for Theorem 4.1 to be applicable. The following preliminary result for diffusions without inert drift may have interest of its own.

Recall that $W_0^{1,2}(D)$ is the closure of the space $C_c^\infty(D)$ of smooth functions with compact support in D under the Sobolev norm $\|u\|_{1,2} := \left(\int_D (|u(x)| + \sum_{i=1}^d |\partial_i u(x)|^2)\,dx\right)^{1/2}$.

Theorem 4.2 *Let $D \subset \mathbb{R}^d$ be a domain (i.e. a connected open set) and $A(x) = (a_{ij}(x))$ be a measurable $d \times d$ matrix-valued function on D that is uniformly elliptic and bounded. Suppose that φ is a function in $W_0^{1,2}(D)$ that is positive on D. Then the minimal diffusion X on D having infinitesimal generator $\mathcal{L} = \frac{1}{2\varphi^2}\sum_{i,j=1}^d \partial_i(\varphi^2 a_{ij}\partial_j)$ is conservative.*

Proof The process X is the symmetric diffusion (with respect to the symmetrizing measure $\varphi(x)^2\,dx$) associated with the Dirichlet form $(\mathcal{E}, \mathcal{F})$ in $L^2(D, \varphi(x)^2\,dx)$, where

$$\mathcal{E}(u, v) = \frac{1}{2}\int_D \sum_{i,j=1}^d a_{ij}(x)\partial_i u(x)\partial_j v(x)\,\varphi(x)^2\,dx,$$

and $\mathcal{F}$ is the closure of $C_c^\infty(D)$ with respect to the norm $\sqrt{\mathcal{E}_1}$; here we define

$$\mathcal{E}_1(u,u) := \mathcal{E}(u,u) + \int_D u(x)^2 \varphi(x)^2\,dx.$$

Let X^0 be the symmetric diffusion in D with respect to Lebesgue measure on D associated with the Dirichlet form $(\mathcal{E}^0, W_0^{1,2}(D))$ in $L^2(D, dx)$, where

$$\mathcal{E}^0(u,v) = \frac{1}{2}\int_D \sum_{i,j=1}^d a_{ij}(x)\partial_i u(x)\partial_j v(x)\,dx.$$

Then by [12, Theorems 2.6 and 2.8], X can be obtained from X^0 through a Girsanov transform using the martingale $dZ_t = \varphi(X_t^0)^{-1} dM_t^\varphi$, where M^φ is the martingale part of the Fukushima decomposition for $\varphi(X_t^0) - \varphi(X_0^0)$. We now conclude from [12, Theorem 2.6(ii)] that X is conservative. □

Recall the definition of $W_0^{1,2}(D)$ stated before Theorem 4.2.

Theorem 4.3 *Suppose that Γ is a constant symmetric positive definite matrix. Let $D \subset \mathbb{R}^d$ be a bounded domain, $\sigma = \sigma(x)$ be a $d \times d$ matrix that is C^1 on $\overline{D}$ so that $A = \sigma^T\sigma$ is uniformly elliptic and bounded on D. Suppose that V is a C^1 function in D such that $\varphi = e^{-V/2} \in W_0^{1,2}(D)$. Then the (minimal) solution of* (4.1) *with initial distribution π is conservative and has π as its stationary distribution.*

Proof Let $(X, \mathbb{P}_x)$ denote the symmetric diffusion process with infinitesimal generator

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

By Theorem 4.2, $(X, \mathbb{P}_x)$ is conservative. Let $K_t = K_0 - \frac{1}{2}\int_0^t (\Gamma\nabla V)(X_s)\,ds$. Define for $t \geq 0$,

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

Clearly $\{M_t, t \geq 0\}$ is a continuous positive $\mathbb{P}_x$-local martingale for every $x \in D$ with respect to the minimal augmented filtration $\{\mathcal{F}_t, t \geq 0\}$ of X. For $n \geq 1$, let $T_n := \inf\{t > 0 : |K_t| \geq 2^n\}$. Since X is conservative, we have $T_\infty := \lim_{n\to\infty} T_n = \infty$, $\mathbb{P}_x$-a.s., for all $x \in D$.

Let $\mathbb{P}$ be the distribution of the symmetric stationary diffusion X with initial probability distribution $c_1 e^{-V}\,dx$ in D. The existence of such a measure follows, for example, from Theorem 4.2. By enlarging the filtration if necessary, let K_0 have the

Gaussian distribution $c_2 \exp(-(\Gamma^{-1} y, y))$, independent of the symmetric diffusion X. Define a new probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = M_t \quad \text{on } \mathcal{F}_t \quad \text{for every } t \geq 0.$$

By the Girsanov theorem, the process

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

is a Brownian motion under $\mathbb{Q}$ up to the time T_∞, so under $\mathbb{Q}$,

$$\begin{cases} dX_t = \sigma(X_t)\, dW_t + \mathbf{b}(X_t)\, dt - \frac{1}{2}(A\nabla V)(X_t)\, dt + K_t\, dt, \\ dK_t = -\frac{1}{2}\Gamma \nabla V(X_t)\, dt, \end{cases} \qquad \text{for } 0 \leq t < T_\infty. \tag{4.9}$$

Note that (X, K) has initial distribution π under $\mathbb{Q}$.

Conversely, given a solution (X, K) of (4.9) under the measure $\mathbb{Q}$, the process X is a conservative $\mathcal{L}_x$-diffusion under the measure

$$\exp\left(-\sigma^{-1}(X_s) K_s \, dW_s - \frac{1}{2} \int_0^t |\sigma^{-1}(X_s) K_s|^2 \, ds \right) d\mathbb{Q}.$$

So to prove that the solution to (4.1) is conservative, it suffices to show that $\mathbb{Q}(T_\infty = \infty) = 1$.

We are going to show that $\{M_t, t \geq 0\}$ is in fact a positive $\mathbb{P}$-martingale. This implies that $\mathbb{Q}(T_\infty = \infty) = 1$ because $\mathbb{Q}(T_\infty > t) = \mathbb{E}_{\mathbb{P}}[M_t] = 1$ for every $t > 0$.

Let $E := D \times \mathbb{R}^d$ and $E_\Delta = E \cup \{\Delta\}$ be the one-point compactification of E. Recall $\mathcal{G}$ is defined by (4.2) with $\mathcal{D}(\mathcal{G}) = C_c^2(E)$. Since (4.5) holds, by the same argument as in the proof of Theorem 4.1, we deduce that π is a stationary measure for some solution $\widehat{\mathbb{Q}}$ on $\mathbb{D}(\mathbb{R}_+, E_\Delta)$ to the martingale problem for $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$. Let $(\widehat{X}, \widehat{K})$ denote the coordinate maps on E_Δ and set $\zeta := \inf\{t > 0 : (\widehat{X}_t, \widehat{K}_t) = \Delta\}$. Then $\{(\widehat{X}_t, \widehat{K}_t), t < \zeta\}$ satisfies the SDE (4.9), and consequently, it has the same distribution as $\{(X_t, K_t), t < T_\infty\}$ under $\mathbb{Q}$.

Note that the matrix σ^{-1} is bounded so there is a constant $c_1 > 0$ such that

$$|\sigma^{-1}(x)v| \leq c_1 |v| \quad \text{for every } x \in \overline{D} \text{ and } v \in \mathbb{R}^d.$$

Since under $\widehat{\mathbb{Q}}$, $\widehat{K}_t$ has the same Gaussian distribution for every $t \geq 0$, there exist c_2 and c_3 such that if $r \leq c_2$, then

$$\mathbb{E}_{\widehat{\mathbb{Q}}}\left[\exp\left(r \left|\sigma^{-1}(\widehat{X}_s)\widehat{K}_s\right|^2\right)\right] \leq \mathbb{E}_{\widehat{\mathbb{Q}}}\left[\exp\left(r c_1^2 \left|\widehat{K}_s\right|^2\right)\right] \leq c_3 \quad \text{for every } s \geq 0.$$

By Jensen's inequality applied with the measure $\frac{1}{t_0}1_{[0,t_0]}(s)\,ds$ with $t_0 \in (0, c_2/3]$ and the function e^x we have

$$\begin{aligned}\mathbb{E}_{\widehat{\mathbb{Q}}}\left[\exp\left(3\int_0^{t_0}\left|\sigma^{-1}(\widehat{X}_s)\widehat{K}_s\right|^2 ds\right)\right] &= \mathbb{E}_{\widehat{\mathbb{Q}}}\left[\exp\left(\frac{1}{t_0}\int_0^{t_0}\left(3t_0\left|\sigma^{-1}(\widehat{X}_s)\widehat{K}_s\right|^2\right) ds\right)\right] \\ &\le \mathbb{E}_{\widehat{\mathbb{Q}}}\left[\frac{1}{t_0}\int_0^{t_0}\exp\left(3t_0\left|\sigma^{-1}(\widehat{X}_s)\widehat{K}_s\right|^2\right) ds\right] \\ &\le \frac{1}{t_0}\int_0^{t_0} c_3\, ds = c_3.\end{aligned}$$

Define $N_t = \int_0^t \sigma^{-1}(X_s)K_s\,dB_s$. This is a martingale with respect to $\mathbb{P}$ and its quadratic variation $\langle N\rangle_t$ is equal to $\int_0^t |\sigma^{-1}(X_s)K_s|^2\,ds$ under both $\mathbb{P}$ and $\mathbb{Q}$. Note that $\exp(-N_t - \frac{1}{2}\langle N\rangle_t)$ is a positive local martingale with respect to $\mathbb{P}$ and hence a $\mathbb{P}$-supermartingale. Recall that T_∞ is the lifetime for (X, K), i.e., $(X_t, K) = \Delta$ for $t \ge T_\infty$, and that by convention, every function f on E is extended to a function on E_Δ by setting $f(\Delta) = 0$. In particular, we have $N_t = N_{t\wedge T_\infty}$. We have

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

Using Cauchy–Schwartz,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left[\exp\left(\int_0^t\left|\sigma^{-1}(X_s)K_s\right|^2 ds\right)\right] &= \mathbb{E}_{\mathbb{P}}\left[e^{\langle N\rangle_t}\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{\langle N\rangle_t}e^{-N_t-\frac{3}{2}\langle N\rangle_t}\right] \\ &\le \left(\mathbb{E}_{\mathbb{Q}}\left[e^{-2N_t}\right]\right)^{1/2}\left(\mathbb{E}_{\mathbb{Q}}\left[e^{3\langle N\rangle_t}\right]\right)^{1/2} \\ &\le \left(\mathbb{E}_{\mathbb{Q}}\left[\exp\left(3\int_0^t\left|\sigma^{-1}(X_s)K_s\right|^2 ds\right)\right]\right)^{1/2} \\ &\le \left(\mathbb{E}_{\widehat{\mathbb{Q}}}\left[\exp\left(3\int_0^t\left|\sigma^{-1}(\widehat{X}_s)\widehat{K}_s\right|^2 ds\right)\right]\right)^{1/2}.\end{aligned}$$

As we observed in the previous paragraph, the last term is bounded if $t \le c_2$. It follows from Novikov's criterion (see [31, Proposition VIII.1.15]) that $\{M_t, t \in [0, c_2]\}$ is a uniformly integrable $\mathbb{P}$-martingale. It follows that $\mathbb{Q}(T_\infty > c_2) = 1$. Using the Markov property, we have $\mathbb{Q}(T_\infty = \infty) = 1$. Consequently, π is a stationary distribution for (X, K) of (4.1). This proves the theorem. □

The next corollary follows immediately from Theorem 4.3 and the fact that the solution of (4.1) depends in a continuous way in its initial starting point (x_0, k_0).

Corollary 4.4 *Under the conditions of Theorem 4.3, for every $(x_0, k_0) \in D \times \mathbb{R}^d$, the minimal solution of (4.1) with initial value (x_0, k_0) is conservative.*

The following remark gives some sufficient conditions for $\varphi \in W_0^{1,2}(D)$ and thus for the condition of Theorem 4.1 to hold with $e^{-V} = \varphi^2$.

Remark 4.5 (i) Let $W^{1,2}(D) := \{f \in L^2(D, dx) : \nabla f \in L^2(D, dx)\}$. It is known (see, for example, [17]) that a function $u \in W^{1,2}(D)$ is in $W_0^{1,2}(D)$ if and only if its quasi-continuous version vanishes quasi-everywhere on ∂D. So in particular, if $u \in W^{1,2}(D)$ and if

$$\lim_{\substack{x \in D \\ x \to z}} u(x) = 0 \text{ for every } z \in \partial D,$$

then $u \in W_0^{1,2}(D)$.

(ii) For any positive bounded function $u \in W_0^{1,2}(D) \cap C^1(D)$, $\varphi = e^{-1/u} \in W_0^{1,2}(D) \cap C^1(D)$ and so Theorem 4.1 is applicable for $V = 2/u$ in view of Theorem 4.3. This follows because $f := \exp(-1/(2x))$ is bounded and Lipschitz continuous on $[0, k]$ for every $k \geq 1$. So by the normal contraction property (cf. [17]), $V = f(u) \in W_0^{1,2}(D)$. In particular, this is the case if u is the first positive eigenfunction for the Dirichlet Laplacian on D (if such a ground state exists).

(iii) The assumption that $e^{-V/2} \in C^1(D) \cap W_0^{1,2}(D)$ is close to being optimal for ensuring that the solution to (4.1) never hits the boundary of the domain D. Consider the case $n = 1$, $D = \mathbb{R}_+$, $\sigma = 1$, $V : \mathbb{R}_+ \to \mathbb{R}_+$, and remove the inert drift so that $dX = -\frac{1}{2}\nabla V(X)\,dt + dB$. Define a function $\Psi : \mathbb{R}_+ \to \mathbb{R}$ by

$$\Psi(x) = \int_1^x \exp(V(y))\,dy.$$

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 $\exp(V)$ is still integrable, this easily implies that this diffusion has a positive probability of reaching 0 in any finite time.

Let us now consider the case where $\exp(V(x)) = C/x^\alpha$ near $x = 0$. In this case, $\frac{d}{dx}e^{-V/2}$ is L^2-locally integrable near 0 if and only if $\alpha > 1$, which is almost the range of parameters for which $\exp(V)$ is no longer integrable at 0, namely, $\alpha \geq 1$.

The remainder of this section is devoted to the analysis of the case when both σ (and hence A) and Γ are the identity matrix. We consider the SDE

$$\begin{cases} dX_t = -\frac{1}{2}\nabla V(X_t)\,dt + K_t\,dt + dB_t, \\ dK_t = -\frac{1}{2}\nabla V(X_t)\,dt\ . \end{cases} \tag{4.10}$$

It is possible to show under very weak additional assumptions that π is the only invariant measure for (4.10). This is not obvious for example in the case where $V(x) = f(d(x, \partial D))$ for some function $f : \mathbb{R}_+ \to \mathbb{R}_+$ which has a singularity at 0 and is such that $f(r) = 0$ for r larger than some (small) constant ε. On the set $\{x : d(x, \partial D) > \varepsilon\} \times \mathbb{R}^d$, the diffusion (4.10) has then a deterministic component $dK = 0$. In particular, this shows that (4.10) is not hypoelliptic and that its transition probabilities are not absolutely continuous with respect to Lebesgue measure.

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

$$(\nabla_\alpha V(x))_k = \frac{\partial^{\ell+1} V(x)}{\partial x_k \partial x_{\alpha_1} \ldots \partial x_{\alpha_\ell}}.$$

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

Assumption 4.6 There exists $x_* \in D$ such that the collection $\{\nabla_\alpha V(x_*)\}_{|\alpha|>0}$ spans all of $\mathbb{R}^d$.

Remark 4.7 In the case where D is smooth and V is of the form $V(x) = f(d(x, \partial D))$ for some smooth function f diverging at 0, Assumption 4.6 is satisfied as soon as there exists a point on the boundary such that its curvature has full rank.

We then have

Proposition 4.8 *If $e^{-V/2} \in C^\infty(D) \cap W_0^{1,2}(D)$ and Assumption 4.6 holds, then $\pi(dx, dy) := c_0 \mathbf{1}_D(x) \exp(-V(x) - |y|^2)\,dx\,dy$ is the unique invariant measure for* (4.10).

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 $\phi : \mathcal{X} \to \mathbb{R}$. The proof of Proposition 4.8 is then based on the following fact, a version of which can be found for example in [18, Thm 3.16], which is a consequence of well-known results from [13, Chapter 4] and [27, Section 6.1.2]: 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) One has $x \in \mathit{supp}\,\pi$ for every invariant probability measure π for the semigroup $\{\mathcal{P}_t\}$,

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 (4.10). It is easy to check that Assumption 4.6 means precisely that the operator $\partial_t - \mathcal{G}$, where $\mathcal{G}$ is the generator of (4.10), satisfies Hörmander's condition [19,28] 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 $\Phi(t, y, K) = (\mathcal{P}_t\phi)\,(y, K)$ is a solution (in the sense of distributions) of the equation $(\partial_t - \mathcal{G})\,\Phi = 0$, 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 invariant measure of (4.10). This will be checked by showing first that (ii) follows from the following two properties:

(iii) for all $y \in \mathcal{X}$, there exists $s(y) \geq 0$ such that $x_* \in supp\, \mathcal{P}_{s(y)}(y, \cdot)$.
(iv) for every neighborhood U of x_* there exists a neighborhood $U' \subset U$ and a time $T_U > 0$ such that $\inf_{y \in U'} \mathcal{P}_t(y, U) > 0$ for every $t < T_U$.

The argument as to why (iii) and (iv) imply (ii) is the following:

Fix an arbitrary neighborhood U of x_* and an arbitrary invariant measure π. By the definition of an invariant measure, one then has

$$\pi(U) = \int_0^\infty \int_{\mathcal{X}} e^{-t} \mathcal{P}_t(y, U)\, \pi(dy)\, dt \geq \int_{\mathcal{X}} \int_0^{T_U} e^{-(s(y)+t)} \mathcal{P}_{s(y)+t}(y, U)\, dt\, \pi(dy)$$
$$= \int_{\mathcal{X}} \int_0^{T_U} \int_{U'} e^{-(s(y)+t)} \mathcal{P}_t(y', U) \mathcal{P}_{s(y)}(y, dy')\, dt\, \pi(dy) > 0.$$

In the last inequality, we used the facts that $\mathcal{P}_{s(y)}(y, U') > 0$ by (iii) and $\mathcal{P}_t(y', U) > 0$ by (iv). Since U was arbitrary, this inequality is precisely (ii).

Since (iv) is satisfied for every SDE with locally Lipschitz coefficients, it remains to check that (iii) is also satisfied. For this, it is enough 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 [33], so we consider the control system

$$\dot{x} = -\frac{1}{2} \nabla V(x) + K + u(t), \quad \dot{K} = -\frac{1}{2} \nabla V(x), \tag{4.11}$$

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 (4.11) 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 $\mathbb{R}^d$. (This fact has already been noted in the paragraph following (4.3).) 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 easily be 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

$$u(t) = \dot{X}_\varepsilon(t) + \frac{1}{2}\nabla V(X_\varepsilon(t)) + \frac{1}{2}\int_0^t \nabla V(X_\varepsilon(s)),$$

and to note that the solution to (4.11) is given by $\left(X_\varepsilon(t), K_0 - \int_0^t \nabla V(X_\varepsilon(s))\right)$. This solution has the desired properties by construction. □

Remark 4.9 The paper [7] also shows uniqueness of the invariant measure for a mechanical system for which hypoellipticity fails to hold on an open set. However, the techniques used there seem to be less straightforward than the argument used here and cannot easily be ported to our setting.

Remark 4.10 Through all of this section, the assumption that $\exp(-V/2) \in W_0^{1,2}(D)$ could have been replaced by the assumptions that $V \in \mathcal{C}^\infty$, $\lim_{x\to\partial D} V(x) = +\infty$, $\int_D \exp(-V(x))\,dx < \infty$, and there exists a constant $C \in \mathbb{R}$ such that the relation

$$\Delta V(x) \le |\nabla V(x)|^2 + C\,(1 + V(x)) \tag{4.12}$$

holds for every $x \in D$. This is because (4.12) ensures that $V(x) + \frac{K^2}{2}$ is a Lyapunov function for the diffusion (4.10) and so guarantees the non-explosion of solutions.

5 Weak convergence

In this section, we will prove that reflecting diffusions with inert drift can be approximated by diffusions with smooth inert drifts.

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) $c_1\delta_D(x) \le \delta(x) \le c_2\delta_D(x)$ for every $x \in \mathbb{R}^d$;
(ii) $\delta \in C^\infty(D)$ and for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $|\alpha| := \sum_{k=1}^d \alpha_k$, there is a constant c_α such that

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

The existence of such a regularized distance function δ is given by [36, Lemma 2.1]. For $n \ge 1$, define

$$V_n(x) := \begin{cases} \exp\left(\frac{1}{n\delta(x)}\right) & \text{for } x \in D, \\ +\infty & \text{for } x \in D^c. \end{cases} \tag{5.1}$$

Observe that $V_n \in C^\infty(D)$ and for every multi-index α, the α-derivative of e^{-V_n} at $x \in D$ tends to zero as $x \to \partial D$. So $e^{-V_n} \in C^\infty(\mathbb{R}^d)$. Observe also that as $n \to \infty$, $e^{-V_n(x)}$ increases to $e^{-1}\,\mathbf{1}_D(x)$.

Suppose that $A = \sigma^T \sigma$ is a uniformly elliptic and bounded matrix having C^2 entries a_{ij} on $\overline{D}$, and ρ is a C^2 function on $\overline{D}$ that is bounded between two positive constants.

We will consider the smooth potential approximations to the symmetric reflecting diffusion process on D associated with generator $\mathcal{L} = \frac{1}{2\rho}\nabla(\rho A \nabla)$ but with an additional generalized inert drift. More precisely, let Γ be a positive definite constant symmetric $d \times d$-matrix. As before, we use $\mathbf{n}(x)$ to denote the unit inward normal vector at $x \in \partial D$ and define $\mathbf{u}(x) := A(x)\mathbf{n}(x)$. We consider the following diffusion process X with "generalized inert drift" K:

$$\begin{cases} dX_t &= \sigma(X_t)\,dB_t + \mathbf{b}(X_t)\,dt + \frac{1}{2}(A\nabla \log \rho)(X_t)dt + \mathbf{u}(X_t)\,dL_t + K_t\,dt, \\ dK_t &= \Gamma\,\mathbf{n}(X_t)\,dL_t + \Gamma\,\nabla \log \rho(X_t)dt, \end{cases} \tag{5.2}$$

where $X_t \in \overline{D}$ for every $t \geq 0$, and L is continuous and non-decreasing with $L_t = \int_0^1 \mathbf{1}_{\partial D}(X_s)\,dL_s$. Here B is a d-dimensional Brownian motion and the drift $\mathbf{b}$ is given by $\mathbf{b} = (b_1, \ldots, b_d)$ with $b_k(x) := \frac{1}{2}\sum_{i=1}^d \partial_i a_{ik}(x)$. When $\rho \equiv 1$, (5.2) becomes

$$\begin{cases} dX_t &= \sigma(X_t)\,dB_t + \mathbf{b}(X_t)\,dt + \mathbf{u}(X_t)\,dL_t + K_t\,dt, \\ dK_t &= \Gamma\,\mathbf{n}(X_t)\,dL_t, \end{cases} \tag{5.3}$$

which is a special case of (2.1).

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

$$dX_t^{(n)} = \sigma(X_t^{(n)})\,dB_t^{(n)} + \mathbf{b}(X_t^{(n)})\,dt + \frac{1}{2}(A\nabla \log \rho)(X_t^{(n)})dt - \frac{1}{2}(A\nabla V_n)(X_t^{(n)})\,dt,$$

with initial probability distribution $\pi_n(dx) := c_n \exp(-V_n(x))\rho(x)\mathbf{1}_D(x)\,dx$, where $B^{(n)}$ is a d-dimensional Brownian motion, and $c_n > 0$ is a normalizing constant so that $\pi_n(\mathbb{R}^d) = 1$. Hence $X^{(n)}$ is the minimal diffusion in D with infinitesimal generator

$$\mathcal{L}^{(n)} = \frac{1}{2\rho(x)e^{-V_n(x)}} \sum_{i,j=1}^d \partial_i \left(\rho(x)e^{-V_n(x)} a_{ij}(x)\partial_j\right),$$

which is symmetric with respect to the measure π_n. Observe that here $V_n - \log \rho$ plays the role of the potential V in Sect. 3. Clearly $e^{(\log \rho - V_n)/2} \in C(D) \cap W_0^{1,2}(D)$ and so by Theorem 4.2, $X^{(n)}$ is conservative and never reaches ∂D. In addition, $X^{(n)}$ is a symmetric diffusion with respect to the measure π_n in D and its Dirichlet form

$(\mathcal{E}^{(n)}, \mathcal{F}^{(n)})$ in $L^2(D, \pi_n(dx))$ is given by (see [36, Lemma 3.5])

$$\mathcal{E}^{(n)}(f, f) := \frac{1}{2} \int_D \sum_{i,j=1}^{d} a_{ij}(x) \partial_i f(x) \partial_j f(x) \pi_n(dx),$$
$$\mathcal{F}^{(n)} := \left\{ f \in L^2(D, \pi_n) : \mathcal{E}^{(n)}(f, f) < \infty \right\}.$$

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

$$r_t(\omega)(s) := \begin{cases} \omega(t-s), & \text{if } 0 \le s \le t, \\ \omega(0), & \text{if } s \ge t. \end{cases} \tag{5.4}$$

Let $\mathbb{P}_n$ denote the law of $X^{(n)}$ with initial distribution π_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. [17, Lemma 5.7.1]). We know that

$$X_t^{(n)} - X_0^{(n)} = M_t^{(n)} + \int_0^t \left(\mathbf{b} + \tfrac{1}{2} A \nabla \log \rho - \tfrac{1}{2} A \nabla V_n \right) (X_s^{(n)})\, ds, \quad t \ge 0, \tag{5.5}$$

where $M_t^{(n)} := \int_0^t \sigma(X_s^{(n)})\, dB_s^{(n)}$ is a square-integrable martingale. On the other hand, by the Lyons–Zheng's forward and backward martingale decomposition (see [17, Theorem 5.7.1]), we have for every $T > 0$,

$$X_t^{(n)} - X_0^{(n)} = \frac{1}{2} M_t^{(n)} - \frac{1}{2} (M_T^{(n)} - M_{T-t}^{(n)}) \circ r_T \quad \text{for } t \in [0, T]. \tag{5.6}$$

Since σ is bounded, for every $T > 0$, the law of $\{M_t^{(n)}, t \ge 0\}$ under $\mathbb{P}_n$ is tight in the space $C([0, T], \mathbb{R}^d)$ and so is $\frac{1}{2}(M_T^{(n)} - M_{T-t}^{(n)}) \circ r_T$. It follows that the laws of $(X^{(n)}, M^{(n)}, (M_T^{(n)} - M_{T-\cdot}^{(n)}) \circ r_T)$ under $\mathbb{P}_n$ are tight in the space $C([0, T], \mathbb{R}^{3d})$ for every $T > 0$. On the other hand, by (5.5)–(5.6),

$$H^{(n)}(t) := -\frac{1}{2} \int_0^t (A \nabla V_n)(X_s^{(n)})\, ds$$
$$= -\frac{1}{2} B_t^{(n)} - \frac{1}{2} (B_T^{(n)} - B_{T-t}^{(n)}) \circ r_T - \int_0^t \mathbf{b}(X_s^{(n)})\, ds - \frac{1}{2} \int_0^t (A \nabla \log \rho)(X_s^{(n)})\, ds.$$

Thus the laws of $(X^{(n)}, M^{(n)}, \int_0^\cdot \mathbf{b}(X_s^{(n)})\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s^{(n)})ds, H^{(n)}, B^{(n)})$ under $\mathbb{P}_n$ are tight with respect to the space $C([0,T],\mathbb{R}^{6d})$ for every $T>0$. By almost the same argument as that for [29, Theorems 4.1 and 4.4], $(X^{(n)}, \mathbb{P}_n)$ converges weakly in $C([0,T],\mathbb{R}^d)$ to a stationary reflecting diffusion X in D with normalized initial distribution $c\rho(x)\,dx$ on D. Passing to a subsequence, if necessary, we conclude that

$$\left(\left(X^{(n)}, M^{(n)}, \int_0^\cdot \mathbf{b}(X_s^{(n)})\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s^{(n)})ds, H^{(n)}, B^{(n)}\right), \mathbb{P}_n\right)$$

converges weakly in $C([0,T],\mathbb{R}^{4d})$ to a process

$$\left(\left(X, M, \int_0^\cdot \mathbf{b}(X_s)\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s)ds, H, B\right), \mathbb{P}\right).$$

Now we apply the Skorokhod representation (see Theorem 3.1.8 in [16]) to construct the processes $\left(X^{(n)}, M^{(n)}, \int_0^\cdot \mathbf{b}(X_s^{(n)})\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s^{(n)})ds, H^{(n)}, B^{(n)}\right)$ and $\left(X, M, \int_0^\cdot \mathbf{b}(X_s)\,ds, \int_0^\cdot (A\nabla\log\rho)(X)ds, H, B\right)$ on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$ in such a way that $\left(X^{(n)}, M^{(n)}, \int_0^\cdot \mathbf{b}(X_s^{(n)})\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s^{(n)})ds,\right.$ $\left.H^{(n)}, B^{(n)}\right)$ converges to $\left(X, M, \int_0^\cdot \mathbf{b}(X_s)\,ds, \int_0^\cdot (A\nabla\log\rho)(X_s)ds, H, B\right)$ a.s., on the time interval $[0,1]$ in the supremum norm. Therefore, in view of (5.5),

$$X_t = X_0 + M_t + \int_0^t \mathbf{b}(X_s)\,ds + \frac{1}{2}\int_0^t (A\nabla\log\rho)(X_s)ds + H_t \quad \text{for every } t\ge 0.$$

By the proof of [29, Theorem 6.1 and Remark 6.2], H is a continuous process of locally finite variation. On the other hand, since X is reflecting diffusion in a Lipschitz domain, by [11] it admits a Skorokhod decomposition. That is,

$$\begin{aligned} X_t = X_0 &+ \int_0^t \sigma(X_s)\,dB_t + \int_0^t \mathbf{b}(X_s)\,ds + \frac{1}{2}\int_0^t (A\nabla\log\rho)(X_s)ds \\ &+ \int_0^t A(X_s)\mathbf{n}(X_s)\,dL_s, \quad t\ge 0, \end{aligned}$$

where B is a Brownian motion with respect to the minimal augmented filtration generated by X_t, $\mathbf{n}$ is the unit inward normal vector field on ∂D, which is well defined a.e. with respect to the surface measure μ on ∂D, and L is a positive continuous additive functional of X with Revuz measure $\frac{1}{2}\rho(x)\mu(dx)$. By the uniqueness of Doob–Meyer decomposition, we have

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

Recall that $V_n - \nabla \log \rho$ plays the role of the potential V in Sect. 3. Define for $t \geq 0$,

$$K_t^{(n)} := K_0^{(n)} - \frac{1}{2}\Gamma \int_0^t \nabla(V_n - \log \rho)(X_s^{(n)})\, ds \text{ and } K_t := K_0 + \Gamma \int_0^t \mathbf{n}(X_s)\, dL_s + \int_0^t \nabla \log \rho(X_s) ds.$$

Set

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

and

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

Let $d\mathbb{Q} = Z_1 d\mathbb{P}$ and $d\mathbb{Q}_n = Z_1^{(n)} d\mathbb{P}$. By the Girsanov theorem, under $\mathbb{Q}_n$, $(X^{(n)}, K^{(n)})$ satisfies the following equation

$$\begin{cases} dX_t^{(n)} = \sigma(X_t^{(n)})\, dW_t^{(n)} + \mathbf{b}(X_t^{(n)}) - \frac{1}{2}A\nabla(V_n - \log \rho)(X_t^{(n)})\, dt + K_t^{(n)}\, dt, \\ dK_t^{(n)} = -\frac{1}{2}\Gamma\, \nabla(V_n - \log \rho)(X_t^{(n)})\, dt, \end{cases}$$

where $W^{(n)}$ is a d-dimensional Brownian motion. On the other hand, by the proofs of Theorems 2.1 and 2.3, (X, K) under $\mathbb{Q}$ is a diffusion with generalized inert drift, and satisfies

$$\begin{cases} dX_t = \sigma(X_t)\, dW_t + \mathbf{b}(X_t)\, dt + \frac{1}{2}A\nabla \log \rho(X_t)dt + (A\mathbf{n})(X_t)\, dL_t + K_t\, dt, \\ dK_t = \Gamma\, \mathbf{n}(X_t)\, dL_t + \Gamma\nabla \log \rho(X)t)dt, \end{cases}$$

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

$$c_n \mathbf{1}_D(x)\rho(x)e^{-V_n(x)-(\Gamma^{-1}y,y)}\,dx\,dy \tag{5.7}$$

and $c\mathbf{1}_D(x)\rho(x)e^{-(\Gamma^{-1}y,y)}\,dx\,dy$, respectively, where $c_n > 0$ and $c > 0$ are normalizing constants chosen to make sure that the corresponding measures are probability measures. We may assume, without loss of generality, that $(X_0^{(n)}, K_0^{(n)})$ converges to (X_0, K_0) almost surely. Since $e^{(\log\rho - V_n)/2} \in C(D)\cap W_0^{1,2}(D)$, we know from Theorem 4.3 that $c_n\mathbf{1}_D(x)\rho(x)e^{-V_n(x)-(\Gamma^{-1}y,y)}\,dx\,dy$ is a stationary measure for $(X^{(n)}, K^{(n)})$ under $\mathbb{Q}_n$.

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) in (5.2) under $\mathbb{Q}$.*

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

Observe that $A(X^{(n)})$ converges to $A(X)$ uniformly on $[0, 1]$ and

$$K_t^{(n)} = K_0^{(n)} + \Gamma\int_0^t A^{-1}(X_s^{(n)})dH_s^{(n)} + \Gamma\int_0^t \nabla\log\rho(X_s^{(n)})ds.$$

Let $|H^{(n)}|_t$ and $|K^{(n)}|_t$ denote the total variation process of $H^{(n)}$ and $K^{(n)}$ over the interval $[0, t]$, respectively. Then there is a constant $c > 0$ independent of $n \geq 1$ such that

$$|K^{(n)}|_t \leq |K_0^{(n)}| + |H^{(n)}|_t + \int_0^t |\Gamma\nabla\log\rho(X_s^{(n)})|ds \quad \text{for every } t \geq 0.$$

On the other hand, we know from the proof of Theorem 6.1 in [29, pp. 58–59] that

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

Hence by Theorem 2.2 of [23], $K^{(n)}$ converges to K in probability with respect to the uniform topology on $C[0, 1]$. This yields, by [23, Theorem 2.2] again, that $\int_0^1 \sigma^{-1}(X_s^{(n)})K_s^{(n)}\,dB_s^{(n)} \to \int_0^1 \sigma^{-1}(X_s)K_s\,dB_s$ 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 $\lim_{n\to\infty} Z_1^{(n)} = Z_1$, $\mathbb{P}$-a.s.

Let Φ be a continuous function on $(C[0,1])^2$ with $0 \leq \Phi \leq 1$. Since $\Phi(X^{(n)}, K^{(n)}) \to \Phi(X, K)$, $\mathbb{P}$-a.s., and $Z_1^{(n)} \to Z_1$, $\mathbb{P}$-a.s., by Fatou's lemma,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[\Phi(X,K)Z_1] &\leq \liminf_{n\to\infty}\mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)},K^{(n)})Z_1^{(n)}]\\ &\leq \limsup_{n\to\infty}\mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)},K^{(n)})Z_1^{(n)}]\end{aligned} \tag{5.8}$$

and

$$\mathbb{E}_{\mathbb{P}}[(1-\Phi)(X,K)Z_1] \le \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[(1-\Phi)(X^{(n)},K^{(n)})Z_1^{(n)}]. \tag{5.9}$$

Summing (5.8) and (5.9) we obtain $\mathbb{E}_{\mathbb{P}}[Z_1] \le \limsup_{n\to\infty} \mathbb{E}_{\mathbb{P}}[Z_1^{(n)}]$. Note that by the proof of Theorem 4.3, $Z^{(n)}$ is a continuous non-negative $\mathbb{P}$-martingale, while by the proof of Theorem 2.1, M is a continuous $\mathbb{P}$-martingale. Hence, $\mathbb{E}_{\mathbb{P}}[Z_1] = 1 = \mathbb{E}_{\mathbb{P}}[Z_1^{(n)}]$ and, therefore the inequalities in (5.8) and (5.9) are in fact equalities. It follows that

$$\begin{aligned}\lim_{n\to\infty} \mathbb{Q}_n[\Phi(X^{(n)},K^{(n)})] &= \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}[\Phi(X^{(n)},K^{(n)})Z_1^{(n)}] = \mathbb{E}_{\mathbb{P}}[\Phi(X,K)Z_1]\\ &= \mathbb{Q}[\Phi(X,K)].\end{aligned}$$

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

Theorem 5.2 *Consider the SDE (5.2) on a bounded Lipschitz domain $D \subset \mathbb{R}^d$. A stationary distribution for the solution (X,K) to (5.2) is*

$$\pi(dx,dy) = c_1 \mathbf{1}_D(x)\rho(x)e^{-(\Gamma^{-1}y,y)}\,dx\,dy,$$

where c_1 is the normalizing constant so that $\pi(\overline{D}\times\mathbb{R}^d)=1$.

Proof Recall the notation from previous proofs in this section. We start $(X^{(n)},K^{(n)})$ with the stationary distribution π_n of Theorem 4.3, namely,

$$\mathbb{Q}_n((X_0^{(n)},K_0^{(n)})\in A) = \pi_n(A) = c_n \int_A e^{-V_n(x)}\rho(x)e^{-(\Gamma^{-1}y,y)}\,dx\,dy, \quad A\subset \overline{D}\times\mathbb{R}^d,$$

where c_n is a normalizing constant. Note that here we have $e^{-V_n(x)}\rho(x)$ in place of $e^{-V(x)}$ in Theorem 4.1. Clearly π_n converge weakly on $\overline{D}\times\mathbb{R}^d$ to the probability π. 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 $t_0\le 1$,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[f(X_{t_0},K_{t_0})\right] &= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}\left[f(X_{t_0}^{(n)},K_{t_0}^{(n)})\right] = \lim_{n\to\infty} \int_{\overline{D}\times\mathbb{R}^d} f(x,y)\pi_n(dx,dy)\\ &= \int_{\overline{D}\times\mathbb{R}^d} f(x,y)\pi(dx,dy).\end{aligned}$$

This shows that π is a stationary distribution for the solution to (5.2). □

Taking $\rho\equiv 1$, Theorem 5.2 shows that $\pi(dx,dy) = c_1\mathbf{1}_D(x)e^{-(\Gamma^{-1}y,y)}\,dx\,dy$, where c_1 is the normalizing constant so that $\pi(\overline{D}\times\mathbb{R}^d)=1$, is a stationary distribution for the solution to (5.3).

6 Irreducibility

In this section, D is a bounded C^2 domain in $\mathbb{R}^d$. We will prove uniqueness of the stationary distribution for (X, K) under assumptions stronger than those in previous sections. 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 μ on $\overline{D} \times \mathbb{R}^d$ and $t_0 > 0$ such that if $\mu(A) > 0$, then for all $(x, y) \in \overline{D} \times \mathbb{R}^d$, $\mathbb{Q}_{x,y}((X_{t_0}, K_{t_0}) \in A) > 0$.

Let Γ be a positive definite constant symmetric $d \times d$-matrix.

Theorem 6.1 *Assume that σ is the identity matrix (consequently $\mathbf{u} \equiv \mathbf{n}$) and $\mathbf{b} = 0$. Then the solution (X, K) to (5.3) 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 $K_t = y_0 + \int_0^t \mathbf{n}(X_s)\, dL_s$. 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 $\mathbb{P}_{x,y}$, with respect to $2d$-dimensional Lebesgue measure.

In order to do so, our main ingredient will be that for $t > 0$ the law of (X_t, K_t) has a component that has a strictly positive density with respect to Lebesgue measure on some open set. The idea behind the proof of this result is that one can find d small balls $\{B_j\}_{j=1}^d$ on the boundary ∂D such that

(i) The normal vector $\mathbf{n}(x)$ is 'almost' parallel to the jth unit vector e_j for $x \in B_j$.
(ii) With positive probability, the process X_s has visited all of the B_j's in chronological order before time t, but has not hit the boundary ∂D anywhere outside of the B_j's.
(iii) The amounts S_j of local time that X has accumulated on the B_j's are 'almost' independent.

This suggests that at time t, the law of K_t has a positive component which 'almost' looks like the law of a random vector with independent components, each of them having a density with respect to Lebesgue measure. It follows that the law of K_t has a component which has a density with respect to d-dimensional Lebesgue measure. Since, as long as X is in the interior of the domain D, it is just Brownian motion with drift and K remains constant, we conclude that the law of (X_t, K_t) has a density with respect to $2d$-dimensional Lebesgue measure.

The detailed proof is broken into several distinct steps. In the first step, we use a support theorem and a controllability argument to show that given any point $z_1 \in D$, $\varepsilon > 0$ and any final time t_0, the process (X_t, K_t) has a positive probability to be in an ε-neighborhood of $(z_1, 0)$ after time $t_0/2$, whatever its initial condition. In the second step, we present a review of excursion theory and show how the path of a reflecting Brownian motion can be decomposed into a collection of excursions and how reflecting Brownian motion up to the first hitting time of some subset $U \subset D$ can be constructed from the excursions of a reflecting Brownian motion conditioned never to hit U by adding a 'last excursion' after a suitably chosen amount of local time spent at the boundary. This construction is then used in the third step to 'stitch together' a reflecting Brownian motion from d independent reflecting Brownian motions Y_t^j. In the final step, we show how to condition each of the Y^j's on hitting the boundary ∂D only in B_j and deduce from this that the local time K_t has a density with respect to Lebesgue measure. We conclude by showing how to combine these results to obtain the desired Harris irreducibility.

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 $\ell(0) = 0$, $g(0) = f(0)$, $\int_0^T \mathbf{1}_D(g(s))\, d\ell_s = 0$, and $g(t) = f(t) + \int_0^t \mathbf{n}(g(s))\, d\ell_s$. It has been proved in [25] 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 [25], 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

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

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., [26] for the foundations of excursion theory in abstract settings and [8] for the special case of excursions of Brownian motion. See also [20] for excursions of reflecting Brownian motion on C^3 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 $(L_t^*, \mathbf{H}_x)$ consisting of a positive continuous additive functional L_t^* and a family of excursion laws $\{\mathbf{H}_x\}_{x\in\partial D}$. 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 $\mathbf{H}_x$ 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, $\mathbf{H}_x$ 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 $\mathbf{H}_x$ 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 $t \geq \zeta$. So $e_t \equiv \Delta$ if $\inf\{r > t : X_r \in \partial D\} = t$. Let $\mathcal{E}_u = \{e_s : s \leq u\}$.

Let $\sigma_t = \inf\{s \geq 0 : L_s^* \geq t\}$ and let I be the set of left endpoints of all connected open components of $(0,\infty)\backslash\{t \geq 0 : X_t \in \partial D\}$. The following is a special case of the exit system formula of [26]. For every $x \in \overline{D}$,

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

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 $\mathbf{H}_x(f) = \int_{\mathcal{C}} f\,d\mathbf{H}_x$.

The normalization of the exit system is somewhat arbitrary, for example, if $(L_t^*, \mathbf{H}_x)$ is an exit system and $c \in (0,\infty)$ is a constant then $(cL_t^*, (1/c)\mathbf{H}_x)$ is also an exit system. Let $\mathbb{P}_y^D$ 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 $\mathbf{H}_x$'s to be standard excursion laws normalized so that for some constant $c_1 \in (0,\infty)$,

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

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 $(L_t, \mathbf{H}_x)$ 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 positive 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 [25] that for any starting point in $D \backslash U$, 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 $(L'_t, \mathbf{H}'_{x,\ell})$ for (X'_t, L'_t) from the closed set $\overline{D} \times [0, \infty)$. We will construct this exit system on the basis of $(L_t, \mathbf{H}_x)$ 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 $\mathbf{H}'_{x,\ell}$ 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 $\mathbf{H}'_{x,\ell} = \mathbf{H}_x$. Let $\widehat{\mathbb{P}}_y^D$ denote the distribution of Brownian motion starting from $y \in D \backslash U$, conditioned to hit ∂D before hitting U, and killed upon exiting D. For $\ell < 1$, we have

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

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

$$\mathbf{H}'_{x,\ell}(A) = \mathbf{H}_x(A \backslash A_*). \tag{6.4}$$

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 $x \in \partial D$,

$$c_3 < \mathbf{H}_x(A_*) < c_4. \tag{6.5}$$

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

$$\mu([s_1, s_2] \times F) = \int_{1 \wedge s_1}^{1 \wedge s_2} \mathbf{H}_{X'_{\sigma'_t}}(F \cap A_*)\, dt. \tag{6.6}$$

Since $\mu([0, \infty) \times \mathcal{C}) < \infty$, the Poisson point process $\mathcal{E}$ may be empty; that is, if the Poisson process is viewed as a random measure, then the support of that measure 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

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

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.

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 $j \ne k$, 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 starting with the uniform distribution on U_j, 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 $x, y \in U_{j+1}$, $j = 1, \dots, d$,

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

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 $c_6 > 0$,

$$\mathbb{P}\left(Z_1^j = Y_0^{j+1} \text{ and } Z_t^j \notin \partial D \text{ for } t \in [0, 1] \,\middle|\, \{Y^k\}_{1 \le k \le j}, \{Z^k\}_{1 \le k \le j-1}\right) \ge c_6. \tag{6.7}$$

The process Z^j does not depend otherwise on $\{Y^k\}_{1 \le k \le d}$ and $\{Z^k\}_{k \ne j}$. 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

$$F_j = \left\{ Z_1^j = Y_0^{j+1} \text{ and } Z_t^j \notin \partial D \text{ for } t \in [0,1] \right\}.$$

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

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

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

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

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

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

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

$$X_t^* = Z_{t-T_1-1-T_2-1-\cdots-T_d}^d$$

for $t \ge T_1 + 1 + T_2 + 1 + \cdots + T_d$.

By construction, X^* is a reflecting Brownian motion in D starting from x_1. Note that in view of (6.7), conditional on $\{X_t^j, t \ge 0\}$, $j = 1, \ldots, d$, 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 $j = 1, \ldots, d$. *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 $y_1, \ldots, y_d \in \partial D$ such that the $\mathbf{n}(y_j)$'s point in d orthogonal directions. Let $C_j = \{\mathbf{z} \in \mathbb{R}^d : \angle(\mathbf{n}(y_j), \mathbf{z}) \le \delta_0\}$, for some $\delta_0 > 0$ so small that for any $\mathbf{z}_j \in C_j$, $j = 1, \ldots, d$, the vectors $\{\mathbf{z}_j\}$ are linearly independent. Let $\delta_1 > 0$ be so small that for every $j = 1, \ldots, d$, 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 $\sigma_t^j = \inf\{s \ge 0 : L_s^j \ge t\}$. It is easy to see that there exists $p_2 > 0$ such that with probability greater than p_2, for every $j = 1, \ldots, d$, we have $X_t^j \notin \partial D \setminus \mathcal{B}(y_j, \delta_1)$, for $t \in [0, \sigma_1^j]$. Let

$$R_j = \sup\{t < T_j : Y_t^j \in \partial D\} \quad \text{and} \quad S_j = L_{R_j}^j.$$

Let F_* be the event that for every $j = 1, \dots, d$, $X_t^j \notin \partial D \backslash \mathcal{B}(y_j, \delta_1)$ for $t \in [0, \sigma_1^j]$ and $S_j < \sigma_1^j$. Then (6.5) shows that $\mathbb{P}_{x_1}(F_*) \ge p_2(1 - e^{-c_3})^d$.

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

$$\begin{aligned} &\Lambda([a_1, b_1], [a_2, b_2], \dots, [a_d, b_d]) \\ &\quad = \{K_{t_1}^1 + K_{t_2}^2 + \dots + K_{t_d}^d : t_k \in [a_k, b_k] \text{ for } 1 \le k \le d\}. \end{aligned}$$

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 $c_7 \prod_{1 \le k \le d}(b_k - a_k)$, and bounded above by $c_8 \prod_{1 \le k \le d}(b_k - a_k)$.

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

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

By (6.5) and (6.6), conditional on $\mathcal{G}$, the random variables $S_j = L^j_{R_j}$, $j = 1, \dots, d$, 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 $K_{S_1}^1 + \dots + K_{S_d}^d$ 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 $K_{S_1}^1 + \dots + K_{S_d}^d$ has a component with a density with respect to d-dimensional Lebesgue measure that is bounded below on U_*.

Define $K_t^* = \Gamma \int_0^t \mathbf{n}(X^* s) d\widetilde{L}_s$ and $T_* = \sum_{j=1}^d T_j$, where $\widetilde{L}$ is the boundary local time for reflecting Brownian motion X^*. 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 component with 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 component with 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 component with a strictly positive 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 $\mathbb{Q}_{x_0, y_0}$, which, by the proof of Theorem 2.1, is the reflecting diffusion with inert drift. This proves the theorem. □

Theorem 6.2 *Consider the SDE (5.3) with σ being the identity matrix,* $\mathbf{b} = 0$, *and* $\mathbf{u} \equiv \mathbf{n}$. *The probability distribution $\pi(dx, dy)$ defined by*

$$\pi(A) = \int_A c_1 \mathbf{1}_D(x) e^{-(\Gamma^{-1}y,y)}\, dx\, dy, \quad A \subset \overline{D} \times \mathbb{R}^d,$$

is the only invariant measure for the solution (X, K) *to* (5.3).

Proof In view of Theorem 5.2, 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 [32]. 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 $\mu \perp \nu$ by assumption, 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 the measure ψ is non-zero. □

References

1. Bass, R.F.: Diffusions and Elliptic Operators. Springer, Berlin (1997)
2. Bass, R.F., Burdzy, K., Chen, Z.-Q.: Uniqueness for reflecting Brownian motion in lip domains. Ann. Inst. Henri Poincare Probab. Statist. **41**, 197–235 (2005)
3. Bass, R.F., Burdzy, K., Chen, Z.-Q.: Pathwise uniqueness for a degenerate stochastic differential equation. Ann. Probab. **35**, 2385–2418 (2007)
4. Benaïm, M., Ledoux, M., Raimond, O.: Self-interacting diffusions. Probab. Theory Related Fields **122**, 1–41 (2002)
5. Benaïm, M., Raimond, O.: Self-interacting diffusions. II. Convergence in law. Ann. Inst. H. Poincaré Probab. Statist. **39**, 1043–1055 (2003)
6. Benaïm, M., Raimond, O.: Self-interacting diffusions. III. Symmetric interactions. Ann. Probab. **33**, 1717–1759 (2005)
7. Bou-Rabee, N., Owhadi, H.: Ergodicity of Langevin processes with degenerate diffusion in momentums. Int. J. Pure Appl. Math. **45**(3), 475–490 (2008)
8. Burdzy, K.: Multidimensional Brownian Excursions and Potential Theory. Longman Sci. Tech., Harlow (1987)
9. Burdzy, K., Hołyst, R., Pruski, Ł.: Brownian motion with inert drift, but without flux: a model. Physica A **384**, 278–284 (2007)
10. Burdzy, K., White, D.: A Gaussian oscillator. Electron. Comm. Probab. **9**, 92–95 (2004)
11. Chen, Z.-Q.: On reflecting diffusion processes and Skorokhod decompositions. Probab. Theory Related Fields **94**, 281–315 (1993)
12. Chen, Z.-Q., Fitzsimmons, P.J., Takeda, M., Ying, J., Zhang, T.-S.: Absolute continuity of symmetric Markov processes. Ann. Probab. **32**, 2067–2098 (2004)
13. Da Prato, G., Zabczyk, J.: Ergodicity for Infinite-Dimensional Systems. Cambridge University Press, Cambridge (1996)
14. Dupuis, P., Ishii, H.: SDEs with oblique reflection on nonsmooth domains. Ann. Probab. **21**, 554–580 (1993)
15. Dupuis, P., Ishii, H.: Correction: SDEs with oblique reflection on nonsmooth domains. Ann. Probab. **36**, 1992–1997 (2008)
16. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
17. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. De Gruyter, New York (1994)
18. Hairer, M., Mattingly, J.C.: Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. Ann. Math. (2) **164**, 993–1032 (2006)
19. Hörmander, L.: The Analysis of Linear Partial Differential Operators. Springer, Berlin (1985)
20. Hsu, P.: On excursions of reflecting Brownian motion. Trans. Am. Math. Soc. **296**, 239–264 (1986)
21. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. Kodansha, Tokyo (1981)

22. Knight, F.: On the path of an inert object impinged on one side by a Brownian particle. Probab. Theory Related Fields **121**, 577–598 (2001)
23. Kurtz, T., Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Probab. **19**, 1035–1070 (1991)
24. Liggett, T.M.: Interacting Particle Systems. Springer, New York (1985)
25. Lions, P.L., Sznitman, A.S.: Stochastic differential equations with reflecting boundary conditions. Comm. Pure Appl. Math. **37**, 511–537 (1984)
26. Maisonneuve, B.: Exit systems. Ann. Probab. **3**(3), 399–411 (1975)
27. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, London (1993)
28. Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Berlin (1995)
29. Pardoux, E., Williams, R.J.: Symmetric reflected diffusions. Ann. Inst. H. Poincaré Probab. Statist. **30**, 13–62 (1994)
30. Pemantle, R.: A survey of random processes with reinforcement. Probab. Surv. **4**, 1–79 (2007)
31. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)
32. Sinai, Y.: Topics in Ergodic Theory. Princeton University Press, Princeton (1994)
33. Stroock, D.W., Varadhan, S.R.S.: 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, pp. 333–359. University of California Press, Berkeley (1972)
34. White, D.: Processes with Inert Drift. Ph.D. Thesis, University of Washington (2005)
35. White, D.: Processes with inert drift. Electr. J. Probab. **12**, 1509–1546 (2007)
36. Williams, R.J., Zheng, W.: On reflecting Brownian motion—a weak convergence approach. Ann. Inst. H. Poincaré Probab. Statist. **26**, 461–488 (1990)