Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 2, 1002–1040
https://doi.org/10.1214/19-AIHP992

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Weak uniqueness and density estimates for SDEs with coefficients depending on some path-functionals

Noufel Frikha[a] and Libo Li[b]

[a]*Sorbonne Université, Sorbonne Paris Cité, CNRS, Laboratoire de Probabilité Statistique et Modélisation, LPSM, F-75005, Paris, France. E-mail: frikha@math.univ-paris-diderot.fr*

[b]*Libo Li, School of Mathematics and Statistics, University of New South Wales, Sydney, Australia. E-mail: libo.li@unsw.edu.au*

Received 24 July 2018; revised 26 March 2019; accepted 2 April 2019

Abstract. In this article, we develop a methodology to prove weak uniqueness for stochastic differential equations with coefficients depending on some path-functionals of the process. As an extension of the technique developed by Bass and Perkins (In *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut* (2009) 47–53) in the standard diffusion case, the proposed methodology allows one to deal with process whose probability laws is singular with respect to the Lebesgue measure. To illustrate our methodology, we prove weak existence and uniqueness in the two following examples: a diffusion process with coefficients depending on its running local time and a diffusion process with coefficients depending on its running maximum. In each example, we also prove the existence of the associated transition density and establish some Gaussian upper-estimates.

Résumé. Dans cet article, nous développons une méthodologie permettant de prouver l'unicité faible pour des équations différentielles stochastiques dont les coefficients dépendent de certaines fonctionnelles de la trajectoire du processus. Dans le prolongement de la technique développée par Bass & Perkins (In *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut* (2009) 47–53) dans le cas des processus de diffusions standards, la méthologie proposée permet de traiter le cas de processus dont la loi est singulière par rapport à la mesure de Lebesgue. Afin d'illustrer notre méthodologie, nous prouvons l'existence et l'unicité faible dans les deux exemples suivants: un processus de diffusion dont les coefficients dépendent du temps local en zéro courant et un processus de diffusion dont les coefficients dépendent du maximum courant. Dans chaque exemple, nous prouvons également l'existence d'une densité de transition associée et établissons des estimées Gaussiennes.

MSC: Primary 60H10; 60G46; secondary 60H30; 35K65

Keywords: Weak uniqueness; Martingale problem; Parametrix expansion; Local time; Maximum; Density estimates

1. Introduction

In the present paper, we investigate the weak existence and uniqueness of an one-dimensional stochastic differential equation (SDE) with coefficients depending on some path-functional A and dynamics given by

$$X_t = x_0 + \int_0^t b\big(X_s, A_s(X)\big)\, ds + \int_0^t \sigma\big(X_s, A_s(X)\big)\, dW_s, \tag{1.1}$$

where $(W_t)_{t\geq 0}$ stands for a one-dimensional Brownian motion and $(A_t(X))_{t\geq 0}$ is an $\mathbb{R}^{d-1}$-valued functional ($d \geq 2$) depending on the path of X in a non-anticipative way, namely, it depends on the path of $X_{\cdot\wedge t}$. Some examples include its local and occupation times, its running maximum or minimum, its first hitting time of a level, its running average, etc. From the point of view of applications, systems of the type (1.1) appear in many fields. Let us mention stochastic Hamiltonian systems where $A_t(X) = \int_0^t F(s, X_s)\, ds$, see e.g. [16] for a general overview, [18] for convergence to equilibrium or [2] for an application to the pricing of Asian options. We also mention [7], where the author constructs a weak solution to the SDE (1.1) with $b \equiv 0$ and $A_t(X) = \max_{0\leq s\leq t} X_s$ is the running maximum of X and investigates an application in mathematical finance.

In the standard multi-dimensional diffusion framework, the martingale approach initiated by Stroock and Varadhan turns out to be particularly powerful when trying to get uniqueness results. It is now well-known that the martingale problem associated to a multi-dimensional diffusion operator is well posed as soon as the drift is bounded measurable and that the diffusion matrix is continuous (with respect to the space variable) and strictly positive, see e.g. Stroock and Varadhan [17]. In the indicated framework, uniqueness is derived from Calderón–Zygmund estimates. Also, when a is Hölder continuous, an analytical approach using Schauder estimates can be applied, see e.g. Friedman [8].

Recently, Bass and Perkins [4] introduced a new technique for proving uniqueness for the martingale problem and illustrated it in the framework of non-degenerate, non-divergence and time-homogeneous diffusion operators under the assumption that the diffusion matrix is strictly positive and Hölder continuous. It has also been recently extended by Menozzi [14] for a class of multi-dimensional degenerate Kolmogorov equations, that is, the case of a multi-dimensional path functional $A = (A_t^1, \dots, A_t^N)_{t\geq 0}$ given by: $A_t^1(X) = \int_0^t F_1(X_s, A_s(X))\, ds$, $A_t^2(X) = \int_0^t F_2(A_s^1, \dots, A_s^N)\, ds, \dots, A_t^N(X) = \int_0^t F_N(A_s^{N-1}, A_s^N)\, ds$, under an assumption of weak Hörmander type on the functions $(F_1, \dots, F_N)$. The approach in the two mentioned papers consists in using a perturbation method for Markov semigroups, known as the parametrix technique, such as exposed in Friedman [8] in the case of uniformly elliptic diffusion. More precisely, the first step of the strategy is to approximate the original system by a simple process obtained by freezing the drift and the diffusion coefficients in the original dynamics, and use the fact that the transition density of such approximation as well as its derivatives can be explicitly estimated. Then, the key ingredient is the *smoothing property* of the underlying *parametrix kernel*, see assumption (H1) (iv) in Section 2.1 for a precise statement. This property reflects the quality of the approximation of the original dynamics. An important remark is that this *smoothing property* is only achieved when the freezing point, that is the point where the coefficients are evaluated in the approximation process, is chosen to be the terminal point in the transition density.

The main purpose of this paper is to develop a technique in order to prove weak uniqueness as well as existence of a transition density for some SDEs with path-functional coefficients where the probability law of the couple $(X_t, A_t(X))$ may be singular with respect to the Lebesgue measure on $\mathbb{R}^d$. The main new feature added here compared to previous works on this topic is that our technique enables us to deal with a process whose probability law is absolutely continuous with respect to a σ-finite measure.

Our methodology can be summed up as follows: suppose that the transition density of the Markov process $(W_t, A_t(W))_{t\geq 0}$ with initial point x exists with respect to a σ-finite measure $\nu(x, dy)$, not necessarily being the Lebesgue measure on $\mathbb{R}^d$, and a chain rule (Itô's) formula for $h(t, X_t, A_t(X))$ is available, where h belongs to a suitable class of functions $\mathcal{D}$ related to the domain of the infinitesimal generator of the Markov process $(W_t, A_t(W))_{t\geq 0}$. Then as soon as the derivatives of the transition density of $(W_t, A_t(W))_{t\geq 0}$ satisfies some good estimates or equivalently if the *parametrix kernel* enjoys a smoothing property with respect to $\nu(x, dy)$, one has the main tools to prove weak uniqueness for the SDE (1.1).

Since the probability law of the process $(X_t, A_t(X))_{t\geq 0}$ may be singular, it is not clear how to select the approximation process and even if this crucial *smoothing property* will be achieved in such context. To be more precise, let us look at the following example. If one considers the couple $(X_t, A_t(X))_{t\geq 0}$, $A_t(X) = L_t^0(X)$ being the local time at point 0 accumulated by X up to time t, then it is easy to see that on $\{T_0 > t\}$, T_0 being the first hitting time of 0 by X, one has $L_t^0(X) = 0$ whereas on $\{T_0 \leq t\}$, the process may accumulate local time so that the probability law of the couple $(X_t, L_t^0(X))$ consists in two parts, one being singular with an atom in the local time part, the other one (hopefully) being absolutely continuous with respect to the Lebesgue measure on $\mathbb{R} \times \mathbb{R}_+$. Hence, we see that for such dynamics the situation is more challenging than in the standard diffusion setting. This new difficulty will be overcome by choosing two independent parametrix kernels, one for each part, and then by proceeding with a non-trivial selection of the freezing point according to the singular measure induced by the approximation process. Even in this singular framework, one is able to prove that the *smoothing property* of the parametrix kernel still holds which is, as previously mentioned, the keystone to prove weak uniqueness for the SDE (1.1). As far as we know these feature appear to be new.

It will be apparent in what follows that our approach is not limited to the one-dimensional SDE case. We believe that the methodology developed here could be used to tackle multi-dimensional examples such as the one studied in [14] (up to to some technicalities). However, we decide to confine our presentation to the one-dimensional framework to foster understanding of the main arguments. As an illustration of our methodology we consider two examples. In the first one, we choose $A_t(X) = L_t^0(X)$, $t \geq 0$, where $L_t^0(X)$ is the local time at 0 accumulated by X up to time t. The bivariate density of Brownian motion and its running local time at 0 can be found e.g. in Karatzas and Shreve [11]. In the second example, we consider the running maximum of the process, that is, $A_t(X) = \max_{0\leq s\leq t} X_s$, $t \geq 0$.

The first part of our main results is an application of the general methodology developed in Section 2. More precisely, in Theorem 3.3 and Theorem 3.7 we prove that weak uniqueness holds for the SDE (1.1) when the path-functional is the local time or the running maximum, under the assumption that *the drift b is bounded measurable and the diffusion coefficient $a = \sigma^2$ is uniformly elliptic, bounded and η-Hölder-continuous, for some $\eta \in (0, 1]$.*

Finally, the strategy developed in this paper can be used not only to prove the existence, but also to retrieve an explicit representation of the transition density (with respect to a σ-finite measure) of $(X_t, A_t(X))_{t\geq 0}$ as an infinite series. However, one has to overcome new technical issues compared to the standard diffusion setting investigated by Friedman [8] or even to the degenerate case considered by Delarue and Menozzi [5]. Leaving this technical discussion to Section 3.3, we only point out that the main difficulty lies in the non-integrable time singularity induced by the mixing of the singular and non-singular parts of the parametrix kernel. In order to overcome this issue, which to our best knowledge appears to be new, the key idea is to use the symmetry of the density of the killed proxy with respect to the initial and terminal points, in order to retrieve the integrability in time of the underlying parametrix kernel. As the second part of our main results, we prove the existence of the transition density for $(X_t, A_t(X))_{t\geq 0}$ as well as its representation in infinite series for the two examples mentioned before, see Theorem 3.4 and Theorem 3.8 below. Some Gaussian upper-estimates are also established.

Notations and useful estimates:

We introduce here some basic notations and definitions used throughout this paper. For a sequence of linear operators $(S_i)_{1\leq i\leq n}$, we define $\prod_{i=1}^n S_i = S_1 \cdots S_n$. We will often use the convention $\prod_{\varnothing} = 1$ which appears when we have for example $\prod_{i=0}^{-1}$. Let $\mathcal{J}$ be a closed subset of $\mathbb{R}^d$ with non-empty interior, we denote by $\mathcal{C}_b^k(\mathcal{J})$, the collection of bounded continuous functions which are k-times continuously differentiable with bounded derivatives in the interior of $\mathcal{J}$. The derivatives at the boundary $\partial\mathcal{J}$ are defined as limits from the interior and it is assumed that they exist (and thus finite). The set $\mathcal{B}_b(\mathcal{J})$ is the collection of real-valued bounded measurable maps defined on $\mathcal{J}$. Furthermore we will use the following notation for time and space variables $\mathbf{s}_p = (s_1, \dots, s_p)$, $\mathbf{z}_p = (z_1, \dots, z_p)$, the differentials $d\mathbf{s}_p = ds_1 \cdots ds_p$, $d\mathbf{z}_p = dz_1 \cdots dz_p$ and for a fixed time $t \geq 0$, we denote by $\Delta_p(t) = \{\mathbf{s}_p \in [0,t]^p : s_{p+1} := 0 \leq s_p \leq s_{p-1} \leq \cdots \leq s_1 \leq t =: s_0\}$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_\ell)$ of length ℓ, we sometimes write $\partial_\alpha f(x) = \partial_{x_{\alpha_1}} \cdots \partial_{x_{\alpha_\ell}} f(x)$, for a vector x. We denote by $y \mapsto g(ct, y)$ the transition density function of the standard Brownian motion with constant diffusion coefficient $\sigma = \sqrt{c}$, i.e. $g(ct, y) = (2\pi ct)^{-1/2} \exp(-y^2/(2tc))$, $y \in \mathbb{R}$. Its derivatives of order i, $i \in \mathbb{N}$, are denoted by $H_i(ct, y) = \partial_y^i g(ct, y)$. For a given $z \in \mathbb{R}^d$, the Dirac measure at point z is denoted by $\delta_z(dx)$. For $a, b \in \mathbb{R}$, we use the notation $a \asymp b$ if there exists a constant $C > 1$ such that $C^{-1} b \leq a \leq Cb$. We denote by $|f|_\infty$ the supremum norm of a function f.

One of the key inequality that will be used intensively in this work is the following: for any $p, q > 0$ and $x \in \mathbb{R}$, $|x|^p e^{-qx^2} \leq (p/(2qe))^{p/2}$. As a direct consequence of this inequality, we have the *space-time inequality*,

$$\forall p, c > 0, \quad |x|^p H_0(ct, x) \leq C t^{p/2} H_0(c't, x) \tag{1.2}$$

which in turn gives the *standard Gaussian estimates* for the derivatives of Gaussian density, namely

$$\forall c, \forall i \in \mathbb{N}, \quad \left|H_i(ct, x)\right| \leq \frac{C}{t^{i/2}} H_0(c't, x) \tag{1.3}$$

for some positive constants C, c'. As these two inequalities are applied on numerous occasions throughout the paper, we will often omit to refer to them explicitly.

2. Abstract framework for weak uniqueness

2.1. *A perturbation formula*

Throughout this section, we assume that there exists a weak solution (X, W), $\{\mathcal{F}_t\}$ to (1.1) for every initial condition x_0 and that the process $Y_t := (X_t, A_t(X))_{t\geq 0}$, starting from the initial point x at time 0, has (almost surely) continuous sample paths living on a closed space $\mathcal{J} \subset \mathbb{R}^d$. The process $(X_t, A_t(X))_{t\geq 0}$ induces a probability measure $\mathbb{P}^x$ (or simply denoted $\mathbb{P}$) on $\Omega = \mathcal{C}([0,\infty), \mathcal{J})$ which is endowed with the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$. We consider the collection of linear maps $(P_t)_{t\geq 0}$ defined by $P_t h(x) := \mathbb{E}[h(X_t, A_t(X))]$ for $h \in \mathcal{B}_b(\mathcal{J})$. It is important to point out that, at this moment, we do not know whether Y is a (strong) Markov process or not. However, one of our main assumption (see assumption (H1) (ii) below) links the process Y to the solution of a martingale problem in the sense of Stroock and Varadhan [17]. As in the standard diffusion case, the strong Markov property will be a consequence of weak uniqueness.

Note that, for the moment, the coefficients b, σ are real-valued Borel measurable functions defined on $\mathcal{J}$ and σ is assumed to be continuous on $\mathcal{J}$. However to obtain the approximation process we will often work with the continuous extension of σ to $\mathbb{R}^d$ by using Tietze's extension theorem and freeze σ at some point $z \in \mathbb{R}^d$. Therefore for some fixed point $z \in \mathbb{R}^d$, we denote the approximation or the proxy process by $\bar{X}^z \equiv \bar{X}$, that is the solution of (1.1) with $b = 0$ and the diffusion coefficient σ evaluated at z. Without going into details, the key idea is to consider the process Y as a perturbation of the proxy $\bar{Y} \equiv \bar{Y}^z = (\bar{X}_t, A_t(\bar{X}))_{t \geq 0}$ whose law is denoted by $\bar{p}_t(x, dy) \equiv \bar{p}^z_t(x, dy)$. Accordingly, we define the collection of linear maps $(\bar{P}_t)_{t\geq 0}$ by $\bar{P}_t h(x) \equiv \bar{P}^z_t h(x) := \mathbb{E}[h(\bar{X}_t, A_t(\bar{X}))] = \int h(y)\bar{p}_t(x, dy)$ for $h \in \mathcal{B}_b(\mathcal{J})$. Note that the linear map $\bar{P}_t$ depends on z and, in general, to indicate that one is working with the approximation process with coefficients frozen at z or functions associated with the approximation process frozen at z, we put a bar on top of the function and, in order to simplify the notation, we will often omit to write the dependence with respect to z. To indicate that the frozen point is the terminal point y of the proxy density, we will put a hat instead of a bar.

We define the shift operators $(\theta_t y)(s) = y(t + s)$, $0 \leq s < \infty$ for $t \geq 0$ and $y \in \mathcal{C}([0, \infty), \mathcal{J})$. For any deterministic time $t \geq 0$, we denote by $\mathbb{Q}_{t,w}$ the regular conditional probability for $\mathbb{P}^x$ given $\mathcal{F}_t$ and let $\mathbb{P}_{t,w}$ be the probability measure defined by $\mathbb{P}_{t,w} := \mathbb{Q}_{t,w} \circ \theta_t^{-1}$. In particular, for every $F \in \mathcal{B}(\mathcal{C}([0, \infty), \mathcal{J}))$, one has $\mathbb{P}^x(\theta_t^{-1} F|\mathcal{F}_t)(w) = \mathbb{Q}_{t,w}(\theta_t^{-1} F) = \mathbb{P}_{t,w}(F)$, for $\mathbb{P}^x$-a.e. $w \in \Omega$. Our proof of weak uniqueness consists in first proving that two solutions of the martingale problem associated to (1.1) have the same one-dimensional marginal distribution and then passing to the full law of the process. Here we do not know that our solutions are part of a (strong) Markov family and in general, as in the standard diffusion case, this will be a consequence of weak uniqueness, or equivalently of the well-posedness of the martingale problem, see [12,17] or [3]. As a consequence we require the following substitute.

Assumption (H0). For every $x \in \mathcal{J}$, for every $t \in [0, \infty)$, there exists a $\mathbb{P}^x$-null event N of $\mathcal{F}_t$ such that, for every $w \notin N$, the probability measure $\mathbb{P}_{t,w}$ can be induced by a weak solution starting from $w(t)$ at time 0.

The above assumption allows one to prove that the finite-dimensional distributions are unique once we know that two weak solutions have the same one-dimensional marginal distributions, see e.g. Chapter 5 [12] or Chapter VI [3].

At this stage, we prefer not to state assumptions on the coefficients b and σ, but rather, we provide the required assumptions on the law of the approximation process $\bar{Y}$. Then we develop the methodology in order to establish a first order expansion of $(P_t)_{t\geq 0}$ and to prove that two solutions of the martingale problem share the same one-dimensional marginal distribution. These assumptions are then checked on the two examples investigated here: the local time at zero and the running maximum.

Assumption (H1). Given the initial condition $x \in \mathcal{J}$ and a fixed point $z \in \mathbb{R}^d$.

(i) (a) The proxy process $\bar{Y}^z$ is a Markov process with continuous sample paths and with infinitesimal generator $\bar{\mathcal{L}} \equiv \bar{\mathcal{L}}^z$ and $\partial_t \bar{P}^z_t h = \bar{\mathcal{L}}^z \bar{P}^z_t h = \bar{P}_t \bar{\mathcal{L}}^z h$ for $h \in \mathrm{Dom}(\bar{\mathcal{L}}^z)$ where

$$\mathrm{Dom}(\bar{\mathcal{L}}^z) := \left\{ h \in \mathcal{C}_b(\mathcal{J}) : \forall x \in \mathcal{J}, \bar{\mathcal{L}}^z h(x) := \lim_{t\downarrow 0} \frac{1}{t}\left[\bar{P}^z_t h - h\right](x) \text{ exists and } \bar{\mathcal{L}}^z h \in \mathcal{C}_b(\mathcal{J}) \right\}.$$

(b) There exists a σ-finite measure $\nu(x, \cdot)$ such that $\nu(\cdot, A)$ is Borel measurable for any $A \in \mathcal{B}(\mathbb{R}^d)$ and satisfying for all $t > 0$, the law of $\bar{Y}^z_t$ is absolutely continuous with respect to $\nu(x, \cdot)$. Moreover, there exists a $\nu(x, dy)$-integrable function $(t, x, y) \mapsto \bar{p}^z_t(x, y)$ satisfying

$$\bar{p}^z_t(x, dy) = \bar{p}^z_t(x, y)\nu(x, dy).$$

In particular, the Markov semi-group can be written $\bar{P}^z_t h(x) = \int h(y)\bar{p}^z_t(x, y)\nu(x, dy)$ for all $h \in \mathcal{B}_b(\mathcal{J})$. Also to make the notation simpler, when $z = y$, we write

$$\hat{p}_t(x, y) := \bar{p}^y_t(x, y)$$

and define accordingly

$$\widehat{P}_t h(x) := \int h(y)\widehat{p}_t(x, y)\nu(x, dy) = \int h(y)\bar{p}^y_t(x, y)\nu(x, dy). \tag{2.1}$$

(ii) There exists a subset of functions $\mathcal{D} \subset \mathcal{C}^{1,0}_b(\mathbb{R}_+ \times \mathcal{J})$ and an operator $\mathcal{L}$ acting on $\mathcal{D}$ such that:

(a) For all $h \in \mathcal{C}^\infty_b(\mathcal{J})$, $(0, \infty) \times \mathcal{J} \ni (t, x) \mapsto \bar{P}^z_t h(x) \in \mathcal{D}$ and for every fixed $t > 0$, $\bar{P}^z_t h \in \mathrm{Dom}(\bar{\mathcal{L}}^z)$.

(b) For all $h \in \mathcal{D}$, the process

$$h(t, y_t) - h(0, x) - \int_0^t \{\partial_1 + \mathcal{L}\}h(s, y_s)\, ds, \quad t \geq 0$$

is a continuous square integrable martingale under $\mathbb{P}^x$.

(c) There exists a *parametrix kernel* $\bar{\theta}_t$ with respect to the measure ν, that is, a measurable map $(t,z,x,y) \mapsto \bar{\theta}_t^z(x,y)$ such that for all $h \in \mathcal{C}_b^\infty(\mathcal{J})$ and all $t>0$

$$\big(\mathcal{L}-\bar{\mathcal{L}}^z\big)\bar{P}_t^z h(x) = \int h(y)\bar{\theta}_t^z(x,y)\nu(x,dy). \tag{2.2}$$

We again simplify the notation by writing $\widehat{\theta}_t(x,y) = \bar{\theta}_t^y(x,y)$ and define

$$\mathcal{S}_t h(x) := \int h(y)\widehat{\theta}_t(x,y)\nu(x,dy) = \int h(y)\bar{\theta}_t^y(x,y)\nu(x,dy). \tag{2.3}$$

(iii) For all $x,y \in \mathcal{J}$ and $t>0$, the maps $(t,z)\mapsto \bar{p}_t^z(x,y)$ and $(t,z)\mapsto \bar{\theta}_t^z(x,y)$ are continuous on $(0,\infty)\times\mathbb{R}^d$.

(iv) For all $t>0$, there exists some $\nu(x,dy)$-integrable functions $p_t^*(x,y)$, $\theta_t^*(x,y)$, a constant $\bar{\zeta}\in\mathbb{R}$ and a positive constant C, eventually depending on t but in a non-decreasing way, such that

$$\forall z\in\mathbb{R}^d, \quad \big|\bar{p}_t^z(x,y)\big| \le p_t^*(x,y), \qquad \big|\bar{\theta}_t^z(x,y)\big| \le \theta_t^*(x,y)$$

and

$$\int p_t^*(x,y)\nu(x,dy) \le C, \qquad \int \theta_t^*(x,y)\nu(x,dy) \le Ct^{\bar{\zeta}}.$$

For the case $z=y$, we assume that the parametrix kernel enjoys the following *smoothing property*: there exists $\zeta > -1$ and a positive constant C, eventually depending on t in a non-decreasing way, such that

$$\forall t>0, \forall x\in\mathcal{J}, \quad \int \big|\widehat{\theta}_t(x,y)\big|\nu(x,dy) \le Ct^{\zeta}. \tag{2.4}$$

(v) For any $h\in\mathcal{C}_b(\mathcal{J})$, one has

$$\lim_{t\downarrow 0}\widehat{P}_t h(x) = \lim_{t\downarrow 0}\int h(y)\bar{p}_t^y(x,y)\nu(x,dy) = h(x).$$

(vi) For any $h\in\mathcal{C}_b(\mathcal{J})$ and $t>0$,

$$\lim_{r\downarrow 0}\widehat{P}_{t+r}h(x) = \widehat{P}_t h(x) \quad \text{and} \quad \lim_{r\downarrow 0}\mathcal{S}_{t+r}h(x) = \mathcal{S}_t h(x)$$

where the operators $\widehat{P}_t$ and $\mathcal{S}_t$ are defined in (2.1) and (2.3) respectively.

Remark 2.1. The set of assumptions (H1) will allow us to prove a perturbation formula of the map P_t around $\widehat{P}_t$, see Theorem 2.2 below. The kernel of $\mathcal{S}_t$ defined above satisfies the smoothing property given by equation (2.4) which is, as mentioned in the introduction, the key point to prove uniqueness in law for equation (1.1). This smoothing property was exploited in [4] and in [14] for some degenerate Kolmogorov equations. The main new feature added here is that we are able to deal with a process that admits a density with respect to a σ-finite measure (with eventually several atoms). In particular, the process can be singular in the sense that it may not admit a transition density with respect to the Lebesgue measure on $\mathcal{J}$.

Assumption (H1) (ii) (b) provides a chain rule formula for the process $Y=(X_t,A_t(X))_{t\ge 0}$ for a suitable class of functions $\mathcal{D}$. The operator $\mathcal{L}$ is identified by means of this chain rule formula. As we will see in Section 3, this assumption will help us to formulate the martingale problem associated to the process Y. This will be used later on in order to establish the existence of a weak solution to the SDE (1.1).

Theorem 2.2. *Assume that* (H1) *holds. Then, for any* $h\in\mathcal{C}_b(\mathcal{J})$,

$$P_T h(x) = \widehat{P}_T h(x) + \int_0^T P_s\mathcal{S}_{T-s}h(x)\,ds.$$

Proof. Let $f \in \mathcal{C}_b^\infty(\mathcal{J})$. Let $r > 0$, by assumption (H1) (ii) (a) and (b) applied to $[0,T] \times \mathcal{J} \ni (t,x) \mapsto h(t,x) = \bar{P}_{T-t+r} f(x) \in \mathcal{D}$, there exists a continuous martingale $(\bar{M}_t)_{0 \le t \le T}$ starting at 0 such that

$$\begin{aligned}\bar{P}_{T-t+r} f(Y_t) &= \bar{P}_{T+r} f(x) + \int_0^t (\partial_s + \mathcal{L}) \bar{P}_{T-s+r} f(Y_s)\, ds + \bar{M}_t \\ &= \int f(y) \bar{p}^z_{T+r}(x,y) \nu(x,dy) + \iint_0^t f(y) \bar{\theta}^z_{T-s+r}(Y_s,y) \nu(Y_s,dy)\, ds + \bar{M}_t.\end{aligned}$$

We now proceed to the *diagonalisation argument*, that is the argument that allows one to select the freezing point z according to the measure $\nu(x,dy)$. For a fixed point $y \in \mathcal{J}$, we consider a sequence of non-negative mollifiers $\delta^z_\varepsilon(y)$, $\varepsilon > 0$ such that $\delta^z_\varepsilon(y) \le C_\varepsilon$ and $(\delta^z_\varepsilon(y))_{\varepsilon > 0}$ converges weakly to the Dirac mass at y as $\varepsilon \downarrow 0$, e.g. we set $\delta^z_\varepsilon(y) = g(\varepsilon, z-y)$. For any $h \in \mathcal{C}_b^\infty(\mathcal{J})$, we apply the above decomposition for $x \mapsto f(x) = \delta^z_\varepsilon(x) h(x)$ and take expectations. We obtain

$$P_t \bar{P}_{T-t+r} \delta^z_\varepsilon h(x) = \int \delta^z_\varepsilon(y) h(y) \bar{p}^z_{T+r}(x,y) \nu(x,dy) + \mathbb{E}\bigg[\iint_0^t \delta^z_\varepsilon(y) h(y) \bar{\theta}^z_{T-s+r}(Y_s,y) \nu(Y_s,dy)\, ds\bigg].$$

We let $t \uparrow T$ and integrate with respect to z over $\mathbb{R}^d$, from (H1) (iv) and Fubini's theorem, we obtain

$$\begin{aligned}&\int P_T \bar{P}_r \delta^z_\varepsilon h(x)\, dz \\ &\quad = \iint \delta^z_\varepsilon(y) h(y) \bar{p}^z_{T+r}(x,y) \nu(x,dy)\, dz + \iint_0^T \mathbb{E}\bigg[\int \delta^z_\varepsilon(y) h(y) \bar{\theta}^z_{T-s+r}(Y_s,y) \nu(Y_s,dy)\bigg] ds\, dz. \qquad (2.5)\end{aligned}$$

We now pass to the limit as $\varepsilon \downarrow 0$ and then $r \downarrow 0$ in (2.5). We first give some useful estimates. By using (H1) (iv), we have

$$\bigg|\int \delta^z_\varepsilon(y) \bar{p}^z_t(x,y)\, dz\bigg| \le p^*_t(x,y), \qquad (2.6)$$

$$\bigg|\int \delta^z_\varepsilon(y) \bar{\theta}^z_t(x,y)\, dz\bigg| \le \theta^*_t(x,y). \qquad (2.7)$$

Let us consider the left-hand side of (2.5). From (2.6), we can apply Fubini's theorem to pass the integral w.r.t dz inside to obtain

$$\begin{aligned}\int \mathbb{E}\big[\bar{P}_r \delta^z_\varepsilon h(Y_T)\big]\, dz &= \mathbb{E}\bigg[\iint \delta^z_\varepsilon(y) h(y) \bar{p}^z_r(Y_T,y) \nu(Y_T,dy)\, dz\bigg] \\ &= \mathbb{E}\bigg[\int \bigg\{\int \delta^z_\varepsilon(y) \bar{p}^z_r(Y_T,y)\, dz\bigg\} h(y) \nu(Y_T,dy)\bigg]. \qquad (2.8)\end{aligned}$$

Then the dominated convergence theorem together with (H1) (iii) allow us to conclude

$$\lim_{\varepsilon \downarrow 0} \int \mathbb{E}\big[\bar{P}_r \delta^z_\varepsilon h(Y_T)\big]\, dz = \mathbb{E}\bigg[\int \bar{p}^y_r(Y_T,y) h(y) \nu(Y_T,dy)\bigg] = \mathbb{E}\big[\widehat{P}_r h(Y_T)\big].$$

Consequently, by letting $r \downarrow 0$ and using (H1) (iv) and (v), we obtain again by the dominated convergence theorem $\lim_{r \downarrow 0} \mathbb{E}[\widehat{P}_r h(Y_T)] = P_T h(x)$.

For the right-hand side of (2.5). Again by (2.6) we can apply Fubini's theorem and then the dominated convergence theorem while having in mind (H1) (iii). This yields

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} \int \bar{P}_{T+r} \delta^z_\varepsilon h(x)\, dz &= \lim_{\varepsilon \downarrow 0} \int \bigg\{\int \delta^z_\varepsilon(y) h(y) \bar{p}^z_{T+r}(x,y)\, dz\bigg\} \nu(x,dy) \\ &= \int h(y) \bar{p}^y_{T+r}(x,y) \nu(x,dy) = \widehat{P}_{T+r} h(x). \qquad (2.9)\end{aligned}$$

Letting $r \downarrow 0$ in the previous equality, by (H1) (vi) we deduce that $\lim_{r \downarrow 0} \lim_{\varepsilon \downarrow 0} \int \bar{P}_{T+r} \delta^z_\varepsilon h(x)\, dz = \widehat{P}_T h(x)$.

We now handle the second term appearing on the right-hand side of (2.5). To pass to the limit as $\varepsilon \downarrow 0$, we first apply Fubini's theorem (using (2.7)) then the dominated convergence theorem using (H1) (iv) and (iii) so that

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0}\iint_0^T \mathbb{E}\bigg[\int \delta_\varepsilon^z(y)h(y)\bar{\theta}^z_{T-s+r}(Y_s,y)\nu(Y_s,dy)\bigg]ds\,dz \\
&\quad=\lim_{\varepsilon\downarrow 0}\mathbb{E}\bigg[\iint_0^T \bigg\{\int h(y)\delta_\varepsilon^z(y)\bar{\theta}^z_{T-s+r}(Y_s,y)\,dz\bigg\}\nu(Y_s,dy)\,ds\bigg] \\
&\quad=\mathbb{E}\bigg[\iint_0^T \lim_{\varepsilon\downarrow 0}\bigg\{\int h(y)\delta_\varepsilon^z(y)\bar{\theta}^z_{T-s+r}(Y_s,y)\,dz\bigg\}\nu(Y_s,dy)\,ds\bigg] \\
&\quad=\mathbb{E}\bigg[\iint_0^T h(y)\bar{\theta}^y_{T-s+r}(Y_s,y)\nu(Y_s,dy)\,ds\bigg]=\mathbb{E}\bigg[\int_0^T \mathcal{S}_{T-s+r}h(Y_s)\,ds\bigg].
\end{aligned}\tag{2.10}$$

To pass to the limit as $r \downarrow 0$, we remark that in (H1) (iv) equation (2.4), $\zeta > -1$, this allow dominated convergence to be used again which yields

$$\lim_{r\downarrow 0}\mathbb{E}\bigg[\int_0^T \mathcal{S}_{T-s+r}h(Y_s)\,ds\bigg]=\mathbb{E}\bigg[\int_0^T \lim_{r\downarrow 0}\mathcal{S}_{T-s+r}h(Y_s)\,ds\bigg]=\mathbb{E}\bigg[\int_0^T \mathcal{S}_{T-s}h(Y_s)\,ds\bigg]$$

where we applied (H1) (vi) for the last equality. The result is valid for $h \in \mathcal{C}_b^\infty(\mathcal{J})$ and an approximation argument completes the proof. □

We are not so far from obtaining a representation of $P_T h$ in infinite series. Once weak existence and uniqueness for the SDE (1.1) is established, this representation will be useful in order to establish the existence of a transition density for the process Y. We will also employ it to derive some Gaussian upper-bound estimates for the density of the couple $(X_T, A_T(X))$.

Corollary 2.1. *Assume that* (H1) *holds and that for any $h \in \mathcal{C}_b(\mathcal{J})$ and $t > 0$, the function $\mathcal{S}_t h$ defined in* (2.3) *belongs to $\mathcal{C}_b(\mathcal{J})$. Then one may iterate the first order formula in Theorem* 2.2 *to obtain*

$$P_T h(x) = \widehat{P}_T h(x) + \sum_{n\ge 1} I_T^n h(x) \tag{2.11}$$

with

$$I_T^n h(x) := \int_{\Delta_n(T)} d\mathbf{s}_n \widehat{P}_{s_n}\mathcal{S}_{s_{n-1}-s_n}\cdots\mathcal{S}_{T-s_1}h(x).$$

Moreover, the series (2.11) *converges absolutely and uniformly for $x \in \mathcal{J}$.*

Proof. We remark that since for all $t > 0$, $x \mapsto \mathcal{S}_t h(x) \in \mathcal{C}_b(\mathcal{J})$, we can iterate the first order expansion in Theorem 2.2 to obtain

$$P_T h(x) = \widehat{P}_T h(x) + \sum_{n=1}^{N-1}\int_{\Delta_n(T)} d\mathbf{s}_n \widehat{P}_{s_n}\mathcal{S}_{s_{n-1}-s_n}\cdots\mathcal{S}_{T-s_1}h(x) + \mathcal{R}_T^N h(x)$$

where the remainder term is given by

$$\mathcal{R}_T^N h(x) = \int_{\Delta_N(T)} d\mathbf{s}_N P_{s_N}\mathcal{S}_{s_{N-1}-s_N}\cdots\mathcal{S}_{T-s_1}h(x).$$

From iterative application of estimate (2.4), the remainder term is bounded by

$$\big|\mathcal{R}_T^N h(x)\big| \le C_T|h|_\infty \int_{\Delta_N(T)} d\mathbf{s}_N \prod_{n=0}^{N-1} C(s_n - s_{n-1})^\zeta \le C_T^N T^{N(1+\zeta)}\frac{\Gamma(1+\zeta)^N}{\Gamma(1+N(1+\zeta))}$$

where $\zeta \mapsto \Gamma(\zeta)$ is the Gamma function. By Lemma A.1 and the asymptotics of the Gamma function at infinity, we clearly see that the remainder goes to zero uniformly in $x \in \mathcal{J}$ as $N \uparrow \infty$. □

2.2. *Weak uniqueness*

Throughout this section, we will assume that (H1) is in force. Our aim is to prove weak uniqueness for the SDE (1.1). The main argument is an extension of the technique introduced by Bass and Perkins [4] in order to investigate the well-posedness of a martingale problem. It notably allows us to deal with singular probability law in the sense that the law of Y_t, $t > 0$, may not be absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$. Moreover, the new contribution of this section compared to the existing literature on this topic is that we identify the main assumptions, namely (H0), (H1) and (H2), required to establish weak uniqueness, thus allowing for a general treatment.

We consider two weak solutions of the SDE (1.1) starting at time 0 from the same initial point $x \in \mathcal{J}$. Denote by $\mathbb{P}_1$ and $\mathbb{P}_2$ the two probability measures induced on the space $(\mathcal{C}([0,\infty),\mathcal{J}), \mathcal{B}(\mathcal{C}([0,\infty),\mathcal{J})))$. We introduce the two resolvents associated to $\mathbb{P}_1$ and $\mathbb{P}_2$, defined for $h \in \mathcal{B}_b(\mathcal{J})$ and $\lambda > 0$, by

$$S^i_\lambda h(x) := \mathbb{E}_{\mathbb{P}_i}\left[\int_0^\infty e^{-\lambda t} h(y_t)\,dt\right] = \int_0^\infty e^{-\lambda t}\mathbb{E}_{\mathbb{P}_i}\big[h(y_t)\big]\,dt, \quad i = 1,2, \qquad S^\Delta_\lambda h(x) := \big(S^1_\lambda - S^2_\lambda\big)h(x)$$

$$\big\|S^\Delta_\lambda\big\| := \sup_{\|h\|_\infty \le 1}\big|S^\Delta_\lambda h\big|.$$

For $z \in \mathbb{R}^d$, the resolvent of the approximation process is defined by

$$\bar{R}^z_\lambda f(x) = \int_0^\infty e^{-\lambda t}\,\bar{P}^z_t h(x)\,dt, \quad \text{for } h \in \mathcal{B}_b(\mathcal{J}).$$

We make the following assumptions.

Assumption (H2). For any $z \in \mathbb{R}^d$ and any $h \in \mathcal{C}^\infty_b(\mathcal{J})$. The map $(t,x) \mapsto \bar{R}_\lambda \bar{P}_t h(x)$ belongs to $\mathcal{D}$ with $\mathcal{L}\bar{R}_\lambda \bar{P}_t h$ being bounded measurable and is such that for $r > 0$

$$(\lambda - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r h = \bar{R}_\lambda(\lambda - \bar{\mathcal{L}})\bar{P}_r h, \tag{2.12}$$

$$(\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r h = \int_0^\infty e^{-\lambda s}(\mathcal{L} - \bar{\mathcal{L}})\bar{P}_{r+s} h\,ds. \tag{2.13}$$

Lemma 2.1. *Under assumption* (H1) *and* (H2), *for* $h \in \mathcal{C}^\infty_b(\mathcal{J})$

(i) $\bar{R}_\lambda(\lambda - \bar{\mathcal{L}})\bar{P}_r h = \bar{P}_r h.$
(ii) $S^i_\lambda(\lambda - \mathcal{L})\bar{R}_\lambda \bar{P}_r h = \bar{R}_\lambda \bar{P}_r h$, *for* $i = 1,2$.

Proof. To see (i), from the very definition of $\bar{R}_\lambda$ and (H1) (i) (a) and (ii) (a), that is, the fact that $(\bar{P}_t)_{t\ge 0}$ is a Markov semi-group with infinitesimal generator $\bar{\mathcal{L}}$ satisfying $\bar{P}_t\bar{\mathcal{L}}\bar{P}_r h = \bar{\mathcal{L}}\bar{P}_t\bar{P}_r h = \bar{\mathcal{L}}\bar{P}_{t+r}h = \partial_t \bar{P}_{t+r}h$, we get

$$\begin{aligned}\bar{R}_\lambda(\lambda - \bar{\mathcal{L}})\bar{P}_r h &= \int_0^\infty e^{-\lambda t}\,\bar{P}_t(\lambda - \bar{\mathcal{L}})\bar{P}_r h\,dt \\ &= \int_0^\infty e^{-\lambda t}(\lambda - \partial_t)\bar{P}_{t+r} h\,dt = \bar{P}_r h,\end{aligned}$$

where we used the continuity of $t \mapsto \bar{P}_{t+r}h$, which is a direct consequence of the continuity of the sample path of $(\bar{X}, A(\bar{X}))$.

To see (ii), let $h \in \mathcal{C}^\infty_b(\mathbb{R}^d)$ since $\bar{R}_\lambda \bar{P}_r h \in \mathcal{D}$ by (H1) (ii) (b) we get

$$e^{-\lambda t}\bar{R}_\lambda \bar{P}_r h(Y_t) = \bar{R}_\lambda \bar{P}_r h(x) + \int_0^t -\lambda e^{-\lambda s}\bar{R}_\lambda \bar{P}_r h(Y_s)\,ds + \int_0^t e^{-\lambda s}\mathcal{L}\bar{R}_\lambda \bar{P}_r h(Y_s)\,ds + \bar{M}_t$$

where $\bar{M}$ is a square integrable martingale. By taking the expectation and using the fact that $\mathcal{L}\bar{R}_\lambda \bar{P}_r h$ is bounded, we have by the Fubini theorem

$$e^{-\lambda t}\mathbb{E}_{\mathbb{P}^i}\big[\bar{R}_\lambda \bar{P}_r h\big(y(t)\big)\big] = \bar{R}_\lambda \bar{P}_r h(x) - \int_0^t e^{-\lambda s}\mathbb{E}_{\mathbb{P}_i}\big[(\lambda - \mathcal{L})\bar{R}_\lambda \bar{P}_r h(y_s)\big]\,ds.$$

Finally, the dominated convergence theorem is used to pass to the limit as $t \uparrow \infty$ in the previous equality and this completes the proof of (ii). $\square$

Let $z \in \mathbb{R}^d$ and $r > 0$. We consider the sequence of non-negative mollifiers $\{\delta_\varepsilon^z(y) = g(\varepsilon, z - y), \varepsilon > 0\}$. Let us first observe that if $h \in \mathcal{C}_b^\infty(\mathcal{J})$ then from assumption (H2) one has

$$\begin{aligned}(\lambda - \mathcal{L})\bar{R}_\lambda \bar{P}_r \delta_\varepsilon^z h(x) &= (\lambda - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r \delta_\varepsilon^z h(x) - (\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r \delta_\varepsilon^z h(x) \\ &= \bar{P}_r \delta_\varepsilon^z h(x) - (\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r \delta_\varepsilon^z h(x).\end{aligned} \tag{2.14}$$

Note that by using (2.13) the second term appearing in the right-hand side of the above equality is given by

$$(\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_r \delta_\varepsilon^z h(x) = \int_0^\infty e^{-\lambda t}(\mathcal{L} - \bar{\mathcal{L}})\bar{P}_{t+r}\delta_\varepsilon^z h(x)\,dt. \tag{2.15}$$

We are now ready to prove weak uniqueness for (1.1).

Theorem 2.3. *Assume that Assumptions* (H0), (H1) *and* (H2) *are satisfied*, *then weak uniqueness holds for the SDE given in* (1.1).

Proof. Let $z \in \mathbb{R}^d$ be the freezing point. The first part of the proof is similar to that of Theorem 2.2. We integrate both sides of (2.14) over $\mathbb{R}^d$ with respect to dz, apply S_λ^Δ and then pass to the limit as $\varepsilon, r \downarrow 0$. For $i = 1, 2$, we first consider $(\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h$ using (2.14) and (2.15), then by using estimates (2.6) and the inequality $|(\mathcal{L} - \bar{\mathcal{L}}^z)\bar{P}_{t+r}^z \delta_\varepsilon^z h(x)| \le C_\varepsilon |h|_\infty r^{\bar{\zeta}}$, we can apply Fubini's theorem so that

$$S_\lambda^i \int (\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz = \int_0^\infty e^{-\lambda t}\mathbb{E}_{\mathbb{P}_i}\left[\int (\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h(y_t)\,dz\right]dt = \int S_\lambda^i (\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz,$$

and from Lemma 2.1 (ii)

$$S_\lambda^\Delta\left(\int (\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz\right) = \int S_\lambda^\Delta (\lambda - \mathcal{L})\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz = 0.$$

Then, from (2.14) we deduce

$$S_\lambda^\Delta\left(\int \bar{P}_r^z \delta_\varepsilon^z h\,dz\right) - S_\lambda^\Delta\left(\int (\mathcal{L} - \bar{\mathcal{L}}^z)\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz\right) = 0.$$

Let us consider the first term in the above expression and take the limit first as $\varepsilon \downarrow 0$ and then as $r \downarrow 0$. For $i = 1, 2$, under (H1), the limit as $\varepsilon \downarrow 0$ can be taken using the dominated convergence theorem while the limit as $r \downarrow 0$ follows from (H1) (v) (similarly to (2.9))

$$\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int_0^\infty e^{-\lambda t}\mathbb{E}_{\mathbb{P}_i}\left[\int \bar{P}_r^z \delta_\varepsilon^z h(y_t)\,dz\right]dt = \lim_{r\downarrow 0}\int_0^\infty e^{-\lambda t}\mathbb{E}_{\mathbb{P}_i}\left[\widehat{P}_r h(y_t)\right]dt = \int_0^\infty e^{-\lambda t}\mathbb{E}_{\mathbb{P}_i}\left[h(y_t)\right]dt.$$

This shows that $\lim_{r,\varepsilon\downarrow 0} S_\lambda^\Delta(\int \bar{P}_r^z \delta_\varepsilon^z h\,dz)(x) = S_\lambda^\Delta h(x)$. For the second term, we first rewrite it using (2.15)

$$\begin{aligned}S_\lambda^i\left(\int (\mathcal{L} - \bar{\mathcal{L}}^z)\bar{R}_\lambda^z \bar{P}_r^z \delta_\varepsilon^z h\,dz\right) &= \mathbb{E}_{\mathbb{P}_i}\left[\int_0^\infty e^{-\lambda u}\left\{\iint_0^\infty e^{-\lambda t}(\mathcal{L} - \bar{\mathcal{L}}^z)\bar{P}_{t+r}^z \delta_\varepsilon^z h(y_u)\,dt\,dz\right\}du\right] \\ &= \mathbb{E}_{\mathbb{P}_i}\left[\int_0^\infty e^{-\lambda u}\left\{\iint_0^\infty e^{-\lambda t}\int \bar{\theta}_{t+r}^z(y_u, y)\delta_\varepsilon^z(y)h(y)\nu(y_u, dy)\,dt\,dz\right\}du\right].\end{aligned}$$

To pass to the limit as $\varepsilon \downarrow 0$, we check that Fubini's and the dominated convergence theorems can be applied to the right hand side of the above equation. We consider the term in the curly brackets. We use (2.7) and then (H1) (iv) to obtain for $i = 1, 2$,

$$\iint_0^\infty e^{-\lambda t}\int \left|\bar{\theta}_{t+r}^z(y_u, y)\right|\delta_\varepsilon^z(y)h(y)\nu(y_u, dy)\,dt\,dz$$

$$= \int_0^\infty e^{-\lambda t} \left\{ \iiint \left|\bar{\theta}^z_{t+r}(y_u, y)\right| \delta^z_\varepsilon(y)\, dz \left|h(y)\right| \nu(y_u, dy) \right\} dt$$
$$\le |h|_\infty \int_0^\infty e^{-\lambda t} \left\{ \int \bar{\theta}^*_{t+r}(y_u, y) \nu(y_u, dy) \right\} dt \le \frac{1}{\lambda} C |h|_\infty r^{\bar{\zeta}}$$

which is $d\mathbb{P}^i \times e^{-\lambda u}\, du$ integrable. More importantly, we observe that, in the first equality above, the term in the curly bracket can be bounded by a constant independent of ε. One can then apply Fubini's theorem and the dominated convergence theorem to obtain

$$\lim_{\varepsilon \downarrow 0} \int_0^\infty e^{-\lambda u} \int_0^\infty e^{-\lambda t} \left\{ \mathbb{E}_{\mathbb{P}_i} \left[\int_{\mathbb{R}^d} \left(\mathcal{L} - \bar{\mathcal{L}}^z\right) \bar{P}^z_{t+r} \delta^z_\varepsilon h(y_u)\, dz \right] \right\} dt\, du$$
$$= \int_0^\infty e^{-\lambda u} \int_0^\infty e^{-\lambda t} \lim_{\varepsilon \downarrow 0} \mathbb{E}_{\mathbb{P}_i} \left[\int_{\mathbb{R}^d} \left(\mathcal{L} - \bar{\mathcal{L}}^z\right) \bar{P}^z_{t+r} \delta^z_\varepsilon h(y_u)\, dz \right] dt\, du$$
$$= \int_0^\infty e^{-\lambda u} \mathbb{E}_{\mathbb{P}_i} \left[\int_0^\infty e^{-\lambda t} \mathcal{S}_{t+r} h(y_u)\, dt \right] du,$$

where the last equality follows from similar arguments as those employed in (2.10). One can now let r goes to zero by using the estimate (2.4) in (H1) (iv) and (vi) to obtain

$$\lim_{r \downarrow 0} \int_0^\infty e^{-\lambda u} \mathbb{E}_{\mathbb{P}_i} \left[\int_0^\infty e^{-\lambda t} \mathcal{S}_{t+r} h(y_u)\, dt \right] du = \int_0^\infty e^{-\lambda u} \mathbb{E}_{\mathbb{P}_i} \left[\int_0^\infty e^{-\lambda t} \mathcal{S}_t h(y_u)\, dt \right] du.$$

By putting the two terms together, we obtain

$$S^\Delta_\lambda h = S^\Delta_\lambda \left(\int_0^\infty e^{-\lambda t} \mathcal{S}_t h\, dt \right)$$

and one can pick λ such that

$$\left| \int_0^\infty e^{-\lambda t} \mathcal{S}_t h\, dt \right| \le |h|_\infty \int_0^\infty e^{-\lambda t} t^\zeta\, dt = |h|_\infty \frac{\Gamma(\zeta)}{\lambda^{1+\zeta}} < \frac{1}{2} |h|_\infty.$$

From the above computation and the definition of $\|S^\Delta_\lambda\|$, we find that

$$\left|S^\Delta_\lambda h\right| = \left| S^\Delta_\lambda \left(\int_0^\infty e^{-\lambda t} \mathcal{S}_t h\, dt \right) \right| \le \frac{1}{2} \left\|S^\Delta_\lambda\right\| |h|_\infty.$$

By an approximation argument, the last inequality remains valid for real-valued bounded continuous functions h supported in $\mathcal{J}$. Taking the supremum over $|h|_\infty \le 1$ yields $\|S^\Delta_\lambda\| \le \frac{1}{2}\|S^\Delta_\lambda\|$ and, since $\|S^\Delta_\lambda\| < \infty$, we conclude that $S^\Delta_\lambda = 0$. Consequently, $\int_0^\infty e^{-\lambda t} \mathbb{E}_{\mathbb{P}_1}[h(y_t)]\, dt = \int_0^\infty e^{-\lambda t} \mathbb{E}_{\mathbb{P}_2}[h(y_t)]\, dt$. By the continuity of sample paths together with the uniqueness of the Laplace transform together we get $\mathbb{E}_{\mathbb{P}_1}[h(y_t)] = \mathbb{E}_{\mathbb{P}_2}[h(y_t)]$ for all $t \ge 0$ if h is bounded continuous. By a monotone class argument, this also extends to bounded measurable functions so that $\mathbb{P}_1$ and $\mathbb{P}_2$ share the same one-dimensional marginal distribution.

Now one can use the standard argument based on regular conditional probabilities to show that the finite dimensional distributions of the process $(Y_t)_{t \ge 0} = (X_t, A_t(X))_{t \ge 0}$ agree under $\mathbb{P}_1$ and $\mathbb{P}_2$. This is where we employ assumption (H0) on regular conditional probability measure. Since this argument is standard, we omit it. The above is sufficent to prove weak uniqueness, see [17], Section 5.4.C in [12] or Section VI.2 in [3]. □

3. Two examples

3.1. *A diffusion process and its running local time*

In this example, we consider the SDE with dynamics

$$X_t = x + \int_0^t b\left(X_s, L^0_s(X)\right) ds + \int_0^t \sigma\left(X_s, L^0_s(X)\right) dW_s \tag{3.1}$$

where $L_s^0(X) := \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}\int_0^s \mathbb{I}_{\{0\le X_u<\epsilon\}}\, d\langle X\rangle_u$ is the local time at 0 accumulated by X at time s. Here $\mathcal{J} = \mathbb{R}\times\mathbb{R}_+$, $A_t(X) = L_t^0(X)$ and $d = 2$. We introduce the following assumptions:

(R-η) The coefficients b and $a = \sigma^2$ are bounded measurable functions defined on $\mathbb{R}\times\mathbb{R}_+$. The diffusion coefficient a is η-Hölder continuous on $\mathbb{R}\times\mathbb{R}_+$.
(UE) There exists some constant $\underline{a} > 0$ such that $\forall (x,\ell)\in\mathbb{R}\times\mathbb{R}_+$, $\underline{a}\le a(x,\ell)$.

Remark 3.1. In (R-η) and (UE), note that the diffusion coefficient $a : \mathcal{J}\to\mathbb{R}$ is defined and assumed to be bounded, continuous and uniformly elliptic on the closed subset $\mathcal{J}\subset\mathbb{R}^2$. To ensure that (H1) (iii) is satisfied and the previously developed machinery can be applied, in the following, wherever needed, we make use of the Tietze extension theorem and consider the bounded, continuous and uniformly elliptic extension of the diffusion coefficient a to $\mathbb{R}^2$, which we also denote by a.

Let $\mathcal{D}$ be the class of function $h\in\mathcal{C}_b^{1,2,1}(\mathbb{R}_+\times\mathbb{R}\backslash\{0\}\times\mathbb{R}_+)\cap\mathcal{C}_b(\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+)$ such that for all $t\ge 0$, $\partial_2 h(t,0+,\ell) = \lim_{x\downarrow 0}\frac{h(t,x,\ell)-h(t,0,\ell)}{x}$ and $\partial_2 h(t,0-,\ell) = \lim_{x\uparrow 0}\frac{h(t,x,\ell)-h(t,0,\ell)}{x}$ exist, are finite and satisfy the following transmission condition:

$$\forall (t,\ell)\in(\mathbb{R}_+)^2,\quad \frac{\partial_2 h(t,0+,\ell)-\partial_2 h(t,0-,\ell)}{2} + \partial_3 h(t,0,\ell) = 0. \tag{3.2}$$

We define the linear operator $\mathcal{L}$ by:

$$\mathcal{L}h(t,x,\ell) = b(x,\ell)\partial_2 h(t,x-,\ell) + \frac{1}{2}a(x,\ell)\partial_2^2 h(t,x-,\ell),\quad (t,x,\ell)\in\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+. \tag{3.3}$$

As mentioned in the Introduction, we need a chain rule formula for the process $(X_t, A_t(X))_{t\ge 0}$ which allows to identify the set of functions $\mathcal{D}$ and the linear operator $\mathcal{L}$. In fact, the set $\mathcal{D}$ described above is precisely the set of functions for which we are able to provide a good characterisation of the martingale problem. Indeed, one has to rely on the following chain rule formula or generalisation of the Itô formula whose proof closely follows the arguments of Theorem 2.2 in Elworthy et al. [6] or Theorem 2.1 in Peskir [15]. Note that here that we are working with the local time at zero whereas the Lévy's definition of local time is considered in [6].

Proposition 3.1 (Generalised Itô's formula). *Assume that $h\in\mathcal{C}^{1,2,1}(\mathbb{R}_+\times\mathbb{R}\backslash\{0\}\times\mathbb{R}_+)\cap\mathcal{C}(\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+)$ satisfies: $\partial_2 h(t,0+,\ell) = \lim_{x\downarrow 0}(h(t,x,\ell)-h(t,0,\ell))/x$ and $\partial_2 h(t,0-,\ell) = \lim_{x\uparrow 0}(h(t,x,\ell)-h(t,0,\ell))/x$ exist and are finite. Then, almost surely for all $t\ge 0$, one has*

$$\begin{aligned} h\big(t,X_t,L_t^0(X)\big) &= h(0,x,0) + \int_0^t \{\partial_1+\mathcal{L}\}h\big(s,X_s,L_s^0(X)\big)\,ds + \int_0^t \sigma\big(X_s,L_s^0(X)\big)\partial_2 h\big(s,X_s-,L_s^0(X)\big)\,dW_s \\ &\quad + \int_0^t \left\{\frac{\partial_2 h(s,0+,L_s^0(X)) - \partial_2 h(s,0-,L_s^0(X))}{2} + \partial_3 h\big(s,0,L_s^0(X)\big)\right\} dL_s^0(X). \end{aligned}$$

3.2. *Weak existence*

To this end, we have identified the set $\mathcal{D}$ and the linear operator $\mathcal{L}$, the weak existence of a solution to (3.1), which is equivalent to the existence of a solution to the following martingale problem, follows from a standard compactness argument.

We will say that a probability measure P on $(\mathcal{C}([0,\infty),\mathbb{R}\times\mathbb{R}_+), \mathcal{B}(\mathcal{C}([0,\infty),\mathbb{R}\times\mathbb{R}_+)))$ endowed with the canonical filtration $(\mathcal{F}_t)_{t\ge 0}$ is a solution to the local martingale problem if $P(y_1(0)=x, y_2(0)=0)=1$, $t\mapsto y_2(t)$ is a non-decreasing process P-$a.s.$ and

$$h\big(y(t)\big) - h\big(y(0)\big) - \int_0^t \mathcal{L}h\big(y(s)\big)\,ds \tag{3.4}$$

is a continuous local martingale for $h\in\mathcal{C}^{2,1}(\mathbb{R}\backslash\{0\}\times\mathbb{R}_+)\cap\mathcal{C}(\mathbb{R}\times\mathbb{R}_+)$ satisfying the transmission condition (3.2). We will say that P is a solution to the martingale problem if (3.4) is a continuous square integrable martingale for every $h\in\mathcal{D}$. Similarly to the standard diffusion case (see e.g. Proposition 5.4.11 in [12]), since the coefficients b and σ are bounded, existence of a solution to the local martingale problem is equivalent to the existence of solution to the (non-local) martingale problem. Moreover, existence of a solution to the local martingale problem is equivalent to the existence of a weak solution to the functional SDE (3.1) and the measure $\mathbb{P}_{t,w} = \mathbb{Q}_{t,w} \circ \theta_t^{-1}$, where $\mathbb{Q}_{t,w}$ is a regular conditional probability for P given $\mathcal{F}_t$, solves the martingale problem for every $w \notin N$, $N \in \mathcal{F}_t$ being a P-null event. Hence, assumption (H0) is satisfied.

We are now in position to state the existence of a weak solution to the SDE (3.1). The lines of reasoning here are rather standard and are based on a compactness argument, see e.g. Theorem 5.4.22 in [12], or Girsanov transform, see e.g. Proposition 5.3.6 in [12], so that we omit the proof of the following result.

Theorem 3.2. *Assume that the coefficients b, $\sigma : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ are bounded and continuous functions. Then, for every $x \in \mathbb{R}$, there exists a weak solution to the SDE* (3.1). *Assume that b is bounded measurable and that a is a bounded continuous function satisfying* (UE), *then the previous result still holds.*

3.3. *Weak uniqueness and representation of the transition density*

We now introduce the proxy process $\bar{X}_t := x_0 + \sigma(z_1) W_t$, $t \geq 0$, obtained from the original process X by removing the drift part and by freezing the diffusion coefficient at $z_1 = (x_1, \ell_1) \in \mathbb{R}^2$ (recall that σ has been extended on $\mathbb{R}^2$ by Tietze's extension theorem). For $h \in \mathcal{C}_b(\mathbb{R} \times \mathbb{R}_+)$, we define

$$\bar{P}_t h(x_0, \ell_0) = \mathbb{E}\big[h\big(\bar{X}_t, \ell_0 + L_t^0(\bar{X})\big)\big] = \mathbb{E}\big[h\big(x_0 + \bar{\sigma} W_t, \ell_0 + L_t^0(\bar{X})\big)\big]$$

where, for sake of simplicity, we set $\bar{\sigma} = \sigma(z_1)$, $\bar{a} = \bar{\sigma}^2$. We also introduce the operator $\bar{\mathcal{L}}$ acting on functions $h \in \mathcal{D}$

$$\bar{\mathcal{L}} h(t, x, \ell) = \frac{1}{2} \bar{a} \partial_2^2 h(t, x-, \ell), \quad (t, x, \ell) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+. \tag{3.5}$$

We now compute the bivariate transition density of the approximation process $(\bar{X}_t, L_t^0(\bar{X}))_{t \geq 0}$ from the joint density of $(W_t, L_t^0(W))$ which is readily available from Karatzas and Shreve [11]. To this end, we denote by $T_0 = \inf\{t \geq 0 : x_0 + \sigma(z_1) W_t = 0\}$ the first hitting time of 0 by the process $(\bar{X}_t)_{t \geq 0}$. Let $h \in \mathcal{C}_b(\mathbb{R} \times \mathbb{R}_+)$. We make use of the following decomposition:

$$\bar{P}_t h(x_0, \ell_0) := \mathbb{E}\big[h(x_0 + \bar{\sigma} W_t, \ell_0) \mathbb{I}_{\{T_0 \geq t\}}\big] + \mathbb{E}\big[h\big(x_0 + \bar{\sigma} W_t, \ell_0 + L_t^0(\bar{X})\big) \mathbb{I}_{\{T_0 < t\}}\big] =: \mathrm{I} + \mathrm{II}.$$

In the first term, the process $\bar{X}$ does not accumulate local time at zero. The bivariate density of $(W_t, \max_{0 \leq s \leq t} W_s)$, see e.g. [12], gives

$$\mathrm{I} = \int_{\mathbb{R} \times \mathbb{R}_+} h(x, \ell) \big\{H_0(\bar{a}t, x - x_0) - H_0(\bar{a}t, x + x_0)\big\} \mathbb{I}_{\{x_0 x \geq 0\}} \, dx \, \delta_{\ell_0}(d\ell).$$

To compute II we make use of the bivariate density of $(W_t, L_t^0(W))_{t \geq 0}$ established in [11]. Conditioning with respect to T_0 and using the strong Markov property of W yield

$$\begin{aligned}
\mathrm{II} &= \int_0^t \mathbb{P}(T_0 \in ds) \mathbb{E}\big[h\big(x_0 + \bar{\sigma} W_t, \ell_0 + L_t^0(\bar{X})\big) | T_0 = s\big] \\
&= \int_0^t \mathbb{P}(T_0 \in ds) \mathbb{E}\big[h\big(\bar{\sigma} W_{t-s}, \ell_0 + \bar{\sigma} L_{t-s}^0(W)\big)\big] \\
&= \int_0^t ds (-H_1)\left(s, \frac{|x_0|}{\bar{\sigma}}\right) \int_{\mathbb{R} \times \mathbb{R}_+} h(\bar{\sigma} x, \ell_0 + \bar{\sigma} \ell) (-H_1)\big(t - s, |x| + \ell\big) \, dx \, d\ell
\end{aligned}$$

where we used the exact expression for the density of the passage time T_0, which is given by $\mathbb{P}(T_0 \in ds) = (-H_1)(s, |x_0|/\bar{\sigma}) \, ds$, $s > 0$ and $x_0 \in \mathbb{R}$. Since the sum of independent passage times is again a passage time, see e.g. page 824 of Karatzas and Shreve [11], one has

$$(-H_1)\big(t, |x| + |y|\big) = \int_0^t ds (-H_1)\big(t - s, |x|\big)(-H_1)\big(s, |y|\big), \quad x, y \neq 0, t > 0. \tag{3.6}$$

Combining these observations with Fubini's theorem and a change of variable yield

$$\mathrm{II} = \int_{\mathbb{R} \times \mathbb{R}_+} h(x, \ell) \frac{1}{\bar{a}} (-H_1)\left(t, \frac{|x| + |x_0| + \ell - \ell_0}{\bar{\sigma}}\right) \mathbb{I}_{\{\ell_0 \leq \ell\}} \, dx \, d\ell.$$

From the expression of I and II, we see that $(\bar{X}_t, \ell_0 + L^0_t(\bar{X}))$ admits a density $(x, \ell) \mapsto \bar{p}_t(x_0, \ell_0, x, \ell)$, for $t > 0$, given by

$$\begin{aligned}
&\bar{p}_t(x_0, \ell_0, dx, d\ell) = \bar{p}_t(x_0, \ell_0, x, \ell)\nu(x_0, \ell_0, dx, d\ell)\\
&\bar{p}_t(x_0, \ell_0, x, \ell) := \bar{f}_t(x_0, x)\mathbb{I}_{\{\ell=\ell_0\}} + \bar{q}_t(x_0, \ell_0, x, \ell)\mathbb{I}_{\{\ell_0<\ell\}}
\end{aligned} \tag{3.7}$$

with

$$\begin{aligned}
&\bar{f}_t(x_0, x) := H_0(\bar{a}t, x - x_0) - H_0(\bar{a}t, x + x_0),\\
&\bar{q}_t(x_0, \ell_0, x, \ell) := -\frac{1}{\bar{a}} H_1\big(t, \big(|x| + |x_0| + \ell - \ell_0\big)/\bar{\sigma}\big),\\
&\nu(x_0, \ell_0, dx, d\ell) := \mathbb{I}_{\{\ell_0<\ell\}}\, dx\, d\ell + \mathbb{I}_{\{x_0 x\geq 0\}}\, dx \delta_{\ell_0}(d\ell).
\end{aligned}$$

Moreover, as already mentioned in assumption (H1) in Section 2.1, we let $\widehat{p}_t(x_0, \ell_0, x, \ell) = \widehat{f}_t(x_0, x)\mathbb{I}_{\{\ell=\ell_0\}} + \widehat{q}_t(x_0, \ell_0, x, \ell)\mathbb{I}_{\{\ell_0<\ell\}}$ with

$$\begin{aligned}
&\widehat{f}_t(x_0, x) := H_0\big(a(x, \ell_0)t, x - x_0\big) - H_0\big(a(x, \ell_0)t, x + x_0\big),\\
&\widehat{q}_t(x_0, \ell_0, x, \ell) := -\frac{1}{a(x, \ell)} H_1\big(t, \big(|x| + |x_0| + \ell - \ell_0\big)/\sigma(x, \ell)\big).
\end{aligned}$$

Hence, we see that both measures $\bar{p}_t$ and $\widehat{p}_t$ consist of two parts. The first part is absolutely continuous with respect to the σ-finite measure $\mathbb{I}_{\{x_0 x\geq 0\}}\, dx\delta_{\ell_0}(d\ell)$. Here the approximation process is obtained by freezing the diffusion coefficient at (x, ℓ_0). This is a natural idea since in this part the process $\bar{X}$ does not accumulate local time at zero and ℓ_0 is both the initial and terminal point of the density. The second part is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R} \times \mathbb{R}_+$. The approximation process is obtained by freezing the diffusion coefficient at the terminal point of the density as in the standard diffusion case.

The first main result of this section establishes weak uniqueness for the SDE (3.1) by proving that assumptions (H1) and (H2) of Section 2.1 are satisfied. Its proof is given in the Appendix.

Theorem 3.3. *For $\eta \in (0, 1]$, under* (R-η) *and* (UE), *weak uniqueness holds for the SDE* (3.1).

In the following, we show that, given $(x_0, \ell_0) \in \mathbb{R} \times \mathbb{R}_+$, the transition density of the process $(X_t, \ell_0 + L^0_t(X))_{t\geq 0}$ is absolutely continuous with respect to the sigma finite measure $\nu(x_0, \ell_0, dx, d\ell)$. Our strategy consists in establishing a representation in infinite series of $P_t h$ from which stems an explicit representation of the density of the couple $(X_t, \ell_0 + L^0_t(X))$, see Theorem 3.4. Though we will not proceed in that direction, we point out that this representation may be useful in order to study the regularity properties of the density, to obtain integration by parts formulas or to derive an unbiased Monte Carlo simulation method. We refer the interested reader to [9] for some results in that direction related to the first hitting times of one-dimensional elliptic diffusions.

To this end, we need to iterate the first step expansion obtained in Theorem 2.2. We recall that

$$\begin{aligned}
\mathcal{S}_t h(x_0, \ell_0) = \int h(x, \ell)\bigg\{ &-\frac{1}{2}\frac{(a(x_0, \ell_0) - a(x, \ell))}{a^2(x, \ell)} H_3\bigg(t, \frac{|x| + |x_0| + \ell - \ell_0}{\sigma(x, \ell)}\bigg)\\
&+ \frac{b(x_0, \ell_0)\mathrm{sign}(x_0)}{a^{\frac{3}{2}}(x, \ell)} H_2\bigg(t, \frac{|x| + |x_0| + \ell - \ell_0}{\sigma(x, \ell)}\bigg)\bigg\}\mathbb{I}_{\{\ell>\ell_0\}}\, dx\, d\ell\\
&+ \int h(x, \ell_0)\bigg\{\frac{1}{2}\big(a(x_0, \ell_0) - a(x, \ell_0)\big)\big\{H_2\big(a(x, \ell_0)t, x - x_0\big) - H_2\big(a(x, \ell_0)t, x + x_0\big)\big\}\\
&- b(x_0, \ell_0)\big\{H_1\big(a(x, \ell_0)t, x - x_0\big) - H_1\big(a(x, \ell_0)t, x + x_0\big)\big\}\bigg\}\mathbb{I}_{\{x x_0\geq 0\}}\, dx
\end{aligned} \tag{3.8}$$

with $\mathrm{sign}(x_0) := -\mathbb{I}_{\{x_0\leq 0\}} + \mathbb{I}_{\{x_0>0\}}$. The dissymetric version of the sign function appears here due to the fact that one has to consider the left-derivative in both operators $\mathcal{L}$ and $\bar{\mathcal{L}}$. For a proof of the previous identity we refer the reader to Section A.1. From this expression, due to the sign function appearing in the first integral in the right-hand side of (3.8), we remark that $(x_0, \ell_0) \mapsto \mathcal{S}_t h(x_0, \ell_0)$ is not continuous at any point $(0, \ell_0)$, $\ell_0 \geq 0$, unless $b(0, \ell_0) = 0$. Hence, a direct application of Corollary 2.1 does not work. In order to circumvent this issue, we proceed as follows. We first consider the drift-less SDE. We then briefly indicate how to proceed in the presence of a bounded measurable drift by means of the Girsanov theorem. From now on, we let $b \equiv 0$.

We will use the notation $\mathcal{S}_t h(x_0,\ell_0) = \int h(x,\ell)\widehat{\theta}_t(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell)$ with

$$\widehat{\theta}_t(x_0,\ell_0,x,\ell) := \begin{cases} -\frac{1}{2}\frac{(a(x_0,\ell_0)-a(x,\ell))}{a^2(x,\ell)} H_3(t, \frac{|x|+|x_0|+\ell-\ell_0}{\sigma(x,\ell)}), & \ell > \ell_0, \\ \frac{1}{2}(a(x_0,\ell_0)-a(x,\ell_0))\{H_2(a(x,\ell_0)t, x-x_0) - H_2(a(x,\ell_0)t, x+x_0)\}, & \ell = \ell_0. \end{cases} \tag{3.9}$$

Since the function $(x_0,\ell_0) \mapsto \mathcal{S}_t h(x_0,\ell_0)$ is continuous on $\mathbb{R}\times\mathbb{R}_+$, by applying Corollary 2.1 we obtain

$$P_T h(x_0,\ell_0) = \widehat{P}_T h(x_0,\ell_0) + \sum_{n\geq 1}\int_{\Delta_n(T)} \widehat{P}_{s_n}\mathcal{S}_{s_n-s_{n-1}}\cdots\mathcal{S}_{T-s_1}h(x_0,\ell_0)\,d\mathbf{s}_n \tag{3.10}$$

with the convention $s_0 = T$.

In order to retrieve the transition density associated to $(P_t)_{t\geq 0}$, we aim to prove an integral representation for the above series. More precisely, our aim is to prove that the right-hand side of (3.10) can be written as $\int h(x,\ell)p_T(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell)$, with an explicit representation for $p_T(x_0,\ell_0,x,\ell)$. We start by an examination of the n-th term of the series expansion.

Before proceeding, we observe that the measure $\nu(x_0,\ell_0,dx,d\ell)$ satisfies a useful convolution type property in the sense that

$$\nu\big(x_0,\ell_0,dx',d\ell'\big)\nu\big(x',\ell',dx,d\ell\big) = u\big(x_0,\ell_0,x,\ell,dx',d\ell'\big)\nu(x_0,\ell_0,dx,d\ell) \tag{3.11}$$

where the measure u is given by

$$u\big(x_0,\ell_0,x,\ell,dx',d\ell'\big) := \begin{cases} \mathbb{I}_{\{\ell_0<\ell'<\ell\}}\,dx'\,d\ell' + \mathbb{I}_{\{x'x_0>0\}}\,dx'\delta_{\ell_0}(d\ell') + \mathbb{I}_{\{xx'>0\}}\,dx'\delta_{\ell}(d\ell'), & \ell_0<\ell, \\ \mathbb{I}_{\{x_0x'>0\}}\,dx'\delta_{\ell_0}(d\ell'), & \ell_0=\ell, x_0x>0. \end{cases} \tag{3.12}$$

Applying repeatedly (3.11) and using Fubini's theorem, we obtain

$$\begin{aligned} &\int_{\Delta_n(T)} d\mathbf{s}_n\,\widehat{P}_{s_n}\mathcal{S}_{s_{n-1}-s_n}\ldots\mathcal{S}_{T-s_1}h(x_0,\ell_0) \\ &\quad= \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\Bigg\{\int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}\times\mathbb{R}_+)^n}\widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1) \\ &\qquad\times\Bigg[\prod_{i=1}^n \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1})u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i)\Bigg]\Bigg\}\nu(x_0,\ell_0,dx,d\ell) \\ &\quad= \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)p_T^n(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell) \end{aligned}$$

where we set

$$p_T^n(x_0,\ell_0,x,\ell) := \begin{cases} \int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}\times\mathbb{R}_+)^n}\widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1) \\ \quad\times[\prod_{i=1}^n \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1})u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i)] & n\geq 1, \\ \widehat{p}_T(x_0,\ell_0,x,\ell) & n=0. \end{cases} \tag{3.13}$$

We are now ready to give a representation for the density of the couple $(X_t, \ell_0 + L_t^0(X))$. As already mentioned in the introduction, we point out that the proof of the convergence of the asymptotic expansion for the transition density is not standard in the current setting. To see this, we note that using (3.9), the standard Gaussian estimate (1.3) and the space-time inequality in (1.2), we obtain the following estimate (for detailed computations, we refer the reader to (A.3) and the estimates just below it)

$$\left|\widehat{\theta}_t(x_0,\ell_0,x,\ell)\right| \le \begin{cases} \frac{C}{t^{\frac{3-\eta}{2}}} H_0(ct,|x|+|x_0|+\ell-\ell_0), & \ell>\ell_0, \\ \frac{C}{t^{1-\frac{\eta}{2}}} H_0(ct,x-x_0), & \ell=\ell_0, \end{cases} \tag{3.14}$$

for some constants $C := C(a)$, $c := c(a) > 1$. Indeed, in the classical diffusion setting, the parametrix expansion of the transition density converges since the order of the singularity in time induced by the parametrix kernel $\widehat{\theta}_t$ is of order $t^{-1+\frac{\eta}{2}}$, which is still integrable near 0. The situation here appears to be more delicate. At first glance, the order of the time singularity in $\widehat{\theta}_t$ consists in two parts. The first part corresponds to the non-singular part of the law of $(\bar{X}_t, \ell_0 + L_t^0(\bar{X}))$. From (3.14), it induces a singularity in time of order $t^{-\frac{3}{2}+\frac{\eta}{2}}$ which is integrable in time after integrating the corresponding kernel $H_0(ct, |x|+|x_0|+\ell-\ell_0)$ over the domain $\mathbb{R}\times(\ell_0,\infty)$. The second part corresponds to the singular part of the law of $(\bar{X}_t, \ell_0 + L_t^0(\bar{X}))$, that is, the one of the proxy process killed when it reaches zero, and is absolutely continuous with respect to the singular measure $dx\delta_{\ell_0}(d\ell)$. Here the situation is standard and the singularity in time appearing in (3.14), namely $t^{-1+\frac{\eta}{2}}$, is integrable near 0.

The main difficulty appears when one wants to control the whole space-time convolution appearing in the right-hand side of (3.13). More precisely, it lies in the cross-terms which are of a different nature, for instance when one considers the convolution of the non-singular part in the kernel $\widehat{\theta}_{T-s_1}$ and the singular part in the kernel $\widehat{\theta}_{s_1-s_2}$. Standard arguments such as those employed in [8] or [14] do not guarantee the convergence of the integral defining (3.13). To overcome this difficulty and to show that the parametrix expansion for the transition density does indeed converge, one has to make use of the key estimate obtained in Lemma A.4.

As our second main result, we prove that the transition density of $(X_t, \ell_0 + L_t^0(X))_{t\ge0}$ exists and satisfies a Gaussian upper bound. We postpone its proof to the Appendix.

Theorem 3.4. *Assume that* (R-η) *and* (UE) *hold for some* $\eta\in(0,1]$. *For* $T>0$ *and* $(x_0,\ell_0)\in\mathbb{R}\times\mathbb{R}_+$ *define the measure*

$$\begin{aligned} p_T(x_0,\ell_0,dx,d\ell) &:= p_T(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell) \\ &= f_T(x_0,\ell_0,x)\mathbb{I}_{\{xx_0\ge0\}}\,dx\delta_{\ell_0}(d\ell) + q_T(x_0,\ell_0,x,\ell)\mathbb{I}_{\{\ell_0<\ell\}}\,dx\,d\ell \end{aligned}$$

with $p_T(x_0,\ell_0,x,\ell) := \sum_{n\ge0} p_T^n(x_0,\ell_0,x,\ell)$ *and*

$$f_T(x_0,\ell_0,x) := \sum_{n\ge0} p_T^n(x_0,\ell_0,x,\ell_0), \qquad q_T(x_0,\ell_0,x,\ell) := \sum_{n\ge0} p_T^n(x_0,\ell_0,x,\ell).$$

Then, both series defining $f_T(x_0,\ell_0,x)$ *and* $q_T(x_0,\ell_0,x,\ell)$ *converge absolutely and uniformly for* $(x_0,\ell_0),(x,\ell)\in\mathbb{R}\times\mathbb{R}_+$. *Moreover, for* $h\in\mathcal{C}_b(\mathbb{R}\times\mathbb{R}_+)$ *the following representation for the semigroup holds*

$$P_Th(x_0,\ell_0) = \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)p_T(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell).$$

Therefore, for all $(x_0,\ell_0)\in\mathbb{R}\times\mathbb{R}_+$, *the function* $(x,\ell)\mapsto p_T(x_0,\ell_0,x,\ell)$ *is the probability density function of the couple* $(X_T, \ell_0 + L_T^0(X))$ *with respect to the* σ*-finite measure* $\nu(x_0,\ell_0,dx,d\ell)$, *where* X_T *is the solution taken at time* T *of the SDE* (3.1) *starting from* x_0 *at time* 0, $L_T^0(X)$ *being its running local time at time* T.

Finally, there exist some constants $C := C(T,a)$, $c := c(a) > 1$ *such that for all* $(x_0,\ell_0),(x,\ell)\in\mathbb{R}\times\mathbb{R}_+$ *and for all* $\beta\in[0,1]$ *the following Gaussian upper-bounds hold*:

$$f_T(x_0,\ell_0,x) \le C\left\{\frac{|x|^\beta}{T^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{T^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cT,x-x_0) \quad \text{and}$$

$$q_T(x_0,\ell_0,x,\ell) \le C\frac{1}{T^{1/2}}H_0\big(cT,|x|+|x_0|+\ell-\ell_0\big).$$

Remark 3.5. The above result establishes the existence of a transition density for the Markov process $(X_t, \ell_0 + L_t^0(X))_{t\ge0}$ where the dynamics is given by (3.1) without drift. In order to add a drift, one can use the Girsanov transform as follows. Let $\mathbb{R}\times\mathbb{R}_+\ni(x,\ell)\mapsto b(x,\ell)$ be a real-valued bounded measurable function. We consider the unique weak solution $\{(X,W),(\Omega,\mathcal{F},\mathbb{P}),(\mathcal{F}_t)_{t\ge0}\}$ of (3.1) and let $\widetilde{W}_t := W_t + \int_0^t \widetilde{b}(X_s, L_s^0(X))\,ds$, with $\widetilde{b}(x,\ell) := b(x,\ell)/\sigma(x,\ell)$.

Then, defining the new probability measure on $\mathcal{F}_T$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \exp\left\{-\int_0^T \widetilde{b}\big(X_s, L_s^0(X)\big)\, dW_s - \frac{1}{2}\int_0^T \widetilde{b}^2\big(X_s, L_s^0(X)\big)\, ds\right\},$$

from Girsanov's theorem, we know that $\{(\widetilde{X}, \widetilde{W}), (\Omega, \mathcal{F}, \mathbb{Q}), (\mathcal{F}_t)_{t\geq 0}\}$, with $\widetilde{X}_t = x_0 + \int_0^t \sigma(\widetilde{X}_s, A_s(\widetilde{X}))\, d\widetilde{W}_s$ is a weak solution to (3.1) with $b \equiv 0$. We also know from Theorem 3.3 that weak uniqueness holds for this equation and from Theorem 3.4 that the couple $(\widetilde{X}_T, A_T(\widetilde{X}))$ admits a density with respect to the measure $\nu(x_0, \ell_0, dx, d\ell)$. Therefore, for any bounded measurable function $h : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$, one has

$$\begin{aligned}
P_T h(x_0, \ell_0) &:= \mathbb{E}_{\mathbb{P}}\big[h\big(X_T, \ell_0 + L_T^0(X)\big)\big] \\
&= \int_{\mathbb{R}\times\mathbb{R}_+} h(x, \ell)\mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}} \,\Big|\, \big(X_T, \ell_0 + L_T^0(X)\big) = (x, \ell)\right]\widetilde{p}_T(x_0, \ell_0, x, \ell)\nu(x_0, \ell_0, dx, d\ell) \\
&= \int_{\mathbb{R}\times\mathbb{R}_+} h(x, \ell)\mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}} \,\Big|\, \big(X_T, \ell_0 + L_T^0(X)\big) = (x, \ell)\right]\widetilde{p}_T(x_0, \ell_0, x, \ell)\mathbb{I}_{\{\ell_0<\ell\}}\, dx\, d\ell \\
&\quad + \int_{\mathbb{R}} h(x, \ell_0)\mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}} \,\Big|\, \big(X_T, \ell_0 + L_T^0(X)\big) = (x, \ell_0)\right]\widetilde{p}_T(x_0, \ell_0, x, \ell)\mathbb{I}_{\{x_0 x\geq 0\}}\, dx. \qquad (3.15)
\end{aligned}$$

We thus deduce that $(X_T, \ell_0 + L_T^0(X))$ admits a density with respect to the measure $\nu(x_0, \ell_0, dx, d\ell)$. Then, one may extend Theorem 2.2 to the class of functions h that are bounded measurable with respect to the space variable x and continuous with respect to the variable ℓ as follows. Indeed, if h is such a function, then from Theorem 174 p.111 of Kestelman [13], there exists a sequence $(h_N)_{N\geq 1}$ of continuous functions on $\mathbb{R} \times \mathbb{R}_+$ such that $\sup_{N\geq 1} |h_N|_\infty \leq |h|_\infty$ and $\lim_{N\uparrow\infty} h_N(\cdot, \ell) = h(\cdot, \ell)$ a.e. and for all $\ell \in \mathbb{R}_+$.

Next, by applying Theorem 2.2 to h_N we obtain

$$P_T h_N(x_0, \ell_0) = \widehat{P}_T h_N(x_0, \ell_0) + \int_0^T P_s \mathcal{S}_{T-s} h_N\, ds.$$

We then pass to the limit in the previous equality. From equations (3.15), (3.8), the identity $\widehat{P}_T h_N(x_0, \ell_0) = \int_{\mathbb{R}\times\mathbb{R}_+} h_N(x, \ell)\widehat{p}_T(x_0, \ell_0, dx, d\ell)$ and the dominated convergence theorem, we obtain that $\lim_N P_T h_N(x_0, \ell_0) = P_T h(x_0, \ell_0)$, $\lim_N \widehat{P}_T h_N(x_0, \ell_0) = \widehat{P}_T h(x_0, \ell_0)$ and $\lim_N \mathcal{S}_{T-s} h_N(x_0, \ell_0) = \mathcal{S}_{T-s} h(x_0, \ell_0)$ for every $(x_0, \ell_0) \in \mathbb{R} \times \mathbb{R}_+$. Hence, using again the dominated convergence theorem with (2.4), we deduce that Theorem 2.2 is valid for h. Since $\mathcal{S}_{T-s} h$ given by (3.8) is bounded measurable with respect to x and continuous with respect to ℓ, one may iterate the first step formula of Theorem 2.2 and obtain the semigroup expansion (2.11) where the function $\mathcal{S}_t h$ is given by (3.8) and $\widehat{\theta}_t$ is defined accordingly. Finally, in order to conclude, one has to repeat the arguments employed in the proof of Theorem 3.4. We omit the remaining technical details.

Remark 3.6. We again point out that the representation in infinite series obtained previously are of great interest. For instance, one may be interested in studying the regularity properties of the density functions $(t, x_0, \ell_0, x, \ell) \mapsto f_t(x_0, \ell_0, x, \ell), q_t(x_0, \ell_0, x, \ell)$, to derive a probabilistic interpretation for $P_t h$ and $p_t(x_0, \ell_0, x, \ell)$ and to obtain some integration by parts formulas, from which stem an unbiased Monte Carlo simulation scheme. We refer e.g. to [9,10] and the references therein for some results in that direction concerning to the first hitting times of one-dimensional elliptic diffusions.

3.4. *A diffusion process with coefficients depending on its running maximum*

We now turn our attention to the following SDE with dynamics

$$X_t = x_0 + \int_0^t b(X_s, M_s)\, ds + \int_0^t \sigma(X_s, M_s)\, dW_s \qquad (3.16)$$

where $M_t = A_t(X) := m_0 \vee \max_{0\leq s\leq t} X_s$, $m_0 \geq x_0$, is the running maximum of the process X at time t. The state space of the process $(X_t, M_t)_{t\geq 0}$ is denoted by the closed set $\mathcal{J} = \{(x, m) \in \mathbb{R}^2 : x \leq m\}$. We define accordingly the collection of linear maps $P_t h(x_0, m_0) = \mathbb{E}[h(X_t, M_t)]$ for $h \in \mathcal{B}_b(\mathcal{J})$ and introduce the following assumptions:

(R-η) The coefficients b and $a = \sigma^2$ are bounded measurable functions defined on $\mathcal{J}$. The diffusion coefficient a is η-Hölder continuous on $\mathcal{J}$.

(UE) There exists some constant $\underline{a} > 0$ such that $\forall (x, m) \in \mathcal{J}$, $\underline{a} \le a(x, m)$.

In the following, just as mentioned in Remark 3.1, in order to ensure that (H1) (iii) is satisfied, we will consider, whenever needed, the bounded, continuous and uniformly elliptic extension of the diffusion coefficient a from the closed subset $\mathcal{J}$ to $\mathbb{R}^2$.

Although the lines of reasoning used in the proof of Theorem 3.7 and Theorem 3.8 are rather similar to those employed in the case of the SDE with its running local time, we decided to include this example in order to illustrate the generality of our framework. Indeed, unlike the local time and also the examples considered so far in the literature by means of the parametrix technique, the running maximum is not a continuous additive functional. This difference is reflected in the definition of the collection of linear maps $(P_t)_{t\ge 0}$, and also in the definition of the approximation process $\bar{Y}$ and the maps $(\bar{P}_t)_{t\ge 0}$, see below in Section 3.5. This allows to define a Markov semigroup for $\bar{Y}$ and later on for Y.

The weak existence of a solution to the SDE (3.16) follows from standard results, see e.g. Chapter 6 in [17] or Theorem 5.4.22 and Remark 5.4.23 in [12], see also Forde [7] for another approach under the assumption that a and b are bounded continuous functions.

In the same spirit as the previous section, we can characterise solutions of the SDE (3.16) in terms of the associated (local) martingale problem. We let $\mathcal{D}$ be the class of functions $h : \mathbb{R}_+ \times \mathcal{J} \to \mathbb{R}$ such that $h \in \mathcal{C}_b^{1,2,1}(\mathbb{R}_+ \times \mathcal{J})$ which satisfies the condition $\partial_3 h(t, m, m) = 0$, $(t, m) \in \mathbb{R}_+ \times \mathbb{R}$. For $h \in \mathcal{D}$, we define the operator

$$\mathcal{L}h(t, x, m) = \frac{1}{2} a(x, m)\partial_2^2 h(t, x, m) + b(x, m)\partial_2 h(t, x, m).$$

By observing that the process $t \mapsto M_t$ increases only on the set $\{t : X_t = M_t\}$ and by applying Itô's lemma we obtain

$$h(t, X_t, M_t) = h(0, x_0, m_0) + \int_0^t (\partial_1 + \mathcal{L})h(s, X_s, M_s)\, ds + \int_0^t \sigma(X_s, M_s)\partial_2 h(s, X_s, M_s)\, dW_s \tag{3.17}$$

for $h \in \mathcal{D}$. If $\mathbb{P}$ be a solution to the martingale problem associated to the operator $\mathcal{L}$ with the class of functions $\mathcal{D}$ then the measure $\mathbb{P}_{t,w} = \mathbb{Q}_{t,w} \circ \theta_t^{-1}$, where $\mathbb{Q}_{t,w}$ is a regular conditional probability for $\mathbb{P}$ given $\mathcal{F}_t$, solves the martingale problem for every $w \notin N$, $N \in \mathcal{F}_t$ being a $\mathbb{P}$-null event so that (H0) is satisfied.

3.5. *Weak uniqueness and representation of the transition density*

We now introduce the proxy process $\bar{X}_t := x_0 + \sigma(z_1)W_t$, $t \ge 0$, obtained from the original process X by removing the drift part and by freezing the diffusion coefficient at $z_1 = (x_1, m_1) \in \mathbb{R}^2$ in (3.16). As already done for the original process Y, for $h \in \mathcal{C}_b(\mathcal{J})$, we define accordingly

$$\bar{P}_t h(x_0, m_0) = \bar{P}_t^{z_1} h(x_0, m_0) = \mathbb{E}\big[h(\bar{X}_t, \bar{M}_t)\big] = \mathbb{E}\Big[h\Big(x_0 + \sigma(z_1)W_t, m_0 \vee \max_{0\le s\le t}\big(x_0 + \sigma(z_1)W_s\big)\Big)\Big],$$

and, with $\bar{\sigma} = \sigma(z_1)$, $\bar{a} = \bar{\sigma}^2$, for $h \in \mathcal{D}$, the operator

$$\bar{\mathcal{L}}h(t, x, m) = \frac{1}{2}\bar{a}\partial_2^2 h(t, x, m), \quad (t, x, m) \in \mathbb{R}_+ \times \mathcal{J}.$$

We now compute the law of the couple $(\bar{X}_t, \bar{M}_t)$. We denote by $T_{m_0} = \inf\{t \ge 0 : x_0 + \bar{\sigma} W_t = m_0\}$ the first hitting time of 0 by the process $(\bar{X}_t)_{t\ge 0}$. We decompose $\bar{P}_t h(x_0, m_0)$ as follows

$$\begin{aligned}\bar{P}_t h(x_0, m_0) &:= \mathbb{E}\big[h(x_0 + \bar{\sigma} W_t, m_0)\mathbb{I}_{\{T_{m_0}\ge t\}}\big] + \mathbb{E}\Big[h\Big(x_0 + \bar{\sigma} W_t, m_0 \vee \max_{0\le s\le t}(x_0 + \bar{\sigma} W_s)\Big)\mathbb{I}_{\{T_{m_0}<t\}}\Big] \\ &=: \mathrm{I} + \mathrm{II}.\end{aligned}$$

From the reflection principle of Brownian motion, see e.g. [12], one derives the law of the proxy process killed when it exits $(-\infty, m_0)$, namely

$$\begin{aligned}\mathrm{I} &:= \mathbb{E}\big[h(x_0 + \bar{\sigma} W_t, m_0)\mathbb{I}_{\{T_{m_0}\ge t\}}\big] \\ &= \int_{\mathbb{R}^2} h(x, m)\big\{H_0(\bar{a}t, x - x_0) - H_0(\bar{a}t, 2m_0 - x - x_0)\big\}\mathbb{I}_{\{x\vee x_0\le m_0\}}\, dx\,\delta_{m_0}(dm).\end{aligned} \tag{3.18}$$

The bivariate density of the couple $(W_t, \max_{0\le s\le t} W_s)$, see e.g. Proposition 2.8.1 in [12], gives

$$\begin{aligned}
\mathrm{II} &:= \mathbb{E}\Big[h\Big(x_0+\bar\sigma W_t, m_0 \vee \max_{0\le s\le t}(x_0+\bar\sigma W_s)\Big)\mathbb{I}_{\{T_{m_0}<t\}}\Big] \\
&= \mathbb{E}\Big[h\Big(x_0+\bar\sigma W_t, \max_{0\le s\le t}(x_0+\bar\sigma W_s)\Big)\mathbb{I}_{\{\max_{0\le s\le t}(x_0+\bar\sigma W_s)\ge m_0\}}\Big] \\
&= \int_{\mathbb{R}^2} h(x,m)(-2H_1)(\bar a t, 2m-x-x_0)\mathbb{I}_{\{x\vee x_0\vee m_0\le m\}}\,dx\,dm. \qquad (3.19)
\end{aligned}$$

Combining I and II, we see that the couple $(\bar X_t, \bar M_t)$ admits a density $(x,m)\mapsto \bar p_t(x_0,m_0,x,m)$, for $t>0$, given by

$$\begin{aligned}
&\bar p_t(x_0,m_0,dx,dm) = \bar p_t(x_0,m_0,x,m)\nu(x_0,m_0,dx,dm)\\
&\bar p_t(x_0,m_0,x,m) := \bar f_t(x_0,x)\mathbb{I}_{\{m=m_0\}} + \bar q_t(x_0,m_0,x,m)\mathbb{I}_{\{m_0<m\}}
\end{aligned}$$

with

$$\begin{aligned}
&\bar f_t(x_0,x) := H_0(\bar a t, x-x_0) - H_0(\bar a t, 2m_0-x-x_0),\\
&\bar q_t(x_0,m_0,x,m) := -2H_1(\bar a t, 2m-x-x_0),\\
&\nu(x_0,m_0,dx,dm) := \mathbb{I}_{\{x\le m\}}\mathbb{I}_{\{m_0<m\}}\,dx\,dm + \mathbb{I}_{\{x\le m_0\}}\,dx\delta_{m_0}(dm).
\end{aligned}$$

Moreover, as already mentioned in assumption (H1) in Section 2.1, we let

$$\begin{aligned}
&\widehat p_t(x_0,m_0,dx,dm) = \widehat p_t(x_0,m_0,x,m)\nu(x_0,m_0,dx,dm)\\
&\widehat p_t(x_0,m_0,x,m) = \widehat f_t(x_0,x)\mathbb{I}_{\{m=m_0\}} + \widehat q_t(x_0,\ell_0,x,\ell)\mathbb{I}_{\{m_0<m\}}
\end{aligned}$$

with

$$\begin{aligned}
&\widehat f_t(x_0,x) := H_0\big(a(x,m_0)t, x-x_0\big) - H_0\big(a(x,m_0)t, 2m_0-x-x_0\big),\\
&\widehat q_t(x_0,m_0,x,m) := -2H_1\big(a(x,m)t, 2m-x-x_0\big).
\end{aligned}$$

We also observe that $(\bar P_t)_{t\ge 0}$ satisfies the Markov semigroup property. The first main result of this section establishes weak uniqueness for the SDE (3.16) by proving that assumptions (H1) and (H2) of Section 2.1 are satisfied. We postpone the proof to the Appendix.

Theorem 3.7. *For* $\eta\in(0,1]$, *under* (R-η) *and* (UE), *weak uniqueness holds for the SDE* (3.16).

We now show that, given $(x_0,m_0)\in\mathcal{J}$, the law of (X_T,M_T) is absolutely continuous with respect to the measure $\nu(x_0,m_0,dx,dm)$. Our strategy is similar to the one employed in the case of the SDE with its running local time, namely, we establish a representation in infinite series of $P_t h$ from which stems an explicit representation of the density of the couple (X_T,M_T), see Theorem 3.8 below. We first recall

$$\begin{aligned}
\mathcal{S}_t h(x_0,m_0) = &\int h(x,m)\Big\{\frac{1}{2}\big(a(x_0,m_0)-a(x,m)\big)(-2H_3)\big(a(x,m)t, 2m-x-x_0\big)\\
&+ b(x_0,m_0)(-2H_2)\big(a(x,m)t, 2m-x-x_0\big)\Big\}\mathbb{I}_{\{x\le m\}}\mathbb{I}_{\{m_0<m\}}\,dx\,dm\\
&+\int h(x,m_0)\Big\{\frac{1}{2}\big(a(x_0,m_0)-a(x,m_0)\big)\big(H_2\big(a(x,m_0)t, x-x_0\big)\\
&- H_2\big(a(x,m_0)t, 2m_0-x-x_0\big)\big)\\
&+ b(x_0,m_0)\big(H_1\big(a(x,m_0)t, 2m_0-x-x_0\big) - H_1\big(a(x,m_0)t, x-x_0\big)\big)\Big\}\mathbb{I}_{\{x<m_0\}}\,dx\\
=&\int h(x,m)\widehat\theta_t(x_0,m_0,x,m)\nu(x_0,m_0,dx,dm) \qquad (3.20)
\end{aligned}$$

with $\widehat{\theta}_t$ given by (we refer the reader to Section A.3 for a proof of the above identity)

$$\widehat{\theta}_t(x_0,m_0,x,m)$$
$$:= \begin{cases} \frac{1}{2}(a(x_0,m_0)-a(x,m))(-2H_3)(a(x,m)t,2m-x-x_0) & \\ \quad + b(x_0,m_0)(-2H_2)(a(x,m)t,2m-x-x_0), & x \le m, m_0 < m, \\ \frac{1}{2}(a(x_0,m_0)-a(x,m_0))(H_2(a(x,m_0)t,x-x_0)-H_2(a(x,m_0)t,2m_0-x-x_0)) & \\ \quad + b(x_0,m_0)(H_1(a(x,m_0)t,2m_0-x-x_0)-H_1(a(x,m_0)t,x-x_0)), & x < m_0, m = m_0. \end{cases}$$

We remark here that the map $\mathcal{J} \ni (x_0,m_0) \mapsto \mathcal{S}_t h(x_0,m_0)$ is continuous. This is different from the case of the local time, where due to the presence of the sign function, we required that $b=0$ in order to ensure the continuity of the map $\mathcal{S}_t h$. One may thus iterate the first step of the expansion obtained in Theorem 2.2. More precisely, applying Corollary 2.1 we obtain

$$P_T h(x_0,m_0) = \widehat{P}_T h(x_0,m_0) + \sum_{n\ge 1} \int_{\Delta_n(T)} \widehat{P}_{s_n} \mathcal{S}_{s_n - s_{n-1}} \cdots \mathcal{S}_{T-s_1} h(x_0,m_0)\, d\mathbf{s}_n \tag{3.21}$$

with the convention $s_0 = T$. We again observe the following convolution type property of the singular measure, namely

$$\nu\big(x_0,m_0,dx',dm'\big)\nu\big(x',m',dx,dm\big) = u\big(x_0,m_0,x,m,dx',dm'\big)\nu(x_0,m_0,dx,dm) \tag{3.22}$$

where we set

$$u\big(x_0,m_0,x,m,dx',dm'\big)$$
$$:= \begin{cases} \mathbb{I}_{\{x'\vee x_0 \vee m_0 < m'\}} \mathbb{I}_{\{m' < m\}}\, dx'\, dm' + \mathbb{I}_{\{x' < m_0\}}\, dx' \delta_{m_0}(dm') & \\ \quad + \mathbb{I}_{\{x' < m\}}\, dx' \delta_m(dm') & m_0 < m, x \le m, \\ \mathbb{I}_{\{x' < m_0\}}\, dx' \delta_{m_0}(dm') & m = m_0, x \le m_0. \end{cases} \tag{3.23}$$

We then examine the n-th term of the series expansion. Using Fubini's theorem and recursively applying (3.22), it can be expressed as

$$\begin{aligned} &\int_{\Delta_n(T)} d\mathbf{s}_n \widehat{P}_{s_n} \mathcal{S}_{s_{n-1}-s_n} \cdots \mathcal{S}_{T-s_1} h(x_0,m_0) \\ &\quad = \int_{\mathbb{R}^2} h(x,m) \Bigg\{ \int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}^2)^n} \widehat{p}_{s_n}(x_0,m_0,x_1,m_1) \\ &\qquad \times \Bigg[\prod_{i=1}^{n} \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,m_i,x_{i+1},m_{i+1}) u(x_0,m_0,x_{i+1},m_{i+1},dx_i,dm_i) \Bigg] \Bigg\} \nu(x_0,m_0,dx,dm) \\ &\quad = \int_{\mathbb{R}^2} h(x,m) p_T^n(x_0,m_0,x,m) \nu(x_0,m_0,dx,dm) \end{aligned}$$

where we set

$$p_T^n(x_0,m_0,x,m)$$
$$:= \begin{cases} \int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}^2)^n} \widehat{p}_{s_n}(x_0,m_0,x_1,m_1) & \\ \times \big[\prod_{i=1}^n \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,m_i,x_{i+1},m_{i+1}) u(x_0,m_0,x_i,m_i,dx_{i+1},dm_{i+1})\big] & n \ge 1, \\ \widehat{p}_T(x_0,m_0,x,m) & n = 0. \end{cases} \tag{3.24}$$

We are now ready to give a representation for the density of the couple $(X_T, M_T(X))$. As already mentioned in the case of the SDE with its running local time, the proof of the convergence of the asymptotic expansion for the transition density is not standard in the current setting. To overcome the main difficulty which again comes from the different nature of the two kernels appearing in $\widehat{\theta}_t$ one has to make use of the key estimate obtained in Lemma A.4. This allows to obtain the convergence of the parametrix expansion for the transition density. Similar Gaussian upper-bounds for this density are also established. The proof is postponed to the Appendix.

Theorem 3.8. *Assume that* (R-η) *and* (UE) *hold for some* $\eta \in (0,1]$. *For* $T > 0$ *and* $(x_0, m_0) \in \mathcal{J}$, *define the measure*

$$\begin{aligned} p_T(x_0, m_0, dx, dm) &:= p_T(x_0, m_0, x, m)\nu(x_0, m_0, dx, dm) \\ &= f_T(x_0, m_0, x)\mathbb{I}_{\{x \le m_0\}}\, dx\, \delta_{m_0}(dm) + q_T(x_0, m_0, x, m)\mathbb{I}_{\{x \le m\}}\mathbb{I}_{\{m_0 < m\}}\, dx\, dm \end{aligned}$$

with $p_T(x_0, m_0, x, m) := \sum_{n \ge 0} p_T^n(x_0, m_0, x, m)$ *and*

$$f_T(x_0, m_0, x) := \sum_{n \ge 0} p_T^n(x_0, m_0, x, m_0), \qquad q_T(x_0, m_0, x, m) := \sum_{n \ge 0} p_T^n(x_0, m_0, x, m).$$

Then, both series defining $f_T(x_0, m_0, x, m)$ *and* $q_T(x_0, m_0, x)$ *converge absolutely and uniformly for* (x_0, m_0), $(x, m) \in \mathcal{J}$. *Moreover for* $h \in \mathcal{C}_b(\mathcal{J})$, *the following representation for the semigroup holds*,

$$P_T h(x_0, m_0) = \int_{\mathbb{R} \times \mathbb{R}_+} h(x, m) p_T(x_0, m_0, dx, dm).$$

Therefore, for all $(x_0, m_0) \in \mathcal{J}$, *the function* $\mathcal{J} \ni (x, m) \mapsto p_T(x_0, m_0, x, m)$ *is the probability density function with respect to the measure* $\nu(x_0, m_0, dx, dm)$ *of the random vector* $(X_T, M_T(X))$, *where* X_T *is the solution taken at time* T *of the SDE* (3.16) *and* $M_T(X) = m_0 \vee \max_{0 \le s \le T} X_s$.

Finally, for some positive $C := C(a, b, T)$, $c := c(a) > 1$, *for all* $(x_0, m_0), (x, m) \in \mathcal{J}$ *and for all* $\beta \in [0,1]$ *the following Gaussian upper-bounds hold*:

$$\begin{aligned} f_T(x_0, m_0, x) &\le C\left\{\frac{|m_0 - x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|m_0 - x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1\right\} H_0(cT, x - x_0), \\ q_T(x_0, m_0, x, m) &\le \frac{C}{T^{\frac{1}{2}}} H_0(cT, 2m - x - x_0). \end{aligned}$$

4. Conclusion

In this paper, we established weak existence and uniqueness for some SDEs with coefficients depending on some path-dependent functionals $(A_t(X))_{t \ge 0}$ under mild assumptions on the coefficients, namely bounded measurable drift and uniformly elliptic Hölder-continuous diffusion coefficient. We illustrated our approach on two examples: an SDE with coefficients depending on its running local time and an SDE with coefficients depending on its running maximum. We established the existence as well as a representation in infinite series of the density for the couple $(X_t, A_t(X))$ in both examples. Some Gaussian upper-bounds are also obtained.

Obviously, a wide variety of Brownian functionals can be investigated. Simple extensions include for instance the case $A_t(X) = (\tau_L \wedge t, X_{\tau_L \wedge t})$, where $\tau_L = \inf\{t \ge 0 : X_t \ge L\}$ is the first hitting time of the barrier L by X or the bivariate functional $A_t(W) = (\min_{0 \le s \le t} X_s, \max_{0 \le s \le t} X_s)$. More challenging extensions could include other type of processes. One notably may consider the case of a skew diffusion with path-dependent coefficients involving its local and occupation times, see Appuhamillage et al. [1] for an expression of the trivariate density $(B_t^{(\alpha)}, L_t^0(B^{(\alpha)}), \Gamma_t^0(B^{(\alpha)}))$, $t \ge 0$, where $(B_t^{(\alpha)})_{t \ge 0}$ is an α-skew Brownian motion or reflected SDEs. This will be developed in future works.

Appendix

A.1. *Proof of Theorem* 3.3

From the Markov property of the Brownian motion W, one can deduce that $(\bar{X}_t, \ell_0 + L_t^0(\bar{X}))_{t \ge 0}$ is a Markov process so that

$$\mathbb{E}\big[f\big(\bar{X}_s + \bar{\sigma}(W_{t+s} - W_s), \ell_0 + L_s^0(\bar{X}) + L_{t+s}^0(\bar{X}) - L_s^0(\bar{X})\big)|\mathcal{F}_s\big] = \bar{P}_t h\big(\bar{X}_s, \ell_0 + L_s^0(\bar{X})\big).$$

From the above expression of $\bar{P}_t h$, the dominated convergence theorem and the heat equation $\partial_t f(t,x) = \frac{1}{2}\bar{a}\partial_x^2 f(t, x-)$ for $f(t,x) = H_0(\bar{a}t, x)$ and $f(t,x) = (-H_1)(t, |x|/\bar{\sigma})$, we obtain for any $t > 0$ and $h \in \mathcal{C}_b(\mathbb{R}^d)$,

$$\partial_t \bar{P}_t h(x_0, \ell_0) = \frac{\bar{a}}{2}\bigg[\int_{\mathbb{R}} h(x, \ell_0)\big\{H_2(\bar{a}t, x - x_0) - H_2(\bar{a}t, x + x_0)\big\}\mathbb{I}_{\{xx_0 \ge 0\}}\, dx$$

$$
\begin{aligned}
&\quad + \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{1}{\bar a}(-H_3)\left(t, \frac{|x|+|x_0|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell_0\le\ell\}}\,dx\,d\ell\Bigg] \\
&= \bar{\mathcal{L}}\bar P_t h(x_0,\ell_0).
\end{aligned}
$$

Moreover, if $h\in\mathcal{D}$, from Proposition 3.1 one deduce that $\partial_t \bar P_t h = \bar P_t \bar{\mathcal{L}} h$, so that assumption (H1) (i) is satisfied. We now prove that (H1) (ii) holds. Since $(t,x_0,\ell_0)\mapsto \bar P_{t+r}h(x_0,\ell_0)\in\mathcal{C}_b^{1,2,1}(\mathbb{R}_+\times\mathbb{R}\backslash\{0\}\times\mathbb{R}_+)\cap\mathcal{C}_b(\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+)$, it suffices to prove that it satisfies the transmission condition (3.2). From (3.7), for $h\in\mathcal{C}_b^{0,1}(\mathbb{R}\times\mathbb{R}_+)$, we obtain

$$
\begin{aligned}
\partial_1\bar P_{t+r}h(0+,\ell_0) &= \int_{\mathbb{R}_+} h(x,\ell_0)(-2H_1)\big(\bar a(t+r),x\big)\,dx \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{1}{\bar a\bar\sigma}(-H_2)\left((t+r),\frac{|x|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell_0\le\ell\}}\,dx\,d\ell, \\
\partial_1\bar P_{t+r}h(0-,\ell_0) &= \int_{\mathbb{R}_-} h(x,\ell_0)(-2H_1)\big(\bar a(t+r),x\big)\,dx \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{1}{\bar a\bar\sigma}H_2\left((t+r),\frac{|x|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell_0\le\ell\}}\,dx\,d\ell, \\
\partial_2\bar P_{t+r}h(0,\ell_0) &= \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{1}{\bar a\bar\sigma}H_2\left((t+r),\frac{|x|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell\ge\ell_0\}}\,dx\,d\ell \\
&\quad + \int_{\mathbb{R}} h(x,\ell_0)\frac{1}{\bar a}H_1\left((t+r),\frac{|x|}{\bar\sigma}\right)dx.
\end{aligned}
$$

Combining the above computations, we obtain the transmission condition

$$
\begin{aligned}
&\frac{1}{2}\big(\partial_1\bar P_{t+r}h(0+,\ell_0)-\partial_1\bar P_{t+r}h(0-,\ell_0)\big) \\
&\quad = \int h(x,\ell_0)(-H_1)\big(\bar a(t+r),|x|\big)\,dx - \int h(x,\ell)\frac{1}{\bar a\bar\sigma}H_2\left((t+r),\frac{|x|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell\ge\ell_0\}}\,dx\,d\ell \\
&\quad = -\partial_2\bar P_{t+r}h(0,\ell_0).
\end{aligned}
$$

Now, in order to obtain $\bar\theta_t^{z_1}(x_0,\ell_0,x,\ell)$, we first integrate the kernel $\bar p_t(x_0,\ell_0,dx,d\ell)$ given by (3.7) against the test function h and apply the difference of the generators given by (3.3) and (3.5). We obtain

$$
\forall t>0,\quad (\mathcal{L}-\bar{\mathcal{L}})\bar P_t h(x_0,\ell_0) = \frac{1}{2}\big(a(x_0,\ell_0)-\bar a\big)\partial^2_{x_0}\bar P_t f(x_0-,\ell_0) + b(x_0,\ell_0)\partial_{x_0}\bar P_t h(x_0-,\ell_0)
$$

with

$$
\begin{aligned}
\partial_1\bar P_t h(x_0-,\ell_0) &= -\int_{\mathbb{R}} h(x,\ell_0)\big\{H_1(\bar a t, x-x_0)-H_1(\bar a t,x+x_0)\big\}\mathbb{I}_{\{xx_0\ge 0\}}\,dx \\
&\quad - \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{\operatorname{sign}(x_0)}{\bar a\bar\sigma}H_2\left(t,\frac{|x|+|x_0|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell\ge\ell_0\}}\,dx\,d\ell
\end{aligned}
\tag{A.1}
$$

and

$$
\begin{aligned}
\partial_1^2\bar P_t h(x_0-,\ell_0) &= \int_{\mathbb{R}} h(x,\ell_0)\big\{H_2(\bar a t,x-x_0)-H_2(\bar a t,x+x_0)\big\}\mathbb{I}_{\{xx_0\ge 0\}}\,dx \\
&\quad - \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\frac{1}{\bar a^2}H_3\left(t,\frac{|x|+|x_0|+\ell-\ell_0}{\bar\sigma}\right)\mathbb{I}_{\{\ell\ge\ell_0\}}\,dx\,d\ell
\end{aligned}
\tag{A.2}
$$

where $\mathrm{sign}(x_0) := -\mathbb{I}_{\{x_0\le 0\}} + \mathbb{I}_{\{x_0\ge 0\}}$ (since we are taking the left-derivative). Hence, we obtain the fact that $(\mathcal{L} - \bar{\mathcal{L}})\bar{P}_t h(x_0,\ell_0) := \int h(x,\ell)\bar{\theta}_t^{z_1}(x_0,\ell_0,x,\ell)\nu(x_0,\ell_0,dx,d\ell)$ with

$$\bar{\theta}_t^{z_1}(x_0,\ell_0,x,\ell) := \begin{cases} -\frac{1}{2}\frac{(a(z_0)-a(z_1))}{\bar{a}^2}H_3(t,\frac{|x|+|x_0|+\ell-\ell_0}{\bar{\sigma}}) - b(z_0)\frac{\mathrm{sign}(x_0)}{\bar{a}^{\frac{3}{2}}}H_2(t,\frac{|x|+|x_0|+\ell-\ell_0}{\bar{\sigma}}), & \ell > \ell_0, \\ \frac{1}{2}(a(z_0)-a(z_1))\{H_2(\bar{a}t, x-x_0) - H_2(\bar{a}t, x+x_0)\} - b(z_0)\{H_1(\bar{a}t, x-x_0) \\ \quad + H_1(\bar{a}t, x+x_0)\}, & \ell = \ell_0. \end{cases}$$

Since $\mathbb{R}^2 \ni z \mapsto a(z)$ is continuous, assumption (H1) (iii) holds. We now prove that (H1) (iv) and (v) are satisfied. Before that, we point out that the inequality

$$(x+x_0)^2 = (x-x_0)^2 + 4xx_0 \ge (x-x_0)^2$$

is valid on the set $\{xx_0 \ge 0\}$ which implies $H_0(ct, x+x_0) \le H_0(ct, x-x_0)$. We first prove (H1) (iv). Using the fact that a is uniformly elliptic and bounded together with the space-time inequality (1.2) and the standard Gaussian estimates (1.3) we can bound $\bar{p}_t^{z_1}(x_0,\ell_0,x,\ell)$ and $\bar{\theta}_t^{z_1}(x_0,\ell_0,x,\ell)$ as follows

$$\left|\bar{p}_t^{z_1}(x_0,\ell_0,x,\ell)\right| \le \begin{cases} Ct^{-\frac{1}{2}}H_0(ct, |x_0|+|x|+\ell-\ell_0), & \ell > \ell_0, \\ CH_0(ct, x_0-x), & x_0x \ge 0, \ell = \ell_0 \end{cases}$$

and

$$\left|\bar{\theta}_t^{z_1}(x_0,\ell_0,x,\ell)\right| \le \begin{cases} C(\frac{1}{t^{\frac{3}{2}}} + \frac{|b|_\infty}{t})H_0(ct, |x_0|+|x|+\ell-\ell_0), & \ell > \ell_0, \\ C(\frac{1}{t} + \frac{|b|_\infty}{t^{\frac{1}{2}}})H_0(ct, x_0-x), & x_0x > 0, \ell = \ell_0. \end{cases}$$

Integrating the two previous estimates against $\nu(x_0,\ell_0,dx,d\ell)$, we conclude that $\bar{\zeta} = -1$ in assumption (H1) (iv). Selecting the freezing point $z_1 = (x,\ell)$ according to the end point of the measure $\nu(x_0,\ell_0,dx,d\ell)$, we obtain

$$\widehat{\theta}_t(x_0,\ell_0,x,\ell) := \bar{\theta}_t^{(x,\ell)}(x_0,\ell_0,x,\ell)$$
$$:= \begin{cases} -\frac{1}{2}\frac{(a(z_0)-a(x,\ell))}{a^2(x,\ell)}H_3(t,\frac{|x|+|x_0|+\ell-\ell_0}{\sigma(x,\ell)}) - b(z_0)\frac{\mathrm{sign}(x_0)}{a^{\frac{3}{2}}(x,\ell)}H_2(t,\frac{|x|+|x_0|+\ell-\ell_0}{\sigma(x,\ell)}), & \ell > \ell_0, \\ \frac{1}{2}(a(z_0)-a(x,\ell))\{H_2(a(x,\ell_0)t, x-x_0) - H_2(a(x,\ell_0)t, x+x_0)\} \\ \quad - b(z_0)\{H_1(a(x,\ell_0)t, x-x_0) + H_1(a(x,\ell_0)t, x+x_0)\}, & \ell = \ell_0. \end{cases}$$

and one has the following bound

$$\left|\widehat{\theta}_t(x_0,\ell_0,x,\ell)\right| \le \begin{cases} C(\frac{1}{t^{\frac{3-\eta}{2}}} + \frac{|b|_\infty}{t})H_0(ct, |x_0|+|x|+\ell-\ell_0), & \ell > \ell_0 \\ C(\frac{1}{t^{1-\frac{\eta}{2}}} + \frac{|b|_\infty}{t^{\frac{1}{2}}})H_0(ct, x_0-x), & x_0x \ge 0, \ell = \ell_0. \end{cases} \tag{A.3}$$

To obtain the previous estimate, we note that on the set $\{\ell > \ell_0\}$, by using uniform ellipticity, boundedness and η-Hölder continuity of the diffusion coefficient a as well as standard estimates for derivatives of the Gaussian density given in (1.3), we obtain

$$\begin{aligned} &\left|\frac{1}{2}\frac{(a(x_0,\ell_0)-a(x,\ell))}{a^2(x,\ell)}H_3\left(t,\frac{|x|+|x_0|+\ell-\ell_0}{\sigma(x,\ell)}\right)\right| \\ &\quad \le C\frac{(|x_0-x|+|\ell_0-\ell|)^\eta}{t^{\frac{3}{2}}}H_0\big(ct, |x|+|x_0|+\ell-\ell_0\big) \\ &\quad \le C\frac{1}{t^{\frac{3-\eta}{2}}}\left(\frac{|x_0|+|x|+\ell-\ell_0}{t^{\frac{1}{2}}}\right)^\eta H_0\big(ct, |x|+|x_0|+\ell-\ell_0\big) \\ &\quad \le C\frac{1}{t^{\frac{3-\eta}{2}}}H_0\big(ct, |x|+|x_0|+\ell-\ell_0\big) \end{aligned}$$

where the last inequality follows from the space-time inequality (1.2). Similarly, to estimate the second term, that is, when $\ell = \ell_0$ and $x_0 x \geq 0$, we have

$$\begin{aligned}
&\left|\frac{1}{2}\big(a(x_0,\ell_0) - a(x,\ell_0)\big)\big\{H_2\big(a(x,\ell_0)t, x - x_0\big) - H_2\big(a(x,\ell_0)t, x + x_0\big)\big\}\right| \\
&\quad \leq C\frac{|x_0 - x|^\eta}{t}\big(H_0(\bar{a}t, x - x_0) + H_0(\bar{a}t, x + x_0)\big) \\
&\quad \leq \frac{C}{t^{1-\frac{\eta}{2}}} H_0(ct, x - x_0)
\end{aligned}$$

where in the last inequality, we have used the fact that on the set $\{x_0 x \geq 0\}$, $(x - x_0)^2 \leq (x + x_0)^2$ and the space time inequality (1.2). The estimates for the term involving $|b|_\infty$ follows directly from the standard Gaussian estimate given in (1.3).

To obtain the estimate (2.4), we integrate the bound (A.3) against $\nu(x_0, \ell_0, dx, d\ell)$. We note that on the set $\{\ell \geq \ell_0\}$, this integral can be estimated through integration by parts with respect to ℓ and using the fact that $t \in (0, T]$, that is

$$\begin{aligned}
&\left(\frac{1}{t^{\frac{3-\eta}{2}}} + \frac{|b|_\infty}{t}\right)\int_{\mathbb{R}}\int_{\ell_0}^\infty H_0\big(ct, |x_0| + |x| + \ell - \ell_0\big)\,dx\,d\ell \\
&\quad \leq C\frac{1}{t^{\frac{3-\eta}{2}}}\int_{\mathbb{R}}\int_0^\infty H_0\big(ct, |x_0| + |x| + \ell\big)\,d\ell\,dx \\
&\quad = \frac{1}{t^{\frac{3-\eta}{2}}}\int_{\mathbb{R}}\int_0^\infty \ell(-H_1)\big(ct, |x_0| + |x| + \ell\big)\,d\ell\,dx \\
&\quad \leq C\frac{1}{t^{\frac{3-\eta}{2}}}\int_0^\infty\int_0^\infty \big(\ell + |x_0|\big)(-H_1)\big(ct, |x_0| + x + \ell\big)\,dx\,d\ell \\
&\quad = C\frac{1}{t^{\frac{3-\eta}{2}}}\int_0^\infty \big(\ell + |x_0|\big)H_0\big(ct, |x_0| + \ell\big)\,d\ell \leq \frac{C}{t^{1-\frac{\eta}{2}}}.
\end{aligned} \tag{A.4}$$

On the set $\{x_0 x \geq 0\} \cap \{\ell = \ell_0\}$, straightforward integration gives

$$\int_{\mathbb{R}}\left(\frac{1}{t^{1-\frac{\eta}{2}}} + \frac{|b|_\infty}{t^{\frac{1}{2}}}\right)H_0(ct, x_0 - x)\,dx \leq \frac{C(1 + |b|_\infty t^{\frac{1-\eta}{2}})}{t^{1-\frac{\eta}{2}}}. \tag{A.5}$$

From the above computations, we conclude that (H1) (iv) is satisfied.

We now prove (H1) (v), namely,

$$\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^2} h(x,\ell)\widehat{p}_\varepsilon(x_0,\ell_0,dx,d\ell) \to h(x_0,\ell_0), \quad h \in \mathcal{C}_b(\mathbb{R}\times\mathbb{R}_+).$$

We consider the change of variable $\ell' = \ell - \ell_0$ and $x = x'$, and decompose $\int_{\mathbb{R}^2} h(x,\ell)\widehat{p}_\varepsilon(x_0,\ell_0,dx,d\ell)$ into the following

$$\int_{\mathbb{R}\times[0,\infty)} h\big(x', \ell_0 + \ell'\big)\big[\widehat{p}_\varepsilon\big(x_0, 0, dx', d\ell'\big) - \bar{p}_\varepsilon\big(x_0, 0, dx', d\ell'\big)\big] + \int_{\mathbb{R}\times[0,\infty)} h\big(x', \ell_0 + \ell'\big)\bar{p}_\varepsilon\big(x_0, 0, dx', d\ell'\big)$$

where the frozen point in $\bar{p}_\varepsilon$ is given by $(x_1, \ell_1) = (x_0, 0)$. It is clear from the continuity of h that the second term converges to $h(x_0, \ell_0)$ as $\varepsilon \downarrow 0$. The first term can be decomposed as follows

$$\begin{aligned}
&\int_{\mathbb{R}\times[0,\infty)} h\big(x', \ell_0 + \ell'\big)\big[\widehat{p}_\varepsilon\big(x_0, 0, dx', d\ell'\big) - \bar{p}_\varepsilon\big(x_0, 0, dx', d\ell'\big)\big] \\
&\quad = \int_{\mathbb{R}} h\big(x', \ell_0\big)\big[H_0\big(a\big(x', \ell_0\big)\varepsilon, x' - x_0\big) - H_0\big(a(x_0, \ell_0)\varepsilon, x' - x_0\big) \\
&\qquad - \big(H_0\big(a\big(x', \ell_0\big)\varepsilon, x' + x_0\big) - H_0\big(a(x_0, \ell_0)\varepsilon, x' + x_0\big)\big)\big]\mathbb{I}_{\{x' x_0 \geq 0\}}\,dx' \\
&\qquad - \int_{\mathbb{R}\times[0,\infty)} h\big(x', \ell' + \ell_0\big)\left[\frac{1}{a(x', \ell' + \ell_0)}H_1\left(\varepsilon, \frac{|x'| + |x_0| + \ell'}{\sigma(x', \ell' + \ell_0)}\right)\right.
\end{aligned}$$

$$
\begin{aligned}
&\quad - \frac{1}{a(x_0,\ell_0)} H_1\left(\varepsilon, \frac{|x'|+|x_0|+\ell'}{\sigma(x_0,\ell_0)}\right)\Bigg] \, dx' \, d\ell' \\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}
$$

To show that the first term vanishes as $\varepsilon \downarrow 0$, we combine the mean value theorem with the Hölder regularity of a and the space-time inequality

$$
|\mathrm{I}| \le C|h|_\infty \int_{\mathbb{R}} \frac{\varepsilon |x'-x_0|^\eta}{\varepsilon} \left\{ H_0\big(c\varepsilon, x'-x_0\big) + H_0\big(c\varepsilon x' + x_0\big) \right\} \mathbb{I}_{\{x'x_0 \ge 0\}} \, dx' \le C|h|_\infty \varepsilon^{\frac{\eta}{2}}.
$$

For the second term, we again combine the mean value theorem with the Hölder regularity of a and the space-time inequality (1.2)

$$
\begin{aligned}
|\mathrm{II}| &\le C|h|_\infty \int_{\mathbb{R}\times[0,\infty)} \big(|x'|+|x_0|+\ell'\big)^\eta \Bigg[H_1\left(\varepsilon, \frac{|x'|+|x_0|+\ell'}{\sigma(x',\ell'+\ell_0)}\right) \\
&\quad + \frac{(|x'|+|x_0|+\ell')^2}{\varepsilon^{\frac{3}{2}}} H_0\left(\varepsilon, \frac{|x'|+|x_0|+\ell'}{\sigma(x',\ell'+\ell_0)}\right)\Bigg] \, dx' \, d\ell' \\
&\le C|h|_\infty \varepsilon^{\frac{\eta}{2}}.
\end{aligned}
$$

We now prove (H1) (vi). To show that $\widehat{P}_{t+r}h(x) \to \widehat{P}_t h(x)$ as $r \downarrow 0$, it is sufficient to apply the dominated convergence together with the fact that h is bounded and that, for fixed $t > 0$ and for all r such that $0 \le r \le T$, $|H_i(c(t+r), x)| \le C(\sqrt{t+T}/t^{\frac{i+1}{2}}) H_0(c(T+t), x)$. Similar arguments yield $\lim_{r\to 0} \mathcal{S}_{t+r}h(x) = \mathcal{S}_t h(x)$.

We now prove that (H2) holds. Let $h \in \mathcal{C}_b^\infty(\mathcal{J})$ and $r > 0$. In order to prove that $(0,\infty) \times \mathcal{J} \ni (r, x_0, \ell_0) \mapsto \bar{R}_\lambda \bar{P}_r h(x_0,\ell_0) = \int_0^\infty e^{-\lambda u} \bar{P}_{u+r} h(x_0,\ell_0)\, du$ belongs to $\mathcal{D}$, we remark that by the Lebesgue differentiation theorem it is enough to check that $\bar{P}_{u+r}h \in \mathcal{D}$ which is exactly assumption (H1) (ii) (a). From (A.1), (A.2) and the fact that the coefficients b and a are bounded, we directly get that the map $(x_0,\ell_0) \mapsto \mathcal{L}\bar{P}_{u+r}h(x_0,\ell_0)$ is bounded (independently of u) so that by Lebesgue's differentiation theorem one gets $\mathcal{L}\bar{R}_\lambda \bar{P}_r h = \int_0^\infty e^{-\lambda u} \mathcal{L}\bar{P}_{u+r} h \, du$.

The identity (2.12) then follows from the Lebesgue differentiation theorem and the fact that $\bar{P}_r h \in \mathcal{D}$ so that $\bar{\mathcal{L}}\bar{P}_{u+r}h = \bar{\mathcal{L}}\bar{P}_u \bar{P}_r h = \bar{P}_u \bar{\mathcal{L}} \bar{P}_r h$. The identity (2.13) follows again from the Lebesgue differentiation theorem. We thus conclude that (H1) and (H2) are satisfied. The proof of the theorem is now complete.

A.2. *Proof of Theorem* 3.4

As previously mentioned, we only prove the result for $b \equiv 0$. In order to include a drift, we refer to Remark 3.5. We examine the n-th term of the series (3.10) and prove an important smoothing property of the kernel. More precisely, let $x = x_{n+1}$, $\ell = \ell_{n+1}$ and $s_0 = T$, we claim the following key inequality

$$
\begin{aligned}
&\left| \int_{(\mathbb{R}\times\mathbb{R}_+)^n} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1) \left\{ \prod_{i=1}^n \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1}) u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i) \right\} \right| \\
&\quad \le \prod_{i=1}^n C(s_{i-1}-s_i)^{-1+\frac{\eta}{2}} \times \left\{ \frac{1}{T^{\frac{1}{2}}} H_0\big(cT, |x|+|x_0|+\ell-\ell_0\big) \mathbb{I}_{\{\ell_0<\ell\}} \right. \\
&\qquad \left. + \left\{ \frac{|x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cT, x-x_0) \mathbb{I}_{\{xx_0\ge 0\}} \mathbb{I}_{\{\ell=\ell_0\}} \right\}
\end{aligned}
\tag{A.6}
$$

for any $\beta \in [0,1]$. From the previous bound and Lemma A.1 we deduce that

$$
\begin{aligned}
&\left| p_T^n(x_0,\ell_0,x,\ell) \right| \\
&\quad \le \left| \int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}\times\mathbb{R}_+)^n} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1) \right. \\
&\qquad \left. \times \left\{ \prod_{i=1}^n \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1}) u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i) \right\} \right|
\end{aligned}
$$

$$
\begin{aligned}
&\le \int_{\Delta_n(T)} d\mathbf{s}_n \prod_{i=1}^n C(s_{i-1}-s_i)^{-1+\frac{\eta}{2}} \times \left\{ \frac{1}{T^{\frac12}} H_0\big(cT, |x|+|x_0|+\ell-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell\}} \right. \\
&\quad \left. + \left\{ \frac{|x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cT, x-x_0)\mathbb{I}_{\{xx_0\ge 0\}}\mathbb{I}_{\{\ell=\ell_0\}} \right\} \\
&= \frac{(CT^{\eta/2}\Gamma(\eta/2))^n}{\Gamma(1+n\eta/2)} \left\{ \frac{1}{T^{\frac12}} H_0\big(cT, |x|+|x_0|+\ell-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell\}} \right. \\
&\quad \left. + \left\{ \frac{|x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cT, x-x_0)\mathbb{I}_{\{xx_0\ge 0\}}\mathbb{I}_{\{\ell=\ell_0\}} \right\}
\end{aligned}
$$

with the convention $s_0 = T$. Hence, from Fubini's theorem, the semigroup series obtained from Corollary 2.1 admits the following integral representation

$$
P_T h(x_0,\ell_0) = \int_{\mathbb{R}\times\mathbb{R}_+} h(x,\ell)\bigg(\sum_{n\ge 0} p_T^n(x_0,\ell_0,x,\ell)\bigg)\nu(x_0,\ell_0,dx,d\ell)
$$

where $p_T^n(x_0,\ell_0,x,\ell)$ is given by (3.13). Moreover, from the above inequality, for any (x_0,ℓ_0), $(x,\ell)\in\mathbb{R}\times\mathbb{R}_+$, one gets the following Gaussian upper bound

$$
\begin{aligned}
&\bigg|\sum_{n\ge 0} p_T^n(x_0,\ell_0,x,\ell)\bigg| \\
&\quad \le C_T \left\{ \frac{1}{T^{\frac12}} H_0\big(cT, |x|+|x_0|+\ell-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell\}} + \left\{ \frac{|x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cT, x-x_0)\mathbb{I}_{\{xx_0\ge 0\}}\mathbb{I}_{\{\ell=\ell_0\}} \right\}
\end{aligned}
$$

where $C_T := \sum_{N\ge 1}(CT^{\eta/2}\Gamma(\eta/2))^N/\Gamma(1+N\eta/2) < \infty$, for some constants $C, c > 1$.

The proof will thus be complete once we prove (A.6). We proceed by induction and show that for $j = 1,\dots,n$, the following estimate holds

$$
\begin{aligned}
&\left| \int_{(\mathbb{R}\times\mathbb{R}_+)^j} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1) \left\{ \prod_{i=1}^j \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1}) u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i) \right\} \right| \\
&\quad \le \prod_{i=1}^j C(s_{n-i}-s_{n-i+1})^{-1+\frac{\eta}{2}} \left\{ \frac{1}{s_{n-j}^{\frac12}} H_0\big(cs_{n-j}, |x_{j+1}|+|x_0|+\ell_{j+1}-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell_{j+1}\}} \right. \\
&\qquad \left. + \left\{ \frac{|x_{j+1}|^\beta}{s_{n-j}^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{s_{n-j}^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cs_{n-j}, x_{j+1}-x_0)\mathbb{I}_{\{x_{j+1}x_0\ge 0\}}\mathbb{I}_{\{\ell_{j+1}=\ell_0\}} \right\}.
\end{aligned}
\tag{A.7}
$$

We start by proving a one step estimate, namely we compute an upper bound for

$$
\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}_+} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1)\widehat{\theta}_{s_{n-1}-s_n}(x_1,\ell_1,x_2,\ell_2) u(x_0,\ell_0,x_2,\ell_2,dx_1,d\ell_1) \\
&\quad = \int_{\mathbb{R}} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_0)\widehat{\theta}_{s_{n-1}-s_n}(x_1,\ell_1,x_2,\ell_1)\mathbb{I}_{\{x_0x_1\ge 0\}}\, dx_1 \mathbb{I}_{\{\ell_2=\ell_0\}}\mathbb{I}_{\{x_2x_0\ge 0\}} \\
&\qquad + \int_{\mathbb{R}\times\mathbb{R}_+} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1)\widehat{\theta}_{s_{n-1}-s_n}(x_1,\ell_1,x_2,\ell_2)\mathbb{I}_{\{x_1x_0\ge 0\}}\, dx_1\delta_{\ell_0}(d\ell_1)\mathbb{I}_{\{\ell_0<\ell_2\}} \\
&\qquad + \int_{\mathbb{R}\times\mathbb{R}_+} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1)\widehat{\theta}_{s_{n-1}-s_n}(x_1,\ell_1,x_2,\ell_2)\mathbb{I}_{\{x_2x_1\ge 0\}}\, dx_1\delta_{\ell_2}(d\ell_1)\mathbb{I}_{\{\ell_0<\ell_2\}} \\
&\qquad + \int_{\mathbb{R}\times\mathbb{R}_+} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1)\widehat{\theta}_{s_{n-1}-s_n}(x_1,\ell_1,x_2,\ell_2)\mathbb{I}_{\{\ell_0\le\ell_1\le\ell_2\}}\, dx_1\, d\ell_1\mathbb{I}_{\{\ell_0<\ell_2\}} \\
&\quad =: A_1\mathbb{I}_{\{\ell_2=\ell_0\}}\mathbb{I}_{\{x_2x_0\ge 0\}} + (A_2+A_3+A_4)\mathbb{I}_{\{\ell_0<\ell_2\}}
\end{aligned}
\tag{A.8}
$$

where we used (3.12). In the following, we present in detail the computation of the estimates for A_i, $i = 1, 2, 3, 4$ and shall be brief in the induction step as the computations are quite similar.

From (3.9), Lemma A.4 of Section A.5 with $r = 2$ and the space-time inequality (1.2), for all $\beta \in [0, 1]$,

$$\begin{aligned}
|A_1| &\le C \int_{\mathbb{R}} \left\{ \frac{|x_1|^\beta}{s_n^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{s_n^{\frac{\beta}{2}}} \wedge 1 \right\} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(c s_n, x_1 - x_0) \\
&\quad \times \left\{ \frac{|x_2|^\beta}{(s_{n-1} - s_n)^{\frac{2-\eta+\beta}{2}}} \wedge \frac{|x_1|^\beta}{(s_{n-1} - s_n)^{\frac{2-\eta+\beta}{2}}} \wedge \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \right\} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \left\{ \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(c s_{n-1}, x_2 - x_0)
\end{aligned} \tag{A.9}$$

where to obtain inequality (A.9), we considered the two cases $s_n \in (0, s_{n-1}/2)$ and $s_n \in (s_{n-1}/2, s_{n-1})$.

Indeed, if $s_n \in (0, s_{n-1}/2)$, one has $(s_{n-1} - s_n) \asymp s_{n-1}$ and by using the inequalities

$$\left\{ \frac{|x_2|^\beta}{(s_{n-1} - s_n)^{\frac{2-\eta+\beta}{2}}} \wedge \frac{|x_1|^\beta}{(s_{n-1} - s_n)^{\frac{2-\eta+\beta}{2}}} \wedge \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \right\} \le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \left\{ \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge 1 \right\} \tag{A.10}$$

$$\left\{ \frac{|x_1|^\beta}{s_n^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{s_n^{\frac{\beta}{2}}} \wedge 1 \right\} \le 1 \tag{A.11}$$

and the semigroup property, one obtains

$$|A_1| \le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \left\{ \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(c s_{n-1}, x_2 - x_0).$$

To obtain the upper bound in (A.9) with $|x_0|^\beta s_{n-1}^{-\beta/2}$, we bound the left-hand side of (A.10) by $|x_1|^\beta (s_{n-1} - s_n)^{-1+\frac{\eta-\beta}{2}}$ and use $|x_1|^\beta \le (|x_1 - x_0|^\beta + |x_0|^\beta)$ to bound $|A_1|$ using two terms. For the first term associated with $|x_1 - x_0|^\beta$, we bound the left-hand side of (A.11) by $|x_0|^\beta s_n^{-\beta/2}$ and for the second term associated with $|x_0|^\beta$, we bound the left hand of (A.11) by 1. This gives

$$\begin{aligned}
|A_1| &\le C \int_{\mathbb{R}} \frac{|x_0|^\beta}{s_n^{\frac{\beta}{2}}} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(c s_n, x_1 - x_0) \frac{1}{s_{n-1}^{\frac{\beta}{2}}} \frac{|x_1 - x_0|^\beta}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\quad + C \int_{\mathbb{R}} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(c s_n, x_1 - x_0) \frac{1}{s_{n-1}^{\frac{\beta}{2}}} \frac{|x_0|^\beta}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{|x_0|^\beta}{s_{n-1}^{\frac{\beta}{2}}} H_0(c s_{n-1}, x_2 - x_0)
\end{aligned}$$

where, in the last inequality above, we applied space-time inequality (1.2) to remove $s_n^{\beta/2}$ in the first term using $|x_0 - x_1|^\beta H(c s_n, x_1 - x_0)$. This proves the desired bound for $s_n \in (0, s_{n-1}/2)$.

For the case $s_n \in (s_{n-1}/2, s_{n-1})$, one has $s_n \asymp s_{n-1}$, and by bounding the left-hand side of (A.11) and (A.10) by $|x_0|^\beta s_n^{-\beta/2}$ and $(s_{n-1} - s_n)^{-1+\eta/2}$ respectively, we obtain

$$\begin{aligned}
|A_1| &\le C \int_{\mathbb{R}} \frac{|x_0|^\beta}{s_n^{\frac{\beta}{2}}} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(c s_n, x_1 - x_0) \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{|x_0|^\beta}{s_{n-1}^{\frac{\beta}{2}}} H_0(c s_{n-1}, x_2 - x_0).
\end{aligned}$$

To obtain the upper bound in (A.9) with $|x_2|^\beta s_{n-1}^{-\beta/2}$, we bound the left hand side of (A.11) by $|x_1|^\beta s_n^{-\frac{\beta}{2}}$ and use $s_{n-1} \asymp s_n$ and $|x_1|^\beta \le (|x_2|^\beta + |x_1 - x_2|^\beta)$ to bound $|A_1|$ by the sum of two terms. For the first term associated with $|x_2|^\beta$, we bound the left-hand side of (A.10) by $(s_{n-1} - s_n)^{-1+\eta/2}$ and for the second term associated with $|x_1 - x_2|^\beta$ we bound the left hand side of (A.10) by $|x_2|^\beta (s_{n-1} - s_n)^{-\beta/2}$. This gives

$$
\begin{aligned}
|A_1| &\le C \int_{\mathbb{R}} \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(cs_n, x_1 - x_0) \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\quad + C \frac{|x_1 - x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \int_{\mathbb{R}} \mathbb{I}_{\{x_1 x_0 \ge 0\}} H_0(cs_n, x_1 - x_0) \frac{|x_2|^\beta}{(s_{n-1} - s_n)^{1-\frac{\eta+\beta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big)\, dx_1 \\
&\le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} H_0(cs_{n-1}, x_2 - x_0)
\end{aligned}
$$

where the last inequality follows from the space-time inequality applied to $|x_1 - x_2|^\beta H_0(c(s_n - s_{n-1}), x_2 - x_1)$.

To estimate A_2, we apply Lemma A.4 with $\beta = 1$, then

$$
\begin{aligned}
|A_2| &\le C \int_{\mathbb{R}} \left\{ \frac{|x_1|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cs_n, x_1 - x_0) \mathbb{I}_{\{x_1 x_0 \ge 0\}} \\
&\quad \times \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} H_0\big(c(s_{n-1} - s_n), |x_2| + |x_1| + \ell_2 - \ell_0\big)\, dx_1 \\
&\le C \left\{ \frac{1}{s_n^{\frac{\beta}{2}} (s_{n-1} - s_n)^{\frac{3-\eta-\beta}{2}}} \wedge \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} \right\} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big) \qquad \text{(A.12)} \\
&\le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big). \qquad \text{(A.13)}
\end{aligned}
$$

To obtain (A.12), we use the fact that

$$
\left\{ \frac{|x_1|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge 1 \right\} \le 1
$$

and convolve the Gaussian kernels using the second inequality in Lemma A.2 to obtain

$$
|A_2| \le C \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} H_0\big(c(s_{n-1} - s_n), |x_2| + |x_0| + \ell_2 - \ell_0\big).
$$

To obtain the upper bound in (A.12) with $s_n^{-\beta/2}(s_{n-1} - s_n)^{-(3-\eta-\beta)/2}$, we note that

$$
\left\{ \frac{|x_1|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{|s_n|^{\frac{\beta}{2}}} \wedge 1 \right\} \le \frac{1}{s_n^{\frac{\beta}{2}}} |x_1|^\beta \le \frac{1}{s_n^{\frac{\beta}{2}}} \big[|x_1| + |x_0| + (\ell_2 - \ell_0)\big]^\beta
$$

and together with the space-time inequality (1.2) gives the $(s_{n-1} - s_n)^{-(3-\eta-\beta)/2}$ term. The result then holds by convolving the Gaussian densities using Lemma A.2. Finally, to obtain (A.13) from (A.12), we again consider two cases. For the case $s_n \in (0, s_{n-1}/2)$, one has $(s_{n-1} - s_n) \asymp s_{n-1}$ and

$$
\frac{1}{s_n^{\frac{\beta}{2}} (s_{n-1} - s_n)^{\frac{3-\eta-\beta}{2}}} \wedge \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} \le \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} \le C \frac{1}{(s_{n-1} - s_n)^{\frac{1-\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}}.
$$

For the case $s_n \in (s_{n-1}/2, s_{n-1})$, one has $s_n \asymp s_{n-1}$. We take $\beta = 1$ in (A.12) so that

$$
\frac{1}{s_n^{\frac{\beta}{2}} (s_{n-1} - s_n)^{\frac{3-\eta-\beta}{2}}} \wedge \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} \le \frac{1}{s_n^{\frac{\beta}{2}} (s_{n-1} - s_n)^{\frac{3-\eta-\beta}{2}}} \le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}}.
$$

To estimate A_3, again by applying Lemma A.4 with $\beta = 1$, one has

$$
\begin{aligned}
|A_3| \le{}& \int_{\mathbb{R}} \frac{C}{s_n^{\frac{1}{2}}} H_0\big(cs_n, |x_1| + |x_0| + \ell_2 - \ell_0\big) \\
&\times \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \left\{ \frac{|x_1|^\beta}{(s_{n-1} - s_n)^{\frac{\beta}{2}}} \wedge 1 \right\} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big) \mathbb{I}_{\{x_2 x_1 \ge 0\}}\, dx_1 \\
\le{}& \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big)
\end{aligned}
\tag{A.14}
$$

where we again separated the computations into the two cases $s_n \in (0, s_{n-1}/2)$ and $s_n \in (s_{n-1}/2, s_{n-1})$. In the case $s_n \in (0, s_{n-1}/2)$, we have $(s_{n-1} - s_n) \asymp s_{n-1}$ and $s_n^{-1/2}$ is removed using the inequality $|x_1| \le |x_1| + |x_0| + \ell_2 - \ell_0$ and the space-time inequality to obtain

$$
\begin{aligned}
|A_3| \le{}& \int_{\mathbb{R}} \frac{C}{s_n^{\frac{1}{2}}} H_0\big(cs_n, |x_1| + |x_0| + \ell_2 - \ell_0\big) \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{|x_1|}{(s_{n-1} - s_n)^{\frac{1}{2}}} \\
&\times H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big) \mathbb{I}_{\{x_2 x_1 \ge 0\}}\, dx_1 \\
\le{}& C \int_{\mathbb{R}} H_0\big(cs_n, |x_1| + |x_0| + \ell_2 - \ell_0\big) \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{(s_{n-1} - s_n)^{\frac{1}{2}}} \\
&\times H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big) \mathbb{I}_{\{x_2 x_1 \ge 0\}}\, dx_1 \\
\le{}& \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big).
\end{aligned}
$$

In the case $s_n \in (s_{n-1}/2, s_{n-1})$, one has $s_n \asymp s_{n-1}$ and it suffices to note that,

$$
\begin{aligned}
|A_3| \le{}& \int_{\mathbb{R}} \frac{C}{s_n^{\frac{1}{2}}} H_0\big(cs_n, |x_1| + |x_0| + \ell_2 - \ell_0\big) \frac{1}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} H_0\big(c(s_{n-1} - s_n), x_2 - x_1\big) \mathbb{I}_{\{x_2 x_1 \ge 0\}}\, dx_1 \\
\le{}& \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big)
\end{aligned}
$$

where the Gaussian densities are convolved using the first inequality in Lemma A.2. Finally, to estimate A_4, by Lemma A.3, one has

$$
\begin{aligned}
|A_4| \le{}& C \int_{\mathbb{R} \times (\ell_0, \ell_2)} \frac{1}{s_n^{\frac{1}{2}}} H_0\big(cs_n, |x_1| + |x_0| + \ell_1 - \ell_0\big) \frac{1}{(s_{n-1} - s_n)^{\frac{3-\eta}{2}}} \\
&\times H_0\big(c(s_{n-1} - s_n), |x_2| + |x_1| + \ell_2 - \ell_1\big)\, dx_1\, d\ell_1 \\
\le{}& \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big).
\end{aligned}
$$

Combining all the previous upper-bounds we finally obtain the first step estimate

$$
\begin{aligned}
&\left| \int_{\mathbb{R} \times \mathbb{R}_+} \widehat{p}_{s_n}(x_0, \ell_0, x_1, \ell_1) \widehat{\theta}_{s_{n-1} - s_n}(x_1, \ell_1, x_2, \ell_2) u(x_0, \ell_0, x_2, \ell_2, dx_1, d\ell_1) \right| \\
&\quad \le \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \frac{1}{s_{n-1}^{\frac{1}{2}}} H_0\big(cs_{n-1}, |x_2| + |x_0| + \ell_2 - \ell_0\big) \mathbb{I}_{\{\ell_0 < \ell_2\}} \\
&\qquad + \frac{C}{(s_{n-1} - s_n)^{1-\frac{\eta}{2}}} \left\{ \frac{|x_2|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{s_{n-1}^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(cs_{n-1}, x_2 - x_0) \mathbb{I}_{\{x_2 x_0 \ge 0\}} \mathbb{I}_{\{\ell_2 = \ell_0\}}
\end{aligned}
\tag{A.15}
$$

for any $\beta \in [0,1]$ and for some positive constants $C := C(a), c$. Now, we assume that the estimate (A.7) is valid at step j and we prove that a similar bound holds at step $j+1$, namely

$$\begin{aligned}
&\left|\int_{(\mathbb{R}\times\mathbb{R}_+)^{j+1}} \widehat{p}_{s_n}(x_0,\ell_0,x_1,\ell_1)\left\{\prod_{i=1}^{j+1}\widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,\ell_i,x_{i+1},\ell_{i+1})u(x_0,\ell_0,x_{i+1},\ell_{i+1},dx_i,d\ell_i)\right\}\right| \\
&\quad\le \prod_{i=1}^{j+1} C(s_{n-i}-s_{n-i+1})^{-1+\frac{\eta}{2}}\left\{\frac{1}{s_{n-(j+1)}^{\frac12}}H_0\big(cs_{n-(j+1)},|x_{j+2}|+|x_0|+\ell_{j+2}-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell_{j+2}\}}\right. \\
&\qquad\left.+\left\{\frac{|x_{j+2}|^\beta}{s_{n-(j+1)}^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{s_{n-(j+1)}^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cs_{n-(j+1)},x_{j+2}-x_0)\mathbb{I}_{\{x_{j+2}x_0\ge 0\}}\mathbb{I}_{\{\ell_{j+2}=\ell_0\}}\right\}.
\end{aligned}\tag{A.16}$$

From (A.7) at step j, the left-hand side of (A.16) is bounded by

$$\begin{aligned}
&\prod_{i=1}^{j} C(s_{n-i}-s_{n-i+1})^{-1+\frac{\eta}{2}}\int_{\mathbb{R}\times\mathbb{R}_+}\left\{\frac{1}{s_{n-j}^{\frac12}}H_0\big(cs_{n-j},|x_{j+1}|+|x_0|+\ell_{j+1}-\ell_0\big)\mathbb{I}_{\{\ell_0<\ell_{j+1}\}}\right. \\
&\quad\left.+\left\{\frac{|x_{j+1}|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge 1\right\}\times H_0(cs_{n-j},x_{j+1}-x_0)\mathbb{I}_{\{x_{j+1}x_0\ge 0\}}\mathbb{I}_{\{\ell_{j+1}=\ell_0\}}\right\} \\
&\quad\times\big|\widehat{\theta}_{s_{n-(j+1)}-s_{n-j}}(x_{j+1},\ell_{j+1},x_{j+2},\ell_{j+2})\big|u(x_0,\ell_0,x_{j+2},\ell_{j+2},dx_{j+1},d\ell_{j+1})
\end{aligned}$$

which is in turn equal to $\prod_{i=1}^{j}C(s_{n-i}-s_{n-i+1})^{-1+\frac{\eta}{2}}\{A_1\mathbb{I}_{\{x_{j+2}x_0\ge0\}}\mathbb{I}_{\{\ell_{j+2}=\ell_0\}}+(A_2+A_3+A_4)\mathbb{I}_{\{\ell_0<\ell_{j+2}\}}\}$ with

$$\begin{aligned}
A_1 &:= \int_{\mathbb{R}}\left\{\frac{|x_{j+1}|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cs_{n-j},x_{j+1}-x_0)\mathbb{I}_{\{x_{j+1}x_0\ge0\}}\frac{1}{(s_{n-(j+1)}-s_{n-j})^{1-\frac{\eta}{2}}} \\
&\quad\times\left\{\frac{|x_{j+2}|^\beta}{(s_{n-(j+1)}-s_{n-j})^{\frac{\beta}{2}}}\wedge\frac{|x_{j+1}|^\beta}{(s_{n-(j+1)}-s_{n-j})^{\frac{\beta}{2}}}\wedge 1\right\} \\
&\quad\times H_0\big(c(s_{n-(j+1)}-s_{n-j}),x_{j+2}-x_{j+1}\big)\mathbb{I}_{\{x_{j+2}x_{j+1}\ge0\}}\,dx_{j+1} \\
&\le \frac{C}{(s_{n-(j+1)}-s_{n-j})^{1-\frac{\eta}{2}}}\left\{\frac{|x_{j+2}|^\beta}{s_{n-(j+1)}^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{s_{n-(j+1)}^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cs_{n-(j+1)},x_{j+2}-x_0)
\end{aligned}$$

where the last inequality follows from considering separately the two cases $s_{n-j}\in(0,s_{n-(j+1)}/2)$ and $s_{n-j}\in(s_{n-(j+1)}/2,s_{n-(j+1)})$ by using similar arguments as those employed to establish the estimate (A.9).

To estimate A_2, by Lemma A.4, for any $\beta\in[0,1]$, one has

$$\begin{aligned}
A_2 &:= \int_{\mathbb{R}}\left\{\frac{|x_{j+1}|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge\frac{|x_0|^\beta}{s_{n-j}^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cs_{n-j},x_{j+1}-x_0)\mathbb{I}_{\{x_{j+1}x_0\ge0\}}\frac{1}{(s_{n-(j+1)}-s_{n-j})^{\frac{3-\eta}{2}}} \\
&\quad\times H_0\big(c(s_{n-(j+1)}-s_{n-j}),|x_{j+2}|+|x_{j+1}|+\ell_{j+2}-\ell_0\big)\,dx_{j+1} \\
&\le\left\{\frac{1}{s_{n-j}^{\frac{\beta}{2}}(s_{n-(j+1)}-s_{n-j})^{\frac{3-\eta-\beta}{2}}}\wedge\frac{1}{(s_{n-(j+1)}-s_{n-j})^{\frac{3-\eta}{2}}}\right\}H_0\big(cs_{n-(j+1)},|x_{j+2}|+|x_0|+\ell_{j+2}-\ell_0\big) \\
&\le\frac{C}{(s_{n-(j+1)}-s_{n-j})^{1-\frac{\eta}{2}}}\frac{1}{s_{n-(j+1)}^{\frac12}}H_0\big(cs_{n-(j+1)},|x_{j+2}|+|x_0|+\ell_{j+2}-\ell_0\big)
\end{aligned}\tag{A.17}$$

where to obtain the last inequality we again separated the computations into the two cases $s_{n-j}\in(0,s_{n-(j+1)}/2)$ and $s_{n-j}\in(s_{n-(j+1)}/2,s_{n-(j+1)})$ which is estimated using similar argument to (A.13).

To estimate A_3, similarly, from Lemma A.4 with $\beta = 1$, one has

$$
\begin{aligned}
A_3 &:= C \int_{\mathbb{R}} \frac{1}{s_{n-j}^{\frac{1}{2}}} H_0\big(cs_{n-j}, |x_{j+1}| + |x_0| + \ell_{j+1} - \ell_0\big) \frac{1}{(s_{n-(j+1)} - s_{n-j})^{1-\frac{\eta}{2}}} \left\{ \frac{|x_j|^\beta}{(s_{n-(j+1)} - s_{n-j})^{\frac{\beta}{2}}} \wedge 1 \right\} \\
&\quad \times H_0\big(c(s_{n-(j+1)} - s_{n-j}), x_{j+2} - x_{j+1}\big) \mathbb{I}_{\{x_{j+1} x_{j+2} \geq 0\}} \, dx_{j+1} \\
&\leq \frac{C}{(s_{n-(j+1)} - s_{n-j})^{1-\frac{\eta}{2}}} \frac{1}{s_{n-(j+1)}^{\frac{1}{2}}} H_0\big(cs_{n-(j+1)}, |x_{j+2}| + |x_0| + \ell_{j+2} - \ell_0\big)
\end{aligned}
$$

where we followed similar arguments as done in order to obtain the inequality (A.14). Finally, from Lemma A.3, one has

$$
\begin{aligned}
A_4 &:= C \int_{\mathbb{R} \times (\ell_0, \ell_j)} \frac{1}{s_{j+1}^{\frac{1}{2}}} H_0\big(cs_{n-j}, |x_{j+1}| + |x_0| + \ell_{j+1} - \ell_0\big) \frac{1}{(s_{n-(j+1)} - s_{n-j})^{\frac{3-\eta}{2}}} \\
&\quad \times H_0\big(c(s_{n-(j+1)} - s_{n-j}), |x_{j+2}| + |x_{j+1}| + \ell_{j+2} - \ell_{j+1}\big) \, dx_{j+1} \, d\ell_{j+1} \\
&\leq \frac{C}{(s_{n-(j+1)} - s_{n-j})^{1-\frac{\eta}{2}}} \frac{1}{s_{n-(j+1)}^{\frac{1}{2}}} H_0\big(cs_{n-(j+1)}, |x_{j+2}| + |x_0| + \ell_{j+2} - \ell_0\big).
\end{aligned}
$$

Hence (A.16) is valid and therefore by induction, the estimate (A.7) holds for $j = 1, \ldots, n$. Now, the Gaussian bound (A.6) follows from (A.7) by taking $j = n$ and applying the change of variable $k = n - i$.

A.3. *Proof of Theorem* 3.7

From the very expression of $\bar{P}_t h$, we remark that $(x_0, m_0) \mapsto \bar{P}_t h(x_0, m_0) \in \mathcal{C}_b^{2,1}(\mathcal{J})$ if $h \in \mathcal{C}_b^\infty(\mathcal{J})$ and satisfies the condition $\partial_2 \bar{P}_t h(m_0, m_0) = \lim_{x_0 \uparrow m_0} \partial_2 \bar{P}_t h(x_0, m_0) = 0$, $m_0 \in \mathbb{R}$. Indeed, from (3.18) and (3.19), for $x_0 < m_0$ and $t > 0$, we obtain

$$
\begin{aligned}
&\partial_{m_0} \bar{P}_t h(x_0, m_0) \\
&\quad = \partial_{m_0} \int_{-\infty}^{m_0} h(x, m_0) \big\{ H_0(\bar{a}t, x - x_0) - H_0(\bar{a}t, 2m_0 - x - x_0) \big\} \, dx \\
&\qquad + \partial_{m_0} \int_{m_0}^{\infty} \int_{-\infty}^{m} h(x, m)(-2H_1)(\bar{a}t, 2m - x - x_0) \, dx \, dm \\
&\quad = \int_{-\infty}^{m_0} \partial_2 h(x, m_0) \big\{ H_0(\bar{a}t, m_0 - x_0) - H_0(\bar{a}t, 2m_0 - x - x_0) \big\} \, dx \\
&\qquad + \int_{-\infty}^{m_0} h(x, m_0)(-2H_1)(\bar{a}t, 2m_0 - x - x_0) \, dx \\
&\qquad + \int_{-\infty}^{m_0} h(x, m_0)(2H_1)(\bar{a}t, 2m_0 - x - x_0) \, dx,
\end{aligned}
$$

and by letting $x_0 \uparrow m_0$, we have $\partial_2 \bar{P}_t h(m_0, m_0) = 0$. Moreover, simple computations that are similar to the case of the local time, namely by using dominated convergence theorem and the heat equation satisfied by the underlying kernels, one can show that $\partial_t \bar{P}_t h(x_0, m_0) = \bar{\mathcal{L}} \bar{P}_t h(x_0, m_0)$, for any $t > 0$ and any $(x_0, m_0) \in \mathcal{J}$. If $h \in \mathcal{D}$, from the chain rule formula (3.17), we get $\partial_t \bar{P}_t h(x_0, m_0) = \bar{P}_t \bar{\mathcal{L}} h(x_0, m_0)$. Then, we write

$$
\begin{aligned}
(\mathcal{L} - \bar{\mathcal{L}}) \bar{P}_t h(x_0, m_0) &= \frac{1}{2} \big(a(x_0, m_0) - a(x_1, m_1)\big) \partial_1^2 \bar{P}_t h(x_0, m_0) + b(x_0, m_0) \partial_1 \bar{P}_t h(x_0, m_0) \\
&= \int_{\mathbb{R}^2} h(x, m) \bar{\theta}_t^{(x_1, m_1)}(x_0, m_0, x, m) \nu(x_0, m_0, dx, dm)
\end{aligned}
$$

where $\nu(x_0, m_0, dx, dm) = \mathbb{I}_{\{x\le m\}}\mathbb{I}_{\{m_0<m\}}\,dx\,dm + \mathbb{I}_{\{x\le m_0\}}\,dx\,\delta_{m_0}(dm)$ and

$$\bar{\theta}_t^{(x_1,m_1)}(x_0, m_0, x, m)$$
$$:= \begin{cases} \frac{1}{2}(a(x_0,m_0) - a(x_1,m_1))(-2H_3)(\bar{a}t, 2m - x - x_0) \\ \quad + b(x_0,m_0)(-2H_2)(\bar{a}t, 2m - x - x_0), & x \le m, m_0 < m, \\ \frac{1}{2}(a(x_0,m_0) - a(x_1,m_1))(H_2(\bar{a}t, x - x_0) - H_2(\bar{a}t, 2m_0 - x - x_0)) & x < m_0, m = m_0 \\ \quad + b(x_0,m_0)(H_1(\bar{a}t, 2m_0 - x - x_0) - H_1(\bar{a}t, x - x_0)). \end{cases}$$

Hence, we see that (H1) (i), (ii), (iii) hold.

We now verify (H1) (iv) for $\bar{\zeta} = -1$. We proceed as in Section A.1. From (UE), there exists positive constants $C, c > 0$ such that for any $t > 0$ and $(x_0, m_0) \in \mathcal{J}$, one has

$$\left|\bar{p}_t(x_0, m_0, x, m)\right| \le C\left\{(-H_1)(ct, 2m - x - x_0)\mathbb{I}_{\{x\le m\}}\mathbb{I}_{\{m<m_0\}} + H_0(ct, x - x_0)\mathbb{I}_{\{m=m_0\}}\right\}$$

and the right-hand side of the above inequality is $\nu(x_0, m_0, dx, dm)$ integrable.

In the same spirit, from (UE), (HR) and the space-time inequality (1.3), we bound $\bar{\theta}_t^{(x_1,m_1)}(x_0, m_0, x, m)$ as follows

$$\left|\bar{\theta}_t^{(x_1,m_1)}(x_0, m_0, x, m)\right| \le \begin{cases} C(\frac{1}{t^{\frac{3}{2}}} + \frac{|b|_\infty}{t})H_0(ct, 2m - x - x_0), & x \le m, m_0 < m, \\ C(\frac{1}{t} + \frac{|b|_\infty}{t^{\frac{1}{2}}})H_0(ct, x - x_0), & x < m_0, m = m_0. \end{cases}$$

By integrating the above with respect to $\nu(x_0, m_0, dx, dm)$ yields

$$\begin{aligned} &\int_{\mathbb{R}^2} \left|\bar{\theta}_t^{(x_1,m_1)}(x_0, m_0, x, m)\right| \nu(x_0, m_0, dx, dm) \\ &\quad \le C\left(\frac{1}{t^{\frac{3}{2}}} + \frac{|b|_\infty}{t}\right)\int_{m_0}^{\infty}\int_{-\infty}^{m} H_0(ct, 2m - x - x_0)\,dx\,dm + C\left(\frac{1}{t} + \frac{|b|_\infty}{t^{\frac{1}{2}}}\right)\int_{-\infty}^{m_0} H_0(ct, x - x_0)\,dx \\ &\quad \le C\left(\frac{1}{t^{\frac{3}{2}}} + \frac{|b|_\infty}{t}\right)\int_{m_0}^{\infty} (m - m_0)H_0(ct, m - m_0)\,dm + \frac{C(1 + |b|_\infty t^{\frac{1}{2}})}{t} \\ &\quad \le \frac{C(1 + |b|_\infty t^{\frac{1}{2}})}{t} \end{aligned}$$

where to obtain the second inequality, we apply integration by parts formula w.r.t. x in the first integral and bound the second integral by one. In the third inequality we applied the space-time inequality (1.2).

Similarly, by selecting the freezing point $z_1 = (x_1, m_1)$ according to the end point of the measure $\nu(x_0, m_0, dx, dm)$, and by noticing that $m - m_0 \le 2m - x - x_0$, for $m_0 \le m$, $x_0 \le x$, we obtain from (HR), (UE) and the space-time inequality (using similar arguments to those given just below (A.3) in the local time case)

$$\left|\widehat{\theta}_t(x_0, m_0, x, m)\right| \le \begin{cases} C(\frac{1}{t^{\frac{3-\eta}{2}}} + \frac{|b|_\infty}{t})H_0(ct, 2m - x - x_0), & x \le m, m_0 < m, \\ C(\frac{1}{t^{1-\frac{\eta}{2}}} + \frac{|b|_\infty}{t^{\frac{1}{2}}})H_0(ct, x - x_0), & x < m_0, m = m_0. \end{cases}$$

The above estimate shows that the order of the time singularity of $\widehat{\theta}_t$ is better than that of $\bar{\theta}_t$. This is because by freezing the diffusion coefficient at the terminal point (x, m), we gained space regularity from the Hölder continuity of the diffusion coefficient which is then used to improve the order of time singularity through the space-time inequality (1.2). This is the *smoothing property* discussed in the introduction since

$$\begin{aligned} \int_{\mathbb{R}^2} \left|\widehat{\theta}_t(x_0, m_0, x, m)\right| \nu(x_0, m_0, dx, dm) &\le C\left(\frac{1}{t^{\frac{3-\eta}{2}}} + \frac{|b|_\infty}{t}\right)\int_{m_0}^{\infty}\int_{-\infty}^{m} H_0(ct, 2m - x - x_0)\,dx\,dm \\ &\quad + C\left(\frac{1}{t^{1-\frac{\eta}{2}}} + \frac{|b|_\infty}{t^{\frac{1}{2}}}\right)\int_{-\infty}^{m_0} H_0(ct, x - x_0)\,dx \end{aligned}$$

$$\le C\left(\frac{1}{t^{\frac{3-\eta}{2}}}+\frac{|b|_\infty}{t}\right)\int_{m_0}^{\infty}(m-m_0)H_0(ct,m-m_0)\,dm+\frac{C_t}{t^{1-\frac{\eta}{2}}}$$
$$\le \frac{C_t}{t^{1-\frac{\eta}{2}}}$$

where to obtain the second last inequality, we have applied integration by parts with respect to x and taken $C_t = C(1+|b|_\infty t^{(1-\eta)/2})$. Hence, we conclude that (H1) (iv) is satisfied with $\bar\zeta = -1$ and $\zeta = -1+\eta/2$.

We now prove (H1) (v). We decompose $\int_{\mathbb{R}^2} h(x,m)\widehat{p}_\varepsilon(x_0,m_0,dx,dm)$ as follows

$$\int_{\mathbb{R}^2} h(x,m)\big[\widehat{p}_\varepsilon(x_0,m_0,dx,dm)-\bar p_\varepsilon(x_0,m_0,dx,dm)\big]+\int_{\mathbb{R}^2} h(x,m)\bar p_\varepsilon(x_0,m_0,dx,dm)$$

with $\bar p_\varepsilon(x_0,m_0,x,m)\nu(x_0,m_0,dx,dm)$, $\bar p_\varepsilon(x_0,m_0,x,m) := \bar f_\varepsilon(x_0,x)\mathbb{I}_{\{m=m_0\}}+\bar q_\varepsilon(x_0,m_0,x,m)\mathbb{I}_{\{m_0<m\}}$ and where the point at which we freeze the coefficient is $z_1=(x_0,m_0)$. From the continuity of h and path continuity of approximation process it is clear that the second term converges to $h(x_0,m_0)$ as $\varepsilon\downarrow 0$. It thus suffices to prove that the first term goes to zero as $\varepsilon\downarrow 0$. By the mean value theorem and the space-time inequality, we obtain

$$\begin{aligned}\big|\widehat{f}_\varepsilon(x_0,x)-\bar f_\varepsilon(x_0,x)\big| &\le C\varepsilon\frac{|x-x_0|^\eta}{\varepsilon}\big\{H_0(c\varepsilon,x-x_0)+H_0(c\varepsilon,2m_0-x-x_0)\big\}\\ &\le C\varepsilon^{\frac{\eta}{2}}H_0(c\varepsilon,x-x_0)\end{aligned}$$

and, similarly,

$$\begin{aligned}\big|\widehat{q}_\varepsilon(x_0,m_0,x,m)-\bar q_\varepsilon(x_0,m_0,x,m)\big| &\le C|x-x_0|\frac{(|x-x_0|^\eta+|m-m_0|^\eta)}{\varepsilon}H_0(c\varepsilon,2m-x-x_0)\\ &\le C\frac{|2m-x-x_0|^{1+\eta}}{\varepsilon}H_0(c\varepsilon,2m-x-x_0)\\ &\le C\varepsilon^{\frac{\eta}{2}}\frac{1}{\varepsilon^{\frac12}}H_0(c\varepsilon,2m-x-x_0).\end{aligned}$$

Integrating the two previous bounds against $\nu(x_0,m_0,dx,dm)$ we obtain

$$\left|\int_{\mathbb{R}^2} h(x,m)\big[\widehat{p}_\varepsilon(x_0,m_0,dx,dm)-\bar p_\varepsilon(x_0,m_0,dx,dm)\big]\right|\le C|h|_\infty\varepsilon^{\frac{\eta}{2}}$$

and we conclude by letting $\varepsilon\downarrow 0$.

In order to prove that assumption (H1) (vi) holds we proceed as in the case of the local time, namely $\widehat{P}_{t+r}h\to\widehat{P}_t h$ as $r\downarrow 0$, follows from the dominated convergence theorem together with the fact that h is bounded and the inequality for fixed $t>0$ and all r such that $0\le r\le T$, $|H_i(c(t+r),x)|\le C(\sqrt{t+T}/t^{\frac{i+1}{2}})H_0(c(T+t),x)$. Similar arguments also yield $\lim_{r\to 0}\mathcal{S}_{t+r}h(x)=S_t h(x)$.

One can verify that assumption (H2) holds by following the same arguments as in the case of the diffusion process and its running local time. The remaining technical details are thus omitted.

A.4. *Proof of Theorem* 3.8

As the technique of proof is very similar to the local time case, we will not present all details in the following computation. In the model of the SDE with its running maximum, we examine the n-th term of the series (3.21) and follow the convention $x=x_{n+1}$, $m=m_{n+1}$ and $s_0=T$. For any $\beta\in[0,1]$, we prove the following key inequality

$$\begin{aligned}&\left|\int_{(\mathbb{R}^2)^n}\widehat{p}_{s_n}(x_0,m_0,x_1,m_1)\left\{\prod_{i=1}^n\widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i,m_i,x_{i+1},m_{i+1})u(x_0,m_0,x_{i+1},m_{i+1},dx_i,dm_i)\right\}\right|\\ &\quad\le\prod_{k=1}^n C(s_{k-1}-s_k)^{-1+\frac{\eta}{2}}\left\{\frac{1}{T^{\frac12}}H_0(cT,2m-x-x_0)\mathbb{I}_{\{x<m\}}\mathbb{I}_{\{\ell_0<m\}}\right.\\ &\qquad\left.+\left\{\frac{|m_0-x|^\beta}{T^{\frac{\beta}{2}}}\wedge\frac{|m_0-x_0|^\beta}{T^{\frac{\beta}{2}}}\wedge 1\right\}H_0(cT,x-x_0)\mathbb{I}_{\{x<m_0\}}\mathbb{I}_{\{m=m_0\}}.\right\}\end{aligned}\tag{A.18}$$

From the previous bound, we deduce that

$$
\begin{aligned}
&\left|\int_{\Delta_n(T)} d\mathbf{s}_n \int_{(\mathbb{R}^2)^n} \widehat{p}_{s_n}(x_0, m_0, x_1, m_1)\left\{\prod_{i=1}^{n} \widehat{\theta}_{s_{n-i}-s_{n-i+1}}(x_i, m_i, x_{i+1}, m_{i+1}) u(x_0, m_0, x_{i+1}, m_{i+1}, dx_i, dm_i)\right\}\right| \\
&\quad \le \int_{\Delta_n(T)} d\mathbf{s}_n \prod_{k=1}^{n} C(s_{k-1}-s_k)^{-1+\frac{\eta}{2}} \left\{\frac{1}{T^{\frac{1}{2}}} H_0(cT, 2m-x-x_0)\mathbb{I}_{\{x<m\}}\mathbb{I}_{\{m_0<m\}}\right. \\
&\qquad \left. + \left\{\frac{|m-x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|m-x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1\right\} H_0(cT, x-x_0)\mathbb{I}_{\{x<m_0\}}\mathbb{I}_{\{m=m_0\}}\right\} \\
&\quad = C(N,T)\left\{\frac{1}{T^{\frac{1}{2}}} H_0(cT, 2m-x-x_0)\mathbb{I}_{\{x<m\}}\mathbb{I}_{\{m_0<m\}}\right. \\
&\qquad \left. + \left\{\frac{|m-x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|m-x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1\right\} H_0(cT, x-x_0)\mathbb{I}_{\{x<m_0\}}\mathbb{I}_{\{m=m_0\}}\right\}
\end{aligned}
$$

where, from Lemma A.1, we have $C(N,T) := \frac{(CT^{\eta/2}\Gamma(\eta/2))^N}{\Gamma(1+N\eta/2)}$. Hence, from Fubini's theorem, the semigroup series obtained from Corollary 2.1 admits the following integral representation

$$
P_T h(x_0, m_0) = \int_{\mathbb{R}^2} h(x,m)\left(\sum_{n\ge 0} p_T^n(x_0, m_0, x, m)\right)\nu(x_0, m_0, dx, dm)
$$

where $p_T^n(x_0, m_0, x, m)$ is given by (3.24). Moreover, from the above inequality, for any $(x_0, m_0), (x, m) \in \mathcal{J}^2$, one gets the following Gaussian upper bounds

$$
\begin{aligned}
\left|p_T(x_0, m_0, x, m)\right| &:= \left|\sum_{n\ge 0} p_T^n(x_0, m_0, x, m)\right| \\
&\le C_T\left\{\frac{1}{\sqrt{T}} H_0(cT, 2m-x-x_0)\mathbb{I}_{\{x\le m\}}\mathbb{I}_{\{m_0<m\}}\right. \\
&\quad \left. + \left\{\frac{|m_0-x|^\beta}{T^{\frac{\beta}{2}}} \wedge \frac{|m_0-x_0|^\beta}{T^{\frac{\beta}{2}}} \wedge 1\right\} H_0(cT, x-x_0)\mathbb{I}_{\{x<m_0\}}\mathbb{I}_{\{m=m_0\}}\right\}
\end{aligned}
\tag{A.19}
$$

where $C_T := \sum_{N\ge 1}(CT^{\eta/2}\Gamma(\eta/2))^N / \Gamma(1+N\eta/2) < \infty$, for some constants $C, c > 1$. Hence it remains to prove (A.18). Since its proof is similar to the proof of (A.6) in the case of local time, we briefly present the guidelines and omit technical details. We note that from Lemma A.4 and the standard Gaussian estimate (1.3), the following estimates hold

$$
\left|\widehat{p}_t(x_0, m_0, x, m)\right| \le \begin{cases} \frac{C}{t^{\frac{1}{2}}} H_0(ct, 2m-x-x_0), & x \le m, m_0 < m, \\ C\{\frac{|m-x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|m-x|^\beta}{t^{\frac{\beta}{2}}} \wedge 1\} H_0(ct, x-x_0), & x < m_0, m = m_0, \end{cases}
\tag{A.20}
$$

and similarly

$$
\left|\widehat{\theta}_t(x_0, m_0, x, m)\right| \le \begin{cases} \frac{C}{t^{\frac{3-\eta}{2}}} H_0(ct, 2m-x-x_0), & x \le m, m_0 < m, \\ \frac{C}{t^{1-\frac{\eta}{2}}}\{\frac{|m-x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|m-x|^\beta}{t^{\frac{\beta}{2}}} \wedge 1\} H_0(ct, x-x_0), & x < m_0, m = m_0 \end{cases}
\tag{A.21}
$$

where $\beta \in [0, 1]$ can be freely chosen. We proceed in a similar fashion to the case of the local time and first compute an upper bound for

$$
\begin{aligned}
&\int_{\mathbb{R}^2} \widehat{p}_s\big(x_0, m_0, x', m'\big)\widehat{\theta}_{t-s}\big(x', m', x, m\big) u\big(x_0, m_0, x, m, dx', dm'\big) \\
&\quad = \int_{\mathbb{R}^2} \widehat{p}_s\big(x_0, m_0, x', m'\big)\widehat{\theta}_{t-s}\big(x', m', x, m\big)\mathbb{I}_{\{x'\vee x_0\vee m_0<m'\}}\mathbb{I}_{\{x\vee x'\vee m'<m\}}\mathbb{I}_{\{m'<m\}}\, dx'dm'
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\mathbb{R}^2}\widehat{p}_s\big(x_0,m_0,x',m'\big)\mathbb{I}_{\{m_0=m'\}}\widehat{\theta}_{t-s}\big(x',m',x,m\big)\mathbb{I}_{\{x\vee x'\vee m'<m\}}\mathbb{I}_{\{x'<m_0\}}\delta_{m_0}\big(dm'\big)\,dx'\\
&+\int_{\mathbb{R}^2}\widehat{p}_s\big(x_0,m_0,x',m'\big)\mathbb{I}_{\{x'\vee x_0\vee m_0<m'\}}\widehat{\theta}_{t-s}\big(x',m',x,m\big)\mathbb{I}_{\{m'=m\}}\mathbb{I}_{\{x'<m\}}\delta_{m}\big(dm'\big)\,dx'\\
&+\int_{\mathbb{R}^2}\widehat{p}_s\big(x_0,m_0,x',m'\big)\widehat{\theta}_{t-s}\big(x',m',x,m\big)\mathbb{I}_{\{x'<m_0\}}\mathbb{I}_{\{x<m'\}}\mathbb{I}_{\{m=m'\}}\delta_{m_0}\big(dm'\big)\,dx'\\
&=:(A_1+A_2+A_3)\mathbb{I}_{\{x\vee x_0\vee m_0<m\}}+A_4\mathbb{I}_{\{x<m_0\}}\mathbb{I}_{\{m_0=m\}}
\end{aligned}
\tag{A.22}
$$

where we used (3.23). From Lemma A.3, (A.20) and (A.21), one directly gets

$$
|A_1|\le\frac{1}{(t-s)^{1-\frac{\eta}{2}}}\frac{C}{\sqrt{t}}H_0(ct,2m-x-x_0).
$$

For the term A_2, we notice that $m_0=m'<m$ and for $\beta=1$, we obtain the following bound

$$
\begin{aligned}
|A_2|&\le\frac{C}{(t-s)^{\frac{3-\eta}{2}}}\int_{\mathbb{R}}\left\{\frac{|m_0-x'|^\beta}{s^{\frac{\beta}{2}}}\wedge\frac{|m_0-x_0|}{s^{\frac{\beta}{2}}}\wedge 1\right\}H_0\big(cs,x'-x_0\big)\\
&\quad\times H_0\big(c(t-s),2m-x-x'\big)\mathbb{I}_{\{x\vee x'\vee m_0<m\}}\mathbb{I}_{\{x'<m_0\}}\,dx'\\
&\le\frac{C}{(t-s)^{1-\frac{\eta}{2}+\frac{1}{2}}}\left\{\frac{|m'-x'+m-x|^\beta}{s^{\frac{\beta}{2}}}\mathbb{I}_{\{s\in(\frac{t}{2},t]\}}+\mathbb{I}_{\{s\in(0,\frac{t}{2})\}}\right\}\\
&\quad\times\int_{\mathbb{R}}H_0\big(cs,x'-x_0\big)H_0\big(c(t-s),2m-x-x'\big)\mathbb{I}_{\{x\vee x'\vee m_0<m\}}\mathbb{I}_{\{x'<m_0\}}\,dx'\\
&\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\frac{1}{t^{\frac{1}{2}}}H_0(ct,2m-x-x_0)
\end{aligned}
$$

where the last inequality follows from the fact that $m'-x'+m-x\le 2m-x-x'$ and the space-time inequality (1.2) in the case $s\in(\frac{t}{2},t)$ and Gaussian convolution together with $(t-s)\asymp t$ if $s\in[0,t/2]$.

For the term A_3, we note that $m_0<m$ and $|m-x|\le|m-x+m_0-x_0|\le|2m-x-x_0|$. Hence, one has

$$
\begin{aligned}
|A_3|&\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\int_{\mathbb{R}}\frac{1}{s^{\frac{1}{2}}}H_0\big(cs,2m-x'-x_0\big)\left\{\frac{|m-x|^\beta}{(t-s)^{\frac{\beta}{2}}}\wedge\frac{|m-x'|^\beta}{(t-s)^{\frac{\beta}{2}}}\wedge 1\right\}\\
&\quad\times H_0\big(c(t-s),x-x'\big)\mathbb{I}_{\{x'\vee x_0\vee m_0<m\}}\mathbb{I}_{\{x'<m\}}\,dx'.
\end{aligned}
$$

On the set $s\in(\frac{t}{2},t)$, by the Gaussian convolution, we obtain

$$
|A_3|\le\frac{C}{s^{\frac{1}{2}}}\frac{1}{(t-s)^{1-\frac{\eta}{2}}}H_0(ct,2m-x-x_0)\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\frac{1}{t^{\frac{1}{2}}}H_0(ct,2m-x-x_0).
$$

On the set $s\in(0,\frac{t}{2})$, take $\beta=1$ and from the space-time inequality, we obtain

$$
\begin{aligned}
|A_3|&\le C\frac{1}{s^{\frac{1}{2}}}\frac{1}{(t-s)^{1-\frac{\eta}{2}}}\int_{\mathbb{R}}\frac{|m-x'|}{(t-s)^{\frac{1}{2}}}H_0\big(cs,2m-x'-x_0\big)H_0\big(c(t-s),x-x'\big)\mathbb{I}_{\{x'\vee x_0\vee m_0<m\}}\mathbb{I}_{\{x'<m\}}\,dx'\\
&\le C\frac{1}{s^{\frac{1}{2}}}\frac{1}{(t-s)^{1-\frac{\eta}{2}}t^{\frac{1}{2}}}\int_{\mathbb{R}}\big|2m-x'-x_0\big|H_0\big(cs,2m-x'-x_0\big)\\
&\quad\times H_0\big(c(t-s),x-x'\big)\mathbb{I}_{\{x'\vee x_0\vee m_0<m\}}\mathbb{I}_{\{x'<m\}}\,dx'\\
&\le C\frac{1}{(t-s)^{1-\frac{\eta}{2}}}\frac{1}{t^{\frac{1}{2}}}H_0(ct,2m-x-x_0).
\end{aligned}
$$

For the term A_4, we are on the set $\{x<m_0\}\cap\{m_0=m\}$ so that from Lemma A.4 one has

$$
|A_4|\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\int_{\mathbb{R}^2}\left\{\frac{|m_0-x_0|^\beta}{s^{\frac{\beta}{2}}}\wedge\frac{|m_0-x'|^\beta}{s^{\frac{\beta}{2}}}\wedge 1\right\}H_0\big(cs,x'-x_0\big)
$$

$$\times \left\{ \frac{|m-x|^\beta}{(t-s)^{\frac{\beta}{2}}} \wedge \frac{|m-x'|^\beta}{(t-s)^{\frac{\beta}{2}}} \wedge 1 \right\} H_0\big(c(t-s), x-x'\big) \mathbb{I}_{\{x'<m_0\}} \mathbb{I}_{\{x<m'\}} \mathbb{I}_{\{m_0=m\}} \delta_{m_0}\big(dm'\big)\, dx'$$
$$\le \frac{C}{(t-s)^{1-\frac{\eta}{2}}} \left\{ \frac{|m-x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|m-x|^\beta}{t^{\frac{\beta}{2}}} \wedge 1 \right\} H_0(ct, x-x_0).$$

In order to obtain the last inequality we again separated the computations into the two cases: $s \in (0, t/2)$ and $s \in (t/2, t)$. For $s \in (0, t/2)$, that is, $t-s \asymp t$, we use the inequality $|m-x|^\beta (t-s)^{-\beta/2} \wedge |m-x'|^\beta (t-s)^{-\beta/2} \wedge 1 \le (t-s)^{-\beta/2}(|m-x_0|^\beta + |x_0-x'|^\beta)$ and decompose the computations into a sum of two terms accordingly. For the term $(t-s)^{-\beta/2}|m-x_0|^\beta$, we use the bound $|m_0-x_0|^\beta s^{-\beta/2} \wedge |m_0-x'|^\beta s^{-\beta/2} \wedge 1 \le 1$ which, after convolution of Gaussian kernels, yield $|A_4| \le C(t-s)^{-1+\eta/2} t^{-\beta/2} |m-x_0|^\beta H_0(ct, x-x_0)$. For the term $(t-s)^{-\beta/2}|x_0-x'|^\beta$, we use the bound $|m_0-x_0|^\beta s^{-\beta/2} \wedge |m_0-x'|^\beta s^{-\beta/2} \wedge 1 \le |m_0-x_0|^\beta s^{-\beta/2} = |m-x_0|^\beta s^{-\beta/2}$ which, after using the space-time inequality and convolution of Gaussian kernels, yield $|A_4| \le C(t-s)^{-1+\eta/2} t^{-\beta/2} |m-x_0|^\beta H_0(ct, x-x_0)$. The bound $|A_4| \le C(t-s)^{-1+\eta/2}\{|m-x|^\beta t^{-\beta/2} \wedge 1\} H_0(ct, x-x_0)$ is easily obtained. We thus obtained the desired bound for $s \in (0, t/2)$. For $s \in [t/2, t)$, that is, $s \asymp t$, we proceed similarly by using the inequality $|m-x_0|^\beta s^{-\beta/2} \wedge |m-x'|^\beta s^{-\beta/2} \wedge 1 \le s^{-\beta/2}(|m-x|^\beta + |x-x'|^\beta)$ and decompose the computations into a sum of two terms accordingly. For the term $s^{-\beta/2}|m-x|^\beta$, we use the bound $|m-x|^\beta (t-s)^{-\beta/2} \wedge |m-x'|^\beta (t-s)^{\beta/2} \wedge 1 \le 1$ together with the convolution of Gaussian kernels and for the term $s^{-\beta/2}|x-x'|^\beta$ we use the bound $|m-x|^\beta (t-s)^{-\beta/2} \wedge |m-x'|^\beta (t-s)^{\beta/2} \wedge 1 \le |m-x|^\beta (t-s)^{-\beta/2}$ together with the space-time inequality and the convolution of Gaussian kernels. We thus obtain $|A_4| \le C(t-s)^{-1+\eta/2} t^{-\beta/2} |m-x|^\beta H_0(ct, x-x')$. The bound $|A_4| \le C(t-s)^{-1+\eta/2} t^{-\beta/2}\{|m-x_0|^\beta \wedge 1\} H_0(ct, x-x')$ is obtained by using the two inequalities $|m-x|^\beta (t-s)^{-\beta/2} \wedge |m-x'|^\beta (t-s)^{\beta/2} \wedge 1 \le 1$ and $|m_0-x_0|^\beta s^{-\beta/2} \wedge |m_0-x'|^\beta s^{\beta/2} \wedge 1 \le |m_0-x_0|^\beta s^{-\beta/2} \wedge 1$. This last inequality completes the bound for A_4.

Therefore gathering the above estimates, the following Gaussian upper-bound holds

$$\left| \int_{\mathbb{R}^2} \widehat{p}_s\big(x_0, m_0, x', m'\big) \widehat{\theta}_{t-s}\big(x', m', x, m\big) u\big(x_0, m_0, x, m, dx', dm'\big) \right|$$
$$\le \begin{cases} \frac{C}{(t-s)^{1-\frac{\eta}{2}}} \frac{1}{t^{\frac{1}{2}}} H_0(ct, 2m-x-x_0), & x \le m, m_0 < m, \\ \frac{C}{(t-s)^{1-\frac{\eta}{2}}} \{\frac{|m-x|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|m-x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge 1\} H_0(ct, x-x_0), & x < m_0, m = m_0, \end{cases}$$

for any $\beta \in [0, 1]$. By an induction argument which is similar to the case of the local time, one gets

$$\big|p_T^n(x_0, m_0, x, m)\big|$$
$$\le \begin{cases} C^n (\int_{\Delta_n(T)} d\mathbf{s}_n \prod_{k=1}^n (s_{k-1}-s_k)^{-1+\frac{\eta}{2}}) \times \frac{1}{\sqrt{T}} H_0(cT, 2m-x-x_0), & x \le m, m_0 < m, \\ C^n (\int_{\Delta_n(T)} d\mathbf{s}_n \prod_{k=1}^n (s_{k-1}-s_k)^{-1+\frac{\eta}{2}}) \times H_0(cT, x-x_0), & x < m_0, m = m_0. \end{cases}$$

We omit the remaining technical details. Hence from Lemma A.1 and the asymptotic property of the Gamma function, the Gaussian upper bound (A.19) for the transition density is valid.

A.5. *Some useful technical results*

Lemma A.1. *Let $b > -1$ and $a \in [0, 1)$. Then for any $t_0 > 0$,*

$$\int_{\Delta_n(t_0)} d\mathbf{t}_n \quad t_n^b \prod_{j=0}^{n-1} (t_j - t_{j+1})^{-a} = \frac{t_0^{b+n(1-a)} \Gamma^n(1-a) \Gamma(1+b)}{\Gamma(1+b+n(1-a))}.$$

Proof. Using the change of variables $s = ut$, one has

$$\int_0^t s^b (t-s)^{-a}\, ds = t^{b+1-a} \int_0^1 u^b (1-u)^{-a}\, du = t^{b+1-a} B(1+b, 1-a)$$

where $(x, y) \mapsto B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt$ stands for the standard Beta function. Using this equality repeatedly, we obtain the statement. □

Lemma A.2. *Let $c > 0$. For any $(x, x_0) \in \mathbb{R}^2$, $0 \le \ell_0 \le \ell$ and $0 < s < t$, one has*

$$\int_{\mathbb{R}} H_0\big(cs, |x'| + |x_0| + \ell - \ell_0\big) H_0\big(c(t-s), x - x'\big) \mathbb{I}_{\{xx' \ge 0\}}\, dx' \le H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big),$$

$$\int_{\mathbb{R}} H_0\big(c(t-s), |x'| + |x| + \ell - \ell_0\big) H_0\big(cs, x_0 - x'\big) \mathbb{I}_{\{x_0 x' \ge 0\}}\, dx' \le H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big).$$

Proof. Using the fact that $\{x'x \ge 0\} = \{x' \ge 0\} \cap \{x \ge 0\} \cup \{x' \le 0\} \cap \{x \le 0\}$ and positivity of the integrand

$$\begin{aligned}
&\int_{\mathbb{R}} H_0\big(cs, |x'| + |x_0| + \ell - \ell_0\big) H_0\big(c(t-s), x - x'\big) \mathbb{I}_{\{xx' \ge 0\}}\, dx' \\
&\quad \le \mathbb{I}_{\{x \ge 0\}} \int_{\mathbb{R}} H_0\big(cs, x' + |x_0| + \ell - \ell_0\big) H_0\big(c(t-s), x - x'\big) \mathbb{I}_{\{x' \ge 0\}}\, dx' \\
&\qquad + \mathbb{I}_{\{x \le 0\}} \int_{\mathbb{R}} H_0\big(cs, -x' + |x_0| + \ell - \ell_0\big) H_0\big(c(t-s), x - x'\big) \mathbb{I}_{\{x' \le 0\}}\, dx' \\
&\quad \le \mathbb{I}_{\{x \ge 0\}} H_0\big(ct, x + |x_0| + \ell - \ell_0\big) + \mathbb{I}_{\{x \le 0\}} H_0\big(ct, |x_0| + \ell - \ell_0 - x\big) \\
&\quad = H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)
\end{aligned}$$

where the last inequality follows from convolution properties of the Gaussian density. The proof of the second inequality follows from similar arguments and is therefore omitted. □

Lemma A.3. *Let $c_1 > 0$. For any $(x, x_0) \in \mathbb{R}^2$, $0 \le \ell_0 \le \ell_2$ and $0 < s < t$, one has*

$$\begin{aligned}
&\int_{\mathbb{R} \times (\ell_0, \ell)} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c_1(t-s), |x| + |x'| + \ell - \ell'\big) \frac{1}{s^{\frac{1}{2}}} H_0\big(c_1 s, |x'| + |x_0| + \ell' - \ell_0\big)\, dx'\, d\ell' \\
&\quad \le \frac{C}{t^{\frac{1}{2}}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)
\end{aligned}$$

for some positive constants C, c independent of t, x_0, ℓ_0 and ℓ. Similarly, for any $(x_0, m_0) \in \mathbb{R}^2$, $x_0 \le m_0$, $m \ge m_0$ and $0 < s < t$, one has

$$\begin{aligned}
&\int_{(m_0, m) \times (-\infty, m')} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c_1(t-s), 2m - x - x'\big) \frac{1}{s^{\frac{1}{2}}} H_0\big(c_1 s, 2m' - x' - x_0\big)\, dx'\, dm' \\
&\quad \le \frac{C}{t^{\frac{1}{2}}} H_0(ct, 2m - x - x_0)
\end{aligned}$$

for some positive constants C, c independent of t, x_0, m_0 and m.

Proof. We will only prove the first bound. The second one follows from similar arguments. For simplicity, we write

$$C_2 := \int_{\mathbb{R} \times (\ell_0, \ell)} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + |x'| + \ell - \ell'\big) \frac{1}{s^{\frac{1}{2}}} H_0\big(cs, |x'| + |x_0| + \ell' - \ell_0\big)\, dx' d\ell'.$$

Let us assume that $|x| + |x_0| + \ell - \ell_0 \le t^{\frac{1}{2}}$. We use the fact that the diagonal estimate is global. For $s \in [\frac{t}{2}, t]$, one has $s \asymp t$ so that

$$\frac{1}{s^{\frac{1}{2}}} H_0\big(cs, |x'| + |x_0| + \ell' - \ell_0\big) \le \frac{C}{t} \le C \frac{1}{\sqrt{t}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)$$

which in turn implies:

$$\begin{aligned}
|C_2| &\le \frac{C}{\sqrt{t}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big) \int_{\mathbb{R} \times (\ell_0, \ell)} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + |x'| + \ell - \ell'\big)\, dx'\, d\ell' \\
&\le \frac{C}{\sqrt{t}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big).
\end{aligned}$$

Similarly, for $s \in [0, \frac{t}{2}]$, one has

$$|C_2| \leq \frac{C}{\sqrt{t}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big).$$

Hence, the claim follows in the diagonal regime $|x| + |x_0| + \ell - \ell_0 \leq t^{\frac{1}{2}}$. We now consider the off-diagonal regime $|x| + |x_0| + \ell - \ell_0 > t^{\frac{1}{2}}$. We write $\mathbb{R} \times (\ell_0, \ell) = D_1 \cup D_2$ where

$$D_1 := \big\{\big(x', \ell'\big) \in \mathbb{R} \times (\ell_0, \ell) : \big|x'\big| + |x_0| + \ell' - \ell_0 \leq |x| - \big|x'\big| + \ell - \ell'\big\},$$
$$D_2 := \big\{\big(x', \ell'\big) \in \mathbb{R} \times (\ell_0, \ell) : \big|x'\big| + |x_0| + \ell' - \ell_0 > |x| - \big|x'\big| + \ell - \ell'\big\}.$$

On the set D_1, we remark $|x| - |x'| + \ell - \ell' \asymp |x| + |x_0| + \ell - \ell_0$, so that

$$\begin{aligned}\frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + \big|x'\big| + \ell - \ell'\big) &\leq \frac{C}{t} \frac{(|x| + |x_0| + \ell - \ell_0)^2}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + |x_0| + \ell - \ell_0\big) \\ &\leq \frac{C}{t^{\frac{1}{2}}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)\end{aligned}$$

where we used the space-time inequality for the last inequality. Hence, one gets

$$\begin{aligned}&\int_{D_1} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + \big|x'\big| + \ell - \ell'\big) \frac{1}{s^{\frac{1}{2}}} H_0\big(cs, \big|x'\big| + |x_0| + \ell' - \ell_0\big)\, dx'\, d\ell' \\ &\quad \leq \frac{C}{t^{\frac{1}{2}}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big).\end{aligned}$$

On the set D_2, one has $|x| - |x'| + \ell - \ell' \asymp |x| + |x_0| + \ell - \ell_0$, so that

$$\begin{aligned}\frac{1}{s^{\frac{1}{2}}} H_0\big(cs, \big|x'\big| + |x_0| + \ell' - \ell_0\big) &\leq \frac{C}{t} \frac{(|x| + |x_0| + \ell - \ell_0)^2}{s^{\frac{1}{2}}} H_0\big(cs, |x| + |x_0| + \ell - \ell_0\big) \\ &\leq \frac{C}{t^{\frac{1}{2}}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)\end{aligned}$$

which in turn yields

$$\begin{aligned}&\int_{D_2} \frac{1}{(t-s)^{\frac{1}{2}}} H_0\big(c(t-s), |x| + \big|x'\big| + \ell - \ell'\big) \frac{1}{s^{\frac{1}{2}}} H_0\big(cs, \big|x'\big| + |x_0| + \ell' - \ell_0\big)\, dx'\, d\ell' \\ &\quad \leq \frac{C}{t^{\frac{1}{2}}} H_0\big(ct, |x| + |x_0| + \ell - \ell_0\big)\end{aligned}$$

and proves the claim. □

Lemma A.4. Local time: *Define* $\bar{f}_t(x_0, x) := H_0(at, x - x_0) - H_0(at, x + x_0)$ *for some given constant* $a > 0$. *For any* $\beta \in [0, 1]$, *there exists* $C,\ c > 1$, *such that for any* $(x_0, x) \in \mathbb{R}^2$ *satisfying* $xx_0 \geq 0$ *and* $r = 0, 2$, *the following estimates hold*:

$$\big|\partial_{x_0}^r \bar{f}_t(x_0, x)\big| \leq \frac{C}{t^{\frac{r}{2}}} \left\{\frac{|x|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge 1\right\} H_0(ct, x - x_0).$$

Running maximum: *Define* $\bar{f}_t(x_0, x) := H_0(at, x - x_0) - H_0(at, 2m_0 - x - x_0)$ *for some given constant* $a > 0$. *For any* $\beta \in [0, 1]$, *there exists* $C,\ c > 1$, *such that for any* $(x, x_0, m_0) \in \mathbb{R}^3$ *satisfying* $x, x_0 \leq m_0$ *and for any* $r = 0, 1, 2$, *one has*

$$\big|\partial_{x_0}^r \bar{f}_t(x_0, x)\big| \leq \frac{C}{t^{\frac{r}{2}}} \left\{\frac{|m_0 - x_0|^\beta}{t^{\frac{\beta}{2}}} \wedge \frac{|m_0 - x|^\beta}{t^{\frac{\beta}{2}}} \wedge 1\right\} H_0(ct, x - x_0), \quad x_0, x \leq m_0.$$

Proof. From the expression of $\bar{f}_t(x_0, x)$, the following estimates for $\bar{f}_t(x_0, x)$ and its derivatives hold

$$\left|\partial_{x_0}^r \bar{f}_t(x_0, x)\right| \le \frac{C}{t^{\frac{r}{2}}}\big(H_0(ct, x - x_0) + H_0(ct, x + x_0)\big), \quad r = 0, 2.$$

Furthermore, for $(x_0, x) \in \mathbb{R}^2$ such that $xx_0 \ge 0$, one has $H_0(ct, x + x_0) \le H_0(ct, x - x_0)$, since in the exponent

$$(x - x_0)^2 + 4x_0(x - x_0) + 4x_0^2 \ge (x - x_0)^2.$$

Hence, we deduce that $|\partial_{x_0}^r \bar{f}_t(x_0, x)| \le Ct^{-\frac{r}{2}} H_0(ct, x - x_0)$. To derive the bounds with the $|x|^\beta$ or $|x_0|^\beta$ terms, we first consider the case where $|x|^2 \le t$, to estimate $\bar{f}_t(x_0, x) = H_0(a(y, \ell)t, x - x_0) - H_0(a(y, \ell)t, x + x_0)$ one applies the mean value theorem to $H_0(a(y, \ell)t, x - x_0)$ with respect to the points x and $-x$ to obtain for some $\theta \in [0, 1]$,

$$\left|\bar{f}_t(x_0, x)\right| = \left|2x\partial_x H_0\big(a(y, \ell)t, x_0 - \theta x + (1 - \theta)x\big)\right| \le C\frac{|x|^\beta}{t^{\frac{\beta}{2}}} H_0(ct, x - x_0)$$

where in the second line we have used the space-time inequality and the fact that $|x|^{1-\beta} \le t^{\frac{1-\beta}{2}}$. For the case that $|x|^2 \ge t$, one directly gets

$$\left|\bar{f}_t(x_0, x)\right| \le C\frac{|x|^\beta}{t^{\frac{\beta}{2}}} H_0(ct, x - x_0).$$

The proof for the second derivatives of $\bar{f}_t(x_0, x)$ as well as the estimates with the $|x_0|^\beta$ term and the estimates for the running maximum case follow similar arguments and details are omitted. □

Acknowledgement

The authors wish to thank Professor Arturo Kohatsu-Higa for his careful readings and valuable comments during the writing of this paper.

References

[1] T. Appuhamillage, V. Bokil, E. Thomann, E. Waymire and B. Wood. Occupation and local times for skew Brownian motion with applications to dispersion across an interface. *Ann. Appl. Probab.* **21** (1) (2011) 183–214. MR2759199 https://doi.org/10.1214/10-AAP691

[2] E. Barucci, S. Polidoro and V. Vespri. Some results on partial differential equations and Asian options. *Math. Models Methods Appl. Sci.* **11** (3) (2001) 475–497. MR1830951 https://doi.org/10.1142/S0218202501000945

[3] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, Berlin, 1997. MR1483890

[4] R. F. Bass and E. A. Perkins. A new technique for proving uniqueness for martingale problems. In *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut* 47–53, 2009.

[5] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259** (6) (2010) 1577–1630. MR2659772 https://doi.org/10.1016/j.jfa.2010.05.002

[6] K. D. Elworthy, A. Truman and H. Zhao. Generalized Itô formulae and space-time Lebesgue–Stieltjes integrals of local times. In *Séminaire de Probabilités XL* 117–136. *Lecture Notes in Math.* **1899**. Springer, Berlin, 2007. MR2409002 https://doi.org/10.1007/978-3-540-71189-6_5

[7] M. Forde. A diffusion-type process with a given joint law for the terminal level and supremum at an independent exponential time. *Stochastic Process. Appl.* **121** (12) (2011) 2802–2817. MR2844541 https://doi.org/10.1016/j.spa.2011.07.009

[8] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, New York, 1964. MR0181836

[9] N. Frikha, A. Kohatsu-Higa and L. Li. Integration by parts formula for killed processes: A point of view of approximation theory. Preprint. 2018. *Electron. J. Probab.* To appear.

[10] M. G. Garroni and J.-L. Menaldi. *Green Functions for Second Order Parabolic Integro-Differential Problems*. *Pitman Research Notes in Mathematics Series* **275**. Longman Scientific & Technical, Harlow, 1992. (Copublished in the United States with, John Wiley & Sons, Inc., New York.) MR1202037

[11] I. Karatzas and S. E. Shreve. Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. *Ann. Probab.* **12** (3) (1984) 819–828. MR0744236

[12] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, 2nd edition. Springer Verlag, Berlin, 1991. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

[13] H. Kestelman. *Modern Theories of Integration*, 2nd revised edition. Dover Publications, Inc., New York, 1960. MR0122951

[14] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16** (2011) 234–250. MR2802040 https://doi.org/10.1214/ECP.v16-1619

[15] G. Peskir. A change-of-variable formula with local time on curves. *J. Theoret. Probab.* **18** (13) (2005) 499–535. MR2167640 https://doi.org/10.1007/s10959-005-3517-6

[16] C. Soize. *The Fokker–Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions*. *Series on Advances in Mathematics for Applied Sciences.* **17**. World Scientific Publishing Co., Inc., River Edge, NJ, 1994. MR1287386 https://doi.org/10.1142/9789814354110

[17] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, Berlin Heidelberg New-York, 1979. MR0532498

[18] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** (2) (2002) 163–198. MR1924934