A probabilistic interpretation of the parametrix method

Vlad Bally, Arturo Kohatsu-Higa

A probabilistic interpretation of the parametrix method *

Vlad Bally
Université Paris-Est Marne-la-Vallée, LAMA (UMR CNRS, UPEMLV, UPEC) and MathRisk, INRIA, F-77454 Marne-la-Vallée, France.
Email: bally@univ-mlv.fr
Arturo Kohatsu-Higa
Ritsumeikan University and Japan Science and Technology Agency
Department of Mathematical Sciences
1-1-1 Nojihigashi,
Kusatsu, Shiga, 525-8577, Japan
E-mail: arturokohatsu@gmail.com

January 9, 2014

Abstract

In this article, we introduce the parametrix method for constructions of fundamental solutions as a general method based on semigroups and difference of generators. This leads to a probabilistic interpretation of the parametrix method that are 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 for solving parabolic partial differential equations (PDE's) 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 regular function in terms of polynomial functions.

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 partial differential equations. It has been successfully extended to other situations for theoretical goals (see e.g [13], [14] and [16]). In [9], the authors consider the parametrix as an analytical method for approximations for continuous diffusions. These analytical approximations may be used as deterministic approximations and is 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 re-expresses the arguments of the method in probabilistic terms, but also to introduce an alternative method of exact simulation. Our abstract approach is based on a semigroup method which clarifies the use of a Taylor-expansion like argument. This leads to the natural emergence of the difference between the generators of the process and its approximation. This approach will allow the direct application of this method to various situations.

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

*This research has been supported by grants of the Japanese government.

the continuous diffusion process with infinitesimal operator L. To fix the ideas, consider the diffusion process X_t solution of the following stochastic differential equation (SDE)

$$dX_t = \sum_{j=1}^{m} \sigma_j(X_t)dW_t^j + b(X_t)dt,\ X_0 = x,$$

where W is a multi-dimensional 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. Here $X_t(x)$ denotes the solution which starts from x, i.e. $X_0(x) = x$. The infinitesimal operator associated to P is defined for $f \in C_c^2(\mathbb{R}^d)$ as

$$Lf(x) = \sum_{i,j} a^{i,j}(x)\partial_{i,j}^2 f(x) + b^i(x)\partial_i f(x), \qquad a := \sigma\sigma^*. \tag{1.1}$$

By the Feyman-Kac formula, one knows that $P_t f$ is the unique solution to $\partial_t u = Lu$ with 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 $[0,T]$, $\pi = \{0 = t_0 < ... < t_n = T\}$, the Euler scheme associated to this time grid is defined as $X_0^\pi(x) = x$

$$X_{t_{k+1}}^\pi = X_{t_k}^\pi + \sum_{j=1}^{m} \sigma_j(X_{t_k}^\pi)(W_{t_{k+1}}^j - W_{t_k}^j) + b(X_{t_k}^\pi)h. \tag{1.2}$$

It is well known (see [19]) that $X^\pi(x)$ is an approximation scheme of $X(x)$ of order one. That is,

$$|\mathbb{E}[f(X_T(x))] - \mathbb{E}[f(X_T^\pi(x))]| \leq C_f \max\{t_{i+1} - t_i; i = 0, ..., n-1\} \tag{1.3}$$

for f measurable and bounded (see [3], [4]) 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

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

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 (1.2). 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 the probabilistic parallel of the parametrix method. This probabilistic interpretation will give an exact simulation method for $\mathbb{E}[f(X_T)]$ 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 :

$$\mathbb{E}[f(X_T)] = e^T \mathbb{E}[f(X_T^\pi) \prod_{k=0}^{J_T - 1} \theta_{\tau_{k+1} - \tau_k}(X_{\tau_k}^\pi, X_{\tau_{k+1}}^\pi)]. \tag{1.4}$$

Here, $\pi := \{\tau_k, k \in \mathbb{N}\}$ are the jump times of a Poisson process $\{J_t; t \in [0,T]\}$ of parameter one and $X_{\tau_k}^\pi$ is the Markov chain with transition kernel given by

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

For $\tau_k \leq t < \tau_{k+1}$ we define $X_t^\pi = X_{\tau_k}^\pi$. 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 the weight function.

Before discussing the nature of the above exact simulation formula, let us remark that another such formula is available for one dimensional processes (see [6]) which is strongly based on the explicit formulas for diffusion process that one may obtain for one dimensional diffusions using the Lamperti formula. Although many elements may be common between these two formulas, the one presented here is different in nature.

In order to motivate the above formula (1.4), let us recall the following basic heuristic argument that leads to the parametrix method

$$P_T f(x) - P_T^z f(x) = -\int_0^T \partial_t P_{T-t} P_t^z f(x) dt = \int_0^T P_{T-t}(L - L^z) P_t^z f(x) dt.$$

The above expression is already an equivalent of a Taylor expansion or order one where the notion of first order derivative is being replaced by $L - L^z$. If we repeat the same procedure again inside the time integral, we will have the expansion of order 2, where the first order term will now become $\int_0^T P_{T-t}^z (L - L^z) P_t^z f(x) dt$. Let us now discuss the probabilistic structure of the above expressions. The first term $P_T^z f(x)$ corresponds to the Euler scheme, $X^{\pi,z}(x)$ with $\pi = \{0, T\}$ with coefficients frozen at z. Now let us suppose that

$$P_{T-t}^z (L - L^z) P_t^z f(x) = P_{T-t}^z (\theta_{T-t} P_t^z (f))(x).$$

For some function $\theta_t : \mathbb{R}^d \to \mathbb{R}^d$. Then the probabilistic interpretation of the above expression corresponds to an Euler scheme $X^{\pi,z}(x)$ with time partition $\pi = \{0, T-t, T\}$ and weight $\theta_{T-t}(X_{T-t}^{\pi,z}(x))$. After proper renormalization, the time integrals appearing in the first order term may be interpreted as the first jump of a Poisson process which will have a uniform law once we now that one jump has happened. The term $P_T^z f(x)$ will therefore correspond to the case when the Poisson process does not jump in the interval $[0, T]$ and similarly the k-th multiple time integral in the parametrix method will be interpreted as the case that the Poisson process has jumped k-times.

We will now discuss two technical points: how to obtain the weight function θ and how to choose z.

Computing θ can be done in two ways: The first is to apply directly the operator $L - L^z$ to the density associated to P_t^z. This will be called the backward method. The second is to compute the dual operator (or carry out an integration by parts), in order to apply the dual operator to the density associated to P_{T-t}^z. This will be called the forward method which is the one that corresponds to the formula introduced in (1.4). Although it may seem the least natural form of obtaining the weight θ, the forward method leads to the usual Euler scheme in the case of diffusions. The backward method will lead to a "backward" (in space) running Euler scheme (for more details, see Section 5.1).

The choice if z is a delicate point in the argument and it is related to the degeneracy of the weight function θ_{T-t}. In fact, it is not difficult to see that in the continuous diffusion case, the weight function θ_{T-t} is related to the derivatives of a Gaussian density and that it degenerates as $T - t$ approaches 0. This rate of degeneration is in general non-integrable in t if we do not make a particular choice of z. Without giving much more detail, the particular choice of z will correspond to a "diagonal" case which will require a technical argument and which will finally lead (in the forward case) to the Euler scheme with random time partition.

Our article is structured as follows: In Section 2, we give the notation used throughout the paper. In Section 3 we provide a general abstract framework base 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 4 we give the analytical form of the forward method which demands less model requirements but as we will see in the examples requires stringent conditions on the coefficients of the diffusion due to hypothesis 4.2. In Subsection 4.1 we give the probabilistic representation and in Subsection 4.2, we give the continuity and differentiability properties of the density functions. This is the usual application of the parametrix method.

In Section 5, we give the backward approach which will require less regularity on the coefficients in the case of diffusions. As in the forward case, we also give the probabilistic interpretation and the regularity results corresponding to the backward method in parallel sections.

In Section 6, we consider our main examples. The first corresponds to the continuous diffusion with uniformly elliptic diffusion coefficient. We see in Subsection 6.1 that in the forward approach we will need that the coefficients have to be smooth. While in Subsection 6.2, we show that in order for the backward approach to be applicable, we only require the coefficients to be Hölder continuous. In Subsection 6.3 we also consider the case of a jump driven SDE where the Lévy measure is locally 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 notations and general definitions

We now give some basic notation and definitions used through this article. For a sequence of operators S_i, $i = 1, ..., n$ 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. $Dom(S)$ 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, $A \leq B$ 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 i-th variable of the function f and similarly for higher order derivatives (e.g

derivatives with respect to a multi-index β, of length $|\beta|$, are denoted by $\partial^\beta f$.

$\delta_a(dx)$ denotes 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 $\mathbb{R}_+ \equiv (0, \infty)$. $1_A(x)$ denotes the indicator function of the set A, $C(A)$ denotes the space of real valued functions continuous in the set A. $C_b(A)$ denotes the space of continuous bounded functions in A. $C_c^\infty(\mathbb{R}^d)$ denotes the space of real valued infinitely differentiable functions with compact support defined on $\mathbb{R}^d$. $C_b^k(\mathbb{R}^d; \mathbb{R}^l)$ denotes the space of $\mathbb{R}^l$-valued bounded functions defined on $\mathbb{R}^d$ with bounded derivatives up to order k. The norm in this space is defined as $\|f\|_{k,\infty} = \max_{|\beta| \leq k} \sup_{x \in \mathbb{R}^d} \left|\partial^\beta f(x)\right|$. In the particular case that $k = 0$ we also use the simplified notation $\|f\|_\infty \equiv \|f\|_{0,\infty}$.

$q_a(y)$ denotes the multi-dimensional Gaussian density at $y \in \mathbb{R}^d$ with mean zero and variance-covariance matrix given by the positive definite matrix a. 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 $H_a^{i,j}(y)$, $i\ j \in \{1, ..., d\}$, denote the multi-dimensional 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 8.2.

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.

3 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. The following is the framework for our work.

Hypothesis 3.1. *$(P_t)_{t\geq 0}$ is a semigroup of linear operators defined on a space containing $C_c^\infty(\mathbb{R}^d)$ with infinitesimal generator L such that $C_c^\infty(\mathbb{R}^d) \subseteq Dom(L)$. $P_t f(x)$ 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 6, P^z stands for the semigroup of a diffusion process with coefficients "frozen" at z. We consider this explicit approximating class in Section 6 but maybe there are other interesting examples of approximating semigroups.

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

Hypothesis 3.2. *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 $C_c^\infty(\mathbb{R}^d) \subseteq Dom(L^z)$. 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 3.3. *$Lf(z) = L^z f(z)$ 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

$$\begin{aligned} \psi_t^x(y) &:= p_t^x(x,y), \\ \phi_t^z(x) &:= p_t^z(x,z). \end{aligned}$$

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 e.g. $L^z\phi_t^z(x) \equiv (L^z\phi_t^z)(x)$.

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,

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

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

$$\hat{S}_t f(y) := \int f(x)(L^x - L^y)\phi_t^y(x) dx. \tag{3.1}$$

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, $i = 0, ..., n+1$, always going from y_0 towards y_{n+1} going from x to y in the forward method and from y to x in the backward method.

The above hypotheses 3.1, 3.2 and 3.3 will be assumed throughout the theoretical part of the article. They will be easily verified in the examples.

4 Forward method

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

Hypothesis 4.1. *$P_t^z g,\ P_t g \in \cap_{x\in\mathbb{R}^d} Dom(L^x)$, $\forall g \in C_c^\infty(\mathbb{R}^d)$, $z \in \mathbb{R}^d$, $t \in [0,T]$.* .

We assume the following two regularity properties for the difference operator S. First, we assume that there exists a probability kernel $S_t(x, dy)$ which represents S_t:

Hypothesis 4.2. *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*

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

Furthermore there exists $\rho \in [0,1)$ and a jointly measurable non-negative function $\gamma_t^1(x,y)$, such that $p_t^x(x,y)\,|\theta_t(x,y)| \leq \frac{C}{t^\rho}\gamma_t^1(x,y)$ for $(t,x,y) \in (0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ with $\sup_{(t,x)\in(0,T]\times\mathbb{R}^d}\int_{\mathbb{R}^d}\gamma_t^1(x,y)dy < \infty$.

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

Hypothesis 4.3. **A.** *Under Hypothesis 4.2, suppose that there exists a jointly measurable function $\gamma = (\gamma^1, \gamma^2)$, $\rho \in [0,1)$ and a constant $C > 0$ such that for any $(t,x,y) \in (0,T] \times \mathbb{R}^d \times \mathbb{R}^d$*

$$p_t^x(x,y)\,|\theta_t(x,y)| \le \frac{C}{t^\rho}\gamma_t^1(x,y), \tag{4.1}$$

$$p_t^x(x,y) \le C\gamma_t^2(x,y). \tag{4.2}$$

The functions γ^i satisfy the following condition.

B. *There exists $r > 0$, $\zeta \ge 1$ 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 $n \in \mathbb{N}$, $\delta_i > 0, i = 0, ..., n-1$ and $(y_0, y_n) \in B(z_0, r) \times B(z_n, r)$ we have

$$\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1}|y_i| \le R\}} \prod_{i=0}^{n-2} \gamma^1_{\delta_i}(y_i, y_{i+1})^\zeta \gamma^2_{\delta_{n-1}}(y_{n-1}, y_n)^\zeta \le C_R^n \xi(\sum_{i=0}^{n-1}\delta_i). \tag{4.3}$$

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 $n \in \mathbb{N}$, $\delta_i > 0, i = 0, ..., n-1$, $(y_0, y_n) \in B(z_0, r) \times B(z_n, r)$ and $\varepsilon > 0$, there exists $R_\varepsilon > 0$ with

$$\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1}|y_i| > R\}} \prod_{i=0}^{n-2} \gamma^1_{\delta_i}(y_i, y_{i+1}) \gamma^2_{\delta_{n-1}}(y_{n-1}, y_n) \le C^n \varepsilon \xi(\sum_{i=0}^{n-1}\delta_i). \tag{4.4}$$

Remark 4.4. *(i) We remark here that Hypothesis 4.2 entail some integration by parts property which will be made clear when dealing with examples in Section 6 (see* (6.3)*). Hypothesis 4.3 is a uniform integrability hypothesis together with a uniform estimate of the tails of the convolution of γ^i, needed in order to obtain the continuity of the fundamental solution. The function $\xi(\varepsilon, \delta)$ represents the time variable in the convolution of the functions γ^i and will lead to the integrability of this term when we need to integrate the expression with respect to the times δ_j. In our examples this function is essentially a constant or polynomial functions in δ.*

(ii) If properties (4.3) *with $\zeta = 1$ and* (4.4) *are satisfied then we have in particular that there exists a constant $C > 0$ and a function $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ such that for every $n \in \mathbb{N}$ and $\delta_i > 0, i = 0, ..., n-1$*

$$\sup_{(y_0,y_n)\in B(z_0,r)\times B(z_n,r)} \int dy_1 \cdots \int dy_{n-1} \prod_{i=0}^{n-2} \gamma^1_{\delta_i}(y_i, y_{i+1}) \gamma^2_{\delta_{n-1}}(y_{n-1}, y_n) \le C^n \xi(\sum_{i=0}^{n-1}\delta_i). \tag{4.5}$$

(iii) In typical examples, we will have $\gamma_t^i(x,y) = q_t(x-y)$, $i = 1, 2$. Later, in Section 6.3, we will give an example where γ^1 and γ^2 are given by different functions. Although this is not essential, this set up will make verifications easier in the case of a stochastic differential equation with jumps.

*(iv) In order to obtain the existence of the density we will use Hypothesis 4.3.***B.** *with $\zeta = 1$. Moreover, in order to obtain the continuity and differentiability of the density, we will need a uniform integrability property which follows if Hypothesis 4.3.***B.** *is satisfied for some $\zeta > 1$ and $\zeta\rho < 1$.*

In the forward method, the expansion will be given in terms of some operator compositions, denoted by I^n, which are defined for $(t_0, x) \in (0,T] \times \mathbb{R}^d$ and $f \in C_c^\infty(\mathbb{R}^d)$ as follows

$$I^n_{t_0}(f)(x) := \begin{cases} \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \left(\prod_{i=0}^{n-1} S_{t_i - t_{i+1}}\right) Q_{t_n} f(x), & \text{if } n \ge 1, \\ Q_{t_0} f(x), & \text{if } n = 0. \end{cases} \tag{4.6}$$

In order to define the kernels generating $I^n(f)$, define for $n \ge 1$, $s_0 = 0 \le s_1 < ... < s_n$, $y_0, ..., y_{n+1} \in \mathbb{R}^d$,

$$H_n((s_i)_{i=0,\dots,n}; (y_i)_{i=0,\dots,n}) := \prod_{i=0}^{n-1} \left(\theta_{s_{i+1}-s_i}(y_i, y_{i+1}) p^{y_i}_{s_{i+1}-s_i}(y_i, y_{i+1})\right),$$

$$J^n_{s_n}(y_0,y_n) := \int_0^{s_n} ds_{n-1}\cdots\int_0^{s_2} ds_1 \int dy_1 \cdots \int dy_{n-1} H_n((s_i)_{i=0,\ldots,n};(y_i)_{i=0,\ldots,n}),$$

$$I^n_{t_0}(x,y) = \int_0^{t_0} ds_n \int p^{y_n}_{t_0-s_n}(y_n,y) J^n_{s_n}(x,y_n)dy_n.$$

The relationship between the time variables is given by $t_0 - t_i = s_i$. The reason for this change will become technically clear later. Furthermore, it facilitates the analysis of regularity of the density and it introduces a "natural" time frame for the stochastic representation to be given in Section 4.1. The following is the main result of this section, which is a Taylor-like expansion of P in terms of Q.

Theorem 4.5. *Suppose that Hypotheses 4.1, 4.2 and 4.3 hold with $\zeta = 1$. Then for every fixed $f \in C^\infty_c(\mathbb{R}^d)$ and $t \in (0,T]$, $I^n_t(f)$ is well defined and the sum $\sum_{n=1}^\infty I^n_t(f)(x)$ converges absolutely and uniformly for $(t,x) \in [0,T]\times\mathbb{R}^d$ and*

$$P_t f(x) = \sum_{n=0}^\infty I^n_t(f)(x). \tag{4.7}$$

Furthermore, for fixed $t \in (0,T]$, $\sum_{n=1}^\infty I^n_t(x,y)$ also converges absolutely and uniformly for $(x,y) \in \mathbb{R}^d\times\mathbb{R}^d$ and we have that there exists a jointly measurable function $p_t(x,y)$ such that $P_t f(x) = \int f(y)p_t(x,y)dy$ where

$$p_t(x,y) = p^x_t(x,y) + \sum_{n=1}^\infty I^n_t(x,y). \tag{4.8}$$

In the case 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. Note that $I^n(f)$ is well defined due to Hypotheses 3.2 and 4.2. In fact, for $f \in C^\infty_c(\mathbb{R}^d)$ and using Hypothesis 4.2, we obtain that $\left(\prod_{i=0}^{n-1} S_{t_i-t_{i+1}}\right) Q_{t_n} f(x)$ is jointly measurable. Furthermore, due to Hypothesis 4.2, we have that $\|S_t f\|_\infty \leq Ct^{-\rho}\|f\|_\infty$ and due to Hypothesis 3.2 we have that $\|Q_t f\|_\infty \leq C\|f\|_\infty$. Therefore, applying these results succesively, we have the following bound with $\rho \in [0,1)$

$$\left\| \left(\prod_{i=0}^{n-1} S_{t_i-t_{i+1}}\right) Q_{t_n} f \right\|_\infty \leq C \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-\rho} \|f\|_\infty .$$

The expansion of order one for P_{t_0} is obtained, using Hypotheses 3.1, 3.2, 4.1 and 4.2 as follows

$$P_{t_0} f(x) - P^x_{t_0} f(x) = \int_0^{t_0} \partial_{t_1}\left(P^x_{t_0-t_1}P_{t_1}f\right)(x)dt_1 = \int_0^{t_0} P^x_{t_0-t_1}(L-L^x)P_{t_1}f(x)dt_1.$$

Note that the Hypothesis 4.2 ensures the finiteness of the above integral. Using the identity in Hypothesis 3.3, $Lg(x) = L^x g(x)$ with $g(y) = P_{t_1}f(y)$ we obtain

$$\begin{aligned} P^x_{t_0-t_1}(L-L^x)P_{t_1}f(x) &= \int (L-L^x)P_{t_1}f(y)P^x_{t_0-t_1}(x,dy) = \int (L^y-L^x)P_{t_1}f(y)P^x_{t_0-t_1}(x,dy) \\ &= S_{t_0-t_1}P_{t_1}f(x). \end{aligned}$$

Therefore, we have the following forward expansion formula

$$P_{t_0} f(x) = P^x_{t_0} f(x) + \int_0^{t_0} S_{t_0-t_1}P_{t_1}f(x)dt_1 = Q_{t_0}f(x) + \int_0^{t_0} S_{t_0-t_1}P_{t_1}f(x)dt_1. \tag{4.9}$$

Notice that the above integral makes sense for $f \in C^\infty_c(\mathbb{R}^d)$: by Hypothesis 4.2

$$|S_{t_0-t_1}P_{t_1}f(x)| \leq C_1(t_0-t_1)^{-\rho}\|P_{t_1}f\|_\infty \leq C_1(t_0-t_1)^{-\rho}\|f\|_\infty \tag{4.10}$$

and this quantity is integrable because $\rho \in (0,1)$.

The next step is to iterate this formula. We set $t_0 = T$ and the first iteration of (4.9) gives

$$\begin{aligned} P_{t_0} f(x) &= Q_{t_0} f(x) + \int_0^{t_0} \int S_{t_0-t_1}(x,dy) P_{t_1} f(y) dt_1 \\ &= Q_{t_0} f(x) + \int_0^{t_0} \int S_{t_0-t_1}(x,dy) \left(Q_{t_1} f(y) + \int_0^{t_1} S_{t_1-t_2} P_{t_2} f(y) dt_2 \right) dt_1 \\ &= Q_{t_0} f(x) + \int_0^{t_0} S_{t_0-t_1} Q_{t_1} f(x) dt_1 + \int_0^{t_0} \int_0^{t_1} S_{t_0-t_1} S_{t_1-t_2} P_{t_2} f(x) dt_2 dt_1. \end{aligned}$$

Repeating the above procedure we obtain

$$P_{t_0} f(x) = \sum_{n=0}^{N} I_{t_0}^n(f)(x) + R_{t_0}^N(f)(x)$$

with

$$R_{t_0}^N(f)(x) = \int_0^{t_0} dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{N-1}} dt_N \left(\prod_{i=0}^{N-1} S_{t_i - t_{i+1}} \right) P_{t_N} f(x).$$

We define for $\rho \in (0,1)$

$$c_n(t,\rho) := \int_0^{t} dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-\rho}.$$

Using repeatedly Hypothesis 4.2 and Lemma 8.1, we obtain that there exists positive constants C and C_1 such that

$$\begin{aligned} \left| R_{t_0}^n(f)(x) \right| &\leq C C_1^{n-1} c_n(T,\rho) \left\| f \right\|_\infty \to 0 \qquad \textit{and} \\ \sum_{n=0}^{\infty} \left| I_{t_0}^n(f)(x) \right| &\leq C \left\| f \right\|_\infty \left(1 + \sum_{n=1}^{\infty} C_1^n c_n(T,\rho) \right) < \infty. \end{aligned}$$

Therefore the series is absolutely and uniformly convergent and the proof of the assertion (4.7) finishes.

Now we prove (4.8). We start going back to the multiple integral (4.6) and we give an explicit expression for the integrand as a space integral. This is a straightforward calculation with the change of variables $t_0 - t_i = s_i$ with $i = 1, ..., n$ and an interchange of the order of integration. This gives

$$I_{t_0}^n(f)(y_0) = \int_0^{t_0} ds_n \int Q_{t_0 - s_n} f(y_n) J_{s_n}^n(y_0, y_n) dy_n.$$

By Hypothesis 4.2, we have

$$|H_n((s_i)_{i=0,\dots,n}; (y_i)_{i=0,\dots,n})| \leq C^n \prod_{i=0}^{n-1} (s_{i+1} - s_i)^{-\rho} \gamma^1_{s_{i+1}-s_i}(y_i, y_{i+1}). \tag{4.11}$$

Furthermore from (4.2) and Hypothesis 4.3 (see Remark 4.4 (ii)) with $\delta_i = s_{i+1} - s_i$ (note that $\sum_{i=0}^{n-1} \delta_i = s_n$)

$$\begin{aligned} \left| I_{t_0}^n(y_0, y_{n+1}) \right| &\leq C^n \xi(t_0) \int_0^{t_0} ds_n \int_0^{s_n} ds_{n-1} \cdots \int_0^{s_2} ds_1 \prod_{i=0}^{n-1} (s_{i+1} - s_i)^{-\rho} \\ &= C^n \xi(t_0) c_n(T,\rho). \end{aligned}$$

Therefore the infinite sum in (4.8) converges absolutely and uniformly for $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$. This gives by duality from (4.7),

$$p_{t_0}(x,y) = p_{t_0}^x(x,y) + \sum_{n=1}^{\infty} I_{t_0}^n(x,y).$$

where the above series converges uniformly with respect to $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$. Finally one proves that as the semigroup P is positive then $p_{t_0}(x,y)$ has to be positive locally in y for fixed (t_0,x) then as $P_{t_0}1=1$ one obtains that $\int p_{t_0}(x,y)dy=1$. □

Remark 4.6. *One may also interpret (4.9) as a weak type formulation of the following integral equation*

$$p_t(x,y)=p_t^x(x,y)+\int_0^t\int p_s(z,y)\theta_{t-s}(x,z)p_{t-s}^x(x,z)dzds. \tag{4.12}$$

This formalization is elegant in the forward case but in the backward case becomes too technically involved and therefore we do not consider it here.

4.1 Probabilistic representation using the forward method

Our aim now is to give a probabilistic representation for the formula (4.7) that may be useful for simulation. In order to easily understand the probabilistic interpretation, it is better to rewrite the integral for $I^n_{t_0}$, using the change of variables $s_k=t_0-t_k$, $k=0,...,n$ and Fubini's theorem to obtain that

$$I^n_{t_0}(f)(x)=\int_0^{t_0}ds_n\int_0^{s_n}ds_{n-1}\cdots\int_0^{s_2}ds_1\left(\prod_{i=0}^{n-1}S_{s_{i+1}-s_i}\right)Q_{t_0-s_n}f(x). \tag{4.13}$$

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 $\tau_0=0$).

We fix $x\in\mathbb{R}^n$ and we consider the following probabilistic scheme on the set $\{J_T=n\}$ for $n\in\mathbb{N}$. Define a partition of $[0,T]$, $\pi:=\{s_i;i=0,...,n+1\}$ where $s_i=\tau_i\wedge T$ and we construct the process $(X_t^\pi)_{t\in[0,T]}$ as follows. We assume that $X_0^\pi=x$ and inductively define $X^\pi_{s_k}$, $k=1,...,n+1$, independent from the process J, using the following conditional probabilities

$$P(X^\pi_{s_{k+1}}\in dy'\mid X^\pi_{s_k}=y)=P^x_{s_{k+1}-s_k}(y,dy')=p^y_{s_{k+1}-s_k}(y,y')dy'. \tag{4.14}$$

The definition of the stochastic process X^π, between time grid points, is given by

$$X^\pi_s=X^\pi_{s_k}\qquad s_k\leq s<s_{k+1}.$$

With these definitions, we have that $S_1f(x)=\mathbb{E}\left[f(X^\pi_{s_1})\theta_{s_1}(x,X^\pi_{s_1})\right]$ and $Q_{s_1}f(x)=\mathbb{E}[f(X^\pi_{s_1})]$. Therefore using these representations, we obtain the probabilistic representation of the integrand in (4.13)

$$\left(\prod_{j=0}^{n-1}S_{s_{j+1}-s_j}\right)Q_{T-s_n}f(x)=\mathbb{E}\left[f(X^\pi_T)\theta_{s_n-s_{n-1}}(X^\pi_{s_{n-1}},X^\pi_{s_n})\cdots\theta_{s_1-s_0}(X^\pi_{s_0},X^\pi_{s_1})\right].$$

Finally, to obtain the probabilistic interpretation for the representation formula (4.7), we need to find the probabilistic representation of the multiple integrals in (4.13). 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 (4.13) 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

$$I^n_T(f)(x)=e^T\mathbb{E}\left[1_{\{J_T=n\}}f(X^\pi_T)\prod_{j=0}^{n-1}\theta_{\tau_{j+1}-\tau_j}(X^\pi_{\tau_j},X^\pi_{\tau_{j+1}})\right]$$

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

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 4.7. *Suppose that Hypotheses 4.1, 4.2 and 4.3 hold with $\zeta = 1$. Then the following probabilistic representations are satisfied for $f \in C_c^\infty(\mathbb{R}^d)$,*

$$P_T f(x) = e^T \mathbb{E}\left[f(X_T^\pi)\Gamma_T(x)\right], \tag{4.15}$$

$$p_T(x,y) = e^T \mathbb{E}\left[\psi_{T-\tau_T}^{X_{\tau_T}^\pi}(y)\Gamma_T(x)\right]. \tag{4.16}$$

Here, $\psi_t^x(y) = p_t^x(x,y)$ and $\Gamma_T(x) \equiv \Gamma_T(x)(\omega)$ is a random variable defined as

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

Remark 4.8. 1. *Note that the above formulas (4.15) and (4.16) are amenable to simulation and they give exact simulation formulas. Furthermore, extensions for bounded measurable functions f can be obtained if limits are taken in* (4.15).

2. *The above representations (4.16) and (4.15) 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 (4.16) becomes*

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

This procedure corresponds to an importance sampling method.

4.2 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 result is usually obtained when the analytical version of the parametrix method is applied. Before doing this, we introduce the following assumption.

Hypothesis 4.9. *$\theta_t(\cdot, z) \in C(\mathbb{R}^d)$, for every $t > 0$ and $z \in \mathbb{R}^d$.*

Theorem 4.10. *Suppose that the Hypotheses 4.1, 4.2, 4.3, for $\zeta \in (1, \rho^{-1})$ and 4.9 are satisfied. We assume that for every $(t,x) \in (0,T] \times \mathbb{R}^d$ the function $y \to p_t^x(x,y)$ is continuous. Then $(t,x,y) \to p_t(x,y)$ is continuous on $(0,T] \times \mathbb{R}^d \times \mathbb{R}^d$.*

Proof. We start proving the continuity of the partial sums in (4.8). For this we define

$$\begin{aligned} I_n^R(t_0, y_0, y_{n+1}) &= \int_0^T ds_n \int_0^{s_{n-1}} ds_{n-1} \cdots \int_0^{s_2} ds_1 \int dy_1 \cdots \int dy_n 1_{\{\sum_{i=1}^{n-1}|y_i| \leq R\}} H_n((s_i)_{i=0,\dots,n}; (y_i)_{i=0,\dots,n}) p_{t_0-s_n}(y_n, y_{n+1}) \\ \bar{I}_n^R(t_0, y_0, y_{n+1}) &= I_n(t_0, y_0, y_{n+1}) - I_n^R(t_0, y_0, y_{n+1}). \end{aligned}$$

Consider fixed $n \in \mathbb{N}$, $(z_0, z_{n+1}) \in \mathbb{R}^d \times \mathbb{R}^d$ and two fixed balls $B(z_0, r)$ and $B(z_{n+1}, r)$ with $r > 0$ given in Hypothesis 4.3. From (4.11) and (4.3) (recall that $\zeta\rho < 1$) it follows from Lemma 8.1, for any fixed $R > 0$

$$\sup_{(y_0, y_{n+1}) \in B(z_0,r) \times B(z_{n+1},r)} \int_0^T ds_n \int_0^{s_n} ds_{n-1} \cdots \int_0^{s_2} ds_1 \int dy_1 \cdots \int dy_n 1_{\{\sum_{i=1}^{n-1}|y_i| \leq R\}} \left| H_n((s_i)_{i=0,\dots,n}; (y_i)_{i=0,\dots,n}) p_{t_0-s_n}^{y_n}(y_n, y_{n+1}) \right|^\zeta < \infty.$$

This guarantees that the functions

$$((s_i)_{i=1,\dots,n}; (y_i)_{i=1,\dots,n}) \to 1_{\{\sum_{i=1}^{n-1}|y_i| \leq R\}} H_n((s_i)_{i=0,\dots,n}; (y_i)_{i=0,\dots,n}) p_{t_0-s_n}^{y_n}(y_n, y_{n+1})$$

are uniformly integrable (for $(y_0, y_{n+1}) \in B(z_0, r) \times B(z_{n+1}, r)$ and $0 \le t_0 \le T$). Then one may interchange the limit $\lim_{(t_0,y_0,y_{n+1})\to(t_0',y_0',y_{n+1}')} I_n^R(t_0, y_0, y_{n+1})$ 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 (4.4) is satisfied. Therefore

$$\lim_{\varepsilon\downarrow 0} \sup_{(y_0,y_{n+1})\in B(z_0,r)\times B(z_{n+1},r)} \bar{I}_n^{R_\varepsilon}(t_0, y_0, y_{n+1}) \le C \lim_{\varepsilon\downarrow 0} \varepsilon\xi(t_0) = 0.$$

Therefore we obtain that

$$\lim_{(t_0,y_0,y_{n+1})\to(t_0',y_0',y_{n+1}')} I_{t_0}^n(y_0, y_{n+1}) = I_{t_0'}^n(y_0', y_{n+1}')$$

and therefore $(t_0, x, y) \to I_{t_0}^n(x, y)$ is continuous. We also have

$$\sum_{n\ge N} \sup_{(y_0,y_{n+1})\in B(z_0,r)\times B(z_{n+1},r)} \left|I_{t_0}^n(y_0, y_{n+1})\right| \le \xi(t_0) \sum_{n\ge N} C^n c_n(t_0, \rho) \xrightarrow[N\to\infty]{} 0$$

and this proves that $(t_0, x, y) \to p_{t_0}(x, y)$ is continuous as it is the uniform limit of continuous functions. $\square$

Now we also give a statement about the first order derivatives of the fundamental solutions. The proof is similar.

Theorem 4.11. *Suppose that the Hypotheses 4.1 and 4.2 are satisfied. Furthermore, we assume that Hypothesis 4.3 is satisfied with $\zeta \in (1, \rho^{-1})$ for $(p_t^x(x, y), \gamma_t^2(x, y))$ and also when these two functions in* (4.2) *and* (4.3), (4.4) *are replaced by $(\nabla_y p_t^x(x, y), \gamma_t^3(x, y))$. Then for every $(t, x) \in (0, T] \times \mathbb{R}^d$ the function $y \to p_t(x, y)$ is one time differentiable. Moreover*

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

Proof. We discuss now the differentiability with respect to y. We have

$$\nabla_{y_{n+1}} I_{t_0}^n(y_0, y_{n+1}) = \int_0^{t_0} ds_n \cdots \int_0^{s_2} ds_1 \int dy_1 \cdots \int dy_n H_n((s_i)_{i=0,\ldots,n}; (y_i)_{i=0,\ldots,n}) \nabla_{y_{n+1}} p_{t_0-s_n}^{y_n}(y_n, y_{n+1}).$$

Therefore the same reasoning as in the proof of Theorem 4.10 shows that $y \to p_{t_0}(x, y)$ is differentiable and

$$\nabla_y p_{t_0}(x, y) = \nabla_y p_{t_0}^x(x, y) + \sum_{n=1}^{\infty} \nabla_y I_{t_0}^n(x, y).$$

$\square$

Remark 4.12.

1. *We remark that the functions γ^2 and γ^3 in order to apply the above theorem do not have to necessarily be the same as the respective functions in Theorem 4.5 or 4.10.*
2. *The reason why the corresponding theorem for the second derivative is not interesting in the present setting is due to the fact that in most cases the corresponding condition would not be satisfied. See e.g. Section 6 for a concrete example.*

5 The backward method: Probabilistic representation using the adjoint semigroup

Usually, in semigroup theory one assumes that for each $t > 0$, P_t maps continuously $L^2(\mathbb{R}^d)$ into itself. Then for every $g \in L^2(\mathbb{R}^d)$ we have

$$\left|\int_{\mathbb{R}^d} g(x)P_t f(x)dx\right| \le \|g\|_{L^2} \|P_t f\|_{L^2} \le C \|g\|_{L^2} \|f\|_{L^2}.$$

Then in that case, one may define the adjoint semigroup by

$$\langle P_t^* g, f\rangle = \int_{\mathbb{R}^d} P_t^* g(y)f(y)dy = \int_{\mathbb{R}^d} g(x)P_t f(x)dx = \langle g, P_t f\rangle$$

where $P_t^* g \in L^2(\mathbb{R}^d)$ is the element which represents the continuous functional $f \to \langle g, P_t f \rangle$ by Riesz representation theorem. Clearly P_t^* is still a semigroup and 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 4.7. Unfortunately, the adjoint version of the arguments given in Section 4 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 $Lg(y) = a(y)\triangle g(y)$ then $L^x g(y) = a(x)\triangle g(y)$ (for more details, see Section 6). Then $L^* g(y) = \triangle(ag)(y)$ and $(L^x)^* g(y) = a(x)\triangle g(y) = L^x g(y)$. So, letting the coefficients of L^* be frozen at $y = x$ does not coincide with $(L^x)^*$ and therefore the previous argument will fail. In order not to confuse the reader we will keep using $*$ to denote adjoint operators while other approximating operators will be denoted by $\hat{}$ (hat) even though they play the role of the approximating operators in the previous sections[1].

A second problem is that the above mentioned continuity of P_t in $L^2(\mathbb{R}^d)$ is difficult to verify in the examples. Instead as in section 4, we will assume that the semigroup P_t is well defined in of $C_b(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$. In order to deal with the adjoint operator, P_t^*, which is not necessarily defined, we will use a somewhat roundabout way (in duality) to use them without mentioning them directly. This will demand slightly stronger hypothesis than the one required for the forward method. This methodology is sometimes called densely defined adjoint operators.

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 that

Hypothesis 5.1. *(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) $\phi_t^z \in Dom(L) \cap (\cap_{y \in \mathbb{R}^d} Dom(L^y))$, *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

$$P_t^z \phi_\varepsilon^z(x) = p_{t+\varepsilon}^z(x,z) = \phi_{t+\varepsilon}^z(x). \tag{5.1}$$

As stated before, we remark that $\hat{Q} \neq Q^*$. In fact, $\hat{Q}$ is obtained through a density whose coefficients are "frozen" at the arrival point of the underlying process. Also note that due to Hypothesis 5.1 then $\left\| \hat{Q}_t f \right\|_\infty \le C_T \|f\|_\infty$ for all $t \in [0,T]$. Furthermore, we make the following hypotheses on the behavior of the "frozen" infinitesimal generators L^z applied to ϕ_t^z.

Before introducing the next two hypotheses, we explain the reasoning behind the notation to follow. In the forward method it was clear that the method went from a departure point x to an arrival point y with transition points y_i, $i = 0, ..., n+1$, $y_0 = x$ 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 y and $y_0 = y$ and $y_{n+1} = x$. Therefore the notation to follow, tries to give this intuition.

Hypothesis 5.2. *We suppose that there exists a continuous function* $\hat{\theta} \in C((0,T] \times \mathbb{R}^d \times \mathbb{R}^d)$ *and* $(L^{y_1} - L^y)\phi_t^y(y_1) = \hat{\theta}_t(y_1, y)\phi_t^y(y_1)$ *is integrable* $P_s(x, dy_1)$ *for all* $(s,x) \in [0,T] \times \mathbb{R}^d$ *and* $(t,y) \in (0,T] \times \mathbb{R}^d$.

[1] 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.

Hypothesis 5.3. *There exists a positive constant C, $\rho \in (0,1)$ and a jointly measurable function $\gamma = (\gamma^1, \gamma^2)$ such that for all $(t,x,y) \in (0,T] \times \mathbb{R}^d \times \mathbb{R}^d$, we have that $\phi_t^y(x) \left|\hat{\theta}_t(x,y)\right| \leq \frac{C}{t^\rho}\gamma_t^1(x,y)$, $p_t^y(x,y) \leq C\gamma_t^2(x,y)$ and they satisfy that $\sup_{(t,y)\in[0,T]\times\mathbb{R}^d} \int \gamma_t^1(x,y)dx < \infty$ and $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int \gamma_t^1(x,y)dy < \infty$. Furthermore, there exists $\zeta \in (1, \rho^{-1})$, such that for every $R > 0$*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int 1_{\{|y|\leq R\}} \gamma_t^1(x,y)^\zeta dy < \infty. \tag{5.2}$$

Hypothesis 5.4. *Under the conditions of Hypothesis 5.3, we suppose that the functions γ^i satisfy the following conditions.*
There exists $r > 0$, $\zeta \geq 1$ 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 $n \in \mathbb{N}$, $\delta_i > 0, i = 0, ..., n-1$ and $(y_0,y_n) \in B(z_0,r) \times B(z_n,r)$ we have

$$\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1}|y_i|\leq R\}} \gamma_{\delta_0}^2(y_0,y_1)^\zeta \prod_{i=1}^{n-1} \gamma_{\delta_i}^1(y_i,y_{i+1})^\zeta \leq C_R^n \xi(\sum_{i=0}^{n-1}\delta_i). \tag{5.3}$$

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 $n \in \mathbb{N}$, $\delta_i > 0, i = 0, ..., n-1$, $(y_0,y_n) \in B(z_0,r) \times B(z_n,r)$ and $\varepsilon > 0$, there exists $R_\varepsilon > 0$ with

$$\int dy_1 \cdots \int dy_{n-1} 1_{\{\sum_{i=1}^{n-1}|y_i|> R_\varepsilon\}} \gamma_{\delta_0}^2(y_0,y_1) \prod_{i=1}^{n-1} \gamma_{\delta_i}^1(y_i,y_{i+1}) \leq C^n \varepsilon \xi(\sum_{i=0}^{n-1}\delta_i). \tag{5.4}$$

Remark 5.5. *In the examples we will consider in the following section $\rho = 1 - \frac{\alpha}{2}$ where α will represent the order of Hölder continuity of the coefficients of the SDE's.*

We define now (recall (3.1) and Hypothesis 5.2)

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

Note that due to Hypothesis 5.3 we have that

$$\left\|\hat{S}_t f\right\|_\infty \leq Ct^{-\rho}\|f\|_\infty \text{ for } t \in (0,T]. \tag{5.5}$$

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

$$\hat{I}_{t_0}^n(g)(y) := \begin{cases} \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \left(\prod_{i=0}^{n-1} \hat{S}_{t_i - t_{i+1}}\right) \hat{Q}_{t_n} g(y), & \text{if } n \geq 1, \\ \hat{Q}_{t_0} g(y), & \text{if } n = 0. \end{cases} \tag{5.6}$$

Furthermore we define the adjoint operators

$$\begin{aligned} \hat{Q}_t^* f(x) &:= \int f(y)p_t^y(x,y)dy, \\ \hat{S}_t^* f(x) &:= \int f(y)\hat{\theta}_t(x,y)p_t^y(x,y)dy. \end{aligned}$$

Note that due to the Hypothesis 5.1 (i) and 5.3 we have that for any $f \in L^\infty$

$$\sup_t \left\|\hat{Q}_t^* f\right\|_\infty \leq C\|f\|_\infty, \tag{5.7}$$

$$\left\|\hat{S}_t^* f\right\|_\infty \leq \frac{C}{t^\rho}\|f\|_\infty. \tag{5.8}$$

As in (4.6), we define the following auxiliary operators

$$\hat{I}^{n,*}_{t_0}(f) := \begin{cases} \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \hat{Q}^*_{t_n} \hat{S}^*_{t_{n-1}-t_n} \cdots \hat{S}^*_{t_0-t_1} f, & n \geq 1, \\ \hat{Q}^*_{t_0} f & n = 0, \end{cases}$$

$$\hat{H}_n((s_i)_{i=0,\ldots,n};(y_i)_{i=0,\ldots,n}) := \prod_{i=0}^{n-1} \left(\hat{\theta}_{s_{i+1}-s_i}(y_{i+1}, y_i) p^{y_i}_{s_{i+1}-s_i}(y_{i+1}, y_i) \right),$$

$$\hat{J}^n_{s_n}(y_0, y_n) := \int_0^{s_n} ds_{n-1} \cdots \int_0^{s_2} ds_1 \int dy_1 \cdots \int dy_{n-1} \hat{H}_n((s_i)_{i=0,\ldots,n};(y_i)_{i=0,\ldots,n}).$$

$$\hat{I}^n_{t_0}(y_0, y_{n+1}) := \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \int dy_1 \cdots \int dy_n \hat{H}_n((t_0 - t_i)_{i=0,\ldots,n};(y_i)_{i=0,\ldots,n}) p^{y_n}_{t_n}(y_{n+1}, y_n).$$

Note that $\langle \hat{I}^{n,*}_{t_0}(f), g \rangle = \langle f, \hat{I}^n_{t_0} g \rangle$ and $\hat{I}^{n,*}(f)(x) = \int_{\mathbb{R}^d} f(y) \hat{I}^{n,*}_{t_0}(y, x) dy$ for $f,\ g \in C^\infty_c(\mathbb{R}^d)$. Our main result in this section is

Theorem 5.6. *Suppose that Hypotheses 5.1, 5.2 and 5.3 hold for $\zeta = 1$. Then for every $g \in C^\infty_c(\mathbb{R}^d)$ 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*

$$P^*_t g(y) = \sum_{n=0}^\infty \hat{I}^n_t(g)(y) \qquad dy - a.s.,\ t \in (0, T]. \tag{5.9}$$

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

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

Finally, $\sum_{n=0}^\infty \hat{I}^n_t(y, x)$ 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 $f \in C^\infty_c(\mathbb{R}^d)$ we have $P_t f(x) = \int_{\mathbb{R}^d} f(y) p_t(x, y) dy$ and it is given by

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

In the case 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 4.5. First, we prove that each integral $\hat{I}^n_{t_0}(g)(y)$ is well defined. Iterating (5.5), we obtain that for any linear operator Q satisfying that $\|Qf\|_\infty \leq C \|f\|_\infty$, then

$$\left\| \left(\prod_{i=0}^{n-1} \hat{S}_{t_i - t_{i+1}} \right) Qf \right\|_\infty \leq C^n \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-\rho} \|f\|_\infty. \tag{5.10}$$

Notice that in view of Hypothesis 5.1-(i) we have $\sup_{t_n \leq t_0} \left\| \hat{Q}_{t_n} g \right\|_\infty \leq C_T \|g\|_\infty$. So, by (5.10) and (8.1), the multiple integral which defines $\hat{I}^n_{t_0}(g)(y)$ in (5.6) is well defined and have explicit upper bounds. Therefore one can also prove the absolute convergence of the sum $\sum_{n=0}^\infty \hat{I}^n_t(g)(y)$. In fact, (see Section 8.1)

$$\sum_{n=0}^\infty \left\| \hat{I}^n_t(g) \right\|_\infty \leq \sum_{n=0}^\infty C^n c_n(T, \rho) < \infty. \tag{5.11}$$

In a similar fashion, note that for $\hat{I}^{n,*}_{t_0}(f) = \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \hat{Q}^*_{t_n} \hat{S}^*_{t_{n-1}-t_n} \cdots \hat{S}^*_{t_0-t_1} f$ a similar argument can be given. In fact, due to (5.7) and (5.8), we have

$$\left\| \hat{Q}^*_{t_n} \hat{S}^*_{t_{n-1}-t_n} \cdots \hat{S}^*_{t_0-t_1} f \right\|_\infty \le C^n \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-\rho} \left\| f \right\|_\infty .$$

Therefore the same considerations as in the case of $\hat{I}^n_t(f)$ apply in this case.

To start with the proof of (5.9), as in the proof of Theorem 4.5, we need to find a one step formula in the "Taylor"-like expansion. In order to do this, we need an approximation argument. We take $\varepsilon > 0$ and we recall that due to Hypotheses 3.2 and 5.1 (iii), we have for each $z \in \mathbb{R}^d$ that $P^z_{T-t}\phi^z_\varepsilon = \phi^z_{T-t+\varepsilon} = p^z_{t+\epsilon}(\cdot, z) \in Dom(L)$ and $Lf(x) = L^x f(x)$. Then from Hypotheses 5.2 and 5.3, we have that for $0 < t < T$,

$$\begin{aligned} \partial_t \left(P_t P^{y_2}_{T-t} \right) \phi^{y_2}_\varepsilon(x) &= P_t(L - L^{y_2}) P^{y_2}_{T-t} \phi^{y_2}_\varepsilon(x) = \int P_t(x, dy_3)(L - L^{y_2}) P^{y_2}_{T-t} \phi^{y_2}_\varepsilon(y_3) \\ &= \int P_t(x, dy_3) \hat{\theta}_{T-t+\varepsilon}(y_3, y_2) \phi^{y_2}_{T-t+\varepsilon}(y_3). \end{aligned}$$

We take $g,\ h \in C^\infty_c(\mathbb{R}^d)$ and we note that due to Hypothesis 5.3

$$\begin{aligned} & \int dx \, |g(x)| \int P_t(x, dy_2) \int dy_1 \, |h(y_1)| \left| \hat{\theta}_{T-t+\varepsilon}(y_2, y_1) \right| \phi^{y_1}_{T-t+\varepsilon}(y_2) \\ \le\ & C_3 (T - t + \varepsilon)^{-\rho} \left\| h \right\|_\infty \int dx \, |g(x)| \int P_t(x, dy_2) \\ \le\ & C_3 (T - t)^{-\rho} \left\| h \right\|_\infty \left\| g \right\|_1 . \end{aligned}$$

The above expression is integrable with respect to $1_{(0,T)}(t)dt$ 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 5.2, 5.3 (see (5.2)) and the fact that $h \in C^\infty_c(\mathbb{R}^d)$, we have that for fixed $t \in [0, T)$ we can take limits as $\varepsilon \to 0$ for $\int dy_1 \, |h(y_1)| \left| \hat{\theta}_{T-t+\varepsilon}(y_2, y_1) \right| \phi^{y_1}_{T-t+\varepsilon}(y_2)$, 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

$$\begin{aligned} & \lim_{\varepsilon \to 0} \int_0^T dt \int dy_1 h(y_1) \int dx g(x) \int P_t(x, dy_2) \hat{\theta}_{T-t+\varepsilon}(y_2, y_1) \phi^{y_1}_{T-t+\varepsilon}(y_2) \\ = & \int_0^T dt \int dy_1 h(y_1) \int dx g(x) \int P_t(x, dy_2) \hat{\theta}_{T-t}(y_2, y_1) \phi^{y_1}_{T-t}(y_2). \end{aligned}$$

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

$$\begin{aligned} \int dy_1 h(y_1) (\langle g, P_T \phi^{y_1}_\varepsilon \rangle - \langle g, P^{y_1}_T \phi^{y_1}_\varepsilon \rangle) &= \int dy_1 h(y_1) \int dx g(x) \int_0^T \partial_t \left(P_t P^{y_1}_{T-t} \right) \phi^{y_1}_\varepsilon(x) dt \\ &= \int_0^T dt \int dy_1 h(y_1) \int dx g(x) \int P_t(x, dy_2) \hat{\theta}_{T-t+\varepsilon}(y_2, y_1) \phi^{y_1}_{T-t+\varepsilon}(y_2). \end{aligned} \tag{5.12}$$

In order to obtain the expansion of order one, we need to take limits in (5.12). To deal with the limit of the left hand side of (5.12) we note that given the assumptions $g,\ h \in C^\infty_c(\mathbb{R}^d)$ and Hypothesis 5.1 (ii), we have

$$\begin{aligned} \lim_{\varepsilon \to 0} \int dy_1 h(y_1) \langle g, P_T \phi^{y_1}_\varepsilon \rangle &= \lim_{\varepsilon \to 0} \int g(y_2) \int P_T(y_2, dw) \int dy_1 h(y_1) \phi^{y_1}_\varepsilon(w) = \langle P_T h, g \rangle \\ \lim_{\varepsilon \to 0} \int dy_1 h(y_1) \left\langle P^{y_1,*}_T g, \phi^{y_1}_\varepsilon \right\rangle &= \lim_{\varepsilon \to 0} \int dy_1 h(y_1) P^{y_1,*}_{T+\varepsilon} g(y_1) = \int dy_1 h(y_1) P^{y_1,*}_T g(y_1). \end{aligned}$$

Therefore taking limits in (5.12), we obtain

$$\begin{aligned}\langle P_T h, g\rangle &= \int dy_1 h(y_1) P_T^{*,y_1} g(y_1) + \int dy_1 h(y_1) \int dx g(x) \int_0^T dt \int P_t(x, dy_2) \hat{\theta}_{T-t}(y_2, y_1) \phi_{T-t}^{y_1}(y_2) \\ &= \left\langle \hat{Q}_T^* h, g \right\rangle + \int_0^T \left\langle P_t \hat{S}_{T-t}^* h, g \right\rangle dt.\end{aligned}$$

Following the same argument as in the proof of Theorem 4.5, we iterate the above first order approximation formula. Before this, we note that $h \in C_c^\infty(\mathbb{R}^d)$ implies that $\hat{S}_{T-t}^* h \in C_b(\mathbb{R}^d)$ due to Hypotheses 3.2 and 5.3. Therefore we have by iteration that

$$\langle P_T h, g\rangle = \left\langle \hat{Q}_T^* h, g \right\rangle + \int_0^T \left\langle \hat{Q}_t^* \hat{S}_{T-t}^* h, g \right\rangle dt + \int_0^T \int_0^t \left\langle P_s \hat{S}_{t-s}^* \hat{S}_{T-t}^* h, g \right\rangle ds dt.$$

Then we continue with this iteration procedure and we need to prove that the residue converges to zero. The residue will be denoted by $\hat{R}_{t_0}^N(g)$ and defined as follows. For $g \in C_c^\infty(\mathbb{R}^d)$, we define the residue terms $\hat{R}_{t_0}^N(g)$ by

$$\hat{R}_{t_0}^N(g)(z) = \int_0^{t_0} dt_1 \cdots \int_0^{t_{N-1}} dt_N P_{t_N} \left(\prod_{i=N-1}^{0} \hat{S}_{t_i - t_{i+1}}^* \right) g(z).$$

In order to prove the formula (5.9), we proceed with $L^\infty(\mathbb{R}^d)$-type estimates. In fact, as a consequence of (5.10) we have

$$\left\| \hat{R}_{t_0}^N(g) \right\|_\infty \leq C_3^N \int_0^{t_0} dt_1 \cdots \int_0^{t_{N-1}} dt_N \prod_{i=0}^{N-1} (t_i - t_{i+1})^{-\rho} \|g\|_\infty .$$

Therefore (5.9) follows due to the estimates on $c_N(t_0, \rho)$ and (5.11).

The proof of the other statements are done in the same way as in the proof of Theorem 4.5.□

Remark 5.7.
1. *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.*
2. *Recall that $L^x \phi_t^y(x) = L\phi_t^y(x)$ so, by the definition of $\hat{\theta}$ and $\hat{S}$, we have*

$$\begin{aligned}\hat{S}_t f(y) &= \int f(x)(L^x - L^y)\phi_t^y(x)dx = \int f(x)(L - L^y)\phi_t^y(x)dx \\ &= \int (L^* - (L^y)^*) f(x) \phi_t^y(x) dx = \int (L^* - (L^y)^*) f(x) p_t^y(x, y) dx \\ &= \int (L^* - (L^y)^*) f(x) p_t^{*,y}(y, x) dx = \int (L^* - (L^y)^*) f(x) \left(P_t^y\right)^* (y, dx)\end{aligned}$$

which is the analogue to (if we exchange y and x as well as L and L^*)

$$S_t f(x) = \int (L^y - L^x) f(y) p_t^x(x, y) dy = \int (L - L^x) f(y) P_t^x(x, dy).$$

which has been used in the previous section. This analogy may suggest that if the generator L is symmetric the two formulas coincide. In fact this is false: consider again, the example of one dimensional SDE's (see also Section 6). Then L is symmetric (i.e. $L = L^*$) if and only if $b = \frac{1}{2}\partial a$ because in this case $L = \frac{1}{2}a\partial^2 + \frac{1}{2}\partial a \partial = \frac{1}{2}\partial\left(a\partial\right)$. But then P_t^x is not symmetric - in order to have this property $P_t^x = \left(P_t^x\right)^*$ we need the condition $b = 0$, making the situation trivial.

5.1 Probabilistic representation 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 4.1, (see (4.14)) we have performed a similar construction.

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 $\tau_0 = 0$).

We fix $x \in \mathbb{R}^n$ and we consider the following probabilistic scheme on the set $\{J_T = n\}$ for $n \in \mathbb{N}$. Define a partition of $[0, T]$, $\pi := \{s_i; i = 0, ..., n+1\}$ where $s_i = \tau_i \wedge T$ we construct a Markov chain $X_t^{*,\pi}$, $t = s_k$ which represents the Euler scheme associated to a partition $\pi := \{s_i; i = 0, ..., n+1\}$ where $0 = s_0 < s_1 < ... < s_{n+1} = T$.

More precisely we define

$$\begin{aligned} X_0^{*,\pi}(y) &= y, \\ P(X_{s_{k+1}}^{*,\pi}(y) \in dy_{k+1} \mid X_{s_k}^{*,\pi}(y) = y_k) &= C_{s_{k+1}-s_k}^{-1}(y_k) P_{s_{k+1}-s_k}^{y_k,*}(y_k, dy_{k+1}) = C_{s_{k+1}-s_k}^{-1}(y_k) \phi_{s_{k+1}-s_k}^{y_k}(y_{k+1}) dy_{k+1}, \\ C_{s_{k+1}-s_k}(y_k) &= \int \phi_{s_{k+1}-s_k}^{y_k}(y_{k+1}) dy_{k+1}. \end{aligned}$$

Then the same arguments as in the previous section give the representation

$$I_T^{*,n}(g)(y) = e^T \mathbb{E}\left[1_{\{J_T = n\}} g(X_T^{*,\pi}(y)) C_{T-\tau_n}(X_{\tau_n}^{*,\pi}(y)) \prod_{j=0}^{n-1} C_{\tau_{j+1}-\tau_j}(X_{\tau_j}^{*,\pi}(y)) \hat{\theta}_{\tau_{j+1}-\tau_j}(X_{\tau_{j+1}}^{*,\pi}(y), X_{\tau_j}^{*,\pi}(y)) \right].$$

We define

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

Sometimes we may use the notation $X_{\tau_j}^{*,\pi}(y)$ to indicate that $X_0^{*,\pi}(y) = y$. The main result in this section is about representations of the adjoint semigroup P^* and its densities.

Theorem 5.8. *Suppose that Hypotheses 5.1, 5.2 and 5.3 (with $\zeta = 1$) hold then the following representation formula is valid for any $g \in C_c^\infty(\mathbb{R}^d)$*

$$P_T^* g(y) = P_T^{*,z} g(y) + e^T \mathbb{E}\left[g(X_T^{*,\pi}(y)) \Gamma_T^*(y) 1_{\{J_T \geq 1\}} \right] = e^T \mathbb{E}\left[g(X_T^{*,\pi}(y)) \Gamma_T^*(y) \right]. \tag{5.13}$$

Theorem 5.9. *Suppose that Hypotheses 5.1, 5.2 and 5.3 (with $\zeta = 1$) hold then the following representation formula for the density is valid*

$$p_T(x, y) = e^T \mathbb{E}\left[p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(y)}(x, X_{\tau_T}^{*,\pi}(y)) \Gamma_T^*(y) \right]. \tag{5.14}$$

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

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

In particular, the above result implies the following expression for the Laplace transform of X_T, for $\alpha \in \mathbb{R}_+^d$ and $\bar{\alpha} = \prod_{i=1}^d \alpha_i > 0$

$$\mathbb{E}[e^{-\alpha \cdot X_T}] = \bar{\alpha}^{-1} e^T \mathbb{E}\left[p_{T-\tau_T}^{X_{\tau_T}^{*,\pi}(Z)}(x, X_{\tau_T}^{*,\pi}(Z)) \Gamma_T^*(Z) \right].$$

Here, Z is a $d-$dimensional random vector with exponential distribution of parameter α. Similarly $X_{\tau_T}^{*,\pi}(Z)$ denotes the Euler scheme which starts from the same exponentially distributed initial point.

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

$$\mathbb{E}\left[g(X_T^\pi) C_{T-\tau_T}(y) \mid \tau_T,\ X_{\tau_T}^\pi = y \right] = \int g(x) p_{T-\tau_T}^{*,y}(y, x) dx$$

so that (5.13) says that $P^*_T(y,dx)=p^*_T(y,x)dx$ with

$$p^*_T(y,x)=e^T\mathbb{E}\left[p^{*,X^{*,\pi}_{\tau_T}(y)}_{T-\tau_T}(X^{*,\pi}_{\tau_T}(y),x)\Gamma^*_T(y)\right]. \tag{5.15}$$

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 (5.15) proves that $P_t(x,dy)$ is absolutely continuous and the density is represented by

$$p_T(x,y)-p^y_T(x,y) \;=\; e^T\mathbb{E}\left[p^{X^{*,\pi}_{\tau_T}(y)}_{T-\tau_T}(x,X^{*,\pi}_{\tau_T}(y))\Gamma^*_T(y)1_{\{J_T\geq 1\}}\right].$$

The representation for P_Th can be obtained integrating $\int h(y)p_T(x,y)dy$ using (5.14). □

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

$$p_T(x,y)-p^y_T(x,y)=e^{\lambda T}\mathbb{E}\left[\lambda^{-J^\lambda_T}p^{X^{*,\pi}_{\tau_{J^\lambda_T}}(y)}_{T-\tau_{J^\lambda_T}}(x,X^{*,\pi}_{\tau_{J^\lambda_T}}(y))\Gamma^*_T(y)1_{\{J^\lambda_T\geq 1\}}\right].$$

We discuss now the regularity of $p_t(x,y)$. We make the following supplementary assumption

Hypothesis 5.10. *For any $z\in\mathbb{R}^d$, $p^z_t(\cdot,z)\in C^1(\mathbb{R}^d)$.*

Then the same arguments as in Section 4.2 prove:

Theorem 5.11. *Suppose that Hypotheses 5.1, 5.2 and 5.10. Furthermore, we assume that Hypotheses 5.3 and 5.4 are satisfied for some $\rho\in[0,1)$ and $\zeta\in(1,\rho^{-1})$ for $(p^x_t(x,y),\gamma^2_t(x,y))$ and also when these two functions are replaced by $(\nabla_y p^x_t(x,y),\gamma^3_t(x,y))$.*
Then $(t,x,y)\to p_t(x,y)$ is continuous on $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ and for every $t\in(0,T],z\in\mathbb{R}^d$, the function $x\to p_t(x,z)$ is one time differentiable. Moreover,

$$p_T(x,y)=e^T\mathbb{E}\left[p^{X^{*,\pi}_{\tau_T}(y)}_{T-\tau_T}(x,X^{*,\pi}_{\tau_T}(y))\hat{\Gamma}_T(y)\right]$$

and

$$\nabla_x p_T(x,y)=\mathbb{E}\left[\nabla_x p^{X^{*,\pi}_{\tau_T}(y)}_{T-\tau_T}(x,X^{*,\pi}_{\tau_T}(y))\hat{\Gamma}_T(y)\right].$$

6 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 multi-dimensional diffusion processes. We will consider first the forward method as in Section 4, where we will see that the smoothness of the coefficients is required in order to find explicit expressions for θ in Hypothesis 4.2 while in the backward method, treated in Section 5, we will only need to require that the diffusion coefficient is Hőlder continuous in order to define $\hat{\theta}$ in Hypothesis 5.3. On the other hand although the backward method recovers the density function, it does not adapt easily to the simulation of all functions of the diffusion process.

6.1 Example 1: The forward method for continuous SDE's with smooth coefficients

We consider the following $d-$dimensional SDE

$$X_t=x+\sum_{j=1}^m\int_0^t\sigma_j(X_s)dW^j_s+\int_0^t b(X_s)ds. \tag{6.1}$$

Here $\sigma_j,\ b:\mathbb{R}^d\to\mathbb{R}^d$ $\sigma_j\in C^2_b(\mathbb{R}^d;\mathbb{R}^d)$ is uniformly elliptic (i.e. $0<\underline{a}I\leq a\leq\overline{a}I$ for $\underline{a},\ \overline{a}\in\mathbb{R}$ with $a=\sigma\sigma^*$), $b\in C^2_b(\mathbb{R}^d;\mathbb{R}^d)$ 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 $Lf(x) = \frac{1}{2}\sum_{i,j} a^{i,j}(x)\partial^2_{i,j}f(x) + \sum_i b^i(x)\partial_i f(x)$ for $f \in C_c^\infty(\mathbb{R}^d)$ and $a^{i,j}(x) = \sum_k \sigma^i_k(x)\sigma^j_k(x)$. Clearly $P_t f(x)$ is jointly measurable and bounded and therefore Hypothesis 3.1 is satisfied. We will consider the following approximation process:

$$X^z_t(x) = x + \sum_{j=1}^m \sigma_j(z)W^j_t + b(z)t$$

which defines the semigroup

$$P^z_t f(x) = \mathbb{E}[f(X^z_t(x))] = \int f(y)q_{ta(z)}(y - x - b(z)t)dy, \tag{6.2}$$

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

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

Therefore Hypotheses 3.2 and 3.3 are clearly satisfied. Hypothesis 4.1 is clearly satisfied as $a^{i,j},\ b^i \in C_b^2(\mathbb{R}^d)$ for $i,\ j \in \{1, ..., d\}$. Now we proceed with the verification of Hypothesis 4.2. Using integration by parts, we have for $f \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned} S_t f(x) &= \int (L^y - L^x)f(y)P^x_t(x, dy) \\ &= \frac{1}{2}\sum_{i,j}\int (a^{i,j}(y) - a^{i,j}(x))q_{ta(x)}(y - x - b(x)t)\partial^2_{i,j}f(y)dy \\ &\quad +\sum_i \int (b^i(y) - b^i(x))q_{ta(x)}(y - x - b(x)t)\partial_i f(y)dy \\ &= \int dy f(y)\left(\frac{1}{2}\sum_{i,j}\partial^2_{i,j}\left((a^{i,j}(y) - a^{i,j}(x))q_{ta(x)}(y - x - b(x)t)\right)\right) \\ &\quad -\sum_i \int dy f(y)\partial_i\left((b^i(y) - b^i(x))q_{ta(x)}(y - x - b(x)t)\right). \end{aligned} \tag{6.3}$$

In view of (8.3), we have

$$\begin{aligned} \partial^2_{i,j}\left((a^{i,j}(y) - a^{i,j}(x))q_{ta(x)}(y - x - b(x)t)\right) &= \theta^{i,j}_t(x, y)q_{ta(x)}(y - x - b(x)t), \\ \partial_i\left((b^i(y) - b^i(x))q_{ta(x)}(y - x - b(x)t)\right) &= \rho^i_t(x, y)q_{ta(x)}(y - x - b(x)t), \end{aligned}$$

where we define

$$\begin{aligned} \theta^{i,j}_t(x, y) &= \partial^2_{i,j}a^{i,j}(y) + \partial_j a^{i,j}(y)h^i_t(x, y) + \partial_i a^{i,j}(y)h^j_t(x, y) + (a^{i,j}(y) - a^{i,j}(x))h^{i,j}_t(x, y), \\ \rho^i_t(x, y) &= \partial_i b^i(y) + (b^i(y) - b^i(x))h^i_t(x, y) \\ h^i_t(x, y) &= H^i_{ta(x)}(y - x - b(x)t), \qquad (6.4) \\ h^{i,j}_t(x, y) &= H^{i,j}_{ta(x)}(y - x - b(x)t). \qquad (6.5) \end{aligned}$$

So we obtain

$$S_t f(x) = \int dy f(y)q_{ta(x)}(y - x - b(x)t)\theta_t(x, y) = \int f(y)\theta_t(x, y)P^x_t(x, dy). \tag{6.6}$$

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 that the Hypotheses 4.2 and 4.3 are verified. We have verified the first part of Hypothesis 4.2 by the definition of θ in (6.6). In order to verify the rest of the conditions in Hypothesis 4.2 we see that by (8.3) and (8.4) with $\alpha = 1$

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

for a constant $C > 1$ and $c \in (0,1)$ and consequently all the conditions in Hypothesis 4.2 are satisfied with $\rho = \frac{1}{2}$ and $\gamma_t^1(x,y) = q_{ct\bar{a}}(y-x)$.

Similarly, Hypothesis 4.3 is satisfied under the the same definition of γ^1 and $\gamma^2 = \gamma^1$, $\zeta = 1$ and $\xi(0,x) = C$ for (4.3) by using that $1_{\{\sum_{i=1}^{n-1}|y_i|\le R\}} \le 1$. For (4.4), one uses that $1_{\{\sum_{i=1}^{n-1}|y_i|>R\}} \le \sum_{i=1}^{n-1}\sum_{j=1}^{d} 1_{\{|y_i^j|>\frac{R}{n\sqrt{d}}\}}$. Next, one performs the change of variables $y_1 = x_1$, $y_i - y_{i+1} = x_{i+1}$ for $i = 1, ..., n-2$ in the integral of (4.4) and use the inequality $1_{\{|y_i^j|>\frac{R}{n\sqrt{d}}\}} \le \frac{n^2 d|y_i^j|^2}{R^2}$ to obtain the following bound

$$(6.7)\quad \sum_{i=1}^{n-1}\sum_{j=1}^{d}\frac{n^2 d}{R^2}\sup_{(y_0,y_n)\in B(z_0,r)\times B(z_n,r)}\int dx_1\cdots\int dx_{n-1}\left|\sum_{k=1}^{i} x_k^j\right|^2 q_{c\delta_0\bar{a}}(x_1-y_0)\prod_{i=1}^{n-2} q_{c\delta_i\bar{a}}(x_{i+1}) q_{c\delta_{n-1}\bar{a}}(y_n - \sum_{i=1}^{n-1} x_i).$$

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 $\left|\sum_{k=1}^{i} x_k^j\right|^2 \le n\sum_{k=1}^{i}|x_k^j|^2$. Then one rewrites the integral in a probabilistic way using Gaussian random variables. This becomes $\mathbb{E}[|Z_k^j|^2/Z_0 + ... + Z_{n-1} = y_n]p_{Z_0+...+Z_{n-1}}(y_n)$ where Z_i is a d-dimensional Gaussian random vector with mean 0 and covariance matrix $c\delta_i\bar{a}I$. 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 (6.7) is bounded by $C\frac{n^4 d^2}{R^2\sqrt{\delta}}(|y_n|^2 + \delta)$ with $\delta = \sum_{i=0}^{n-1}\delta_i$. Therefore the condition (4.4) will be satisfied taking $R_\varepsilon = \varepsilon^{-\frac{1}{2}}$ and $\xi(\varepsilon,\delta) = \varepsilon(\delta^{-\frac{1}{2}} + \delta^{\frac{1}{2}})$ and the upper bound in (4.4) becomes $C2^n\xi(\varepsilon,\delta)$. 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 (8.4), $|\nabla_y p_t^x(x,y)| \le \frac{C}{t^{\frac{1}{2}}} q_{ct\bar{a}}(y-x)$. 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 $\gamma_t^i(x,y) = t^{\frac{1}{2}(1-\zeta^{-1})} q_{ct\bar{a}}(y-x)$ and $\xi(x) = C$ in order to obtain (4.3). One also obtains (4.4) as in the proof of continuity. Therefore the Hypotheses in Theorem 4.11 are satisfied.

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

$$X_{t_{k+1}}^\pi = X_{t_{k+1}-t_k}^z(x)\Big|_{z=x=X_{t_k}^\pi},$$

for $k = 0, ..., n$. Here $X_0^\pi = x$ and the noise used for $X_{t_{k+1}-t_k}^z(x)$ is independent of $X_{t_j}^\pi$ for all $j = 0, ..., k$ and of the Poisson process J.

Theorem 6.1. *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* $\bar{a} \ge a \ge \underline{a}$. *Define*

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

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

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

and therefore

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

where $(X_t^\pi)_{t\in\pi}$ is the Euler scheme with $X_0^\pi = x$ and random partition $\pi = \{\tau_i; i = 0, ..., \tau_{J_T}\} \cup \{T\}$ where $0 = \tau_0 < ... < \tau_{J_T} \leq T$ where the random times τ_i are the associated jump times of the Poisson process J 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 (6.5).

Proof. As a consequence of Theorems 4.7, 4.10 and 4.11 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 hypothesis in the next section are verified). $\square$

6.2 Example 2: The backward method for continuous SDE's 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 (6.1). For further references on this matter, see [18]. The approximating semigroup is the same as in the previous section and is given by (6.2). Therefore we have, as before,

$$\begin{aligned} p^z_t(x, y) &= q_{ta(z)}(y - x - b(z)t), \\ \phi^z_t(x) &= q_{ta(z)}(z - x - b(z)t). \end{aligned}$$

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 3.1, 3.2 and 3.3 are satisfied. Similarly, Hypothesis 5.1 can be easily verified. We will now check Hypothesis 5.2. We define

$$\begin{aligned} \hat{\theta}_t(x, z) &= \frac{1}{2}\sum_{i,j}(a^{i,j}(x) - a^{i,j}(z))\hat{h}^{i,j}_t(x, z) - \sum_i (b^i(x) - b^i(z))\hat{h}^i_t(x, z), \\ \hat{h}^i_t(x, z) &= H^i_{ta(z)}(z - x - b(z)t) \\ \hat{h}^{i,j}_t(x, z) &= H^{i,j}_{ta(z)}(z - x - b(z)t). \end{aligned}$$

So that, by (8.3)

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

Using (8.4) and the Hölder continuity of $a^{i,j}$ we obtain

$$\begin{aligned} & \left|(a^{i,j}(x) - a^{i,j}(z))\partial^2_{i,j}\phi^z_t(x)\right| \\ \leq\ & C\,|x - z|^\alpha \left|\partial^2_{i,j}\phi^z_t(x)\right| \leq C(|z - x - b(z)t|^\alpha + \|b\|^\alpha_\infty t^\alpha) \left|\partial^2_{i,j} q_{ta(z)}(z - x - b(z)t)\right| \\ \leq\ & Ct^{-(1-\frac{\alpha}{2})} q_{\overline{a}t}(z - x - b(z)t). \end{aligned}$$

And using (8.4)(ii) with $\alpha = 0$ we obtain

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

Finally we have

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

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

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

We conclude that the Hypothesis 5.2 is verified. The verification of Hypothesis 5.3 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 6.2. *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*

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

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

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

6.3 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 SDE's 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 [17].

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 non negative measure on $\mathbb{R}_+ := (0,\infty)$ which satisfies

Hypothesis 6.3. $\nu(\mathbb{R}_+) = \infty$ *and* $C_\nu := \int_{\mathbb{R}_+} c\nu(dc) < \infty$.

We refer the reader to [12] for notations 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

Hypothesis 6.4. $\int_0^\infty e^{-u}\eta_\nu(\frac{u}{s})du \geq C_* s^h \int_0^\infty e^{-u}\eta_\nu(u)du \ \forall s \geq s_*$.

For example, if $\nu(dc) = 1_{(0,1]}(c)c^{-(1+\beta)}dc$ with $0 < \beta < 1$ then Hypothesis 6.3 is satisfied and Hypothesis 6.4 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

$$\begin{aligned} V_t &= \int_0^t \int_{\mathbb{R}_+\times\mathbb{R}} cN(dx, dc, ds), \\ Z_t &= \int_0^t \int_{\mathbb{R}_+\times\mathbb{R}} xN(dx, dc, ds), \\ N_\nu(dx, ds) &= \int_{\mathbb{R}_+} N(dx, dc, ds). \end{aligned}$$

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 6.5. *Assume Hypothesis 6.3. Let $h : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ be such that $\left|\int_{\mathbb{R}\times\mathbb{R}_+}(e^{i\theta h(x,c)}-1)q_c(x)dxd\nu(c)\right| < \infty$. 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*

$$\mathbb{E}[\exp(i\theta U_t(h))] = \exp\left(t\int_{\mathbb{R}\times\mathbb{R}_+}(e^{i\theta h(x,c)}-1)q_c(x)dxd\nu(c)\right).$$

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 6.3. 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 6.3 we have that

$$\mathbb{E}[Z_t^2] = t\int_{\mathbb{R}\times\mathbb{R}_+}|u|^2\,q_c(u)\nu(dc)du = t\int_{\mathbb{R}_+}c\nu(dc) < \infty. \tag{6.8}$$

Therefore Z is a Lévy process of finite variance. $N_\nu(dx,ds)$ is a Poisson random measure with compensator $\mu_\nu(dx)ds := \int_{\mathbb{R}_+}q_c(x)\nu(dc)dxds$ 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,

$$\begin{array}{llcl}(E_\nu) & X_t^\nu(x) & = & x+\int_0^t\int_{\mathbb{R}}\sigma(X_{s-}^\nu(x))u\widetilde{N}_\nu(ds,du)\\ (E_\nu^z) & X_t^{\nu,z}(x) & = & x+\int_0^t\int_{\mathbb{R}}\sigma(z)u\widetilde{N}_\nu(ds,du).\end{array}$$

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

Hypothesis 6.6. *(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 [5]). Therefore, from now on, we suppose that a unique weak solution to (E_ν) exists so that $P_t^\nu f(x) = \mathbb{E}\left[f(X_t^\nu(x))\right]$ is a semigroup with infinitesimal operator (note that $\int u\mu_\nu(du) = 0$)

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

Therefore Hypothesis 3.1 is clearly satisfied.

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

$$L^{\nu,z}f(x) = \int_{\mathbb{R}}(f(x+\sigma(z)u)-f(x))\mu_\nu(du). \tag{6.9}$$

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 5. In order to proceed with the verification of Hypothesis 3.2 we need to prove the following auxiliary lemma.

Lemma 6.7. *Suppose that Hypotheses 6.3 and 6.4 holds for some $h > 0$. Then for every $p > 0$ there exists a constant C such that for every $t > 0$*

$$\mathbb{E}\left[V_t^{-p}\right]\le Ct^{-\frac{p}{h}}. \tag{6.10}$$

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

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

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

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

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

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

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

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

with

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

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

Now we can verify Hypothesis 3.2. 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 6.5 and the previous Lemma which is used together with Lemma 8.2 in order to obtain the needed uniform integrability properties.

Corollary 6.8. *Suppose that Hypotheses 6.3 and 6.6 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*

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

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 3.2 is satisfied and $\phi_t^y(x) = \mathbb{E}\left[q_{\sigma^2(z)V_t}(x-y)\right]$. Furthermore as it is usually the case Hypotheses 3.3 and 5.1-(0) are trivially satisfied. For Hypothesis 5.1-(i), one only needs to apply (8.2)-(ii). Hypothesis 5.1-(ii) follows from the joint continuity of $p_t^z(x,y)$ and Hypothesis 5.1-(iii) follows from the regularity of $p_t^z(x,y)$ as stated in the above Corollary 6.8 and (6.8).

We are now ready to proceed and verify Hypotheses 5.2 and 5.3. We have by (6.9), $\int uq_c(u)du = 0$ and properties of convolution that

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

In particular Hypothesis 5.2 holds with

$$(6.11)\qquad \hat{\theta}_t(x,y)=\frac{1}{\mathbb{E}\left[q_{\sigma^2(y)V_t}(x-y)\right]}\left\{\int_{\mathbb{R}_+}\mathbb{E}\left[q_{\sigma^2(x)c+\sigma^2(y)V_t}(x-y)-q_{\sigma^2(y)c+\sigma^2(y)V_t}(x-y)\right]\nu(dc)\right\}.$$

Theorem 6.9. *Suppose that Hypotheses 6.3, 6.4 and 6.6 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*

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

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

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

Proof. We have already verified Hypotheses 5.1, 5.2 and 5.10. It remains to verify Hypothesis 5.3. For this, we have to estimate

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

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

$$b\geq b'\qquad\Rightarrow\qquad\frac{a+cb}{a+cb'}\leq\frac{b}{b'}.$$

From this inequality, we obtain

$$(6.12)\qquad \begin{aligned}\frac{s''}{s'}&\leq\frac{\overline{\sigma}^2}{\underline{\sigma}^2}\qquad\text{and}\\ \frac{|s''-s'|}{s'}&\leq\frac{c\left|\sigma^2(x)-\sigma^2(y)\right|}{\sigma^2(x)c+\sigma^2(y)V_t}\leq\frac{cC_\alpha\left|a\right|^\alpha}{\underline{\sigma}^2(c+V_t)}.\end{aligned}$$

where C_α is the Hölder constant of σ^2. Finally, from Lemma 8.3 and (6.12) this gives

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

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

$$(6.13)\qquad \left|\hat{\theta}_t(x,y)\right|\phi_t^y(x)\leq C\int_{\mathbb{R}_+}\mathbb{E}\left[(c+V_t)^{-(1-\frac{\alpha}{2})}q_{\frac{\underline{\sigma}^2}{2}(c+V_t)}(x-y)\right]c\nu(dc).$$

A first step is to obtain estimates for the right hand of the above inequality, so as to be able to define γ^1. For this, we define

$$(6.14)\qquad \begin{aligned}g_t(x,y)&=\int_{\mathbb{R}_+}\mathbb{E}\left[V_t^{-(1-\frac{\alpha}{2})}q_{\frac{\underline{\sigma}^2}{2}(c+V_t)}(x-y)\right]\overline{\nu}(dc),\\ \overline{\nu}(dc)&=\frac{1(c>0)}{C_\nu}c\nu(dc),\qquad C_\nu=\int_{\mathbb{R}_+}c\nu(dc).\end{aligned}$$

We denote

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

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

$$\gamma_t^1(x,y) = t^{\frac{\chi}{h}} g_t(x,y)$$

and we notice that by (6.13)

$$\left|\hat{\theta}_t(x,y)\right| \phi_t^y(x) \leq C g_t(x,y) = C t^{-\frac{\chi}{h}} \gamma_t^1(x,y) = C t^{-\rho} \gamma_t^1(x,y).$$

We also define $\gamma_t^2(x,y) := \mathbb{E}\left[q_{\sigma^2(y)V_t}(x-y)\right] = p_t^y(x,y)$, and we use Lemma 8.2 in order to define γ^3

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

With these definitions we need to check that (5.3) and (5.4) holds. We verify the former the latter being similar to the line of proof in (6.7) if one uses (6.16) at the end of the calculation. To verify (5.3), it is enough to prove that for $n \in \mathbb{N}, \delta_i > 0, i = 1, ..., n$

$$\sup_{y_0, y_{n+1} \in \mathbb{R}^d} \int dy_1 \cdots \int dy_n \prod_{i=0}^n g_{\delta_i}(y_i, y_{i+1})^\zeta \leq \frac{C^n}{(\delta_1 + \cdots + \delta_n)^{\frac{1}{2h}}} \prod_{i=1}^n \delta_i^{-\frac{\chi}{2h}} \tag{6.15}$$

where C is a constant which depends on p, h and s_* which appear in Hypothesis 6.4 and in (6.10). Notice first that for every $a > 0$ and $x \in \mathbb{R}$ one has for a positive constant C,

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

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

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

We consider $(V_t^i)_{t\geq 0}, i = 1, ..., n$ to be independent copies of $(V_t)_{t\geq 0}$ and we write

$$\begin{aligned}
&\int dy_1 \cdots \int dy_n \prod_{i=0}^n g_{\delta_i}(y_i, y_{i+1})^\zeta \\
\leq\ & C^{-n} \mathbb{E}\left[\prod_{i=1}^n (V_{\delta_i}^i)^{-\chi} \int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) \int dy_1 \cdots \int dy_n \prod_{i=1}^n q_{\frac{\overline{\sigma}^2}{2\zeta}(c_i + V_{\delta_i}^i)}(y_i - y_{i+1})\right] \\
=\ & C^{-n} \mathbb{E}\left[\prod_{i=1}^n (V_{\delta_i}^i)^{-\chi} \int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) q_{\frac{\overline{\sigma}^2}{2\zeta}\sum_{i=1}^n (c_i + V_{\delta_i}^i)}(y_0 - y_{n+1})\right] \\
\leq\ & C^{-n} \left(\mathbb{E}\left[\prod_{i=1}^n (V_{\delta_i}^i)^{-2\chi\zeta}\right]\right)^{1/2} \left(\mathbb{E}\left[\int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) q^2_{\frac{\overline{\sigma}^2}{2\zeta}\sum_{i=1}^n (c_i + V_{\delta_i}^i)}(y_0 - y_{n+1})\right]\right)^{1/2}.
\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}\left[\int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) q^2_{\frac{\overline{\sigma}^2}{2\zeta}\sum_{i=1}^n (c_i + V_{\delta_i}^i)}(y_0 - y_{n+1})\right] \\
&= \mathbb{E}\left[\int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) q^2_{\frac{\overline{\sigma}^2}{2\zeta}(\sum_{i=1}^n c_i + V_{\delta_1 + \cdots + \delta_n})}(y_0 - y_{n+1})\right] \\
&\leq \mathbb{E}\left[\int \overline{\nu}(dc_1) \cdots \int \overline{\nu}(dc_n) \frac{2\zeta}{\overline{\sigma}^2(\sum_{i=1}^n c_i + V_{\delta_1 + \cdots + \delta_n})}\right] \\
&\leq \frac{2\zeta}{\overline{\sigma}^2} \mathbb{E}\left[\frac{1}{V_{\delta_1 + \cdots + \delta_n}}\right] \leq \frac{C\zeta}{\overline{\sigma}^2(\delta_1 + \cdots + \delta_n)^{\frac{1}{h}}}.
\end{aligned} \tag{6.16}$$

The last inequality is a consequence of (6.10). Again by (6.10)

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

so (6.15) 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 moreover we consider two independent processes $V_t = \int_0^t cn_\nu(ds, dc)$ and $V_t' = \int_0^t cn_\nu'(ds, dc)$ where n_v and n_ν' are two independent Poisson random measures of intensity $c\nu(dc)ds$. We denote by $(\Delta_j)_{j\in\mathbb{N}}$ and $(\Delta_j')_{j\in\mathbb{N}}$ two independent sequences of independent standard normal random variables. Finally, we also define a sequence of independent random variables $(Z_j)_{j\in\mathbb{N}}$ with $Z_j \sim \overline{\nu}(dc)$ given in (6.14). All the above r.v.'s are also independent between themselves Then we construct the Euler scheme associated to the random grid $\pi \equiv \pi(\omega) = (0 < \tau_1 < \cdots < \tau_{J_T} < T)$ as

$$X_t^{*,\pi}(y) = y + \sum_{i=0}^{J_t-1}\Delta_i(V_{\tau_{i+1}} - V_{\tau_i})^{\frac{1}{2}}\sigma(X_{\tau_i}^{*,\pi}(y)).$$

Notice that

$$X_{\tau_{i+1}}^{*,\pi}(y) = X_{\tau_i}^{*,\pi}(y) + \Delta_i(V_{\tau_{i+1}} - V_{\tau_i})^{\frac{1}{2}}\sigma(X_{\tau_i}^{*,\pi}(y))$$

so that,

$$P(X_{\tau_{i+1}}^{*,\pi}(y) \in dy_{i+1} \mid \{\tau_i; i \in \mathbb{N}\},\ X_{\tau_i}^{*,\pi}(y) = y_i) = p_{\tau_{i+1}-\tau_i}^{y_i}(y_{i+1}, y_i)dy_{i+1} = P_{\tau_{i+1}-\tau_i}^{y_i,*}(y_i, dy_{i+1}).$$

Then we define

$$\Gamma_T^\pi(y) = \prod_{i=1}^{J_T-1}\Bigg(q_{\sigma^2(X_{\tau_{i+1}}^{*,\pi}(y))Z_i+\sigma^2(X_{\tau_i}^{*,\pi}(y))\left(V_{\tau_{i+1}}'-V_{\tau_i}'\right)}(X_{\tau_{i+1}}^{*,\pi}(y) - X_{\tau_i}^{*,\pi}(y))$$
$$-q_{\sigma^2(X_{\tau_i}^{*,\pi}(y))Z_i+\sigma^2(X_{\tau_i}^{*,\pi}(y))\left(V_{\tau_{i+1}}'-V_{\tau_i}'\right)}(X_{\tau_{i+1}}^{*,\pi}(y) - X_{\tau_i}^{*,\pi}(y))\Bigg)\frac{C_\nu}{Z_i\phi_{\tau_{i+1}-\tau_i}^{X_{\tau_i}^{*,\pi}(y)}(X_{\tau_{i+1}}^{*,\pi}(y))}.$$

Corollary 6.10. *Under the Hypothesis of Theorem 6.9, we have*

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

and

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

The drawback of the present probabilistic representation as an expression for simulation is that we still need an explicit expression or a supplementary simulation method for $\phi_t^z(x)$.

6.3.1 Another probabilistic representation

In this section, we provide an alternative probabilistic representation which may be preferable for simulation purposes. First note that using the mean value theorem, we can rewrite (6.11) as

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

Here we define

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

Therefore we have that if we consider $U \sim Unif[\underline{\sigma}^2, \bar{\sigma}^2]$ independent of all other random variables, we can represent $\hat{\theta}_t(x,y)\phi_t^y(x)$ as

$$\hat{\theta}_t(x,y)\phi_t^y(x) = \frac{C_\nu}{2}\left(\bar{\sigma}^2 - \underline{\sigma}^2\right)\int_{\mathbb{R}} sgn_\sigma(x,y)\mathbb{E}\left[1_{\{\sigma^2(x)\wedge\sigma^2(y)\le U\le\sigma^2(x)\vee\sigma^2(y)\}} h^{1,1}_{Uc+\sigma^2(y)V_t}(x,y) q_{Uc+\sigma^2(y)V_t}(x-y)\right]\overline{\nu}(dc).$$

Here $h^{1,1}$ is the Hermite polynomial defined in (6.5). In this case the approximating Markov chain is defined as $Y_0^{*,\pi}(y) = y$

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

The corresponding weight is given by

$$\Gamma^\pi_T(y) = \left(\frac{C_\nu}{2}\left(\bar{\sigma}^2 - \underline{\sigma}^2\right)\right)^{J_T}\prod_{i=1}^{J_T} sgn_\sigma(Y^{*,\pi}_{\tau_{i+1}}(y), Y^{*,\pi}_{\tau_i}(y))1_{\{\sigma^2(Y^{*,\pi}_{\tau_{i+1}}(y))\wedge\sigma^2(Y^{*,\pi}_{\tau_i}(y))\le U_i\le\sigma^2(Y^{*,\pi}_{\tau_{i+1}}(y))\vee\sigma^2(Y^{*,\pi}_{\tau_i}(y))\}}$$
$$h^{1,1}_{U_iZ_i+\sigma^2(Y^{*,\pi}_{\tau_i}(y))V_t}\left(Y^{*,\pi}_{\tau_{i+1}}(y), Y^{*,\pi}_{\tau_i}(y)\right).$$

A similar result to Corollary 6.10 is also satisfied.

6.3.2 Examples of Lévy measures

We conclude this section with two examples of Lévy measures that satisfy Hypotheses 6.3 and 6.4.

Example 6.11. *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 6.3 are satisfied in this example. First of all, we consider Hypothesis 6.3. Clearly $\nu(\mathbb{R}_+) = \infty$, and if $\rho > 1$ then $\int c d\nu(c) < \infty$.

Now we verify Hypothesis 6.4. One has $\eta_\nu(a) = card\{k : c_k > a\} = [a^{-\frac{1}{\rho}}] - 1(a^{-\frac{1}{\rho}} \in \mathbb{N})$ for $a > 0$. We define $\eta'_\nu(a) = a^{-\frac{1}{\rho}}$. Then clearly η'_ν satisfies the Hypothesis 6.4 with $h = \frac{1}{\rho}$. Furthermore, $s^{-\frac{1}{\rho}}\eta_\nu(\frac{u}{s}) - \eta'_\nu(u)$ converges uniformly to zero as $s \to \infty$. Then as $\eta_\nu \le \eta'_\nu$ then Hypothesis 6.4 is verified for η_ν with $h = \frac{1}{\rho}$. So we may use Corollary 6.10 or Theorem 6.9 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

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

Example 6.12. *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 6.4 holds with $h = \beta \in [0,1)$. Therefore Corollary 6.10 or Theorem 6.9 can be applied for α- Hölder coefficient σ with $\alpha \in (2(1-\beta), 1)$. One may also compute*

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

so we have the following asymptotic behaviour around 0

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

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

7 Some Conclusions and Final Remarks

We have given probabilistic representations for semigroup operators that may have various applications. In fact, there are many directions of generalization which are possible. 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 non trivial operators.

Other extensions may include the study of regularity problems of the density using stochastic representations under a variation of the Malliavin Calculus.

In fact, recent efforts in infinite dimensional analysis for stochastic differential equations are concentrated in obtaining probabilistic methods for the regularity of the density of the solution to stochastic systems under weak conditions on the coefficients. Such is the case in the results of [2], [7], [8] and [11] which require some Hölder property of the coefficients in order to obtain the existence of the density.

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. In fact, one of the future directions is to study the regularity of the law associated with solutions of SDE's driven by stable processes with Hölder type coefficients, This topic will be treated in future works.

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. Our intention in this article was to provide in a setting as simple as possible the probabilistic representation formulas in such as a way that the reader may be able to see possible uses of these formulas. We have not written our results in full generality so as not to cloud our presentation in technical arguments. This will be done in future work with the purpose of solving specific problems.

8 Appendix

8.1 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 8.1. *Let $a \in [0,1)$ and $b > -1$. Then we have*

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

In particular, for $b=0$

$$c_n(t_0,a) := c_n(t_0,a,b) \le t_0^{n(1-a)} \frac{\Gamma^n(1-a)}{[1+n(1-a)]!} \qquad for \qquad n \ge (1-a)^{-1}. \tag{8.1}$$

Proof. Let $b > -1$ and $0 < a < 1$ 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

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

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 (8.1). $\square$

8.2 Some properties of Gaussian type kernels

In this section, we introduce some preliminary estimates concerning Gaussian kernels. We consider a d dimensional square symmetric non negative definite matrix a. We assume that $0 < \underline{a}I \leq a \leq \overline{a}I$ 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 variance a is denoted by

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

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:

$$\begin{aligned} (i) \qquad q_s(y) &\leq \left(\frac{t}{s}\right)^{\frac{d}{2}} q_t(y) \qquad \forall s < t \\ (ii) \qquad \rho_a^{-\frac{d}{2}} q_{\underline{a}}(y) &\leq q_a(y) \leq \rho_a^{\frac{d}{2}} q_{\overline{a}}(y). \end{aligned} \tag{8.2}$$

We define for $a \in R^{d\times d}$, 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

$$\partial_i q_a(y) = H_a^i(y) q_a(y) \qquad \partial_{i,j}^2 q_a(y) = H_a^{i,j}(y) q_a(y). \tag{8.3}$$

We will use the following basic estimates:

Lemma 8.2. *For $\alpha \in [0,1]$ we have for all $i, j \in \{1, ..., d\}$, $y \in \mathbb{R}^d$ and $t > 0$*

$$\begin{aligned} (i) \qquad |y|^\alpha \left|\partial_{i,j}^2 q_{ta}(y)\right| &\leq C_a \frac{1}{t^{1-\frac{\alpha}{2}}} q_{\frac{t\overline{a}}{2}}(y) \qquad \textit{and} \\ (ii) \qquad |y|^\alpha \left|\partial_i q_{ta}(y)\right| &\leq C_a' \frac{1}{t^{\frac{1-\alpha}{2}}} q_{\frac{t\overline{a}}{2}}(y). \end{aligned} \tag{8.4}$$

with

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

Proof. We have

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

so that

$$|y|^\alpha \left|\partial_{ij}^2 q_{ta}(y)\right| \leq \frac{1}{\underline{a}t^{1-\frac{\alpha}{2}}} \frac{|y|^\alpha}{t^{\frac{\alpha}{2}}} \left(\frac{|y|^2}{\underline{a}t} + 1\right) q_{ta}(y).$$

We use (8.2) and we obtain

$$q_{ta}(y) \leq \rho_a^{\frac{d}{2}} q_{t\overline{a}}(y) = (2\rho_a)^{\frac{d}{2}} \exp\left(-\frac{1}{4t\overline{a}}|y|^2\right) q_{\frac{1}{2}t\overline{a}}(y).$$

We may find a constant c_α such that $v^\lambda e^{-v} \leq c_\alpha$ for every $0 \leq \lambda \leq 2 + \alpha$. 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

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

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

Lemma 8.3. *Let $0 < s' \leq s''$ and $y \in \mathbb{R}$ then*

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

Proof. Using the fact that q solves the heat equation, we have

$$\begin{aligned} |q_{s''}(y) - q_{s'}(y)| &\leq \int_{s'}^{s''} |\partial_s q_s(y)|\, ds = \frac{1}{2} \int_{s'}^{s''} \left| \partial_a^2 q_s(y) \right| ds \leq \frac{1}{2} \int_{s'}^{s''} (\frac{1}{s} + \frac{a^2}{s^2}) q_s(y) ds \\ &\leq C \int_{s'}^{s''} \frac{1}{s} q_{\frac{s}{2}}(y) ds. \end{aligned}$$

In order to obtain the last inequality we have used the following elementary inequality. For every $\theta, u > 0$ one has $v^\theta e^{-\frac{y^2}{u}} \leq C_{\theta,u} < \infty, \forall y \in R$ so that

$$\frac{y^2}{s} q_s(y) = \frac{y^2}{s} e^{-\frac{y^2}{4s}} 2 q_{\frac{s}{2}}(y) \leq C q_{\frac{s}{2}}(y).$$

Moreover for $s < t$, one has $q_s(y) \leq \sqrt{\frac{t}{s}} q_t(y)$. This gives

$$\begin{aligned} \int_{s'}^{s''} \frac{1}{s} q_{\frac{s}{2}}(y) ds &\leq \sqrt{\frac{s''}{s'}} q_{\frac{s''}{2}}(y) \int_{s'}^{s''} \frac{1}{s} ds \leq \sqrt{\frac{2s''}{s'}} q_{s''}(y) \ln \left(\frac{s''}{s'} \right) \\ &= \sqrt{\frac{2s''}{s'}} q_{s''}(y) \ln \left(1 + \left(\frac{s''}{s'} - 1 \right) \right) \leq \sqrt{\frac{2s''}{s'}} q_{s''}(y) \left(\frac{s'' - s'}{s'} \right). \end{aligned}$$

From here the result follows. □

References

[1] Applebaum, D., Lévy processes and stochastic calculus. Cambridge University Press, 2009.

[2] Bally, V., and Caramellino, L., Regularity of probability laws by using an interpolation method. Preprint arXiv:1211.0052v1, 2012.

[3] Bally, V., and Talay, D., The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function, Probab. Theory Related Fields, 104, 1996, 1, 43–60, doi=10.1007/BF01303802.

[4] Bally, V., and Talay, D., The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density, Monte Carlo Methods Appl., 2, 1996, 2, 93–128, doi=10.1515/mcma.1996.2.2.93.

[5] Bass, R., Stochastic differential equations with jumps , Probability Surveys, 1, 2004, 1-19 .

[6] Beskos, A. and Roberts, G.O., Exact simulation of diffusions Ann. Appl. Probab., 15, 2005, 2422-2444.

[7] Debussche, A., and Fournier, N., Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. Journal of Functional Analysis, To appear 2013.

[8] Debussche, A., and Romito, M., Existence of densities for the 3D Navier–Stokes equations driven by Gaussian noise, 2012, Preprint arXiv:1203.0417.

[9] Foschi P., Pascucci A., and Corielli F., Parametrix approximation of diffusion transition densities SIAM J. Financial Math., vol.1 pp.833-867, 2010.

[10] Friedman, A., Partial differential equations of parabolic type, Dover Publications, Inc., 1964.

[11] Fournier, N., and Printems, J., Absolute continuity for some one dimensional processes, Bernoulli, 16(2), 343-360, 2010.

[12] Ikeda, N., and Watanabe, S., Stochastic Differential Equations and Diffusion Processes, North-Holand/Kodansha, 1989.

[13] Konakov, V., Menozzi S., and Molchanov, S., Explicit parametrix and local limit theorems for some degenerate diffusion processes. Annales de l'Institut Henri Poincaré (Série B). 46-4, 908-923, 2010.

[14] Konakov, V. and Menozzi, S., Weak Error for stable driven SDEs: expansion of the densities. Journal of Theoretical Probability. 24-2, 454-478, 2011.

[15] Léandre, R., Malliavin calculus of Bismut type without probability. Proc. Indian Acad. Sci. (Math. Sci.), 116, 507518, 2006.

[16] Menozzi, S., Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. Electronic Communications in Probability, 16, 234-250, 2011.

[17] Ressel, P., A short proof of Schoenberg's theorem. Proc. Amer. Math. Soc. 57, 66-68, 1976.

[18] Stroock, D.W. and Varadhan, S.R.S. Multidimensional Diffusion Processes. Springer- Verlag, Berlin, 1979.

[19] Talay, D., and Tubaro, L., Expansion of the global error for numerical schemes solving stochastic differential equations. Stochastic Analysis and Applications, 8(4):94–120, 1990.