# Brownian Motion on Some Spaces with Varying Dimension*

**Zhen-Qing Chen** and **Shuwen Lou**

October 4, 2017

**Abstract**

In this paper we introduce and study Brownian motion on a class of state spaces with varying dimension. Starting with a concrete case of such state spaces that models a big square with a flag pole, we construct a Brownian motion on it and study how heat propagates on such a space. We derive sharp two-sided global estimates on its transition density functions (also called heat kernel). These two-sided estimates are of Gaussian type, but the measure on the underlying state space does not satisfy volume doubling property. Parabolic Harnack inequality fails for such a process. Nevertheless, we show Hölder regularity holds for its parabolic functions. We also derive the Green function estimates for this process on bounded smooth domains. Brownian motion on some other state spaces with varying dimension are also constructed and studied in this paper.

**AMS 2010 Mathematics Subject Classification**: Primary 60J60, 60J35; Secondary 31C25, 60H30, 60J45

**Keywords and phrases**: Space of varying dimension, Brownian motion, Laplacian, transition density function, heat kernel estimates, Hölder regularity, Green function

## 1 Introduction

Brownian motion takes a central place in modern probability theory and its applications, and is a basic building block for modeling many random phenomena. Brownian motion has intimate connections to analysis since its infinitesimal generator is Laplacian operator. Brownian motion in Euclidean spaces has been studied by many authors in depth. Brownian motion on manifolds and on fractals has also been investigated vigorously, and is shown to have intrinsic interplay with the geometry of the underlying spaces. See [24, 27, 28, 29] and the references therein. In most of these studies, the underlying metric measure spaces are assumed to satisfy volume doubling (VD) property. For Brownian motion on manifolds with walk dimension 2, a remarkable fundamental result obtained independently by Grigor'yan [20] and Saloff-Coste [30] asserts that the following are equivalent: (i) two-sided Aronson type Gaussian bounds for heat kernel, (ii) parabolic Harnack equality, and (iii) VD and Poincaré inequality. This result is then extended to strongly local Dirichlet forms on metric measure space in [9, 31, 32] and to graphs in [17]. For Brownian motion on fractals with walk dimension larger than 2, the above equivalence still

*Research partially supported by NSF Grant DMS-1206276.

holds but one needs to replace (iii) with (iii') VD, Poincaré inequality and a cut-off Sobolev inequality; see [3, 4, 1].

Recently, analysis on non-smooth spaces has attracted lots of interest. In real world, there are many objects having varying dimension. It is natural to study Brownian motion and "Laplacian operator" on such spaces. A simple example of spaces with varying dimension is a large square with a thin flag pole. Mathematically, it is modeled by a plane with a vertical line installed on it:

$$\mathbb{R}^2 \cup \mathbb{R}_+ := \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0 \text{ or } x_1 = x_2 = 0 \text{ and } x_3 > 0\right\}. \tag{1.1}$$

Here and in the sequel, we use $:=$ as a way of definition and denote $[0, \infty)$ by $\mathbb{R}_+$. Spaces with varying dimension arise in many disciplines including statistics, physics and engineering (e.g. molecular dynamics, plasma dynamics). See, for example, [26, 34] and the references therein.

The goal of this paper is to construct and study Brownian motion and Laplacian on spaces of varying dimension, in particular, to investigate how heat propagates on such spaces. Intuitively, Brownian motion on space $\mathbb{R}^2 \cup \mathbb{R}$ of (1.1) should behave like a two-dimensional Brownian motion when it is on the plane, and like a one-dimensional Brownian motion when it is on the vertical line (flag pole). However the space $\mathbb{R}^2 \cup \mathbb{R}$ is quite singular in the sense that the base $O$ of the flag pole where the plane and the vertical line meet is a singleton. A singleton would never be visited by a two-dimensional Brownian motion, which means Brownian motion starting from a point on the plane will never visit $O$. Hence there is no chance for such a process to climb up the flag pole. To circumvent this difficulty, we collapse or short (imagine putting an infinite conductance on) a small closed disk $\overline{B(0, \varepsilon)} \subset \mathbb{R}^2$ centered at the origin into a point $a^*$ and consider the resulting Brownian motion with darning on the collapsed plane, for which $a^*$ will be visited. The notion of Brownian motion with darning is coined in [12] and its potential theory has been studied in details in [11] and [13, Sections 2-3]. Through $a^*$ we put a vertical pole and construct Brownian motion with varying dimension on $\mathbb{R}^2 \cup \mathbb{R}_+$ by joining together the Brownian motion with darning on the plane and the one-dimensional Brownian motion along the pole. It is possible to construct this process rigorously via Poisson point process of excursions. But we find that the most direct way to construct BMVD is by using a Dirichlet form approach, which will be carried out in Section 2.

To be more precise, the state space of BMVD on $E$ is defined as follows. Fix $\varepsilon > 0$ and $p > 0$. Denote by $B_\varepsilon$ the closed disk on $\mathbb{R}^2$ centered at $(0, 0)$ with radius $\varepsilon$. Let $D_\varepsilon := \mathbb{R}^2 \setminus B_\varepsilon$. By identifying $B_\varepsilon$ with a singleton denoted by $a^*$, we can introduce a topological space $E := D_\varepsilon \cup \{a^*\} \cup \mathbb{R}_+$, with the origin of $\mathbb{R}_+$ identified with $a^*$ and a neighborhood of $a^*$ defined as $\{a^*\} \cup (V_1 \cap \mathbb{R}_+) \cup (V_2 \cap D_\varepsilon)$ for some neighborhood $V_1$ of 0 in $\mathbb{R}^1$ and $V_2$ of $B_\varepsilon$ in $\mathbb{R}^2$. Let $m_p$ be the measure on $E$ whose restriction on $\mathbb{R}_+$ and $D_\varepsilon$ is the Lebesgue measure multiplied by $p$ and 1, respectively. In particular, we have $m_p(\{a^*\}) = 0$. Note that the measure $m_p$ also depends on $\varepsilon$, the radius of the hole $B_\varepsilon$.

**Definition 1.1.** Let $\varepsilon > 0$ and $p > 0$. A Brownian motion with varying dimensions (BMVD in abbreviation) with parameters $(\varepsilon, p)$ on $E$ is an $m_p$-symmetric diffusion $X$ on $E$ such that

(i) its part process in $\mathbb{R}_+ \setminus \{a^*\}$ or $D_\varepsilon$ has the same law as standard Brownian motion killed upon leaving $\mathbb{R}_+ \setminus \{0\}$ or $D_\varepsilon$, respectively;

(ii) it admits no killings on $a^*$.

It follows from the $m_p$-symmetry of $X$ and the fact $m_p(\{a^*\}) = 0$ that BMVD $X$ spends zero Lebesgue amount of time at $a^*$.

Recall that a Markov process on $E$ is said to be a Feller process if its transition semigroup $\}P_t; t \geq 0\}$ is strongly continuous in $(C_\infty(E), \|\cdot\|_\infty)$, where $C_c(E)$ is the space of continuous functions on $E$ that vanishes at infinity and $\|f\|_\infty := \sup_{x\in E}|f(x)|$. We say a Markov process on $E$ has strong Feller property if for every $t > 0$, $P_t$ maps a bounded measurable function on $E$ to a bounded continuous function on $E$. The first part of the following theorem will be established in Section 2, while its second part is a consequence of Theorem 1.3.

**Theorem 1.2.** (i) *For every $\varepsilon > 0$ and $p > 0$, BMVD with parameters $(\varepsilon, p)$ exists and is unique in law.*

(ii) *The BMVD is a Feller process having strong Feller property.*

We point out that BMVD on $E$ can start from every point in $E$. We further characterize the $L^2$-infinitesimal generator $\mathcal{L}$ of BMVD $X$ in Section 2, which can be viewed as the Laplacian operator on this singular space. We show that $u \in L^2(E; m_p)$ is in the domain of the generator $\mathcal{L}$ if and only if $\Delta u$ exists as an $L^2$-integrable function in the distributional sense when restricted to $D_\varepsilon$ and $\mathbb{R}_+$, and $u$ satisfies zero-flux condition at $a^*$; see Theorem 2.3 for details. It ican be shown (see Remark 3.3 below) that BMVD $X$ has a transition density function $p(t, x, y)$ with respect to the measure $m_p$ that is continuous in $y$ for every $t > 0$ and $y \in E$. In fact, it follows from Theorem 6.3 that $p(t, x, y)$ is (locally) jointly Hölder continuous in $(t, x, y)$. The kernel $p(t, x, y)$ is also called the fundamental solution (or heat kernel) for $\mathcal{L}$. Note that $p(t, x, y)$ is symmetric in $x$ and $y$. The main purpose of this paper is to investigate how the BMVD $X$ propagates in $E$; that is, starting from $x \in E$, how likely $X_t$ travels to position $y \in E$ at time $t$. This amounts to study the properties of $p(t, x, y)$ of $X$. In this paper, we will establish the following sharp two-sided estimates on $p(t, x, y)$ in Theorem 1.3 and Theorem 1.4. To state the results, we need first to introduce some notations. Throughout this paper, we will denote the geodesic metric on $E$ by $\rho$. Namely, for $x, y \in E$, $\rho(x, y)$ is the shortest path distance (induced from the Euclidean space) in $E$ between $x$ and $y$. For notational simplicity, we write $|x|_\rho$ for $\rho(x, a^*)$. We use $|\cdot|$ to denote the usual Euclidean norm. For example, for $x, y \in D_\varepsilon$, $|x - y|$ is the Euclidean distance between $x$ and $y$ in $\mathbb{R}^2$. Note that for $x \in D_\varepsilon$, $|x|_\rho = |x| - \varepsilon$. Apparently,

$$\rho(x, y) = |x - y| \wedge (|x|_\rho + |y|_\rho) \quad \text{for } x, y \in D_\varepsilon \tag{1.2}$$

and $\rho(x, y) = |x| + |y| - \varepsilon$ when $x \in \mathbb{R}_+$ and $y \in D_\varepsilon$ or vice versa. Here and in the rest of this paper, for $a, b \in \mathbb{R}$, $a \wedge b := \min\{a, b\}$.

We now fix $\varepsilon > 0$ and $p > 0$. The following are the main results of this paper.

**Theorem 1.3.** *Let $T > 0$ be fixed. There exist positive constants $C_i$, $1 \leq i \leq 14$ so that the transition density $p(t, x, y)$ of BMVD satisfies the following estimates when $t \in (0, T]$:*

(i) *For $x \in \mathbb{R}_+$ and $y \in E$,*

$$\frac{C_1}{\sqrt{t}} e^{-C_2 \rho(x,y)^2/t} \leq p(t, x, y) \leq \frac{C_3}{\sqrt{t}} e^{-C_4 \rho(x,y)^2/t}. \tag{1.3}$$

(ii) *For* $x, y \in D_\varepsilon \cup \{a^*\}$, *when* $|x|_\rho + |y|_\rho < 1$,

$$\begin{aligned} & \frac{C_5}{\sqrt{t}} e^{-C_6 \rho(x,y)^2/t} + \frac{C_5}{t}\left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-C_7|x-y|^2/t} \\ \leq \quad & p(t,x,y) \leq \frac{C_8}{\sqrt{t}} e^{-C_9 \rho(x,y)^2/t} + \frac{C_8}{t}\left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-C_{10}|x-y|^2/t}; \end{aligned} \tag{1.4}$$

*and when* $|x|_\rho + |y|_\rho \geq 1$,

$$\frac{C_{11}}{t} e^{-C_{12}\rho(x,y)^2/t} \leq p(t,x,y) \leq \frac{C_{13}}{t} e^{-C_{14}\rho(x,y)^2/t}. \tag{1.5}$$

Since $p(t,x,y)$ is symmetric in $(x,y)$, the above two cases cover all the cases for $x, y \in E$. Theorem 1.3 shows that the transition density function $p(t,x,y)$ of BMVD on $E$ has one-dimensional character when at least one of $x, y$ is in the pole (i.e. in $\mathbb{R}_+$); it has two-dimensional character when both points are on the plane and at least one of them is away from the pole base $a^*$. When both $x$ and $y$ are in the plane and both are close to $a^*$, $p(t,x,y)$ exhibits a mixture of one-dimensional and two-dimensional characters; see (1.4). Theorem 1.3 will be proved through Theorems 4.6-4.8.

The large time heat kernel estimates for BMVD are given by the next theorem, which are very different from the small time estimates.

**Theorem 1.4.** *There exist positive constants* $C_i$, $15 \leq i \leq 29$, *so that the transition density* $p(t,x,y)$ *of BMVD satisfies the following estimates for* $t \geq 8$.

(i) *For* $x, y \in D_\varepsilon \cup \{a^*\}$,

$$\frac{C_{15}}{t} e^{-C_{16}\rho(x,y)^2/t} \leq p(t,x,y) \leq \frac{C_{17}}{t} e^{-C_{18}\rho(x,y)^2/t}.$$

(ii) *For* $x \in \mathbb{R}_+$, $y \in D_\varepsilon \cup \{a^*\}$,

$$\frac{C_{19}}{t}\left(1 + \frac{|x| \log t}{\sqrt{t}}\right) e^{-C_{20}\rho(x,y)^2/t} \leq p(t,x,y) \leq \frac{C_{21}}{t}\left(1 + \frac{|x| \log t}{\sqrt{t}}\right) e^{-C_{22}\rho(x,y)^2/t}$$

*when* $|y|_\rho \leq 1$, *and*

$$\begin{aligned} & \frac{C_{23}}{t}\left(1 + \frac{|x|}{\sqrt{t}} \log\left(1 + \frac{\sqrt{t}}{|y|_\rho}\right)\right) e^{-C_{24}\rho(x,y)^2/t} \leq p(t,x,y) \\ \leq \; & \frac{C_{25}}{t}\left(1 + \frac{|x|}{\sqrt{t}} \log\left(1 + \frac{\sqrt{t}}{|y|_\rho}\right)\right) e^{-C_{26}\rho(x,y)^2/t} \qquad \text{when } |y|_\rho > 1. \end{aligned}$$

(iii) *For* $x, y \in \mathbb{R}_+$,

$$\begin{aligned} & \frac{C_{27}}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-C_{28}|x-y|^2/t} + \frac{C_{27}}{t}\left(1 + \frac{(|x|+|y|)\log t}{\sqrt{t}}\right) e^{-2(x^2+y^2)/t} \\ \leq \quad & p(t,x,y) \\ \leq \quad & \frac{C_{29}}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-C_{30}|x-y|^2/t} + \frac{C_{29}}{t}\left(1 + \frac{(|x|+|y|)\log t}{\sqrt{t}}\right) e^{-(x^2+y^2)/2t}. \end{aligned}$$

Theorem 1.4 will be proved through Theorems 5.14, 5.15 and 5.17. Note that in Theorems 1.3 and 1.4, the constants $C_k$, $1 \leq k \leq 29$, may depend on $\varepsilon$ and $p$. It is an interesting but delicate problem to study their quantative dependence on $(\varepsilon, p)$. Clearly, as $\varepsilon \to 0$, BMVD starting from a point in $D_\varepsilon$ will take longer and longer time to reach the base point $a^*$ and hence to climb up the pole. In the limit, the disk $B_\varepsilon$ shrinks to a singleton $(0,0) \in \mathbb{R}^2$ and the process starts from the plane is just Brownian motion on the plane $\mathbb{R}^2$ that will never visit the point $(0,0)$, while the process starts from the pole is just the reflected Brownian motion on $[0, \infty)$. We do not pursue this problem in this paper. See, however, Proposition 4.3 and the paragraph following it on the effect of the parameters $(\varepsilon, p)$ on the behavior of BMVD $X$. See also Remark 4.5 for an open problem on using PDE approach.

Due to the singular nature of the space, the standard Nash inequality and Davies method for obtaining heat kernel upper bound do not give sharp bound for our BMVD. We can not employ either the methods in [20, 30, 9, 31, 32] on obtaining heat kernel estimates through volume doubling and Poincaré inequality or the approach through parabolic Harnack inequality. In fact, $(E, m_p)$ does not have volume doubling property, and we will show the parabolic Harnack inequality fails for BMVD $X$; see Proposition 2.1 and Remark 4.9(iii). Hence a new approach is needed to study the heat kernel of BMVD. A key role is played by the "signed radial process" of BMVD, which we can analyze and derive its two-sided heat kernel estimates. From it, by exploring the rotational symmetry of BMVD, we can obtain short time sharp two-sided heat kernel estimates by the following observation. The sample paths of BMVD starting at $x$ reach $y$ at time $t$ in two possible ways: with or without passing through $a^*$. The probability of the first scenario is given exactly by the probability transition density function of killed Brownian motion in $D_\varepsilon$. The probability of the second scenario can be computed by employing the strong Markov property of BMVD at the first hitting time of the pole base $a^*$ and reducing it to the signed radial process of BMVD, exploring the symmetry of BMVD starting from $a^*$. The large time heat kernel estimates are more delicate. For large time estimate, the key is to obtain the correct on-diagonal estimate. This is done through some delicate analysis of BMVD and Bessel process on the plane. As a corollary of the sharp two-sided heat kernel estimates, we find that the usual form of the parabolic Harnack inequality fails for parabolic functions of BMVD. Nevertheless, we will show in Section 6 that joint Hölder regularity holds for bounded parabolic functions of $X$.

Let $X^D$ be the part process of BMVD killed upon exiting a bounded connected $C^{1,1}$ open subset $D$ of $E$. Denote by $p_D(t, x, y)$ its transition density function. Using the Green function estimates and boundary Harnack inequality for absorbing Brownian motion in Euclidean spaces, we can derive sharp two-sided estimates on Green function

$$G_D(x, y) := \int_0^\infty p_D(t, x, y) dt$$

for BMVD in $D$. Recall that an open set $D \subset \mathbb{R}^d$ is called to be $C^{1,1}$ if there exist a localization radius $R_0 > 0$ and a constant $\Lambda_0 > 0$ such that for every $z \in \partial D$, there exists a $C^{1,1}-$function $\phi = \phi_z : \mathbb{R}^{d-1} \to \mathbb{R}$ satisfying $\phi(0) = 0$, $\nabla\phi(0) = (0, \dots, 0)$, $\|\nabla\phi\|_\infty \leq \Lambda_0$, $|\nabla\phi(x) - \nabla\phi(z)| \leq \Lambda|x - z|$ and an orthonormal coordinate system $CS_z$ : $y = (y_1, \dots, y_d) := (\widetilde{y}, y_d)$ with its origin at $z$ such that

$$B(z, R_0) \cap D = \{y \in B(0, R_0) \text{ in } CS_z : y_d > \phi(\widetilde{y})\}.$$

For the state space $E$, an open set $D \subset E$ will be called $C^{1,1}$ in $E$, if $D \cap (\mathbb{R}_+ \setminus \{a^*\})$ is a $C^{1,1}$ open set in $\mathbb{R}_+$, and $D \cap D_\varepsilon$ is a $C^{1,1}$ open set in $\mathbb{R}^2$.

**Theorem 1.5.** *Suppose $D$ is a bounded $C^{1,1}$ domain of $E$ that contains $a^*$. Let $G_D(x,y)$ be the Green function of BMVD $X$ killed upon exiting $D$. Then for $x \neq y$ in $D$, we have*

$$G_D(x,y) \asymp \begin{cases} \delta_D(x) \wedge \delta_D(y), & x, y \in D \cap \mathbb{R}_+; \\ \\ (\delta_D(y) \wedge 1)\,(\delta_D \wedge 1) + \ln\left(1 + \frac{\delta_{D\cap D_\varepsilon}(x)\delta_{D\cap D_\varepsilon}(y)}{|x-y|^2}\right), & x, y \in D \cap D_\varepsilon; \\ \\ \delta_D(x)\delta_D(y), & x \in D \cap \mathbb{R}_+,\ y \in D \cap D_\varepsilon. \end{cases}$$

*Here $\delta_D(x) := \mathrm{dist}_\rho(x, \partial D) := \inf\{\rho(x,z) : z \notin D\}$ and $\delta_{D\cap D_\varepsilon}(x) := \inf\{\rho(x,z) : z \notin D \cap D_\varepsilon\}$.*

For two positive functions $f$ and $g$, $f \asymp g$ means that $f/g$ is bounded between two positive constants. In the following, we will also use notation $f \lesssim g$ (respectively, $f \gtrsim g$) to mean that there is some constant $c > 0$ so that $f \le cg$ (respectively, $f \ge cg$).

The above space $E$ of varying dimension is special. It serves as a toy model for further study on Brownian motion on more general spaces of varying dimension. Another two examples of spaces of varying dimension and BMVD on them are given and studied in Section 8 of this paper. Even for this toy model, several interesting and non-trivial phenomena have arisen. The heat kernel estimates on spaces of varying dimension are quite delicate. They are of Gaussian type but they are not of the classical Aronson Gaussian type. The different dimensionality is also reflected in the heat kernel estimates for BMVD. Even when both points $x$ and $y$ are on the plane, the heat kernel $p(t,x,y)$ exhibits both one-dimensional and two-dimensional characteristics depending on whether both points are close to the base point $a^*$ or not. In addition, both the Euclidean distance $|x-y|$ and the geodesic distance $\rho(x,y)$ between the two points $x$ and $y$ play a role in the kernel estimates. Heat kernel estimates have been studied in [22] for Brownian motion in $\mathbb{R}^2$ penetrating some fractal fields that are of positive capacity with respect to the Brownian motion in the plane. The upper and lower bound heat kernel estimates obtained in [22] do not match as they are of different forms. The spaces of varying dimension considered in this paper is singular in the sense that the original intersection of the line and plane is a point that is polar to Brownian motion on the plane. As far as we know, this is the first paper that is devoted to the detailed study of heat propagation on singular spaces of varying dimension and their related potential theory. Moreover, the lower and upper bound heat kernel estimates obtained in this paper are sharp in the sense that they are of the same form, and they hold for all time $t > 0$. Our approach is mainly probabilistic. For other related work and approaches on Markov processes living on spaces with possibly different dimensions, we refer the reader to [18, 23, 25] and the references therein.

The rest of the paper is organized as follows. Section 2 gives a Dirichlet form construction and characterization of BMVD, as well as its infinitesimal generator. Nash-type inequality for $X$ is given in Section 3. In Section 4, we present small time heat kernel estimates for $X$, while the large time estimates are given in Section 5. Hölder continuity of parabolic functions of $X$ is established in Section 6. Section 7 is devoted to the two-sided sharp estimates for Green function of $X$ in bounded $C^{1,1}$ domains in $E$ that contain the pole base $a^*$. BMVD on a large square with multiple vertical flag poles or with an arch are studied in Sections 8.

For notation convenience, in this paper we set

$$\overline{p}_D(t,x,y) := p(t,x,y) - p_D(t,x,y), \tag{1.6}$$

where $D$ is a domain of $E$ and $p_D(t,x,y)$ is the transition density of the part process killed upon exiting $D$. In other words, for any non-negative function $f \geq 0$ on $E$,

$$\int_E \overline{p}_D(t,x,y) f(y) m_p(dy) = \mathbb{E}_x \left[ f(X_t); t \geq \tau_D \right], \tag{1.7}$$

where $\tau_D := \inf\{t \geq 0 : X_t \notin D\}$. Thus while $p_D(t,x,y)$ gives the probability density that BMVD starting from $x$ hits $y$ at time $t$ without exiting $D$, $\overline{p}_D(t,x,y)$ is the probability density for BMVD starting from $x$ leaves $D$ before ending up at $y$ at time $t$.

We use $C_c^\infty(E)$ to denote the space of continuous functions with compact support in $E$ so that there restriction to $D_\varepsilon$ and $\mathbb{R}_+$ are smooth on $\overline{D_\varepsilon}$ and on $\mathbb{R}_+$, respectively. We also follow the convention that in the statements of the theorems or propositions $C, C_1, \cdots$ denote positive constants, whereas in their proofs $c, c_1, \cdots$ denote positive constants whose exact value is unimportant and may change from line to line.

# 2 Preliminaries

Throughout this paper, we denote the Brownian motion with varying dimension by $X$ and its state space by $E$. In this section, we will construct BMVD using Dirichlet form approach. For the definition and basic properties of Dirichlet forms, including the relationship between Dirichlet form, $L^2$-infinitesimal generator, resolvents and semigroups, we refer the reader to [12] and [19].

For a connected open set $D \subset \mathbb{R}^d$, $W^{1,2}(D)$ is the collection of $L^2(D; dx)$-integrable functions whose first order derivatives (in the sense of distribution) exist and are also in $L^2(D; dx)$. Define

$$\mathcal{E}^0(f,g) = \frac{1}{2} \int_D \nabla f(x) \cdot \nabla g(x) dx, \qquad f, g \in W^{1,2}(D).$$

It is well known that when $D$ is smooth, $(\mathcal{E}^0, W^{1,2}(D))$ is a regular Dirichlet form on $L^2(\overline{D}; dx)$ and its associated Hunt process is the (normally) reflected Brownian motion on $\overline{D}$. Moreover, every function $f$ in $W^{1,2}(D)$ admits a quasi-continuous version on $\overline{D}$, which will still be denoted by $f$. A quasi-continuous function is defined quasi-everywhere (q.e. in abbreviation) on $\overline{D}$. When $d = 1$, by Cauchy-Schwartz inequality, every function in $W^{1,2}(D)$ is 1/2-Hölder on $\overline{D}$. Denote by $W_0^{1,2}(D)$ the $\mathcal{E}_1^0$-closure of $C_c^\infty(D)$, where $\mathcal{E}_1^0(f,f) := \mathcal{E}^0(u,u) + \int_D u(x)^2 dx$. It is known that for any open set $D \subset \mathbb{R}^d$, $(\mathcal{E}^0, W_0^{1,2}(D))$ is a regular Dirichlet form on $L^2(D; dx)$ associated with the absorbing Brownian motion in $D$.

For a subset $A \subset E$, we define $\sigma_A := \inf\{t > 0, X_t \in A\}$ and $\tau_{{}_A} := \inf\{t \geq 0 : X_t \notin A\}$. Similar notations will be used for other stochastic processes. We will use $B_\rho(x,r)$ (resp. $B_e(x,r)$) to denote the open ball in $E$ under the path metric $\rho$ (resp. in $\mathbb{R}_+$ or $\mathbb{R}^2$ under the Euclidean metric) centered at $x \in E$ with radius $r > 0$.

A measure $\mu$ on $E$ is said to have (resp. local) volume doubling property if there exists a constant $C > 0$ so that $\mu(B_\rho(x, 2r)) \leq C\mu(B_\rho(x,r))$ for all $x \in E$ and every $r > 0$ (resp. $r \in (0,1]$).

**Proposition 2.1.** *For any $p > 0$, volume doubling property fails for measure $m_p$.*

*Proof.* Note that for small $r > 0$ and $x_0 \in D_\varepsilon$ with $|x_0|_\rho = r$, $m_p(B_\rho(x_0, r)) = \pi r^2$ while $m_p(B_\rho(x_0, 2r)) = 2\varepsilon r + r^2 + r$. Thus there does not exist any constant $C > 0$ so that $m_p(B_\rho(x, 2r)) \leq C m_p(B_\rho(x,r))$ for all $x \in E$ and every $r \in (0,1]$. □

The following is an extended version of Theorem 1.2(i).

**Theorem 2.2.** *For every $\varepsilon > 0$ and $p > 0$, BMVD $X$ on $E$ with parameter $(\varepsilon, p)$ exists and is unique. Its associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E; m_p)$ is given by*

$$\mathcal{F} = \left\{ f : f|_{D_\varepsilon} \in W^{1,2}(D_\varepsilon),\ f|_{\mathbb{R}_+} \in W^{1,2}(\mathbb{R}_+),\ \text{and } f(x) = f(0) \text{ q.e. on } \partial D_\varepsilon \right\}, \tag{2.1}$$

$$\mathcal{E}(f,g) = \frac{1}{2}\int_{D_\varepsilon} \nabla f(x) \cdot \nabla g(x) dx + \frac{p}{2}\int_{\mathbb{R}_+} f'(x) g'(x) dx. \tag{2.2}$$

*Proof.* Let $\mathcal{E}$ and $\mathcal{F}$ be defined as above. Let $u_1(x) = \mathbb{E}^x\left[e^{-\sigma_{B_\varepsilon}}\right]$ when $x \in D_\varepsilon$ and $u_1(x) = \mathbb{E}^x\left[e^{-\sigma_{D_\varepsilon}}\right]$ when $x \in \mathbb{R}_+$. It is known that $u_1|_{D_\varepsilon} \in W^{1,2}(D_\varepsilon)$, $u_1|_{\mathbb{R}_+} \in W^{1,2}(\mathbb{R}_+)$, $u_1(x) = 1$ for $x \in \partial D_\varepsilon$ and $u_1(0) = 1$. Hence $u_1 \in \mathcal{F}$.

*Existence:* Let

$$\mathcal{F}^0 = \left\{ f : f|_{\mathbb{R}^2} \in W_0^{1,2}(D_\varepsilon),\ f|_{\mathbb{R}_+} \in W_0^{1,2}(\mathbb{R}_+) \right\}.$$

Then $\mathcal{F}$ is the linear span of $\mathcal{F}^0 \cup \{u_1\}$. It is easy to check that $(\mathcal{E}, \mathcal{F})$ is a strongly local regular Dirichlet form on $L^2(E; m)$. So there is an $m_p$-symmetric diffusion process $X$ on $E$ associated with it. Using the Dirichlet form characterization, the part process of $X$ killed upon hitting $a^*$ is an absorbing Brownian motion in $D_\varepsilon$ or $\mathbb{R}_+$, depending on the starting point. So $X$ is a BMVD on $E$. Moreover, $X$ is conservative; that is, it has infinite lifetime.

*Uniqueness:* Conversely, if $X$ is a BMVD, it suffices to check from definition that its associated Dirichlet form $(\mathcal{E}^*, \mathcal{F}^*)$ in $L^2(E; m_p)$ has to be $(\mathcal{E}, \mathcal{F})$ given in (2.1)-(2.2). Indeed, since $a^*$ is non-polar for $X$, for all $u \in \mathcal{F}^*$,

$$H^1_{a^*} u(x) := \mathbb{E}^x\left[e^{-\sigma_{a^*}} u(X_{\sigma_{a^*}})\right] = u(a^*)\mathbb{E}^x[e^{-\sigma_{a^*}}] \in \mathcal{F} \cap \mathcal{F}^*$$

and $u - H^1_{a^*}u(x) \in \mathcal{F}^0$. Thus $\mathcal{F}^* \subset \mathcal{F}$. On the other hand, since the part process of $X$ killed upon hitting $a^*$ has the same distribution as the absorbing Brownian motion on $D_\varepsilon \cup (0, \infty)$, which has Dirichlet form $(\mathcal{E}, \mathcal{F}^0)$ on $L^2(E \setminus \{a^*\}; m_p)$, we have $\mathcal{F}^0 \subset \mathcal{F}^*$. It follows that $\mathcal{F} \subset \mathcal{F}^*$ and therefore $\mathcal{F} = \mathcal{F}^*$. Since $X$ is a diffusion that admits no killings inside $E$, its Dirichlet form $(\mathcal{E}^*, \mathcal{F}^*)$ is strongly local. Let $\mu_{\langle u \rangle} = \mu^c_{\langle u \rangle}$ denote the energy measure associated with $u \in \mathcal{F}^*$; see [12, 19]. Then by the strong locality of $\mu^c_{\langle u \rangle}$ and the $m_p$-symmetry of $X$, we have for every bounded $u \in \mathcal{F}^* = \mathcal{F}$,

$$\begin{aligned}
\mathcal{E}^*(u,u) &= \frac{1}{2}\left(\mu^c_{\langle u \rangle}(E \setminus \{a^*\}) + \mu^c_{\langle u \rangle}(a^*)\right) = \frac{1}{2}\mu^c_{\langle u \rangle}(E \setminus \{a^*\}) = \frac{1}{2}\mu^c_{\langle u \rangle}(D_\varepsilon) + \frac{1}{2}\mu^c_{\langle u \rangle}(\mathbb{R}_+) \\
&= \frac{1}{2}\int_{D_\varepsilon} |u(x)|^2 dx + \frac{p}{2}\int_{\mathbb{R}_+} |u'(x)|^2 dx = \mathcal{E}(u,u).
\end{aligned}$$

This proves $(\mathcal{E}^*, \mathcal{F}^*) = (\mathcal{E}, \mathcal{F})$. □

Following [12], for any $u \in \mathcal{F}$, we define its flux $\mathcal{N}_p(u)(a^*)$ at $a^*$ by

$$\mathcal{N}_p(u)(a^*) = \int_E \nabla u(x) \cdot \nabla u_1(x) m_p(dx) + \int_E \Delta u(x) u_1(x) m_p(dx),$$

which by the Green-Gauss formula equals

$$\int_{\partial B_\varepsilon} \frac{\partial u(x)}{\partial \vec{n}} \sigma(dx) - p u'(a^*).$$

Here $\vec{n}$ is the unit inward normal vector field of $B_\varepsilon$ at $\partial B_\varepsilon$. The last formula justifies the name of "flux" at $a^*$.

Let $\mathcal{L}$ be the $L^2$-infinitesimal generator of $(\mathcal{E}, \mathcal{F})$, or equivalently, the BMVD $X$, with domain of definition $\mathcal{D}(\mathcal{L})$. It can be viewed as the "Laplacian" on the space $E$ of varying dimension.

**Theorem 2.3.** *A function $u \in \mathcal{F}$ is in $\mathcal{D}(\mathcal{L})$ if and only if the distributional Laplacian $\Delta u$ of $u$ exists as an $L^2$-integrable function on $E \setminus \{a^*\}$ and $u$ has zero flux at $a^*$. Moreover, $\mathcal{L}u = \frac{1}{2}\Delta u$ on $E \setminus \{a^*\}$ for $u \in \mathcal{D}(\mathcal{L})$.*

*Proof.* By the Dirichlet form characterization, $u \in \mathcal{D}(\mathcal{L})$ if and only if $u \in \mathcal{F}$ and there is some $f \in L^2(E; m_p)$ so that

$$\mathcal{E}(u, v) = -\int_E f(x)v(x)m_p(dx) \quad \text{for every } v \in \mathcal{F}.$$

In this case, $\mathcal{L}u := f$. The above is equivalent to

$$\frac{1}{2}\int_E \nabla u(x) \cdot \nabla v(x) m_p(dx) = -\int_E f(x)v(x)m_p(dx) \quad \text{for every } v \in C_c^\infty(E \setminus \{a^*\}) \tag{2.3}$$

and

$$\frac{1}{2}\int_E \nabla u(x) \cdot \nabla u_1(x) m_p(dx) = -\int_E f(x)u_1(x)m_p(dx). \tag{2.4}$$

Equation (2.3) implies that $f = \frac{1}{2}\Delta u \in L^2(E; m_p)$, and (2.4) is equivalent to $\mathcal{N}_p(u)(a^*) = 0$. □

# 3 Nash Inequality and Heat Kernel Upper Bound Estimate

In the rest of this paper, we fix $\varepsilon > 0$ and $p > 0$. Recall that $D_\varepsilon := \mathbb{R}^2 \setminus \overline{B_e(0, \varepsilon)}$. In the following, if no measure is explicitly mentioned in the $L^p$-space, it is understood as being with respect to the measure $m_p$; for instance, $L^p(E)$ means $L^p(E; m_p)$.

**Lemma 3.1.** *There exists $C_1 > 0$ so that*

$$\|f\|_{L^2(E)}^2 \leq C_1 \left(\mathcal{E}(f, f)^{1/2}\|f\|_{L^1(E)} + \mathcal{E}(f, f)^{\ 1/3}\|f\|_{L^1(E)}^{4/3}\right) \qquad \text{for every } f \in \mathcal{F}.$$

*Proof.* Since $D_\varepsilon \subset \mathbb{R}^2$ and $\mathbb{R}_+$ are smooth domains, we have by the classical Nash's inequality,

$$\|f\|_{L^2(D_\varepsilon)}^2 \leq c\|\nabla f\|_{L^2(D_\varepsilon)} \|f\|_{L^1(D_\varepsilon)} \quad \text{for } f \in W^{1,2}(D_\varepsilon) \cap L^1(D_\varepsilon),$$

and

$$\|f\|_{L^2(\mathbb{R}_+)}^3 \leq C\|f\|_{L^1(\mathbb{R}_+)}^2 \|f'\|_{L^2(\mathbb{R}_+)} \quad \text{for } f \in W^{1,2}(\mathbb{R}_+) \cap L^1(\mathbb{R}_+).$$

The desired inequality now follows by combining these two inequalities. □

The Nash-type inequality in Lemma 3.1 immediately implies that BMVD $X$ on $E$ has a symmetric density function $p(t, x, y)$ with respect to the measure $m_p$ and that the following on-diagonal estimate by [10, Corollary 2.12] holds.

**Proposition 3.2.** *There exists $C_2 > 0$ such that*

$$p(t, x, y) \leq C_2\left(\frac{1}{t} + \frac{1}{t^{1/2}}\right) \qquad \text{for all } t > 0 \text{ and } x, y \in E.$$

**Remark 3.3.** Note that by [10, Corollary 2.12], one initially only gets that the above estimate holds for a.e. $x, y \in E$. But it can be shown that $p(t,x,y)$ can be chosen so that for every $t>0$ and $y \in E$, $x \mapsto p(t,x,y)$ is continuous on $E$, and so the estimate in Proposition 3.2 holds for every $t>0$ and every $x, y \in E$. Indeed, since $X$ moves like Brownian motion in Euclidean spaces before hitting $a^*$, it is easy to verify that for each $t>0$, $(x,y) \mapsto p(t,x,y)$ is continuous in $(E \setminus \{a^*\}) \times (E \setminus \{a^*\})$. For each $t>0$ and fixed $y \in E \setminus \{a^*\}$,

$$\int_E p(t/2,y,z)^2 m_p(dz) = p(t,y,y) < \infty.$$

So by the Dirichlet form theory, $x \mapsto p(t,x,y) = \int_E p(t/2,x,z)p(t/2,z,y)m_p(dz)$ is $\mathcal{E}$-quasi-continuous on $E$. Since $a^*$ is non-polar for $X$, $x \mapsto p(t,x,y)$ is continuous at $a^*$, and hence is continuous on $E$. By the symmetry and Chapman-Kolmogorov equation again, we conclude that

$$x \mapsto p(t,x,a^*) = \int_E p(t/2,x,z)p(t/2,z,a^*)m_p(dz)$$

is continuous on $E$. Consequently, $p(t,x,y)$ is well defined pointwisely on $(0,\infty) \times E \times E$ so that for each fixed $t>0$ and $y \in E$, $p(t,x,y)$ is a continuous function in $x \in E$. Hence the estimate in Proposition 3.2 holds pointwise. Moreover, since for each $y \in E$, $(t,x) \mapsto p(t,x,y)$ is a parabolic function for BMVD $X$, Theorem 6.3 implies that $p(t,x,y)$ is in fact locally jointly Hölder continuous in $(t,x,y)$. □

We can use Davies method to get an off-diagonal upper bound estimate.

**Proposition 3.4.** *There exist $C_3, C_4 > 0$ such that*

$$p(t,x,y) \le C_3 \left( \frac{1}{t} + \frac{1}{t^{1/2}} \right) e^{-C_4 \rho(x,y)^2/t} \qquad \text{for all } t>0 \text{ and } x,y \in E.$$

*Proof.* Fix $x_0, y_0 \in E$, $t_0 > 0$. Set a constant $\alpha := \rho(y_0,x_0)/4t_0$ and $\psi(x) := \alpha |x|_\rho$. Then we define $\psi_n(x) = \psi(x) \wedge n$. Note that for $m_p$-a.e. $x \in E$,

$$e^{-2\psi_n(x)} |\nabla e^{\psi_n(x)}|^2 = |\nabla \psi_n(x)|^2 = |\alpha|^2 \, \mathbf{1}_{\{|x|_\rho \le \frac{n}{|\alpha|}\}}(x) \le \alpha^2.$$

Similarly, $e^{2\psi_n(x)} |\nabla e^{-\psi_n(x)}|^2 \le \alpha^2$. By [10, Corollary 3.28],

$$p(t,x,y) \le c \left( \frac{1}{t} + \frac{1}{t^{1/2}} \right) \exp\left( -|\psi(y) - \psi(x)| + 2t|\alpha|^2 \right). \tag{3.1}$$

Taking $t = t_0, x = x_0$ and $y = y_0$ in (3.1) completes the proof. □

As indicated by the statements of Theorems 1.3 and 1.4, the upper bound estimate in Proposition 3.4 is not sharp.

# 4 Signed Radial Process and Small Time Estimate

In order to get the sharp two-sided heat kernel estimates, we consider the radial process of $X$. Namely, we project $X$ to $\mathbb{R}$ by applying the following mapping from $E$ to $\mathbb{R}$:

$$u(x) = \begin{cases} -|x|, & x \in \mathbb{R}_+; \\ \\ |x|_\rho, & x \in D_\varepsilon. \end{cases} \tag{4.1}$$

We call $Y_t := u(X_t)$ the signed radial process of $X$. Observe that $u \in \mathcal{F}_{\rm loc}$, where $\mathcal{F}_{\rm loc}$ denotes the local Dirichlet space of $(\mathcal{E}, \mathcal{F})$, whose definition can be found, for instance, in [12, 19]. By Fukushima decomposition [19, Chapter 5],

$$Y_t - Y_0 = u(X_t) - u(X_0) = M_t^{[u]} + N_t^{[u]}, \quad \mathbb{P}_x\text{-a.s. for q.e. } x \in E,$$

where $M_t^{[u]}$ is a local martingale additive functional of $X$, and $N_t^{[u]}$ is a continuous additive functional of $X$ locally having zero energy. We can explicitly compute $M^{[u]}$ and $N^{[u]}$. For any $\psi \in C_c^\infty(E)$,

$$\begin{aligned}
\mathcal{E}(u,\psi) &= \frac{1}{2}\int_{D_\varepsilon} \nabla|x|\cdot\nabla\psi dx + \frac{p}{2}\int_{\mathbb{R}_+}(-1)\psi' dx \\
&= -\frac{1}{2}\int_{D_\varepsilon} \operatorname{div}\left(\frac{x}{|x|}\right)\psi dx - \frac{1}{2}\int_{\partial B_e(0,\varepsilon)} \psi(0)\frac{\partial|x|}{\partial\vec{n}}\sigma(dx) + \frac{p\psi(0)}{2} \\
&= -\frac{1}{2}\int_{D_\varepsilon}\frac{1}{|x|}\psi dx - \frac{2\pi\varepsilon - p}{2}\,\psi(0) \\
&= -\int_E \psi(x)\nu(dx),
\end{aligned}$$

where $\vec{n}$ is the outward pointing unit vector normal of the surface $\partial B_e(0,\varepsilon)$, $\sigma$ is the surface measure on $\partial B_e(0,\varepsilon) \subset \mathbb{R}^2$, and

$$\nu(dx) := \frac{1}{|2x|}\mathbf{1}_{D_\varepsilon}(x)dx + \frac{2\pi\varepsilon - p}{2}\,\delta_{\{a^*\}}.$$

Recall that we identify $0 \in \mathbb{R}_+$ with $a^*$. It follows from [19, Theorem 5.5.5] that

$$dN_t^{[u]} = \frac{1}{2(u(X_t)+\varepsilon)}\mathbf{1}_{\{X_t\in D_\varepsilon\}}dt + (2\pi\varepsilon - p)dL_t^0(X), \tag{4.2}$$

where $L_t^0(X)$ is the positive continuous additive functional of $X$ having Revuz measure $\frac{1}{2}\delta_{\{a^*\}}$. We call $L^0$ the local time of $X$ at $a^*$. Next we compute $\langle M^{[u]}\rangle$, the predictable quadratic variation process of local martingale $M^{[u]}$). Let $u_n = (-n)\vee u \wedge n$, and it immediately follows $u_n \in \mathcal{F}$. Let $\mathcal{F}_b$ denote the space of bounded functions in $\mathcal{F}$. By [19, Theorem 5.5.2], the Revuz measure $\mu_{\langle u_n\rangle}$ for $\langle M^{[u_n]}\rangle$ can be calculated as follows. For any $f \in \mathcal{F}_b \cap C_c(E)$,

$$\int_E f(x)\mu_{\langle u_n\rangle}(dx) = 2\mathcal{E}(u_n f, u_n) - \mathcal{E}(u_n^2, f) = \int_E f(x)|\nabla u_n(x)|^2 m_p(dx),$$

which shows that

$$\mu_{\langle u_n\rangle}(dx) = |\nabla u_n(x)|^2 m_p(dx) = \mathbf{1}_{B_\rho(a^*,n)} m_p(dx).$$

By the strong local property of $(\mathcal{E},\mathcal{F})$, we have $\mu_{\langle u\rangle} = \mu_{\langle u_n\rangle}$ on $B_\rho(a^*,n)$. It follows that $\mu_{\langle u\rangle}(dx) = m_p(dx)$. Thus by [12, Proposition 4.1.9], $\langle M^{[u]}\rangle_t = t$ for $t \geq 0$ and so $B_t := M_t^{[u]}$ is a one-dimensional Brownian motion. Combining this with (4.2), we conclude

$$\begin{aligned}
dY_t &= dB_t + \frac{1}{2(Y_t+\varepsilon)}\mathbf{1}_{\{X_t\in D_\varepsilon\}}dt + (2\pi\varepsilon - p)dL_t^0(X) \\
&= dB_t + \frac{1}{2(Y_t+\varepsilon)}\mathbf{1}_{\{Y_t>0\}}dt + (2\pi\varepsilon - p)dL_t^0(X).
\end{aligned} \tag{4.3}$$

We next find the SDE for the semimartingale $Y$. The semi-martingale local time of $Y$ is denoted as $L_t^0(Y)$, that is,

$$L_t^0(Y) := \lim_{\delta\downarrow 0}\frac{1}{\delta}\int_0^t 1_{[0,\delta)}(Y_s)d\langle Y\rangle_s = \lim_{\delta\downarrow 0}\frac{1}{\delta}\int_0^t 1_{[0,\delta)}(Y_s)ds, \tag{4.4}$$

where $\langle Y\rangle_t = t$ is the quadratic variation process of the semimartingale $Y$.

**Proposition 4.1.** $L_t^0(Y) = 4\pi\varepsilon L_t^0(X)$.

*Proof.* By computation analogous to that for $Y_t = u(X_t)$, one can derive using Fukushima's decomposition for $v(X_t) := |X_t|_\rho$, that

$$dv(X_t) = d\widetilde{B_t} + \frac{1}{2(|X_t|_\rho + \varepsilon)}\mathbf{1}_{\{X_t\in D_\varepsilon\}}dt + (2\pi\varepsilon + p)dL_t^0(X),$$

where $\widetilde{B}$ is a one-dimensional Brownian motion. Observe that $v(X_t) = |Y_t|$. Thus we have

$$d|Y_t| = d\widetilde{B_t} + \frac{1}{2(Y_t + \varepsilon)}\mathbf{1}_{\{Y_t>0\}}dt + (2\pi\varepsilon + p)dL_t^0(X).$$

On the other hand, by Tanaka's formula, we have

$$\begin{aligned} d|Y_t| &= \mathrm{sgn}(Y_t)dY_t + dL_t^0(Y) \\ &= \mathrm{sgn}(Y_t)dB_t + \mathrm{sgn}(Y_t)\frac{1}{2(Y_t+\varepsilon)}\mathbf{1}_{\{Y_t>0\}}dt + \mathrm{sgn}(Y_t)(2\pi\varepsilon - p)dL_t^0(X) + dL_t^0(Y) \\ &= \mathrm{sgn}(Y_t)dB_t + \frac{1}{2(Y_t+\varepsilon)}\mathbf{1}_{\{Y_t>0\}}dt + (2\pi\varepsilon - p)\mathrm{sgn}(Y_t)dL_t^0(X) + dL_t^0(Y), \end{aligned}$$

where $\mathrm{sgn}(x) := \mathbf{1}_{\{x<0\}} - \mathbf{1}_{\{x\le 0\}}$. Since the decomposition of a continuous semi-martingale as the sum of a continuous local martingale and a continuous process with finite variation is unique, one must have

$$(2\pi\varepsilon + p)L_t^0(X) = \mathrm{sgn}(Y_t)(2\pi\varepsilon - p)L_t^0(X) + L_t^0(Y). \tag{4.5}$$

The local time $L_t^0(X)$ increases only when $Y_t = 0$. Therefore

$$(2\pi\varepsilon + p)L_t^0(X) = -(2\pi\varepsilon - p)L_t^0(X) + L_t^0(Y),$$

and so $4\pi\varepsilon L_t^0(X) = L_t^0(Y)$. □

The semi-martingale local time in (4.4) is non-symmetric in the sense that it only measures the occupation time of $Y_t$ in the one-sided interval $[0,\delta)$ instead of the symmetric interval $(-\delta,\delta)$. One can always relate the non-symmetric semi-martingale local time $L^0(Y)$ to the symmetric semi-martingale local time $\widehat{L}^0(Y)$ defined by

$$\widehat{L}_t^0(Y) := \lim_{\delta\downarrow 0}\frac{1}{2\delta}\int_0^t 1_{(-\delta,\delta)}(Y_s)d\langle Y\rangle_s = \lim_{\delta\downarrow 0}\frac{1}{2\delta}\int_0^t 1_{(-\delta,\delta)}(Y_s)ds.$$

**Lemma 4.2.** $\widehat{L}_t^0(Y) = \frac{2\pi\varepsilon+p}{4\pi\varepsilon}L_t^0(Y) = (2\pi\varepsilon + p)L_t^0(X)$.

*Proof.* Viewing $|Y_t| = |-Y_t|$ and applying Tanaka's formula to the semimartingale $-Y$, we can derive in a way analogous to the computation leading to (4.5) that

$$(2\pi\varepsilon + p)L_t^0(X) = -\mathrm{sgn}(-Y_t)(2\pi\varepsilon - p)L_t^0(X) + L_t^0(-Y) = (2\pi\varepsilon - p)L_t^0(X) + L_t^0(-Y).$$

Thus we get $2pL_t^0(X) = L_t^0(-Y)$, which yields

$$\widehat{L}_t^0(Y) = \frac{1}{2}\left(L_t^0(Y) + L_t^0(-Y)\right) = \frac{1}{2}\left(4\pi\varepsilon L_t^0(X) + 2pL_t^0(X)\right) = (2\pi\varepsilon + p)L_t^0(X).$$

□

Lemma 4.2 together with (4.3) gives the following SDE characterization for the signed radial process $Y$, which tells us precisely how $X$ moves after hitting $a^*$.

**Proposition 4.3.**

$$dY_t = dB_t + \frac{1}{2(Y_t + \varepsilon)}\mathbf{1}_{\{Y_t>0\}}dt + \frac{2\pi\varepsilon - p}{2\pi\varepsilon + p}d\widehat{L}_t^0(Y). \tag{4.6}$$

Let $\beta = \frac{2\pi\varepsilon - p}{2\pi\varepsilon + p}$. SDE (4.6) says that $Y$ is a skew Brownian motion with drift on $\mathbb{R}$ with skew parameter $\beta$. It follows (see [29]) that starting from $a^*$, the process $Y$ (resp. $X$) has probability $(1-\beta)/2 = \frac{p}{2\pi\varepsilon+p}$ to enter $(-\infty, 0)$ (resp. $\mathbb{R}_+$) and probability $(1+\beta)/2 = \frac{2\pi\varepsilon}{2\pi\varepsilon+p}$ to enter $(0, \infty)$ (resp. $D_\varepsilon$).

SDE (4.6) has a unique strong solution; see, e.g., [5]. So $Y$ is a strong Markov process on $\mathbb{R}$. The following is a key to get the two-sided sharp heat kernel estimate on $p(t, x, y)$ for BMVD $X$.

**Proposition 4.4.** *The one-dimensional diffusion process $Y$ has a jointly continuous transition density function $P^{(Y)}(t, x, y)$ with respect to the Lebesgue measure on $\mathbb{R}$. Moreover, for every $T \geq 1$, there exist constants $C_i > 0$, $1 \leq i \leq 4$, such that the following estimate holds:*

$$\frac{C_1}{\sqrt{t}}e^{-C_2|x-y|^2/t} \leq p^{(Y)}(t, x, y) \leq \frac{C_3}{\sqrt{t}}e^{-C_4|x-y|^2//t}, \quad (t, x, y) \in (0, T] \times \mathbb{R} \times \mathbb{R}. \tag{4.7}$$

*Proof.* Let $\beta := \frac{2\pi\varepsilon - p}{2\pi\varepsilon + p}$ and $Z$ be the skew Brownian motion

$$dZ_t = dB_t + \beta\widehat{L}_t^0(Z),$$

where $\widehat{L}_t^0(Z)$ is the symmetric local time of $Z$ at 0. The diffusion process $Y$ can be obtained from $Z$ through a drift perturbation (i.e. Girsanov transform). The transition density function $p_0(t, x, y)$ of $Z$ is explicitly known and enjoys the two-sided Aronson-type Gaussian estimates (4.7); see, e.g., [29]. One can further verify that

$$|\nabla_x p_0(t, x, y)| \leq c_1 t^{-1} \exp(-c_2|x - y|^2/t),$$

from which one can deduce (4.7) by using the same argument as that for Theorem A in Zhang [35, §4]. □

**Remark 4.5.** As mentioned earlier, estimate (4.7) for $p^{(Y)}(t,x,y)$ is the key to derive two-sided estimates on the transition density function $p(t,x,y)$ of BMVD over finite time intervals. It follows from Proposition 4.3 that $p^{(Y)}(t,x,y)$ is the fundamental solution of the following partial differential equation on $\mathbb{R}$:

$$\frac{\partial}{\partial t}u(t,x)=\frac{1}{2}\left(\frac{\partial^2}{\partial x^2}+\frac{1}{x+\varepsilon}\mathbf{1}_{\{x>0\}}\frac{\partial}{\partial x}\right)u(t,x) \quad \text{with} \quad \frac{\partial}{\partial x}u(t,0+)=\frac{2\pi\varepsilon}{p}\frac{\partial}{\partial x}u(t,0-). \tag{4.8}$$

Here $\frac{\partial}{\partial x}u(t,0+)$ and $\frac{\partial}{\partial x}u(t,0-)$ denote the right and left derivative of $x\mapsto u(t,x)$ at $x=0$, respectively. An interesting question is whether one can use PDE method to solve (4.8) and get a more precise estimates on $p^{(Y)}(t,x,y)$ that gives the explicit dependence on $\varepsilon$ and $p$. If this can be done, then by exploring the rotation invariance of the BMVD $X$ starting from $a^*$ as in the proof of Theorems 4.7 and 4.8 below, one may be able to sharpen the heat kernel estimates on $p(t,x,y)$ in Theorems 1.3 and 1.4 with explicit information on the dependence of $\varepsilon$ and $p$ . So far, we are unable to use PDE method to solve (4.8) explicitly.

Proposition 4.4 immediately gives the two-sided estimates on the transition function $p(t,x,y)$ of $X$ when $x,y\in\mathbb{R}_+$ since $X_t=-Y_t$ when $X_t\in\mathbb{R}_+$.

**Theorem 4.6.** *For every $T\geq 1$, there exist $C_i>0$, $5\leq i\leq 8$, such that the following estimate holds:*

$$\frac{C_5}{\sqrt{t}}e^{-C_6|x-y|^2/t}\leq p(t,x,y)\leq \frac{C_7}{\sqrt{t}}e^{-C_8|x-y|^2/t} \qquad \text{for } t\in(0,T] \text{ and } x,y\in\mathbb{R}_+.$$

Let $A$ be any rotation of the plane around the pole. Using the fact that starting from $a^*$, $AX_t$ has the same distribution as $X_t$, we can derive estimates for $p(t,x,y)$ for other $x,y\in E$. The next result gives the two-sided estimates on $p(t,x,y)$ when $x\in\mathbb{R}$ and $y\in D_\varepsilon$.

**Theorem 4.7.** *For every $T\geq 1$, there exist constants $C_i>0$, $9\leq i\leq 12$, such that for all $x\in\mathbb{R}_+$, $y\in D_\varepsilon$ and $t\in[0,T]$,*

$$\frac{C_9}{\sqrt{t}}e^{-C_{10}\rho(x,y)^2/t}\leq p(t,x,y)\leq \frac{C_{11}}{\sqrt{t}}e^{-C_{12}\rho(x,y)^2/t}.$$

*Proof.* We first note that in this case by the symmetry of $p(t,x,y)$,

$$p(t,x,y)=p(t,y,x)=\int_0^t \mathbb{P}_y(\sigma_{a^*}\in ds)p(t-s,a^*,x).$$

By the rotational invariance of two-dimensional Brownian motion, $\mathbb{P}_y(\sigma_{a^*}\in ds)$ only depends on $|y|_\rho$, therefore so does $y\mapsto p(t,x,y)$. For $x\in\mathbb{R}_+$ and $y\in D_\varepsilon$, set $\widetilde{p}(t,x,r):=p(t,x,y)$ for $r=|y|_\rho$. For all $a>b>0$ and $x\in\mathbb{R}_+$,

$$\begin{aligned}\int_a^b p^{(Y)}(t,-|x|,y)dy &= \mathbb{P}_{-|x|}(a\leq Y_t\leq b)=\mathbb{P}_x(X_t\in D_\varepsilon \text{ with } a\leq |X_t|_\rho\leq b)\\ &=\int_{y\in D_\varepsilon:a\leq|y|_\rho\leq b}p(t,x,y)m_p(dy)=\int_{y\in D_\varepsilon:a+\varepsilon\leq|y|\leq b+\varepsilon}p(t,x,y)m_p(dy)\\ &=\int_a^b 2\pi(r+\varepsilon)\widetilde{p}(t,x,r)dr.\end{aligned}$$

This implies when $x \in \mathbb{R}_+$, $y \in D_\varepsilon$,

$$p^{(Y)}(t, -|x|, |y|_\rho) = 2\pi(|y|_\rho + \varepsilon)\widetilde{p}(t, x, |y|_\rho) = 2\pi(|y|_\rho + \varepsilon)p(t, x, y). \tag{4.9}$$

We thus have by Proposition 4.4 that

$$\frac{c_1}{\sqrt{t}} e^{-c_2\rho(x,y)^2/t} \le p(t, x, y) \le \frac{c_3}{\sqrt{t}} e^{-c_4\rho(x,y)^2/t} \quad \text{for } x \in \mathbb{R}_+ \text{ and } y \in D_\varepsilon \text{ with } |y|_\rho < 1. \tag{4.10}$$

When $|y|_\rho > 1$, we first have

$$p(t, x, y) = \frac{1}{2\pi(|y|_\rho + \varepsilon)} p^{(Y)}(t, -|x|, |y|_\rho) \lesssim \frac{1}{(|y|_\rho + \varepsilon)\sqrt{t}} e^{-c_3\rho(x,y)^2/t} \le \frac{1}{\sqrt{t}} e^{-c_3\rho(x,y)^2/t},$$

while since $\rho(x, y) \ge |y|_\rho > 1$,

$$\begin{aligned} p(t, x, y) &= \frac{1}{2\pi(|y|_\rho + \varepsilon)} p^{(Y)}(t, -|x|, |y|_\rho) \gtrsim \frac{1}{(|y|_\rho + \varepsilon)\sqrt{t}} e^{-c_4\rho(x,y)^2/t} \\ &\gtrsim \frac{1}{\sqrt{t}} \frac{\sqrt{t}}{\sqrt{T}\rho(x, y)} e^{-c_4\rho(x,y)^2/t} \gtrsim \frac{1}{\sqrt{t}} e^{-(c_4+1)\rho(x,y)^2/t}. \end{aligned} \tag{4.11}$$

This completes the proof. □

Theorems 4.6 and 4.7 establish Theorem 1.3(i). We next consider part (ii) of Theorem 1.3 when both $x$ and $y$ are in $D_\varepsilon$.

**Theorem 4.8.** *For every $T \ge 1$, there exist constants $C_i > 0$, $13 \le i \le 22$, such that for all $t \in [0, T]$ and $x, y \in D_\varepsilon$, the following estimates hold.*
*When* $\max\{|x|_\rho, |y|_\rho\} \le 1$,

$$\begin{aligned} &\frac{C_{13}}{\sqrt{t}} e^{-C_{14}\rho(x,y)^2/t} + \frac{C_{13}}{t} \left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right) \left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-C_{15}|x-y|^2/t} \le p(t, x, y) \\ &\le \frac{C_{16}}{\sqrt{t}} e^{-C_{17}\rho(x,y)^2/t} + \frac{C_{16}}{t} \left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right) \left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-C_{18}|x-y|^2/t}; \end{aligned} \tag{4.12}$$

*and when* $\max\{|x|_\rho, |y|_\rho\} > 1$,

$$\frac{C_{19}}{t} e^{-C_{20}\rho(x,y)^2/t} \le p(t, x, y) \le \frac{C_{21}}{t} e^{-C_{22}\rho(x,y)^2/t}. \tag{4.13}$$

*Here $|\cdot|$ and $|\cdot|_\rho$ denote the Euclidean metric and the geodesic metric in $D_\varepsilon$, respectively.*

*Proof.* For $x \in D_\varepsilon$ and $t \in (0, T]$, note that

$$p(t, x, y) = \overline{p}_{D_\varepsilon}(t, x, y) + p_{D_\varepsilon}(t, x, y), \tag{4.14}$$

where

$$\overline{p}_{D_\varepsilon}(t, x, y) = \int_0^t p(t - s, a^*, y)\mathbb{P}_x(\sigma_{\{a^*\}} \in ds). \tag{4.15}$$

As mentioned in the proof for Theorem 4.7, $p(t-s,a^*,y)$ is a function in $y$ depending only on $|y|_\rho$. Therefore so is $y\mapsto \overline{p}_{D_\varepsilon}(t,x,y)$. Set $\widetilde{p}_{D_\varepsilon}(t,x,r):=\overline{p}_{D_\varepsilon}(t,x,y)$ for $r=|y|_\rho$. For any $b>a>0$,

$$\begin{aligned}\mathbb{P}_x\left(\sigma_{a^*}<t,\, X_t\in D_\varepsilon \text{ with } a\le |X_t|_\rho\le b\right) &= \int_{a\le |y|_\rho\le b}\overline{p}_{D_\varepsilon}(t,x,y)m_p(dy)\\ &= 2\pi\int_a^b (r+\varepsilon)\widetilde{p}_{D_\varepsilon}(t,x,r)dr.\end{aligned}$$

On the other hand,

$$\begin{aligned}&\mathbb{P}_x\left(\sigma_{a^*}<t,\, X_t\in D_\varepsilon \text{ with } a\le |X_t|_\rho\le b\right)\\ =\ &\mathbb{P}^{(Y)}_{|x|_\rho}\left(\sigma_{a^*}<t,\, Y_t>0 \text{ with } a\le |Y_t|_\rho\le b\right)=\int_0^t\left(\int_a^b p^{(Y)}(t-s,0,r)dr\right)\mathbb{P}^{(Y)}_{|x|_\rho}(\sigma_0\in ds).\end{aligned}$$

It follows that

$$2\pi(r+\varepsilon)\widetilde{p}_{D_\varepsilon}(t,x,r)=\int_0^t p^{(Y)}(t-s,0,r)\mathbb{P}^{(Y)}_{|x|_\rho}(\sigma_{\{0\}}\in ds)=p^{(Y)}(t,-|x|_\rho,r).$$

In other words,

$$\overline{p}_{D_\varepsilon}(t,x,y)=\frac{1}{2(|y|_\rho+\varepsilon)}p^{(Y)}(t,-|x|_\rho,|y|_\rho). \tag{4.16}$$

It is known that the Dirichlet heat kernel $p_{D_\varepsilon}(t,x,y)$ enjoys the following two-sided estimates:

$$\frac{c_1}{t}\left(1\wedge\frac{|x|_\rho}{\sqrt{t}}\right)\left(1\wedge\frac{|y|_\rho}{\sqrt{t}}\right)e^{-c_2|x-y|^2/t}\le p_{D_\varepsilon}(t,x,y)\le \frac{c_3}{t}\left(1\wedge\frac{|x|_\rho}{\sqrt{t}}\right)\left(1\wedge\frac{|y|_\rho}{\sqrt{t}}\right)e^{-c_4|x-y|^2/t} \tag{4.17}$$

for $t\in(0,T]$ and $x,y\in D_\varepsilon$. We now consider two different cases.

*Case* (i): $\max\{|x|_\rho,|y|_\rho\}\le 1$ and $t\in(0,T]$. In this case, it follows from (4.14)-(4.17) and Proposition 4.4 that

$$\begin{aligned}&\frac{c_5}{\sqrt{t}}e^{-c_6(|x|_\rho+|y|_\rho)^2/t}+\frac{c_5}{t}\left(1\wedge\frac{|x|_\rho}{\sqrt{t}}\right)\left(1\wedge\frac{|y|_\rho}{\sqrt{t}}\right)e^{-c_7|x-y|^2/t}\le p(t,x,y)\\ &\qquad\le \frac{c_8}{\sqrt{t}}e^{-c_9(|x|_\rho+|y|_\rho)^2/t}+\frac{c_8}{t}\left(1\wedge\frac{|x|_\rho}{\sqrt{t}}\right)\left(1\wedge\frac{|y|_\rho}{\sqrt{t}}\right)e^{-c_{10}|x-y|^2/t}.\end{aligned} \tag{4.18}$$

Observe that

$$(|x|_\rho+|y|_\rho)^2/t\asymp \rho(x,y)^2/t \quad \text{ if } |x|_\rho\wedge|y|_\rho\le\sqrt{t}. \tag{4.19}$$

When $|x|_\rho\wedge|y|_\rho>\sqrt{t}$, for $a>0$, $b>0$,

$$\begin{aligned}&\frac{1}{\sqrt{t}}e^{-a(|x|_\rho+|y|_\rho)^2/t}+\frac{1}{t}\left(1\wedge\frac{|x|_\rho}{\sqrt{t}}\right)\left(1\wedge\frac{|y|_\rho}{\sqrt{t}}\right)e^{-b|x-y|^2/t}\\ \asymp\ &\frac{1}{\sqrt{t}}e^{-a(|x|_\rho+|y|_\rho)^2/t}+\left(\frac{1}{\sqrt{t}}+\frac{1}{t}\right)e^{-b|x-y|^2/t}\\ =\ &\frac{1}{\sqrt{t}}\left(e^{-a(|x|_\rho+|y|_\rho)^2/t}+e^{-b|x-y|^2/t}\right)+\frac{1}{t}e^{-b|x-y|^2/t}\end{aligned} \tag{4.20}$$

The desired estimate (4.12) now follows from (4.18)-(4.20) and the fact (1.2).

*Case* (ii): $\max\{|x|_\rho, |y|_\rho\} > 1$ and $t \in (0, T]$. By the symmetry of $p(t, x, y)$ in $x$ and $y$, in this case we may and do assume $|y|_\rho > 1 > \sqrt{t/T}$. It then follows from (4.16)-(4.17), Proposition 4.4 and (4.11) that

$$\frac{c_{11}}{t}e^{-c_{12}(|x|_\rho+|y|_\rho)^2/t} + \frac{c_{11}}{t}\left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-c_{13}|x-y|^2/t} \le p(t, x, y)$$
$$\le \frac{c_{14}}{t}e^{-c_{15}(|x|_\rho+|y|_\rho)^2/t} + \frac{c_{14}}{t}\left(1 \wedge \frac{|x|_\rho}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|_\rho}{\sqrt{t}}\right) e^{-c_{16}|x-y|^2/t}. \quad (4.21)$$

When $|x|_\rho \wedge |y|_\rho \le \sqrt{t}$, the lower bound estimate (4.13) follows from (4.21) and (4.19), while the upper bound estimate (4.13) follows from Proposition 3.4. Whereas when $|x|_\rho \wedge |y|_\rho > \sqrt{t}$, the desired estimate (4.13) follows from (4.21) and (1.2). This completes the proof of the theorem. □

**Remark 4.9.** (i) One cannot expect to rewrite the estimate of (4.12) as $t^{-1}e^{-c\rho(x,y)^2/t}$. A counterexample is that $x = y = a^*$, in which case $x$ and $y$ can be viewed as either on $\mathbb{R}$ or on $D_\varepsilon$, therefore both Proposition 4.4 and Theorem 4.7 have already confirmed that $p(t, x, y) \asymp t^{-1/2}$, which is consistent with the (4.12).

(ii) The Euclidean distance appearing in (4.12) cannot be replaced with the geodesic distance. To see this, take $x = (\varepsilon + t^{-1/2}, 0)$ and $y = (-\varepsilon - t^{-1/2}, 0)$ in $D_\varepsilon$. The estimate of (4.12) is comparable with $t^{-1/2} + t^{-1}\exp(-\varepsilon^2/t)$, but if we replaced $|x - y|$ with $\rho(x, y)$, it would be comparable with $t^{-1/2} + t^{-1}$. For fixed $\varepsilon$, as $t \downarrow 0$, $t^{-1/2} + t^{-1}\exp(-\varepsilon^2/t) \sim t^{-1/2}$, but $t^{-1/2} + t^{-1} \sim t^{-1}$.

(iii) Theorem 1.3 also shows that the parabolic Harnack inequality fails for $X$. For a precise statement of the parabolic Harnack inequality, see, for example, [20, 30, 32]. For $s \in (0, 1]$, take some $y \in D_\varepsilon$ such that $|y|_\rho = \sqrt{s}$. Set $Q_+ := (3s/2, 2s) \times B_\rho(y, 2\sqrt{s})$ and $Q_- := (s/2, s) \times B_\rho(y, 2\sqrt{s})$. Let $u(t, x) := p(t, x, y)$. It follows from Theorem 1.3 that $\sup_{Q_+} u \asymp s^{-1/2} + s^{-1} \asymp s^{-1}$ and $\inf_{Q_-} u \asymp s^{-1/2}$. Clearly there does not exist any positive constant $C > 0$ so that $\sup_{Q_+} u \le C \inf_{Q_-} u$ holds for all $s \in (0, 1]$. This shows that parabolic Harnack inequality fails for $X$.

**Proof of Theorem 1.2(ii).** It follows from Remark 3.3, Theorem 1.3 and the Lebesgue dominated convergence theorem that for every bounded function $f$ on $E$ and $t > 0$,

$$P_t f(x) := \mathbb{E}_x[f(X_t)] = \int_E p(t, x, y) f(y) m_p(dy)$$

is continuous on $E$. For $f \in C_\infty(E)$, we have by Theorem 1.3 again that $P_t f \in C_\infty(E)$. Clearly, since for every $x \in E$,

$$\lim_{t\to 0} P_t(x) = \lim_{t\to 0} \mathbb{E}_x[f(X_t)] = \mathbb{E}_x\left[\lim_{t\to 0} f(X_t)\right] = f(x),$$

we have $\lim_{t\to 0} \|P_t f - f\|_\infty = 0$ for every $f \in C_\infty(E)$. This proves that the BMVD is a Feller process having strong Feller property. □

## 5 Large time heat kernel estimates

Recall that $X$ denotes the BMVD process on $E$ and its signed radial process defined by (4.3) is denoted by $Y$. In this section, unless otherwise stated, it is always assumed that $T \geq 8$ and $t \in [T, \infty)$. With loss of generality, we assume that the radius $\varepsilon$ of the "hole" $B(0, \varepsilon)$ satisfies $\varepsilon \leq 1/4$. We begin with the following estimates for the distribution of hitting time of a disk by a two-dimensional Brownian motion, which follow directly from [21, §5.1, Case (a), $\alpha = 1$] and a Brownian scaling.

**Proposition 5.1** (Grigor'yan and Saloff-Coste [21])**.** *Let $X$ be a Brownian motion on $\mathbb{R}^2$ and $K$ be the closed ball with radius $\varepsilon$ centered at the origin.*

(i) *If $0 < t < 2|x|^2$ and $|x| \geq 1 + \varepsilon$, then*

$$\frac{c}{\log |x|} \exp\left(-C|x|^2/t\right) \leq \mathbb{P}_x(\sigma_K \leq t) \leq \frac{C}{\log |x|} \exp\left(-c|x|^2/t\right), \tag{5.1}$$

*for some positive constants $C > c > 0$.*

(ii) *If $t \geq 2|x|^2$ and $|x| \geq 1 + \varepsilon$, then*

$$\mathbb{P}_x(\sigma_K \leq t) \asymp \frac{\log \sqrt{t} - \log |x|}{\log \sqrt{t}}, \tag{5.2}$$

*and*

$$\partial_t \mathbb{P}_x(\sigma_K \leq t) \asymp \frac{\log |x|}{t(\log t)^2}. \tag{5.3}$$

Our first goal is to establish an upper bound estimate on $\int_0^t p(s, a^*, a^*)ds$ in Proposition 5.4. This will be done through two Propositions by using the above hitting time estimates.

**Proposition 5.2.** *$p(t, a^*, a^*)$ is decreasing in $t \in (0, \infty)$.*

*Proof.* This follows from

$$\begin{aligned}
\frac{d}{dt}p(t, a^*, a^*) &= \frac{d}{dt}\int_E p(t/2, a^*, x)^2 m_p(dx) \\
&= \int_E \left(\frac{\partial}{\partial t}p(\cdot, a^*, x)(t/2)\right) p(t/2, a^*, x) m_p(dx) \\
&= \int_E \mathcal{L}_x p(t/2, a^*, x) p(t/2, a^*, x) m_p(dx) \\
&= -\mathcal{E}\left(p(t/2, a^*, x), p(t/2, a^*, x)\right) \leq 0.
\end{aligned}$$

□

**Proposition 5.3.** *There exists some constant $C_1 > 0$ such that*

$$p(t, a^*, a^*) \leq C_1 \frac{\log t}{t} \quad \textit{for } t \in [8, \infty).$$

*Proof.* For $t \geq 8$ and $x \in D_\varepsilon$ with $1 < |x|_\rho < \sqrt{t/3}$, by Proposition 5.2,

$$p(t, x, a^*) = \int_0^t \mathbb{P}_x(\sigma_{a^*} \in ds) p(t-s, a^*, a^*) \geq p(t, a^*, a^*) \mathbb{P}_x(\sigma_{a^*} \leq t) \asymp p(t, a^*, a^*) \left(1 - \frac{\log |x|}{\log \sqrt{t}}\right),$$

where the " $\asymp$ " is due to (5.3). Therefore,

$$\begin{aligned} 1 \geq \mathbb{P}_{a^*}\left(X_t \in D_\varepsilon \text{ with } 1 < |X_t|_\rho < \sqrt{t/4}\right) &= \int_{D_\varepsilon \cap \{1 < |x|_\rho < \sqrt{t/4}\}} p(t, a^*, x) dx \\ &\geq c_1 p(t, a^*, a^*) \int_{1+\varepsilon}^{\sqrt{t/4}+\varepsilon} \left(1 - \frac{\log r}{\log \sqrt{t}}\right) r dr \\ &\geq c_2\, p(t, a^*, a^*)\, \frac{t}{\log t}. \end{aligned}$$

By selecting $C_2$ large enough, the above yields the desired estimate for $p(t, a^*, a^*)$ for $t \geq 8$. $\square$

**Proposition 5.4.** *There exists some $C_2 > 0$ such that*

$$\int_0^t p(s, a^*, a^*) ds \leq C_2 \log t, \qquad \text{for all } t \geq 4.$$

*Proof.* For $t \geq 8$ and $x \in D_\varepsilon$ with $1 < |x|_\rho < \sqrt{t/2}$, we have by (5.3),

$$p(t, x, a^*) \geq \int_{t/2}^t p(t-s, a^*, a^*) \mathbb{P}_x(\sigma_{a^*} \in ds) \asymp \frac{\log |x|}{t (\log t)^2} \int_0^{t/2} p(s, a^*, a^*) ds.$$

Thus by using polar coordinate,

$$\begin{aligned} 1 \quad &\geq \quad \mathbb{P}_{a^*}\left(X_t \in D_\varepsilon \text{ with } 1 < |X_t|_\rho < \sqrt{t/2}\right) = \int_{\{x \in D_\varepsilon : 1 < |x|_\rho < \sqrt{t/2}\}} p(t, a^*, x) m_p(dx) \\ &\gtrsim \quad \int_1^{\sqrt{t/2}} \frac{\log(r+\varepsilon)}{t (\log t)^2} (r+\varepsilon) dr \cdot \int_0^{t/2} p(s, a^*, a^*) ds \\ &= \quad \frac{1}{t (\log t)^2} \int_{1+\varepsilon}^{\sqrt{t/2}+\varepsilon} r \log r dr \cdot \int_0^{t/2} p(s, a^*, a^*) ds \\ &\asymp \quad \frac{t \log t}{t (\log t)^2} \int_0^{t/2} p(s, a^*, a^*) ds \\ &= \quad \frac{1}{\log t} \int_0^{t/2} p(s, a^*, a^*) ds. \end{aligned}$$

This yields the desired estimate. $\square$

The above proposition suggests that $p(t, a^*, a^*) \leq c/t$ for $t \in [4, \infty)$. However, in order to prove this rigorously, we first compute the upper bounds for $p(t, a^*, x)$ for different regions of $x$, and then use the identity $p(t, a^*, a^*) = \int_E p(t/2, a^*, x)^2 m_p(dx)$ to obtain the sharp upper bound estimate for $p(t, a^*, a^*)$.

**Proposition 5.5.** *There exists $C_3 > 0$ such that for all $t \geq 8$ and $x \in D_\varepsilon$ with $1 < |x|_\rho < \sqrt{t/2}$,*

$$p(t, a^*, x) \leq \frac{C_3}{t} \log\left(\frac{\sqrt{t}}{|x|}\right).$$

*Proof.* By Proposition 5.3, (5.2) and Proposition 5.4,

$$\begin{aligned} p(t, x, a^*) &= \int_0^{t/2} p(t-s, a^*, a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) + \int_{t/2}^t p(t-s, a^*, a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) \\ &\leq c_1 \left(\frac{\log t}{t}\mathbb{P}_x(\sigma_{a^*} \leq t/2) + \frac{\log |x|}{t(\log t)^2}\int_0^{t/2} p(s, a^*, a^*)ds\right) \\ &\leq c_2 \left(\frac{\log t}{t}\frac{\log\left(\sqrt{t}/|x|\right)}{\log \sqrt{t}} + \frac{\log |x|}{t \log t}\right) \\ &\leq c_3 \left(\frac{1}{t}\log\left(\frac{\sqrt{t}}{|x|}\right) + \frac{1}{t}\right) \leq \frac{c_4}{t}\log\left(\frac{\sqrt{t}}{|x|}\right). \end{aligned}$$

□

The following asymptotic estimate for the distribution of Brownian hitting time of a disk from [33, Theorem 2] will be used in the next proposition.

**Lemma 5.6** (Uchiyama [33])**.** *Let $(B_t)_{t\geq 0}$ be the standard two-dimensional Brownian motion and $\sigma_r := \inf\{t > 0, |B_t| \leq r\}$. Denote by $p_{r,x}(t)$ the probability density function of $\sigma_r$ with $B_0 = x$. For every $r_0$, uniformly for $|x| \geq r_0$, as $t \to \infty$,*

$$p_{r_0,x}(t) = \frac{\log\left(\frac{1}{2}e^{c_0}|x|^2\right)}{t\left(\log t + c_0\right)^2}\exp\left(-\frac{|x|^2}{2t}\right) + \begin{cases} O\left(\frac{1+(\log(|x|^2/t))^2}{|x|^2(\log t)^3}\right) & \text{for } |x|^2 \geq t, \\ \frac{2\gamma \log(t/|x|^2)}{t(\log t)^3} + O\left(\frac{1}{t(\log t)^3}\right) & \text{for } |x|^2 < t, \end{cases}$$

*where $c_0$ is a positive constant only depending on $r_0$ and $\gamma = -\int_0^\infty e^{-u}\log u du$ is the Euler constant.*

**Proposition 5.7.** *There exists $C_4, C_5 > 0$ such that for all $t \geq 8$ and all $x \in D_\varepsilon$ ,*

$$p(t, a^*, x) \leq \begin{cases} C_4\left(\frac{1}{t}e^{-C_5|x|^2/t} + \frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2(\log t)^2}\right) & \text{when } |x|_\rho > \sqrt{t}, \\ \frac{C_4}{t}\left(e^{-C_5|x|^2/t} + \frac{1}{(\log t)^2}\right) & \text{when } \sqrt{t}/2 \leq |x|_\rho \leq \sqrt{t}. \end{cases}$$

*Proof.* Note that

$$\begin{aligned} p(t, a^*, x) &= \int_0^t p(t-s, a^*, a^*)\mathbb{P}^x(\sigma_{a^*} \in ds) \\ &= \int_0^{t/2} p(t-s, a^*, a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) + \int_{t/2}^t p(t-s, a^*, a^*)\mathbb{P}_x(\sigma_{a^*} \in ds). \qquad (5.4) \end{aligned}$$

By the monotonicity of $p(t,a^*,a^*)$ established in Proposition 5.2, estimate (5.1) and Proposition 5.3,

$$\begin{aligned}
\int_0^{t/2} p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\in ds) &\le p(t/2,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\le t/2) \\
&\overset{(5.1)}{\le} p(t/2,a^*,a^*)\frac{c_1}{\log|x|}e^{-c_2|x|^2/t} \\
&\le \frac{c_1\log t}{t\log|x|}e^{-c_2|x|^2/t} \\
&\le \frac{c_3}{t}e^{-c_2|x|^2/t}. \qquad (5.5)
\end{aligned}$$

On the other hand, by Proposition 5.4 and Lemma 5.6, for $|x|_\rho\ge\sqrt{t}$,

$$\begin{aligned}
\int_{t/2}^{t} p(t-s,a^*,a^*)\mathbb{P}^x(\sigma_{a^*}\in ds) &\le \sup_{s\in[t/2,t]} p_{\varepsilon,x}(s)\cdot\int_{t/2}^{t} p(t-s,a^*,a^*)ds \\
&\le c_4\left(\frac{\log|x|}{t(\log t)^2}e^{-|x|^2/(2t)}+\frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2(\log t)^3}\right)\log t \\
&= c_4\left(\frac{\log|x|}{t\log t}e^{-|x|^2/(2t)}+\frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2(\log t)^2}\right) \\
&\le \frac{c_4}{t}e^{-c_5|x|^2/t}+c_4\frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2(\log t)^2}, \qquad (5.6)
\end{aligned}$$

while for $\sqrt{t}/2\le|x|_\rho\le\sqrt{t}$,

$$\begin{aligned}
\int_{t/2}^{t} p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\in ds) &\le \sup_{s\in[t/2,t]} p_{\varepsilon,x}(s)\cdot\int_{t/2}^{t} p(t-s,a^*,a^*)ds \\
&\le c_6\left(\frac{\log|x|}{t(\log t)^2}e^{-|x|^2/(2t)}+\frac{1}{t(\log t)^3}\right)\log t \\
&\le \frac{c_6}{t}e^{-c_5|x|^2/t}+\frac{c_6}{t(\log t)^2}. \qquad (5.7)
\end{aligned}$$

The desired estimate now follows from (5.4)-(5.7). □

**Proposition 5.8.** *There exists $C_6>0$ such that*

$$p(t,a^*,x)\le\frac{C_6\log t}{t}e^{-x^2/(2t)} \qquad \textit{for all } t\ge 8 \textit{ and } x\in\mathbb{R}_+.$$

*Proof.* Starting from $x\in\mathbb{R}_+$, BMVD $X$ on $E$ runs like a one-dimensional Brownian motion before hitting $a^*$. Thus by the known formula for the first passage distribution for one-dimensional Brownian motion,

$$\mathbb{P}_x(\sigma_{a^*}\in dt)=\frac{1}{\sqrt{2\pi t^3}}xe^{-x^2/(2t)}dt \quad \text{for } x\in(0,\infty). \qquad (5.8)$$

This together with Propositions 5.2–5.4 and a change of variable $s = x^2/r$ gives

$$\begin{aligned} p(t,a^*,x) &= \int_0^{t/2} p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\in ds) + \int_{t/2}^t p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\in ds) \\ &\le p(t/2,a^*,a^*)\int_0^{t/2} \frac{1}{\sqrt{2\pi s^3}} x e^{-x^2/(2s)}ds + \int_{t/2}^t p(t-s,a^*,a^*)\frac{1}{\sqrt{2\pi s^3}} x e^{-x^2/(2s)}ds \\ &\lesssim \frac{\log t}{t}\int_{2x^2/t}^\infty \frac{1}{\sqrt{r}}e^{-r/2}dr + \frac{x}{\sqrt{t^3}}e^{-x^2/t}\int_{t/2}^t p(t-s,a^*,a^*)ds \\ &\lesssim \frac{\log t}{t}e^{-x^2/t} + \frac{1}{t}e^{-x^2/(2t)}\cdot \log t \\ &\lesssim \frac{\log t}{t}e^{-x^2/(2t)}. \end{aligned}$$

□

We are now in a position to establish the following on-diagonal upper bound estimate at $a^*$.

**Theorem 5.9.** *There exists $C_7 > 0$ such that*

$$p(t,a^*,a^*) \le C_7\left(t^{-1/2}\wedge t^{-1}\right) \quad \textit{for all } t\in(0,\infty).$$

*Proof.* For $t \ge 8$, we have

$$\begin{aligned} p(t,a^*,a^*) &= \int_E p(t/2,a^*,x)^2 m_p(dx) \\ &= \left(\int_{\mathbb{R}_+} + \int_{D_\varepsilon\cap\{0<|x|_\rho<1\}} + \int_{D_\varepsilon\cap\{1<|x|_\rho<\sqrt{t/2}\}} + \int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t/2}\}}\right) p(t/2,a^*,x)^2 m_p(dx). \end{aligned} \tag{5.9}$$

It follows from Proposition 5.8 that

$$\int_{\mathbb{R}_+} p(t/2,a^*,x)^2 m_p(dx) \lesssim \left(\frac{\log t}{t}\right)^2 p\int_0^\infty e^{-x^2/(2t)}dx \asymp \frac{c_1(\log t)^2}{t^{3/2}}. \tag{5.10}$$

By Proposition 3.4,

$$\begin{aligned} \int_{D_\varepsilon\cap\{0<|x|_\rho<1\}} p(t/2,a^*,x)^2 m_p(dx) &\le \left(\sup_{x\in D_\varepsilon:0<|x|_\rho<1} p(t/2,a^*,x)\right)^2 m_p(D_\varepsilon\cap\{0<|x|_\rho<1\}) \\ &\lesssim \left(\frac{1}{\sqrt{t}}\right)^2 = \frac{1}{t}. \end{aligned} \tag{5.11}$$

In view of Proposition 5.7,

$$\begin{aligned} &\int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t/2}\}} p(t/2,a^*,x)^2 m_p(dx) \\ \le\ & \int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t/2}\}} \left(\frac{c_1}{t}e^{-c_2|x|^2/t}\right)^2 m_p(dx) + \int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t}\}} c_1\left(\frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2\left(\log t\right)^2}\right)^2 m_p(dx) \\ &+ \int_{D_\varepsilon\cap\{\sqrt{t/2}\le|x|_\rho\le\sqrt{t}\}} c_1\left(\frac{1}{t(\log t)^2}\right)^2 m_p(dx). \end{aligned} \tag{5.12}$$

Using polar coordinate,

$$\int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t/2}\}} \left(\frac{c_1}{t}e^{-c_2|x|^2/t}\right)^2 m_p(dx) = c_3\int_{\sqrt{t/2}+\varepsilon}^\infty \frac{r}{t^2}e^{-c_2r^2/t}dr \le \frac{c_4}{t}, \tag{5.13}$$

while

$$\begin{aligned}
&\int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t}\}} \left(\frac{\left(\log\left(|x|^2/t\right)\right)^2}{|x|^2\left(\log t\right)^2}\right)^2 m_p(dx) = 2\pi\int_{\sqrt{t}+\varepsilon}^\infty \frac{r\left(\log(r^2/t)\right)^4}{r^4(\log t)^4}dr \\
\overset{u=r/\sqrt{t}}{\le}\quad & 2\pi\int_1^\infty \frac{(\log u)^4}{u^3t^{3/2}(\log t)^4}\sqrt{t}\,du = \frac{2\pi}{t(\log t)^4}\int_1^\infty\frac{(\log u)^4}{u^3}\,du = \frac{c_5}{t(\log t)^4},
\end{aligned} \tag{5.14}$$

and

$$\int_{D_\varepsilon\cap\{\sqrt{t/2}\le|x|_\rho\le\sqrt{t}\}} \left(\frac{1}{t(\log t)^2}\right)^2 m_p(dx) \asymp \frac{1}{t(\log t)^4} \lesssim \frac{1}{t}. \tag{5.15}$$

Hence it follows from (5.12)–(5.15) that

$$\int_{D_\varepsilon\cap\{|x|_\rho>\sqrt{t/2}\}} p(t/2,a^*,x)^2 m_p(dx) \le \frac{c_6}{t}. \tag{5.16}$$

By Proposition 5.5 and using polar coordinates,

$$\begin{aligned}
&\int_{D_\varepsilon\cap\{1<|x|_\rho<\sqrt{t/2}\}} p(t/2,a^*,x)^2 m_p(dx) \\
\lesssim\quad & \frac{1}{t^2}\int_{x\in D_\varepsilon\cap\{1<|x|_\rho<\sqrt{t/2}\}} \left(\log\left(\frac{\sqrt{t}}{|x|}\right)\right)^2 m_p(dx) \\
=\quad & \frac{2\pi}{t^2}\int_{1+\varepsilon}^{\sqrt{t/2}+\varepsilon} r\left(\log\left(\frac{\sqrt{t}}{r}\right)\right)^2 dr \\
\overset{u=r/\sqrt{t}}{=}\quad & \frac{2\pi}{t}\int_{(1+\varepsilon)/\sqrt{t}}^{(\sqrt{t/2}+\varepsilon)/\sqrt{t}} u(\log u)^2du \le \frac{c_7}{t}.
\end{aligned} \tag{5.17}$$

Combining (5.10),(5.11), (5.16) and (5.17), we conclude that $p(t,a^*,a^*)\le c_8/t$ for $t\ge 8$. On the other hand, taking $x=y=0=a^*$ in Theorem 4.6 yields that $p(t,a^*,a^*)\le c_9t^{-1/2}$ for $t\in(0,8]$. This completes the proof of the theorem. □

**Theorem 5.10.** *There is a constant $C_8\ge 1$ so that*

$$C_8^{-1}\left(t^{-1/2}\wedge t^{-1}\right) \le p(t,a^*,a^*) \le C_8\left(t^{-1/2}\wedge t^{-1}\right) \quad \textit{for all } t\in(0,\infty).$$

*Proof.* In view of Theorem 5.9, it remains to establish the lower bound estimate. By Cauchy-Schwartz inequality, for $M\ge 1$ to be determined later,

$$\begin{aligned}
p(t,a^*,a^*) &= \int_E p(t/2,a^*,x)^2 m_p(dx) \ge \int_{\{x\in E:|x|_\rho\le M\sqrt{t}\}} p(t/2,a^*,x)^2 m_p(dx) \\
&\ge \frac{1}{m_p(\{x\in E:|x|_\rho\le M\sqrt{t}\})}\left(\int_{\{x\in E:|x|_\rho\le M\sqrt{t}\}} p(t/2,a^*,x)m_p(dx)\right)^2 \\
&\gtrsim \left(t^{-1/2}\wedge t^{-1}\right)\mathbb{P}_{a^*}(|X_t|_\rho\le M\sqrt{t})^2.
\end{aligned} \tag{5.18}$$

We claim that by taking $M$ large enough, $\mathbb{P}_{a^*}(|X_t|_\rho \le M\sqrt{t}) \ge 1/2$ for every $t > 0$, which will then give the desired lower bound estimate on $p(t, a^*, a^*)$. Recall the signed radial process $Y = u(X)$ of $X$ from (4.1) satisfies SDE (4.6). For any $a > 0$ and $\delta \in (0, \varepsilon)$, let $Z^{\delta,a}$ and $\widetilde{Z}^{\delta,-a}$ be the pathwise unique solution of the following SDEs; see [5, Theorem 4.3]:

$$Z_t^{\delta,a} = a + B_t + \int_0^t \frac{1}{Z_s^{\delta,a} + \delta} \mathbf{1}_{\{Z_s^{\delta,a} > 0\}} ds + \widehat{L}_t^0(Z^{\delta,a}), \tag{5.19}$$

$$\widetilde{Z}_t^{\delta,-a} = -a + B_t + \int_0^t \frac{1}{\widetilde{Z}_s^{\delta,-a} - \delta} \mathbf{1}_{\{\widetilde{Z}_s^{\delta,-a} < 0\}} ds - \widehat{L}_t^0(\widetilde{Z}^{\delta,-a}), \tag{5.20}$$

where $B$ is the Brownian motion in (4.6), and $\widehat{L}^0(Z^{\delta,a})$, $\widehat{L}^0(\widetilde{Z}^{\delta,-a})$ are the symmetric local times of $Z^{\delta,a}$ and $Z^{\delta,-a}$ at 0, respectively. Denote by $Y^a$ and $Y^{-a}$ the pathwise solutions of (4.6) with $Y_0^a = a$ and $Y_0^{-a} = -a$, respectively. By comparison principle from [5, Theorem 4.6], we have with probability one that $Y_t^a \le Z_t^{\delta,a}$ for all $t \ge 0$ and $Y_t^{-a} \ge Z_t^{\delta,-a}$ for all $t \ge 0$. On the other hand, there are unique solutions to

$$dZ_t^a = a + B_t + \int_0^t \frac{1}{Z_s^a} ds, \qquad dZ_t^{-a} = -a + B_t + \int_0^t \frac{1}{Z_s^{-a}} ds.$$

In fact, $Z^a$ and $-Z^{-a}$ are both two-dimensional Bessel processes on $(0, \infty)$ starting from $a$. They have infinite lifetimes and never hits 0. By [5, Theorem 4.6] again, diffusion processes $Z^{\delta,a}$ is decreasing in $\delta$, and $\widetilde{Z}^{\delta,-a}$ is increasing in $\delta$. It is easy to see from the above facts that $\lim_{\delta\to 0} Z_t^{\delta,a} = Z_t^a$ and $\lim_{\delta\to 0} \widetilde{Z}_t^{\delta,-a} = Z_t^{-a}$. Consequently, with probability one,

$$Y_t^a \le Z_t^a \quad \text{and} \quad Y_t^{-a} \ge Z_t^{-a} \qquad \text{for every } t \ge 0.$$

In particular, we have for every $t > 0$, $\mathbb{P}(Y_t^a \ge M\sqrt{t}) \le \mathbb{P}(Z_t^a \ge M\sqrt{t})$ and

$$\mathbb{P}(Y_t^{-a} \le -M\sqrt{t}) \le \mathbb{P}(Z_t^{-a} \le -M\sqrt{t}) = \mathbb{P}(Z_t^a \ge M\sqrt{t}).$$

Let $W$ be two-dimensional Brownian motion. Then we have from the above that for every $t > 0$,

$$\mathbb{P}_a(Y_t \ge M\sqrt{t}) + \mathbb{P}_{-a}(Y_t \le -M\sqrt{t}) \le 2\mathbb{P}_{(a,0)}(|W_t| \ge M\sqrt{t}).$$

Passing $a \to 0$ yields, by the Brownian scaling property, that

$$\mathbb{P}_{a^*}(|X_t|_\rho \ge M\sqrt{t}) = \mathbb{P}_0(|Y_t| \ge M\sqrt{t}) \le 2\mathbb{P}_0(|W_t| \ge M\sqrt{t}) = 2\mathbb{P}_0(|W_1| \ge M),$$

which is less than 1/2 by choosing $M$ large. This completes the proof of the theorem. □

In the next two propositions, we use the two-sided estimate for $p(t, a^*, a^*)$ as well as Markov property of $X$ to get two-sided bounds for $p(t, a^*, x)$ for different regions of $x$. We first record an elementary lemma that will be used later.

**Lemma 5.11.** *For every $x > 0$,*

$$\frac{1}{1+x} e^{-x^2/2} \le \int_x^\infty e^{-y^2/2} dy \le \frac{e\pi}{1+x} e^{-x^2/2}.$$

*Proof.* Define $\phi(x) = \int_x^\infty e^{-y^2/2}dy - \frac{1}{1+x}e^{-x^2/2}$. Then $\phi'(x) = -\frac{x}{(1+x)^2}e^{-x^2/2} < 0$. Since $\lim_{x\to\infty}\phi(x) = 0$, we have $\phi(x) > 0$ for every $x > 0$. This establishes the lower bound estimate of the lemma. For the upper bound, note that for $x \in (0,1)$,

$$\int_x^\infty e^{-y^2/2}\,dy \le \frac{1}{2}\int_{-\infty}^\infty e^{-y^2/2}\,dy = \sqrt{\frac{\pi}{2}} \le \sqrt{e\pi/2}\,e^{-x^2/2},$$

while for every $x > 0$, using a change of variable $y = x + z$,

$$\int_x^\infty e^{-y^2/2}\,dy \le e^{-x^2/2}\int_0^\infty e^{-xz}\,dz = x^{-1}e^{-x^2/2}.$$

This establishes the upper bound estimate of the lemma. □

**Proposition 5.12.** *There exist constants $C_i > 0$, $9 \le i \le 10$, so that for all $x \in \mathbb{R}_+$ and $t \ge 8$,*

$$\frac{C_9}{t}\left(1 + \frac{|x|\log t}{\sqrt{t}}\right)e^{-2x^2/t} \le p(t,a^*,x) \le \frac{C_{10}}{t}\left(1 + \frac{|x|\log t}{\sqrt{t}}\right)e^{-x^2/2t}.$$

*Proof.* Observing that when $x \in \mathbb{R}_+$, we have by (5.8),

$$\begin{aligned} p(t,a^*,x) = p(t,x,a^*) &= \int_0^t p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) \\ &= \int_0^{t/2} p(t-s,a^*,a^*)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds + \int_{t/2}^t p(t-s,a^*,a^*)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds. \quad (5.21) \end{aligned}$$

By Theorem 5.10, Lemma 5.11 and a change of variable $r = x/\sqrt{s}$,

$$\begin{aligned} \int_0^{t/2} p(t-s,a^*,a^*)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds &\asymp \frac{1}{t}\int_0^{t/2}\frac{x}{s^{3/2}}e^{-x^2/2s}\,ds = \frac{2}{t}\int_{x\sqrt{2/t}}^\infty e^{-r^2/2}\,dr \\ &\asymp \frac{1}{t}\frac{1}{1+(x/\sqrt{t})}e^{-x^2/t}, \end{aligned}$$

while

$$\begin{aligned} \frac{x}{\sqrt{2\pi t^3}}e^{-x^2/t}\int_0^{t/2} p(r,a^*,a^*)dr &\le \int_{t/2}^t p(t-s,a^*,a^*)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/2s}\,ds \\ &\le \frac{2x}{\sqrt{\pi t^3}}e^{-x^2/2t}\int_0^{t/2} p(r,a^*,a^*)dr. \end{aligned}$$

This, Theorem 5.10 and (5.21) yields the desired result. □

**Proposition 5.13.** *There exist $C_i > 0$, $11 \le i \le 14$ such that*

$$\frac{C_{11}}{t}e^{-C_{12}|x|_\rho^2/t} \le p(t,a^*,x) \le \frac{C_{13}}{t}e^{-C_{14}|x|_\rho^2/t} \qquad \textit{for all } t \ge 8 \textit{ and } x \in D_\varepsilon.$$

*Proof.* When $t \in [1,8]$, the estimates follows from Theorem 4.8. So it remains to establish the estimates for $t > 8$. We do this by considering three cases. Note that $p(t,a^*,x) = p(t,x,a^*)$.
*Case 1.* $1 \le |x|_\rho < 2\sqrt{t}$. We have by Theorem 5.10 and Proposition 5.1,

$$\begin{aligned} p(t,x,a^*) &= \int_0^{t/2} p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) + \int_{t/2}^t p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*} \in ds) \\ &\asymp \frac{1}{t}\mathbb{P}_x(\sigma_{a^*} \le t/2) + \frac{\log|x|}{t(\log t)^2}\int_0^{t/2} p(s,a^*,a^*)ds \\ &\asymp \frac{1}{t}\left(1 - \frac{\log|x|}{\log\sqrt{t}}\right) + \frac{\log|x|}{t(\log t)^2}\int_0^{t/2}\left(\frac{1}{\sqrt{s}} \wedge \frac{1}{s}\right)ds \\ &\asymp \frac{1}{t}\left(1 - \frac{\log|x|}{\log\sqrt{t}}\right) + \frac{\log|x|}{t\log t} \asymp \frac{1}{t}. \end{aligned}$$

*Case 2.* $|x|_\rho \ge 2\sqrt{t}$. In the following computation, $B_\rho(x,r) := \{y \in E : \rho(x,y) < r\}$, and $B_e(x,r) := \{y \in E : |y-x| < r\}$. We denote by $\{W; \mathbb{P}^0_x, x \in \mathbb{R}^2\}$ two-dimensional Brownian motion and $p^0(t,x,y) = (2\pi t)^{-1}\exp(-|x-y|^2/2t)$ its transition density. Since

$$p(t,x,a^*) = \mathbb{E}_x\left[p(t-\sigma_{B_\rho(a^*,2)}, X_{\sigma_{B_\rho(a^*,2)}}, a^*); \sigma_{B_\rho(a^*,2)} < t\right],$$

it follows from Case 1, Theorem 4.6 and Theorem 4.8 that there is $c_1 \ge 1$ so that

$$\begin{aligned} p(t,x,a^*) &\lesssim \mathbb{E}_x\left[\left(t-\sigma_{B_\rho(a^*,2)}\right)^{-1} e^{-\frac{(2+2\varepsilon)^2}{2c_1(t-\sigma_{B_\rho(a^*,2)})}}; \sigma_{B_\rho(a^*,2)} < t\right] \\ &= \mathbb{E}^0_x\left[\left(t-\sigma_{B_e(0,2+\varepsilon)}\right)^{-1} e^{-\frac{(2+2\varepsilon)^2}{2c_1(t-\sigma_{B_e(0,2+\varepsilon)})}}; \sigma_{B_e(0,2+\varepsilon)} < t\right] \\ &\lesssim \mathbb{E}^0_x\left[\left(c_1t-\sigma_{B_e(0,2+\varepsilon)}\right)^{-1} e^{-\frac{(2+\varepsilon)^2}{2(c_1t-\sigma_{B_e(0,2+\varepsilon)})}}; \sigma_{B_e(0,2+\varepsilon)} < t\right] \\ &\le \mathbb{E}^0_x\left[\left(c_1t-\sigma_{B_e(0,2+\varepsilon)}\right)^{-1} e^{-\frac{(2+\varepsilon)^2}{2(c_1t-\sigma_{B_e(0,2+\varepsilon)})}}; \sigma_{B_e(0,2+\varepsilon)} < c_1t\right] \\ &\asymp \mathbb{E}^0_x\left[p^0(c_1t-\sigma_{B_e(0,2+\varepsilon)}), W_{\sigma_{B_e(0,2+\varepsilon)}}, 0); \sigma_{B_e(0,2+\varepsilon)} < c_1t\right] \\ &= p^0(c_1t,x,0) = (2\pi c_1 t)^{-1}\exp(-|x|^2/2c_1t). \end{aligned}$$

Similarly, for the lower bound estimate, it follows from Case 1, Theorem 4.6 and Theorem 4.8

that there is $c_2 \in (0,1]$ so that

$$\begin{aligned}
p(t,x,a^*) &\gtrsim \mathbb{E}_x\left[\left(t-\sigma_{B_\rho(a^*,2)}\right)^{-1} e^{-\frac{2^2}{2c_2(t-\sigma_{B_\rho(a^*,2)})}}; \sigma_{B_\rho(a^*,2)} < t\right] \\
&\geq \mathbb{E}^0_x\left[\left(t-\sigma_{B_e(0,2+\varepsilon)}\right)^{-1} e^{-\frac{2^2}{2c_2(t-\sigma_{B_e(0,2+\varepsilon)})}}; \sigma_{B_e(0,2+\varepsilon)} < c_2 t\right] \\
&\gtrsim \mathbb{E}^0_x\left[\left(c_2 t-\sigma_{B_e(0,2+\varepsilon)}\right)^{-1} e^{-\frac{(2+\varepsilon)^2}{2(c_2 t-\sigma_{B_e(0,2+\varepsilon)})}}; \sigma_{B_e(0,2+\varepsilon)} < c_2 t\right] \\
&\asymp \mathbb{E}^0_x\left[p^0(c_2 t-\sigma_{B_e(0,2+\varepsilon)}), W_{\sigma_{B_e(0,2+\varepsilon)}}, 0); \sigma_{B_e(0,2+\varepsilon)} < c_2 t\right] \\
&= p^0(c_2 t, x, 0) = (2\pi c_2 t)^{-1}\exp(-|x|^2/2c_2 t).
\end{aligned}$$

Realizing that $|x|_\rho > 2\sqrt{t} > 4\sqrt{2}$ implies that $|x|_\rho \asymp |x|$, we get the desired estimates in this case.

*Case 3.* $0 < |x|_\rho < 1$. Note that by Proposition 3.2,

$$\int_{y\in D_\varepsilon\cap B_\rho(a^*,2)} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) \leq \int_{D_\varepsilon\cap B_\rho(a^*,2)} \left(\frac{c_1}{\sqrt{t}}\right)^2 m_p(dy) \asymp \frac{1}{t}, \tag{5.22}$$

while by Cases 1 and 2 above,

$$\int_{D_\varepsilon\cap B_\rho(a^*,2)^c} p(t/2,x,y)p(t/2,a^*,y)m_p(dy) \leq \sup_{y\in D_\varepsilon\cap B_\rho(a^*,2)^c} p(t/2,a^*,y) \lesssim \frac{1}{t}. \tag{5.23}$$

On the other hand, by Theorem 5.10 again,

$$\begin{aligned}
\int_{\mathbb{R}_+} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) &= \mathbb{E}_x\left[p(t/2,X_{t/2},a^*); X_{t/2}\in\mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[p(t/2,X_{t/2},a^*); \sigma_{a^*} < t/2 \text{ and } X_{t/2}\in\mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_{a^*}[p(t/2,X_{t/2-s},a^*)]|_{s=\sigma_{a^*}}; \sigma_{a^*} < t/2 \text{ and } X_{t/2}\in\mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[p(t-\sigma_{a^*},a^*,a^*); \sigma_{a^*} < t/2 \text{ and } X_{t/2}\in\mathbb{R}_+\right] \\
&\asymp \frac{1}{t}\,\mathbb{P}_x\left(\sigma_{a^*} \leq t/2 \text{ and } X_{t/2}\in\mathbb{R}_+\right) \leq \frac{1}{t}.
\end{aligned} \tag{5.24}$$

The estimates (5.22), (5.23) and (5.24) imply that

$$\begin{aligned}
p(t,a^*,x) = &\int_{D_\varepsilon\cap B_\rho(a^*,2)} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) + \int_{D_\varepsilon\cap B^c_\rho(a^*,2)} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) \\
&+ \int_{\mathbb{R}_+} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) \lesssim \frac{1}{t}.
\end{aligned}$$

On the other hand, there is a constant $c_3 > 0$ so that $\mathbb{P}_x(\sigma_{a^*} \leq 1) \geq c_3$ for all $x \in D_\varepsilon$ with $|x|_\rho \leq 1$. Hence we have by Theorem 5.10 that for $t \geq 2$ and $x \in D_\varepsilon$ with $|x|_\rho \leq 1$,

$$p(t,a^*,x) = p(t,x,a^*) \geq \int_0^1 p(t-s,a^*,a^*)\mathbb{P}_x(\sigma_{a^*}\in ds) \gtrsim \frac{1}{t}\mathbb{P}_x(\sigma_{a^*}\leq 1) \gtrsim \frac{1}{t}.$$

In conclusion, we have $p(t,a^*,x) \asymp \frac{1}{t}$ for $t \geq 4$ and $x \in D_\varepsilon$ with $|x|_\rho \leq 1$. This completes the proof of the proposition. □

We are now in the position to derive estimates on $p(t,x,y)$ for $(x,y) \in D_\varepsilon \times D$ and $t \in [8,\infty)$, by using the two-sided estimate of $p(t,x,a^*)$ and the Markov property of $X$.

**Theorem 5.14.** *There exist constants $C_i > 0$, $15 \le i \le 18$, such that the following estimate holds:*

$$\frac{C_{15}}{t} e^{-C_{16}\rho(x,y)^2/t} \le p(t,x,y) \le \frac{C_{17}}{t} e^{-C_{18}\rho(x,y)^2/t}, \quad (t,x,y) \in [8,\infty) \times D_\varepsilon \times D_\varepsilon.$$

*Proof.* As before, denote by $\{W; \mathbb{P}^0_x, x \in \mathbb{R}^2\}$ two-dimensional Brownian motion and $p^0(t,x,y) = (2\pi t)^{-1}\exp(-|x-y|^2/2t)$ its transition density. We first note that, as a special case of [36, Theorem 1.1(a)], there are constants $c_1 > c_2 > 0$ so that for $t \ge 1$ and $x, y \in D_\varepsilon$,

$$(|x|_\rho \wedge 1)\,(|y|_\rho \wedge 1)\, t^{-1} e^{-c_1|x-y|^2/t} \lesssim p^0_{D_\varepsilon}(t,x,y) \lesssim (|x|_\rho \wedge 1)\,(|y|_\rho \wedge 1)\, t^{-1} e^{-c_2|x-y|^2/t}.$$

It follows that there is $c_3 \in (0,1]$ so that

$$p^0_{D_\varepsilon}(t,x,y) \gtrsim p^0_{D_\varepsilon}(c_3 t,x,y) \quad \text{for every } t \ge 1 \text{ and } x,y \in D_\varepsilon. \tag{5.25}$$

We will prove the theorem by considering two different cases.

*Case 1.* $\max\{|x|_\rho, |y|_\rho\} > 1$. Without loss of generality, we assume $|y|_\rho > 1$. In this case, it is not hard to verify that $\rho(x,y) \asymp |x-y|$. Recall from (1.6)-(1.7) that for $x,y \in D_\varepsilon$,

$$\overline{p}_{D_\varepsilon}(t,x,y) := p(t,x,y) - p_{D_\varepsilon}(t,x,y) = \mathbb{E}_x\left[p(t-\sigma_{a^*}, a^*, y); \sigma_{a^*} < t\right].$$

By Proposition 5.13 and the assumption that $\varepsilon \in (0,1/4]$, there are constants $c_4 \ge 1$ so that for every $x,y \in D_\varepsilon$ with $|y|_\rho > 1$

$$\begin{aligned}
\overline{p}_{D_\varepsilon}(t,x,y) &\lesssim \int_0^t \frac{1}{t-s} e^{-\frac{(|y|_\rho+3\varepsilon)^2}{2c_4(t-s)}}\, \mathbb{P}_x\left(\sigma_{a^*} \in ds\right) \\
&\le \int_0^{c_1 t} \frac{1}{c_4 t - s} e^{-\frac{(|y|_\rho+2\varepsilon)^2}{2(c_4 t-s)}}\, \mathbb{P}^0_x\left(\sigma_{B_e(0,\varepsilon)} \in ds\right) \\
&\lesssim \mathbb{E}^0_x\left[p^0(c_4 t - s, W_{\sigma_{B_e(0,\varepsilon)}}, y); \sigma_{B_e(0,\varepsilon)} < c_4 t\right] \\
&\le p^0(c_4 t, x, y).
\end{aligned} \tag{5.26}$$

Similarly, there is a constant $c_5 \in (0, c_3]$ so that

$$\begin{aligned}
\overline{p}_{D_\varepsilon}(t,x,y) &\gtrsim \int_0^t \frac{1}{t-s} e^{-\frac{(|y|_\rho-\varepsilon)^2}{2c_5(t-s)}}\, \mathbb{P}_x\left(\sigma_{a^*} \in ds\right) \\
&\ge \int_0^{c_5 t} \frac{1}{c_5 t - s} e^{-\frac{|y|_\rho^2}{2(c_5 t-s)}}\, \mathbb{P}^0_x\left(\sigma_{B_e(0,\varepsilon)} \in ds\right) \\
&\gtrsim \mathbb{E}^0_x\left[p^0(c_5 t - s, W_{\sigma_{B_e(0,\varepsilon)}}, y); \sigma_{B_e(0,\varepsilon)} < c_5 t\right] \\
&= \overline{p}^0_{D_\varepsilon}(c_5 t, x, y).
\end{aligned} \tag{5.27}$$

Since $p_{D_\varepsilon}(t,x,y) = p^0_{D_\varepsilon}(t,x,y)$ and $c_4 \ge 1$, we have from (5.26) that

$$p(t,x,y) = p_{D_\varepsilon}(t,x,y) + \overline{p}_{D_\varepsilon}(t,x,y) \lesssim p^0(t,x,y) + p^0(c_4 t,x,y) \lesssim p^0(c_4 t,x,y) \lesssim t^{-1} e^{-\rho(x,y)^2/2c_4 t}.$$

On the other hand, we have by (5.25) and (5.27) that for every $t \geq 1$ and $x, y \in D_\varepsilon$ satisfying $|x|_\rho + |y|_\rho > 2$,

$$p(t,x,y) \gtrsim p^0_{D_\varepsilon}(c_5 t, x, y) + \overline{p}^0_{D_\varepsilon}(c_5 t, x, y) = p^0(c_5 t, x, y) \gtrsim t^{-1} e^{-|x-y|^2/2c_5 t} \gtrsim t^{-1} e^{-\rho(x,y)^2/c_5 t},$$

where the last "$\gtrsim$" is due to the fact that $|x-y|/\sqrt{2} \leq \rho(x,y)$, which can be verified easily from the assumptions that $|x|_\rho + |y|_\rho > 2$ and $\varepsilon \leq 1/4$. This establishes the desired two-sided estimates in this case.

*Case 2.* $\max\{|x|_\rho, |y|_\rho\} \leq 1$. In this case, it suffices to show that $p(t,x,y) \asymp t^{-1}$ for $t \geq 8$. The proof is similar to that of Case 3 of Proposition 5.13. By Proposition 3.2, for $t \geq 8$,

$$\int_{y \in D_\varepsilon \cap B_\rho(a^*,2)} p(t/2,x,z)p(t/2,z,y)m_p(dy) \leq \int_{D_\varepsilon \cap B_\rho(a^*,2)} \left(\frac{c_6}{\sqrt{t}}\right)^2 m_p(dy) \asymp \frac{1}{t}, \tag{5.28}$$

while by Case 1,

$$\int_{D_\varepsilon \cap B(a^*,2)^c} p(t/2,x,z)p(t/2,z,y)m_p(dz) \leq \sup_{z \in D_\varepsilon \cap B(a^*,2)^c} p(t/2,y,z) \lesssim \frac{1}{t}. \tag{5.29}$$

On the other hand, by Proposition 5.13, for $t \geq 8$,

$$\begin{aligned}
\int_{\mathbb{R}_+} p(t/2,x,z)p(t/2,z,y)m_p(dz) &= \mathbb{E}_x\left[p(t/2, X_{t/2}, y); X_{t/2} \in \mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[p(t/2, X_{t/2}, y); \sigma_{a^*} < t/2 \text{ and } X_{t/2} \in \mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_{a^*}[p(t/2, X_{t/2-s}, y)]|_{s=\sigma_{a^*}}; \sigma_{a^*} < t/2 \text{ and } X_{t/2} \in \mathbb{R}_+\right] \\
&= \mathbb{E}_x\left[p(t-\sigma_{a^*}, a^*, y); \sigma_{a^*} < t/2 \text{ and } X_{t/2} \in \mathbb{R}_+\right] \\
&\asymp \frac{1}{t}\,\mathbb{P}_x\left(\sigma_{a^*} \leq t/2 \text{ and } X_{t/2} \in \mathbb{R}_+\right) \leq \frac{1}{t}.
\end{aligned} \tag{5.30}$$

The estimates (5.28)-(5.30) imply that

$$\begin{aligned}
p(t,a^*,x) = &\int_{D_\varepsilon \cap B_\rho(a^*,2)} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) + \int_{D_\varepsilon \cap B^c_\rho(a^*,2)} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) \\
&+ \int_{\mathbb{R}_+} p(t/2,a^*,y)p(t/2,y,x)m_p(dy) \\
\lesssim &\frac{1}{t}.
\end{aligned}$$

On the other hand, there is a constant $c_7 > 0$ so that $\mathbb{P}_x(\sigma_{a^*} \leq 1) \geq c_7$ for all $x \in D_\varepsilon$ with $|x|_\rho \leq 2$. Hence we have by Proposition 5.13 that for $t \geq 2$ and $x \in D_\varepsilon$ with $|x|_\rho \leq 1$,

$$p(t,x,y) \geq \int_0^1 p(t-s,a^*,y)\mathbb{P}_x(\sigma_{a^*} \in ds) \gtrsim \frac{1}{t}\mathbb{P}_x(\sigma_{a^*} \leq 1) \gtrsim \frac{1}{t}.$$

Therefore we have $p(t,x,y) \asymp \frac{1}{t}$ for $t \geq 8$ and $x, y \in D_\varepsilon$ with $|x|_\rho + |y|_\rho \leq 2$. This completes the proof of the theorem. □

**Theorem 5.15.** *There exist constants $C_i > 0$, $19 \le i \le 26$, such that the following estimates hold for $(t, x, y) \in [4, \infty) \times \mathbb{R}_+ \times D_\varepsilon$: when $|y|_\rho < 1$,*

$$\frac{C_{19}}{t}\left(1 + \frac{|x|\log t}{\sqrt{t}}\right) e^{-C_{20}\rho(x,y)^2/t} \le p(t,x,y) \le \frac{C_{21}}{t}\left(1 + \frac{|x|\log t}{\sqrt{t}}\right) e^{-C_{22}\rho(x,y)^2/t}; \tag{5.31}$$

*while for $|y|_\rho \ge 1$,*

$$\begin{aligned} & \frac{C_{23}}{t}\left(1 + \frac{|x|}{\sqrt{t}}\log\left(1 + \frac{\sqrt{t}}{|y|_\rho}\right)\right) e^{-C_{24}\rho(x,y)^2/t} \\ \le \quad & p(t,x,y) \le \frac{C_{25}}{t}\left(1 + \frac{|x|}{\sqrt{t}}\log\left(1 + \frac{\sqrt{t}}{|y|_\rho}\right)\right) e^{-C_{26}\rho(x,y)^2/t}. \end{aligned} \tag{5.32}$$

*Proof.* First, note that by Proposition 5.13 and Theorem 4.8,

$$\frac{1}{t}e^{-c_1|y|_\rho^2/t} \lesssim p(t,a^*,y) \lesssim \frac{1}{t}e^{-c_2|y|_\rho^2/t} \quad \text{for } t \ge 1 \text{ and } y \in D_\varepsilon \tag{5.33}$$

and

$$\frac{1}{\sqrt{t}}e^{-c_3|y|_\rho^2/t} \lesssim p(t,a^*,y) \lesssim \frac{1}{\sqrt{t}}e^{-c_4|y|_\rho^2/t} \quad \text{for } t \le 1 \text{ and } y \in D_\varepsilon \text{ with } |y|_\rho \le 1. \tag{5.34}$$

By (5.8),

$$p(t,x,y) = \int_0^t p(t-s,a^*,y)\mathbb{P}_x(\sigma_{a^*} \in ds) = \int_0^t p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds. \tag{5.35}$$

It follows from (5.33) and Lemma 5.11 that for every $y \in D_\varepsilon$ and $t \ge 4$,

$$\begin{aligned} & \int_0^{t/2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \lesssim -\frac{1}{t}e^{-c_2|y|_\rho^2/t}\int_{s=0}^{t/2} e^{-\frac{|x|^2}{2s}}\,d\left(\frac{|x|}{\sqrt{s}}\right) \\ \lesssim \quad & \frac{1}{t}e^{-c_2|y|_\rho^2/t}\frac{1}{1+|x|/\sqrt{t}}e^{-|x|^2/t} \le \frac{1}{t}e^{-c_3\rho(x,y)^2/t}. \end{aligned} \tag{5.36}$$

Similarly, we have

$$\int_0^{t/2} p(t-s,a^*,y)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \gtrsim \frac{1}{t}e^{-(|x|^2+c_4|y|_\rho^2)/t} \gtrsim \frac{1}{t}e^{-c_5\rho(x,y)^2/t}. \tag{5.37}$$

We now consider two cases depending on the range of the values of $|y|_\rho$.
*Case 1.* $y \in D_\varepsilon$ with $|y|_\rho < 1$. In this case, we have by (5.33)

$$\begin{aligned} & \int_{t/2}^{t-1} p(t-s,a^*,y)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \lesssim \int_{t/2}^{t-1}\frac{1}{t-s}e^{-c_2|y|_\rho^2/(t-s)}\frac{|x|}{\sqrt{s^3}}e^{-x^2/(2s)}ds \\ \lesssim \quad & \frac{|x|}{t^{3/2}}e^{-|x|^2/2t}\int_{t/2}^{t-1}\frac{1}{t-s}ds \lesssim \frac{|x|\log t}{t^{3/2}}e^{-|x|^2/2t}. \end{aligned}$$

Similarly, we have

$$\int_{t/2}^{t-1} p(t-s,a^*,y)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \gtrsim \frac{|x|\log t}{t^{3/2}}e^{-|x|^2/t}.$$

On the other hand, by (5.34),

$$\int_{t-1}^{t} p(t-s,a^*,y)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \lesssim \frac{|x|}{\sqrt{t^3}}e^{-x^2/(2t}\int_{t-1}^{t}\frac{1}{\sqrt{t-s}}e^{-c_4|y|_\rho^2/(t-s)}ds$$
$$\lesssim \frac{|x|}{\sqrt{t^3}}e^{-x^2/(2t)}\int_{t-1}^{t}\frac{1}{\sqrt{t-s}}ds \lesssim \frac{|x|}{\sqrt{t^3}}e^{-x^2/(2t)},$$

and similarly

$$\int_{t-1}^{t} p(t-s,a^*,y)\frac{x}{\sqrt{2\pi s^3}}e^{-x^2/(2s)}ds \gtrsim \frac{|x|}{\sqrt{t^3}}e^{-x^2/t}.$$

These estimates together with (5.35)-(5.37) establishes (5.31).

*Case 2.* $y \in D_\varepsilon$ with $|y|_\rho > 1$. Note that by (5.8),

$$\begin{aligned} p(t,x,y) &= \int_0^t p(t-s,a^*,y)\mathbb{P}_x(\sigma_{a^*}\in ds) \qquad (5.38)\\ &= \int_0^{t/2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds + \int_{t/2}^{t} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds. \end{aligned}$$

By Theorem 4.8, (5.33) and a change of variable $r = |y|_\rho/\sqrt{t-s}$, we have

$$\begin{aligned} \int_{t/2}^{t} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds &\lesssim \int_{t/2}^{t}\frac{1}{t-s}e^{-c_6|y|_\rho^2/(t-s)}\frac{|x|}{\sqrt{s^3}}e^{-x^2/2s}ds\\ &\lesssim \frac{|x|}{t^{3/2}}e^{-x^2/2t}\int_{t/2}^{t}\frac{1}{t-s}e^{-c_6|y|_\rho^2/(t-s)}ds\\ &= \frac{|x|}{t^{3/2}}e^{-x^2/2t}\int_{|y|_\rho/\sqrt{2/t}}^{\infty}\frac{2}{r}e^{-c_6r^2}dr. \end{aligned}$$

Note that for each fixed $a>0$, $\int_\lambda^\infty r^{-1}e^{-ar^2}dr = \int_\lambda^1 r^{-1}e^{-ar^2}dr + \int_1^\infty r^{-1}e^{-c_2r^2}dr$ is comparable to $\log(1/\lambda)$ when $0<\lambda\le 1/2$. For $\lambda \ge 1/2$, by Lemma 5.11

$$\int_\lambda^\infty r^{-1}e^{-ar^2}dr \le 2\int_\lambda^\infty e^{-ar^2}dr = \frac{2}{\sqrt{2a}}\int_{\sqrt{2a}\lambda}^\infty e^{-s^2/2}ds \lesssim \frac{1}{1+\sqrt{a}\lambda}e^{-a\lambda^2} \le e^{-a\lambda^2}$$

and

$$\int_\lambda^\infty r^{-1}e^{-ar^2}dr \gtrsim \int_\lambda^\infty e^{-2ar^2}dr = \frac{1}{2\sqrt{a}}\int_{2\sqrt{a}\lambda}^\infty e^{-s^2/2}ds \gtrsim \frac{1}{1+\sqrt{a}\lambda}e^{-2a\lambda^2} \gtrsim e^{-3a\lambda^2}$$

Hence we have

$$\log(1+\lambda^{-1})e^{-3a\lambda^2} \lesssim \int_\lambda^\infty r^{-1}e^{-ar^2}dr \le \log(1+\lambda^{-1})e^{-a\lambda^2} \quad \text{for any } \lambda>0. \qquad (5.39)$$

Thus we have

$$\begin{aligned} \int_{t/2}^{t} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds &\lesssim \frac{|x|}{t^{3/2}}e^{-x^2/2t}\log\left(1+\frac{\sqrt{t}}{|y|_\rho}\right)e^{-2c_6|y|_\rho^2/t}\\ &\le \frac{|x|}{t^{3/2}}\log\left(1+\frac{\sqrt{t}}{|y|_\rho}\right)e^{-c_7\rho(x,y)^2/t}. \end{aligned}$$

Similarly, we have

$$\int_{t/2}^{t} p(t-s, a^*, y) \frac{|x|}{\sqrt{2\pi s^3}} e^{-x^2/2s} ds \gtrsim \frac{|x|}{t^{3/2}} \log\left(1 + \frac{\sqrt{t}}{|y|_\rho}\right) e^{-c_8 \rho(x,y)^2/t}.$$

These together with (5.36)-(5.37) establishes (5.32). □

We will need the following elementary lemma.

**Lemma 5.16.** *For every $c > 0$, there exists $C_{27} \geq 1$ such that for every $t \geq 3$ and $0 < y \leq \sqrt{t}$,*

$$C_{27}^{-1} \log t \leq \int_2^t \frac{1}{s}\left(1 + \frac{y \log s}{\sqrt{s}}\right) e^{-c|y|^2/s} ds \leq C_{27} \log t.$$

*Proof.* By a change of variable $r = y/\sqrt{s}$,

$$\int_2^t \frac{y \log s}{s^{3/2}} e^{-c|y|^2/s} ds \leq \log t \int_2^t \frac{y}{s^{3/2}} e^{-c|y|^2/s} ds = 2 \log t \int_{y/\sqrt{t}}^{y/\sqrt{2}} e^{-cr^2} dr \lesssim \log t,$$

while since $0 < y \leq \sqrt{t}$,

$$\int_2^t \frac{1}{s} e^{-c|y|^2/s} ds \asymp \int_2^t \frac{1}{s} ds \asymp \log t.$$

This proves the lemma. □

**Theorem 5.17.** *There exist constants $C_i > 0$, $28 \leq i \leq 31$, such that the following estimate holds for all $(t, x, y) \in [8, \infty) \times \mathbb{R}_+ \times \mathbb{R}_+$:*

$$\begin{aligned} &\frac{C_{28}}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-C_{29}|x-y|^2/t} + \frac{C_{28}}{t}\left(1 + \frac{(|x|+|y|)\log t}{\sqrt{t}}\right) e^{-2(x^2+y^2)/t} \leq p(t,x,y) \\ &\leq \frac{C_{30}}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-C_{31}|x-y|^2/t} + \frac{C_{30}}{t}\left(1 + \frac{(|x|+|y|)\log t}{\sqrt{t}}\right) e^{-(x^2+y^2)/2t}. \end{aligned} \tag{5.40}$$

*Proof.* When either $x = a^*$ or $y = a^*$, this has been established in Proposition 5.12 so we assume $|x| \wedge |y| > 0$. For simplicity, denote $\mathbb{R}_+ \setminus \{a^*\}$ by $(0, \infty)$. Since

$$p_{(0,\infty)}(t,x,y) = (2\pi t)^{-1/2}\left(e^{-|x-y|^2/2t} - e^{-|x+y|^2/2t}\right) = (2\pi t)^{-1/2} e^{-|x-y|^2/2t}\left(1 - e^{-2xy/t}\right),$$

there are constants $c_1 > c_2 > 0$ so that

$$\frac{1}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-c_1|x-y|^2/t} \lesssim p_{(0,\infty)}(t,x,y) \lesssim \frac{1}{\sqrt{t}}\left(1 \wedge \frac{|x|}{\sqrt{t}}\right)\left(1 \wedge \frac{|y|}{\sqrt{t}}\right) e^{-c_2|x-y|^2/t}. \tag{5.41}$$

for all $t > 0$ and $x, y \in \mathbb{R}_+$. Note also

$$p(t,x,y) = p_{(0,\infty)}(t,x,y) + \int_0^t p(t-s, a^*, y) \mathbb{P}_x(\sigma_{a^*} \in ds). \tag{5.42}$$

We prove this theorem by considering two cases.
*Case 1.* $|x| \wedge |y| \geq \sqrt{t}$. In this case , $p(t,x,y) \geq p_{(0,\infty)}(t,x,y) \gtrsim t^{-1/2} e^{-c_3|x-y|^2/t}$. Thus we have by Proposition 3.4,

$$\frac{1}{\sqrt{t}} e^{-c_3|x-y|^2/t} \lesssim p(t,x,y) \lesssim \frac{1}{\sqrt{t}} e^{-c_4|x-y|^2/t}.$$

So estimates (5.40) holds in this case.

*Case 2.* $0 < |x| \wedge |y| < \sqrt{t}$. Without loss of generality, we may and do assume $|y| < \sqrt{t} \wedge |x|$. By (5.8),

$$\int_0^t p(t-s,a^*,y)\mathbb{P}_x(\sigma_{a^*} \in ds) = \int_0^t p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds. \tag{5.43}$$

By Proposition 5.12 and Lemma 5.11

$$\begin{aligned}
\int_0^{t/2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \quad &\lesssim \quad \int_0^{t/2} \frac{1}{t-s}\left(1+\frac{|y|\log(t-s)}{\sqrt{t-s}}\right)e^{-y^2/2(t-s)}\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \\
&\lesssim \quad \frac{1}{t}\left(1+\frac{|y|\log t}{\sqrt{t}}\right)\int_0^{t/2}\frac{|x|}{s^{3/2}}e^{-x^2/2s}ds \\
&\overset{r=|x|/\sqrt{s}}{\lesssim} \quad \frac{1}{t}\left(1+\frac{|y|\log t}{\sqrt{t}}\right)\int_{\sqrt{2}|x|/\sqrt{t}}^{\infty} e^{-r^2/2}dr \\
&\lesssim \quad \frac{1}{t}\left(1+\frac{|y|\log t}{\sqrt{t}}\right)\frac{1}{1+|x|/\sqrt{t}}e^{-x^2/t} \\
&\asymp \quad \frac{1}{t}\left(1+\frac{|y|\log t}{\sqrt{t}}\right)e^{-(x^2+y^2)/t},
\end{aligned}$$

while by Proposition 5.12 and Lemma 5.16,

$$\begin{aligned}
\int_{t/2}^{t-2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \quad &\lesssim \quad \int_{t/2}^{t-2} \frac{1}{t-s}\left(1+\frac{|y|\log(t-s)}{\sqrt{t-s}}\right)e^{-y^2/2(t-s)}\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \\
&\lesssim \quad \frac{|x|}{t^{3/2}}e^{-x^2/2t}\int_{t/2}^{t-2}\frac{1}{t-s}\left(1+\frac{|y|\log(t-s)}{\sqrt{t-s}}\right)e^{-y^2/2(t-s)}ds \\
&\overset{r=t-s}{=} \quad \frac{|x|}{t^{3/2}}e^{-x^2/2t}\int_2^{t/2}\frac{1}{r}\left(1+\frac{|y|\log r}{\sqrt{r}}\right)e^{-y^2/2r}dr \\
&\lesssim \quad \frac{|x|}{t^{3/2}}e^{-x^2/2t}\log t \asymp \frac{|x|\log t}{t^{3/2}}e^{-(x^2+y^2)/2t}.
\end{aligned}$$

A similar calculation shows

$$\int_0^{t/2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \gtrsim \frac{1}{t}\left(1+\frac{|y|\log t}{\sqrt{t}}\right)e^{-2(x^2+y^2)/t}$$

and

$$\int_{t/2}^{t-2} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \gtrsim \frac{|x|}{t^{3/2}}e^{-x^2/t}\log t \asymp \frac{|x|\log t}{t^{3/2}}e^{-(x^2+y^2)/t}.$$

By Theorem 4.6,

$$\begin{aligned}
\int_{t-2}^{t} p(t-s,a^*,y)\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \quad &\lesssim \quad \int_{t-2}^{t}\frac{1}{\sqrt{t-s}}e^{-c_5y^2/(t-s)}\frac{|x|}{\sqrt{2\pi s^3}}e^{-x^2/2s}ds \\
&\lesssim \quad \frac{|x|}{\sqrt{t^3}}e^{-x^2/2t}\int_{t-2}^{t}\frac{1}{\sqrt{t-s}}ds \\
&\lesssim \quad \frac{|x|}{t^{3/2}}e^{-(x^2+y^2)/2t}.
\end{aligned}$$

These estimates together with (5.41)-(5.43) establish the theorem. □

Theorem 5.17 together with Theorems 5.14 and 5.15 gives Theorem 1.4.

## 6 Hölder regularity of Parabolic Functions

As we noted in Remark 4.9(iii), parabolic Harnack principle fails for the BMVD $X$. However we show in this section that Hölder regularity holds for the parabolic functions of $X$. In the elliptic case (that is, for harmonic functions instead of parabolic functions), this kind of phenomenon has been observed for solutions of SDEs driven by multidimensional Lévy processes with independent coordinate processes; see [6].

To show the Hölder-continuity of parabolic functions of $X$, we begin with the following two lemmas.

**Lemma 6.1.** *There exist $C_1 > 0$ and $0 < C_2 \le 1/2$ such that for every $x_0 \in E$ and $R > 0$,*

$$p_{B(x_0,R)}(t, x, y) \ge \frac{1}{2}\, p(t, x, y) \qquad \text{for } t \in (0, C_1/(R \vee 1)^2] \text{ and } x, y \in B(x_0, C_2R).$$

*Proof.* By Theorem 1.3, there exist constants $c_i > 0$, $1 \le i \le 4$, such that for all $t \le 1$ and $x, y \in E$,

$$\frac{c_3}{\sqrt{t}} e^{-c_4\rho(x,y)^2/t} \le p(t, x, y) \le \frac{c_1}{t} e^{-c_2\rho(x,y)^2/t}. \tag{6.1}$$

We choose $0 < c_5 < 1/2$ sufficiently small such that

$$\frac{(1 - c_5)^2}{(2c_5)^2} \ge \frac{c_2}{c_4}. \tag{6.2}$$

As $t \mapsto t^{-1}e^{-c_0/t}$ is increasing in $t \in (0, 1/c_0]$, we have for $0 < t \le 1/(c_2(1 - c_5)^2R^2)$ and $x, y \in B(x_0, c_5R)$,

$$\begin{aligned}
\overline{p}_{B(x_0,c_5R)}(t, x, y) &:= \mathbb{E}_x[p(t - \tau_{B(x_0,c_5R)}, X_{B(x_0,c_5R)}, y); \tau_{B(x_0,c_5R)} < t] \\
&\lesssim \mathbb{E}_x[(t - \tau_{B(x_0,c_5R)})^{-1}e^{-c_2((1-c_5)R)^2/(t-\tau_{B(x_0,c_5R)})}; \tau_{B(x_0,c_5R)} < t] \\
&\le t^{-1}e^{-c_2((1-c_5)R)^2/t} \lesssim e^{-c_2(1-c_5)^2R^2/2t},
\end{aligned}$$

while

$$p(t, x, y) \ge \frac{c_3}{\sqrt{t}} e^{-c_4(2c_5R)^2/t} \overset{(6.2)}{\ge} \frac{c_3}{\sqrt{t}} e^{-c_2(1-c_5)^2R^2/2t}.$$

Hence there is $c_6 \le 1/(c_2(1 - c_5)^2)$ so that $p(t, x, y) \ge \frac{1}{2}\overline{p}_{B(x_0,c_5R)}(t, x, y)$ for every $R > 0$, $x_0 \in E$, $0 < t \le c_6/(R\vee 1)^2$ and $x, y \in B(x_0, c_5R)$. This proves the lemma as $p_{B(x_0,c_5R)}(t, x, y) = p(t, x, y) - \overline{p}_{B(x_0,c_5R)}(t, x, y)$. □

Let $Z_s = (V_s, X_s)$ be the space-time process of $X$ where $V_s = V_0 + s$. In the rest of this section,

$$Q(t, x, R) := (t, t + R^2) \times B_\rho(x, R).$$

For any Borel measurable set $A \subset Q(t, x, R)$, we use $|A|$ to denote its measure under the product measure $dt \times m_p(dx)$.

**Lemma 6.2.** *Fix $R_0 \geq 1$. There exist constants $0 < C_3 \leq 1/2$ and $C_4 > 0$ such that for all $0 < R \leq R_0$, $x_0 \in E$, $x \in B_\rho(x_0, C_3R)$ and any $A \subset Q(0, x_0, C_3R)$ with $\frac{|A|}{|Q(0,x_0,C_3R)|} \geq \frac{1}{3}$,*

$$\mathbb{P}_{(0,x)}(\sigma_A < \tau_R) \geq C_4, \tag{6.3}$$

*where $\tau_R = \tau_{Q(0,x_0,R)} = \inf\{t \geq 0 : X_t \notin B_\rho(x_0, R)\} \wedge R^2$ and $\sigma_A := \inf\{t \geq 0 : (V_t, X_t) \in A\}$.*

*Proof.* Let $C_1$ and $C_2$ be the constants in Lemma 6.1. Define $C_3 = (C_1/R_0^3)^{1/2} \wedge C_2$. For $x_0 \in E$ and $R \in (0, R_0]$, denote by $X^{B_\rho(x_0,R)}$ the subprocess of $X$ killed upon exiting the ball $B_\rho(x_0, R)$ and $p_{B_\rho(x_0,R)}$ its transition density with respect to the measure $m_p$. As $|(0, C_3^2R^2/6) \times B_\rho(x_0, C_3R)| = |Q(0, x_0, C_3R)|/6$ and $|A| \geq |Q(0, x_0, C_3R)|/3$, we have

$$|\{(t, x) \in A : t \in [C_3^2R^2/6, C_3^2R^2]\}| \geq |Q(0, x_0, C_3R)|/6.$$

For $s > 0$, let $A_s := \{x \in E : (s, x) \in A\}$. Note that

$$\begin{aligned}
\mathbb{E}_x \int_0^{\tau_R} \mathbf{1}_A(s, X_s)ds &= \mathbb{E}_x \int_0^{\tau_R \wedge (C_3R)^2} \mathbf{1}_A(s, X^{B_\rho(x_0,R)})ds \\
&= \int_0^{C_3^2R^2} \mathbb{P}_x \left( X_s^{B_\rho(x_0,R)} \in A_s \right) ds \\
&= \int_0^{C_3^2R^2} \int_{A_s} p_{B_\rho(x_0,R)}(s, x, y)m_p(dy)ds \\
&\geq \int_{C_3^2R^2/6}^{C_3^2R^2} p_{B_\rho(x_0,R)}(s, x, y)m_p(dy)ds.
\end{aligned} \tag{6.4}$$

We now consider two cases.
*Case 1.* $m_p(B_\rho(x_0, R)) > pR/6$. In this case, we have by (6.4), Lemma 6.1 and (6.1) that

$$\begin{aligned}
\mathbb{E}_x \int_0^{\tau_R} \mathbf{1}_A(s, X_s)ds &\gtrsim \int_{C_3^2R^2/6}^{C_3^2R^2} \int_{A_s} \frac{1}{\sqrt{t}} e^{-c_2\rho(x,y)^2/t} m_p(dy)ds \\
&\gtrsim \frac{1}{R} |\{(t, x) \in A : t \in [C_3^2R^2/6, C_3^2R^2]\}| \\
&\gtrsim |Q(0, x_0, C_3R)|/R \gtrsim R^2.
\end{aligned}$$

*Case 2.* $m_p(B_\rho(x_0, R)) \leq pR/6$. In this case, $x_0$ must be in $D_\varepsilon$ with $\rho(x_0, a^*) \geq \frac{5R}{6}$ and so $m_p(B_\rho(x_0, R)) \geq (5R/6)^2$. Thus we have by (6.4) and Theorem 1.3(iii),

$$\begin{aligned}
\mathbb{E}_x \int_0^{\tau_R} \mathbf{1}_A(s, X_s)ds &\gtrsim \int_{C_3^2R^2/6}^{C_3^2R^2} \int_{A_s} \frac{1}{t} e^{-c_2\rho(x,y)^2/t} m_p(dy)ds \\
&\gtrsim \frac{1}{R^2} |\{(t, x) \in A : t \in [C_3^2R^2/6, C_3^2R^2]\}| \\
&\gtrsim |Q(0, x_0, C_3R)|/R^2 \gtrsim R^2.
\end{aligned}$$

Thus in both cases, there is a constant $c_0 > 0$ independent of $x_0$ and $R \in (0, R_0]$ so that

$$\mathbb{E}_x \int_0^{\tau_R} \mathbf{1}_A(s, X_s)ds \geq c_0 R^2. \tag{6.5}$$

On the other hand,

$$\begin{aligned}\mathbb{E}_x \int_0^{\tau_R} \mathbf{1}_A(s, X_s) ds &= \int_0^{\infty} \mathbb{P}_x \left( \int_0^{\tau_R} \mathbf{1}_A(s, X_s) ds > u \right) du \\ &= \int_0^{R^2} \mathbb{P}_x \left( \int_0^{\tau_R} \mathbf{1}_A(s, X_s) ds > u \right) du \\ &\le R^2\, \mathbb{P}_x(\sigma_A < \tau_R).\end{aligned}$$

The desired estimate now follows from this and (6.5). □

**Theorem 6.3.** *For every $R_0 > 0$, there are constants $C = C(R_0) > 0$ and $\beta \in (0,1)$ such that for every $R \in (0, R_0]$, $x_0 \in E$, and every bounded parabolic function $q$ in $Q(0, x_0, 2R)$, it holds that*

$$|q(s,x) - q(t,y)| \le C\|q\|_{\infty,R}\, R^{-\beta} \left( |t-s|^{1/2} + \rho(x,y) \right)^{\beta} \tag{6.6}$$

*for every $(s,x)$, $(t,y) \in Q(0, x_0, R/4)$, where $\|q\|_{\infty,R} := \sup_{(t,y)\in(0,4R^2]\times B_\rho(x_0,2R)} |q(t,y)|$.*

*Proof.* With loss of generality, assume $0 \le q(s) \le \|q\|_{\infty,R} = 1$. We first assume $x_0 = a^*$ and show that (6.6) holds for all $(s,x),(t,y) \in Q(0,x_0,R)$ (instead of $Q(0,x_0,R/4)$). Let $C_3 \in (0,1/2]$ and $C_4 \in (0,1)$ be the constants in Lemma 6.2. Let

$$\eta = 1 - C_4/4 \ge 3/4 \quad \text{and} \quad \gamma = C_3/2 \le 1/4.$$

Note that for every $(s,x) \in Q(0,a^*,R)$, $q$ is parabolic in $Q(s,x,R) \subset Q(0,a^*,2R)$. We will show by induction that $\sup_{Q(s,x,\gamma^k R)} |q| - \inf_{Q(s,x,\gamma^k R)} |q| \le \eta^k$ for all integer $k$. For notation convenience, we denote $Q(s,x,\gamma^k R)$ by $Q_k$. Define $a_i = \inf_{Q_i} q$, $b_i = \sup_{Q_i} q$. Clearly, $b_i - a_i \le 1 \le \eta^i$ for all $i \le 0$. Now suppose $b_i - a_i \le \eta^i$ for all $i \le k$ and we wil show that $b_{k+1} - a_{k+1} \le \eta^{k+1}$. Observe that $Q_{k+1} \subset Q_k$ and so $a_k \le q \le b_k$ on $Q_{k+1}$. Define

$$A' := \left\{ z \in \left(s + (\gamma^{k+1}R)^2, s + (C_3\gamma^k R)^2\right) \times B_\rho(x, C_3\gamma^k R) : q(z) \le (a_k + b_k)/2 \right\},$$

which is a subset of $Q_k$. Note that

$$\left| \left(s + (\gamma^{k+1}R)^2, s + (C_3\gamma^k R)^2\right) \times B_\rho(x, \gamma^k R) \right| = \frac{3}{4}(C_3\gamma^k R)^2\, m_p(B_\rho(x, C_3\gamma^k R)).$$

We may suppose $|A'| \ge \frac{1}{2}(C_3\gamma^k R)^2\, m_p(B(x, C_3\gamma^k R))$; otherwise we consider $1-q$ instead of $q$. Let $A$ be a compact subset of $A'$ such that $|A| \ge \frac{1}{2}(C_3\gamma^k R)^2\, m_p(B(x, C_3\gamma^k R))$. For any given $\varepsilon > 0$, pick $z_1 = (t_1, x_1), z_2 \in Q_{k+1}$ so that $q(z_1) \ge b_{k+1} - \varepsilon$ and $q(z_1) \le a_{k+1} + \varepsilon$. Note that

$Z_{\tau_k} \in \partial Q_k$ as BMVD $X_t$ has continuous sample paths. So by Lemma 6.2,

$$\begin{aligned} b_{k+1} - a_{k+1} - 2\varepsilon &\leq q(z_1) - q(z_2) \\ &= \mathbb{E}_{z_1}\left[q(Z_{\sigma_A \wedge \tau_k}) - q(z_2)\right] \\ &= \mathbb{E}_{z_1}\left[q(Z_{\sigma_A}) - q(z_2); \sigma_A < \tau_k\right] + \mathbb{E}_{z_1}\left[q(Z_{\tau_k}) - q(z_2); \tau_k < \sigma_A\right] \\ &\leq \left(\frac{a_k + b_k}{2} - a_k\right) \mathbb{P}_{z_1}\left(\sigma_A < \tau_k)\right) + (b_k - a_k)\mathbb{P}_{z_1}(\sigma_A > \tau_k) \\ &= (b_k - a_k)\left(1 - \mathbb{P}_{z_1}(\sigma_A < \tau_k)/2\right) \\ &\leq \eta^k(1 - C_4/2) \\ &\leq \eta^{k+1}. \end{aligned}$$

Since $\varepsilon$ is arbitrary, we get $b_{k+1} - a_{k+1} \leq \eta^{k+1}$. This proves that $b_k - a_k \leq \eta^k$ for all integer $k$.

For $z = (s, x)$ and $w = (t, y)$ in $Q(0, a^*, R)$ with $s \leq t$, let $k$ be the smallest integer such that $|z - w| := |t - s|^{1/2} + \rho(x, y) \leq \gamma^k R$. Then

$$|q(z) - q(w)| \leq \eta^k = \gamma^{k \log \eta / \log \gamma} \leq \left(\frac{|z - w|}{\gamma R}\right)^{\log \eta / \log \gamma}.$$

This establishes (6.6) for $x_0 = a^*$ and for every $(s, x), (t, y) \in Q(0, a^*, R)$ with $\beta = \log \eta / \log \gamma$. Note that $\beta \in (0, 1)$ since $0 < \gamma < \eta < 1$.

For general $x_0 \in E$, we consider two cases based on the distance $\rho(x, a^*)$:

*Case 1.* $|x|_\rho < R/2$. In this case, $Q(0, x_0, R/4) \subset Q(0, a^*, 3R/4) \subset Q(0, a^*, 3R/2) \subset Q(0, x_0, 2R)$. By what we have established above,

$$|q(s, x) - q(t, y)| \leq C(R_0)\|q\|_{\infty, R}\, R^{-\beta} \left(|t - s|^{1/2} - \rho(x, y)\right)^\beta \quad \text{for } (s, x), (t, y) \in Q(0, x_0, R/4).$$

*Case 2.* $|x|_\rho \geq R/2$. Since $a^* \notin Q(0, x_0, R/2)$, it follows from the classical results for Brownian motion in $\mathbb{R}^d$ with $d = 1$ and $d = 2$ that for every $(s, x), (t, y) \in Q(0, x_0, R/4)$

$$\begin{aligned} |q(s, x) - q(t, y)| &\leq C(R_0)\|q\|_{\infty, R/2} R^{-\beta} \left(|t - s|^{1/2} + \rho(x, y)\right)^\beta \\ &\leq C(R_0)\|q\|_{\infty, R}\ R^{-\beta} \left(|t - s|^{1/2} + \rho(x, y)\right)^\beta. \end{aligned}$$

This completes the proof of the theorem. □

# 7 Green Function Estimates

In this section, we establish two-sided bounds for the Green function of BMVD $X$ killed upon exiting a bounded connected $C^{1,1}$ open set $D \subset E$. Recall that the Green function $G_D(x, y)$ is defined as follows:

$$G_D(x, y) = \int_0^\infty p_D(t, x, y)dt,$$

where $p_D(t, x, y)$ is the transition density function of the subprocess $X^D$ with respect to $m_p$. We assume $a^* \in D$ throughout this section, as otherwise, due to the connectedness of $D$, either

$D \subset \mathbb{R}_+$ or $D \subset D_\varepsilon$. Therefore $G_D(x,y)$ is just the standard Green function of a bounded $C^{1,1}$ domain for Brownian motion in one-dimensional or two-dimensional spaces, whose two-sided estimates are known, see [16]. It is easy to see from

$$p_D(t,x,y) = p(t,x,y) - \mathbb{E}_x[p(t-\tau_D, X_{\tau_D}, y); \tau_D < t],$$

that $p_D(t,x,y)$ is jointly continuous in $(t,x,y)$.

Recall that for any bounded open set $U \subset E$, $\delta_U(\cdot) := \rho(\cdot, \partial U)$ denotes the $\rho$-distance to the boundary $\partial U$. For notational convenience, we set $U_1 := D \cap (\mathbb{R}_+ \setminus \{a^*\})$ and $U_2 := D \cap D_\varepsilon$. Note that $a^* \in \partial U_1 \cap \partial U_2$ and $U_1 = (0,b)$ for some $b > 0$.

The following theorem gives two-sided Green function estimates for $X$ in bounded $C^{1,1}$ domains.

**Theorem 7.1.** *Let $G_D(x,y)$ be the Green function of $X$ killed upon exiting $D$, where $D$ is a connected bounded $C^{1,1}$ domain of $E$ containing $a^*$. We have for $x \neq y$ in $D$,*

$$G_D(x,y) \asymp \begin{cases} \delta_D(x) \wedge \delta_D(y), & x \in U_1 \cup \{a^*\},\ y \in U_1 \cup \{a^*\}; \\ \delta_D(x)\delta_D(y) + \ln\left(1 + \frac{\delta_{U_2}(x)\delta_{U_2}(y)}{|x-y|^2}\right), & x \in U_2,\ y \in U_2; \\ \delta_D(x)\delta_D(y), & x \in U_1 \cup \{a^*\},\ y \in U_2. \end{cases}$$

*Proof.* We first show that $G_D(x,a^*)$ is a bounded positive continuous function on $D$. By Theorem 1.3, there is a constant $c_1 > 0$ so that for every $x \in D$,

$$\begin{aligned} \mathbb{P}_x(\tau_D < 1) &\geq \mathbb{P}_x(X_1 \in E \setminus D) = \int_{\mathbb{R}_+ \cap D^c} p(1,x,z) m_p(dz) + \int_{D_\varepsilon \cap D^c} p(1,x,z) m_p(dz) \\ &\geq c_1. \end{aligned} \tag{7.1}$$

Thus $\mathbb{P}_x(\tau_D \geq 1) \leq 1 - c_1$ for every $x \in D$. By the strong Markov property of $X$, there are constants $c_2, c_3 > 0$ so that $\mathbb{P}_x(\tau_D \geq t) \leq c_2 e^{-c_3 t}$ for every $x \in D$ and $t > 0$. For $t \geq 2$ and $x, y \in D$, we thus have by Theorem 1.3,

$$p_D(t,x,y) = \int_D p_D(t-1,x,z) p_D(1,z,y) m_p(dz) \leq c_4 \int_D p_D(t-1,x,z) m_p(dz) \leq c_5 e^{-c_3 t}.$$

By Theorem 1.3 again, we conclude that

$$G_D(x,a^*) = \int_0^2 p_D(t,x,y)dt + \int_2^\infty p_D(t,x,y)dt$$

converges and is a bounded positive continuous function in $x \in D$. In particular, $G_D(a^*,a^*) < \infty$.

We further note that $x \mapsto G_D(x,a^*)$ is a harmonic function in $U_1$ and so it is a linear function. As it vanishes at $b := \partial D \cap \mathbb{R}_+$, we have

$$G_D(x,a^*) = c_6 |b - x| \asymp \delta_D(x) \quad \text{for } x \in U_1. \tag{7.2}$$

(i) Assume $x, y \in U_1 \cup \{a^*\}$ and $x \neq y$. If $x = a^*$ or $y = a^*$, the desired estimate holds in view of (7.2). Thus we assume neither $x$ nor $y$ is $a^*$. By the strong Markov property of $X$,

$$G_D(x,y) = G_{U_1}(x,y) + \mathbb{E}_x[G_D(X_{\sigma_{a^*}}, y); \sigma_{a^*} < \tau_D] = G_{U_1}(x,y) + G_D(a^*, y)\mathbb{P}_x(\sigma_{a^*} < \tau_D).$$

Since $x \mapsto \mathbb{P}_x(\sigma_{a^*} < \tau_D)$ is a harmonic function in $U_1$ that vanishes at $b$, by the same reasoning as that for (7.2), we have

$$\mathbb{P}_x(\sigma_{a^*} < \tau_D) = c_7|x-b| \asymp \delta_D(x) \quad \text{for } x \in U_1. \tag{7.3}$$

Thus

$$G_D(a^*, y)\mathbb{P}_x(\sigma_{a^*} < \tau_D) \asymp \delta_D(x)\delta_D(x) \text{for } x, y \in U_1.$$

On the other hand, it is known, see, e.g., [16, (29) on p.45], that

$$G_{U_1}(x,y) = G_{(0,b)}(x,y) = \begin{cases} 2x(b-y)/b & \text{if } 0 < x < y < b, \\ 2y(b-x)/b & \text{if } 0 < y < x < b. \end{cases} \tag{7.4}$$

By symmetry of the Green function $G_D(x,y)$, without loss of generality, we may and do assume that $0 < x < y < b$. By considering the cases of $x < b/2$ and $x \geq b/2$ separately, we conclude from the above estimates that

$$G_D(x,y) \asymp G_{U_1}(x,y) + \delta_D(x)\delta_D(y) \asymp \delta_D(x) \wedge \delta_D(y).$$

(ii) Assume that $x, y \in U_2$. By the strong Markov property of $X$,

$$G_D(x,y) = G_{U_2}(x,y) + \mathbb{E}_x[G_D(X_{\sigma_{a^*}}, y); \sigma_{a^*} < \tau_D] = G_{U_2}(x,y) + G_D(a^*, y)\mathbb{P}_x(\sigma_{a^*} < \tau_D).$$

Since both $y \mapsto G_D(a^*, y)$ and $x \mapsto \mathbb{P}_x(\sigma_{a^*} < \tau_D)$ are bounded positive harmonic functions on $D \cap D_\varepsilon$ that vanishes on $D_\varepsilon \cap \partial D$, it follows from the boundary Harnack inequality for Brownian motion in $\mathbb{R}^2$ that

$$G_D(a^*, y) \asymp \delta_D(y) \quad \text{and} \quad \mathbb{P}_x(\sigma_{a^*} < \tau_D) \asymp \delta_D(x) \tag{7.5}$$

This combined with the Green function estimates of $G_{U_2}(x,y)$ (see [16]) yields

$$G_D(x,y) \asymp \ln\left(1 + \frac{\delta_{U_2}(x)\delta_{U_2}(y)}{|x-y|^2}\right) + \delta_D(x)\delta_D(y).$$

(iii) We now consider the last case that $x \in U_1 \cup \{a^*\}$ and $y \in U_2$. When $x = a^*$, the desired estimates follows from (7.5) and so it remains to consider $x \in U_1$ and $y \in U_2$. By the strong Markov property of $X$, (7.3) and (7.5),

$$G_D(x,y) = \mathbb{E}_x[G_D(X_{\sigma_{a^*}}, y); \sigma_{a^*} < \tau_D] = G_D(a^*, y)\mathbb{P}_x(\sigma_{a^*} < \tau_D) \asymp \delta_D(x)\delta_D(y).$$

This completes the proof of the theorem. □

# 8 Some other BMVD

In this section, we present two more examples of spaces with varying dimension that are variations of $\mathbb{R}^2 \cup \mathbb{R}$ considered in previous sections. The existence and uniqueness of BMVD on these spaces can be established in a similar way as Theorem 2.2 in Section 2. We will concentrate on two-sided estimates on the transition density function on fixed time intervals of these two BMVD. One can further study their large time heat kernel estimates. Due to the space limitation, we will not pursue it in this paper.

## 8.1 A square with multiple flag poles

In this subsection, we study the BMVD on a large square with multiple flag poles of infinite length. The state space $E$ is defined as follows. Let $k \geq 2$ and $\{z_j; 1 \leq j \leq k\}$ be $k$ points in $\mathbb{R}^2$ that have distance at least 4 between each other. Fix a finite sequence $\{\varepsilon_j; 1 \leq j \leq k\} \subset (0, 1/2)$ and a sequence of positive constants $\vec{p} := \{p_j; 1 \leq j \leq k\}$. For $1 \leq j \leq k$, denote by $B_j$ the closed disk on $\mathbb{R}^2$ centered at $z_j$ with radius $\varepsilon_j$. Clearly, the distance between any two distinct balls is at least 3. Let $D_\varepsilon = \mathbb{R}^2 \setminus (\cup_{1 \leq i \leq k} B_i)$. For $1 \leq j \leq k$, denote by $L_j$ the half-line $\{(z_j, w) \in \mathbb{R}^3 : w > 0\}$. By identifying each closed ball $B_j$ with a singleton denoted by $a_j^*$, we equip the space $E := D_\varepsilon \bigcup \{a_1^*, \cdots, a_k^*\} \bigcup (\cup_{i=1}^k L_i)$ with induced topology from $\mathbb{R}^2$ and the half-lines $L_j$, $1 \leq j \leq k$, with the endpoint of the half-line $L_i$ identified with $a_i^*$ and a neighborhood of $a_i^*$ defined as $\{a_i^*\} \cup (V_1 \cap L_i) \cup (V_2 \cap D_\varepsilon)$ for some neighborhood $V_1$ of 0 in $\mathbb{R}$ and $V_2$ of $B_i$ in $\mathbb{R}^2$. Let $m_{\vec{p}}$ be the measure on $E$ whose restriction on $D_\varepsilon$ is the two-dimensional Lebesgue measure, and whose restriction on $L_j$ is the one-dimensional Lebesgue measure multiplied by $p_j$ for $1 \leq j \leq k$. So in particular, $m_{\vec{p}}(\{a_i^*\}) = 0$ for all $1 \leq i \leq k$. We denote the geodesic distance on $E$ by $\rho$.

Similar to Definition 1.1, BMVD on the plane with multiple half lines is defined as follows.

**Definition 8.1.** Given a finite sequence $\vec{\varepsilon} := \{\varepsilon_j; 1 \leq j \leq k\} \subset (0, 1/2)$ and a sequence of positive constants $\vec{p} := \{p_j; 1 \leq j \leq k\}$. A Brownian motion with varying dimension with parameters $(\vec{\varepsilon}, \vec{p})$ on $E$ is an $m_{\vec{p}}$-symmetric diffusion $X$ on $E$ such that

(i) its subprocess in $L_i$, $1 \leq i \leq k$, or $D_\varepsilon$ has the same law as that of standard Brownian motion killed upon leaving $\mathbb{R}_+$ or $D_\varepsilon$;

(ii) it admits no killings at $a_i^*$ for every $1 \leq i \leq k$.

Recall that the endpoint of the half-line $L_i$ is identified with $a_i^*$, $1 \leq i \leq k$. Similar to Theorem 2.2, we have the following theorem stating the existence and uniqueness of the planary BMVD $X$ with multiple half lines.

**Theorem 8.2.** *For each $k \geq 2$, every $\vec{\varepsilon} := \{\varepsilon_j; 1 \leq j \leq k\} \subset (0, 1/2)$ and $\vec{p} := \{p_j; 1 \leq j \leq k\} \subset (0, \infty)$, BMVD $X$ on $E$ with parameter $(\vec{\varepsilon}, \vec{p})$ exists and is unique. Its associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E; m_p)$ is given by*

$$\begin{aligned}
\mathcal{F} &= \left\{ f : f|_{\mathbb{R}^2} \in W^{1,2}(\mathbb{R}^2);\ f|_{L_i} \in W^{1,2}(\mathbb{R}),\ f|_{B_i} = f|_{L_i}(a_i^*) \text{ for } 1 \leq i \leq k \right\}, \\
\mathcal{E}(f, g) &= \frac{1}{2} \int_{D_\varepsilon} \nabla f(x) \cdot \nabla g(x) dx + \sum_{i=1}^{k} \frac{p_i}{2} \int_{L_i} f'(x) g'(x) dx.
\end{aligned}$$

It is not difficult to see that BMVD $X$ has a continuous transition density $p(t, x, y)$ with respect to the measure $m_{\vec{p}}$.

**Proposition 8.3.** *There exist constants $C_1, C_2 > 0$ such that*

$$p(t, x, y) \leq C_1 \left( \frac{1}{t} + \frac{1}{t^{1/2}} \right) e^{-C_2 \rho(x,y)^2 / t} \qquad \text{for all } x, y \in E \text{ and } t > 0.$$

*Proof.* By an exactly the same argument as that for Proposition 3.1, we can establish Nash-type inequality for $X$. From it, the off-diagonal upper bound can be derived using Davies' method as in Propositions 3.2 and 3.4. □

The following theorem gives two-sided bounds for the transition density function $p(t,x,y)$ when $t \in (0,T]$ for each fixed $T>0$.

**Theorem 8.4.** *Let $T \geq 2$ be fixed. There exist positive constants $C_j$, $3 \leq j \leq 16$ so that the transition density $p(t,x,y)$ of BMVD $X$ on $E$ satisfies the following estimates when $t \in (0,T]$.*

(i) *For $x,y \in D_\varepsilon \cap B_\rho(a_i^*,1)$ for some $1 \leq i \leq k$,*

$$\frac{C_3}{\sqrt{t}} e^{-C_4 \rho(x,y)^2/t} + \frac{C_3}{t}\left(1 \wedge \frac{\rho(x,a_i^*)}{\sqrt{t}}\right)\left(1 \wedge \frac{\rho(y,a_i^*)}{\sqrt{t}}\right) e^{-C_5|x-y|^2/t} \leq p(t,x,y)$$
$$\leq \frac{C_6}{\sqrt{t}} e^{-C_7 \rho(x,y)^2/t} + \frac{C_6}{t}\left(1 \wedge \frac{\rho(x,a_i^*)}{\sqrt{t}}\right)\left(1 \wedge \frac{\rho(y,a_i^*)}{\sqrt{t}}\right) e^{-C_8|x-y|^2/t};$$

(ii) *For some $1 \leq i \leq k$, $x \in L_i$ and $y \in L_i \cup B_\rho(a_i^*,1)$, or $x \in L_i \cup B_\rho(a_i^*,1)$ and $y \in L_i$,*

$$\frac{C_9}{\sqrt{t}} e^{-C_{10}\rho(x,y)^2/t} \leq p(t,x,y) \leq \frac{C_{11}}{\sqrt{t}} e^{-C_{12}\rho(x,y)^2/t};$$

(iii) *For all other cases,*

$$\frac{C_{13}}{t} e^{-C_{14}\rho(x,y)^2/t} \leq p(t,x,y) \leq \frac{C_{15}}{t} e^{-C_{16}\rho(x,y)^2/t}. \tag{8.1}$$

*Proof.* The idea of the proof is to reduce it to the heat kernel for BMVD on plane with one vertical half-line. For an open subset $D$ of $E$, we use $X^D$ to denote the subprocess of $X$ killed upon leaving $D$ and $p_D(t,x,y)$ the transition density function of $X^D$ with respect to $m_{\vec{p}}$.

Let $C_1 > 0$ and $C_2 \in (0,1/2)$ be the constants in Lemma 6.1.

(i) We first show that the desired estimates hold for any $x,y \in E$ with $\rho(x,y) < 2C_2$ and for every $t \in (0,T]$. In this case, let $z_0 \in E$ so that $\{x,y\} \in B_\rho(z_0,C_2)$. Since $\rho(a_i^*,a_j^*) > 3$ for $i \neq j$, without loss of generality we may and do assume that $a_1^*$ is the base that is closest to $z$ and so $\min_{2\leq j\leq k} \rho(z_0,a_j^*) > 3/2$. We have by Lemma 6.1 that

$$p_{B_\rho(z_0,1)}(t,w,z) \geq \frac{1}{2} p_1(t,w,z) \quad \text{ for } t \in (0,C_1] \text{ and } w,z \in B_\rho(z_0,C_2), \tag{8.2}$$

where $p_1(t,x,y)$ stands for the transition density function of BMVD on the plane with one vertical halfline $L_1$ at base $a_1^*$ and vertical half line $L_1$. This together with Theorem 1.3 in particular implies that there is a constant $c_1 > 0$ so that

$$p_{B_\rho(z_0,1)}(t,w,z) \geq c_1 \quad \text{for every } t \in [C_1/2, C_1] \text{ and } w,z \in B_\rho(z_0,C_2).$$

It thus follows from the Chapman-Kolmogorov's equation that there is a constant $c_2 > 0$ so that

$$p_{B_\rho(z_0,1)}(t,w,z) \geq c_2 \quad \text{for every } t \in [C_1,T] \text{ and } w,z \in B_\rho(z_0,C_2). \tag{8.3}$$

This together with (8.2) implies that there is a constant $c_3 > 0$ so that

$$p(t,w,z) \geq p_{B_\rho(z_0,1)}(t,w,z) \geq c_3 p_1(t,w,z) \quad \text{ for } t \in (0,T] \text{ and } w,z \in B_\rho(z_0,C_2). \tag{8.4}$$

On the other hand, we have by Proposition 3.4 and the fact that $s \mapsto s^{-1}e^{-a^2/s}$ is increasing in $(0, a^2)$ and decreasing in $(a^2, \infty)$ that for $t \in (0, T]$ and $\rho(x, y) < 2C_2 < 1$,

$$\begin{aligned}\bar{p}_{B_\rho(z_0,1)}(t, x, y) &:= \mathbb{E}_x[p(t - \tau_{B_\rho(z_0,1)}, X_{\tau_{B_\rho(z_0,1)}}, y); \tau_{B_\rho(z_0,1)} < t] \\ &\lesssim t^{-1}e^{-c_4/t} \lesssim e^{-c_5/t} \lesssim e^{-c_6\rho(x,y)^2/t}.\end{aligned}$$

Consequently,

$$p(t, x, y) = p_{B_\rho(z_0,1)}(t, x, y) + \bar{p}_{B_\rho(z_0,1)}(t, x, y) \lesssim p_1(t, x, y) + e^{-c_6\rho(x,y)^2/t}. \tag{8.5}$$

This together with (8.4) and Theorem 1.3 establishes the desired estimate of the theorem for any $x, y \in E$ with $\rho(x, y) < 2C_2$ and $t \in (0, T]$.

(ii) Note that $p(t, x, y)$ is symmetric in $x$ and $y$. It suffices to consider the case when $x \in L_i$ and $y \in L_i \cup B_\rho(a_i^*, C_2)$ for some $1 \le i \le k$ with $\rho(x, y) \ge 2C_2$. Without loss of generality, we may and do assume $i = 1$ and $\rho(x, a_1^*) < \rho(y, a_1^*)$ if $y \in L_1$. Let $z_0 \in L_1$ with $\rho(z_0, a_1^*) = \rho(y, a_1^*) + C_2$. By the strong Markov property of $X$, (8.4)-(8.5) and Theorem 1.3, we have for $t \in (0, T]$,

$$\begin{aligned}p(t, x, y) &= \mathbb{E}_x[p(t - \sigma_{z_0}, z_0, y); \sigma_{z_0} < t] \\ &\ge c_3\mathbb{E}_x[p_1(t - \sigma_{z_0}, z_0, y); \sigma_{z_0} < t] \\ &= c_3 p_1(t, x, y) \gtrsim t^{-1/2}\, e^{-c_7\rho(x,y)^2/y},\end{aligned}$$

and

$$\begin{aligned}p(t, x, y) &= \mathbb{E}_x[p(t - \sigma_{z_0}, z_0, y); \sigma_{z_0} < t] \\ &\overset{(8.5)}{\lesssim} \mathbb{E}_x[p_1(t - \sigma_{z_0}, z_0, y); \sigma_{z_0} < t] + \mathbb{E}_x[e^{-c_6C_2^2/(t-\sigma_{z_0})}; \sigma_{z_0} < t] \\ &\le p_1(t, x, y) + e^{-c_6C_2^2/t}\mathbb{P}_x(\sigma_{z_0} < t) \\ &\lesssim t^{-1/2}\, e^{-c_8\rho(x,y)^2/y} + e^{-c_6C_2^2/t}e^{-|x-z_0|^2/t} \\ &\lesssim t^{-1/2}\, e^{-c_9\rho(x,y)^2/y}.\end{aligned}$$

In the second to last inequality, we used crossing estimate for one-dimensional Brownian motion and Lemma 5.11. The above two estimates give the desired estimates.

(iii) Let $U_1 = D_\varepsilon \setminus \cup_{j=1}^k B_\rho(a_j^*, C_2)$. There are three remaining cases:

(a) $x \in L_i \cup B_\rho(a_i^*, C_2)$ and $y \in L_j \cup B_\rho(a_j^*, C_2)$ for $i \ne j$;

(b) $x \in L_i \cup B_\rho(a_i^*, C_2)$ for some $1 \le i \le k$ and $y \in U_1$ with $\rho(x, y) \ge 2C_2$;

(c) $x, y \in U_1$ with $\rho(x, y) \ge 2C_2$.

We claim that (8.1) holds for all these three cases. The upper bound in (8.1) holds due to Proposition 3.4 so it remains to establish the lower bound.

It follows from the Dirichlet heat kernel estimate for Brownian motion in $C^{1,1}$-domain [36, 14] that in case (c), for any $t \in (0, T]$,

$$p(t, x, y) \ge p_D(t, x, y) \gtrsim t^{-1}e^{-c_{10}|x-y|^2/t} \ge t^{-1}e^{-c_{10}\rho(x,y)^2/t}. \tag{8.6}$$

For case (b), without loss of generality, we assume $i = 1$. Define $u_1(w) = -\rho(w, a_1^*)$ for $w \in L_1$ and $u_1(w) = \rho(w, a_1^*)$ analogous to (4.1). Let $Y = u(X)$ and $\tau_1 := \inf\{t > 0 : Y_t \geq 2C_2\}$. Then $Y_t^{(-\infty,2C_2)} := Y_t$ for $t \in [0, \tau_1)$, and $Y_t^{(-\infty,2C_2)} := \partial$ for $t \geq \tau_1$ has the same distribution as the killed radial process radial process $Y$ in Section 4 for BMVD on plane with one vertical half-line. By Proposition 4.3 and the arguments similar to Proposition 4.4 of this paper and that of [14], one can show that $Y_t^{(-\infty,2C_2)}$ has a transition density function $p^0(t, w, z)$ with respect to the Lebesgue measure on $\mathbb{R}$ and it has the following two-sided estimates:

$$t^{-1/2}\left(1 \wedge \frac{|w|}{\sqrt{t}}\right)\left(1 \wedge \frac{|z|}{\sqrt{t}}\right)e^{-c_{11}|w-z|^2/t} \lesssim p^0(t, w, z) \lesssim t^{-1/2}\left(1 \wedge \frac{|w|}{\sqrt{t}}\right)\left(1 \wedge \frac{|z|}{\sqrt{t}}\right)e^{-c_{12}|w-z|^2/t}$$

for $t \in (0, T]$ and $x, y \in (-\infty, 2C_2)$. Thus we have for $t \in (0, T]$, $x \in L_1 \cup B_\rho(a_1^*, C_2)$ and $y \in U_1$ with $\rho(x, y) \geq 2C_2$,

$$\begin{aligned}
p(t, x, y) &\geq \int_{D_\varepsilon \cap (B_\rho(a_1^*, 6C_2/4) \setminus B_\rho(a_1^*, 5C_2/4))} p(t/2, x, z)p(t/2, z, y)m_{\vec{p}}(dz) \\
&\gtrsim t^{-1}e^{-c_{13}\rho(a_1^*, y)^2/t} \int_{D_\varepsilon \cap (B_\rho(a_1^*, 6C_2/4) \setminus B_\rho(a_1^*, 5C_2/4))} p(t/2, x, z)m_{\vec{p}}(dz) \\
&\geq t^{-1}e^{-c_{13}\rho(a_1^*, y)^2/t} \int_{5C_2/4}^{6C_2/4} p^0(t/2, u_1(x), w)dw \\
&\gtrsim t^{-1}e^{-c_{13}\rho(a_1^*, y)^2/t} \int_{5C_2/4}^{6C_2/4} t^{-1/2}e^{-c_{11}(6C_2/4 - u_1(x))^2/t}dw \\
&\gtrsim t^{-1}e^{-c_{13}\rho(a_1^*, y)^2/t}e^{-c_{14}(6C_2/4 - u_1(x))^2/t} \\
&\gtrsim t^{-1}e^{-c_{15}\rho(x, y)^2/t},
\end{aligned} \tag{8.7}$$

where in the second inequality it was used that $\rho(z, y) \leq 5\rho(a_1^*, y)/2$, and in the last two inequalities we used the fact that $|6C_2/4 - u_1(x)| \geq C_2/2$ and $\rho(x, y) \geq 2C_2$.

Now for case (a) when $x \in L_i \cup B_\rho(a_i^*, C_2)$ and $y \in L_j \cup B_\rho(a_j^*, C_2)$ for $i \neq j$, let $z_0 \in D_\varepsilon$ so that both $\rho(z, a_i^*)$ and $\rho(z, a_j^*)$ take values within $(\rho(a_i^*, a_j^*)/3, 2\rho(a_i^*, a_j^*)/3)$. We then have by (8.7) that for all $t \in (0, T]$,

$$\begin{aligned}
p(t, x, y) &\geq \int_{D_\varepsilon \cap B_\rho(z_0, C_2)} p(t/2, x, z)p(t/2, z, y)m_{\vec{p}}(dz) \\
&\gtrsim t^{-1}e^{-c_{16}\rho(x, z_0)^2/t}\, t^{-1}e^{-c_{16}\rho(y, z_0)^2/t}\, m_{\vec{p}}(D_\varepsilon \cap B_\rho(z_0, C_2)) \\
&\gtrsim t^{-1}e^{-c_{17}\rho(x, y)^2/t},
\end{aligned}$$

where in the last inequality, we used the fact that $\rho(x, y) \geq 3$. This completes the proof that the lower bound in (8.1) holds for all three cases (a)-(c). The theorem is now proved. □

## 8.2 A large square with an arch

In this subsection, we study Brownian motion on a large square with an arch. The state space $E$ is defined as follows. Let $z_1, z_2 \in \mathbb{R}^2$ with $|z_1 - z_2| \geq 6$. Fix constants $0 < \varepsilon_1, \varepsilon_2 < 1/2$ and $p > 0$. For $i = 1, 2$, denote by $B_i$ the closed disk on $\mathbb{R}^2$ centered at $z_i$ with radius $\varepsilon_i$. Let $D_\varepsilon = \mathbb{R}^2 \setminus (B_1 \cup B_2)$. We short $B_i$ into a singleton denoted by $a_i^*$. Denote by $L$ a one dimensional arch with two endpoints $a_1^*$ and $a_2^*$. Without loss of generality, we assume $L$ is isometric to an closed interval $[-b, b]$ for some $b \geq 4$. We equip the space $E := D_\varepsilon \cup \{a_1^*, a_2^*\} \cup L$ with the Riemannian distance $\rho$ induced from $D_\varepsilon$ and $L$, analogous to the last example of a large square with multiple flag poles. Let $m_p$ be the measure on $E$ whose restriction on $L$ and $D_\varepsilon$ is the arch length measure and the Lebesgue measure multiplied by $p$ and 1, respectively. In particular, we have $m_p(\{a_1^*, a_2^*\}) = 0$. As before, BMVD on $E$ is defined as follows.

**Definition 8.5.** Given $0 < \varepsilon_1, \varepsilon_2 < 1/2$ and $p > 0$, BMVD on $E$ with parameters $(\varepsilon_1, \varepsilon_2, p)$ on $E$ is an $m_p$-symmetric diffusion $X$ on $E$ such that

(i) its subprocess process in $L$ or $D_\varepsilon$ has the same distribution as the one-dimensional or two-dimensional Brownian motion killed upon leaving $L$ or $D_\varepsilon$, respectively.

(ii) it admits no killings at $\{a_1^*, a_2^*\}$.

Similar to Theorem 2.2, we have the following.

**Theorem 8.6.** *For every $0 < \varepsilon_1, \varepsilon_2 < 1/2$ and $p > 0$, BMVD $X$ on $E$ with parameter $(\varepsilon_1, \varepsilon_2, p)$ exists and is unique. Its associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E; m_p)$ is given by*

$$\begin{aligned} \mathcal{F} &= \left\{ f : f|_{\mathbb{R}^2} \in W^{1,2}(\mathbb{R}^2),\ f|_L \in W^{1,2}(L),\ f|_{B_i} = f|_L(a_i^*),\ i = 1, 2 \right\}, \\ \mathcal{E}(f, g) &= \frac{1}{2} \int_{D_\varepsilon} \nabla f(x) \cdot \nabla g(x) dx + \frac{p}{2} \int_L f'(x) g'(x) dx. \end{aligned}$$

It is easy to see that BMVD $X$ has a continuous transition density function $p(t, x, y)$ with respect to the measure $m_p$. Similar to that for Proposition 3.1, Propositions 3.2 and 3.4, using the classical Nash's inequality for one- and two-dimensional Brownian motion and Davies method, one can easily establish the following.

**Proposition 8.7.** *Let $T \geq 2$. There exist $C_1, C_2 > 0$ such that*

$$p(t, x, y) \leq C_1 \left( \frac{1}{t} + \frac{1}{t^{1/2}} \right) e^{-C_2 \rho(x,y)^2/t} \qquad \text{for all } x, y \in E, t \in (0, T].$$

The next theorem gives short time sharp two-sided estimates on $p(t, x, y)$.

**Theorem 8.8.** *Let $T \geq 2$ be fixed. There exist positive constants $C_i$, $3 \leq i \leq 16$, so that the transition density $p(t, x, y)$ of BMVD $X$ on $E$ satisfies the following estimates when $t \in (0, T]$:*

(i) *For $x \in L$ and $y \in E$,*

$$\frac{C_3}{\sqrt{t}} e^{-C_4 \rho(x,y)^2/t} \leq p(t, x, y) \leq \frac{C_5}{\sqrt{t}} e^{-C_6 \rho(x,y)^2/t}. \tag{8.8}$$

(ii) *For $x, y \in D_\varepsilon \cup \{a_1^*, a_2^*\}$, when $\rho(x, a_i^*) + \rho(y, a_i^*) < 1$ for some $i = 1, 2$,*

$$\begin{aligned} & \frac{C_7}{\sqrt{t}} e^{-C_8 \rho(x,y)^2/t} + \frac{C_7}{t} \left( 1 \wedge \frac{\rho(x, a_i^*)}{\sqrt{t}} \right) \left( 1 \wedge \frac{\rho(y, a_i^*)}{\sqrt{t}} \right) e^{-C_9 |x-y|^2/t} \\ \leq\ & p(t, x, y) \leq \frac{C_{10}}{\sqrt{t}} e^{-C_{11} \rho(x,y)^2/t} + \frac{C_{10}}{t} \left( 1 \wedge \frac{\rho(x, a_i^*)}{\sqrt{t}} \right) \left( 1 \wedge \frac{\rho(y, a_i^*)}{\sqrt{t}} \right) e^{-C_{12} |x-y|^2/t}; \end{aligned} \tag{8.9}$$

*otherwise,*

$$\frac{C_{13}}{t} e^{-C_{14} \rho(x,y)^2/t} \leq p(t, x, y) \leq \frac{C_{15}}{t} e^{-C_{16} \rho(x,y)^2/t}. \tag{8.10}$$

*Proof.* This theorem can be established by a similar consideration as that for Theorem 8.4. Here we only give a brief sketch. Let $C_1 > 0$ and $C_2 \in (0, 1/2)$ be the constants in Lemma 6.1.

*Case 1.* $\rho(x, y) < 2C_2$. The desired estimates can be obtained in a similar way as that for *Case 1* in the proof of Theorem 8.4, by using Lemma 6.1.

*Case 2.* $\rho(x, y) \geq 2C_2$. Due to the upper bound estimate in Theorem 8.7, it suffices to show the following lower bound estimate hold: there exists some $c_1, c_2 > 0$, such that

$$p(t, x, y) \geq c_1 e^{-c_2 \rho(x,y)^2/t} \quad \text{for all } t \in (0, T]. \tag{8.11}$$

We divided its proof into three subcases.

(i) Both $x$ and $y$ are in $L \cup B_\rho(a_1^*, C_2) \cup B_\rho(a_2^*, C_2)$. Without loss of generality, we assume $x \in L \cup B_\rho(a_1^*, C_2)$ and $x$ is closer to $a_1^*$ if both $x$ and $y$ are on the arch $L$. Denote by $\ell(w, z)$ the arch length in $L$ between two points $w, z \in L$, and $U_1 := L \cup B_\rho(a_1^*, 3C_2) \cup B_\rho(a_2^*, 3C_2)$. We define a modified signed radial process $Y_t = u(X_t^{U_1})$, where

$$u(z) := \begin{cases} \rho(z, a_1^*) & \text{if } z \in D_\varepsilon \cap B_\rho(a_1^*, 3C_2), \\ -\ell(a_1^*, z) & \text{if } z \in L, \\ -\ell(a_1^*, a_2^*) - \rho(z, a_2^*) & \text{if } z \in D_\varepsilon \cap B_\rho(a_2^*, 3C_2). \end{cases}$$

By a similar argument as that for Section 4, we can show that $Y$ is a subprocess of a skew Brownian motion on $\mathbb{R}$ with skewness at points 0 and $\ell(a_1^*, a_2^*)$ and bounded drift killed upon leaving $I = (-\ell(a_1^*, a_2^*) - 2C_2, 2C_2)$. The desired lower bound estimate for $p(t, x, y)$ can be derived from the Dirichlet heat kernel estimate for the one-dimensional diffusion $Y$.

(ii) Both $x, y \in U_2 := D_\varepsilon \setminus (B_\rho(a_1^*, C_2/2) \cup B_\rho(a_2^*, C_2/2))$. In this case, the desired lower bound estimate for $p(t, x, y)$ follows from the Dirichlet heat kernel estimate for two-dimensional Brownian motion in $C^{1,1}$ domain $U_2$.

(iii) $x \in L \cup B_\rho(a_1^*, C_2/2) \cup B_\rho(a_2^*, C_2/2)$ and $y \in D_\varepsilon \setminus (B_\rho(a_1^*, 2C_2) \cup B_\rho(a_2^*, 2C_2))$. Without loss of generality, we assume $x$ is closer to $a_1^*$ than to $a_2^*$. Let

$$D_3 := \{z \in D_\varepsilon : C_2/2 \leq \rho(z, a_1^*) \leq C_2\}.$$

Note that $\rho(x, y) \geq (\rho(x, z) + \rho(y, z))/5$ for $z \in D_3$. By Markov property,

$$p(t, x, y) \geq \int_{D_3} p(t/2, x, z) p(t/2, z, y) m_p(dz).$$

The desired lower bound for $p(t, x, y)$ follows from the results obtained in (i) and (ii).

□

**Acknowledgement.** We thank the referees for helpful comments.

# References

[1] S. Andres and M. T. Barlow. Energy inequalities for cutoff-functions and some applications. *J. Reine Angew. Math.* **699** (2015), 183–215.

[2] D. G. Aronson, Non-negative solutions of linear parabolic equations, *Ann. Scuola Norm Sup. Pisa (3)* **22** (1968), 607-694.

[3] M. T. Barlow and R. F. Bass. Stability of parabolic Harnack inequalities. *Trans. Amer. Math. Soc.* **356** (2003) no. 4, 1501–1533.

[4] M. T. Barlow, R. F. Bass and T. Kumagai. Stability of parabolic Harnack inequalities on metric measure spaces. *J. Math. Soc. Japan* **58** (2006), 485–519.

[5] R. F. Bass and Z.-Q. Chen, One-dimensional stochastic differential equations with singular and degenerate coefficients, *Sankhya* **67** (2005), 19-45.

[6] R. F. Bass and Z.-Q. Chen, Regularity of harmonic functions for a class of singular stable-like processes. *Math. Z.* **266** (2010), 489-503.

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

[8] R. F. Bass and D. A. Levin, Harnack inequalities for jump processes. *Potential Anal.* **17** (2002), 375-388.

[9] M. Biroli and U. A. Mosco, Saint-Venant type principle for Dirichlet forms on discontinuous media. *Ann. Mat. Pura Appl.,* **169** (1995), 125-181.

[10] E. A. Carlen, S. Kusuoka and D. W. Stroock, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré. Probab. Statist.* **23** (1987) 245-287.

[11] Z.-Q. Chen, *Topics on Recent Developments in the Theory of Markov Processes.* `http://www.math.washington.edu/~zchen/RIMS_lecture.pdf`

[12] Z.-Q. Chen and M. Fukushima, *Symmetric Markov processes, Time Change and Boundary Theory.* Princeton University Press, 2011.

[13] Z.-Q. Chen, M. Fukushima and S. Rohde, Chordal Komatu-Loewner equation and Brownian motion with darning in multiply connected domains. *Trans. Amer. Math. Soc.* **368** (2016), 4065-4114.

[14] S. Cho, P. Kim and H. Park, Two-sided estimates on Dirichlet heat kernels for time-dependent parabolic operators with singular drifts in $C^{1,\alpha}$-domains. *J. Diff. Equations* **252** (2012), 1101-1145.

[15] T. Coulhon and A. Grigor'yan, On-diagonal lower bounds for heat kernels and Markov chains, *Duke Math. J.* **89** (1997), 133-199.

[16] K. L. Chung and Z. Zhao, *Form Brownian Motion to Schrödinger's Equation.* Springer, 1995.

[17] T. Delmotte, Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana,* **15** (1999), 181-232.

[18] S. N. Evans and R. B. Sowers, Pinching and twisting Markov processes. *Ann. Probab.* **31** (2003), 486-527.

[19] M. Fukushima, Y. Oshima and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes, 2nd rev. and ext. ed.* De Gruyter, Berlin and New York, 2010.

[20] A. Grigor'yan. Heat kernel upper bounds on a complete non-compact manifold. *Revista Math. Iberoamericana* **10** (1994), 395–452.

[21] A. Grigor'yan and L. Saloff-Coste, Hitting probabilities for Brownian motion on Riemannian manifolds, *J. Math. Pures et Appl.* **81** (2002) 115-142.

[22] B. M. Hambly and T. Kumagai, Diffusion processes on fractal fields: heat kernel estimates and large deviations. *Probab. Theory Related Fields* **127** (2003), 305-352.

[23] W. Hansen, Darning and gluing of diffusions. arXiv:1601.07044v2

[24] E. P. Hsu, *Stochastic Analysis on Manifolds.* American Mathematical Society, 2002.

[25] T. Kumagai, Brownian motion penetrating fractals: an application of the trace theorem of Besov spaces. *J. Funct. Anal.* **170** (2000), 69-92.

[26] J. S. Liu and C. Sabatti, Simulated sintering: Markov chain Monte Carlo with spaces of varying dimensions. *Bayesian statistics, 6 (Alcoceber, 1998)*, 389-413, Oxford Univ. Press, New York, 1999.

[27] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus.* Springer, 1988.

[28] P. Mörters and Y. Peres, *Brownian motion. With an appendix by Oded Schramm and Wendelin Werner.* Cambridge University Press, 2010.

[29] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, 1991.

[30] L. Saloff-Coste. A note on Poincaré, Sobolev, and Harnack inequalities. *Inter. Math. Res. Notices* **2** (1992), 27–38.

[31] K.-T. Sturm, Analysis on local Dirichlet spaces II. Gaussian upper bounds for the fundamental solutions of parabolic Harnack equations. *Osaka J. Math.,* **32** (1995), 275-312.

[32] K.-T. Sturm, Analysis on local Dirichlet spaces III. The parabolic Harnack inequality. *J. Math. Pures Appl.,* **75** (1996), 273-297.

[33] K. Uchiyama, Asymptotic estimates of the distribution of Brownian hitting time of a disc. *J. Theor. Probab.* **25** (2012), 450-463. Erratum, *J. Theor. Probab.* **25** (2012), 910-911.

[34] C. Xia, *Optimal Control of Switched Systems with Dimension-Varying State Spaces.* Ph.D Dissertation in Electrical Engineering, Univ. of California at Los Angeles, 2007.

[35] Q. S. Zhang, Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u)+B\nabla-u_t=0$. *Manuscripta Math.* **93** (1997), 381-390.

[36] Q. S. Zhang, The global behavior of heat kernels in exterior domains. *J. Funct. Anal.* **200** (2003), 160-176.

**Zhen-Qing Chen**

Department of Mathematics, University of Washington, Seattle, WA 98195, USA
E-mail: `zqchen@uw.edu`

**Shuwen Lou**

Department of Mathematics, Statistics, and Computer Science,
University of Illinois at Chicago, Chicago, IL 60607, USA
E-mail: `slou@uic.edu`