The Annals of Applied Probability
2015, Vol. 25, No. 6, 3095–3138
DOI: 10.1214/14-AAP1068

A PROBABILISTIC INTERPRETATION OF THE PARAMETRIX METHOD[1]

BY VLAD BALLY AND ARTURO KOHATSU-HIGA

Université Paris-Est Marne-la-Vallée and INRIA, and Ritsumeikan University and Japan Science and Technology Agency

In this article, we introduce the parametrix technique in order to construct fundamental solutions as a general method based on semigroups and their generators. This leads to a probabilistic interpretation of the parametrix method that is amenable to Monte Carlo simulation. We consider the explicit examples of continuous diffusions and jump driven stochastic differential equations with Hölder continuous coefficients.

1. Introduction. The parametrix technique for solving parabolic partial differential equations (PDEs) is a classical method in order to expand the fundamental solution of such an equation in terms of a basic function known as the parametrix. This is the parallel of the Taylor expansion of a smooth function in terms of polynomials.

The concept of order of the polynomial in the classical Taylor expansion is replaced by multiple integrals whose order increases as the expansion becomes more accurate. This method has been successfully applied to many equations and various situations. Its success is due to its flexibility as it can be applied to a wide variety of PDEs. It has been successfully extended to other situations for theoretical goals (see, e.g., [9–12] and [14]). In [6], the authors consider the parametrix as an analytical method for approximations for continuous diffusions. These analytical approximations may be used as deterministic approximations and are highly accurate in the cases where the sum converges rapidly. In general, higher order integrals are difficult to compute and, therefore, this becomes a limitation of the method.

The goal of the present paper is to introduce a general probabilistic interpretation of the parametrix method based on semigroups, which not only reexpresses the arguments of the method in probabilistic terms, but also to introduce an alternative method of simulation with no approximation error. This leads to the natural emergence of the difference between the generators of the process and its parametrix in the same manner as the concept of derivative appears in the classical Taylor expansion.

Received June 2013; revised June 2014.

[1]Supported by grants from the Japanese government.

MSC2010 subject classifications. Primary 35K10, 35K15, 65C20; secondary 65C05, 65C30.

Key words and phrases. Parametrix, stochastic differential equations, density, Monte Carlo methods.

Let us explain the above statement in detail. The first step in the Monte Carlo approach for approximating the solution of the parabolic partial differential equation $\partial_t u = Lu$ is to construct the Euler scheme which approximates the continuous diffusion process with infinitesimal operator L. To fix the ideas, consider the diffusion process $X_t \equiv X_t(x)$ solution of the following stochastic differential equation (SDE):

$$dX_t = \sum_{j=1}^{m} \sigma_j(X_t)\, dW_t^j + b(X_t)\, dt, \qquad t \in [0, T], X_0 = x_0, \tag{1.1}$$

where W is a multidimensional Brownian motion and $\sigma_j, b : \mathbb{R}^d \to \mathbb{R}^d$ are smooth functions. We denote by $P_t f(x) = \mathbb{E}[f(X_t(x))]$ the semigroup associated to this diffusion process. The infinitesimal generator associated to P is defined for $f \in C_c^2(\mathbb{R}^d)$ as

$$Lf(x) = \frac{1}{2} \sum_{i,j} a^{i,j}(x) \partial_{i,j}^2 f(x) + b^i(x) \partial_i f(x), \qquad a := \sigma\sigma^*. \tag{1.2}$$

By the Feynman–Kac formula, one knows that $u(t, x) := P_t f(x)$ is the unique solution to $\partial_t u = Lu$ satisfying the initial condition $u(0, x) = f(x)$. Therefore, the goal is to approximate X first and then the expectation in $P_t f(x) = \mathbb{E}[f(X_t(x))]$ using the law of large numbers which leads to the Monte Carlo method.

Now, we describe some stochastic approximation methods for X. Given a partition of $[0, T]$, $\pi = \{0 = t_0 < \cdots < t_n = T\}$, the Euler scheme associated to this time grid is defined as $X_0^\pi(x) = x$

$$\begin{aligned} X_{t_{k+1}}^\pi(x) &= X_{t_k}^\pi(x) + \sum_{j=1}^{m} \sigma_j\big(X_{t_k}^\pi(x)\big)\big(W_{t_{k+1}}^j - W_{t_k}^j\big) \\ &\quad + b\big(X_{t_k}^\pi(x)\big)(t_{k+1} - t_k). \end{aligned} \tag{1.3}$$

It is well known (see [17]) that $X^\pi \equiv X^\pi(x)$ is an approximation scheme of X of order one. That is, there exists a constant $C_f(x)$ such that

$$\begin{aligned} &\big|\mathbb{E}\big[f\big(X_T(x)\big)\big] - \mathbb{E}\big[f\big(X_T^\pi(x)\big)\big]\big| \\ &\qquad \le C_f(x) \max\{t_{i+1} - t_i; i = 0, \ldots, n-1\} \end{aligned} \tag{1.4}$$

for f measurable and bounded (see [2, 3]) and under strong regularity assumptions on the coefficients σ_j and b.

Roughly speaking, the parametrix method is a deterministic method with the following intuitive background: in short time the diffusion $X_t(x_0)$ is close to the diffusion with coefficients "frozen" in the starting point x_0. So one may replace the operator L by the operator L^{x_0} defined as

$$L^{x_0} f(x) = \frac{1}{2} \sum_{i,j} a^{i,j}(x_0) \partial_{i,j}^2 f(x) + b^i(x_0) \partial_i f(x)$$

and one may replace the semigroup P_t by the semigoup $P_t^{x_0}$ associated to L^{x_0}. Clearly, this is the same idea as the one which leads to the construction of the Euler scheme (1.3). In fact, notice that the generator of the one step (i.e., $\pi = \{0, T\}$) Euler scheme $X^\pi(x_0)$ is given by L^{x_0}.

The goal of the present article is to give a probabilistic representation formula based on the parametrix method. This formula will lead to simulation procedures for $\mathbb{E}[f(X_T)]$ with no approximation error which are based on the weighted sample average of Euler schemes with random partition points given by the jump times of an independent Poisson process. In fact, the first probabilistic representation formula (forward formula) we intend to prove is the following:

$$\mathbb{E}[f(X_T)] = e^T \mathbb{E}\left[f(X_T^\pi) \prod_{k=0}^{J_T-1} \theta_{\tau_{k+1}-\tau_k}(X_{\tau_k}^\pi, X_{\tau_{k+1}}^\pi)\right]. \tag{1.5}$$

Here, $\tau_0 = 0$ and $\pi := \{\tau_k, k \in \mathbb{N}\}$, are the jump times of a Poisson process $\{J_t; t \in [0, T]\}$ of parameter one and X^π is a continuous time Markov process which satisfies

$$P\big(X_{\tau_{k+1}}^\pi \in dy | \{\tau_k, k \in \mathbb{N}\}, X_{\tau_k}^\pi = x_0\big) = P_{\tau_{k+1}-\tau_k}^{x_0}(x_0, dy) = p_{\tau_{k+1}-\tau_k}^{x_0}(x_0, y)\, dy.$$

In the particular case discussed above, then X^π corresponds in fact to an Euler scheme with random partition points. $\theta_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a weight function to be described later.

Before discussing the nature of the above probabilistic representation formula, let us remark that a similar formula is available for one-dimensional diffusion processes (see [5]) which is strongly based on explicit formulas that one may obtain using the Lamperti formula. Although many elements may be common between these two formulations, the one presented here is different in nature.

In order to motivate the above formula (1.5), let us give the following basic heuristic argument that leads to the forward parametrix method:

$$\begin{aligned} P_t f(x) - P_t^x f(x) &= \int_0^t \partial_s \big(P_{t-s}^x P_s f\big)(x)\, dt \\ &= \int_0^t P_{t-s}^x \big(L - L^x\big) P_s f(x)\, ds. \end{aligned} \tag{1.6}$$

Here, we suppose that $P_s f \in \mathrm{Dom}(L - L^x)$. The above expression is already an equivalent of a Taylor expansion of order one where the notion of first-order derivative is being replaced by $L - L^x$. Its iteration will lead to higher order Taylor expansions. Another way of looking at this is to consider (1.6) as a Volterra equation in $Pf(x)$. This will be our point of view here.

In fact, if t is considered as the time variable and considering (1.6) as an equation, one sees, after an application of the integration by parts on the diffferential operator $L - L^x$, that $\{P_t f; t \in [0, T]\}$ for $f \in C_c^\infty$ can also be considered as a solution of the following Volterra type linear equation:

$$
\begin{aligned}
P_t f(x) &= P_t^x f(x) + \int_0^t \int p_{t-s}^x(x, y_1)\big(L - L^x\big) P_s f(y_1)\, dy_1\, ds \\
&= P_t^x f(x) + \int_0^t \int \theta_{t-s}(x, y_1) p_{t-s}^x(x, y_1) P_s f(y_1)\, dy_1\, ds.
\end{aligned}
\tag{1.7}
$$

Here, we have used the function θ, defined as

$$
\begin{aligned}
&\theta_{t-s}(x, y_1) p_{t-s}^z(x, y_1)|_{z=x} \\
&\quad = \big(L - L^z\big)^* p_{t-s}^z(x, \cdot)(y_1)|_{z=x} \\
&\quad = \frac{1}{2} \sum_{i,j} \partial_{i,j}\big(\big(a^{i,j}(\cdot) - a^{i,j}(z)\big) p_{t-s}^z(x, \cdot)\big)(y_1) \\
&\qquad - \sum_i \partial_i\big(\big(b^i(\cdot) - b^i(z)\big) p_{t-s}^z(x, \cdot)\big)(y_1)\Big|_{z=x}.
\end{aligned}
\tag{1.8}
$$

Equation (1.7) can be iterated in $P_t f$ in order to obtain a series expansion of the solution which is the equivalent of the Taylor expansion of $P_t f$.

We may note that the second term in this expansion for $t = T$ can be rewritten using the Euler scheme with partition points $\pi = \{0, T - s, T\}$ as

$$
\begin{aligned}
&\int_0^T \int \theta_{T-s}(x, y_1) p_{T-s}^x(x, \cdot)(y_1) P_s^{y_1} f(y_1)\, dy_1\, ds \\
&\quad = \int_0^T \mathbb{E}\big[f\big(X_T^\pi(x)\big)\theta_{T-s}\big(x, X_{T-s}^\pi(x)\big)\big]\, ds.
\end{aligned}
\tag{1.9}
$$

This is the first step toward the construction of what we call the forward parametrix method.[2] It requires the regularity of the coefficients and it is based on the usual Euler scheme for the sde (1.1). Note that (1.9) will be associated with the term $J_T = 1$ in (1.5). In fact, if there is only one jump of the Poisson process J in the interval $[0, T]$, then the distribution of the jump is uniform in the interval $[0, T]$. This leads to the probabilistic interpretation of the time integral in (1.9).

Let us now discuss an alternative to the above method which requires less regularity conditions on the coefficients of (1.1). This method will be called the backward parametrix method and it is obtained by duality arguments as follows. That is, consider for two functions $f, g \in C_c^\infty(\mathbb{R}^d)$ the pairing $\langle f, P_t^* g\rangle$. Then we will use

[2]This also explains the logic behind the choice of variables in the integrals. Through the rest of the paper y_i will denote the integrating variables in the order given by the corresponding Euler scheme. Similar rule will apply with the times t_i, $i \in \mathbb{N}$. For example, we will have that $t_1 = t - s_1$ in (1.6).

the approximating semigroup $\hat{Q}_t g(y) := (P_t^y)^* g(y) = \int g(x) p_t^y(x, y)\, dx$. A similar heuristic argument gives

$$
\begin{aligned}
P_t^* g(y) - (P_t^y)^* g(y) &= \int_0^t \int (L - L^y) p_{t-s}^y(\cdot, y)(y_1) P_s^* g(y_1)\, dy_1\, ds \\
&= \int_0^t \int \hat{\theta}_{t-s}(y_1, y) p_{t-s}^y(y_1, y) P_s^* g(y_1)\, dy_1\, ds.
\end{aligned} \tag{1.10}
$$

Note that in this case the operator $L - L^y$ is applied to the density function $p_{t-s}^y(\cdot, y)$ with the coefficients frozen at y, therefore, no derivative of the coefficients is needed in this approach. In fact,

$$
\hat{\theta}_{t-s}(y_1, y) p_{t-s}^z(y_1, y)|_{z=y} \tag{1.11}
$$

$$
\begin{aligned}
&= (L - L^z) p_{t-s}^z(\cdot, y)(y_1)|_{z=y} \\
&= \left\{ \frac{1}{2} \sum_{i,j} (a^{i,j}(y_1) - a^{i,j}(z)) \partial_{i,j} \right. \\
&\qquad \left. + \sum_i (b^i(y_1) - b^i(z)) \partial_i \right\} p_t^z(\cdot, y)(y_1) \bigg|_{z=y}.
\end{aligned} \tag{1.12}
$$

As before, we can obtain a probabilistic representation. In this case, one has to be careful with the time direction. In fact, due to the symmetry of the density function $p_{t-s}^y(y_1, y)$ one interprets it as the density of the Euler scheme at y_1 started at y. Therefore, the sign of the drift has to be changed leading to what we call the backward parametrix method. In the particular case that f is a density function, it will be interpreted as a "backward running" Euler scheme from T to 0 with random initial point with density f. The test function g is replaced by a Dirac delta at the initial point of the diffusion x_0. See Section 6 for precise statements.

Therefore, the behavior of forward and backward methods are different. In fact, the forward method applies when the coefficients are regular. In many applied situations, one may have coefficients which are just Hölder continuous and, therefore, the forward method does not apply. In that case, one may apply the backward method which demands less regularity. For this reason, the treatment of the forward method and the backward method are essentially different and they are treated separately. Issues related to simulation will be discussed in another article.

Our article is structured as follows: In Section 2, we give the notation used throughout the paper. In Section 3, we discuss the existence, uniqueness and regularity properties of the solution of the linear Volterra equations of the type (1.7) or (1.10) which will be applicable to both probabilistic representation formulas to be discussed later. In Section 4, we provide a general abstract framework based on semigroups for which our two approaches (forward and backward) can be applied. The main hypotheses applying to both methods are given in this section. In Section 5, we give the analytical form of the forward method. In Section 5.1, we give the probabilistic representation, and in Section 5.2, we give the continuity and differentiability properties of the density functions. This is the usual application of the parametrix method.

In Section 6, we give the backward approach which was first introduced in [13]. We also give the probabilistic interpretation and the regularity results corresponding to the backward method in parallel sections.

In Section 7, we consider our main examples. The first corresponds to the continuous diffusion with uniformly elliptic diffusion coefficient. We see in Section 7.1 that in the forward approach we need the coefficients to be smooth. While in Section 7.2, we show that in order for the backward approach to be applicable, we only require the coefficients to be Hölder continuous. In Section 7.3, we also consider the case of a jump driven SDE where the Lévy measure is of stable type in a neighborhood of zero. This example is given with two probabilistic interpretations.

We close with some conclusions: an Appendix and the References section.

2. Some notation and general definitions. We now give some basic notation and definitions used through this article. For a sequence of operators S_i, $i = 1, \dots, n$ which do not necessarily commute, we define $\prod_{i=1}^{n} S_i = S_1 \cdots S_n$ and $\prod_{i=n}^{1} S_i = S_n \cdots S_1$. We will denote by I, the identity matrix or identity operator and S^* will denote the adjoint operator of S. $\mathrm{Dom}(S)$ denotes the domain of the operator S. If the operator S is of integral type, we will denote its associated measure $S(x, dy)$ so that $Sf(x) = \int f(y)S(x, dy)$. All space integrals will be taken over $\mathbb{R}^d$. For this reason, we do not write the region of integration which we suppose clearly understood. Also in order to avoid long statements, we may refrain from writing often where the time and space variables take values supposing that they are well understood from the context.

In general, indexed products where the upper limit is negative are defined as 1 or I. In a similar fashion, indexed sums where the upper limit is negative are defined as zero.

As it is usual, $A \le B$ for two matrices A and B, denote the fact that $A - B$ is positive definite. Components of vectors or matrices are denoted by superscript letters. When the context makes it clear we denote by $\partial_i f$ the partial derivative operator with respect to the ith variable of the function f and similarly for higher order derivatives. For example, derivatives with respect to a multi-index β, of length $|\beta|$, are denoted by $\partial_\beta f$. Time derivatives will be denoted by ∂_t.

We denote by $\delta_a(dx)$ the point mass measure concentrated in $\{a\}$, $B(x, r)$ denotes the ball of center $x \in \mathbb{R}^d$ and radius $r > 0$, $[x]$ denotes the ceiling or smallest integer function for $x \in \mathbb{R}$ and $\mathbb{R}_+ \equiv (0, \infty)$. The indicator function of the set A is denoted by $1_A(x)$, $C(A)$ denotes the space of real valued functions continuous in the set A. The space of real valued measurable bounded functions defined on A is denoted by $L^\infty(A)$. Similarly, the space of continuous bounded functions in A is denoted by $C_b(A)$. The space of real valued infinitely differentiable functions with compact support defined on $\mathbb{R}^d$ is denoted by $C_c^\infty(\mathbb{R}^d)$. The space of $\mathbb{R}^l$-valued bounded functions defined on $\mathbb{R}^d$ with bounded derivatives up to order k is denoted by $C_b^k(\mathbb{R}^d;\mathbb{R}^l)$. The norm in this space is defined as $\|f\|_{k,\infty} = \max_{|\beta|\le k}\sup_{x\in\mathbb{R}^d}|\partial_\beta f(x)|$. In the particular case that $k=0$, we also use the simplified notation $\|f\|_\infty \equiv \|f\|_{0,\infty}$.

The multidimensional Gaussian density at $y\in\mathbb{R}^d$ with mean zero and covariance matrix given by the positive definite matrix a is denoted by $q_a(y)$. Sometimes we abuse the notation denoting by $q_t(y)$, for $y\in\mathbb{R}^d$, $t>0$ the Gaussian density corresponding to the variance–covariance matrix tI. Similarly, $H_a^i(y)$ and $H_a^{i,j}(y)$ for $i,j\in\{1,\dots,d\}$, denote the multidimensional version of the Hermite polynomials of order one and two. Exact definitions and some of the properties of Gaussian densities used throughout the article are given in Section A.2.

Constants will be denoted by C or c, we will not give the explicit dependence on parameters of the problem unless it is needed in the discussion. As it is usual, constants may change from one line to the next although the same symbol may be used.

In the notation throughout the article, we try to denote by x the starting point of the diffusion and y the arrival point with z being the parameter value where the operator L^z is frozen at. In the forward method, z will be the starting point x and in the backward method z will be the arrival point y. Due to the iteration procedure, many intermediate points will appear which will be generally denoted by y_i, $i=0,\dots,n$, always going from $y_0=x$ toward $y_n=y$ in the forward method and from $y_0=y$ to $y_n=x$ in the backward method. As stated previously, the time variables will be evolving forward in the sense of the Euler scheme if they are denoted by t_i, $i=0,\dots,n$ from $t_0=0$ to $t_n=t$ or backwardly if denoted by s_i, $i=0,\dots,n$ from $s_0=t$ to $s_n=0$.

3. A functional linear equation. In this section, we consider a functional equation of Volterra type which will include both equations (1.7) and (1.10). Therefore, this represents and abstract framework which includes the forward and backward method.

We consider a jointly measurable functions $a:(0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ and we define the operator

$$U_a f(t,x) = \int_0^t \int_{\mathbb{R}^d} f(s,y)a_{t-s}(x,y)\,dy\,ds.$$

Our aim is to solve the equation

$$f = g + U_a f \tag{3.1}$$

and to study the regularity of the solution. Formally, the unique solution is obtained by iteration and given by $H_a g := \sum_{n=0}^\infty U_a^n g$. In order to make this calculation mathematically sound, we have to study the convergence of the series. For this, we consider the iterations of the operator U_a. We define U_a^0 to be the identity operator, $U_a^1 = U_a$ and we define by recurrence $U_a^n = U_a^{n-1} U_a$.

LEMMA 3.1. *If* $I_a(t,x) := \int_{\mathbb{R}^d} |a_t(x,y)|\,dy \in L^\infty([0,T] \times \mathbb{R}^d)$ *and* $g \in L^\infty([0,T] \times \mathbb{R}^d)$, *then the equation* (3.1) *has a unique solution in the space* $\mathcal{A} := \{f \in L^\infty([0,T] \times \mathbb{R}^d); \lim_{N\to\infty} \|U_a^N f\|_\infty = 0\}$.

PROOF. In fact,

$$\|U_a g(t,\cdot)\|_\infty \le \int_0^t \|g(s,\cdot)\|_\infty \|I_a(t-s,\cdot)\|_\infty\, ds.$$

Then by induction it follows that

$$\|U_a^n g\|_\infty \le \frac{T^n \|I_a\|_\infty^n}{n!} \|g\|_\infty.$$

This means that the infinite sum $H_a g := \sum_{n=0}^\infty U_a^n g$ converges absolutely and, therefore, is well defined in $L^\infty([0,T] \times \mathbb{R}^d)$. Furthermore, it is easy to see that the sum is a solution of equation (3.1) satisfying $H_a g \in \mathcal{A}$.

For any solution f of (3.1), one obtains by iteration that

$$f = \sum_{n=0}^N U_a^n g + U_a^N f.$$

Therefore, if $f \in \mathcal{A}$ then f satisfies $f = H_a g$. From here, one obtains the uniqueness. $\square$

Unfortunately, in our case $\int |a_t(x,y)|\,dy$ blows up as $t \to 0$ and we center our discussion on this matter. We see from (1.8) and (1.11) that the rate of divergence is determined by the regularity of the coefficients. We will call this regularity index ρ in what follows. In order to introduce our main assumption, we define for a function $\beta : (0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$, the class of functions Γ_β such that there exists a positive constant C which satisfies the following inequality for every $n \in \mathbb{N}$, $y_0, y_{n+1} \in \mathbb{R}^d$ and every $\delta_i > 0$, $i = 1,\dots,n$ with $s(\delta) := \sum_{i=1}^n \delta_i$,

$$\int dy_1 \cdots \int dy_n \prod_{i=0}^n \gamma_{\delta_i}(y_i, y_{i+1}) \le C^{n+1} \beta_{s(\delta)}(y_0, y_{n+1}). \tag{3.2}$$

HYPOTHESIS 3.2. *There exists a positive constant* C *and a function* $\gamma \in \Gamma_\beta$ *such that* $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int |\gamma_t(x,y)|\,dy < \infty$ *and* $\rho \in [0,1)$ *such that for every* $(t,x,y) \in (0,T] \times \mathbb{R}^d \times \mathbb{R}^d$

$$|a_t(x,y)| \le \frac{C}{t^\rho} \gamma_t(x,y). \tag{3.3}$$

Furthermore, there exists a function β *such that* $\gamma \in \Gamma_\beta$.

If we define $C_\beta(t,x) := \int_0^t ds \int dy\, \beta_{t-s}(x,y)$ for $(t,x) \in [0,T] \times \mathbb{R}^d$, we will also assume that

HYPOTHESIS 3.3. *C_β is a bounded function.*

We denote[3] $D_T = \{((t,x),(s,y)) : 0 < s \le t \le T, x, y \in \mathbb{R}^d\}$. To the function a, we associate the function $A : D_T \to \mathbb{R}$ defined by

$$A\big((t,x),(s,y)\big) = a_{t-s}(x,y).$$

Then we define the operator $U_a : L^\infty([0,T] \times \mathbb{R}^d) \to L^\infty([0,T] \times \mathbb{R}^d)$ by

$$U_a f(t,x) = \int_0^t ds \int dy\, f(s,y) A\big((t,x),(s,y)\big) = \int_0^t ds \int dy\, f(s,y) a_{t-s}(x,y).$$

In fact, note that if $f \in L^\infty([0,T] \times \mathbb{R}^d)$ then

$$\big|U_a f(t,x)\big| \le \|f\|_\infty \int_0^t ds \int dy \frac{|\gamma_{t-s}(x,y)|}{(t-s)^\rho} \le C \|f\|_\infty t^{1-\rho}.$$

Note in particular that this estimate implies that $U_a^n f$ is well defined for $f \in L^\infty([0,T] \times \mathbb{R}^d)$. We also define for $0 < s \le t \le T$ and $x, y \in \mathbb{R}^d$

$$\begin{aligned} &A_1\big((t,x),(s,y)\big) = A\big((t,x),(s,y)\big),\\ &A_n\big((t,x),(s,y)\big) = \int_s^{s_0} ds_1 \int dy_1 \int_{s_1}^{s_0} ds_2 \int dy_2 \cdots \int_{s_{n-2}}^{s_0} ds_{n-1} \int dy_{n-1} \\ &\qquad \times \prod_{i=0}^{n-1} A\big((s_i, y_i),(s_{i+1}, y_{i+1})\big) \end{aligned} \tag{3.4}$$

with the convention that $s_0 = t$, $y_0 = x$ and $s_n = s$, $y_n = y$, $n \ge 2$. Notice that we have that A_n is finite. That is, we have the following.

LEMMA 3.4. *Assume Hypothesis 3.2, then there exists a constant $C(T,\rho)$ such that*

$$\big|A_n\big((t,x),(s,y)\big)\big| \le C\beta_{t-s}(x,y) \times \frac{C^n(T,\rho)}{[1+n\rho]!}, \tag{3.5}$$

$$A_{n+1}\big((t,x),(s,y)\big) = \int_s^t ds_1 \int dy_1\, A_n\big((t,x),(s_1,y_1)\big) A\big((s_1,y_1),(s,y)\big). \tag{3.6}$$

[3]Notice that according to our remark about the meaning of the variables in the integrals of (1.9), this order of time is reversed with respect to the order in space.

PROOF. We use Hypothesis 3.2 and we obtain (with $s_0 = t, y_0 = x, s_n = s, y_n = y$)

$$
\begin{aligned}
&\left|A_n\big((t,x),(s,y)\big)\right| \\
&\quad \le \int_s^{s_0} ds_1 \int dy_1 \int_{s_1}^{s_0} ds_2 \int dy_1 \cdots \int_{s_{n-2}}^{s_0} ds_{n-1} \int dy_{n-1} \\
&\qquad \times \prod_{i=0}^{n-1} (s_i - s_{i+1})^{-\rho} \gamma_{s_i - s_{i+1}}(y_i, y_{i+1}) \\
&\quad \le C^n \beta_{s_0 - s_n}(y_0, y_n) \int_s^{s_0} ds_1 \int_{s_1}^{s_0} ds_2 \cdots \int_{s_{n-2}}^{s_0} ds_{n-1} \prod_{i=0}^{n-1} (s_i - s_{i+1})^{-\rho} \\
&\quad \le C^n \beta_{t-s}(x,y)(t-s)^{n(1-\rho)} \frac{\Gamma^n(\rho)}{[1+n\rho]!} = \beta_{t-s}(x,y) \frac{C^n(T,\rho)}{[1+n\rho]!}
\end{aligned}
$$

the last inequality being a consequence of the change of variables $s_i = s_0 - t_i$ and Lemma A.1 where we have set $C(T,\rho) = CT^{1-\rho}\Gamma(\rho)$. □

Now that A_n is well defined we can now give an explicit formula for U_a^n.

LEMMA 3.5. *Assume Hypotheses* 3.2 *and* 3.3. *Let* $f \in L^\infty([0,T] \times \mathbb{R}^d)$ *then*

$$
U_a^n f(t,x) = \int_0^t ds \int dy A_n\big((t,x),(s,y)\big) f(s,y). \tag{3.7}
$$

PROOF. For $n = 1$, this is true by the definition of U_a. Suppose that this is true for n and let us prove it for $n+1$. By (3.6),

$$
\begin{aligned}
&U_a^{n+1} f(t,x) \\
&\quad = U_a^n U_a f(t,x) = \int_0^t du \int dz A_n\big((t,x),(u,z)\big) U_a f(u,z) \\
&\quad = \int_0^t du \int dz A_n\big((t,x),(u,z)\big) \int_0^u dv \int dw A\big((u,z),(v,w)\big) f(v,w) \\
&\quad = \int_0^t dv \int dw f(v,w) \int_v^t du \int dz A_n\big((t,x),(u,z)\big) A\big((u,z),(v,w)\big) \\
&\quad = \int_0^t dv \int dw f(v,w) \int dz A_{n+1}\big((t,x),(v,w)\big).
\end{aligned}
$$

So (3.7) is proved. The integrability of the above expressions follows from Lemma 3.4 and Hypothesis 3.3. □

The main estimate in this section is the following. For this, we define

$$
C_T(\rho) = \sum_{n=0}^{\infty} \frac{C^n(T,\rho)}{[1+n\rho]!}.
$$

THEOREM 3.6. (A) *Assume that Hypotheses* 3.2 *and* 3.3 *hold true*. *Then the series*

$$S_a\big((t,x),(s,y)\big)=\sum_{n=1}^{\infty}A_n\big((t,x),(s,y)\big) \tag{3.8}$$

is absolutely convergent and

$$\big|S_a\big((t,x),(s,y)\big)\big|\le C_T(\rho)\beta_{t-s}(x,y).$$

(B) *Moreover, for* $f\in L^{\infty}([0,T]\times\mathbb{R}^d)$ *the series*

$$H_a f(t,x):=\sum_{n=0}^{\infty}U_a^n f(t,x)$$

is absolutely convergent and

$$\big|H_a f(t,x)\big|\le C_T(\rho)C_\beta(t,x)\|f\|_\infty.$$

Finally,

$$H_a f(t,x)=\int_0^t\int f(s,y)S_a\big((t,x),(s,y)\big)\,dy\,ds.$$

(C) *Let* $f,g\in L^{\infty}([0,T]\times\mathbb{R}^d)$ *such that*

$$f=g+U_a f.$$

Then $f=H_a g$.

PROOF. From Lemma 3.4, (3.5), we conclude that the series $S_a((t,x),(s,y))=\sum_{n=1}^{\infty}A_n((t,x),(s,y))$ is absolutely convergent and $|S_a((t,x),(s,y))|\le C_T(\rho)\beta_{t-s}(x,y)$. We consider now a function $f\in L^{\infty}([0,T]\times\mathbb{R}^d)$. As a consequence of (3.5),

$$U_a^n f(t,x)\le\frac{C^n(T,\rho)}{[1+n\rho]!}\|f\|_\infty\int_0^t\int dy\,\beta_{t-s}(x,y)\le\frac{C^n(T,\rho)}{[1+n\rho]!}C_\beta(t,x)\|f\|_\infty.$$

It follows that the series $H_a f(t,x):=\sum_{n=0}^{\infty}U_a^n f(t,x)$ is absolutely convergent and $|H_a f(t,x)|\le C_T(\rho)C_\beta(t,x)\|f\|_\infty$. Furthermore, from the above estimates it is clear that we can exchange integrals and sums in order to prove that $H_a g$ is a solution to the equation (3.1). For any given bounded solution f to (3.1), we obtain by iteration of the equation that the solution has to be $H_a g$ and, therefore, we get the uniqueness. □

We give now a corollary with the study of the fundamental solution. This will be used in order to obtain the density functions corresponding to the operators appearing in (1.7) and (1.10). The proof follows directly from the statements and method of proof of Theorem 3.6. For this, we define

$$\mathcal{M} := \left\{ G : (0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}; \int_0^t ds \int dz \left| G_s(z,y)\beta_{t-s}(x,z) \right| < \infty, \right.$$
$$\left. \forall y \in \mathbb{R}^d, \forall t \in [0,T] \right\}.$$

Furthermore, for $G \in \mathcal{M}$ and $g \in L^\infty([0,T] \times \mathbb{R}^d)$, we define $Gg(t,x) = \int dy\, g(y) G_t(x,y)$.

COROLLARY 3.7. *Assume that Hypotheses 3.2 and 3.3 hold true. Then S_a is the fundamental solution to the equation $f = g + U_a f$. That is, for any $g \in L^\infty([0,T] \times \mathbb{R}^d)$, the solution can be written as*

$$f(t,x) = g(t,x) + \int_0^t ds \int dy\, g(s,y) S_a\bigl((t,x),(s,y)\bigr).$$

Furthermore, consider the equation $f = Gg + U_a f$ where $g \in L^\infty([0,T] \times \mathbb{R}^d)$ and $G \in \mathcal{M}$. Then the unique solution can be written as $f(t,x) = Gg(t,x) + \int dy\, g(y) \bar{S}_a((t,x),(0,y))$ where $\bar{S}_a$ is given by the following uniform absolutely convergent infinite sum:

$$\begin{aligned} \bar{S}_a\bigl((t,x),(0,y)\bigr) &= \int_0^t ds \int dz\, G_s(z,y) S_a\bigl((t,x),(s,z)\bigr) \\ &= \sum_{n=1}^{\infty} \int_0^t ds \int dz\, G_s(z,y) A_n\bigl((t,x),(s,z)\bigr). \end{aligned} \tag{3.9}$$

$\bar{S}_a$ is usually called the fundamental solution of the equation $f = Gg + U_a f$. We will now start discussing regularity properties of the solution. We have to replace hypothesis (3.2) by a slightly stronger hypothesis which will lead to uniform integrability.

HYPOTHESIS 3.8. *Given some functions $\gamma : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ and $G \in \mathcal{M}$. Assume that there exists $r > 0$, $\zeta > 1$ and a function $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ such that the following holds:*

(i) *For every $z_0, z_n \in \mathbb{R}^d$ and $R > 0$ there exists a constant $C_R \equiv C_R(z_0,z_n) > 0$ such that for every $n \geq 2$, $\delta_i > 0, i = 0,\dots,n-1$ and $(y_0,y_n) \in B(z_0,r) \times B(z_n,r)$ we have*

$$\begin{aligned} &\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1} |y_i| \leq R\}} \prod_{i=0}^{n-2} \gamma_{\delta_i}(y_i,y_{i+1})^\zeta \left| G_{\delta_{n-1}}(y_{n-1},y_n) \right|^\zeta \\ &\quad \leq C_R^n \xi\left(\sum_{i=0}^{n-1} \delta_i \right). \end{aligned} \tag{3.10}$$

(ii) *For every* $z_0, z_n \in \mathbb{R}^d$ *there exists a constant* $C \equiv C(z_0, z_n) > 0$ *such that for every* $n \in \mathbb{N}$, $\delta_i > 0, i = 0, \dots, n-1$, $(y_0, y_n) \in B(z_0, r) \times B(z_n, r)$ *and* $\varepsilon > 0$, *there exists* $R_\varepsilon > 0$ *with*

$$\begin{aligned}\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1} |y_i| > R_\varepsilon\}} \prod_{i=0}^{n-2} \gamma_{\delta_i}(y_i, y_{i+1}) G_{\delta_{n-1}}(y_{n-1}, y_n) \\ \le C^n \varepsilon \xi\left(\sum_{i=0}^{n-1} \delta_i\right).\end{aligned} \tag{3.11}$$

The reason for both conditions should be clear. The first one, gives a uniform integrability condition on compact sets. The second condition states that the measure of the complement of the compact set $\{\sum_{i=1}^{n-1} |y_i| \le R\}$ is sufficiently small.

LEMMA 3.9. *Assume Hypotheses 3.2 and 3.3. Suppose that Hypothesis 3.8 holds for some* $\zeta \in (1, \rho^{-1})$ *and* γ *given in Hypothesis 3.2. Furthermore, assume that* $(t, x, y) \to (G_t(x, y), a_t(x, y))$ *is continuous in* $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. *Then* $(t, x, y) \to \bar{S}_a((t, x), (0, y))$ *is continuous.*

PROOF. First recall that (3.9) is a uniform absolutely convergent sum, therefore, it is enough to prove the joint continuity of each term in the sum. Each term is divided in two integrals on disjoint sets. The first, on a compact set, is uniformly integrable because

$$\begin{aligned}\sup_{|y_0||y_n| \le K} \int_s^{s_0} ds_1 \int dy_1 \int_{s_1}^{s_0} ds_2 \int dy_1 \cdots \int_{s_{n-2}}^{s_0} ds_{n-1} \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1} |y_i| \le R\}} \\ \times \prod_{i=0}^{n-2} \left|A\big((s_i, y_i), (s_{i+1}, y_{i+1})\big)\right|^\zeta \left|G_{s_n}(y_{n-1}, y_n)\right|^\zeta < \infty\end{aligned}$$

so that we have uniform integrability for the integrand. Then one may interchange the limit $\lim_{(s_0, y_0, y_{n+1}) \to (s_0', y_0', y_{n+1}')}$ from outside to inside the integral for fixed n and R.

The argument now finishes fixing $\varepsilon > 0$ and, therefore, there exists R_ε such that (3.11) is satisfied. Therefore,

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} \sup_{(y_0, y_n) \in B(z_0, r) \times B(z_n, r)} \int_s^{s_0} ds_1 \int dy_1 \int_{s_1}^{s_0} ds_2 \int dy_1 \cdots \\ \times \int_{s_{n-2}}^{s_0} ds_{n-1} \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1} |y_i| \ge R_\varepsilon\}} \\ \times \prod_{i=0}^{n-2} \left|A\big((s_i, y_i), (s_{i+1}, y_{i+1})\big)\right| \left|G_{s_n}(y_{n-1}, y_n)\right| \le C \lim_{\varepsilon \downarrow 0} \varepsilon \xi(s_0) = 0.\end{aligned}$$

This gives the continuity of the partial sums and then of the series itself due to the uniform convergence in (3.9). □

We discuss now the differentiability properties.

THEOREM 3.10. *Assume Hypotheses 3.2 and 3.3 and suppose that* $\bar{G}_t(x, y) := \nabla_y G_t(x, y)$ *exists for all* $(t, x, y) \in (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. *Furthermore, assume that Hypothesis 3.8 is satisfied with* $(\gamma, \bar{G})$ *replacing* (γ, G) *then the application* $y \to \bar{S}_a((t, x), (0, y))$ *is differentiable for* $t > 0$ *and* $y \in \mathbb{R}^d$ *and the sum below converges absolutely and uniformly for* $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\nabla_y \bar{S}_a\big((t, x), (0, y)\big) = \sum_{n=1}^{\infty} \int_0^t ds \int dz\, \nabla_y G_s(z, y) A_n\big((t, x), (s, z)\big).$$

The proof is done in a similar way as the proof of Lemma 3.9 using the definition of derivative.

4. Abstract framework for semigroup expansions. In this section, we introduce a general framework which will be used in order to obtain a Taylor-like expansion method for Markovian semigroups.

HYPOTHESIS 4.1. $(P_t)_{t\geq 0}$ *is a semigroup of linear operators defined on a space containing* $C_c^\infty(\mathbb{R}^d)$ *with infinitesimal generator* L *such that* $C_c^\infty(\mathbb{R}^d) \subseteq \mathrm{Dom}(L)$. $P_t f(x)$ *is jointly measurable and bounded in the sense that* $\|P_t f\|_\infty \leq \|f\|_\infty$ *for all* $f \in C_c^\infty(\mathbb{R}^d)$ *and* $t \in [0, T]$.

The first goal of this article is to give an expansion for $P_T f(x)$ for fixed $T > 0$ and $f \in C_c^\infty(\mathbb{R}^d)$ based on a parametrized semigroup of linear operators $(P_t^z)_{t\geq 0}$, $z \in \mathbb{R}^d$.

In the case of continuous diffusions to be discussed in Section 7, P^z stands for the semigroup of a diffusion process with coefficients "frozen" at z. We consider an explicit approximating class in the diffusion case in Section 7 given by the Euler–Maruyama scheme.

Our hypothesis on $(P_t^z)_{t\geq 0}$ are:

HYPOTHESIS 4.2. *For each* $z \in \mathbb{R}^d$, $(P_t^z)_{t\geq 0}$ *is a semigroup of linear operators defined on a space containing* $C_c^\infty(\mathbb{R}^d)$ *with infinitesimal generator* L^z *such that* $C_c^\infty(\mathbb{R}^d) \subseteq \mathrm{Dom}(L^z)$. *We also assume that* $P_t^z f(x) = \int f(y) p_t^z(x, y)\, dy$ *for any* $f \in C_c^\infty(\mathbb{R}^d)$, $(x, z) \in \mathbb{R}^d \times \mathbb{R}^d$ *and a jointly measurable probability kernel* $p^z \in C((0, T] \times \mathbb{R}^d \times \mathbb{R}^d)$.

The link between L and L^z is given by the following hypothesis.

HYPOTHESIS 4.3. *$Lf(z) = L^z f(z)$ for every $f \in C_c^\infty(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$.*

To simplify notation, we introduce $Q_t f(x) := P_t^x f(x)$, noticing that $(Q_t)_{t\geq 0}$ is no longer a semigroup but it still satisfies that $\|Q_t f\|_\infty \leq \|f\|_\infty$ for all $t \in [0, T]$. We will use the following notation in the forward and backward method, respectively

$$\psi_t^x(y) := p_t^x(x, y),$$
$$\phi_t^z(x) := p_t^z(x, z).$$

The reason for using the above notation is to clarify to which variables of $p_t^z(x, y)$ an operator applies to. This is the case of, for example, $L^z\phi_t^z(x) \equiv (L^z\phi_t^z)(x)$.

The expansion we want to obtain can be achieved in two different ways. One will be called the forward method and the other called the backward method. In any of these methods, the expansion is done based on the semigroup $(P_t^z)_{t\geq 0}$, $z \in \mathbb{R}^d$. In the classical Taylor-like expansion one needs to use polynomials as basis functions. In the forward method, these polynomials will be replaced by products (or compositions) of the following basic operator S,

$$S_t f(x) := \int \big(L^y - L^x\big) f(y)\psi_t^x(y)\, dy, \qquad f \in \bigcap_{x\in\mathbb{R}^d} \mathrm{Dom}\big(L^x\big).$$

In the backward method, a similar role is played by the operator

$$\hat{S}_t f(y) := \int f(x)\big(L^x - L^y\big)\phi_t^y(x)\, dx. \tag{4.1}$$

The above Hypotheses 4.1, 4.2 and 4.3 will be assumed throughout the theoretical part of the article. They will be easily verified in the examples.

5. Forward method. We first state the assumptions needed in order to implement the forward method.

HYPOTHESIS 5.1. $P_t^z g$, $P_t g \in \bigcap_{x\in\mathbb{R}^d} \mathrm{Dom}(L^x)$, $\forall g \in C_c^\infty(\mathbb{R}^d)$, $z \in \mathbb{R}^d$, $t \in [0, T]$.

We assume the following two regularity properties for the difference operator S.

HYPOTHESIS 5.2. *There exists a jointly measurable real valued function $\theta : (0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, such that for all $f \in C_c^\infty(\mathbb{R}^d)$ we have that*

$$S_t f(x) = \int f(y)\theta_t(x, y) P_t^x(x, dy) = \int f(y)\theta_t(x, y) p_t^x(x, y)\, dy,$$
$$(t, x) \in (0, T] \times \mathbb{R}^d.$$

We assume that the function $a_t(x, y) = \theta_t(x, y)p_t^x(x, y)$ verifies the Hypotheses 3.2 and 3.3 and that $G_t(x, y) = p_t^x(x, y) \in \mathcal{M}$.

Note that the above hypothesis implies that the operator S can be extended to the space of bounded functions.

HYPOTHESIS 5.3. *For the functions (a, γ) and the constant $\rho \in [0, 1)$ satisfying Hypothesis 3.2 and $G_t(x, y) = p_t^x(x, y) \in \mathcal{M}$ we assume that the Hypothesis 3.8 is satisfied for some $\zeta \in (1, \rho^{-1})$.*

REMARK 5.4. We remark here that Hypothesis 5.2 entails some integration by parts property which will be made clear when dealing with examples in Section 7 [see (7.3)].

Define for $(s_0, x) \in (0, T] \times \mathbb{R}^d$ and $f \in C_c^\infty(\mathbb{R}^d)$ the following integral operator:

$$
\begin{aligned}
&I_{s_0}^n(f)(x) \\
&\quad := \begin{cases} \displaystyle\int_0^{s_0} ds_1 \cdots \int_0^{s_{n-1}} ds_n \left(\prod_{i=0}^{n-1} S_{s_i - s_{i+1}}\right) Q_{s_n} f(x), & \text{if } n \geq 1, \\ Q_{s_0} f(x), & \text{if } n = 0. \end{cases}
\end{aligned}
\tag{5.1}
$$

We denote by A_n the kernels associated to $a_t(x, y)$ defined in (3.4). Then using the change of variables $t_i = s_0 - s_i$ we obtain the following representation $I_{s_0}^n(f)(x) = \int f(y) I_{s_0}^n(x, y)\,dy$ with

$$
\begin{aligned}
&I_{t_{n+1}}^n(x, y) \\
&\quad = \int_0^{t_{n+1}} dt_n \int p_{t_{n+1}-t_n}^{y_n}(y_n, y) A_n\big((t_{n+1}, x), (t_{n+1} - t_n, y_n)\big)\,dy_n.
\end{aligned}
\tag{5.2}
$$

The following is the main result of this section, which is a Taylor-like expansion of P based on Q.

THEOREM 5.5. *Suppose that Hypotheses 5.1 and 5.2 hold. Then for every $f \in C_c^\infty(\mathbb{R}^d)$ and $t \in (0, T]$, $I_t^n(f)$ is well defined and the sum $\sum_{n=1}^\infty I_t^n(f)(x)$ converges absolutely and uniformly for $(t, x) \in [0, T] \times \mathbb{R}^d$. Moreover,*

$$
P_t f(x) = \sum_{n=0}^\infty I_t^n(f)(x).
\tag{5.3}
$$

Then for fixed $t \in (0, T]$, $\sum_{n=1}^\infty I_t^n(x, y)$ also converges absolutely and uniformly for $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and we have that $P_t f(x) = \int f(y) p_t(x, y)\,dy$ where

$$
p_t(x, y) = p_t^x(x, y) + \sum_{n=1}^\infty I_t^n(x, y).
\tag{5.4}
$$

Furthermore, suppose that $P_t f(x) \geq 0$ for $f \geq 0$ and $P_t 1 = 1$ for all $t \geq 0$. Then $p_t(x, y)$ is a density function.

PROOF. The linear equation on $P_t f$ is obtained, using Hypotheses 4.1, 4.2, 5.1 and 5.2 as follows:

$$P_t f(x) - P_t^x f(x) = \int_0^t \partial_{s_1}\big(P_{t-s_1}^x P_{s_1} f\big)(x)\,ds_1 = \int_0^{t_0} P_{t_0-s_1}^x\big(L - L^x\big)P_{s_1} f(x)\,ds_1.$$

Note that Hypothesis 5.2 ensures the finiteness of the above integral.

Using the identity in Hypothesis 4.3, $Lg(x) = L^x g(x)$ with $g(y) = P_{s_1} f(y)$, we obtain

$$\begin{aligned} P_{t-s_1}^x\big(L - L^x\big)P_{s_1} f(x) &= \int \big(L - L^x\big)P_{s_1} f(y) P_{t-s_1}^x(x, dy) \\ &= \int \big(L^y - L^x\big)P_{s_1} f(y) P_{t-s_1}^x(x, dy) \\ &= S_{t-s_1} P_{s_1} f(x). \end{aligned}$$

Therefore, we have the following equation:

$$P_t f(x) = P_t^x f(x) + \int_0^t ds_1 \int dy\, P_{s_1} f(y)\theta_{t-s_1}(x, y) p_{t-s_1}^x(x, y). \tag{5.5}$$

This is equation (3.1) with $a_t(x, y) = \theta_t(x, y) p_t^x(x, y)$. Therefore, due to Hypotheses 5.1 and 5.2 we obtain that the hypotheses needed for the application of Lemma 3.4 and Theorem 3.6 are satisfied. Therefore, we obtain that $\sum_{n=0}^{\infty} I_t^n(f)(x)$ converges absolutely and uniformly and is the unique solution of (5.5).

Corollary 3.7 gives (5.4). Finally, one proves that as the semigroup P is positive then $p_t(x, y)$ has to be positive locally in y for fixed (t, x) then as $P_t 1 = 1$ one obtains that $\int p_t(x, y)\,dy = 1$. $\square$

5.1. *Probabilistic representation using the forward method.* Our aim now is to give a probabilistic representation for the formula (5.3) that may be useful for simulation.

HYPOTHESIS 5.6. *There exists a continuous Markov process $X^\pi = \{X_t^\pi; t \in [0, T]\}$ such that $X_0^\pi = x$ and for any $t > s$*

$$P\big(X_t^\pi \in dy' | X_s^\pi = y\big) = P_{t-s}^y\big(y, dy'\big) = p_{t-s}^y\big(y, y'\big)\,dy'. \tag{5.6}$$

With this assumption, we have that $S_1 f(x) = \mathbb{E}[f(X_{t_1}^\pi)\theta_{t_1}(x, X_{t_1}^\pi)]$ and $Q_{t_1} f(x) = \mathbb{E}[f(X_{t_1}^\pi)]$. Therefore, using these representations, we obtain the probabilistic representation of the integrand in (5.1):

$$\left(\prod_{j=0}^{n-1} S_{t_{j+1}-t_j}\right) Q_{T-t_n} f(x) = \mathbb{E}\big[f\big(X_T^\pi\big)\theta_{t_n-t_{n-1}}\big(X_{t_{n-1}}^\pi, X_{t_n}^\pi\big)\cdots\theta_{t_1-t_0}\big(X_{t_0}^\pi, X_{t_1}^\pi\big)\big].$$

Finally, to obtain the probabilistic interpretation for the representation formula (5.3), we need to find the probabilistic representation of the multiple integrals in (5.1).

For this, we consider a Poisson process $(J_t)_{t\geq 0}$ of parameter $\lambda = 1$ and we denote by τ_j, $j \in \mathbb{N}$, its jump times (with the convention that $\tau_0 = 0$). Conditionally to $J_T = n$, the jump times are distributed as the order statistics of a sequence n independent uniformly distributed random variables on $[0, T]$. Therefore, the multiple integrals in (5.1) can be interpreted as the expectation taken with respect to these jump times given that $J_T = n$. Therefore, for $n \geq 1$ we have

$$I_T^n(f)(x) = e^T \mathbb{E}\left[1_{\{J_T=n\}} f\left(X_T^\pi\right) \prod_{j=0}^{n-1} \theta_{\tau_{j+1}-\tau_j}\left(X_{\tau_j}^\pi, X_{\tau_{j+1}}^\pi\right) \right],$$

where (with a slight abuse of notation), π denotes the random time partition of $[0, T]$, $\pi \equiv \pi(\omega) = \{\tau_i(\omega) \wedge T; i = 0, \dots, J_T(\omega) + 1\}$.

From now on, in order to simplify the notation, we denote $\tau_T \equiv \tau_{J_T}$. Given the above discussion, we have the main result for this section.

THEOREM 5.7. *Suppose that Hypotheses 5.1, 5.2 and 5.6 hold. Recall that* $\psi_t^x(y) = p_t^x(x, y)$ *and define* $\Gamma_T(x) \equiv \Gamma_T(x)(\omega)$ *as*

$$\Gamma_T(x) = \begin{cases} \displaystyle\prod_{j=0}^{J_T-1} \theta_{\tau_{j+1}-\tau_j}\left(X_{\tau_j}^\pi, X_{\tau_{j+1}}^\pi\right), & \text{if } J_T \geq 1, \\ 1, & \text{if } J_T = 0. \end{cases}$$

Then the following probabilistic representations are satisfied for $f \in C_c^\infty(\mathbb{R}^d)$:

$$P_T f(x) = e^T \mathbb{E}\left[f\left(X_T^\pi\right)\Gamma_T(x)\right], \tag{5.7}$$

$$p_T(x, y) = e^T \mathbb{E}\left[\psi_{T-\tau_T}^{X_{\tau_T}^\pi}(y)\Gamma_T(x)\right]. \tag{5.8}$$

REMARK 5.8. 1. Extensions for bounded measurable functions f can be obtained if limits are taken in (5.7).

2. The above representations (5.8) and (5.7) may be obtained using a Poisson process of arbitrary parameter $\lambda > 0$ instead of $\lambda = 1$. In fact, if we denote $\{J_t^\lambda, t \geq 0\}$ a Poisson process, by τ_i^λ the jump times and by π_λ the corresponding random time grid. Then the formula (5.8) becomes

$$p_T(x, y) = e^{\lambda T} \mathbb{E}\left[\lambda^{-J_T^\lambda} \psi_{T-\tau_{J_T^\lambda}}^{X_{\tau_T}^{\pi_\lambda}}(y)\Gamma_T(x)\right].$$

5.2. *Regularity of the density using the forward method.* Now that we have obtained the stochastic representation, we will discuss the differentiability of $p_T(x, y)$ with respect to y. This type of property is also proved when the analytical version of the parametrix method is discussed in the particular case of fundamental solutions of parabolic PDEs (see, e.g., Chapter 1 in [7]).

THEOREM 5.9. *Suppose that the Hypotheses 5.1, 5.2, 5.3 are satisfied. Furthermore assume that* $(t, x, y) \to (p_t^x(x, y), a_t(x, y))$ *is continuous in* $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. *Then* $(t, x, y) \to p_t(x, y)$ *is continuous on* $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$.

THEOREM 5.10. *Suppose that the Hypotheses 5.1, 5.2 and 5.3 are satisfied. Furthermore, we assume that* $y \to p_t^x(x, y)$ *is differentiable for all* $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ *and that Hypothesis 3.8 is satisfied for* $\nabla_y p_t^x(x, y)$ *instead of* G. *Then for every* $(t, x) \in (0, T] \times \mathbb{R}^d$, *the function* $y \to p_t(x, y)$ *is differentiable. Moreover,*

$$\nabla_y p_T(x, y) = e^T \mathbb{E}\big[\nabla_y p_{T-\tau_T}^{X_{\tau_T}^\pi}(X_{\tau_T}^\pi, y)\Gamma_T(x)\big].$$

6. The backward method: Probabilistic representation using the adjoint semigroup. We will now solve equation (1.7) in dual form. We start with a remark in order to beware the reader about the nonapplicability of the forward method directly to the dual problem.

Usually, in semigroup theory one assumes that for each $t > 0$, P_t maps continuously $L^2(\mathbb{R}^d)$ into itself. Then P_t^* can be defined and it is still a semigroup which has as infinitesimal operator L^* defined by $\langle L^* g, f\rangle = \langle g, Lf\rangle$ for f, $g \in C_c^\infty(\mathbb{R}^d)$. Assume, for the sake of the present discussion, that for every $x \in \mathbb{R}^d$, P_t^x maps continuously $L^2(\mathbb{R}^d)$ into itself and we define $P_t^{x,*} \equiv (P_t^x)^*$ and $L^{x,*}$ by $\langle P_t^{x,*} g, f\rangle = \langle g, P_t^x f\rangle$ and $\langle L^{x,*} g, f\rangle = \langle g, L^x f\rangle$ for $f, g \in C_c^\infty(\mathbb{R}^d)$.

Our aim is to obtain for P^* a representation which is similar to the one obtained for P in Theorem 5.7. Unfortunately, the adjoint version of the arguments given in Section 5 do not work directly. In fact, if $P_t^{x,*}$ denotes the adjoint operator of P_t^x then the relation $Lg(x) = L^x g(x)$ does not imply $L^* g(x) = (L^x)^* g(x)$. To make this point clearer, take, for example, the case of a one-dimensional diffusion process with infinitesimal operator $Lg(y) = a(y)\Delta g(y)$ then $L^x g(y) = a(x)\Delta g(y)$ (for more details, see Section 7). Then $L^* g(y) = \Delta(ag)(y)$ and $(L^x)^* g(y) = a(x)\Delta g(y) = L^x g(y)$. So, letting the coefficients of L^* be frozen at $y = x$ does not coincide with $(L^x)^*$ and, therefore, the previous argument will fail.[4] In order not to confuse the reader, we will keep using the superscript $*$ to denote adjoint operators while other approximating operators will be denoted by the superscript $\hat{}$ (hat).

Note that in the diffusion case, proving that for each $t > 0$, P_t maps continuously $L^2(\mathbb{R}^d)$ into itself is not easy in general. Therefore, instead of adding this as a hypothesis, we will make additional hypotheses on the approximation process.

[4]Note that if we wanted to freeze coefficients as in the forward method one may be lead to the study of the operator $L^{*,z} g(y) = a(z)\Delta g(y) + 2\langle\nabla a(z), \nabla g(y)\rangle + g(y)\Delta a(z)$. Although this may have an interest in itself, we do not pursue this discussion here as this will again involve derivatives of the coefficients while in this section we are pursuing a method which may be applied when the coefficients are Hölder continuous.

This issue will demand us to introduce hypotheses that did not have any counterpart in the forward method. Still, once the linear Volterra equation is obtained, the arguments are parallel and we will again use the results in Section 3. Let us introduce some notation and the main hypotheses. We define the linear operator

$$\hat{Q}_t f(y) := (P_t^y)^* f(y) = \int f(x) p_t^y(x, y)\, dx = \int f(x)\phi_t^y(x)\, dx.$$

We assume the following.

HYPOTHESIS 6.1. (0) *$P_t f(x) = \int f(y) P_t(x, dy)$ for all $f \in C_b(\mathbb{R}^d)$. In particular, P_t is an integral operator.*

(i) *$\int p_t^y(x, y)\, dy < \infty$, for all $x \in \mathbb{R}^d$ and $\int p_t^y(x, y)\, dx < \infty$ for all $y \in \mathbb{R}^d$.*

(ii) *$\lim_{\varepsilon \to 0} P_{T+\varepsilon}^{z,*} g(w) = P_T^{z,*} g(w)$ and $\lim_{\varepsilon \to 0} \int h(z)\phi_\varepsilon^z(w)\, dz = h(w)$ for all $(z, w) \in \mathbb{R}^d \times \mathbb{R}^d$ and for $g, h \in C_c^\infty(\mathbb{R}^d)$.*

(iii) *$\phi_t^z \in \mathrm{Dom}(L) \cap (\bigcap_{y \in \mathbb{R}^d} \mathrm{Dom}(L^y))$, for all $(t, z) \in (0, T] \times \mathbb{R}^d$.*

With these definitions and hypotheses, we have by the semigroup property of P^z, that

$$P_t^z \phi_\varepsilon^z(x) = p_{t+\varepsilon}^z(x, z) = \phi_{t+\varepsilon}^z(x). \tag{6.1}$$

As stated before, we remark that $\hat{Q} \neq Q^*$. In fact, $\hat{Q}$ is defined through a density whose coefficients are "frozen" at the arrival point of the underlying process. Also note that due to Hypothesis 6.1 then $\|\hat{Q}_t f\|_\infty \le C_T \|f\|_\infty$ for all $t \in [0, T]$.

Before introducing the next two hypotheses, we explain the reasoning behind the notation to follow. In the forward method, it was clear that the dynamical system expressed through the transition densities went from a departure point x to an arrival point y with transition points y_i, $i = 0, \dots, n+1$, $y_0 = x$ and $y_{n+1} = y$. In the backward method, the situation is reversed. The initial point for the method is y, the arrival point is x and $y_0 = y$ and $y_{n+1} = x$. The notation to follow tries to give this intuition.

HYPOTHESIS 6.2. *We suppose that there exists a continuous function $\hat{\theta} \in C((0, T] \times \mathbb{R}^d \times \mathbb{R}^d)$ such that $(L^{y_1} - L^y)\phi_t^y(y_1) = \hat{\theta}_t(y_1, y)\phi_t^y(y_1)$. Moreover, we assume that $\hat{\theta}_t(y_1, y)\phi_t^y(y_1)$ is integrable $P_s(x, dy_1)$ for all $(s, x) \in [0, T] \times \mathbb{R}^d$ and $(t, y) \in (0, T] \times \mathbb{R}^d$.*

Define the function $\hat{a}_t(x, y) := \hat{\theta}_t(y, x) p_t^x(y, x) = \hat{\theta}_t(y, x)\phi_t^x(y)$.

HYPOTHESIS 6.3. *Assume that the function $\hat{a}$ satisfies the Hypotheses 3.2 and 3.3 with $G_t(x, y) = p^x(y, x) \in \mathcal{M}$. Furthermore, we assume that the corresponding function γ satisfies $\sup_{(t,y) \in [0,T] \times \mathbb{R}^d} \int |\gamma_t(x, y)|\, dx < \infty$ and that there exists $\zeta \in (1, \rho^{-1})$ such that for every $R > 0$*

$$\sup_{(t,y) \in [0,T] \times \mathbb{R}^d} \int 1_{\{|x| \le R\}} |\gamma_t(x, y)|^\zeta\, dx < \infty. \tag{6.2}$$

HYPOTHESIS 6.4. *For the function $\hat{a}_t(x,y)$ we assume that Hypothesis 3.8 is satisfied for some $\zeta \in (1,\rho^{-1})$.*

We define now [recall (4.1) and Hypothesis 6.2]

$$\hat{S}_t f(y) := \int f(x)\hat{a}_t(y,x)\,dx.$$

For $g \in C_c^\infty(\mathbb{R}^d)$, we define

$$\hat{I}^n_{s_0}(g)(y) \tag{6.3}$$
$$:= \begin{cases} \int_0^{s_0} ds_1 \cdots \int_0^{s_{n-1}} ds_n \left(\prod_{i=0}^{n-1} \hat{S}_{s_i - s_{i+1}}\right)\hat{Q}_{s_n} g(y), & \text{if } n \geq 1, \\ \hat{Q}_{s_0} g(y), & \text{if } n = 0. \end{cases}$$

Furthermore, we define the adjoint operators

$$\hat{Q}^*_t f(x) := \int f(y) p^y_t(x,y)\,dy,$$
$$\hat{S}^*_t f(x) := \int f(y)\hat{a}_t(y,x)\,dy.$$

Note that due to the Hypotheses 6.1(i) and 6.3 we have that for any $f \in L^\infty$

$$\sup_t \|\hat{Q}^*_t f\|_\infty \leq C\|f\|_\infty, \tag{6.4}$$

$$\|\hat{S}^*_t f\|_\infty \leq \frac{C}{t^\rho}\|f\|_\infty. \tag{6.5}$$

As in (5.1), we define the following auxiliary operators for

$$\hat{I}^{n,*}_t(f) := \begin{cases} \int_0^t dt_n \cdots \int_0^{t_2} dt_1 \hat{Q}^*_{t_1}\hat{S}^*_{t_2-t_1}\cdots \hat{S}^*_{t-t_n} f, & n \geq 1, \\ \hat{Q}^*_t f, & n = 0, \end{cases}$$

$$\hat{I}^n_t(y_0, y_{n+1})$$
$$:= \int_0^t dt_n \cdots \int_0^{t_2} dt_1 \int dy_1 \cdots \int dy_n\, \hat{A}_n\big((t,y_0),(t-t_n,y_n)\big) p^{y_n}_{t_n}(y_{n+1},y_n).$$

Here, $\hat{A}_n$ denotes the same function defined in (3.4) where $\hat{a}$ is used instead of a. Note that $\langle \hat{I}^{n,*}_{t_0}(f), g\rangle = \langle f, \hat{I}^n_{t_0} g\rangle$ and $\hat{I}^{n,*}_t(f)(x) = \int f(y)\hat{I}^{n,*}_t(y,x)\,dy$ for $f,g \in C_c^\infty(\mathbb{R}^d)$. Our main result in this section is:

THEOREM 6.5. *Suppose that Hypotheses 6.1, 6.2 and 6.3. Then for every $g \in C_c^\infty(\mathbb{R}^d)$ the sum $\sum_{n=0}^\infty \hat{I}^n_t(g)(y)$ converges absolutely and uniformly for $(t,y) \in$*

$(0, T] \times \mathbb{R}^d$ *and the following representation formula is satisfied*:

$$P_t^* g(y) = \sum_{n=0}^{\infty} \hat{I}_t^n(g)(y), \qquad dy\text{-a.s.}, t \in (0, T]. \tag{6.6}$$

The above equality is understood in the following weak sense $\langle P_t^* g, h\rangle = \langle g, P_t h\rangle = \sum_{n=0}^{\infty} \langle \hat{I}_t^n(g), h\rangle$ *for all* $(g, h) \in C_c^\infty(\mathbb{R}^d) \times C_c^\infty(\mathbb{R}^d)$. *Furthermore*, $\sum_{n=0}^{\infty} \hat{I}_t^{n,*}(f)(x)$ *converges absolutely and uniformly for* $x \in \mathbb{R}^d$ *and fixed* $f \in C_c^\infty(\mathbb{R}^d)$, $t \in (0, T]$ *and it satisfies*

$$P_t f(x) = \sum_{n=0}^{\infty} \hat{I}_t^{n,*}(f)(x), \qquad dx\text{-a.s.}$$

Finally, $\sum_{n=0}^{\infty} \hat{I}_t^n(y, x)$ *converges absolutely and uniformly for* $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ *for fixed* $t > 0$ *and there exists a jointly measurable function* $p_t(x, y)$ *such that we have that for* $f \in C_c^\infty(\mathbb{R}^d)$ *we have* $P_t f(x) = \int f(y) p_t(x, y)\, dy$ *and it is given by*

$$p_t(x, y) = p_t^y(x, y) + \sum_{n=1}^{\infty} \hat{I}_t^n(y, x).$$

Furthermore, *suppose that* $P_t f(x) \geq 0$ *for* $f \geq 0$ *and* $P_t 1 = 1$ *for all* $t \geq 0$. *Then* $p_t(x, y)$ *is a density function.*

PROOF. Many of the arguments are similar to the proof of Theorem 5.5. In fact, we first establish the Volterra equations satisfied by P_t^*. In order to do this, we need an approximation argument. We fix $\varepsilon > 0$ and we recall that due to Hypotheses 4.2 and 6.1(iii), we have for each $z \in \mathbb{R}^d$ that $P_{T-s}^z \phi_\varepsilon^z = \phi_{T-s+\varepsilon}^z = p_{s+\varepsilon}^z(\cdot, z) \in \mathrm{Dom}(L)$ and $Lf(x) = L^x f(x)$. Then from Hypotheses 6.2 and 6.3, we have that for $0 < s < T$,

$$\begin{aligned}
\partial_s\big(P_s P_{T-s}^{y_0}\big)\phi_\varepsilon^{y_0}(x) &= P_s\big(L - L^{y_1}\big) P_{T-s}^{y_0} \phi_\varepsilon^{y_0}(x) \\
&= \int P_s(x, dy_1)\big(L - L^{y_0}\big) P_{T-s}^{y_0} \phi_\varepsilon^{y_0}(y_1) \\
&= \int P_s(x, dy_1) \hat{a}_{T-s+\varepsilon}(y_0, y_1).
\end{aligned}$$

We take $g, h \in C_c^\infty(\mathbb{R}^d)$ and we note that due to Hypothesis 6.3

$$\begin{aligned}
&\int dx \big|g(x)\big| \int P_s(x, dy_1) \int dy_0 \big|h(y_0)\big| \big|\hat{a}_{T-s+\varepsilon}(y_0, y_1)\big| \\
&\quad \leq C_3(T - s + \varepsilon)^{-\rho} \|h\|_\infty \int dx \big|g(x)\big| \int P_s(x, dy_1) \\
&\quad \leq C_3(T - s)^{-\rho} \|h\|_\infty \|g\|_1.
\end{aligned}$$

The above expression is integrable with respect to $1_{(0,T)}(s)\, ds$ for $\rho \in (0, 1)$. Therefore this ensures that Fubini–Tonelli's theorem can be applied and multiple integrals appearing in any order will be well defined.

Furthermore, by Hypotheses 6.2, 6.3 [see (6.2)] and the fact that $h \in C_c^\infty(\mathbb{R}^d)$, we have that for fixed $s \in [0, T)$ we can take limits as $\varepsilon \to 0$ for $\int dy_0 |h(y_0)| \times |\hat{a}_{T-s+\varepsilon}(y_0, y_1)|$, and that the uniform integrability property is satisfied. Therefore, we finally obtain that the following limit exists, is finite and the integration order can be exchanged so that

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_0^T dt \int dy_0\, h(y_0) \int dx\, g(x) \int P_s(x, dy_1)\hat{a}_{T-s+\varepsilon}(y_0, y_1)\\ = \int_0^T dt \int dy_0\, h(y_0) \int dx\, g(x) \int P_s(x, dy_1)\hat{a}_{T-s}(y_0, y_1).\end{aligned}$$

From the previous argument, the following sequence of equalities are valid and the limit of the right-hand side below exists:

$$\begin{aligned}&\int dy_0\, h(y_0)\bigl(\langle g, P_T\phi_\varepsilon^{y_0}\rangle - \langle g, P_T^{y_0}\phi_\varepsilon^{y_0}\rangle\bigr)\\ &\quad = \int dy_0\, h(y_0)\int dx\, g(x) \int_0^T \partial_t\bigl(P_s P_{T-s}^{y_0}\bigr)\phi_\varepsilon^{y_0}(x)\, dt \\ &\quad = \int_0^T dt \int dy_0\, h(y_0) \int dx\, g(x) \int P_s(x, dy_1)\hat{a}_{T-s+\varepsilon}(y_0, y_1).\end{aligned} \tag{6.7}$$

In order to obtain the linear Volterra type equation, we need to take limits in (6.7). To deal with the limit of the left-hand side of (6.7), we note that given the assumptions $g, h \in C_c^\infty(\mathbb{R}^d)$ and Hypothesis 6.1(ii), we have

$$\begin{aligned}\lim_{\varepsilon\to 0}\int dy_0\, h(y_0)\langle g, P_T\phi_\varepsilon^{y_0}\rangle &= \lim_{\varepsilon\to 0}\int g(y_1)\int P_T(y_1, dw)\int dy_0\, h(y_0)\phi_\varepsilon^{y_0}(w)\\ &= \langle P_T h, g\rangle,\\ \lim_{\varepsilon\to 0}\int dy_0\, h(y_0)\langle P_T^{y_0,*} g, \phi_\varepsilon^{y_0}\rangle &= \lim_{\varepsilon\to 0}\int dy_0\, h(y_0) P_{T+\varepsilon}^{y_0,*} g(y_0)\\ &= \int dy_0\, h(y_0) P_T^{y_0,*} g(y_0).\end{aligned}$$

Therefore, taking limits in (6.7), we obtain

$$\begin{aligned}\langle P_T h, g\rangle &= \int dy_0\, h(y_0) P_T^{*,y_0} g(y_0)\\ &\qquad + \int dy_0\, h(y_0) \int dx\, g(x)\int_0^T ds \int P_s(x, dy_1)\hat{a}_{T-s}(y_0, y_1)\\ &= \langle \hat{Q}_T^* h, g\rangle + \int_0^T \langle P_s \hat{S}_{T-s}^* h, g\rangle\, ds.\end{aligned}$$

Rewriting this equation with the adjoint of a densely defined operator, we obtain the Volterra-type equation

$$P_T^* g(x) = \hat{Q}_T g(x) + \int_0^T ds \int dy\, P_s^* g(y)\hat{a}_{T-s}(x, y).$$

This equation has a solution due to the results in Section 3 as we have made the necessary hypotheses to apply the results of Corollary 3.7. Therefore, it follows that (6.6) is the unique solution of the above equation. The proof of the other statements are done in the same way as in the proof of Theorem 5.5. □

REMARK 6.6. The previous proof is also valid with weaker conditions on g and h. For example, $g \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ and $h \in C_b(\mathbb{R}^d)$ will suffice with an appropriate change of hypothesis.

6.1. *Probabilistic representation and regularity using the backward method.* We deal now with the representation of the density associated with the semigroup P_T. We recall that in the Section 5.1 [see (5.6)] we have performed a similar construction.

HYPOTHESIS 6.7. *There exists a continuous Markov process* $\{X_t^{*,\pi}(y), t \in [0,T]\}$, $y \in \mathbb{R}^d$ *such that* $X_0^{*,\pi}(y) = y$ *and for any* $t > s$ *we have*

$$\begin{aligned} P\big(X_t^{*,\pi}(y) \in dy_2 | X_s^{*,\pi}(y) = y_1\big) &= C_{t-s}^{-1}(y_1) P_{t-s}^{y_1,*}(y_1, dy_2) \\ &= C_{t-s}^{-1}(y_1) \phi_{t-s}^{y_1}(y_2)\, dy_2, \\ C_{t-s}(y_1) &:= \int \phi_{t-s}^{y_1}(y_2)\, dy_2. \end{aligned}$$

Let $(J_t)_{t\geq 0}$ be a Poisson process of parameter $\lambda = 1$ and we denote by τ_j, $j \in \mathbb{N}$, its jump times (with the convention that $\tau_0 = 0$). Then the same arguments as in the previous section give the representation

$$\begin{aligned} &I_T^{*,n}(g)(y) \\ &\quad = e^T \mathbb{E}\Bigg[1_{\{J_T = n\}} g\big(X_T^{*,\pi}(y)\big) C_{T-\tau_n}\big(X_{\tau_n}^{*,\pi}(y)\big) \\ &\qquad \times \prod_{j=0}^{n-1} C_{\tau_{j+1}-\tau_j}\big(X_{\tau_j}^{*,\pi}(y)\big) \hat{\theta}_{\tau_{j+1}-\tau_j}\big(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\big)\Bigg]. \end{aligned}$$

We define

$$\begin{aligned} &\Gamma_T^*(y) \\ &\quad = \begin{cases} C_{T-\tau_{J_T}}\big(X_{\tau_{J_T}}^{*,\pi}(y)\big) \displaystyle\prod_{j=0}^{J_T-1} C_{\tau_{j+1}-\tau_j}\big(X_{\tau_j}^{*,\pi}(y)\big) \hat{\theta}_{\tau_{j+1}-\tau_j}\big(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\big), \\ \qquad \text{if } J_T \geq 1, \\ C_T(y), \\ \qquad \text{if } J_T = 0. \end{cases} \end{aligned}$$

Sometimes we may use the notation $X^{*,\pi}_{\tau_j}(y)$ to indicate that $X^{*,\pi}_0(y) = y$. The main result in this section is about representations of the adjoint semigroup P^* and its densities.

THEOREM 6.8. *Suppose that Hypotheses 6.1, 6.2 and 6.7 hold then the following representation formula is valid for any $g \in C_c^\infty(\mathbb{R}^d)$*:

$$
\begin{aligned}
P_T^* g(y) &= P_T^{*,z} g(y) + e^T \mathbb{E}\big[g\big(X_T^{*,\pi}(y)\big)\Gamma_T^*(y) 1_{\{J_T \ge 1\}}\big] \\
&= e^T \mathbb{E}\big[g\big(X_T^{*,\pi}(y)\big)\Gamma_T^*(y)\big].
\end{aligned}
\tag{6.8}
$$

THEOREM 6.9. *Suppose that Hypotheses 6.1, 6.2 and 6.7 hold then the following representation formula for the density is valid*:

$$
p_T(x,y) = e^T \mathbb{E}\big[p_{T-\tau_T}^{X^{*,\pi}_{\tau_T}(y)}\big(x, X^{*,\pi}_{\tau_T}(y)\big)\Gamma_T^*(y)\big]. \tag{6.9}
$$

In particular, let Z be a random variable with density $h \in L^1(\mathbb{R}^d; \mathbb{R}_+)$ then we have

$$
P_T h(x) = e^T \mathbb{E}\big[p_{T-\tau_T}^{X^{*,\pi}_{\tau_T}(Z)}\big(x, X^{*,\pi}_{\tau_T}(Z)\big)\Gamma_T^*(Z)\big].
$$

PROOF. Using the definition of $X^{*,\pi}$ we have for $g \in C_c^\infty(\mathbb{R}^d)$ (we recall that $\tau_T \equiv \tau_{J_T}$)

$$
\mathbb{E}\big[g\big(X_T^{*,\pi}\big)C_{T-\tau_T}(y) | \tau_T, X^{*,\pi}_{\tau_T} = y\big] = \int g(x) p^{*,y}_{T-\tau_T}(y,x)\,dx
$$

so that (6.8) says that $P_T^*(y,dx) = p_T^*(y,x)\,dx$ with

$$
p_T^*(y,x) = e^T \mathbb{E}\big[p_{T-\tau_T}^{*,X^{*,\pi}_{\tau_T}(y)}\big(X^{*,\pi}_{\tau_T}(y), x\big)\Gamma_T^*(y)\big]. \tag{6.10}
$$

Notice that $p_T^*(y,x) = p_T(x,y)$ so the above equality says that $P_T(x,dy) = p_T^*(y,x)\,dy$ with $p_T^*(y,x)$ given in the previous formula. We conclude that the representation formula (6.10) proves that $P_t(x,dy)$ is absolutely continuous and the density is represented by

$$
p_T(x,y) - p_T^y(x,y) = e^T \mathbb{E}\big[p_{T-\tau_T}^{X^{*,\pi}_{\tau_T}(y)}\big(x, X^{*,\pi}_{\tau_T}(y)\big)\Gamma_T^*(y) 1_{\{J_T \ge 1\}}\big].
$$

The representation for $P_T h$ can be obtained by integrating $\int h(y) p_T(x,y)\,dy$ using (6.9). □

As before, we also have that the following generalized formulas with a general Poisson process with parameter λ are valid:

$$
p_T(x,y) - p_T^y(x,y) = e^{\lambda T} \mathbb{E}\big[\lambda^{-J_T^\lambda} p_{T-\tau_{J_T^\lambda}}^{X^{*,\pi}_{\tau_{J_T^\lambda}}(y)}\big(x, X^{*,\pi}_{\tau_{J_T^\lambda}}(y)\big)\Gamma_T^*(y) 1_{\{J_T^\lambda \ge 1\}}\big].
$$

We discuss now the regularity of $p_t(x,y)$.

THEOREM 6.10. *Suppose Hypotheses 6.1, 6.2 and 6.3.*

(i) *Furthermore assume that* $(t,x,y) \to (p_t^y(x,y), \hat{a}_t(x,y))$ *is continuous in* $(0,T] \times \mathbb{R}^d \times \mathbb{R}^d$. *Then* $(t,x,y) \to p_t(x,y)$ *is continuous on* $(0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$. *Moreover,*

$$p_T(x,y) = e^T \mathbb{E}\big[p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\big(x, X_{\tau_T}^{*,\pi}(y)\big)\hat{\Gamma}_T(y)\big].$$

(ii) *Furthermore, we assume that* $x \to p_t^y(x,y)$ *is differentiable for all* $(t,y) \in \mathbb{R}_+ \times \mathbb{R}^d$ *and that the Hypothesis 3.8 is satisfied for* $\nabla_x p_t^y(x,y)$ *instead of* G. *Then the function* $x \to p_t(x,z)$ *is one time differentiable. Moreover,*

$$\nabla_x p_T(x,y) = \mathbb{E}\big[\nabla_x p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\big(x, X_{\tau_T}^{*,\pi}(y)\big)\hat{\Gamma}_T(y)\big].$$

7. Examples: Applications to stochastic differential equations. In this section, we will consider the first natural example for our previous theoretical developments, that is, the case of multidimensional diffusion processes. The forward method will need smooth coefficients and the backward method will require Hölder continuous coefficients.

7.1. *Example* 1: *The forward method for continuous SDE's with smooth coefficients.* We consider the following d-dimensional SDE:

$$X_t = x + \sum_{j=1}^m \int_0^t \sigma_j(X_s)\,dW_s^j + \int_0^t b(X_s)\,ds. \tag{7.1}$$

Here, $\sigma_j, b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma_j \in C_b^2(\mathbb{R}^d;\mathbb{R}^d)$ is uniformly elliptic (i.e., $0 < \underline{a}I \le a \le \overline{a}I$ for $\underline{a}, \overline{a} \in \mathbb{R}$ with $a = \sigma\sigma^*$), $b \in C_b^2(\mathbb{R}^d;\mathbb{R}^d)$ and W is a m-dimensional Wiener process. Under these conditions, there exists a unique pathwise solution to the above equation. Then we define the semigroup $P_t f(x) = \mathbb{E}[f(X_t)]$ which has infinitesimal generator given by $Lf(x) = \frac{1}{2}\sum_{i,j} a^{i,j}(x)\partial_{i,j}^2 f(x) + \sum_i b^i(x)\partial_i f(x)$ for $f \in C_c^\infty(\mathbb{R}^d)$ and $a^{i,j}(x) = \sum_k \sigma_k^i(x)\sigma_k^j(x)$. Clearly, $P_t f(x)$ is jointly measurable and bounded and, therefore, Hypothesis 4.1 is satisfied. We will consider the following approximation process:

$$X_t^z(x) = x + \sum_{j=1}^m \sigma_j(z) W_t^j + b(z)t,$$

which defines the semigroup

$$P_t^z f(x) = \mathbb{E}\big[f\big(X_t^z(x)\big)\big] = \int f(y) q_{ta(z)}\big(y - x - b(z)t\big)\,dy, \tag{7.2}$$

for $f \in C_c^\infty(\mathbb{R}^d)$, with jointly continuously differentiable probability kernel $p_t^z(x,y) = q_{ta(z)}(y-x-b(z)t)$. Furthermore, its associated infinitesimal operator [for $f \in C_c^2(\mathbb{R}^d)$] is given by

$$L^x f(y) = \frac{1}{2}\sum_{i,j} a^{i,j}(x)\partial_{i,j}^2 f(y) + \sum_i b^i(x)\partial_i f(y).$$

Therefore, Hypotheses 4.2 and 4.3 are clearly satisfied. Hypothesis 5.1 is clearly satisfied as $a^{i,j}, b^i \in C_b^2(\mathbb{R}^d)$ for $i, j \in \{1,\dots,d\}$. Now we proceed with the verification of Hypothesis 5.2. Using integration by parts, we have for $f \in C_c^\infty(\mathbb{R}^d)$

$$\begin{aligned}
S_t f(x) &= \int (L^y - L^x) f(y) P_t^x(x,dy) \\
&= \frac{1}{2}\sum_{i,j}\int \big(a^{i,j}(y) - a^{i,j}(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\partial_{i,j}^2 f(y)\,dy \\
&\quad + \sum_i \int \big(b^i(y) - b^i(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\partial_i f(y)\,dy \qquad (7.3)\\
&= \int dy\, f(y)\bigg(\frac{1}{2}\sum_{i,j}\partial_{i,j}^2\big(\big(a^{i,j}(y) - a^{i,j}(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\big)\bigg) \\
&\quad - \sum_i \int dy\, f(y)\partial_i\big(\big(b^i(y) - b^i(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\big).
\end{aligned}$$

In view of (A.3), we have

$$\begin{aligned}
\partial_{i,j}^2\big(\big(a^{i,j}(y) - a^{i,j}(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\big) &= \theta_t^{i,j}(x,y) q_{ta(x)}\big(y-x-b(x)t\big), \\
\partial_i\big(\big(b^i(y) - b^i(x)\big) q_{ta(x)}\big(y-x-b(x)t\big)\big) &= \rho_t^i(x,y) q_{ta(x)}\big(y-x-b(x)t\big),
\end{aligned}$$

where we define for the Hermite polynomials H (see Section A.2)

$$\begin{aligned}
\theta_t^{i,j}(x,y) &= \partial_{i,j}^2 a^{i,j}(y) + \partial_j a^{i,j}(y) h_t^i(x,y) + \partial_i a^{i,j}(y) h_t^j(x,y) \\
&\quad + \big(a^{i,j}(y) - a^{i,j}(x)\big) h_t^{i,j}(x,y), \\
\rho_t^i(x,y) &= \partial_i b^i(y) + \big(b^i(y) - b^i(x)\big) h_t^i(x,y), \\
h_t^i(x,y) &= H_{ta(x)}^i\big(y-x-b(x)t\big), \qquad (7.4)\\
h_t^{i,j}(x,y) &= H_{ta(x)}^{i,j}\big(y-x-b(x)t\big). \qquad (7.5)
\end{aligned}$$

So we obtain

$$\begin{aligned}
S_t f(x) &= \int dy\, f(y) q_{ta(x)}\big(y-x-b(x)t\big)\theta_t(x,y) \\
&= \int f(y)\theta_t(x,y) P_t^x(x,dy). \qquad (7.6)
\end{aligned}$$

Therefore, we have that

$$\theta_t(x,y)=\frac{1}{2}\sum_{i,j}\theta_t^{i,j}(x,y)-\sum_i \rho_t^i(x,y).$$

Now, we verify Hypotheses 5.2 and 5.3. We have verified the first part of Hypothesis 5.2 by the definition of θ in (7.6). In order to verify the rest of the conditions in Hypothesis 5.2, we see that by (A.3) and (A.4) with $\alpha=1$

$$p_t^x(x,y)\big|\theta_t(x,y)\big|\le C\big(\|a\|_{2,\infty}+\|b\|_{1,\infty}\big)\frac{1}{t^{1/2}}q_{ct\overline{a}}(y-x)$$

for a constant $C>1$ and $c\in(0,1)$, and consequently all the conditions in Hypothesis 5.2 are satisfied with $\rho=\frac{1}{2}+\rho_0$ and $\gamma_t(x,y)=\beta_t(x,y)=t^{\rho_0}q_{ct\overline{a}}(y-x)$. Here, $\rho_0\in(\frac{\zeta-1}{2},\frac{1}{2})$.

Similarly, Hypothesis 5.3 is satisfied under the $\xi(x)=C$ for (3.10) by using that $1_{\{\sum_{i=1}^{n-1}|y_i|\le R\}}\le 1$. For (3.11), one uses that

$$1_{\{\sum_{i=1}^{n-1}|y_i|>R\}}\le\sum_{i=1}^{n-1}\sum_{j=1}^{d}1_{\{|y_i^j|>R/(n\sqrt{d})\}}.$$

Next, one performs the change of variables $y_1=x_1$, $y_i-y_{i+1}=x_{i+1}$ for $i=1,\dots,n-2$ in the integral of (3.11) and use the inequality $1_{\{|y_i^j|>R/(n\sqrt{d})\}}\le\frac{n^2d|y_i^j|^2}{R^2}$ to obtain the following bound:

$$\begin{aligned}&\sum_{i=1}^{n-1}\sum_{j=1}^{d}\frac{n^2d}{R^2}\sup_{(y_0,y_n)\in B(z_0,r)\times B(z_n,r)}\int dx_1\cdots\int dx_{n-1}\left|\sum_{k=1}^{i}x_k^j\right|^2 q_{c\delta_0\overline{a}}(x_1-y_0)\\&\qquad\times\prod_{i=1}^{n-2}q_{c\delta_i\overline{a}}(x_{i+1})q_{c\delta_{n-1}\overline{a}}\left(y_n-\sum_{i=1}^{n-1}x_i\right).\end{aligned}\tag{7.7}$$

Without loss of generality, using a further change of variables $z_1=x_1-y_0$, we may consider the case where $y_0=0$. Next, we use the inequality $|\sum_{k=1}^{i}x_k^j|^2\le n\sum_{k=1}^{i}|x_k^j|^2$. Then one rewrites the integral in a probabilistic way using Gaussian random variables. This becomes $\mathbb{E}[|Z_k^j|^2/Z_0+\cdots+Z_{n-1}=y_n]p_{Z_0+\cdots+Z_{n-1}}(y_n)$ where Z_i is a d-dimensional Gaussian random vector with mean 0 and covariance matrix $c\delta_i\overline{a}I$. The conditional variance can be computed explicitly and the density can be bounded by its maximum value (i.e., $y_n=0$). Finally, we obtain that (7.7) is bounded by $C\frac{n^4d^2}{R^2\sqrt{\delta}}(|y_n|^2+\delta)$ with $\delta=\sum_{i=0}^{n-1}\delta_i$. Therefore, condition (3.11) will be satisfied taking $R_\varepsilon=\varepsilon^{-1/2}$ and $\xi(\delta)=\delta^{-1/2}+\delta^{1/2}$ and the upper bound in (3.11) becomes $C2^n\xi(\delta)$. We leave the details of the calculation for the reader. Therefore, the existence of the density follows.

In order to obtain further regularity, we need to verify the uniform integrability condition for $\zeta \in (1, \rho^{-1})$. In this case, we first note that due to (A.4), $|\nabla_y p_t^x(x, y)| \le \frac{C}{t^{1/2}} q_{ct\overline{a}}(y - x)$. Therefore, we may choose any $\rho \in (\frac{1}{2}, \frac{2}{3})$ and let $\zeta = \frac{1}{2(1-\rho)} > 1$. Finally, we define $\gamma_t(x, y) = t^{(1/2)(1-\zeta^{-1})} q_{ct\overline{a}}(y - x)$ and $\xi(x) = C$ in order to obtain (3.10). One also obtains (3.11) as in the proof of continuity. Therefore, the hypotheses in Theorem 5.10 are satisfied.

Now, we give the description of the stochastic representation. Given a Poisson process with parameter $\lambda = 1$ and jump times $\{\tau_i, i = 0, \ldots\}$. Given that $J_T = n$ and $t_i := \tau_i \wedge T$ we define the process $(X^\pi_{t_i})_{i=0,\ldots,n+1}$ for $\pi = \{t_i; i = 0, \ldots, n+1\}$, with $0 = t_0 < t_1 < \cdots < t_n \le t_{n+1} = T$ is then defined as compositions of $X^z(x)$ as follows:

$$X^\pi_{t_{k+1}} = X^z_{t_{k+1}-t_k}(x)|_{z=x=X^\pi_{t_k}},$$

for $k = 0, \ldots, n$. Here $X^\pi_0 = x$ and the noise used for $X^z_{t_{k+1}-t_k}(x)$ is independent of $X^\pi_{t_j}$ for all $j = 0, \ldots, k$ and of the Poisson process J.

THEOREM 7.1. *Suppose that $a \in C_b^2(\mathbb{R}^d; \mathbb{R}^d \times \mathbb{R}^d)$, $b \in C_b^2(\mathbb{R}^d; \mathbb{R}^d)$ and $\overline{a} \ge a \ge \underline{a}$. Define*

$$\Gamma_T(x) = \begin{cases} \displaystyle\prod_{j=0}^{J_T-1} \theta_{\tau_{j+1}-\tau_j}(X^\pi_{\tau_j}, X^\pi_{\tau_{j+1}}), & \text{if } J_T \ge 1, \\ 1, & \text{if } J_T = 0. \end{cases}$$

Then for any $f \in C_c^\infty(\mathbb{R}^d)$ we have

$$P_T f(x) = e^T \mathbb{E}[f(X^\pi_T)\Gamma_T(x)]$$

and, therefore,

$$p_T(x, y) = e^T \mathbb{E}[p^{X^\pi_{\tau_T}}_{T-\tau_T}(X^\pi_{\tau_T}, y)\Gamma_T(x)],$$

where $(X^\pi_t)_{t\in\pi}$ is the Euler scheme with $X^\pi_0 = x$ and random partition $\pi = \{\tau_i; i = 0, \ldots, \tau_{J_T}\} \cup \{T\}$ where $0 = \tau_0 < \cdots < \tau_{J_T} \le T$ where the random times $\{\tau_i\}_i$ are the associated jump times of the simple Poisson process J, independent of X^π with $\mathbb{E}[J_T] = T$. Moreover, $(t, x, y) \to p_t(x, y)$ is continuous on $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ and for every $t > 0$ the function $(x, y) \to p_t(x, y)$ is continuously differentiable. We also have

$$\partial_{y^i} p_T(x, y) = e^T \mathbb{E}\left[h^i_{T-\tau_T}(X^\pi_{\tau_T}, y) p^{X^\pi_{\tau_T}}_{T-\tau_T}(X^\pi_{\tau_T}, y) \prod_{j=0}^{J_T-1} \theta_{\tau_{j+1}-\tau_j}(X^\pi_{\tau_j}, X^\pi_{\tau_{j+1}})\right],$$

where h^i is defined in (7.5).

PROOF. As a consequence of Theorems 5.7, 5.9 and 5.10, we obtain most of the mentioned results. The fact that $y \to p_t(x, y)$ is continuously differentiable will follow from the backward method concerning the adjoint semigroup that we present in the following section (since a is differentiable it is also Hölder continuous so the hypotheses in the next section are verified). □

7.2. *Example* 2: *The backward method for continuous SDEs with Hölder continuous coefficients*. In this section, we will assume the same conditions as in the previous section except the regularity hypothesis on a and b. We will assume that a is a Hölder continuous function of order $\alpha \in (0, 1)$ and b is a bounded measurable function. We suppose the existence of a unique weak solution to (7.1). For further references on this matter, see [16]. The approximating semigroup is the same as in the previous section and is given by (7.2). Therefore we have, as before,

$$p_t^z(x, y) = q_{ta(z)}\big(y - x - b(z)t\big),$$
$$\phi_t^z(x) = q_{ta(z)}\big(z - x - b(z)t\big).$$

In this case, note that for fixed $z \in \mathbb{R}^d$, ϕ^z is a smooth density function and therefore $C_t(x) = 1$. Furthermore, as in the previous section, Hypotheses 4.1, 4.2 and 4.3 are satisfied. Similarly, Hypothesis 6.1 can be easily verified. We will now check Hypothesis 6.2. We define

$$\hat{\theta}_t(x, z) = \frac{1}{2}\sum_{i,j}\big(a^{i,j}(x) - a^{i,j}(z)\big)\hat{h}_t^{i,j}(x, z) - \sum_i\big(b^i(x) - b^i(z)\big)\hat{h}_t^i(x, z),$$
$$\hat{h}_t^i(x, z) = H_{ta(z)}^i\big(z - x - b(z)t\big),$$
$$\hat{h}_t^{i,j}(x, z) = H_{ta(z)}^{i,j}\big(z - x - b(z)t\big)$$

so that, by (A.3),

$$\begin{aligned}\big(L^x - L^z\big)\phi_t^z(x) &= \frac{1}{2}\sum_{i,j}\big(a^{i,j}(x) - a^{i,j}(z)\big)\partial_{i,j}^2 q_{ta(z)}\big(z - x - b(z)t\big)\\ &\quad - \sum_i\big(b^i(x) - b^i(z)\big)\partial_i q_{ta(z)}\big(z - x - b(z)t\big)\\ &= \hat{\theta}_t(x, z) q_{ta(z)}\big(z - x - b(z)t\big).\end{aligned}$$

Using (A.4) and the Hölder continuity of $a^{i,j}$, we obtain

$$\begin{aligned}&\big|\big(a^{i,j}(x) - a^{i,j}(z)\big)\partial_{i,j}^2\phi_t^z(x)\big|\\ &\quad\le C|x - z|^\alpha\big|\partial_{i,j}^2\phi_t^z(x)\big|\\ &\quad\le C\big(\big|z - x - b(z)t\big|^\alpha + \|b\|_\infty^\alpha t^\alpha\big)\big|\partial_{i,j}^2 q_{ta(z)}\big(z - x - b(z)t\big)\big|\\ &\quad\le Ct^{-(1-\alpha/2)} q_{\overline{a}t}\big(z - x - b(z)t\big).\end{aligned}$$

And using (A.4)(ii) with $\alpha = 0$, we obtain

$$\left|\left(b^i(x) - b^i(z)\right)\partial_i \phi_t^z(x)\right| \leq \frac{2}{t^{1/2}} \|b\|_\infty q_{t\overline{a}}\left(z - x - b(z)t\right).$$

Finally, we have

$$\left|\hat{\theta}_t(x,z)\right| \leq \frac{C}{t^{1-\alpha/2}}\left(1 + \|b\|_\infty\right) q_{t\overline{a}}\left(z - x - b(z)t\right).$$

We also have $\phi_t^z(x) \leq C q_{t\overline{a}}(z - x - b(z)t)$ so we obtain

$$\phi_t^z(x)\left|\hat{\theta}_t(x,z)\right| \leq \frac{C}{t^{1-\alpha/2}}\left(1 + \|b\|_\infty\right) q_{2t\overline{a}}\left(z - x - b(z)t\right).$$

We conclude that Hypothesis 6.2 is verified. The verification of Hypothesis 6.3 is done like in the previous section using $\rho \in (\frac{2-\alpha}{2}, \frac{3-\alpha}{3})$ and $\zeta = (3 - \alpha - 2\rho)^{-1} \in (1, \rho^{-1})$. Therefore, we have the following result.

PROPOSITION 7.2. *Suppose that a is Hölder continuous of order $\alpha \in (0,1)$, $\overline{a} \geq a \geq \underline{a}$ and b is measurable and bounded. Then*

$$p_T(x,y) = e^T \mathbb{E}\left[p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\left(x, X_{\tau_T}^{*,\pi}(y)\right) \prod_{j=0}^{J_T-1} \hat{\theta}_{\tau_{j+1}-\tau_j}\left(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\right)\right],$$

where $X^{*,\pi}(y)$ is the Euler scheme with $X_0^{*,\pi} = y$ and drift coefficient $-b$. Moreover, $(t,x,y) \to p_t(x,y)$ is continuous on $(0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$ and for every $(t,y) \in (0,\infty) \times \mathbb{R}^d$ the function $x \to p_t(x,y)$ is continuously differentiable. Moreover,

$$\begin{aligned}
&\partial_{x^i} p_T(x,y) \\
&\quad = -e^T \mathbb{E}\Bigg[\hat{h}^i_{T-\tau_T}\left(x, X_{\tau_T}^{*,\pi}(y)\right) p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\left(x, X_{\tau_T}^{*,\pi}(y)\right) \\
&\qquad\qquad \times \prod_{j=0}^{J_T-1} \hat{\theta}_{\tau_{j+1}-\tau_j}\left(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\right)\Bigg].
\end{aligned}$$

7.3. *Example* 3: *One-dimensional Lévy driven SDE with Hölder type coefficients*. Although we may consider various other situations where the forward and the backward method can be applied and to test their limits, we prefer to concentrate in this section on the backward method for a one-dimensional jump type SDEs driven by a Lévy process of a particular type: we assume that the intensity measure of the Lévy process is a mixture of Gaussian densities. This a quite general class as it can be verified from Schoenberg's theorem; see [15].

For this, let $N(dx, dc, ds)$ denote the Poisson random measure associated with the compensator given by $q_c(x)\,dx\,\nu(dc)\,ds$ where ν denotes a nonnegative measure on $\mathbb{R}_+ := (0,\infty)$ which satisfies the following.

HYPOTHESIS 7.3. *$\nu(\mathbb{R}_+) = \infty$ and $C_\nu := \int_{\mathbb{R}_+} c\nu(dc) < \infty$.*

We refer the reader to [8] for notation and detailed definitions on Poisson random measures. Therefore, heuristically speaking, x stands for the jump size which arises from a Gaussian distribution with random variance obtained from the measure ν.

We define $\eta_\nu(u) := \nu(u, \infty)$ and we assume that there exists some $s_* \geq 0$ and $h, C_* > 0$ such that we have the following.

HYPOTHESIS 7.4. *$\int_0^\infty e^{-u}\eta_\nu(\frac{u}{s})\,du \geq C_* s^h \int_0^\infty e^{-u}\eta_\nu(u)\,du \ \forall s \geq s_*$.*

For example, if $\nu(dc) = 1_{(0,1]}(c)c^{-(1+\beta)}\,dc$ with $0 < \beta < 1$ then Hypothesis 7.3 is satisfied and Hypothesis 7.4 is satisfied with $h = \beta$.

$\widetilde{N}(dx, dc, ds) = N(dx, dc, ds) - q_c(x)\,dx\,\nu(dc)\,ds$ denotes the compensated Poisson random measure. We also define the following auxiliary processes and driving process Z:

$$V_t = \int_0^t \int_{\mathbb{R}_+\times\mathbb{R}} cN(dx, dc, ds),$$

$$Z_t = \int_0^t \int_{\mathbb{R}_+\times\mathbb{R}} xN(dx, dc, ds),$$

$$N_\nu(dx, ds) = \int_{\mathbb{R}_+} N(dx, dc, ds).$$

With a slight variation of some classical proofs (see, e.g., Chapter 2 in [1]) one can obtain the following generalization of the Lévy–Khinchine formula.

PROPOSITION 7.5. *Assume Hypothesis 7.3. Let $h : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ be such that $|\int_{\mathbb{R}\times\mathbb{R}_+}(e^{i\theta h(x,c)} - 1)q_c(x)\,dx\,d\nu(c)| < \infty$. Then the stochastic process $U_t(h) := \int_0^t\int_{\mathbb{R}_+\times\mathbb{R}} h(x, c)N(dx, dc, ds)$ has independent increments with characteristic function given by*

$$\mathbb{E}\bigl[\exp\bigl(i\theta U_t(h)\bigr)\bigr] = \exp\biggl(t\int_{\mathbb{R}\times\mathbb{R}_+}\bigl(e^{i\theta h(x,c)} - 1\bigr)q_c(x)\,dx\,d\nu(c)\biggr)$$

the density of Z_t at y can be written as $\mathbb{E}[q_{V_t}(y)]$.

PROOF. The first part of the proof is classical, while in order to obtain the representation for the density of Z_t, one takes $h(x, c) = x$ to obtain the characteristic function associated with Z_t under Hypothesis 7.3. On the other hand, one only needs to compute the characteristic function associated with the density function $\mathbb{E}[q_{V_t}(y)]$ to finish the proof. □

Notice that due to Hypothesis 7.3 we have that

$$\mathbb{E}[Z_t^2] = t\int_{\mathbb{R}\times\mathbb{R}_+} |u|^2 q_c(u)\nu(dc)\,du = t\int_{\mathbb{R}_+} c\nu(dc) < \infty. \tag{7.8}$$

Therefore, Z is a Lévy process of finite variance. $N_\nu(dx, ds)$ is a Poisson random measure with compensator $\mu_\nu(dx)\,ds := \int_{\mathbb{R}_+} q_c(x)\nu(dc)\,dx\,ds$ and we denote by $\widetilde{N}_\nu(dx, ds)$ the compensated Poisson random measure. Then we consider the solution of the following stochastic differential equation driven by Z and its corresponding approximation obtained after freezing the jump coefficient. That is,

$$X_t^\nu(x) = x + \int_0^t \int_{\mathbb{R}} \sigma\big(X_{s-}^\nu(x)\big)u\widetilde{N}_\nu(ds, du), \tag{E_ν}$$

$$X_t^{\nu,z}(x) = x + \int_0^t \int_{\mathbb{R}} \sigma(z)u\widetilde{N}_\nu(ds, du). \tag{E_ν^z}$$

We assume that $\sigma : \mathbb{R} \to \mathbb{R}$ verifies the following conditions.

HYPOTHESIS 7.6. (i) *There exists* $\underline{\sigma}, \overline{\sigma} > 0$ *such that* $\underline{\sigma} \le \sigma(x) \le \overline{\sigma}$ *for all* $x \in \mathbb{R}$.

(ii) *There exists* $\alpha \in (0, 1]$ *such that* $|\sigma(x) - \sigma(y)| \le C_\alpha |x - y|^\alpha$.

If $\alpha = 1$, then (E_ν) has a unique solution. Here, rather than entering into the discussion of existence and uniqueness results for other values of $\alpha \in (0, 1]$, we refer the reader to a survey article by Bass and the references therein (see [4]). Therefore, from now on, we suppose that a unique weak solution to (E_ν) exists so that $P_t^\nu f(x) = \mathbb{E}[f(X_t^\nu(x))]$ is a semigroup with infinitesimal operator [note that $\int u\mu_\nu(du) = 0$]

$$L^\nu f(x) = \int_{\mathbb{R}} \big(f\big(x + \sigma(x)u\big) - f(x)\big)\mu_\nu(du).$$

Therefore, Hypothesis 4.1 is clearly satisfied.

Similarly, $X^{\nu,z}(x)$, defines a semigroup $P_t^\nu f(x) = \mathbb{E}[f(X_t^{\nu,z}(x))]$ with infinitesimal operator

$$L^{\nu,z} f(x) = \int_{\mathbb{R}} \big(f\big(x + \sigma(z)u\big) - f(x)\big)\mu_\nu(du). \tag{7.9}$$

Our aim is to give sufficient conditions in order that the law of $X_t^\nu(x)$ is absolutely continuous with respect to the Lebesgue measure and to represent the density $p_t(x, y)$ using the backward method as introduced in Section 6. In order to proceed with the verification of Hypothesis 4.2, we need to prove the following auxiliary lemma.

LEMMA 7.7. *Suppose that Hypotheses 7.3 and 7.4 holds for some $h > 0$. Then for every $p > 0$ there exists a constant C such that for every $t > 0$*

$$\mathbb{E}[V_t^{-p}] \leq Ct^{-p/h}. \tag{7.10}$$

PROOF. Recall that the Laplace transform of V_t is given by

$$\mathbb{E}[e^{-aV_t}] = \exp\left(-t\int_{\mathbb{R}_+}(1-e^{-ac})\nu(dc)\right).$$

We use the change $s' = sV_t$ and we obtain

$$\int_0^\infty s^{p-1}e^{-sV_t}\,ds = c_pV_t^{-p}$$

with $c_p = \int_0^\infty s^{p-1}e^{-s}\,ds$. It follows that

$$c_p\mathbb{E}[V_t^{-p}] = \int_0^\infty s^{p-1}\mathbb{E}[e^{-sV_t}]\,ds = \int_0^\infty s^{p-1}\exp\left(-t\int_{\mathbb{R}_+}(1-e^{-sc})\nu(dc)\right)ds.$$

For $s > s_*$ we have using the integration by parts formula and the change of variables $sc = u$,

$$\int_0^\infty(1-e^{-sc})\nu(dc) = \int_0^\infty due^{-u}\eta_\nu\left(\frac{u}{s}\right) \geq C_*s^h\int_0^\infty due^{-u}\eta_\nu(u) =: s^h\alpha_\nu$$

with $\alpha_\nu \in \mathbb{R}_+$. Therefore, again by change of variables, we have that

$$\begin{aligned}\int_{s_*}^\infty s^{p-1}\exp\left(-t\int_{\mathbb{R}_+}(1-e^{-sc})\nu(dc)\right)ds &\leq \int_{s_*}^\infty s^{p-1}e^{-ts^h\alpha_\nu}\,ds \\ &\leq t^{-p/h}C(\nu,p,h)\end{aligned}$$

with

$$C(\nu,p,h) = h^{-1}\int_0^\infty u^{-(1-p/h)}e^{-u\alpha_\nu}\,du < \infty.$$

Since $\int_0^{s_*}s^{p-1}\,ds = \frac{1}{p}s_*^p$ the conclusion follows by taking $s_* = t^{-1/h}$. $\square$

Now we can verify Hypothesis 4.2. For this, we need to compute as explicitly as possible the density $p_t^z(x,y)$ of the law of $X_t^{\nu,z}(x)$. In fact, the following is a corollary of Proposition 7.5 and the previous lemma which is used together with Lemma A.2 in order to obtain the needed uniform integrability properties.

COROLLARY 7.8. *Suppose that Hypotheses 7.3 and 7.6 are verified. Then the law of $X_t^{\nu,z}(x)$ is absolutely continuous with respect to the Lebesgue measure with strictly positive continuous density given by*

$$p_t^z(x,y) = \mathbb{E}[q_{\sigma^2(z)V_t}(x-y)].$$

Therefore, for each fixed $(t,z) \in (0,T] \times \mathbb{R}$, we have that $p_t^z \in C_b^2(\mathbb{R}\times\mathbb{R})$ and $p_t^z(x,y)$ is jointly continuous in (t,z,x,y).

Note that due to the above result Hypothesis 4.2 is satisfied and $\phi_t^y(x) = \mathbb{E}[q_{\sigma^2(z)V_t}(x-y)]$. Furthermore, as it is usually the case Hypotheses 4.3 and 6.1(0) are trivially satisfied. For Hypothesis 6.1(i), one only needs to apply Corollary 7.8. Hypothesis 6.1(ii) follows from the joint continuity of $p_t^z(x,y)$ and Hypothesis 6.1(iii) follows from the regularity of $p_t^z(x,y)$ as stated in the above Corollary 7.8 and (7.8).

We are now ready to proceed and verify Hypotheses 6.2 and 6.3. We have by (7.9), $\int uq_c(u)\,du = 0$ and properties of convolution that

$$\begin{aligned}
(L^{\nu,x} - L^{\nu,z})\phi_t^z(x) &= \int_{\mathbb{R}_+\times\mathbb{R}} \big(\phi_t^z(x+\sigma(x)u) - \phi_t^z(x+\sigma(z)u)\big)q_c(u)\nu(dc)\,du \\
&= \int_{\mathbb{R}_+\times\mathbb{R}} \big(\mathbb{E}\big[q_{\sigma^2(z)V_t}(x-z+\sigma(x)u)\big] \\
&\qquad - \mathbb{E}\big[q_{\sigma^2(z)V_t}(x-z+\sigma(z)u)\big]\big)q_c(u)\nu(dc)\,du \\
&= \int_{\mathbb{R}} \mathbb{E}\big[q_{\sigma^2(x)c+\sigma^2(z)V_t}(x-z) - q_{\sigma^2(z)c+\sigma^2(z)V_t}(x-z)\big]\nu(dc).
\end{aligned}$$

In particular, Hypothesis 6.2 holds with

$$\begin{aligned}
\hat{\theta}_t(x,y) &= \frac{1}{\mathbb{E}[q_{\sigma^2(y)V_t}(x-y)]} \\
&\quad\times \left\{\int_{\mathbb{R}_+} \mathbb{E}\big[q_{\sigma^2(x)c+\sigma^2(y)V_t}(x-y) - q_{\sigma^2(y)c+\sigma^2(y)V_t}(x-y)\big]\nu(dc)\right\}.
\end{aligned} \tag{7.11}$$

THEOREM 7.9. *Suppose that Hypotheses 7.3, 7.4 and 7.6 hold with $h > 1 - \frac{\alpha}{2}$. Then the law of $X_T^\nu(x)$ is absolutely continuous with respect to the Lebesgue measure and its density $p_T(x,y)$ satisfies*

$$p_T(x,y) = e^T\mathbb{E}\left[p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\big(x, X_{\tau_T}^{*,\pi}(y)\big)\prod_{j=0}^{J_T-1}\hat{\theta}_{\tau_{j+1}-\tau_j}\big(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\big)\right],$$

where $X_t^{*,\pi}(y)$ is the Euler scheme given in the backward method starting at $X_0^{*,\pi}(y) = y$. Moreover, $(t,x,y) \to p_t(x,y)$ is continuous on $(0,\infty)\times\mathbb{R}\times\mathbb{R}$ and for every $(t,y) \in (0,\infty)\times\mathbb{R}$ the function $x \to p_t(x,y)$ is differentiable and

$$\partial_x p_T(x,y) = e^T\mathbb{E}\left[\partial_x p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}\big(x, X_{\tau_T}^{*,\pi}(y)\big)\prod_{j=0}^{J_T-1}\hat{\theta}_{\tau_{j+1}-\tau_j}\big(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)\big)\right].$$

PROOF. We have already verified Hypotheses 6.1, 6.2 and the differentiability of p_t^z. It remains to verify the hypotheses in Lemma 3.9 and Theorem 3.10. For this, we have to estimate

$$\big|\hat{\theta}_t(x,y)\big|\phi_t^y(x) \le \int_{\mathbb{R}_+} \mathbb{E}\big[\big|q_{\sigma^2(x)c+\sigma^2(y)V_t}(x-y) - q_{\sigma^2(y)c+\sigma^2(y)V_t}(x-y)\big|\big]\nu(dc).$$

Let us denote $a = x - y$ and

$$s' = \sigma^2(y)c + \sigma^2(y)V_t, \qquad s'' = \sigma^2(x)c + \sigma^2(y)V_t.$$

We assume that $s' \le s''$ (the other case is similar) and note the inequality (with $a, b, b', c > 0$)

$$b \ge b' \quad \Rightarrow \quad \frac{a + cb}{a + cb'} \le \frac{b}{b'}.$$

From this inequality, we obtain

$$\begin{aligned} &\frac{s''}{s'} \le \frac{\overline{\sigma}^2}{\underline{\sigma}^2} \quad \text{and} \\ &\frac{|s'' - s'|}{s'} \le \frac{c|\sigma^2(x) - \sigma^2(y)|}{\sigma^2(x)c + \sigma^2(y)V_t} \le \frac{cC_\alpha |a|^\alpha}{\underline{\sigma}^2(c + V_t)}, \end{aligned} \tag{7.12}$$

where C_α is the Hölder constant of σ^2. Finally, from Lemma A.3 and (7.12) this gives

$$\begin{aligned} |q_{s''}(a) - q_{s'}(a)| &\le CC_\alpha \frac{\overline{\sigma}^2}{\underline{\sigma}^2} \frac{c|a|^\alpha}{\underline{\sigma}^2(c + V_t)} q_{\overline{\sigma}^2(c+V_t)}(a) \\ &\le CC_\alpha \frac{\overline{\sigma}^{2+\alpha}}{\underline{\sigma}^4} \frac{c}{(c + V_t)^{1-\alpha/2}} q_{(\overline{\sigma}^2/2)(c+V_t)}(a). \end{aligned}$$

Returning to our main proof, we obtain (with $C = CC_\alpha \overline{\sigma}^{2+\alpha} \underline{\sigma}^{-4}$)

$$\begin{aligned} &|\hat{\theta}_t(x, y)| \phi_t^y(x) \\ &\quad \le C \int_{\mathbb{R}_+} \mathbb{E}[(c + V_t)^{-(1-\alpha/2)} q_{(\overline{\sigma}^2/2)(c+V_t)}(x - y)] c\nu(dc). \end{aligned} \tag{7.13}$$

A first step is to obtain estimates for the right-hand side of the above inequality, so as to be able to define γ. For this, we define

$$\begin{aligned} &g_t(x, y) = \int_{\mathbb{R}_+} \mathbb{E}[V_t^{-(1-\alpha/2)} q_{(\overline{\sigma}^2/2)(c+V_t)}(x - y)] \overline{\nu}(dc), \\ &\overline{\nu}(dc) = \frac{1(c > 0)}{C_\nu} c\nu(dc), \qquad C_\nu = \int_{\mathbb{R}_+} c\nu(dc). \end{aligned} \tag{7.14}$$

We denote

$$\chi = \left(1 - \frac{\alpha}{2}\right)\zeta + \frac{\zeta - 1}{2} \quad \text{and} \quad \rho = \frac{\chi}{h}.$$

Since $1 - \frac{\alpha}{2} < h$, there exists $\zeta \in (1, \rho^{-1})$ with $\rho \in (0, 1)$. We fix such a ζ. We define now

$$\gamma_t(x, y) := t^{\chi/h} g_t(x, y)$$

and we notice that by (7.13)

$$|\hat{\theta}_t(x,y)|\phi_t^y(x) \le C g_t(x,y) = C t^{-\chi/h}\gamma_t(x,y) = C t^{-\rho}\gamma_t(x,y).$$

We also define $G_t(x,y) := \mathbb{E}[q_{\sigma^2(y)V_t}(x-y)] = p_t^y(x,y)$, and we use Lemma A.2 in order to define γ^3

$$\begin{aligned}|\partial_x p_t^y(x,y)| &\le \mathbb{E}\left[\frac{|x-y|}{\sigma^2(y)V_t} q_{\sigma^2(y)V_t}(x-y)\right] \le C\mathbb{E}\left[V_t^{-1/2} q_{\sigma^2(y)V_t}(x-y)\right] \\ &=: C t^{-1/(2h)}\gamma_t^3(x,y).\end{aligned}$$

With these definitions, we need to check that (3.10) and (3.11) holds. We verify the former as the latter is similar to (7.7) if one uses (7.16) at the end of the calculation. To verify (3.10), it is enough to prove that for $n \in \mathbb{N}$, $\delta_i > 0$, $i = 1,\dots,n$

$$\sup_{y_0, y_{n+1}\in\mathbb{R}^d} \int dy_1 \cdots \int dy_n \prod_{i=0}^n \gamma_{\delta_i}(y_i, y_{i+1})^\zeta \le \frac{C^n}{(\delta_1+\cdots+\delta_n)^{1/(2h)}}, \tag{7.15}$$

where C is a constant which depends on ζ, p, h and s_* which appear in Hypothesis 7.4 and in (7.10). Notice first that for every $a > 0$ and $x \in \mathbb{R}$ one has for a positive constant C,

$$(q_a(x))^\zeta = C a^{-(\zeta-1)/2} q_{a/\zeta}(x).$$

Using Hölder's inequality and the definition of χ, we obtain

$$\begin{aligned} g_{\delta_i}(y_i,y_{i+1})^\zeta &\le \int_{\mathbb{R}_+} \mathbb{E}\left[V_{\delta_i}^{-(1-\alpha/2)\zeta} q_{(\overline{\sigma}^2/2)(c+V_{\delta_i})}(y_i-y_{i+1})^\zeta\right]\overline{\nu}(dc) \\ &\le C\int_{\mathbb{R}_+} \mathbb{E}\left[V_{\delta_i}^{-\chi} q_{(\overline{\sigma}^2/2(\zeta))(c+V_{\delta_i})}(y_i-y_{i+1})\right]\overline{\nu}(dc).\end{aligned}$$

We consider $(V_t^i)_{t\ge0}$, $i=1,\dots,n$ to be independent copies of $(V_t)_{t\ge0}$ and we write

$$\begin{aligned}&\int dy_1\cdots\int dy_n \prod_{i=0}^n g_{\delta_i}(y_i,y_{i+1})^\zeta \\ &\quad\le C^{-n}\mathbb{E}\Bigg[\prod_{i=1}^n (V_{\delta_i}^i)^{-\chi}\int\overline{\nu}(dc_1)\cdots\int\overline{\nu}(dc_n)\int dy_1\cdots\int dy_n \\ &\qquad\qquad \times\prod_{i=1}^n q_{(\overline{\sigma}^2/(2\zeta))(c_i+V_{\delta_i}^i)}(y_i-y_{i+1})\Bigg] \\ &\quad= C^{-n}\mathbb{E}\Bigg[\prod_{i=1}^n (V_{\delta_i}^i)^{-\chi}\int\overline{\nu}(dc_1)\cdots\end{aligned}$$

$$\times \int \overline{\nu}(dc_n) q_{(\overline{\sigma}^2/(2\zeta))\sum_{i=1}^n (c_i + V^i_{\delta_i})}(y_0 - y_{n+1})\bigg]$$

$$\leq C^{-n}\left(\mathbb{E}\left[\prod_{i=1}^n (V^i_{\delta_i})^{-2\chi\zeta}\right]\right)^{1/2}$$

$$\times \left(\mathbb{E}\left[\int \overline{\nu}(dc_1)\cdots\int \overline{\nu}(dc_n) q^2_{(\overline{\sigma}^2/(2\zeta))\sum_{i=1}^n (c_i + V^i_{\delta_i})}(y_0 - y_{n+1})\right]\right)^{1/2}.$$

Notice that V is a Lévy processes, therefore, $\sum_{i=1}^n V^i_{\delta_i}$ has the same law as $V_{\delta_1+\cdots+\delta_n}$ so

$$\mathbb{E}\left[\int \overline{\nu}(dc_1)\cdots\int \overline{\nu}(dc_n) q^2_{(\overline{\sigma}^2/(2\zeta))\sum_{i=1}^n (c_i + V^i_{\delta_i})}(y_0 - y_{n+1})\right]$$

$$\begin{aligned}
&= \mathbb{E}\left[\int \overline{\nu}(dc_1)\cdots\int \overline{\nu}(dc_n) q^2_{(\overline{\sigma}^2/(2\zeta))(\sum_{i=1}^n c_i + V_{\delta_1+\cdots+\delta_n})}(y_0 - y_{n+1})\right] \\
&\leq \mathbb{E}\left[\int \overline{\nu}(dc_1)\cdots\int \overline{\nu}(dc_n)\frac{2\zeta}{\overline{\sigma}^2(\sum_{i=1}^n c_i + V_{\delta_1+\cdots+\delta_n})}\right] \\
&\leq \frac{2\zeta}{\overline{\sigma}^2}\mathbb{E}\left[\frac{1}{V_{\delta_1+\cdots+\delta_n}}\right] \leq \frac{C\zeta}{\overline{\sigma}^2(\delta_1+\cdots+\delta_n)^{1/h}}.
\end{aligned} \tag{7.16}$$

The last inequality is a consequence of (7.10). Again by (7.10),

$$\left(\mathbb{E}\left[\prod_{i=1}^n (V^i_{\delta_i})^{-2\chi\zeta}\right]\right)^{1/2} = \prod_{i=1}^n (\mathbb{E}[(V^i_{\delta_i})^{-2\chi\zeta}])^{1/2} \leq C^n \prod_{i=1}^n \delta_i^{-\chi\zeta/h}$$

so (7.15) is proved. □

We give now a probabilistic representation for the density of the solution of (E_ν). We consider the Poisson process J of parameter $\lambda = 1$ with jump times $\{\tau_i; i \in \mathbb{N}\}$, and a sequence of i.i.d. standard normal random variables $(\Delta_j)_{j\in\mathbb{N}}$.

First, note that using the mean value theorem, we can rewrite (7.11) as

$$\hat{\theta}_t(x,y)\phi^y_t(x) = C_\nu \int_{\mathbb{R}_+}\int_{\underline{\sigma}^2}^{\overline{\sigma}^2} 1_{\{\sigma^2(x)\wedge\sigma^2(y)\leq u\leq \sigma^2(x)\vee\sigma^2(y)\}}\operatorname{sgn}_\sigma(x,y)$$
$$\times\, \mathbb{E}[\partial_t q_{uc+\sigma^2(y)V_t}(x-y)]\,du\,\overline{\nu}(dc).$$

Here, we define

$$\operatorname{sgn}_\sigma(x,y) = \begin{cases} 1, & \text{if } \sigma^2(x) > \sigma^2(y), \\ -1, & \text{if } \sigma^2(x) \leq \sigma^2(y). \end{cases}$$

Therefore, we have that if we consider $U_i \sim \mathrm{Unif}[\underline{\sigma}^2, \overline{\sigma}^2]$ an i.i.d. sequence of random variables independent of all other random variables, we can represent

$\hat{\theta}_t(x,y)\phi_t^y(x)$ as

$$\begin{aligned}
&\hat{\theta}_t(x,y)\phi_t^y(x)\\
&\quad = \frac{C_\nu}{2}(\bar{\sigma}^2 - \underline{\sigma}^2)\\
&\qquad \times \int_{\mathbb{R}} \operatorname{sgn}_\sigma(x,y)\mathbb{E}\big[1_{\{\sigma^2(x)\wedge\sigma^2(y)\le U\le\sigma^2(x)\vee\sigma^2(y)\}}\\
&\qquad\qquad \times h^{1,1}_{Uc+\sigma^2(y)V_t}(x,y)q_{Uc+\sigma^2(y)V_t}(x-y)\big]\overline{\nu}(dc).
\end{aligned}$$

Here, $h^{1,1}$ is the Hermite polynomial defined in (7.5). In this case, the approximating Markov chain is defined as $Y_0^{*,\pi}(y) = y$

$$Y^{*,\pi}_{\tau_{i+1}}(y) = Y^{*,\pi}_{\tau_i}(y) + \Delta_i\big(U_i Z_i + \sigma^2\big(Y^{*,\pi}_{\tau_i}(y)\big)(V_{\tau_{i+1}} - V_{\tau_i})\big)^{1/2}.$$

The corresponding weight is given by

$$\begin{aligned}
\Gamma_T^\pi(y) &= \left(\frac{C_\nu}{2}(\bar{\sigma}^2 - \underline{\sigma}^2)\right)^{J_T}\\
&\qquad \times \prod_{i=1}^{J_T} \operatorname{sgn}_\sigma\big(Y^{*,\pi}_{\tau_{i+1}}(y), Y^{*,\pi}_{\tau_i}(y)\big)\\
&\qquad\qquad \times 1_{\{\sigma^2(Y^{*,\pi}_{\tau_{i+1}}(y))\wedge\sigma^2(Y^{*,\pi}_{\tau_i}(y))\le U_i\le\sigma^2(Y^{*,\pi}_{\tau_{i+1}}(y))\vee\sigma^2(Y^{*,\pi}_{\tau_i}(y))\}}\\
&\qquad\qquad \times h^{1,1}_{U_i Z_i + \sigma^2(Y^{*,\pi}_{\tau_i}(y))V_t}\big(Y^{*,\pi}_{\tau_{i+1}}(y), Y^{*,\pi}_{\tau_i}(y)\big).
\end{aligned}$$

COROLLARY 7.10. *Under the hypothesis of Theorem* 7.9, *we have*

$$p_T(x,y) = e^T\mathbb{E}\big[p_{T-\tau_T}^{Y^{*,\pi}_{\tau_T}(y)}\big(x, Y^{*,\pi}_{\tau_T}(y)\big)\Gamma_T^\pi(y)\big],$$
$$\partial_x p_T(x,y) = e^T\mathbb{E}\big[\partial_x p_{T-\tau_T}^{Y^{*,\pi}_{\tau_T}(y)}\big(x, Y^{*,\pi}_{\tau_T}(y)\big)\Gamma_T^\pi(y)\big].$$

7.3.1. *Examples of Lévy measures.* We conclude this section with two examples of Lévy measures that satisfy Hypotheses 7.3 and 7.4.

EXAMPLE 7.11. Let $c_k = k^{-\rho}$ for some $\rho > 1$ and define the discrete measure $\nu(dc) = \sum_{k=1}^\infty \delta_{c_k}(dc)$.

We verify that all the hypotheses required in Section 7.3 are satisfied in this example. First of all, we consider Hypothesis 7.3. Clearly, $\nu(\mathbb{R}_+) = \infty$, and if $\rho > 1$ then $\int c\, d\nu(c) < \infty$.

Now we verify Hypothesis 7.4. One has $\eta_\nu(a) = \operatorname{card}\{k : c_k > a\} = [a^{-1/\rho}] - 1(a^{-1/\rho} \in \mathbb{N})$ for $a > 0$. We define $\eta'_\nu(a) = a^{-1/\rho}$. Then clearly η'_ν satisfies the Hypothesis 7.4 with $h = \frac{1}{\rho}$. Furthermore, $s^{-1/\rho}\eta_\nu(\frac{u}{s}) - \eta'_\nu(u)$ converges uniformly to zero as $s \to \infty$. Then as $\eta_\nu \le \eta'_\nu$ then Hypothesis 7.4 is verified for η_ν with $h = \frac{1}{\rho}$. So we may use Corollary 7.10 or Theorem 7.9 for equations with α-Hölder coefficient σ with $\alpha > \frac{2(\rho-1)}{\rho}$ and the Lévy measure $\mu_\nu(du) = q_\nu(u)\,du$ with

$$q_\nu(u) = \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{\infty} k^{\rho/2} e^{-k^\rho u^2/2}.$$

EXAMPLE 7.12. We consider the measure $\nu(dc) = 1_{[0,1]}(c) c^{-(1+\beta)}\,dc$ with $\frac{1}{2} < \beta < 1$. Then $\nu(\mathbb{R}_+) = \infty$ and $\int c\,d\nu(c) < \infty$. One has $\eta_\nu(a) = \frac{1}{\beta}(a^{-\beta} - 1)$ for $a \in (0,1)$ so that Hypothesis 7.4 holds with $h = \beta \in [0,1)$. Therefore, Corollary 7.10 or Theorem 7.9 can be applied for an α-Hölder coefficient σ with $\alpha \in (2(1-\beta), 1)$. One may also compute

$$q_\nu(u) = \frac{2^{2\beta}}{u^{1-\beta}\sqrt{2\pi}} \int_{u^2/2}^{\infty} y^{-1-2\beta} e^{-y}\,dy$$

so we have the following asymptotic behavior around 0 for q_ν:

$$\lim_{u \to 0} u^{2\beta} q_\nu(u) = \frac{2^{2\beta}}{\sqrt{2\pi}} \int_0^{\infty} y^{\beta-1} e^{-y}\,dy < \infty.$$

Therefore, the Lévy measure generated by this example is of stable-like behavior around 0.

8. Some conclusions and final remarks. The parametrix method has been a successful method in the mathematical analysis of fundamental solutions of PDEs and we wanted to show the reader the possibility of other directions of possible generalization. One of them is to use the current set-up to introduce stochastic processes representing a variety of different operators which are generated by a parametrized operator L^z. Therefore, allowing the stochastic representation for various nontrivial operators.

The adjoint method we introduced here seems to allow for the analysis of the regularity of the density requiring Hölder continuity of the coefficients through an explicit expression of the density.

Finally, the stochastic representation can be used for simulation purposes. In that case, the variance of the estimators explode due to the instability of the weight function θ in the forward method or $\hat{\theta}$ in the backward method. In fact, the representations presented here have a theoretical infinite variance although the mean is finite. In that respect, the way that the Poisson process and the exponential jump times appear maybe considered somewhat arbitrary. In fact, one can think of various other representations which may lead to variance reduction methods. Preliminary simulations show that different interpretations of the time integrals in the parametrix method may lead to finite variance simulation methods. Many of these issues will be taken up in future work.

APPENDIX

A.1. On some Beta type coefficients. For $t_0 \in \mathbb{R}$, $a \in [0,1)$, $b > -1$ and $n \in \mathbb{N}$, define

$$c_n(t_0,a,b) := \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n t_n^b \prod_{j=0}^{n-1} (t_j - t_{j+1})^{-a}.$$

LEMMA A.1. *Let $a \in [0,1)$ and $b > -1$. Then we have*

$$c_n(t_0,a,b) \le t_0^{b+n(1-a)} \frac{\Gamma(1+b)\Gamma^n(1-a)}{[1+b+n(1-a)]!} \qquad \text{for } n \ge \frac{1-b}{1-a}.$$

In particular, for $b=0$,

$$c_n(t_0,a) := c_n(t_0,a,b) \le t_0^{n(1-a)} \frac{\Gamma^n(1-a)}{[1+n(1-a)]!} \qquad \text{for } n \ge (1-a)^{-1}. \tag{A.1}$$

PROOF. Let $b > -1$ and $0 \le a < 1$ and use the change of variable $s = ut$ so that

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

where $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} \, dt$ is the standard Beta function and $b+1-a > -1$. Using this repeatedly, we obtain

$$\begin{aligned} c_n(t_0,a,b) &= t_0^{b+n(1-a)} \prod_{i=0}^{n-1} B\bigl(1+b+i(1-a), 1-a\bigr) \\ &= t_0^{b+n(1-a)} \frac{\Gamma(1+b)\Gamma^n(1-a)}{\Gamma(1+b+n(1-a))}. \end{aligned}$$

The last equality being a consequence of the identity $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$. The function $\Gamma(x)$ is increasing for $x \ge 2$ so the result follows. Letting $b = 0$, we get (A.1). $\square$

A.2. Some properties of Gaussian type kernels. In this section, we introduce some preliminary estimates concerning Gaussian kernels. We consider a d dimensional square symmetric nonnegative definite matrix a. We assume that $0 < \underline{a} I \le a \le \overline{a} I$ for $\underline{a}, \overline{a} \in \mathbb{R}$ and we define $\rho_a := \frac{\overline{a}}{\underline{a}}$. The Gaussian density of mean zero and covariance matrix a is denoted by

$$q_a(y) = \frac{1}{(2\pi)^{d/2}\sqrt{\det a}} \exp\biggl(-\frac{1}{2}\langle a^{-1}y, y\rangle\biggr).$$

For a strictly positive real number λ, we abuse the notation, denoting by $q_\lambda(y) \equiv q_a(y)$ for $a = \lambda I$ where I is the identity matrix. In particular, q_1 is the standard Gaussian kernel on $\mathbb{R}^d$. Then we have the following immediate inequalities:

$$\text{(A.2)}\qquad \begin{aligned} &\text{(i)} \quad q_s(y) \le \left(\frac{t}{s}\right)^{d/2} q_t(y) \qquad \forall s < t,\\ &\text{(ii)} \quad \rho_a^{-d/2} q_{\underline{a}}(y) \le q_a(y) \le \rho_a^{d/2} q_{\overline{a}}(y). \end{aligned}$$

We define for $a \in \mathbb{R}^{d\times d}$, the Hermite polynomials in $\mathbb{R}^d$ as

$$H_a^i(y) = -(a^{-1}y)^i, \qquad H_a^{i,j}(y) = (a^{-1}y)^i (a^{-1}y)^j - (a^{-1})^{i,j}.$$

Direct computations give

$$\text{(A.3)}\qquad \partial_i q_a(y) = H_a^i(y) q_a(y), \qquad \partial_{i,j}^2 q_a(y) = H_a^{i,j}(y) q_a(y).$$

We will use the following basic estimates.

LEMMA A.2. *For $\alpha \in [0,1]$, we have for all $i, j \in \{1,\dots,d\}$, $y \in \mathbb{R}^d$ and $t > 0$,*

$$\text{(A.4)}\qquad \begin{aligned} &\text{(i)} \quad |y|^\alpha \left|\partial_{i,j}^2 q_{ta}(y)\right| \le C_a \frac{1}{t^{1-\alpha/2}} q_{t\overline{a}/2}(y) \quad \textit{and}\\ &\text{(ii)} \quad |y|^\alpha \left|\partial_i q_{ta}(y)\right| \le C_a' \frac{1}{t^{(1-\alpha)/2}} q_{t\overline{a}/2}(y) \end{aligned}$$

with

$$C_a = (2\rho_a)^{d/2} \underline{a}^{-1} (4\overline{a})^{\alpha/2} (4\rho_a + 1), \qquad C_a' = \underline{a}^{-1}(4\overline{a})^{(1+\alpha)/2}(2\rho_a)^{d/2}.$$

PROOF. We have

$$\left|H_{ta}^{i,j}(y)\right| \le \frac{|y|^2}{\underline{a}^2 t^2} + \frac{1}{\underline{a}t}$$

so that

$$|y|^\alpha \left|\partial_{i,j}^2 q_{ta}(y)\right| \le \frac{1}{\underline{a}t^{1-\alpha/2}} \frac{|y|^\alpha}{t^{\alpha/2}} \left(\frac{|y|^2}{\underline{a}t} + 1\right) q_{ta}(y).$$

We use (A.2) and we obtain

$$q_{ta}(y) \le \rho_a^{d/2} q_{t\overline{a}}(y) = (2\rho_a)^{d/2} \exp\left(-\frac{|y|^2}{4t\overline{a}}\right) q_{t\overline{a}/2}(y).$$

We may find a constant c_α such that $v^\lambda e^{-v} \le c_\alpha$ for every $0 \le \lambda \le 2+\alpha$. Using this inequality twice with $\lambda = \frac{2+\alpha}{2}$ and $\lambda = \frac{\alpha}{2}$ and for $v = \frac{1}{4t\overline{a}}|y|^2$, we obtain

$$|y|^\alpha \left|\partial_{i,j}^2 q_{ta}(y)\right| \le \frac{(2\rho_a)^{d/2}}{\underline{a}t^{1-\alpha/2}} (4\overline{a})^{\alpha/2} (4\rho_a + 1) q_{t\overline{a}/2}(y).$$

The proof of (ii) is similar. $\square$

LEMMA A.3. *Let* $0 < s' \le s''$ *and* $y \in \mathbb{R}$ *then*

$$|q_{s''}(y) - q_{s'}(y)| \le \sqrt{\frac{2s''}{s'}} q_{s''}(y) \left(\frac{s'' - s'}{s'} \right).$$

PROOF. Using the fact that q solves the heat equation, we have using (A.2) and (A.4)

$$\begin{aligned} |q_{s''}(y) - q_{s'}(y)| &\le \int_{s'}^{s''} |\partial_s q_s(y)|\, ds = \frac{1}{2} \int_{s'}^{s''} |\partial_a^2 q_s(y)|\, ds \le C \int_{s'}^{s''} \frac{1}{s} q_{s/2}(y)\, ds \\ &\le \sqrt{\frac{s''}{s'}} q_{s''/2}(y) \int_{s'}^{s''} \frac{1}{s}\, ds \le \sqrt{\frac{2s''}{s'}} q_{s''}(y) \ln\left(\frac{s''}{s'} \right) \\ &= \sqrt{\frac{2s''}{s'}} q_{s''}(y) \ln\left(1 + \left(\frac{s''}{s'} - 1 \right) \right) \le \sqrt{\frac{2s''}{s'}} q_{s''}(y) \left(\frac{s'' - s'}{s'} \right). \end{aligned} \tag{A.5}$$

From here, the result follows. □

Acknowledgments. A. Kohatsu-Higa would like to thank L. Li for comments on an earlier version of this article.

REFERENCES

[1] APPLEBAUM, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge Univ. Press, Cambridge. MR2512800

[2] BALLY, V. and TALAY, D. (1996). The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields* **104** 43–60. MR1367666

[3] BALLY, V. and TALAY, D. (1996). The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.* **2** 93–128. MR1401964

[4] BASS, R. F. (2004). Stochastic differential equations with jumps. *Probab. Surv.* **1** 1–19. MR2095564

[5] BESKOS, A. and ROBERTS, G. O. (2005). Exact simulation of diffusions. *Ann. Appl. Probab.* **15** 2422–2444. MR2187299

[6] CORIELLI, F., FOSCHI, P. and PASCUCCI, A. (2010). Parametrix approximation of diffusion transition densities. *SIAM J. Financial Math.* **1** 833–867. MR2736914

[7] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ. MR0181836

[8] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[9] KONAKOV, V. and MAMMEN, E. (2000). Local limit theorems for transition densities of Markov chains converging to diffusions. *Probab. Theory Related Fields* **117** 551–587. MR1777133

[10] KONAKOV, V. and MAMMEN, E. (2002). Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.* **8** 271–285. MR1931967

[11] KONAKOV, V. and MENOZZI, S. (2011). Weak error for stable driven stochastic differential equations: Expansion of the densities. *J. Theoret. Probab.* **24** 454–478. MR2795049

[12] KONAKOV, V., MENOZZI, S. and MOLCHANOV, S. (2010). Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 908–923. MR2744877

[13] MCKEAN, H. P. JR. and SINGER, I. M. (1967). Curvature and the eigenvalues of the Laplacian. *J. Differential Geom.* **1** 43–69. MR0217739

[14] MENOZZI, S. (2011). Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16** 234–250. MR2802040

[15] RESSEL, P. (1976). A short proof of Schoenberg's theorem. *Proc. Amer. Math. Soc.* **57** 66–68. MR0405007

[16] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes. Grundlehren der mathematischen Wissenschaften* **233**. Springer, Berlin. MR0532498

[17] TALAY, D. and TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8** 483–509. MR1091544

LAMA (UMR CNRS, UPEMLV, UPEC)
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE
AND
MATHRISK
INRIA
F-77454 MARNE-LA-VALLÉE
FRANCE
E-MAIL: bally@univ-mlv.fr

DEPARTMENT OF MATHEMATICAL SCIENCES
RITSUMEIKAN UNIVERSITY
AND JAPAN SCIENCE AND TECHNOLOGY AGENCY
1-1-1 NOJIHIGASHI
KUSATSU, SHIGA, 525-8577
JAPAN
E-MAIL: arturokohatsu@gmail.com