Regularity of the density of a stable-like driven sde with Hölder continuous coefficients

Arturo Kohatsu-Higa*
Department of Mathematical Sciences.
Ritsumeikan University and Japan Science and Technology
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan

Libo Li *†
Department of Mathematics and Statistics
University of New South Wales
Sydney, Australia

Abstract

In this article we use the backward parametrix method in order to prove the existence and regularity of the the transition density associated to the solution process of a stable-like driven stochastic differential equation (sde) with Hölder continuous coefficients. The method of proof uses the parametrix method on the Gaussian component of a subordinated Brownian motion. This analysis which can be generalized also provides a stochastic representation of the density which is potentially useful for other applications.

Keywords: SDE's with jumps, Lévy process, Parametrix method, stochastic representation
AMS Subject Classification (2010): Primary: 60J35, Secondary: 60J75, 60J22

*This research benefited from the support of grants from the Japanese government. The authors would like to thank V. Bally for fruitful discussions.

†Parts of this paper was written while the author was a research associate at Ritsumeikan University and the Japan Science and Technology Agency, Japan.

1 Introduction

Lévy driven stochastic differential equations (sde's) with irregular coefficients is a recent topic of interest in stochastic analysis due to its various applications. In this work, we consider the problem of existence and regularity of the transition densities of the solution to a stable-like driven sde with Hölder continuous coefficients. In this case, it was remarked by Debussche and Fournier [6], that Malliavin calculus is not applicable due to the lack of differentiability of the coefficients. Using different techniques, the authors of [6] show the existence of the density.

To be precise, [6] considers the following sde with jumps

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dZ_s. \tag{1}$$

Here, $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ and Z is a d-dimensional Lévy process with associated Lévy measure ν, which satisfies some stable like assumptions. In particular, ν can be taken to be the multidimensional symmetric α-stable measure.

The main theorem in [6] hinges on an application of Lemma 2.1 within, which is an existence result, that claims that a measure on $\mathbb{R}^d$ exhibits a density and the density function belongs to a Besov space of order less than one, whenever it satisfies certain regularity conditions given in (2.4) of the previously cited lemma. By using Lemma 2.1, the authors have shown that the transition density of (1) exists and in heuristic terms, one may say that the transition density is Hölder in $L^1(\mathbb{R}^d)$-norm of an order depending on the parameters of the problem. For an exact description, we refer the reader to [6].

Konakov and Menozzi [12], also used the parametrix method to study the transition density of sde's of the form (1). The driving process, Z is assumed to be a symmetric stable process with Lévy triple given by $(0, \nu, \gamma)$, where ν is the Lévy measure.

Under some regularity assumptions on the coefficients, the densities of the sde given in (1) exist and they become the fundamental solution to the backward Kolmogorov equation. By using the fact that the densities satisfy the backward Kolmogorov equation, the authors in [12] derived using the parametrix method, a *formal* asymptotic expansion of the transition densities with respect to the transition densities of the so call *parametrix*. By showing that the *formal* asymptotic expansion converges under their hypotheses, the authors were able to obtain an explicit representation of the transition densities as the limit of a convergence sequence of partial sums.

Our approach relies on the recent work of Bally and Kohatsu-Higa [1], where the authors considered the parametrix method from a probabilistic point of view and developed systematically the *forward method* and the *backward method*, which can be used to obtain asymptotic expansions of the semigroup of the process X, without prior knowledge of the existence of the transition densities. Given that the infinite series converges, one retrieves the transition densities by identification. The authors also provide a probabilistic interpretation of the asymptotic expansion of the transition densities, which enables one to find probabilistic representations that may lead to exact *Monte Carlo* type simulation methods in order to estimate the transition densities. Furthermore, one may also see in the structure of this probabilistic representation formula, a weight that is usually obtained through infinite dimensional integration by parts formulas as in Malliavin Calculus in the smooth coefficient case. In that paper, consideration is given to sde's driven by jump processes which have stable measures around zero with compact support. This assumption is taken in order to avoid the moment explosion problem of stable processes.

In the present work we will concentrate on applying the *backward parametrix method* developed in [1] to the study of the existence and regularity with respect to the initial and final points of the transition density of the following process,

$$X_t = X_0 + \int_{(0,t]} b(X_{s-})ds + \int_{(0,t]} \sigma(X_{s-})dZ_s, \quad t \in [0, T]. \tag{2}$$

Where the process $Z = B_V$ is an Brownian motion subordinated by an independent general subordinator V. As a particular case, one obtains results for the case where the driving process Z is a symmetric, rotationally invariant α stable process.

We assume that b is bounded continuous and σ is Hölder continuous. Besides the difference in the drift coefficient hypotheses, our results differ from the main result in [6] in that we provide an *explicit* representation of the transition density. In comparison with the method applied in [12], where an explicit expression can also be obtained, our method does not assume a priori that the density exists and is able to deal with the fact that the σ is Hölder. The use of the subordinated Brownian motion has the advantage that the same technique can be easily extended to the case of jump diffusions processes, once we suppose the uniform elliptic assumptions holds for both the diffusion coefficient and the and jump coefficient σ. Using this technique also allows us to consider far more general subordinator classes and not just the explicit cases of stable or tempered stable driving processes.

The main difficulty at hand is that one needs to have estimates of the increments of the associated Euler scheme which may seem difficult to obtain. This difficulty is solved as we base our estimates on Gaussian estimates together with inverse moment estimates for stable like subordinator. This leads to moment estimates which have to be dealt with by considering two cases, small and big jumps. In each case, delicate estimates have to be obtained and combinatorics in the respective iteration of the Euler scheme have to be considered.

The paper is organized as follows. In section 3, we introduce the assumptions of this paper, the construction of the stable-like processes based on the subordinated Brownian motion and our main results in Theorem 3.9. For ease of understanding, we include in Remark 3.10, heuristic arguments to shed light on the method of proof and roughly explain the reasons behind the conditions that are required in Theorem 3.9.

In section 4, we show in details, the backward parametrix method developed (in a stochastic setting) in Bally and Kohatsu [1] can be applied to the current setting. In particular, in subsection 4.2, we derive in detail the *'formal'* expansion of the semigroup of X and in subsection 4.3, a sketch of the overall strategy of the paper will be provided to guide the reader through the overall strategy employed in the rest of the paper.

Having obtained the *'formal'* expansion of the semigroup, section 5 is devoted to show that the formal expansion indeed converges, which give rise to a representation of the semigroup associated with X. By using the representation of the semigroup derived in section 5, we show in section 6, that the density of the process X exists and can be represented as an convergent series (see (11)).

Once the density $p_t(x, y)$ of X_t has shown to exist, the regularity properties of the density $p_t(x, y)$ are studied in section 7. In section 7, the density $p_t(x, y)$ is shown to be differentiable in x, locally Hölder in y and continuous in t in Theorem 7.2, Theorem 7.3 and Theorem 7.6 respectively. The joint continuity of the density $p_t(x, y)$ in (x, y, t) for $t > 0$ is then obtained in Corollary 7.12 by combining Theorem 7.2, Theorem 7.3 and Theorem 7.6.

To conclude, in section 8, in similar spirit to [1], we provide several stochastic representations of the density function. Finally in the appendix, we gather some auxiliary lemmas.

In the last period of writing the present article, independent from our work the related articles by Knopova and Kulik [10] and more recently Huang [8] have appeared. We briefly explain the difference between the current work and [8, 10].

In [10], the existence and time regularity of the density is studied under Hölder regularity hypotheses on the coefficients and the driving process is the α-stable process with $\alpha \in (0, 1)$. The parametrix methodology is also used in [10], but the structure of the α-stable driving process is strongly used. In [10] some regularity of the drift coefficient is assumed while we only assumed the drift is bounded continuous.

In [8], the larger class of tempered stable processes was considered using the parametrix method. The main aim was to obtain upper and lower bounds for the transition density. The assumptions within [8] differs in that the drift is assumed to be Lipschitz continuous and the class of Lévy measures considered within do not cover the Brownian motion subordinated by the Lamperti-stable subordinator discussed in Example 3.4. Regularity properties of the density were not studied in [8].

In contrast, the current work considers both regularity in space and time of the density. We remark that both [10] and [8] do not use the Schoenberg theorem which leads to the subordinated expression of the driving process. Our approach of using the subordinated Brownian motion enables one to study the diffusion and the jump component simultaneously using only Gaussian estimates and inverse moment estimates of the subordinator. Therefore, further generalization can be obtained by considering subordinated Gaussian processes using a larger class of subordinators, given that the inverse moments of the subordinators can be estimated.

The probabilistic representations are considered, which are also not discussed in [10] and [8].

2 Notations and Definitions

In the rest of the paper, we write $\{e_i\}_{1,\dots,d}$ for the standard basis of $\mathbb{R}^d$, $B(x,\epsilon)$ for the open ball of size ϵ centred at the point $x \in \mathbb{R}^d$ and $\mathbb{S}_d$ for the d-dimensional unit sphere. All vectors are assumed to be column vectors unless stated otherwise and vector norms are denoted by $|\cdot|$. Time variables will be usually taken on the on the interval $(0,T]$ and we often write $\mathbb{R}_+ := (0,\infty)$. Sometimes we prove some convergence results for series uniformly in time and space by finding converging upper bounds. We will then say that a series converges uniformly if the convergence is uniform in all the space variables and for compact sets in time which do not include a neighbourhood of 0.

For $k \in \mathbb{N} \cup \{\infty\}$, we denote by $\|f\|_k$, the L^k norm taken on the space variables of the function $f : \mathbb{R}^d \to \mathbb{R}$. The space of real valued functions which are k-times differentiable with compact support will be denoted by $C_c^k(\mathbb{R}^d)$. If the sub-index c is replaced by b then the compact support property is replaced by the boundedness property. The space of bounded real valued functions will be denoted by $B(\mathbb{R}^d)$ and the space of real valued continuous functions which go to zero as $|x| \to \infty$ is denoted by $C_0(\mathbb{R}^d)$. Note that if $f \in C_0(\mathbb{R}^d)$ then $\|f\|_\infty := \sup_{x\in\mathbb{R}^d} |f(x)| < \infty$.

The multi-dimensional Gaussian density with covariance matrix A and mean 0 will be denote by $q(A,\cdot)$. We denote the $d \times d$ identity matrix by I and notation-wise and we do not make any particular distinction between the multi-dimensional case and the one dimensional case, that is for $v \in \mathbb{R}_+$ we shall write $q(v,\cdot)$ instead of $q(vI,\cdot)$.

Here, A is a positive definite matrix and for such matrices we use the notation $\|A\|_2$ to denote the spectral norm of the matrix A and $\|A\|_1 := \max_j \sum_i |A_{ij}|$. We remark that these norms are equivalent to the Frobenius norm of the matrix A defined as $\|A\|_F := \sqrt{\mathrm{Tr}\,(A^T A)}$ where A^T denotes the transpose of the matrix A. Expectations are denoted by $\mathbb{E}$ and the expectation conditioned to the random vector X is denoted by $\mathbb{E}[\,\cdot\,|X]$.

Finally, to ease notation used in the computations, when integrating over $\mathbb{R}^d$, we will omit the writing the domain of integration. Constants are denoted by C and as it is usual these constants may change its exact value from one line to the next without any further remark. They may depend on the dimension and the uniform ellipticity assumptions and the Hölder continuity assumption on the coefficients without any further comment.

To avoid changing the notation, we will keep using x, y for the variables in functions which do not necessarily mean that they correspond to the departing and arrival point for the transition function of X. See also the Glossary (section 11) for further notation which will be introduced through this presentation.

3 Assumptions, Method and Results

Definition 3.1. *Let B be a d-dimensional Wiener process and V be a subordinator. We say that the subordinated process $\{Z_t := B_{V_t};\ t \in [0,T]\}$ is a α-stable like process if the subordinator V satisfies* *the following hypotheses.*

Hypotheses 3.2. The Lévy measure μ of the subordinator V satisfies the following properties: There exists $\alpha \in (0,1)$ such that
(i) For every $j < \alpha$, $\int_{(1,\infty)} c^j \mu(dc) < \infty$.
(ii) For every $t > 0$, there exists constants C, A independent of t, and a positive concave increasing function m, such that for any $s > 0$

$$\mathbb{E}[e^{-sV_t}] \le Ce^{-s^\alpha m(s)tA}.$$

(iii) For every $w \in (\alpha, 1)$ and $s > 0$, there exists a constant C_w,

$$\int_{(0,1]} c^w e^{-sc}\, \mu(dc) \le C_w s^{\alpha - w}.$$

(iv) The Lévy measure μ is such that $\mu(\mathbb{R}_+) = \infty$, $\mu(\mathbb{R}_-) = 0$.

Note that by identification, the Lévy measure of the α-stable like process Z is given by $\nu(du) = \int_{\mathbb{R}_+} q(c,u)\mu(dc)du$.

Lemma 3.3. *As consequences of Hypotheses 3.2, we have*
(i) for every $0 < j < \alpha$, $\mathbb{E}[V_T^j]$ is finite,
(ii) the integral, $\int_{(0,1]} c\, \mu(dc)$ is finite.

Proof. By Theorem 25.3. in [15], Hypotheses 3.2. (i) implies that for every $0 < j < \alpha$, $\mathbb{E}(V_T^j) < \infty$.

For the second claim, we note that Hypotheses 3.2. (iii) implies that $\int_{(0,1]} c^r\, \mu(dc) < \infty$ for all $r > \alpha$. In fact, using the fact that $e^{-sc} \in (e^{-1}, 1)$ and that $c^{r-w} \le 1$ for $w \in (\alpha, r \wedge 1)$, we obtain that the integral is finite. □

For ease of presentation, we set $\widehat{\mu}(dc) := \big(\mathbb{1}_{(0,1]}(c)c + \mathbb{1}_{(1,\infty)}(c)\big)\, \mu(dc)$, which is a finite measure, as it is integrable around zero due to Lemma 3.3. (ii) and at infinity due to Hypotheses 3.2 (i).

Example 3.4. *Subordinators for which the above Hypotheses 3.2 are satisfied, include α-stable, tempered α-stable, and Lamperti α-stable subordinators (see Caballero et al. [5]) for $\alpha \in (0,1)$. We show here that the Lamperti α-stable subordinators with Lévy measure $\mu(dc) = \frac{e^c}{(e^c-1)^{1+\alpha}} \mathbb{1}_{\{c>0\}} dc$ satisfies Hypotheses 3.2.*

Firstly, by l'Hôpital's rule, the function $\frac{c^{1+\alpha}e^c}{(e^c-1)^{1+\alpha}}$ belongs to $C_0(\mathbb{R}_+)$ and thus there exist a constant k_α depending on $\alpha \in (0,1)$ such that $\mu(dc) \le \frac{k_\alpha}{c^{1+\alpha}} dc$, which shows that Hypotheses 3.2 (i), (iii) are satisfied. Secondly, from integration by parts formula, the Lévy symbol of the Lamperti α-stable subordinator satisfies

$$\psi(s) = \frac{\Gamma(s+\alpha)}{\Gamma(s)} \frac{\Gamma(1-\alpha)k}{\alpha} \ge s^\alpha \left(\frac{s}{s+\alpha}\right)^{1-\alpha} \frac{\Gamma(1-\alpha)k}{\alpha}$$

where the last inequality follows from the lower bound for ratio of Gamma functions given in Wendel [20]. This shows that Hypotheses 3.2 (ii) is satisfied with $m(s) = \left(\frac{s}{s+\alpha}\right)^{1-\alpha}$ and $A = \frac{\Gamma(1-\alpha)k}{\alpha}$.

It follows from the definition of μ that $\mu(\mathbb{R}_-) = 0$ and

$$\int_{(0,1]} \frac{e^c}{(e^c-1)^{1+\alpha}} dc = -\alpha^{-1}(e^c-1)^{-\alpha}\Big|_0^1 = \infty,$$

which shows that Hypotheses 3.2 (iv) is satisifed.

From the Schoenberg's theorem [16], we see that the above construction is flexible and can be greatly generalized by appropriately choosing the Lévy measure μ of the subordinator V. It is with this thought that we have introduced the assumptions on the measure μ.

In the rest of the paper, we chose to present the results under the following assumptions. For $\alpha \in (0,1)$, we assume the following conditions on the coefficients.

Hypotheses 3.5.
(i) $b : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function for $\frac{1}{2} < \alpha < 1$ and $b = 0$ for $0 < \alpha \leq \frac{1}{2}$.
(ii) $a := \sigma\sigma^T : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is globally k-Hölder continuous, for $k \in (0,1]$.
(iii) There exists constants $0 < \underline{a} < \overline{a}$, such that, for all $x \in \mathbb{R}^d$,

$$\underline{a} I \leq a(x) \leq \overline{a} I$$

(iv) a weak solution exists

As mentioned in [6], the existence of solutions to (2) under their hypotheses is not known although they conjecture the existence of weak solution and refer to results in Jacod [9].

In the case of stable process, that is given that V is a true α-stable subordinator, in the case where X is 1-dimensional and $b = 0$, we can refer to Bass [2], or Komatsu [11] and Pragarauskas and Zanzotto [14] or Zanzotto [22] for the existence and uniqueness of weak solution respectively. We have e.g. that if σ is uniformly elliptic and Hölder of order k with $k\alpha > 1$ then pathwise uniqueness and weak existence is satisfied. For $d > 1$, we may refer to Williams [21] for existence results. For more specialized results, we may cite e.g. Bass et al. [3] or Fournier [7].

In the current setting, we assume Hypotheses 3.5. (iv) only for completeness and the ease of presentation. The reader can refer to the works of Bass and Tang [4] on martingale problems for stable-like process, where with relative ease the existence and uniqueness of a weak solution can be derived by modifying their techniques (see Remark 4.1 in [4]).

The main difference between our assumptions (Hypotheses 3.5) from those of [12] is that both the drift and noise coefficients are assumed to be less regular. Note also that in [12], the drift is also assumed to be zero for $2\alpha \in (0,1]$.

Comparing to the assumptions of [6], the main difference is that we assume only that the drift is bounded continuous instead of bounded and Hölder continuous, and as a consequence we obtain less refined existence results for the density of X_t comparing to those of [6]. The assumptions on the noise coefficient in [6] and Hypotheses 3.5 are similar, except that we assume the stronger uniformly ellipticity condition, while in [6] the density exists on the set where the coefficient σ is invertible, and the Besov property will depend on the rate of degeneration of the coefficient σ.

3.1 Main Result

Let us denote by $\mathcal{L}$ the generator of X and its dual operator by $\mathcal{L}^*$. For $f \in C^2(\mathbb{R}^d)$, denote by $\nabla_x^2 f$ the Hessian Matrix of f. Then the infinitesimal generators are given by

$$\begin{aligned}
&\mathcal{L}f(x) = \mathcal{L}^x f(x) \\
&\mathcal{L}^z f(x) := b(z)^T \nabla_x f(x) + \int_{\mathbb{R}^d\times\mathbb{R}_+} \{f(x+\sigma(z)y) - f(x)\}\, q(c,y)dy\mu(dc).
\end{aligned} \tag{3}$$

Note that for all $\alpha \in (0,1)$ the term

$$\int_{\mathbb{R}^d\times\mathbb{R}_+} \nabla_x f(x)^T \sigma(z) y q(c,y) \mathbb{1}_{\{|y|\leq 1\}} dy\mu(dc)$$

is zero due to the symmetry of the Gaussian density $q(c,\cdot)$. The parametrix process that we will use is the *'frozen'* process given by

$$\widehat{X}^z_t(x) = x + \sigma(z) Z_t.$$

The generator of the parametrix is given by

$$\widehat{\mathcal{L}}^z f(x) = \int_{\mathbb{R}^d\times\mathbb{R}_+} \{f(x+a(z)y)-f(x)\}\, q(c,y)dy\mu(dc), \tag{4}$$

and furthermore, its density $\widehat{X}_t^z$ can be explicitly be given as follows.

Lemma 3.6. *The transition density of $\widehat{X}_t^z(x) = x + \sigma(z)Z_t$ is given by*

$$\widehat{p}_t^z(x,y) = \mathbb{E}[q(a(z)V_t, x-y)]$$

for $(x,y,z,t)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}_+$. In particular, this density is smooth with respect to $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$ and

$$\nabla_x \widehat{p}_t^z(x,y) = -\mathbb{E}\left[(a(z)V_t)^{-1}(x-y)q(a(z)V_t,x-y)\right]. \tag{5}$$

Remark 3.7. The term $\widehat{p}_t^z(x,y) = \mathbb{E}[q(a(z)V_t, x-y)]$ is the density of the frozen process, which is the subordinated Brownian motion $B_{a(z)V_t}$. Any upper bound available for this quantity can be shown to become an upper bound for the corresponding density of the solution X to the stochastic differential equation. If V is a stable subordinator, then the upper bounds for the density are well known and are readily available in Sztonyk [18] or Watanabe [19], that is for $\alpha\in(0,1)$,

$$\widehat{p}_t^z(x,y) \le ct^{-d/2\alpha}(1+t^{-1/2\alpha}|y-x|)^{-2\alpha-\gamma},$$

where γ is a constant such that, given the Lévy measure ν of $Z = B_V$ we have

$$\nu(B(x,r)) \le cr^\gamma \qquad \forall x\in\mathbb{S}_d \qquad r\le\frac{1}{2}.$$

One can obtain similar bounds to the above by bounding the term $\mathbb{E}[q(a(z)V_t, x-y)]$ using inverse moment estimates of V and estimates of e^{-x}.

The function $\widehat{\theta}$ is defined for all $(z_1,z_2,t)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}_+$ as

$$\widehat{\theta}_t(z_2,z_1)\widehat{p}_t^{z_1}(z_2,z_1) := (\mathcal{L}^z - \widehat{\mathcal{L}}^{z_1})(\widehat{p}_t^{z_1}(\cdot,z_1))(z_2)\Big|_{z=z_2}. \tag{6}$$

and by Lemma 9.1, we can write

$$\begin{aligned}\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y) &:= b(x)^T\nabla_x\widehat{p}_t^y(x,y) && (7)\\ &\quad + \int_{\mathbb{R}_+}\mathbb{E}\big[q(a(y)V_t + a(x)c, x-y) - q(a(y)V_t + a(y)c, x-y)\big]\mu(dc). && (8)\end{aligned}$$

For every $(x,y,t)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}_+$, we define $I_t^n(y,x)$ as

$$I_t^n(y,x) := \int_0^t \cdots \int_0^{t_{n-1}} dt_n\dots dt_1\; K(t_n,\dots,t_1,t;x,y) \tag{9}$$

$$K(t_n,\dots,t_1,t;x,y) := \int_{\mathbb{R}^n} \widehat{p}_{t_n}^{z_n}(x,z_n)\prod_{i=0}^{n-1}\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)\widehat{p}_{t_i-t_{i+1}}^{z_i}(z_{i+1},z_i)dz_{i+1}. \tag{10}$$

where $x = z_{n+1}$, $y = z_0$, $t = t_0$.

Remark 3.8. *The reason for the choice of the order of indices in the multiple time integral in (9) is due to the order in the probabilistic representation to be given in section 8.*

Theorem 3.9. *Assume Hypotheses 3.5 and let Z be a stable like process as described in Hypotheses 3.2. Then for $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$, the following properties are satisfied*
(i) (Theorem 6.1) The density, $p_t(x, y)$ of X_t for $X_0 = x$, exists and has the following representation

$$p_t(x,y) = \sum_{n=0}^{\infty} I_t^n(y,x). \tag{11}$$

Here, the sum converges absolutely and uniformly.
(ii) (Theorem 7.2) For $\alpha > \frac{1}{2}$, the density $p_t(x, y)$ of X_t is Lipschitz continuous in $x \in \mathbb{R}^d$. Furthermore it is differentiable with respect to x and

$$\nabla_x p_t(x,y) = \sum_{n=0}^{\infty} \nabla_x I_t^n(y,x).$$

Here, the sum converges absolutely and uniformly.
(iii) (Theorem 7.3 and Corollary 7.4) The density, $p_t(x, y)$ of X_t is locally β-Hölder continuous in $y \in \mathbb{R}^d$ for $\beta \in (0, (2\alpha - 1) \wedge k)$ if $\alpha > \frac{1}{2}$, and $\beta \in (0, 2\alpha \wedge k)$ if $\alpha < \frac{1}{2}$.
(iv) (Theorem 7.6) The density, $p_t(x, y)$ of X_t is continuous in $t \in \mathbb{R}_+$.

As pointed out in the introduction, an advantage of using the subordinated Brownian motion is that it allows one to study, by using the same technique, the existence and regularity of transition densities of the sde (2) with a non-trivial diffusion component. However, in that case, the diffusion component will dominate the situation and the parameter α will play little role in the computations.

Therefore, we will not present the computations but only point out that if one wishes to work with a non-trial diffusion component then under Hypotheses 3.11, the main results in Theorem 3.9 can be re-stated with $\alpha = 1$ and k replaced by $k \wedge k'$ (where k' is the Hölder exponent for the coefficient associated with the diffusion component). While regularity results can be formulated with a little to no modification, see Remark 7.5 and Remark 7.11.

Before proceeding to the rest of the paper, we explain in the following remark, in heuristic terms, why do we encounter certain conditions relating the Hölder exponent of σ to the parameter α related to the 2α-stable like process Z. We also demonstrate the reason why if we were to consider the jump diffusion case, the diffusion component will dominate the situation and the computations will effectively be reduced to that of the diffusions.

Remark 3.10. To demonstrate the computations in the jump diffusion case, we consider equation (2), but with a non-trivial diffusion component, that is

$$X_t = X_0 + \int_{(0,t]} b(X_{s-})ds + \int_{(0,t]} \zeta(X_{s-})dB_s + \int_{(0,t]} \sigma(X_{s-})dZ_s, \quad t \in [0,T]. \tag{12}$$

and we work under the following assumption,

Hypotheses 3.11.
(i) $b : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function.
(ii) $a := \sigma\sigma^T$ and $e := \zeta\zeta^T : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are globally k and k'-Hölder continuous, for $k, k' \in (0, 1]$.
(iii) There exists constants $0 < \underline{a} < \overline{a}$ and $0 < \underline{e} < \overline{e}$, such that, for all $x \in \mathbb{R}^d$.

$$\underline{a}I \le a(x) \le \overline{a}I \quad \text{and} \quad \underline{e}I \le e(x) \le \overline{e}I$$

The existence and uniqueness of the solution to the sde (2) with a non-trivial diffusion component (under Hypotheses 3.11) follows directly from the results of Stroock [17] with $m(x, du) = \int_0^\infty q(ca(x), u)\mu(dc)du$. The reader is also referred to Theorem 5.1 in a recent survey by Bass [**?**].

For ease of computation, we assume in this remark that V is a α-stable subordinator and $d = 1$. It is not hard to see that $\widehat{p}_t^z(x, y) = \mathbb{E}[q(a(z)V_t + e(z)t, x - y)]$ is the density of the process $\widehat{X}_t^z = x + \zeta(z)W_t + \sigma(z)B_{V_t}$, where B and W are independent Brownian motions, whom are independent of V. Since V is independent of B, one can carry out a calculation conditioned on $\sigma(V_s; s \geq 0)$. In this setting, the generators of $X_t^z := x + b(z)t + \zeta(z)W_t + \sigma(z)B_{V_t}$ and $\widehat{X}^z$ with V *'fixed'* behave locally like the generators for continuous diffusion process (this part of the argument is not rigorous) and therefore $(\mathcal{L}^{y_1} - \widehat{\mathcal{L}}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x}$ should behave like

$$b(x)\nabla_x \partial_x p_t^y(x, y) + \frac{1}{2}(e(x) - e(y))\partial_{xx}\widehat{p}_t^y(x, y) + \frac{1}{2}(a(x) - a(y))\partial_{xx}\widehat{p}_t^y(x, y)$$

We observe that by using Lemma 9.2 (iii). The first term associated with the drift can be estimated as follows

$$|b(x)\nabla_x \widehat{p}_t^y(x, y)| \leq C\mathbb{E}\big[(V_t + t)^{-\frac{1}{2}} q(C(V_t + t), x - y)\big],$$

and again by Lemma 9.2 (ii) and (iii) the last two term associated with the diffusion and the jump can be estimated using the k-Hölder property of σ and k'-Hölder property of ζ by

$$|a(x) - a(y)||\partial_{xx}\widehat{p}_t^y(x, y)| \leq C\mathbb{E}\big[(V_t + t)^{-(1-\frac{k}{2})} q(C(V_t + t), x - y)\big].$$

• If there is a non-trivial diffusion component, then by noticing $(V_t + t)^{-r} \leq t^{-r}$ for $r > 0$,

$$(\mathcal{L}^{y_1} - \widehat{\mathcal{L}}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x} \leq C\big[t^{-\frac{1}{2}} + t^{-(1-\frac{k'}{2})} + t^{-(1-\frac{k}{2})}\big]\mathbb{E}\big[q(C(V_t + t), x - y)\big].$$

The space integrals, appearing in (10), will be used to deal with the Gaussian terms above. While, to deal with the time integrals appearing in (9) and to show that the series (11) converges, we will apply Corollary 9.9. It is then easy to see that the convergence does not depend on the stable parameter α and the result is identical to the diffusion case.

• If there is no diffusion component, that is $\zeta = 0$, then one is forced to compute the inverse moments of V and by the self-similarity property of V, that is $t^{\frac{1}{\alpha}} V_1 = V_t$ in law, we expect that $(\mathcal{L}^{y_1} - \widehat{\mathcal{L}}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x}$ is upper bounded by

$$Ct^{-\frac{2-k}{2\alpha}}\mathbb{E}\big[V_1^{-\frac{2-k}{2}} q(\bar{a}t^{\frac{1}{\alpha}} V_1, x - y)\big] + Ct^{-\frac{1}{2\alpha}}\mathbb{E}\big[V_1^{-\frac{1}{2}} q(\bar{a}t^{\frac{1}{\alpha}} V_1, x - y)\big].$$

Roughly speaking, the space integrals, appearing in (10), will be used to deal with the terms involving $q(a(y)t^{\frac{1}{\alpha}} V_1, x - y)$ and the finiteness of the inverse moments of V_1 is treated using results from subsection 9.5. While, to deal with the time integrals appearing in (9) and to show that the series (11) converges, (done in Theorem 6.1), we will apply Corollary 9.9. This gives rise to the conditions $\frac{1}{2} < \alpha$ (drift condition) and $1 - \frac{k}{2} < \alpha$, which are conditions required for $t^{-\frac{1}{2\alpha}}$ and $t^{-\frac{2-k}{2\alpha}}$ to be $\mathbb{1}_{(0,T]}dt$ integrable. The condition $\frac{1}{2} < \alpha$ appears difficult to remove, while the condition $1 - \frac{k}{2} < \alpha$ can in fact be removed by careful using the effects of the jumps as is done in Lemma 9.15. Therefore, the assumption that $b = 0$, whenever $\alpha \leq \frac{1}{2}$ is essentially due to the fact that in the case where $b \neq 0$, to show convergence, Corollary 9.9 is only applicable if $\alpha > \frac{1}{2}$.

The drift condition for $\alpha \leq \frac{1}{2}$ may be removed by assuming further regularity on b and changing the parametrix process. We do not do that here as we are interested in minimal conditions.

In short, in order to show that the series in (11) converges, suitable estimates for each term in (9) are required. This will imply the need to estimate the basic function $\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)$ given in (6), which is obtained using inverse moments estimates of V and $V + c$ computed in Lemma 9.14 and Lemma 9.15. These estimates will lead to estimates of the integral which are in the form given in Corollary 9.9, which is used to obtain uniform convergence of the series in (11). We point out here that the difficulty is in carefully controlling the estimate of the inverse moments of V and $V + c$, as the convergence of partial sum (11) depends heavily on these estimates.

Of course, the above arguments are heuristic, but they are a good guideline as to what conditions one will expect in the computations.

As pointed out in Remark 3.10, that under Hypotheses 3.11, the diffusion component dominates the situation and one is no longer required to compute the inverse moments of V. The parameter α then plays little role in the convergence of the series and one falls back to the diffusion case. Therefore, to avoid needless repetition, in the rest of the paper, we work under Hypotheses 3.5

Remark 3.12. In Theorem 3.9 (ii) and (iii), the Lipschitz coefficient is independent of $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and the local Hölder constant is independent of $x \in \mathbb{R}^d \times \mathbb{R}^d$, and are both locally bounded in $t \in \mathbb{R}_+$. The joint continuity of $p_t(x, y)$ in $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$ is obtained in Corollary 7.12 by combining Theorem 3.9 (ii) (iii) and (iv). The fact that one obtains less regularity in y than in x is inherent to the parametrix method. In fact, looking at (10), one notes that the regularity on x depends only on $p_{t_n}^{z_n}(x, z_n)$ while the regularity on y depends on $\widehat{\theta}_{t-t_1}(z_1, y)\widehat{p}_{t-t_1}^{y}(z_1, y)$.

4 The parametrix method

In this section, we show in detail that the backward parametrix method developed in [1], which serves to obtain an expansion for a given semigroup operator around the semigroup of the parametrix, i.e. $\widehat{P}_t^z$ for a given $z \in \mathbb{R}^d$, can be applied in the current setting. This is done because, in the current work, the properties of the basic function $\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)$ do not fall perfectly into the framework developed in [1] and one can not simply refer to their results.

4.1 Some definitions and preliminary estimates

For $f \in B(\mathbb{R}^d)$, we define $P_t f(x) = \mathbb{E}(f(X_t))$ where X is the unique weak solution to (1). We defined the parametrix given by $\widehat{P}_t^z f(x) = \int f(y)\widehat{p}_t^z(x, y)dy$ and note that $(\widehat{P}_t^z)_{t\in\mathbb{R}_+}$ is a semigroup for fixed $z \in \mathbb{R}^d$. For fixed $y, z \in \mathbb{R}^d$ and $f \in \mathrm{Dom}(\widehat{P}_t^y) \cap \mathrm{Dom}(P_t)$, we obtain using the fundamental theorem for operators,

$$P_t f(y) - \widehat{P}_t^z f(y) = \int_0^t \partial_{t_1}(\widehat{P}_{t-t_1}^z P_{t_1})f(y)dt_1 = \int_0^t \widehat{P}_{t-t_1}^z(\mathcal{L} - \mathcal{L}^z)P_{t_1}f(y)dt_1, \tag{13}$$

where we let $\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y) := (\mathcal{L}^{y_1} - \mathcal{L}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x}$. Following the notation in [1], for $f \in B(\mathbb{R}^d)$, we introduce the following operators,

$$\widehat{S}_t^* f(x) := \int f(y)\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)dy$$
$$Q_t^* f(x) := \int f(y)\widehat{p}_t^y(x, y)dy,$$

whenever the integrals are well defined. Recall from (8) that

$$\begin{aligned}\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y) &= b(x)^T\nabla_x\widehat{p}_t^y(x, y)\\ &\quad + \int_{\mathbb{R}_+} \mathbb{E}\big[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\big]\mu(dc).\end{aligned}$$

Before showing that by using (13), an asymptotic expansion for $P_t f$ can be obtained, we give some preliminary estimates for $\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)$ and the operator $\widehat{S}_t^*$. Note that the function $\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)$ can be decomposed into three components which we denote by

$$\begin{aligned}D(t; x, y) &:= b(x)^T\nabla_x\widehat{p}_t^y(x, y)\\ J_{(0,1]}(t; x, y) &:= \int_{(0,1)} \mathbb{E}\big[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\big]\mu(dc)\\ J_{(1,\infty)}(t; x, y) &:= \int_{[1,\infty)} \mathbb{E}\big[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\big]\mu(dc).\end{aligned}$$

The function D is part that is associated with the drift, and $J_{(0,1]}$ and $J_{(1,\infty)}$ are the parts associated with the small jumps and large jumps respectively.

Lemma 4.1. *For* $(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times (0,T]$,

$$\widehat{p}^y_t(x,y) \leq C\mathbb{E}[q(\overline{a}V_t, x-y)]$$

$$|D(t;x,y)| \leq C\mathbb{E}\big[V_t^{-\frac{1}{2}} q(2\overline{a}V_t, x-y)\big] \tag{14}$$

$$|J_{(0,1]}(t;x,y)| \leq C\int_{(0,1]} \mathbb{E}\big[(V_t+c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_t+c), x-y)\big]\widehat{\mu}(dc) \tag{15}$$

$$|J_{(1,\infty)}(t;x,y)| \leq C\int_{(1,\infty)} \mathbb{E}\left[q(2\overline{a}(V_t+c), x-y)\right]\widehat{\mu}(dc), \tag{16}$$

where $\widehat{\mu}(dc) = \left(\mathbb{1}_{(0,1]}(c)c + \mathbb{1}_{(1,\infty)}(c)\right)\mu(dc)$.

Proof. The first inequality follows from Lemma 3.6 and Lemma 9.2 (i). The bound for $D(t;x,y)$ is obtained by applying Lemma 9.2.(ii) to (5) and using the assumption that b is bounded which gives

$$|\nabla_x \mathbb{E}\left[q(a(y)V_t, x-y)\right]| \times |b(x)| \leq C\mathbb{E}\big[(\overline{a}V_t)^{-\frac{1}{2}} q(2\overline{a}V_t, x-y)\big].$$

The bound for $J_{(0,1]}(t;x,y)$ follows from Lemma 9.4. For $J_{(1,\infty)}(t;x,y)$, we see from (6) that by using triangular inequality and Lemma 9.2 (i), for $1 < c < \infty$,

$$\begin{aligned}&\int_{(1,\infty)} \big|\mathbb{E}[q(a(y)V_t + a(x)c, x-y) - q(a(y)V_t + a(y)c, x-y)]\big|\mu(dc)\\ &\quad\leq C\int_{(1,\infty)} \mathbb{E}\left[q(2\overline{a}(V_t+c), x-y)\right]\mu(dc),\end{aligned}$$

and with this we conclude the proof. $\square$

Lemma 4.2. *For a fixed* $(y,t) \in \mathbb{R}^d \times \mathbb{R}_+$, *the function* $x \longrightarrow \widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)$ *belongs to* $C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ *and the function* $\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)$ *satisfies the following inequalities for any* $w \in (\alpha, 1 \wedge (\alpha + 1 - \frac{k}{2}))$

$$\max\left(\int |\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)|dx, \int |\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)|dy\right) \leq C(t^{-\frac{1}{2\alpha}} + t^{-\frac{\widehat{\lambda}}{\alpha}} + 1) \tag{17}$$

$$|\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)| \leq C\mathbb{E}[V_t^{-\frac{1+d}{2}} + V_t^{-(1-\frac{k}{2}+\frac{d}{2})} + V_t^{-\frac{d}{2}}] \tag{18}$$

$$\leq C(t^{-\frac{1+d}{2\alpha}} + t^{-\frac{2-k-d}{2\alpha}} + t^{-\frac{d}{2\alpha}}) \tag{19}$$

where $\widehat{\lambda} := \alpha - w + 1 - \frac{k}{2} \in (\alpha - \frac{k}{2}, 1 - \frac{k}{2})$.

Proof. Inequality (19) follows directly from Lemma 4.1 (using the fact that $V_t + c \geq V_t$), and Lemma 9.14. To show that $\widehat{\theta}_t(\cdot,y)\widehat{p}^y_t(\cdot,y) \in L^1(\mathbb{R}^d)$, we see that from (14), Fubini-Tonelli theorem and Lemma 9.14,

$$\int |D(t;x,y)|dx \leq C\mathbb{E}[V_t^{-\frac{1}{2}}] \leq Ct^{-\frac{1}{2\alpha}}.$$

While from (15) and (16), we apply the Fubini-Tonelli theorem and Lemma 9.15 with $\eta = 0$, $\lambda = 1 - \frac{k}{2}$, $\delta = 0$ and $\gamma = \frac{1}{2}$ to obtain that for any $w \in (\alpha, 1 \wedge (\alpha + \lambda))$

$$\begin{aligned}\int |J_{(0,1]}(t;x,y)|dx + \int |J_{(1,\infty)}(t;x,y)|dx &\leq \int_{\mathbb{R}_+} \mathbb{E}[(V_t+c)^{-(1-\frac{k}{2})}\mathbb{1}_{\{c\leq 1\}} + \mathbb{1}_{\{c>1\}}]\widehat{\mu}(dc)\\ &\leq Ct^{-\frac{\widehat{\lambda}}{\alpha}}\end{aligned}$$

where $\widehat{\lambda} := \alpha - w + 1 - \frac{k}{2} \in (\alpha - \frac{k}{2}, 1 - \frac{k}{2})$ with $\lambda = 1 - \frac{k}{2}$. The integral against dy is computed similarly.

To show that $J_{(0,1]}(t;\cdot,y), J_{(1,\infty)}(t;\cdot,y) \in C_0(\mathbb{R}^d)$, we see from (15) and (16) that the integrand with respect to $\mu(dc) \times \mathbb{P}$ in (8) can be bounded above by $C(cV_t^{-\frac{2-k+d}{2}} \mathbb{1}_{(0,1]}(c) + V_t^{-\frac{d}{2}} \mathbb{1}_{(1,\infty)}(c))$, which is independent of $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, and it is $\mu(dc) \times \mathbb{P}$-integrable due to the finiteness of $\widehat{\mu}$ and Lemma 9.14. The result then holds by applying dominated convergence theorem to interchange the integral in $\mu(dc) \times \mathbb{P}$ and the limit in x. Similar arguments show that $D(t;\cdot,y) \in C_0(\mathbb{R}^d)$. □

Lemma 4.3. *If $f \in L^1(\mathbb{R}^d)$, then for $t > 0$, $\widehat{S}_t^* f \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$.*

Proof. To show that $\widehat{S}_t^* f(x) := \int f(y) \widehat{\theta}_t(x,y) \widehat{p}_t^y(x,y) dy \in L^1(\mathbb{R}^d)$, we see that from Fubini-Tonelli theorem and (17),

$$\int |\widehat{S}_t^* f(x)| dx \leq \int |f(y)| \int |\widehat{\theta}_t(x,y) \widehat{p}_t^y(x,y)| dx dy \leq C \|f\|_1 (t^{-\frac{1}{2\alpha}} + t^{-\frac{\widehat{\lambda}}{\alpha}} + 1).$$

To show that $\widehat{S}_t^* f$ belongs to $C_0(\mathbb{R}^d)$, we note that from (19), the dominated convergence theorem can be applied to interchange the limit in x and integral in dy. The result then follows from Lemma 4.2. □

4.2 Finite order expansion of the Semigroup

Following the method exposed in [1], in this section, we show in detail, how an expansion of the semigroup $(P_t)_{t>0}$ of X can be obtained.

Theorem 4.4. *Let $h, g \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, then*

$$\langle P_t h, g \rangle = \langle \sum_{n=0}^{N} I_t^n(h), g \rangle + \langle R_t^N(h), g \rangle,$$

where for $n = 0$, we set $I_t^n(h) := Q_t^* h$ and for $n \geq 1$,

$$I_t^n(h)(x) = \int_0^{t_0} dt_1 \cdots \int_0^{t_{n-1}} dt_n \, Q_{t_n}^* \widehat{S}_{t_{n-1}-t_n}^* \ldots \widehat{S}_{t_0-t_1}^* h(x) = \int h(y) I_t^n(y,x) dy$$
$$R_t^N(h)(x) = \int_0^{t_0} dt_1 \cdots \int_0^{t_{N-1}} dt_N \, P_{t_N} \widehat{S}_{t_{N-1}-t_N}^* \ldots \widehat{S}_{t_0-t_1}^* h(x).$$

Proof. From Lemma 3.6 and Lemma 9.14, we see that $\widehat{p}_\epsilon^y(\cdot,y)$ is bounded and one can write

$$\begin{aligned}(P_t - \widehat{P}_t^y) \widehat{p}_\epsilon^y(\cdot,y) &= \int_0^t dt_1 P_{t_1} (\mathcal{L} - \mathcal{L}^y) \widehat{P}_{t-t_1}^y \widehat{p}_\epsilon^y(\cdot,y) \\ &= \int_0^t dt_1 \int P_{t_1}(\cdot,dz) \, \widehat{\theta}_{T-t_1+\epsilon}(z,y) \widehat{p}_{T-t_1+\epsilon}^y(z,y).\end{aligned}$$

Therefore, taking the limit as $\epsilon \to 0$, we have for all $h, g \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$,

$$\lim_{\epsilon \to 0} \int dy h(y) \int dx g(x) \, (P_t - \widehat{P}_t^y) p_\epsilon^y(x,y) \tag{20}$$
$$= \lim_{\epsilon \to 0} \int dy h(y) \int dx g(x) \int_0^t dt_1 \int P_{t_1}(x,dz) \, \widehat{\theta}_{t-t_1+\epsilon}(z,y) \widehat{p}_{t-t_1+\epsilon}^y(z,y) \tag{21}$$

To show that (21) is well defined, we have by (17), and (78) in the appendix,

$$\int dy|h(y)| \int dx|g(x)| \int P_{t_1}(x,dz)\ |\widehat{\theta}_{t-t_1+\epsilon}(z,y)p^y_{t-t_1+\epsilon}(z,y)| \tag{22}$$

$$\leq C(t-t_1+\epsilon)^{-(\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha})}\|h\|_\infty \int dx|g(x)| \int P_{t_1}(x,dz) \leq C(t-t_1)^{-(\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha})}\|g\|_1\|h\|_\infty \tag{23}$$

Recall that $\lambda = 1-\frac{k}{2}$ and $\widehat{\lambda} := \alpha - w + 1 - \frac{k}{2}$, which for $\alpha > \frac{1}{2}$ and $w \in (\alpha\vee\lambda, 1\wedge(\alpha+\lambda))$ ($\forall(\alpha,k)\in(0,1)^2$, if $b=0$) satisfies that $\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha} < 1$ and therefore (23) is integrable with respect to $\mathbb{1}_{(0,t]}dt_1$. This implies that for $\alpha > \frac{1}{2}$, the Fubini-Tonelli theorem can be used in (21) to freely exchange the order of integration, and the limit as $\epsilon\to 0$ can be introduced within the integral against dt_1 by dominated convergence theorem (the integrand is bounded by (23)). To exchange the limit as $\epsilon \to 0$ and the integral against $dy\times dx\times P_{t_1}(x,dz)$, we see that the integrand in (22) can bounded above using (19) and the fact that V is increasing. Therefore, by dominated convergence theorem and Lemma 9.8 (continuity of $\widehat{\theta}_t(z,y)\widehat{p}^y_t(z,y)$ in t), we can rewrite (21) as

$$\lim_{\epsilon\to 0}\int_0^t dt_1 \langle\ P_{t_1}\widehat{S}^*_{t-t_1+\epsilon}h, g\ \rangle = \int_0^t dt_1 \lim_{\epsilon\to 0}\langle\ P_{t_1}\widehat{S}^*_{t-t_1+\epsilon}h, g\ \rangle = \int_0^t dt_1 \langle\ P_{t_1}\widehat{S}^*_{t-t_1}h, g\ \rangle$$

To treat (20), we see that from Lemma 3.6, we have $\int \widehat{p}^y_t(x,y)dy \leq C$ and

$$\int dy|h(y)| \int dx|g(x)| \int P_t(x,dw)|\widehat{p}^y_\epsilon(w,y)| \leq C\|g\|_1\|h\|_\infty$$

This implies that Fubini-Tonelli theorem can be applied to interchange the order of integration against $dy\times dx\times P_t(x,dw)$. Then by using the fact that $\widehat{p}^z_t(x,y) \leq \mathbb{E}(V_t^{-\frac{d}{2}})$, we see that the dominated convergence theorem can be applied to treat the limit in ϵ, together with the fact that $\widehat{p}^y_t(x,y)$ is the transition density of a Feller process and $h,g\in C_0(\mathbb{R}^d)$, we have

$$\lim_{\epsilon\to 0}\int dyh(y)\int dxg(x)\int P_t(x,dw)\widehat{p}^y_\epsilon(w,y) = \int dxg(x)\int P_t(x,dw)\lim_{\epsilon\to 0}\int dyh(y)\widehat{p}^y_\epsilon(w,y) = \langle P_th, g\rangle$$

$$\lim_{\epsilon\to 0}\int dyh(y)\int dxg(x)\widehat{p}^y_{t+\epsilon}(x,y) = \int dyh(y)\int dxg(x)\widehat{p}^y_t(x,y) = \langle Q^*_th, g\rangle\,.$$

From the above arguments, we see that (20) and (21) can be rewritten into,

$$\langle P_th, g\rangle = \langle Q^*_th, g\rangle + \int_0^t \langle\ P_{t_1}\widehat{S}^*_{t-t_1}h, g\ \rangle\, dt_1.$$

From Lemma 4.3, we know that $\widehat{S}^*_{t-t_1}h \in C_0(\mathbb{R}^d)\cap L^1(\mathbb{R}^d)$ and by iteration we can derive the expansion of the semigroup in the weak sense. □

4.3 Outline of the proofs

Having derived the finite order expansion of the semigroup $(P_t)_{t>0}$ associated with the process X. For the convenience of the reader, we present in this section an overview of the strategy deployed in the rest of the paper to study the existence and regularity of the density of X.

Representation of the semigroup: As shown in the previous section, by using the method in [1], we have for any fixed $N\in\mathbb{N}$ and for any $g,h\in C_0(\mathbb{R}^d)\cap L^1(\mathbb{R}^d)$,

$$\langle\ P_th, g\ \rangle = \langle\ \sum_{n=0}^N I^n_t(h), g\ \rangle + \langle\ R^N_t(h), g\ \rangle\,.$$

Here $I^n_t(h)(x) := \int I^n_t(y,x)h(y)dy$. By showing in Theorem 5.1 that the reminder $\|R^N_t(h)\|_\infty \to 0$ as $N\to\infty$, we have

$$\langle\ P_th, g\ \rangle = \lim_{N\to\infty}\langle\ \sum_{n=0}^N I^n_t(h), g\ \rangle \tag{24}$$

We show in Theorem 5.1 that the partial sum $\sum_{n=0}^{N} I_t^n(h)(x)$ converges absolutely and uniformly, and we denote the limit by $\sum_{n=0}^{\infty} I_t^n(h)(x)$, which together with (24) shows that

$$\langle \sum_{n=1}^{\infty} I_t^n(h), g \rangle = \langle P_t h, g \rangle .$$

This implies $P_t h(x) = \sum_{n=0}^{\infty} I_t^n(h)(x)$ for almost all $x \in \mathbb{R}^d$.
Existence of the density: For $h \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, we have that for almost for all $x \in \mathbb{R}^d$

$$P_t(h)(x) = \lim_{N\to\infty} \sum_{n=0}^{N} I_t^n(h)(x) = \lim_{N\to\infty} \int \sum_{n=0}^{N} I_t^n(y,x)h(y)dy.$$

To obtain a representation of the density, we show in Theorem 6.1 that the sequence of partial sums $\sum_{n=0}^{N} I_t^n(y,x)$ converges absolutely and uniformly. This implies

$$|\sum_{n=0}^{N} I_t^n(y,x)h(y)| < C|h(y)| \sum_{n=0}^{\infty} |I_t^n(y,x)| \le C_t|h(y)|,$$

where the constant C_t is independent of (x,y). Therefore by dominated convergence theorem,

$$P_t(h)(x) = \int h(y) \lim_{N\to\infty} \sum_{n=0}^{N} I_t^n(y,x)\, dy,$$

for all $h \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, and from which we identify the form of the density to be $p_t(x,y) = \lim_{N\to\infty} \sum_{n=0}^{N} I_t^n(y,x)$.
Regularity of the density: Having shown that the density of X exists and has the representation

$$p_t(x,y) = \lim_{N\to\infty} \sum_{n=0}^{N} I_t^n(y,x),$$

we study the regularity of the density $p_t(x,y)$ using the above representation. To show that the density $p_t(x,y)$ is Lipschitz in x, we show that, for each n, $I_t^n(y,x)$ is Lipschitz in x with Lipschitz constant L_t^n independent of y and that the sequence of partial sums of the Lipschitz constants is convergent and finite, i.e.

$$|p_t(x,y) - p_t(x^*,y)| \le |x - x^*| \lim_{N\to\infty} \sum_{n=0}^{N} L_t^n.$$

The assumptions in Theorem 7.2 allows one to apply Corollary 9.9 to show that $\lim_{N\to\infty} \sum_{n=0}^{N} L_t^n$ is finite. The method of proof for Hölder continuity in y and (Hölder)-continuity in t are similar. By combining continuity in each variable, we obtain the joint continuity of the density in Corollary 7.12.

The density $p_t(x,y)$ is Lipschitz continuous in x and to examine differentiability with respect to x, we notice that by Rademacher theorem, it is sufficient to compute the gradient of the density. In Theorem 7.2, one first identifies $\nabla_x I_t^n(x,y)$ and then shows that the partial sums $\sum_{n=0}^{N} \nabla_x I_t^n(y,x)$ converge absolutely and uniformly, which implies

$$\nabla_x p_t(x,y) = \nabla_x \lim_{N\to\infty} \sum_{n=0}^{N} I_t^n(y,x) = \lim_{N\to\infty} \sum_{n=0}^{N} \nabla_x I_t^n(y,x)$$

for $(y,t) \in \mathbb{R}^d \times \mathbb{R}_+$.

5 Study of the Remainder

To obtain the asymptotic expansion of $P_t(h)$, for $h \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ we show that the reminder $|R_t^N(h)(x)|$ goes to zero as $N \to \infty$. To estimate the reminder in Theorem 4.4 and study the convergence of the expansion, we need to estimate the terms $|\widehat{\theta}_t(x,y)|\widehat{p}_t^y(x,y)$ and $\widehat{p}_t^y(x,y)$, using Lemma 4.1 and Lemma 4.2.

Roughly speaking, the validity of the asymptotic expansion (24) depends on the inverse moments of V_t which are estimated in subsection 9.5.

Theorem 5.1. *The sequence of partial sums $\sum_{n=0}^N I_t^n(h)$ converges uniformly as $N \to \infty$. Furthermore $P_t(h) = \sum_{n=0}^\infty I_t^n(h)$.*

Proof. One divides the proof in two cases: (i) If $b \neq 0$, $k \in (0,1)$ and $\alpha > \frac{1}{2}$, then for any $g, h \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, we have that for all $n, N \in \mathbb{N}$, $I_t^n(h)$ and $R_t^N(h)$ are well defined and (24) is valid.
(ii) if $b = 0$ (the drift is zero), then for any $(\alpha, k) \in (0,1)^2$ and $g, h \in C_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, for all $n, N \in \mathbb{N}$, $I_t^n(h)$ and $R_t^N(h)$ are well defined and (24) is valid.

To do this we need to find upper bounds for the operators used in the definitions of I^n and R^N. In fact, one easily obtains that $|Q^* g(x)| \leq \|g\|_\infty$, $|P_t h(x)| \leq \|h\|_\infty$ and from (17) we have that for any $w \in (\alpha, 1)$,

$$|\widehat{S}_t^* h(x)| \leq C\|h\|_\infty \big(t^{-\frac{1}{2\alpha}} + t^{-\frac{\widehat{\lambda}}{\alpha}}\big), \tag{25}$$

where $\widehat{\lambda} := \alpha - w + 1 - \frac{k}{2}$ and $\lambda = 1 - \frac{k}{2}$. From here, one obtains the estimates:

$$\begin{aligned} |I_t^n(h)(x)| &\leq C^n \|h\|_\infty \int_0^{t_0} \cdots \int_0^{t_{n-1}} dt_n \ldots dt_1 \prod_{i=0}^{n-1} \big((t_i - t_{i+1})^{-\frac{1}{2\alpha}} + (t_i - t_{i+1})^{-\frac{\widehat{\lambda}}{\alpha}}\big) \\ &\leq C^n \|h\|_\infty \int_0^{t_0} \cdots \int_0^{t_{n-1}} dt_n \ldots dt_1 \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})}, \end{aligned} \tag{26}$$

where the last inequality above follows from (78). In order to apply Corollary 9.9 with $\gamma_1 = \gamma_2 = \frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha}$, we see that for any $w \in (\alpha \vee \lambda, 1 \wedge (\alpha + \lambda))$ we have that $\widehat{\lambda} \in (0, \alpha)$ and therefore $\gamma_1 = \gamma_2 \in (0,1)$. From here, we obtain that there exists a constant C such that $|I_t^n(h)(x)| \leq \frac{C^n t^{(1-\gamma_1)n}}{\Gamma(1+n(1-\gamma_1))}$.

The reminder is estimated similarly to give the existence of a constant C such that $|R_t^N(h)(x)| \leq \frac{C^N}{\Gamma(1+N(1-\gamma_1))} \to 0$. Furthermore, note that in general, given $y \geq 0$, the series

$$\sum_n \frac{y^{(1-\gamma_1)n}}{\Gamma(1 + \rho + n(1-\gamma_1))}$$

converges uniformly for y in compact sets. In fact, without loss of generality one can consider $\rho = 0$ as the Gamma function is increasing for values of n large enough. Then the above sum converges if one uses the Stirling approximation for the Gamma function.

Therefore the arguments at the beginning of subsection 4.3 can be carried out and the result follows. Clearly if $b = 0$ then the first term in (26) disappears and we obtain (ii). $\square$

6 Existence of Density

In this subsection, we examine the existence of the density of the process X. The main goal is to show that the partial sum $\sum_{n=0}^N I_t^n(y, x)$ converges uniformly.

For ease of computation and compactness in the formulas, we enlarge the domain of integration and let δ_0 be the Dirac measure at zero then the result of Lemma 4.1 can be rewritten as

$$
\begin{aligned}
D(t;x,y) &\leq C\int_{(0,1]\times\mathbb{R}} \mathbb{E}\big[V_t^{-\frac{1}{2}} q(2\overline{a}(V_t+\bar{c}),x-y)\big]\,\widehat{\mu}(dc)\times\delta_0(d\bar{c})\\
J_{(0,1)}(t;x,y) &\leq C\int_{(0,1]\times\mathbb{R}} \mathbb{E}\big[(V_t+c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_t+c),x-y)\big]\,\widehat{\mu}(dc)\times\delta_0(d\bar{c})\\
J_{[1,\infty)}(t;x,y) &\leq C\int_{(1,\infty)\times\mathbb{R}} \mathbb{E}\big[q(2\overline{a}(V_t+c),x-y)\big]\,\widehat{\mu}(dc)\times\delta_0(d\bar{c}).
\end{aligned}
$$

In order to combine the three above estimates for the drift and jumps together to obtain a single Gaussian upper bound, we introduce an auxiliary $\frac{1}{2}$-Bernoulli random variable η, which is independent of V. For every c, $\bar{c}$, and for $\lambda,\delta,\gamma\in\mathbb{R}_+$, we set

$$
\mathbb{W}^{\lambda,\delta,\gamma}_{\eta,V_t}(c) := (1-\eta)\Big[\mathbb{1}_{(0,1]}(c)(c+V_t)^{-\lambda}+\mathbb{1}_{(1,\infty)}(c)(c+V_t)^{-\delta}\Big]+\eta V_t{}^{-\gamma} \tag{27}
$$

$$
\mathbb{V}_{\eta,V_t}(c,\bar{c}) := \overline{a}(V_t+\eta\bar{c}+(1-\eta)c) \tag{28}
$$

which are both greater than zero. Then from Lemma 4.1 applied to (7) we obtain

$$
|\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)| \leq C\int_{\mathbb{R}_+\times\mathbb{R}} \mathbb{E}\left[q(2\mathbb{V}_{\eta,V_t}(c,\bar{c}),x-y)\mathbb{W}^{\lambda,\delta,\gamma}_{\eta,V_t}(c)\right]\,\widehat{\mu}(dc)\times\delta_0(d\bar{c}).
$$

Before proceeding with the proofs and results, we introduce some notations and simplifications that will be used through out the rest of this article. Using the fact that V_{t-s} has the same distribution as V_t-V_s, we replace $V_{t_i-t_{i+1}}$ by $\Delta_{t_{i+1}}V := V_{t_i}-V_{t_{i+1}}$ in the right hand side of (8) and for $i=0,...,n-1$, $z_0=y$ and $z_{n+1}=x$. Given a sequence of independent $\frac{1}{2}$-Bernoulli random variable $(\eta_i)_{i=0,\ldots,n-1}$, we can write for $\lambda=1-\frac{k}{2}$, $\delta=0$ and $\gamma=\frac{1}{2}$,

$$
\begin{aligned}
&|\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)p^{z_i}_{t_i-t_{i+1}}(z_{i+1},z_i)|\\
&\quad\leq C\int_{\mathbb{R}_+\times\mathbb{R}} \mathbb{E}\big[q(2\mathbb{V}_{\eta_i,\Delta_{t_{i+1}}V}(c_i,\bar{c}_i),z_i-z_{i+1})\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\,\widehat{\mu}(dc_i)\times\delta_0(d\bar{c}_i).
\end{aligned} \tag{29}
$$

Note that from the independent increment property of V, and the independence of the sequence $(\eta_i)_{i=0,\ldots,n-1}$ we are free to exchange the order of expectation, integrals and products when we consider (10).

Theorem 6.1. *The partial sum $\sum_{i=0}^N I^n_t(y,x)$ converges absolutely and uniformly and the density $p_t(x,y)$ is given by*

$$
\begin{aligned}
p_t(x,y) &= \sum_{i=0}^{\infty} I^n_t(y,x)\\
I^n_t(y,x) &= \int_0^t\cdots\int_0^{t_{n-1}} dt_n\ldots dt_1\, K(t_n,\ldots,t_1,t;x,y)
\end{aligned} \tag{30}
$$

where $K(t_n,\ldots,t_1,t;x,y)$ is defined in (10).

Proof. The proof shares various estimates with the proof of Theorem 5.1. To show that the series of partial sums converge absolutely and uniformly, one needs to estimate $|I^n_t(y,x)|$ and this is done in two steps.

Step 1. One first estimates the space integral $|K(t_n,\ldots,t_1,t;x,y)|$ defined in (10). Since the integrand in (29) is positive and from the independent increment property of V, and independence of $(\eta_i)_{i=1,\ldots n}$, we are free to exchange the order of expectation, integrals and products after estimating

(10) using (29). From the convolution property of Gaussian densities, i.e. integrating with respect to the variables $z_i, i = 1, ..., n$, we obtain for $\lambda = 1 - \frac{k}{2}$, $\delta = 0$ and $\gamma = \frac{1}{2}$,

$$
\begin{aligned}
&|K(t_n, \ldots, t_1, t; x, y)| \\
&\leq C^n \mathbb{E}\Big[\int_{(\mathbb{R}_+\times\mathbb{R})^n} q\Big(2\bar{a}V_{t_n} + 2\sum_{i=0}^{n-1} \mathbb{V}_{\eta_i, \Delta_{t_{i+1}}V}(c_i, \bar{c}_i), z_0 - z_{n+1}\Big) \prod_{i=0}^{n-1} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\, \widehat{\mu}(dc_i) \times \delta_0(d\bar{c}_i)\Big] \\
&\leq C^n \mathbb{E}\Big[\int_{(\mathbb{R}_+\times\mathbb{R})^n} \Big(\sum_{i=0}^{n-1}(\eta_i c_i + (1-\eta_i)\bar{c}_i) + \sum_{i=0}^{n} \Delta_{t_{i+1}}V\Big)^{-\frac{d}{2}} \prod_{i=0}^{n-1} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\, \widehat{\mu}(dc_i) \times \delta_0(d\bar{c}_i)\Big] \\
&\leq C^n \mathbb{E}\Big[\Big(\sum_{i=0}^{n} \Delta_{t_{i+1}}V\Big)^{-\frac{d}{2}} \prod_{i=0}^{n-1} \int_{\mathbb{R}_+} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\, \widehat{\mu}(dc_i)\Big]
\end{aligned}
$$

where $V_{t_{n+1}} := 0$. Using Lemma 9.12 and Hypotheses 3.2, the above expectation is given by

$$
\begin{aligned}
&= C^n \Gamma\left(\frac{d}{2}\right)^{-1} \int_{\mathbb{R}_+} s_n^{\frac{d}{2}-1} \mathbb{E}\left[e^{-s_n V_{t_n}}\right] \Big[\prod_{i=0}^{n-1} \int_{\mathbb{R}_+} \mathbb{E}\big[e^{-s_n \Delta_{t_{i+1}}V}\ \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\widehat{\mu}(dc_i)\Big]\, ds_n \\
&= C^n \Gamma\left(\frac{d}{2}\right)^{-1} \int_{\mathbb{R}_+} s_n^{\frac{d}{2}-1} e^{-s_n^\alpha m(s_n) t_n A} \prod_{i=0}^{n-1} \Big[\int_{\mathbb{R}_+} \mathbb{E}\big[e^{-s_n \Delta_{t_{i+1}}V}\ \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\widehat{\mu}(dc_i)\Big]\, ds_n. \qquad (31)
\end{aligned}
$$

To compute the expectations in the product, we note that from (27) the terms can be computed using Lemma 9.15 with $\lambda = 1 - \frac{k}{2}$, $\delta = 0$ and $\gamma = \frac{1}{2}$ which gives that

$$
\int_{\mathbb{R}_+} \mathbb{E}\big[e^{-s_n\Delta_{t_{i+1}}V} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\widehat{\mu}(dc_i) \leq C e^{-M^{\alpha-1} s_n^\alpha m(2s_n) A (t_i - t_{i+1})} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})}.
$$

Since m is increasing, then $m(s) \leq m(2s)$, which gives us

$$
\begin{aligned}
|K(t_n, \ldots, t_1, t; x, y)| &\leq C^n \Big(\int_{\mathbb{R}_+} s_n^{\frac{d}{2}-1} e^{-M^{\alpha-1} s_n^\alpha m(s_n) A \sum_{i=0}^{n-1}(t_i - t_{i+1})} ds_n\Big) \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})} \\
&\leq C^n t^{-\frac{d}{2\alpha}} \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})} \qquad (32)
\end{aligned}
$$

where the last inequality follows from Lemma 9.13.

Step 2. After having computed the estimate for the space integral in (31), we now compute the time integral. That is,

$$
\int_0^{t_0} \cdots \int_0^{t_{n-1}} dt_n \ldots dt_1 |K(t_n, \ldots, t_1, t; x, y)| \leq C^n t^{-\frac{d}{2\alpha}} \int_0^{t_0} \cdots \int_0^{t_{n-1}} dt_n \ldots dt_1 \prod_{i=0}^{n-1} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})}
$$

where $\widehat{\lambda} = \alpha - w + 1 - \frac{k}{2}$ and $\lambda = 1 - \frac{k}{2}$. Proceeding as in the proof of Theorem 5.1 we obtain the result. $\square$

7 Regularity of the Density

The goal of this section is to prove the regularity of the density with respect to the variables $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$. In particular, in order to prove the joint continuity we will prove equicontinuity of any one of the three variables with respect to the other two. The proofs of Theorem 7.2, Theorem 7.3 and Theorem 7.6 are similar to the proof of Theorem 6.1, therefore not all the details of the proof are given and we shall refer to the corresponding parts of the proof of Theorem 6.1.

7.1 Differentiability with respect to the initial point

Lemma 7.1. *For $t>0$, the density $\widehat{p}_t^y(\cdot,y)$ is Lipschitz continuous, more explicitly*

$$|\widehat{p}_t^y(x^*,y)-\widehat{p}_t^y(x,y)|\le C|x^*-x|\mathbb{E}\big[V_t^{-\frac{1}{2}}\left(q(2\overline{a}V_t,x^*-y)+q(2\overline{a}V_t,x-y)\right)\big].$$

Proof. Using Lemma 3.6, the algebraic identity $a^2-b^2=(a-b)(a+b)$ and Lemma 9.2. (iv), we have

$$\begin{aligned}\widehat{p}_t^y(x^*,y)-\widehat{p}_t^y(x,y)&=C\mathbb{E}\Big[(V_t^d\det(a(y)))^{1/2}\\&\times\left(q\left(2V_ta(y),x^*-y\right)-q\left(2V_ta(y),x-y\right)\right)\left(q\left(2V_ta(y),x^*-y\right)+q\left(2V_ta(y),x-y\right)\right)\Big].\end{aligned}$$

Next, applying Lemma 9.3 and Hypotheses 3.5. (iii),

$$|\widehat{p}_t^y(x^*,y)-\widehat{p}_t^y(x,y)|\le C|x^*-x|\mathbb{E}\big[V_t^{-\frac{1}{2}}\left(q(2\overline{a}V_t,x^*-y)+q(2\overline{a}V_t,x-y)\right)\big].$$

□

Theorem 7.2. *For $\alpha>\frac{1}{2}$ and $t>0$, the density $p_t(x,y)$ satisfies the following properties:*
(i) It is globally Lipschitz in x, where the Lipschitz constant is independent of (x,y) and locally bounded in t.
(ii) It is differentiable in x and

$$\partial_{x_i}p_t(x,y)=\lim_{N\to\infty}\sum_{n=0}^{N}\partial_{x_i}I_t^n(y,x)$$

where for a fixed $n\in\mathbb{N}$ and $i=1,\dots,d$

$$\partial_{x_i}I_t^n(y,x)=\int_0^t\cdots\int_0^{t_{n-1}}dt_n\dots dt_1\int\partial_{x_i}\widehat{p}_{t_n}^{z_n}(x,z_n)\prod_{i=0}^{n-1}\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)\widehat{p}_{t_i-t_{i+1}}^{z_i}(z_{i+1},z_i)\,dz_{i+1}.$$

Proof. (i) For notational simplicity, we set

$$K'(t_n,...,t_1,t;x,y):=\int\mathbb{E}\big[V_{t_n}^{-\frac{1}{2}}q(4\overline{a}V_{t_n},x-z_n)\big]\prod_{i=0}^{n-1}\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)\widehat{p}_{t_i-t_{i+1}}^{z_i}(z_{i+1},z_i)dz_{i+1}.$$

From Lemma 7.1 and (9),

$$|I_t^n(y,x)-I_t^n(y,x^*)|\le C|x^*-x|\int_0^t\cdots\int_0^{t_{n-1}}dt_n\dots dt_1\left(K'(t_n,...,t_1,t;x,y)+K'(t_n,...,t_1,t;x^*,y)\right).$$

We consider first the space integral. The proof is along similar lines as in the proof of Theorem 6.1, except that in the present case we have an extra $V_{t_n}^{-\frac{1}{2}}$ term. Following the same computations as in Step 1. of Theorem 6.1, recalling the notation $\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)$ in (27) for $\lambda=1-\frac{k}{2}$, $\delta=0$ and $\lambda=\frac{1}{2}$,

$$\begin{aligned}&|K'(t_n,...,t_1,t;x,y)|\\&\le C^n\mathbb{E}\Big[V_{t_n}^{-\frac{1}{2}}\int_{(\mathbb{R}_+\times\mathbb{R})^n}\big(\sum_{i=0}^{n}\Delta_{t_{i+1}}V\big)^{-\frac{d}{2}}\prod_{i=0}^{n-1}\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\;\widehat{\mu}(dc_i)\times\delta_0(d\bar{c}_i)\Big]\\&\le C^n\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}s^{\frac{1}{2}-1}s_n^{\frac{d}{2}-1}\mathbb{E}\big[e^{-(s+s_n)V_{t_n}}\big]\Big[\prod_{i=0}^{n-1}\int_{\mathbb{R}_+}\mathbb{E}\big[e^{-s_n\Delta_{t_{i+1}}V}\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\widehat{\mu}(dc_i)\Big]ds_nds.\end{aligned}\tag{33}$$

Recall that under Hypotheses 3.2 (ii) , there exists a positive increasing concave function m and constant positive constant A such that

$$\mathbb{E}[e^{-sV_t}] \le Ce^{-s^\alpha m(s)tA}.$$

Next, since $0<\alpha<1$ and m is a positive increasing concave function, one can apply the Jensen's inequality, to obtain $(x+y)^\alpha m(x+y) \ge z(x^\alpha m(2x)+y^\alpha m(2y))$ for any $a,b>0$ and $0<z\le 2^{\alpha-1}$. Thus from Hypotheses 3.2 (ii),

$$\mathbb{E}\big[e^{-(s+s_n)V_{t_n}}\big] \le Ce^{-(s+s_n)^\alpha m(s+s_n)t_nA} \le Ce^{-z(s_n^\alpha m(2s_n)+s^\alpha m(2s))t_nA}.$$

Then one can further bound (33) as follows,

$$\begin{aligned}|K'(t_n,...,t_1,t;x,y)| \le C\Big(&\int_{\mathbb{R}_+} s^{\frac{1}{2}-1}e^{-zs^\alpha m(2s)t_nA}ds\Big)\\ &\times \int_{\mathbb{R}_+} s_n^{\frac{d}{2}-1}e^{-zs_n^\alpha m(2s_n)t_nA}\Big[\prod_{i=0}^{n-1}\int_{\mathbb{R}_+}\mathbb{E}\big[e^{-s_n\Delta_{t_{i+1}}V}\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i,\Delta_{t_{i+1}}V}(c_i)\big]\,\widehat{\mu}(dc_i)\Big]\,ds_n.\end{aligned}$$

The first term above is computed using Lemma 9.13, while the second term is computed the same way as in (31) and (26). This shows that for any $w\in(\alpha,1\wedge(\alpha+\lambda))$

$$|I_t^n(y,x)-I_t^n(y,x^*)| \le C^n|x^*-x|\int_0^t\dots\int_0^{t_{n-1}}dt_n\dots dt_1\, t^{-\frac{d}{2\alpha}}t_n^{-\frac{1}{2\alpha}}\prod_{i=0}^{n-1}(t_i-t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha})} \tag{34}$$

where $\widehat{\lambda}=\alpha-w+1-\frac{k}{2}$ with $\lambda=1-\frac{k}{2}$. The above multiple time integral is finite for $\alpha>\frac{1}{2}$ and $w\in(\alpha\vee\lambda,1\wedge(\alpha+\lambda))$ (see the proof of Theorem 6.1).

From (34) and Corollary 9.9, the partial sum involving $|I_t^n(y,x)-I_t^n(y,x^*)|$ converges absolutely and uniformly in $(x,y,t)\in\mathbb{R}^d\times\mathbb{R}^d\times(0,T]$ to a constant locally bounded in $t>0$. This shows that for a fixed t, the density $p_t(x,y)$ is globally Lipschitz in x.

(ii) By Rademacher's theorem, the density $p_t(x,y)$ is differentiable in x almost everywhere and the derivative is uniquely determined by the gradient $\nabla_x p_t(x,y)$. In the following, we give an explicit representation of the gradient $\nabla_x p_t(x,y)$. For any $f:\mathbb{R}^d\to\mathbb{R}$, we write $\Delta_{i,h}f(x)=f(x+he_i)-f(x)$, and by setting $x^*=x+e_ih$ for $h\neq 0$ in the proof of (i) (in particular from (34)), we observe that

$$\Big|\frac{1}{h}\Delta_{i,h}K(t_n,\dots,t_1,t;x,y)\Big| \le Ct^{-\frac{d}{2\alpha}}t_n^{-\frac{1}{2\alpha}}\prod_{i=0}^{n-1}(t_i-t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha})},$$

and we deduce for $\lambda=1-\frac{k}{2}$, $\alpha>\frac{1}{2}$ and $w\in(\alpha\vee\lambda,1\wedge(\alpha+\lambda))$ that the dominated convergence theorem can be applied to interchange the time integrals and the limit in h to obtain

$$\partial_{x_i}I_t^n(y,x)=\int_0^t\dots\int_0^{t_{n-1}}dt_n\dots dt_1\,\lim_{h\to0}\frac{1}{h}\Delta_{i,h}K(t_n,\dots,t_1,t;x,y).$$

In the following, we show that

$$\lim_{h\to0}\frac{1}{h}\Delta_{i,h}K(t_n,\dots,t_1,t;x,y)=\int\lim_{h\to0}\frac{1}{h}\Delta_{i,h}\widehat{p}^{z_n}_{t_n}(x,z_n)\prod_{i=0}^{n-1}\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)\widehat{p}^{z_i}_{t_i-t_{i+1}}(z_{i+1},z_i)\,dz_{i+1}.$$

From the proof of Theorem 6.1, for every fixed n and $t_n,\dots,t_0$, the measure

$$\widehat{\nu}_{t_0,\dots,t_n}(z_n)dz_n := \Big(\int\prod_{i=0}^{n-1}\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1},z_i)\widehat{p}^{z_i}_{t_i-t_{i+1}}(z_{i+1},z_i)dz_{n-1}\dots dz_1\Big)\,dz_n$$

is a finite signed measure on $\mathbb{R}^d$. Therefore, to exchange the limit in h and the space integral, it is sufficient to show that for all $h \neq 0$, there exists $p > 1$ such that $h^{-1}\Delta_{i,h} p^{z_n}_{t_n}(x, z_n)$ belongs to $L^p(|\widehat{\nu}_{t_0,\dots,t_n}(z_n)|dz_n)$. For $r > 0$, using the Jensen's inequality and Lemma 7.1, we can write

$$\begin{aligned}
&\int |h^{-1}\Delta_{i,h}\widehat{p}^{z_n}_{t_n}(x, z_n)|^{1+r}|\widehat{\nu}_{t_0,\dots,t_n}(z_n)|dz_n \\
&\leq C \int \mathbb{E}\big[V_{t_n}^{-\frac{1}{2}}\left(q(4\overline{a}V_{t_n}, x + he_i - z_n) + q(4\overline{a}V_{t_n}, x - z_n)\right)\big]^{1+r}|\widehat{\nu}_{t_0,\dots,t_n}(z_n)|dz_n \\
&\leq C \int \mathbb{E}\big[V_{t_n}^{-\frac{1+(d+1)r}{2}}\left(q(4(1 + r)^{-1}\overline{a}V_{t_n}, x + he_i - z_n) + q(4(1 + r)^{-1}\overline{a}V_{t_n}, x - z_n)\right)\big]|\widehat{\nu}_{t_0,\dots,t_n}(z_n)|dz_n
\end{aligned}$$

where in the last inequality, we use Lemma 9.2 (iv). By repeating the arguments given in the proof of (i), we see that for $x \in \mathbb{R}^d$ and $h \neq 0$,

$$\int \mathbb{E}\big[V_{t_n}^{-(\frac{1+(d+1)r}{2})}q((1+r)^{-1}\overline{a}V_{t_n}, x-z_n)\big]\,\widehat{\nu}_{t_0,\dots,t_n}(dz_n) \leq C^N t_n^{-\frac{1+(d+1)r}{2\alpha}} t^{-\frac{d}{2\alpha}} \prod_{i=0}^{n-1}(t_i-t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha}\vee\frac{1}{2\alpha})} < \infty.$$

That is the function $h^{-1}\Delta_{i,h}p^{z_n}_{t_n}(x, z_n)$ belongs to $L^{1+r}(|\widehat{\nu}_{t_0,\dots,t_n}(z_n)|dz_n)$ for $\tau > 0$ small enough.

By (34) and Corollary 9.9, one can conclude that the sequence of partial sums $\sum_{n=0}^{N} \partial_{x_i} I^n_t(y, x)$ converges absolutely and uniformly, which implies

$$\lim_{N\to\infty} \sum_{n=0}^{N} \partial_{x_i} I^n_t(y, x) = \partial_{x_i} \lim_{N\to\infty} \sum_{n=0}^{N} I^n_t(y, x) = \partial_{x_i} p_t(x, y)$$

for $(t, y) \in (0, T] \times \mathbb{R}^d$. □

7.2 Hölder continuity with respect to the terminal point

To prove that the density $p_t(x, y)$ is locally β-Hölder continuous in y for $\beta \in (0, k)$, we need to prove that the basic function $\widehat{\theta}_t(x, y)\widehat{p}^y_t(x, y)$ which appears in $I^n_t(y, x)$ (see (10)) is locally β-Hölder continuous y. The proof of this fact is somewhat long and the computations are delicate due to the expression (8). Therefore we postpone the proof to Lemma 9.6 in the Appendix. After one realizes this, we can use the ideas in Theorem 6.1. Therefore we assume the result in Lemma 9.6 and put our current results in a similar framework as in (27) and (29). Recall that η is a $\frac{1}{2}$-Bernoulli random variable independent of V and λ_i, γ_i and δ_i for $i = 1, 2$ are given in Lemma 9.6. That is, $\lambda_1 := 1 - \frac{k}{2} + \frac{\beta}{2} > 0$, $\lambda_2 := 1 - k + \frac{\beta}{2} > 0$ and $\delta_1 := \frac{\beta}{2}$, $\delta_2 := 0$, $(\gamma_1, \gamma_2) := (\frac{\beta}{2}, \frac{1+\beta}{2})$. Recall also the notation defined (27) and (28), we write

$$\begin{aligned}
\sum_{i=1}^{2} \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta,V_t}(c) &= \sum_{i=1}^{2} \left((1 - \eta)\left(\mathbb{1}_{(0,1]}(c)(V_t + c)^{-\lambda_i} + \mathbb{1}_{(1,\infty)}(c)(V_t + c)^{-\delta_i}\right) + \eta V_t^{-\gamma_i}\right) \qquad (35)\\
\mathbb{V}_{\eta,V_t}(c, \bar{c}) &= \overline{a}(V_t + \eta\bar{c} + (1 - \eta)c),
\end{aligned}$$

and we recall $G_{V_t}(x)$ defined in Lemma 9.6,

$$G_{V_t}(x) := \int_{\mathbb{R}_+\times\mathbb{R}} \mathbb{E}\Big[q(8\mathbb{V}_{\eta,V_t}(c, \bar{c}), x) \sum_{i=1}^{2} \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta,V_t}(c)\Big]\,\widehat{\mu}(dc) \times \delta_0(d\bar{c}).$$

Theorem 7.3. *For $\alpha > \frac{1}{2}$, the density $p_t(x, y)$ is locally Hölder continuous in y, with Hölder exponent $\beta \in (0, (2\alpha - 1) \wedge k)$, and the Hölder constant is independent of x and locally bounded in $(t, y) \in (0, T] \times \mathbb{R}^d$.*

Proof. From Theorem 6.1 and triangular inequality, we have

$$|p_t(x,y) - p_t(x,y^*)| \le \sum_{n=0}^{\infty} |I_t^n(y,x) - I_t^n(y^*,x)|.$$

Therefore, it is sufficient to show that $I_t^n(y,x)$ is Hölder continuous in y and the sum of its the Hölder constants converges uniformly. Similarly to the proof of Theorem 6.1, we separate the estimations into computations for the time and the space integrals. We proceed by setting (note that $t_0 = t$)

$$\begin{aligned} &\overline{K}(t_n, \dots, t_1, t; x, y) \\ &:= \int \widehat{p}_{t_n}^{z_n}(x, z_n) G_{\Delta_{t_1} V}(z_1 - y) \prod_{i=1}^{n-1} \widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i) \widehat{p}_{t_i - t_{i+1}}^{z_i}(z_{i+1}, z_i) dz_{i+1} dz_1 \end{aligned} \tag{36}$$

and by Lemma 9.6, we have for $\beta < k$

$$|I_t^n(y,x) - I_t^n(y^*,x)| \le C|y - y^*|^\beta \int_0^t \dots \int_0^{t_{n-1}} dt_n \dots dt_1 \; (\overline{K}(t_n, \dots, t_1, t; x, y) + \overline{K}(t_n, \dots, t_1, t; x, y^*)).$$

Therefore, it is sufficient to show that

$$\sum_{n=1}^{\infty} \int_0^t \dots \int_0^{t_{n-1}} dt_n \dots dt_1 \; (\overline{K}(t_n, \dots, t_1, t; x, y) + \overline{K}(t_n, \dots, t_1, t; x, y^*)) < \infty.$$

Furthermore, the estimate for the above integrals are identical to Theorem 6.1, except we replace $\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_0, \Delta_{t_1} V}(c_0)$ in Theorem 6.1 by $\sum_{i=1}^2 \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta_0, \Delta_{t_1} V}(c_0)$ as defined in (35).

$$\overline{K}(t_n, \dots, t_1, t; x, y) \le C^n \; t_0^{-\frac{d}{2\alpha}} (t_0 - t_1)^{-\gamma^*} \prod_{i=1}^{n-1} (t_i - t_{i+1})^{-(\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha})}, \tag{37}$$

where for $i = 1, 2$ and $w_i \in (\alpha \vee \lambda_i, 1 \wedge (\alpha + \lambda_i))$,

$$\begin{aligned} (\widehat{\lambda}_1, \dots, \widehat{\lambda}_6) &:= \left(\frac{\alpha + \lambda_1 - \omega_1}{\alpha}, \frac{\alpha + \lambda_2 - \omega_2}{\alpha}, \frac{\delta_1}{\alpha}, \frac{\delta_2}{\alpha}, \frac{\gamma_1}{\alpha}, \frac{\gamma_2}{\alpha} \right) \\ \gamma^* &:= \max(\widehat{\lambda}_1, \dots, \widehat{\lambda}_6) = \widehat{\lambda}_1 \vee \widehat{\lambda}_2 \vee \widehat{\lambda}_6. \end{aligned}$$

In order to obtain convergence through Corollary 9.9, we need the condition that $\gamma^* < 1$ as well as $\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha} \in (0,1)$. Using the fact that $\beta < k$, the parameters w_1 and w_2 can always be chosen such that $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ is strictly smaller than one and the condition that $\widehat{\lambda}_3 \vee \widehat{\lambda}_6 < 1$ is equivalent to $\beta < 2\alpha - 1$. ☐

Note that the parameters (γ_1, γ_2) appearing in $\widehat{\lambda}_5$ and $\widehat{\lambda}_6$, as well as $\frac{1}{2\alpha}$ appearing in $\frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha}$ are related to the fact that $b \neq 0$. Therefore these parameters do not appear in the particular case that $b = 0$ and there are less restrictions to be satisfied. This leads to the following result.

Corollary 7.4. *In the case, where $b = 0$ then the density $p_t(x,y)$ is locally β-Hölder continuous in y with $\beta \in (0\,,\, 2\alpha \wedge k)$.*

Proof. The proof is similar to that of Theorem 7.3. From Lemma 9.6, we required that $\beta < k$.

The difference here is that since $b = 0$, one need not to consider $\widehat{\lambda}_5$ and $\widehat{\lambda}_6$. Therefore γ^* given in (37) is $\max(\widehat{\lambda}_1, \dots, \widehat{\lambda}_4) = \widehat{\lambda}_1 \vee \widehat{\lambda}_2 \vee \widehat{\lambda}_3$. The condition $\widehat{\lambda}_1 < 1$ and $\widehat{\lambda}_2 < 1$ are always satisfied as discussed in the proof of Theorem 7.3. Therefore, in this case, we require $\beta < 2\alpha$ (i.e. $\widehat{\lambda}_3 < 1$) and $\beta < k$. ☐

Remark 7.5. In the case that there is a non-trivial diffusion component, the parameter α plays little role, and Theorem 7.3 and Corollary 7.4 can be combined and restated as the following. For $\alpha \in (0,1)$ and $b \neq 0$, the density $p_t(x,y)$ is locally Hölder continuous in y, with Hölder exponent $\beta \in (0, k \wedge k')$, and the Hölder constant is independent of x and locally bounded in $(t,y) \in \mathbb{R}_+ \times \mathbb{R}^d$.

7.3 (Hölder) Continuity in time

To prove that the density $p_t(x,y)$ is locally Hölder continuous in t, we need to show that the basic function $\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y)$ which appears in $I_t^n(y,x)$ is Hölder continuous in t. For the ease of presentation we shall again postpone the proof to Lemma 9.8 in the Appendix.

In the following, our results are stated without the diffusion component. In the case where there is a non-trivial diffusion component Theorem 7.6 and Corollary 7.8 can be combined and the reader is refer to Remark 7.11.

Theorem 7.6. *For $\alpha > \frac{1}{2}$, the density $p_t(x,y)$ is continuous in $t \in (0,T]$. In fact, for given $s,t \in (0,T]$, there exists constant $C_{s,t}$ independent of $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and locally bounded for $s,t \in (0,T]$ such that*

$$|p_t(x,y) - p_s(x,y)| \le C_{s,t}\big(\mathbb{E}\left[|V_t - V_s|^\beta\right] + |t-s|^{(1-\frac{\widehat{\lambda}}{\alpha})\wedge(1-\frac{1}{2\alpha})}\big),$$

where $\beta \in (0, \frac{k}{2+k} \wedge \frac{2\alpha-1}{1+2\alpha}) \cap (0,\alpha)$, and $\widehat{\lambda} = \alpha - w + 1 - \frac{k}{2}$ for $w \in (\alpha \vee (1-\frac{k}{2}), 1 \wedge (\alpha + (1-\frac{k}{2})))$.

Before proceeding to the proof, we want to point out that if V is a true α-stable subordinator, then $\mathbb{E}\left[|V_t - V_s|^\beta\right] = |t-s|^{\frac{\beta}{\alpha}}$ and the density $p_t(x,y)$ is clearly locally Hölder continuous in time with exponent $l \in (0, \frac{k}{\alpha(2+k)} \wedge \frac{2\alpha-1}{\alpha(1+2\alpha)} \wedge (1-\frac{\widehat{\lambda}}{\alpha}) \wedge (1-\frac{1}{2\alpha}))$. In the case where V is an stable-like subordinators, one can obtain similar results by applying moment estimates results of Luschgy and Pagès [13], if in addition to Hypotheses 3.2, the Lévy measure μ of V satisfies the following hypothesis.

Hypotheses 7.7. The Lévy measure μ of V is such that
(i) there exists φ, a regularly varying function at zero with index $-\alpha$ and $y \in (0,1]$ such that $\int_{(x,\infty)} \mu(dc) \le \varphi(x)$ for $x \in (0,y]$.
(ii) The function $x^\alpha \varphi(x)$ is locally bounded for $0 < x \le y$.

Proof. First, recall that the density $p_t(x,y)$ is given in (30). Furthermore, we assume without loss of generality that $t > s$. The quantity that needs to be computed is

$$\begin{aligned}|I_t^n(x,y) - I_s^n(x,y)| \le& \int_s^t \dots \int_0^{t_{n-1}} dt_1 \dots dt_n |K(t_n,\dots,t_1,t;x,y)| \\ &+ \int_0^s \dots \int_0^{t_{n-1}} dt_1 \dots dt_n |K(t_n,\dots,t_1,t;x,y) - K(t_n,\dots,t_1,s;x,y)|. \quad (38)\end{aligned}$$

To compute the first term, we set $\gamma := \frac{\widehat{\lambda}}{\alpha} \vee \frac{1}{2\alpha}$ and from (32) and Corollary 9.9, we have

$$\begin{aligned}\int_s^t \dots \int_0^{t_{n-1}} dt_1 \dots dt_n |K(t_n,\dots,t_1,t;x,y)| &\le \int_s^t \dots \int_0^{t_{n-1}} dt_1 \dots dt_n\, t^{-\frac{d}{2\alpha}} \prod_{i=0}^{n-1}(t_i - t_{i+1})^{-\gamma} \\ &\le \frac{\Gamma(1-\gamma_1)^{n-1} t^{-\frac{d}{2\alpha}}}{\Gamma(1+(n-1)(1-\gamma_1))} \int_s^t t_1^{(n-1)(1-\gamma)}(t-t_1)^{-\gamma} dt_1 \\ &\le \frac{\Gamma(1-\gamma_1)^{n-1} T^{(1-\gamma)(n-1)}}{t^{\frac{d}{2\alpha}}\Gamma(1+(n-1)(1-\gamma_1))} \frac{(t-s)^{1-\gamma}}{1-\gamma}. \quad (39)\end{aligned}$$

the only condition we require here is that $\gamma < 1$, which is the basic assumption required for existence of density in Theorem 6.1.

One considers now the second term in (38), the idea of the proof is similar to Theorem 6.1 and we shall point out only the differences. In the current case, we refer to the notations in Lemma 9.8 and set $t_0 = s$ and $(\lambda_1, \lambda_2) = (1-\frac{k}{2}, 1-\frac{k}{2}+\frac{\beta}{1-\beta})$, $(\delta_1,\delta_2) = (0, \frac{\beta}{1-\beta})$ and $(\gamma_1,\gamma_2) = (\frac{1}{2}, \frac{1}{2}+\frac{\beta}{1-\beta})$.

The proof is similar to that of Theorem 7.3. In place of $\overline{K}$, we work with

$$\int \widehat{p}_{t_n}^{z_n}(x, z_n) G^{H_\Lambda}_{V_s - V_{t_1}}(z_1 - y) \prod_{i=1}^{n-1} \widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i) \widehat{p}_{t_i - t_{i+1}}^{z_i}(z_{i+1}, z_i) dz_{i+1} dz_1.$$

where $G^{H_\Lambda}_{V_s - V_{t_1}}(z_1 - y)$ is defined in Lemma 9.8.

The difference between the current proof and the ones in Theorem 6.1 or Theorem 7.3 (see (36)) is that we first apply Lemma 9.8 and by replace $\mathbb{W}^{\lambda,\delta,\gamma}_{\eta_0, \Delta_{t_1} V}(c_0)$ by $\sum_{i=1}^{2} \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta_0, V_s - V_{t_1}}(c_0)$. Then we convolute the Gaussian densities and from the independent increment property of V, and the fact that the paths of V are increasing, we have

$$\begin{aligned}
&|K(t_n, \dots t_1, t; x, y) - K(t_n, \dots t_1, s; x, y)| \\
&\leq \int_0^1 \mathbb{E}\Big[|V_t - V_s|^\beta \big(vV_t + (1-v)V_s\big)^{-\frac{d}{2}} \int_{\mathbb{R}_+} \sum_{i=1}^{2} \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta_0, V_s - V_{t_1}}(c_0) \widehat{\mu}(dc_0) \prod_{i=1}^{n-1} \int_{\mathbb{R}_+} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i, \Delta_{t_{i+1}} V}(c_i) \widehat{\mu}(dc_i)\Big] dv \\
&\leq C^n \mathbb{E}\left[|V_t - V_s|^\beta\right] \mathbb{E}\big[V_s^{-\frac{d}{2}} \int_{\mathbb{R}_+} \sum_{i=1}^{2} \mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta_0, V_s - V_{t_1}}(c_0) \widehat{\mu}(dc_0) \prod_{i=1}^{n-1} \int_{\mathbb{R}_+} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta_i, \Delta_{t_{i+1}} V}(c_i) \widehat{\mu}(dc_i)\big] \\
&\leq C^n \mathbb{E}\left[|V_t - V_s|^\beta\right] s^{-\frac{d}{2\alpha}} (s - t_1)^{-\bar{\gamma}} \prod_{i=1}^{n-1} (t_i - t_{i+1})^{-\frac{\widehat{\lambda}}{2\alpha} \vee \frac{1}{2\alpha}}
\end{aligned}$$

where $\bar{\gamma}$ is given by

$$\bar{\gamma} := \max\left(\frac{\alpha - w_1 + \lambda_1}{\alpha}, \frac{\alpha - w_2 + \lambda_2}{\alpha}, \frac{\delta_1}{\alpha}, \frac{\delta_2}{\alpha}, \frac{\gamma_1}{\alpha}, \frac{\gamma_2}{\alpha}\right) = \max\left(\frac{\gamma_2}{\alpha}, \frac{\alpha - w_2 + \lambda_2}{\alpha}\right), \tag{40}$$

and $w_i \in (\alpha \vee \lambda_i, 1 \wedge (\lambda_i + \alpha))$. In order to apply Corollary 9.9, we require $\bar{\gamma} < 1$, which is equivalent to $\beta < \frac{2\alpha - 1}{1 + 2\alpha}$ and $1 - \frac{k}{2} + \frac{\beta}{1 - \beta} < w_1 < 1$. The parameter w_1 can be chosen so as to satisfy the previous condition if and only if $\beta < \frac{k}{2+k}$. Therefore by combining all the conditions, we require $\alpha > \frac{1}{2}$ and $\beta \in (0, \frac{k}{2+k} \wedge \frac{2\alpha - 1}{1 + 2\alpha}) \cap (0, \alpha)$. □

Corollary 7.8. *If the drift term $b = 0$, then the density function $p_t(x, y)$ satisfies*

$$|p_t(x, y) - p_s(x, y)| \leq C_{s,t}\big(\mathbb{E}\left[|V_t - V_s|^\beta\right] + |t - s|^{(1 - \frac{\widehat{\lambda}}{\alpha})}\big),$$

for $\beta \in (0, \frac{k}{2+k}) \cap (0, \alpha)$, and $\widehat{\lambda} = \alpha - w + 1 - \frac{k}{2}$ with $w \in (\alpha \vee (1 - \frac{k}{2}), 1 \wedge (\alpha + (1 - \frac{k}{2})))$.

Proof. Given that $b = 0$, then γ in (39) reduces to $\frac{\widehat{\lambda}}{\alpha}$ and one need not to consider γ_1 and γ_2 in (40). The condition $\bar{\gamma} < 1$ reduces to $\alpha - w_1 + \lambda_1 < \alpha$ and as explained in the proof of Theorem 7.6, the parameter w_1 can be chosen such that $\alpha - w_1 + \lambda_1 < \alpha$, whenever $\beta \in (0, \frac{k}{2+k}) \cap (0, \alpha)$. □

Corollary 7.9. *If the Lévy measure μ of the subordinator V satisfies Hypotheses 7.7, then the density $p_t(x, y)$ is locally Hölder continuous in $t \in \mathbb{R}_+$ with Hölder exponent l, then*
(i) for $\alpha > \frac{1}{2}$, $l \in (0, \frac{k}{\alpha(2+k)} \wedge \frac{2\alpha - 1}{\alpha(1 + 2\alpha)} \wedge (1 - \frac{\widehat{\lambda}}{\alpha}) \wedge (1 - \frac{1}{2\alpha}))$,
(ii) for $b = 0$, $l \in (0, \frac{k}{\alpha(2+k)} \wedge (1 - \frac{\widehat{\lambda}}{\alpha}))$.
where $\widehat{\lambda} = \alpha - w + 1 - \frac{k}{2}$ with $w \in (\alpha \vee (1 - \frac{k}{2}), 1 \wedge (\alpha + (1 - \frac{k}{2})))$.

Proof. In the following, we prove only (i), since the proof of (ii) is similar. From Hypotheses 3.2 (iii), we deduce that the Blumenthal-Getoor index of V is α and by stationary increment, we have for $\beta \in (0, \alpha \wedge \frac{k}{2+k} \wedge \frac{2\alpha - 1}{1 + 2\alpha})$,

$$\mathbb{E}[|V_t - V_s|^\beta] = \mathbb{E}[|V_{t \vee s - s \wedge t}|^\beta].$$

Then under Hypotheses 7.7, one can apply Theorem 2 (b) in [13] to show that for fixed $s > 0$ and $\beta \in (0, \alpha \wedge \frac{k}{2+k} \wedge \frac{2\alpha-1}{1+2\alpha})$, as $|t-s| \to 0$,

$$\mathbb{E}[|V_t - V_s|^\beta] = O\left(|t-s|^{\frac{\beta}{\alpha}} \left[|t-s|^{\frac{q}{r}} \varphi(|t-s|^{\frac{1}{\alpha}})^{\frac{q}{r}} + |t-s|^{\frac{p}{q}} \varphi(|t-s|^{\frac{1}{\alpha}})^{\frac{p}{q}}\right]\right)$$

where $r \in (\alpha, 1]$, $q \in [p, \alpha]$ and the term $\left[|t-s|^{\frac{q}{r}} \varphi(|t-s|^{\frac{1}{\alpha}})^{\frac{q}{r}} + |t-s|^{\frac{p}{q}} \varphi(|t-s|^{\frac{1}{\alpha}})^{\frac{p}{q}}\right]$ is locally bounded for $|t-s|$ small enough. Therefore, by combining with Theorem 7.6, we can conclude that the density $p_t(x,y)$ is locally Hölder continuous in time with exponent $l \in (0, \frac{\beta}{\alpha} \wedge (1 - \frac{\widehat{\lambda}}{\alpha}) \wedge (1 - \frac{1}{2\alpha}))$. □

Remark 7.10. In the case where V is a α-stable, tempered α-stable or Lamperti α-stable subordinator, Hypotheses 7.7 is satisfied with $\varphi(x) := \frac{C}{\alpha x^\alpha}$ for $x \in (0,1]$. In all three cases, Theorem 2 (b) in [13], can be applied to show that for $\beta < \alpha$

$$\mathbb{E}[|V_t - V_s|^\beta] = O(|t-s|^{\frac{\beta}{\alpha}}),$$

as $|t-s| \to 0$.

Remark 7.11. In the case where there is a non-trivial diffusion component, Theorem 7.6 and Corollary 7.8 can be combined and state as the following. For $b \neq 0$ and $\alpha \in (0,1)$, given $s, t \in \mathbb{R}_+$, there exists constant $C_{s,t}$ independent of $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and locally bounded for $s, t \in \mathbb{R}_+$ such that

$$|p_t(x,y) - p_s(x,y)| \leq C_{s,t}\big(\mathbb{E}\left[|V_t - V_s|^\beta\right] + |t-s|^{\frac{k}{2} \wedge \frac{k'}{2} \wedge \frac{1}{2}}\big),$$

where $\beta \in (0, \frac{k}{2+k} \wedge \frac{k'}{2+k'} \wedge \frac{1}{3}) \cap (0, \alpha)$.

Corollary 7.12. *The density $p_t(x,y)$ is jointly continuous in $(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times (0,T]$.*

Proof. It is sufficient to combine Theorem 7.2, Theorem 7.3 and Theorem 7.6. □

8 Stochastic Representation

One of the interesting contributions of the working paper of Bally and Kohatsu [1] is that once one has obtained an asymptotic expansion of the semigroup and its density, it is possible to rewrite it as an expectation, which maybe used for other applications. Let us introduce the setup given in [1]. This is also the case here if we suppose that the subordinator V can be exactly simulated.

Hypotheses 8.1. There exist a continuous time Markov chain $(\widehat{X}^z_s)_{s \geq 0}$ such that $\widehat{X}^z_0 = z$ and the transition density of $\widehat{X}^z$ is given by

$$\mathbb{P}\big(\widehat{X}^z_u \in dx \,\big|\, \widehat{X}^z_s = y\big) = H^{-1}_{u-s}(y) \widehat{p}^y_{u-s}(x,y) dx.$$

$$H_{u-s}(y) := \int \widehat{p}^y_{u-s}(x,y) dx.$$

Let N be a Poisson process independent of the Markov chain $\widehat{X}$ with parameter ξ and jump times given by $(T_i)_{i \in \mathbb{N}}$. Given the jump times of the Poisson process N, we define

$$\Gamma^*_T(z) := \begin{cases} H_{T-T_{N_T}}(\widehat{X}^z_{T_{N_T}}) \prod_{k=0}^{N_t - 1} H_{T_{k+1}-T_j}(\widehat{X}^z_{T_j}) \widehat{\theta}_{T_{k+1}-T_k}(\widehat{X}^z_{T_{k+1}}, \widehat{X}^z_{T_k}) & N_T \geq 1 \\ H_T(z) & N_T = 0. \end{cases}$$

Let $\tau_T := T_{N_T}$ be the last jump of the Poisson process N before T.

Theorem 8.2 (Bally and Kohatsu [1]). *Under Hypotheses 8.1, for every $g \in C_c^\infty(\mathbb{R}^d)$ and h a probability density function, we have*

$$P_T^* g(y) = e^{\xi T}\mathbb{E}\left[\xi^{-N_T} g(\widehat{X}_{\tau_T}^y)\Gamma_T^*(y)\right]$$
$$P_T h(x) = e^{\xi T}\mathbb{E}\left[\xi^{-N_T} \widehat{p}_{T-\tau_T}^{\widehat{X}_{\tau_T}^Z}(x, \widehat{X}_{\tau_T}^Z)\Gamma_T^*(Z)\right]$$
$$p_T(x,y) = e^{\xi T}\mathbb{E}\left[\xi^{-N_T} \widehat{p}_{T-\tau_T}^{\widehat{X}_{\tau_T}^y}(x, \widehat{X}_{\tau_T}^y)\Gamma_T^*(y)\right]$$

where Z is a random variable with density h, independent of the Markov chain $\widehat{X}$ and the Poisson process N.

8.1 A probabilistic representation

To obtain a probabilistic representation of the density, which may be suitable for Monte-Carlo simulation, we derive first a probabilistic representation of the basic functions $\widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i)\widehat{p}_{t_i-t_{i+1}}^{z_i}(z_{i+1}, z_i)$. We work in one dimension and for notational simplicity, given $n \in \mathbb{N}$ and fixed $t > t_1 > \dots t_n > 0$, we set for $i = 0, \dots, n-1$, $\Delta_{t_{i+1}} V =: V_{t_i} - V_{t_{i+1}}$ and

$$L(c_i, \Delta_{t_{i+1}}V, z_{i+1}, z_i) := \left(\frac{q(a(z_i)\Delta_{t_{i+1}}V + c_i a(z_{i+1}), z_{i+1} - z_i)}{q(a(z_i)\Delta_{t_{i+1}}V + c_i a(z_i), z_{i+1} - z_i)} - 1\right).$$

Note that L can be written explicitly as it is based on Gaussian density functions.

Let N be a Poisson process with parameter λ and jump times given by $(\tau_i)_{i\in\mathbb{N}}$ and $\tau_0 = 0$. For $i = 0, \dots, n-1$, we define the following family of mutually independent r.v.'s

$$N_i \sim \mathcal{N}(0,1), \qquad \text{law}(C_i^1)(dc) = \frac{\mathbb{1}_{(0,1]}(c)}{\widehat{\mu}((0,1])}\widehat{\mu}(dc), \qquad \text{law}(C_i^2)(dc) = \frac{\mathbb{1}_{(1,\infty)}(c)}{\widehat{\mu}((1,\infty))}\widehat{\mu}(dc),$$
$$\eta_i \sim \text{Bernoulli}(\frac{1}{2}), \qquad \widehat{\eta}_i \sim \text{Bernoulli}(\frac{1}{2}).$$

where the random variables $(\widehat{\eta}_i)_i$ is used to combine the small and large jumps, while $(\eta_i)_i$ is used to combine the jumps and the drift. For simplification, let

$$\widehat{L}_{b,C_i^1,C_i^2,\widehat{\eta}_i,\eta_i}(\Delta_{t_{i+1}}V, z_{i+1}, z_i) :=$$
$$\frac{2(1-\eta_i)b(z_{i+1})z_{i+1}}{a(z_i)\Delta_{t_{i+1}}V} + 4\eta_i\left[(1-\widehat{\eta}_i)\frac{\widehat{\mu}((0,1])}{C_i^1}L(C_i^1, \Delta_{t_{i+1}}V, z_{i+1}, z_i) + \widehat{\eta}_i\widehat{\mu}((1,\infty))L(C_i^2, \Delta_{t_{i+1}}V, z_{i+1}, z_i)\right]$$
$$U_{C_i^2,C_i^1,\widehat{\eta}_i,\eta_i}(z_i, \Delta_{t_{i+1}}V) := a(z_i)(\Delta_{t_{i+1}}V + \eta_i(C_i^1(1-\widehat{\eta}_i) + C_i^2\widehat{\eta}_i)).$$

We define the random variables $Z_{t_{i+1}}$ recursively as

$$Z_{t_0} := y \quad \text{and} \quad Z_{t_{i+1}} := Z_{t_i} + \sigma(Z_{t_i})(\Delta_{t_{i+1}}V + \eta_i(C_i^1(1-\widehat{\eta}_i) + C_i^2\widehat{\eta}_i))^{\frac{1}{2}}N_i.$$

Theorem 8.3.

$$p_t(x,y) = e^{\lambda t}\mathbb{E}\Big(\lambda^{-N_t} q(a(Z_{\tau_1})V_{\tau_1}, x - Z_{\tau_1}) \prod_{i=0}^{N_t - 1} \widehat{L}_{b,C_i^1,C_i^2,\widehat{\eta}_i,\eta_i}(\Delta_{\tau_{N_t - i}}V, Z_{\tau_{N_t-i}}, Z_{\tau_{N_t-i+1}})\Big),$$

Proof. In the expression (8) of $\widehat{\theta}_{t_i-t_{i+1}}(z_{i+1}, z_i)\widehat{p}_{t_i-t_{i+1}}^{z_i}(z_{i+1}, z_i)$, we consider first the integral against $\mu(dc)$ and split it into integrals on $(0,1]$ and $(1,\infty)$. For $j = 1, 2$, by using the independence between

V and $(C_i^j)_{i=1,\dots,n}$, we can write

$$\begin{aligned}
&\int_{\mathbb{R}_+} \mathbb{E}\big(q(a(z_i)V_t + a(z_{i+1})c, z_{i+1} - z_i) - q(a(z_i)V_t + a(z_i)c, z_{i+1} - z_i)\big)\mu(dc) \\
&= \mathbb{E}\big(\frac{\widehat{\mu}((0,1])}{C_i^1} L(C_i^1, \Delta_{t_{i+1}}V, z_{i+1}, z_i) q(a(z_i)(\Delta_{t_{i+1}}V + C_i^1), z_{i+1} - z_i)\big) \\
&\quad + \mu((1,\infty))\mathbb{E}\big(L(C_i^2, \Delta_{t_{i+1}}V, z_{i+1}, z_i) q(a(z_i)(\Delta_{t_{i+1}}V + C_i^2), z_{i+1} - z_i)\big).
\end{aligned}$$

The next step is to combine the Gaussian densities. For a fixed $n \in \mathbb{N}$, using that fact that the sequence $(\eta_i)_{i=1,\dots,n}$ is independent of V and $(C_i^j)_{i=1,\dots,n}$ for $j = 1, 2$,

$$\begin{aligned}
&= \mathbb{E}\left(2(1-\widehat{\eta})\frac{\widehat{\mu}((0,1])}{C_i^1} L(C_i^1, \Delta_{t_{i+1}}V, z_{i+1}, z_i) q(a(z_i)(\Delta_{t_{i+1}}V + C_i^1(1-\widehat{\eta}_i)), z_{i+1} - z_i)\right) \\
&\quad + \mathbb{E}\left(2\widehat{\eta}\widehat{\mu}((1,\infty)) L(C_i^2, \Delta_{t_{i+1}}V, z_{i+1}, z_i) q(a(z_i)(\Delta_{t_{i+1}}V + C_i^2\widehat{\eta}_i), z_{i+1} - z_i)\right) \\
&= \mathbb{E}\Big(\Big(2(1-\widehat{\eta}_i)\frac{\widehat{\mu}((0,1])}{C_i^1} L(C_i^1, \Delta_{t_{i+1}}V, z_{i+1}, z_i) + 2\widehat{\eta}_i\widehat{\mu}((1,\infty)) L(C_i^2, \Delta_{t_{i+1}}V, z_{i+1}, z_i)\Big) \\
&\quad \times q(a(z_i)(\Delta_{t_{i+1}}V + C_i^2\widehat{\eta}_i + C_i^1(1-\widehat{\eta}_i)), z_{i+1} - z_i)\Big).
\end{aligned}$$

The next step is to use make use of η_i to combine the integral with respect to μ with the drift term. By using independence, the basic function $\widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i)\widehat{p}^{z_i}_{t_i - t_{i+1}}(z_{i+1}, z_i)$ can be written into

$$\widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i)\widehat{p}^{z_i}_{t_i - t_{i+1}}(z_{i+1}, z_i) = \mathbb{E}\left(\widehat{L}_{b, C_i^1, C_i^2, \widehat{\eta}_i, \eta_i}(\Delta_{t_{i+1}}V, z_{i+1}, z_i) q(U_{C_i^2, C_i^1, \widehat{\eta}_i, \eta_i}(z_i, \Delta_{t_{i+1}}V), z_{i+1} - z_i)\right).$$

For a fixed $n \in \mathbb{N}$, one considers first the space integral $a(t_n, \dots, t_1, t; x, y)$ in (9). Using the independence assumption between $(N_i)_{i=1,\dots,n}$, and the increments of V, $(\eta_i)_{i=1,\dots,n}$, $(\widehat{\eta}_i)_{i=1,\dots,n}$ and $(C_i^j)_{i=1,\dots n}$ for $j = 1, 2$, we can write using the tower property of conditional expectations

$$\begin{aligned}
&K(t_n, \dots, t_1, t; x, y) \\
&= \mathbb{E}\Big[\int q(a(z_n)V_{t_n}, x - z_n) \prod_{i=0}^{n-1} \widehat{L}_{b, C_i^1, C_i^2, \widehat{\eta}_i, \eta_i}(\Delta_{t_{i+1}}V, z_{i+1}, z_i) q(U_{C_i^2, C_i^1, \widehat{\eta}_i, \eta_i}(z_i, \Delta_{t_{i+1}}V), z_{i+1} - z_i) dz_{i+1}\Big] \\
&= \mathbb{E}\Big[q(a(Z_{t_n})V_{t_n}, x - Z_{t_n}) \prod_{i=0}^{n-1} \widehat{L}_{b, C_i^1, C_i^2, \widehat{\eta}_i, \eta_i}(\Delta_{t_{i+1}}V, Z_{t_{i+1}}, Z_{t_i})\Big]. \qquad (41)
\end{aligned}$$

Next, note that each term $I_t^n(y, x)$ in the partial sum in (11) can be rewritten into

$$I_t^n(y, x) = \mathbb{P}(N_t = n)\lambda^{-n} e^{\lambda t}\mathbb{E}\big[K(\tau_1, \dots, \tau_n, t; x, y) \,\big|\, N_t = n\big]. \qquad (42)$$

The second equality, follows from that fact that given $N_t = n$, the conditional distribution of the ordered jump times $(\tau_1, \dots, \tau_n)$ is given by $\frac{n!}{t^n}\mathbb{1}_{\{t > t_1 > \dots > t_n\}} dt_n \dots dt_1$. Note that the indices are inverted in that t_i corresponds to the realization of the jump τ_{n-i+1} and $\Delta_{\tau_{n-i}}V := V_{\tau_{n-i+1}} - V_{\tau_{n-i}}$.

Using the fact that the Poisson process N is independent from V, $(C_i^1)_{i=1,\dots,n}$, $(C_i^2)_{i=1,\dots,n}$, $(\widehat{\eta}_i)_{i=1,\dots,n}$ and $(\eta_i)_{=1,\dots,n}$, we see from (41) that

$$\begin{aligned}
&K(\tau_1, \dots, \tau_n, t; x, y) = K(t_n, \dots, t_1, t; x, y)\Big|_{t_i = \tau_{n-i+1}} \\
&= \mathbb{E}\Big[q(a(Z_{t_n})V_{t_n}, x - Z_{t_n}) \prod_{i=1}^{n-1} \widehat{L}_{b, C_i^1, C_i^2, \widehat{\eta}_i, \eta_i}(\Delta_{t_{i+1}}V, Z_{t_{i+1}}, Z_{t_i}) \,\Big|\, N_t = n, (\tau_i)_{i=1\dots n}\Big]\Big|_{t_i = \tau_{n-i+1}} \\
&= \mathbb{E}\Big[q(a(Z_{\tau_1})V_{\tau_1}, x - Z_{\tau_1}) \prod_{i=0}^{n-1} \widehat{L}_{b, C_i^1, C_i^2, \widehat{\eta}_i, \eta_i}(\Delta_{\tau_{n-i}}V, Z_{\tau_{n-i}}, Z_{\tau_{n-i+1}}) \,\Big|\, N_t = n, (\tau_i)_{i=1\dots n}\Big]
\end{aligned}$$

Finally, by substituting the above in (42), and by using the tower property of conditional expectations and the law of total probability, we have the result. □

9 Appendix

Lemma 9.1. *For every $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$*

$$(\mathcal{L}^{y_1} - \widehat{\mathcal{L}}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x} = b(x)^T \nabla_x \widehat{p}_t^y(x, y) + \int_{\mathbb{R}_+} \mathbb{E}\left[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\right]\mu(dc).$$

Proof. By applying the generators (3) and (4) to $\widehat{p}_t^y(\cdot, y)(x)$ and Lemma 3.6, we obtain

$$\begin{aligned}
&(\mathcal{L}^{y_1} - \widehat{\mathcal{L}}^y)(\widehat{p}_t^y(\cdot, y))(x)\Big|_{y_1=x} \\
&= b(x)^T \nabla_x \widehat{p}_t^y(x, y) \\
&\quad + \mathbb{E}\Big[\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} q(a(y)V_t, x + \sigma(x)u - y) - q(a(y)V_t, x + \sigma(y)u - y)q(c, u)du\mu(dc)\Big].
\end{aligned}$$

The final result will follow from the explicit calculation for $r = x, y$ of the convolution

$$\int_{\mathbb{R}} q(a(y)V_t, x + \sigma(r)u - y)q(c, u)du = q(a(y)V_t + a(r)c, x - y).$$

□

9.1 Estimates and Derivatives for the Gaussian type densities

Lemma 9.2. *Suppose M is a positive definite matrix and $y \in \mathbb{R}^d$. The Gaussian density $q(M, y)$ satisfies the following properties*
(i) Suppose $\underline{C}I \le M \le \overline{C}I$, for $\underline{C}, \overline{C} \in \mathbb{R}_+$ then for any $y \in \mathbb{R}^d$

$$\left(\frac{\overline{C}}{\underline{C}}\right)^{-\frac{d}{2}} q(\underline{C}, y) \le q(M, y) \le \left(\frac{\overline{C}}{\underline{C}}\right)^{\frac{d}{2}} q(\overline{C}, y).$$

(ii) For any $p > 0$, there exists a constant $C(p)$, independent of $t > 0$, such that for any $y \in \mathbb{R}^d$

$$\left(\frac{|y|^{2+p}}{t^{1+\frac{p}{2}}}\right) q(t, y) = \left(\frac{|y|^{2+p} e^{-|y|^2/4t}}{t^{1+\frac{p}{2}}}\right) q(2t, y) \le C(p) q(2t, y). \tag{43}$$

(iii) For any $y \in \mathbb{R}^d$, $i, j = 1, \ldots, d$ and $t > 0$, if $\underline{C}I \le M \le \overline{C}I$, then

$$|\nabla_y q(Mt, y)| \le Ct^{-\frac{1}{2}} q(\overline{C}t, y), \qquad \|\nabla_y^2 q(Mt, y)\|_F \le Ct^{-1} q(\overline{C}t, y).$$

(iv) Given a positive definite matrix M and $p > 0$ then

$$q(M, x)^p = ((2\pi)^d \det(M))^{-\frac{p-1}{2}} p^{-\frac{d}{2}} q(Mp^{-1}, x).$$

Proof. (i) and (ii) have been proven in [1]. For (iii), we have that

$$|\nabla_y q(Mt, y)| \le \|(Mt)^{-1}\|_F |y| q(\overline{C}t, y) \le (\underline{C}t)^{-\frac{1}{2}} q(\overline{C}t, y) \tag{44}$$

where the inequality follows from (i), Cauchy-Schwartz inequality and Lemma 9.11. For the second derivative,

$$\begin{aligned}
\|\nabla_y^2 q(Mt, y)\|_F &= \|(Mt)^{-1} yy^T (Mt)^{-1} + (Mt)^{-1}\|_F q(\overline{C}t, y) \\
&\le \left\{\|(Mt)^{-1}\|_F \|yy^T\|_F \|(Mt)^{-1}\|_F + \|(Mt)^{-1}\|_F\right\} q(\overline{a}t, y) \\
&\le \left\{(\underline{C}t)^{-2}|y|^2 + (\underline{C}t)^{-1}\right\} q(\overline{C}t, y) \\
&\le C(\underline{C}t)^{-1} q(\overline{C}t, y)
\end{aligned}$$

(iv) is proved by explicit calculation. □

Lemma 9.3. *Let M_0 and M_1 be positive definite $d \times d$ matrices so that for any $v \in [0,1]$ we have that for all $z \in \mathbb{R}^d$, $\underline{C}I \leq M_v \leq \overline{C}I$ for $M_v := vM_1 + (1-v)M_0$. Then for any $x \in \mathbb{R}^d$ there exists a constant C_d which depends only on d such that*

$$|q(M_1, x) - q(M_0, x)| \leq C_d \left(\frac{\overline{C}}{\underline{C}}\right)^{\frac{d}{2}} \left[1 + \underline{C}^{-1}|x|^2\right] \underline{C}^{-1}\|M_1 - M_0\|_2 q(\overline{C}, x).$$

Let M be a positive definite matrix such that for all $z \in \mathbb{R}^d$, $\underline{C}I \leq M \leq \overline{C}I$. Define $v^* \equiv v^*(x,y) =$ $\operatorname{argmin}\{\|vx + (1-v)y\|; v \in [0,1]\}$. Then we have for any $x, y \in \mathbb{R}^d$,

$$|q(M, x) - q(M, y)| \leq C_d|x - y|\underline{C}^{-\frac{1}{2}} q(\overline{C}, v^*x + (1-v^*)y).$$

Proof. The proof is based on the analysis of the function $v \to q(M_v, x)$. That is, by using the fundamental theorem of calculus, we obtain that

$$\partial_v q(M_v, x) = \left[-\operatorname{Tr}\left(M_v^{-1}(M_1 - M_0)\right) + x^T M_v^{-1}(M_1 - M_0)M_v^{-1}x\right] q(M_v, x) \tag{45}$$

Furthermore, standard inequalities for the trace of a matrix through Frobenius norms and properties of spectral norms together with Lemma 9.2 gives

$$\begin{aligned}
|q(M_1, x) - q(M_0, x)| &\leq \int_0^1 |\partial_v q(M_v, x)|\, dv \\
&\leq \int_0^1 \left[\operatorname{Tr}\left(M_v^{-1}(M_1 - M_0)\right) + \|M_v^{-1}\|_2^2 \|M_1 - M_0\|_2 |x|^2\right] q(M_v, x) dv \\
&\leq \left(\frac{\overline{C}}{\underline{C}}\right)^{\frac{d}{2}} \int_0^1 \left[\|M_v^{-1}\|_F \|M_1 - M_0\|_F + \underline{C}^{-2}\|M_1 - M_0\|_2 |x|^2\right] q(\overline{C}, x) dv.
\end{aligned}$$

The proof of the second assertion is similar. □

Lemma 9.4. *For $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$, $c > 0$*

$$|q(a(y)V_t + a(x)c, x - y) - q(a(y)(V_t + c), x - y)| \leq Cc(V_t + c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_t + c), x - y)$$

Proof. The proof follows by direct application of Lemma 9.3. In fact, taking $M_1 = a(y)V_t + a(x)c$, $M_0 = a(y)(V_t + c)$, $\overline{C} = \overline{a}(V_t + c)$, $\underline{C} = \underline{a}(V_t + c)$, we have by applying Lemma 9.2

$$\begin{aligned}
&|q(a(y)V_t + a(x)c, x - y) - q(a(y)(V_t + c), x - y)| \\
&\leq C_d \left(\frac{\overline{a}}{\underline{a}}\right)^{\frac{d}{2}} \left[(\underline{a}(V_t + c))^{-1} c|x - y|^k + \underline{a}(V_t + c)^{-2} c|x - y|^{2+k}\right] q(\overline{a}(V_t + c), x - y) \\
&\leq Cc(V_t + c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_t + c), x - y).
\end{aligned}$$

□

Lemma 9.5. *Given $y, y^* \in \mathbb{R}^d$ and $\bar{y} \in \overline{B(y^*, |y - y^*|)}$, suppose $|y^* - y|^2 \leq v$, then for any $\epsilon > 0$*

$$q(Cv, x - \bar{y}) \leq \left(1 + \frac{1}{\epsilon}\right)^{d/2} e^{\frac{\epsilon}{C}} q(C_\epsilon v, x - y^*).$$

Here $C_\epsilon := C(1 + \frac{1}{\epsilon})$.

Proof. Using Young's inequality, we have that for any $\epsilon > 0$, $|x|^2 - (1+\epsilon)|y|^2 \leq (1 + \frac{1}{\epsilon})|x - y|^2$, we obtain that

$$(1 + \frac{1}{\epsilon})|x - y^* - (\bar{y} - y^*)|^2 \geq |x - y^*|^2 - (1+\epsilon)|\bar{y} - y^*|^2 \geq |x - y^*|^2 - (1+\epsilon)|y - y^*|^2$$

On the set $|y^* - y|^2 \leq v$, we have that $(1 + \frac{1}{\epsilon})|x - y^* - (\bar{y} - y^*)|^2 \geq |x - y^*|^2 - (1+\epsilon)v$ and therefore $e^{-\frac{|x-\bar{y}|^2}{2Cv}} \leq e^{-(1+\frac{1}{\epsilon})^{-1}\frac{|x-y^*|^2}{2Cv}} e^{\frac{\epsilon}{C}}$. □

9.2 Preliminary estimates

In the following, we prove that $\widehat{\theta}(x,y)\widehat{p}_t^y(x,y)$ is locally Hölder continuous in y. In the case where there is a non-trivial diffusion component the Hölder parameter β in Lemma 9.6 takes values in $(0, k\wedge k')$.

Lemma 9.6. *The function $\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y)$ is locally β-Hölder continuous in y with $\beta\in(0,k)$. More explicitly,*

$$|\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y)-\widehat{\theta}_t(x,y^*)\widehat{p}_t^{y^*}(x,y^*)|\le C_{y,y^*}|y-y^*|^\beta(G_{V_t}(x-y)+G_{V_t}(x-y^*))$$

where we set $\lambda_1:=1-\frac{k}{2}+\frac{\beta}{2}$, $\lambda_2:=1-k+\frac{\beta}{2}$ and $\delta_1:=\frac{\beta}{2}$, $\delta_2:=0$, $(\gamma_1,\gamma_2)=(\frac12+\frac{\beta}{2}-\frac{k}{2},\frac{1+\beta}{2})$ and using the notation introduced in (27) *and* (28)*, we set*

$$G_{V_t}(x):=\int_{\mathbb{R}_+\times\mathbb{R}}\mathbb{E}\Big[q(8\mathbb{V}_{\eta,V_t}(c,\bar{c}),x)\sum_{i=1}^2\mathbb{W}_{\eta,V_t}^{\lambda_i,\delta_i,\gamma_i}(c)\Big]\,\widehat{\mu}(dc)\times\delta_0(d\bar{c}),$$

which is finite by Hypotheses 3.2.

Proof. Given $y,y^*\in\mathbb{R}^d$, we use (8) to divide the analysis in various terms,

$$\begin{aligned}|\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y)-\widehat{\theta}_t(x,y^*)\widehat{p}_t^{y^*}(x,y^*)|\le\mathbb{E}\Big[&\big|b(x)^T\nabla_xq(a(y)V_t,x-y)-b(x)^T\nabla_xq(a(y^*)V_t,x-y^*)\big|\\&+\Big|\int_{\mathbb{R}_+}q(a(y)V_t+a(x)c,x-y)-q(a(y)(V_t+c),x-y)\mu(dc)\\&-\int_{\mathbb{R}_+}q(a(y^*)V_t+a(x)c,x-y^*)-q(a(y^*)(V_t+c),x-y^*)\mu(dc)\Big|\Big].\end{aligned}\tag{46}$$

The method is to break the computations on different sets. We will use frequently Lemma 9.3 and Lemma 9.2 without any further mention.

The drift term of (46) on the set $|y-y^*|^2>V_t$

For any $y,y^*\in\mathbb{R}^d$, we compute the term within the expectation of the first term in (46).

$$|b(x)|\times|\nabla_xq(a(y)V_t,x-y)|\le CV_t^{-\frac12}q(2\overline{a}V_t,x-y)\le C|y-y^*|^\beta V_t^{-\frac{1+\beta}{2}}q(2\overline{a}V_t,x-y).\tag{47}$$

This implies that on the set $|y-y^*|^2>V_t$,

$$|b(x)|\times|\nabla_xq(a(y)V_t,x-y)-\nabla_xq(a(y^*)V_t,x-y^*)|\le C|y-y^*|^\beta V_t^{-\frac{1+\beta}{2}}\Big[q(2\overline{a}V_t,x-y)+q(2\overline{a}V_t,x-y^*)\Big].$$

The drift term of (46) on the set $|y-y^*|^2\le V_t$

From (5), we have that the difference will be analyzed term by term as follows

$$\begin{aligned}&|\nabla_xq(a(y)V_t,x-y)-\nabla_xq(a(y^*)V_t,x-y^*)|\le A_1+A_2\\&A_1:=CV_t^{-1}|(a(y)^{-1}-a(y^*)^{-1})(x-y)+a(y^*)^{-1}(y^*-y)|q(a(y)V_t,x-y)\\&A_2:=CV_t^{-1}|a(y^*)^{-1}(x-y^*)||q(a(y)V_t,x-y)-q(a(y^*)V_t,x-y^*)|.\end{aligned}$$

For A_1 it is enough to note that for invertible matrices A and B,

$$A^{-1}-B^{-1}=A^{-1}(B-A)B^{-1}.\tag{48}$$

and we will use this property repeatedly through the proof. Therefore using Lemma 9.2

$$A_1 \leq CV_t^{-1}(|y-y^*|^k V_t^{\frac{1}{2}} + |y-y^*|)q(a(y)V_t, x-y).$$

Finally using $|y-y^*|^2 \leq V_t$ we have that

$$A_1 \leq C|y-y^*|^\beta (V_t^{-\frac{1-k+\beta}{2}} + V_t^{-\frac{1+\beta}{2}})q(\overline{a}V_t, x-y). \tag{49}$$

A_2 is analyzed using the following properties; (i) uniform elliptic assumption, (ii) Lemma 9.3 (as done in the proof of Lemma 9.4), and (iii) Lemma 9.5 in order to replace $x-\alpha^*y-(1-\alpha^*)y^*$ by $x-y$ in the second Gaussian density and then finally apply Lemma 9.2 (ii). That is,

$$\begin{aligned}
A_2 &\leq CV_t^{-1}|x-y^*|\big(|q(a(y^*)V_t, x-y^*) - q(a(y)V_t, x-y^*)| + |q(a(y)V_t, x-y^*) - q(a(y)V_t, x-y)|\big) \\
&\leq CV_t^{-1}|x-y^*|\big(|y-y^*|^k q(2\overline{a}V_t, x-y^*) + V_t^{-\frac{1}{2}}|y-y^*|q(\overline{a}V_t, x-\alpha^*y-(1-\alpha^*)y^*)\big) \\
&\leq CV_t^{-\frac{1}{2}}\big(|y-y^*|^k + V_t^{-\frac{1}{2}}|y-y^*|\big)q(4\overline{a}V_t, x-y^*) \\
&\leq C|y-y^*|^\beta\big(V_t^{-\frac{1-k+\beta}{2}} + V_t^{-\frac{1+\beta}{2}}\big)q(4\overline{a}V_t, x-y^*)
\end{aligned} \tag{50}$$

where the last inequality follows from Lemma 9.5.

Analysis of the integrals in (46) **on** $c \in (0,1]$ **and** $|y-y^*|^2 > (V_t+c)$

It is sufficient to consider only the integrand of the second term in (46) with the following decomposition

$$|q(a(y)V_t + a(x)c, x-y) - q(a(y)(V_t+c), x-y)| + |q(a(y^*)V_t + a(x)c, x-y^*) - q(a(y^*)(V_t+c), x-y^*)|.$$

Both terms both are treated similarly. In fact for the firs term, we have from Lemma 9.4

$$\begin{aligned}
|q(a(y)V_t + a(x)c, x-y) - q(a(y)(V_t+c), x-y)| &\leq C\,c(V_t+c)^{-(1-\frac{k}{2})}q(2\overline{a}(V_t+c), x-y) \\
&\leq Cc|y-y^*|^\beta(V_t+c)^{-(1-\frac{k-\beta}{2})}q(2\overline{a}(V_t+c), x-y).
\end{aligned}$$

Therefore, we see that the integral of this term against $\mu(dc)$ can be bounded by

$$C|y-y^*|^\beta \int_{(0,1]} (V_t+c)^{-(1-\frac{k-\beta}{2})}\big(q(2\overline{a}(V_t+c), x-y) + q(2\overline{a}(V_t+c), x-y^*)\big)\widehat{\mu}(dc), \tag{51}$$

which is integrable by Lemma 3.3 (ii) and Lemma 9.14.

Analysis of the integrals in (46) **on** $c \in (1,\infty)$ **and** $|y-y^*|^2 > (V_t+c)$

From triangular inequality and Lemma 9.2, (i),

$$\begin{aligned}
|q(a(y)V_t + a(x)c, x-y) - q(a(y)(V_t+c), x-y)| &\leq Cq(\overline{a}(V_t+c), x-y) \\
&\leq C|y-y^*|^\beta(V_t+c)^{-\frac{\beta}{2}}q(\overline{a}(V_t+c), x-y)
\end{aligned}$$

and this implies that the integral against $\mu(dc)$ can be bounded by

$$C|y-y^*|^\beta \int_{[1,\infty)} (V_t+c)^{-\frac{\beta}{2}}\big(q(\overline{a}(V_t+c), x-y) + q(\overline{a}(V_t+c), x-y^*)\big)\widehat{\mu}(dc). \tag{52}$$

Analysis of the integral in (46) **on** $c \in (1,\infty)$ **and** $|y-y^*|^2 \le (V_t+c)$

Again we look only at the integrand in (46) and decompose it as follows,

$$\begin{aligned}&\big|q(a(y)V_t+a(x)c,x-y)-q(a(y^*)V_t+a(x)c,x-y^*)\big|\\&\quad+\big|q(a(y)(V_t+c),x-y)-q(a(y^*)(V_t+c),x-y^*)\big|=I_1+I_2.\end{aligned}$$

Firstly, the term I_1 can be upper bounded as follows

$$\begin{aligned}I_1\le&\big|q(a(y)V_t+a(x)c,x-y)-q(a(y^*)V_t+a(x)c,x-y)\big|\\&+\big|q(a(y^*)V_t+a(x)c,x-y)-q(a(y^*)V_t+a(x)c,x-y^*)\big|=I_{1,1}+I_{1,2}.\end{aligned}$$

These terms are estimated using Lemma 9.3 and Lemma 9.2. In fact, one obtains

$$\begin{aligned}I_{1,1}&\le C\left[1+(V_t+c)^{-1}|x-y|^2\right](V_t+c)^{-1}|y-y^*|^kV_tq(\overline{a}(V_t+c),x-y)\\&\le CV_t(V_t+c)^{-1}|y-y^*|^kq(2\overline{a}(V_t+c),x-y)\\&\le C|y-y^*|^kq(2\overline{a}(V_t+c),x-y).\\&\le C_{y,y^*}|y-y^*|^\beta q(2\overline{a}(V_t+c),x-y).\\I_{1,2}&\le C|y-y^*|(V_t+c)^{-\frac{1}{2}}q(\overline{a}(V_t+c),\alpha^*y+(1-\alpha^*)y^*-x)\\&\le C|y-y^*|^\beta(V_t+c)^{-\frac{\beta}{2}}q(2\overline{a}(V_t+c),x-y^*).\end{aligned}$$

Note that Lemma 9.5 has been used in the last inequality. This shows that

$$I_1\le C_{y,y^*}|y-y^*|^\beta\big((V_t+c)^{-\frac{\beta}{2}}+1\big)q(2\overline{a}(V_t+c),x-y^*).$$

The second term I_2 can be estimated similarly. Therefore on this set, we see that the integral against $\mu(dc)$ can be bounded by

$$C_{y,y_*}|y-y^*|^\beta\int_{(1,\infty)}\big((V_t+c)^{-\frac{\beta}{2}}+1\big)q(2\overline{a}(V_t+c),x-y^*)\widehat{\mu}(dc). \tag{53}$$

The analysis of the integral in (46) **on** $c\in(0,1]$ **and** $|y-y^*|^2\le(V_t+c)$

We have to consider again the integrand with respect to $\mu(dc)$ in (46). The problem here is how to obtain a factor of the order c as in the case $c\in(0,1]$ so that one may have integrability through $\widehat{\mu}(dc)$. On the other hand, the analysis has to be made so that terms of the type $|y-y^*|$ appear in the analysis for which we can use Lemma 9.2 in order to obtain the Hölder property. For this analysis, we consider the following decomposition of the integrand in (46).

$$\begin{aligned}A_1:=&q(a(y)V_t+a(x)c,x-y)-q(a(y)V_t+a(x)c,x-y^*)-q(a(y)(V_t+c),x-y)\\&+q(a(y)(V_t+c),x-y^*)\\A_2:=&q(a(y)V_t+a(x)c,x-y^*)-q(a(y)(V_t+c),x-y^*)-q(a(y^*)V_t+a(x)c,x-y^*)\\&+q(a(y^*)(V_t+c),x-y^*).\end{aligned}$$

The analysis of A_2 will follow along the same lines as in the case $c\in[0,1]$ and $|y-y^*|^2\ge(V_t+c)$ because we will obtain the factor c by naturally analyzing terms in pairs. The term A_1 will require first, an application of the fundamental theorem of calculus before decomposing it again in order to obtain the required integrability factor c. In fact,

$$\begin{aligned}A_1=&\int_0^1(x-vy-(1-v)y^*)^T\Big\{(a(y)V_t+a(x)c)^{-1}q(a(y)V_t+a(x)c,x-vy-(1-v)y^*)\\&-(a(y)(V_t+c))^{-1}q(a(y)(V_t+c),x-vy-(1-v)y^*)\Big\}(y-y^*)dv.\end{aligned}$$

We further decompose the term within brackets by taking differences term by term to obtain using (48), the decomposition

$$A_{1,1} := -\,(a(y)(V_t+c))^{-1}(a(y)-a(x))(a(y)V_t+a(x)c)^{-1}cq(a(y)V_t+a(x)c, x-vy-(1-v)y^*)$$
$$A_{1,2} := (a(y)(V_t+c))^{-1}\left(q(a(y)V_t+a(x)c, x-vy-(1-v)y^*) - q(a(y)(V_t+c), x-vy-(1-v)y^*)\right).$$

We see that in $A_{1,1}$, the factor c has already been obtained. In $A_{1,2}$, an application of Lemma 9.3 will give the required factor c as it always appears in the difference between the covariance matrices being considered. We now give each estimate using Lemma 9.2, Lemma 9.3 and Lemma 9.5 repeatedly.

$$\begin{aligned}&|(x-vy-(1-v)y^*)^T A_{1,1}| \vee |(x-vy-(1-v)y^*)^T A_{1,2}|\\ &\le C\|a(y)-a(x)\|_F (V_t+c)^{-3/2} cq(4\overline{a}(V_t+c), x-y).\end{aligned}$$

Therefore using that $|y-y^*|^{1-\beta} \le (V_t+c)^{\frac{1-\beta}{2}}$, we have

$$\begin{aligned}A_1 &\le C\int_0^1 \|a(y)-a(x)\|_F (V_t+c)^{-3/2} cq(4\overline{a}(V_t+c), x-y)dv\ |y-y^*|\\ &\le C\|a(y)-a(x)\|_F (V_t+c)^{-(1+\frac{\beta}{2})} c|y-y^*|^\beta q(4\overline{a}(V_t+c), x-y)\end{aligned}$$

To continue, we apply Lemma 9.2 (ii), to obtain

$$\begin{aligned}A_1 &\le C(V_t+c)^{\frac{k}{2}}(V_t+c)^{-(1+\frac{\beta}{2})} c|y-y^*|^\beta q(8\overline{a}(V_t+c), x-y)\\ &\le C(V_t+c)^{-1+\frac{k}{2}-\frac{\beta}{2}} c|y-y^*|^\beta q(8\overline{a}(V_t+c), x-y^*)\end{aligned} \tag{54}$$

Similarly, we perform the estimation for A_2 using the same ideas and the proof of Lemma 9.3. To do this, we let $A_v^1 := a(y)V_t + (va(x)+(1-v)a(y))c$ and $A_v^2 := a(y^*)V_t + (va(x)+(1-v)a(y^*))c$. Then we perform the calculation as follows

$$\begin{aligned}A_2 = -c\int_0^1 &\left[\mathrm{Tr}\left((A_v^1)^{-1}(a(y)-a(x))\right) + (x-y^*)^T (A_v^1)^{-1}(a(y)-a(x))(A_v^1)^{-1}(x-y^*)\right] q(A_v^1, x-y^*)\\ &- \left[\mathrm{Tr}\left((A_v^2)^{-1}(a(y^*)-a(x))\right) + (x-y^*)^T (A_v^2)^{-1}(a(y^*)-a(x))(A_v^2)^{-1}(x-y^*)\right] q(A_v^2, x-y^*)dv.\end{aligned}$$

As with A_1, now that we have obtained the integrability factor c. Next, we have to analyze the differences term by term in order to obtain $|y-y^*|$. That is,

$$\begin{aligned}\|(A_v^1)^{-1} - (A_v^2)^{-1}\|_F &\le C\|(A_v^1)^{-1}\|_F \|(A_v^2)^{-1}\|_F \|a(y)-a(y^*)\|_F (V_t+c)\\ &\le C(V_t+c)^{-1}\|a(y)-a(y^*)\|_F \le C(V_t+c)^{-1}|y-y^*|^k\\ &\le C(V_t+c)^{-(1-\frac{k-\beta}{2})}|y-y^*|^\beta.\end{aligned}$$

Furthermore, using Lemma 9.3, Lemma 9.2 and $\|a(y)-a(x)\|_F \le C(|x-y^*|^k + |y-y^*|^k)$, we obtain

$$\begin{aligned}A_2 \le& Cc\left\{(V_t+c)^{-(1-\frac{k-\beta}{2})}|y-y^*|^\beta(|x-y^*|^k + |y-y^*|^k) + (V_t+c)^{-1}|y-y^*|^k\right\} q(\overline{a}(V_t+c), x-y^*)\\ &+ Cc(V_t+c)^{-1}|x-y^*|^k|y-y^*|^k q(2\overline{a}(V_t+c), x-y^*)\\ \le& Cc|y-y^*|^\beta\left\{(V_t+c)^{-(1-k+\frac{\beta}{2})} + (V_t+c)^{-(1-\frac{k-\beta}{2})}\right\} q(4\overline{a}(V_t+c), x-y^*).\end{aligned} \tag{55}$$

The result then follows by putting all together all the estimates in (47) (49), (50)∼(55) and considering the smallest and biggest powers of V_t or V_t+c. □

Remark 9.7. *The choice of the condition $|y-y^*|^2 > V_t + c$ (or V_t) cannot be altered due to Lemma 9.5.*

The following lemma gives the Hölder continuity property in $V_t - V_s$ for the function $\widehat{\theta}_t(x,y)\widehat{p}_t^y(x,y)$. In the case where there is a non-trivial diffusion component, the parameter β in Lemma 9.8 again takes value in $(0,\alpha)$. This is to guarantee the finiteness of the β-moment of V.

Lemma 9.8. *The function $\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)$ is continuous in $t\in\mathbb{R}_+$. Furthermore, for any $\beta\in(0,\alpha)$ there exists constant C independent of x,y,t such that for $s\le t$*

$$|\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)-\widehat{\theta}_s(x,y)\widehat{p}^y_s(x,y)|\le C\mathbb{E}\big[|V_t-V_s|^{\beta}G^{H_\Lambda}_{V_t-V_s}(x-y)\big]$$

where Λ is a random variable independent of all other random variables and it has the law $\frac{1}{3}(\delta_0(dv)+\delta_1(dv)+\mathbb{1}_{(0,1)}dv)$. Furthermore we set for $v\in[0,1]$

$$\begin{aligned}\mathbb{V}_{\eta,H_v}(c,\bar{c})&:=\overline{a}(H_v+\eta\bar{c}+(1-\eta)c),\\ G^{H_v}_{V_t-V_s}(x)&:=\int_{\mathbb{R}_+\times\mathbb{R}}\mathbb{E}\big[q(8\mathbb{V}_{\eta,H_v}(c,\bar{c}),x)\sum_{i=1}^{2}\mathbb{W}^{\lambda_i,\delta_i,\gamma_i}_{\eta,V_t-V_s}(c)\big]\widehat{\mu}(dc)\times\delta_0(d\bar{c}).\end{aligned}$$

Here $(\lambda_1,\lambda_2):=(1-\frac{k}{2},1-\frac{k}{2}+\frac{\beta}{1-\beta})$, $(\delta_1,\delta_2):=(0,\frac{\beta}{1-\beta})$, $(\gamma_1,\gamma_2):=(\frac{1}{2},\frac{1}{2}+\frac{\beta}{1-\beta})$ and $H_v\equiv H_v(t,s):=vV_t+(1-v)V_s$.

Proof. Given $s,t\in(0,T]$, we assume without loss of generality that $s\le t$ and we compute from (8).

$$\begin{aligned}|\widehat{\theta}_t(x,y)\widehat{p}^y_t(x,y)-\widehat{\theta}_s(x,y)\widehat{p}^y_s(x,y)|&\le\mathbb{E}\Big[\big|b(x)^T\nabla_x q(a(y)V_t,x-y)-b(x)^T\nabla_x q(a(y)V_s,x-y)\big|\\&+\Big|\int_{\mathbb{R}_+}q(a(y)V_t+a(x)c,x-y)-q(a(y)(V_t+c),x-y)\mu(dc)\\&-\int_{\mathbb{R}_+}\left[q(a(y)V_s+a(x)c,x-y)-q(a(y)(V_s+c),x-y)\right]\mu(dc)\Big|\Big].\end{aligned}\tag{56}$$

The method is to break the computations on different sets like Lemma 9.6. Although many estimates are obtained in a similar fashion as in the proof of Lemma 9.6, we do them here for the sake of completeness.

The drift term in (56) on the set $|V_t-V_s|^{1-\beta}\le V_s$

We compute using (5) as follows; first, use the fact that b is bounded and the triangular inequality to obtain,

$$\begin{aligned}&|b(x)^T\nabla_x q(a(y)V_t,x-y)-b(x)^T\nabla_x q(a(y)V_s,x-y)|\\&\le C|V_t^{-1}a(y)^{-1}(x-y)q(a(y)V_t,x-y)-V_s^{-1}a(y)^{-1}(x-y)q(a(y)V_s,x-y)|\\&\le C\left(|V_t^{-1}-V_s^{-1}|q(a(y)V_t,x-y)+V_s^{-1}|q(a(y)V_t,x-y)-q(a(y)V_s,x-y)|\right)\underline{a}^{-1}|x-y|.\end{aligned}$$

Then to compute $|q(a(y)V_t,x-y)-q(a(y)V_s,x-y)|$ we make use of (5) and

$$\underline{a}(vV_t+(1-v)V_s)\le A_v:=a(y)(vV_t+(1-v)V_s)\le\overline{a}(vV_t+(1-v)V_s).\tag{57}$$

We follow the proof of Lemma 9.3 and we use a combination of (57) and Lemma 9.2 (ii), to obtain

$$\begin{aligned}&|q(a(y)V_t,x-y)-q(a(y)V_s,x-y)|\\&\le C|V_t-V_s|\int_0^1\left[\|A_v^{-1}\|_F\|a(y)\|_F+\|A_v^{-1}\|_2^2\|a(y)\|_2|x-y|^2\right]q(A_v,x-y)dv\\&\le C\int_0^1|V_t-V_s|(vV_t+(1-v)V_s)^{-1}q(2\overline{a}(vV_t+(1-v)V_s),x-y)dv.\end{aligned}\tag{58}$$

By using the above, we can write

$$\begin{aligned}&|b(x)^T\nabla_x q(a(y)V_t,x-y)-b(x)^T\nabla_x q(a(y)V_s,x-y)|\\&\le CV_s^{-\frac{3}{2}}|V_t-V_s|q(\overline{a}V_t,x-y)+C\int_0^1\frac{|x-y||V_t-V_s|}{V_s(vV_t+(1-v)V_s)}q(2\overline{a}(vV_t+(1-v)V_s),x-y)dv\\&\le CV_s^{-\frac{1}{2}}|V_t-V_s|^{\beta}q(\overline{a}V_t,x-y)+CV_s^{-\frac{1}{2}}|V_t-V_s|^{\beta}\int_0^1q(4\overline{a}(vV_t+(1-v)V_s),x-y)dv\end{aligned}\tag{59}$$

where in the last inequality we have used the fact $|V_t - V_s|^{1-\beta} \le V_s$, Lemma 9.2 (ii), the fact that v is positive and $V_t > V_s$.

The drift term of (56) on the set $|V_t - V_s|^{1-\beta} \ge V_s$

It follows directly from the triangular inequality on (5), the fact that $|V_t - V_s|^{1-\beta} \ge V_s$ and $V_t > V_s$.

$$\begin{aligned} &|b(x)^T \nabla_x q(a(y)V_t, x-y)| + |b(x)^T \nabla_x q(a(y)V_s, x-y)| \\ &\le C|V_t - V_s|^\beta V_s^{-(\frac{1}{2}+\frac{\beta}{1-\beta})}(q(\overline{a}V_t, x-y) + q(\overline{a}V_s, x-y)). \end{aligned} \tag{60}$$

The jump term of (56) on the set $c \in (1, \infty)$ and $|V_t - V_s|^{1-\beta} \le V_s + c$

The integrand in (56) is decomposed as follows,

$$\big|q(a(y)V_t + a(x)c, x-y) - q(a(y)V_s + a(x)c, x-y)\big| + \big|q(a(y)(V_t + c), x-y) - q(a(y)(V_s + c), x-y)\big|$$

We compute only the first term, since the method is similar for the second term. By application of Lemma 9.3 and Lemma 9.2 (ii) (similar arguments as in (58)), we obtain

$$\begin{aligned} &\big|q(a(y)V_t + a(x)c, x-y) - q(a(y)V_s + a(x)c, x-y)\big| \\ &\le C|V_t - V_s| \int_0^1 (vV_t + (1-v)V_s + c)^{-1} q(2\overline{a}(v(V_t - V_s) + V_s + c), x-y)dv \\ &\le C|V_t - V_s|^\beta \int_0^1 q(2\overline{a}(v(V_t - V_s) + V_s + c), x-y)dv, \end{aligned} \tag{61}$$

where the last inequality is obtained by using $V_t > V_s$ and then $|V_t - V_s|^{1-\beta} \le V_s + c$.

The jump term of (56) on the set $c \in (1, \infty)$ and $|V_t - V_s|^{1-\beta} \ge V_s + c$

It follows directly from triangular inequality and the fact that $|V_t - V_s|^{1-\beta} \ge V_s + c$,

$$\begin{aligned} &\big|q(a(y)V_t + a(x)c, x-y)| + |q(a(y)V_s + a(x)c, x-y)\big| + \big|q(a(y)(V_t + c), x-y)| + |q(a(y)(V_s + c), x-y)\big| \\ &\le C|V_t - V_s|^\beta (V_s + c)^{-\frac{\beta}{1-\beta}} \left(q(\overline{a}(V_t + c), x-y) + q(\overline{a}(V_s + c), x-y)\right). \end{aligned} \tag{62}$$

The jump part of (56) on the set $c \in (0, 1]$ and $|V_t - V_s|^{1-\beta} \ge V_s + c$

The integrand is decomposed in the following way

$$\big|q(a(y)V_t + a(x)c, x-y) - q(a(y)(V_t + c), x-y)\big| + \big|q(a(y)V_s + a(x)c, x-y) - q(a(y)(V_s + c), x-y)\big|$$

and again we compute only the first term. Due to Lemma 9.4, we have

$$\begin{aligned} &|q(a(y)V_t + a(x)c, x-y) - q(a(y)(V_t + c), x-y)| \le C\, c(V_t + c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_t + c), x-y) \\ &\le C\, c|V_t - V_s|^\beta (V_s + c)^{-(1-\frac{k}{2}+\frac{\beta}{1-\beta})} q(2\overline{a}(V_t + c), x-y). \end{aligned} \tag{63}$$

The jump term of (56) on the set $c \in (0, 1]$ and $|V_t - V_s|^{1-\beta} \le V_s + c$

Without loss of generality assume $s \le t$ and $v, v^* \in (0, 1)$, we set $A_{s,v} := a(y)V_s + (va(x) + (1-v)a(y))\, c$ and note that

$$\begin{aligned} &\underline{a}\,(V_s + c) \le A_{s,v} \le \overline{a}\,(V_s + c) \\ &\underline{a}\,(v^*V_t + (1-v^*)V_s + c) \le v^* A_{t,v} + (1-v^*)A_{s,v} \le \overline{a}\,(v^*V_t + (1-v^*)V_s + c) \end{aligned} \tag{64}$$

and $A_{s,v} \leq A_{t,v}$. By using (45), we can write

$$\begin{aligned}
&\big|(q(a(y)V_t + a(x)c, x-y) - q(a(y)V_t + a(y)c, x-y)) - (q(a(y)V_s + a(x)c, x-y) - q(a(y)(V_s+c), x-y))\big| \\
&\leq \int_0^1 \big|c\,\mathrm{Tr}\left(A_{s,v}^{-1}(a(x)-a(y))\right) q(A_{s,v}, x-y) - c\,\mathrm{Tr}\left(A_{t,v}^{-1}(a(x)-a(y))\right) q(A_{t,v}, x-y)\big| \\
&\quad + \big|c\left[(x-y)^T A_{s,v}^{-1}(a(x)-a(y))A_{s,v}^{-1}(x-y)\right] q(A_{s,v}, x-y) \\
&\quad - c\left[(x-y)^T A_{t,v}^{-1}(a(x)-a(y))A_{t,v}^{-1}(x-y)\right] q(A_{t,v}, x-y)\big| dv. \qquad (65)
\end{aligned}$$

Using the fact that Tr is a linear function, the first integrand can be bounded by,

$$\begin{aligned}
&\big|c\,\mathrm{Tr}\left(A_{s,v}^{-1}(a(x)-a(y))\right) q(A_{s,v}, x-y) - c\,\mathrm{Tr}\left(A_{t,v}^{-1}(a(x)-a(y))\right) q(A_{t,v}, x-y)\big| \\
&= |c\,\mathrm{Tr}\left(\left(A_{s,v}^{-1} - A_{t,v}^{-1}\right)(a(x)-a(y))\right) q(A_{s,v}, x-y) \\
&\quad + c\,\mathrm{Tr}\left(A_{t,v}^{-1}(a(x)-a(y))\right)(q(A_{s,v}, x-y) - q(A_{t,v}, x-y))| \qquad (66) \\
&\leq cC\|A_{s,v}^{-1} - A_{t,v}^{-1}\|_F |x-y|^k q(A_{s,v}, x-y) + cC\|A_{t,v}^{-1}\|_F |x-y|^k |q(A_{s,v}, x-y) - q(A_{t,v}, x-y)|.
\end{aligned}$$

Furthermore, noting that $\|A_{s,v}^{-1} - A_{t,v}^{-1}\|_F \leq \|A_{s,v}^{-1}\|_F \|A_{s,v} - A_{t,v}\|_F \|A_{t,v}^{-1}\|_F$,

$$\begin{aligned}
&\leq cC\|A_{s,v}^{-1}\|_F \|A_{t,v}^{-1}\|_F |V_t - V_s| |x-y|^k q(A_{s,v}, x-y) \\
&\quad + cC\|A_{t,v}^{-1}\|_F |x-y|^k |q(A_{s,v}, x-y) - q(A_{t,v}, x-y)| \\
&\leq cC|V_t - V_s|^\beta (V_s + c)^{-(1-\frac{k}{2})} q(2\overline{a}(V_s + c), x-y) \\
&\quad + cC\|A_{t,v}^{-1}\|_F |x-y|^k |q(A_{s,v}, x-y) - q(A_{t,v}, x-y)|. \qquad (67)
\end{aligned}$$

To compute the second term above, we use (64) and again applying the result of Lemma 9.3 in a similar fashion as in (58) we obtain

$$|q(A_{s,v}, x-y) - q(A_{t,v}, x-y)| \leq C \int_0^1 \frac{|V_t - V_s| q(2\overline{a}(v^* V_t + (1-v^*)V_s + c), x-y)}{(v^* V_t + (1-v^*)V_s + c)} dv^*. \qquad (68)$$

To finish, we notice that the second term in (67) can be bounded above by

$$cC|V_t - V_s|^\beta (V_s + c)^{-(1-\frac{k}{2})} \int_0^1 q(4\overline{a}(v^* V_t + (1-v^*)V_s + c), x-y) dv^*. \qquad (69)$$

Here we have used the fact that the process V is increasing. Using the triangular inequality, the second term of the integrand in (65) can be estimated by

$$\begin{aligned}
&\leq cC|x-y|^2 \|A_{s,v}^{-1}(a(x)-a(y))A_{s,v}^{-1} - A_{t,v}^{-1}(a(x)-a(y))A_{t,v}^{-1}\|_F q(A_{s,v}, x-y) \\
&\quad + cC|x-y|^{2+k} \|A_{t,v}^{-1}\|_F^2 |q(A_{s,v}, x-y) - q(A_{t,v}, x-y)|. \qquad (70)
\end{aligned}$$

To compute the first term above, we notice that

$$\begin{aligned}
&\|A_{s,v}^{-1}(a(x)-a(y))A_{s,v}^{-1} - A_{t,v}^{-1}(a(x)-a(y))A_{t,v}^{-1}\|_F \\
&= \|A_{s,v}^{-1}(a(x)-a(y))A_{s,v}^{-1} - A_{t,v}^{-1}(a(x)-a(y))A_{s,v}^{-1} + A_{t,v}^{-1}(a(x)-a(y))A_{s,v}^{-1} - A_{t,v}^{-1}(a(x)-a(y))A_{t,v}^{-1}\|_F
\end{aligned}$$

and by triangular inequality, uniform ellipticity and $V_s < V_t$, we obtain the next three inequalities respectively,

$$\begin{aligned}
&\leq C\|A_{s,v}^{-1} - A_{t,v}^{-1}\|_F |x-y|^k \left(\|A_{s,v}^{-1}\|_F + \|A_{t,v}^{-1}\|_F\right) \qquad (71) \\
&\leq C\|A_{s,v}^{-1} - A_{t,v}^{-1}\|_F |x-y|^k \left((V_s + c)^{-1} + (V_t + c)^{-1}\right) \\
&\leq C\|A_{s,v}^{-1} - A_{t,v}^{-1}\|_F |x-y|^k (V_s + c)^{-1}.
\end{aligned}$$

By combining the above computations, (70) can be bounded above by

$$
\begin{aligned}
&\le cC|x-y|^{2+k}\|A_{s,v}^{-1}\|_F\|A_{t,v}^{-1}\|_F|V_t-V_s|(V_s+c)^{-1}q(A_{s,v},x-y)\\
&\quad + cC|x-y|^{2+k}\|A_{t,v}^{-1}\|_F^2|q(A_{s,v},x-y)-q(A_{t,v},x-y)|\\
&\le cC|x-y|^{2+k}(V_s+c)^{-3}|V_t-V_s|q(\overline{a}(V_s+c),x-y)\\
&\quad + cC|x-y|^{2+k}\|A_{t,v}^{-1}\|_F^2|q(A_{s,v},x-y)-q(A_{t,v},x-y)|. \qquad (72)
\end{aligned}
$$

Using Lemma 9.2 (ii) and the fact that $|V_t-V_s|^{1-\beta}\le V_s+c$, the first term above is bounded by

$$
cC|V_t-V_s|^{\beta}(V_s+c)^{-(1-\frac{k}{2})}q(2\overline{a}(V_s+c),x-y), \qquad (73)
$$

and for the second term in (72), we first use (68) and the $|x-y|^2$ term (leaving $|x-y|^k$ to be treated separately) is treated using Lemma 9.2 (ii), then the $\|A_{t,v}^{-1}\|_F^2$ term is treated using (64) and the fact that for $v^*\in[0,1]$, $v^*A_{t,v}+(1-v^*)A_{s,v}\le A_{t,v}$.

The last inequality below follows from application of Lemma 9.2 (ii), the fact that $|V_t-V_s|^{1-\beta}\le V_s+c$ and $V_t>V_s$. That is,

$$
\begin{aligned}
&c|x-y|^{2+k}\|A_{t,v}^{-1}\|_F^2|q(A_{s,v},x-y)-q(A_{t,v},x-y)|\\
&\le cC|x-y|^k|V_t-V_s|\int_0^1(V_t+c)^{-2}q(4\overline{a}(v^*V_t+(1-v^*)V_s+c),x-y)dv^*\\
&\le cC|V_t-V_s|^{\beta}(V_s+c)^{-(1-\frac{k}{2})}\int_0^1 q(8\overline{a}(v^*V_t+(1-v^*)V_s+c),x-y)dv^*. \qquad (74)
\end{aligned}
$$

By combining (67), (69), (73), (74) and applying Lemma 9.2 (ii), we obtain the following upper bound for the jump part of (56) on the set $c\in(0,1]$, $|V_t-V_s|^{1-\beta}\le V_s+c$.

$$
cC|V_t-V_s|^{\beta}(V_s+c)^{-1+\frac{k}{2}}\left(q(8\overline{a}(V_s+c),x-y)+\int_0^1 q(8\overline{a}(v^*V_t+(1-v^*)V_s+c),x-y)dv^*\right) \quad (75)
$$

Finally, the claim of the lemma is obtained by combining (59)$\sim$(63) and (75) (and by applying Lemma 9.2 (i) whenever necessary to obtain the factor of 8 in the variance). ☐

9.3 Beta estimates and convergence results

Without lost of generality, we assume $t_0>1$,

Corollary 9.9. *For $\rho>-1$, $\gamma_1<1$ and $\gamma_2<1$, then for n large enough,*

$$
\int_0^{t_0}\cdots\int_0^{t_{n-1}}dt_n\ldots dt_1\,t_n^{\rho}(t_0-t_1)^{-\gamma_2}\prod_{i=1}^{n-1}(t_i-t_{i+1})^{-\gamma_1}\le t_0^{\rho}\Gamma(1+\rho)\frac{(t_0^{(1-\gamma_1)\vee(1-\gamma_2)}\Gamma(1-\gamma_1\vee\gamma_2))^n}{\Gamma(1+\rho+n(1-\gamma_1\wedge\gamma_2))}.
$$

Proof. The result essentially follows from the following observation. For $1-j>0$ and $1+\rho>0$ and setting $s=ut$,

$$
\int_0^t s^{\rho}(t-s)^{-j}ds=t^{\rho+1-j}B(1+\rho,1-j)=t^{\rho+1-j}\frac{\Gamma(1+\rho)\Gamma(1-j)}{\Gamma(1+\rho+1-j)}. \qquad (76)
$$

One can then conclude by iteration. That is,

$$
\begin{aligned}
&\int_0^{t_0}\cdots\int_0^{t_{n-1}}dt_n\ldots dt_1\,t_n^{\rho}(t_0-t_1)^{-\gamma_2}\prod_{i=1}^{n-1}(t_i-t_{i+1})^{-\gamma_1}\\
&=\frac{\Gamma(1+\rho)\Gamma(1-\gamma_1)^{n-1}}{\Gamma(1+\rho+(n-1)(1-\gamma_1))}\int_0^t dt_1\,t_1^{\rho+(n-1)(1-\gamma_1)}(t_0-t_1)^{-\gamma_2}\\
&=t_0^{\rho+(n-1)(1-\gamma_1)+(1-\gamma_2)}\frac{\Gamma(1+\rho)\Gamma(1-\gamma_1)^{n-1}\Gamma(1-\gamma_2)}{\Gamma(1+\rho+(n-1)(1-\gamma_1)+(1-\gamma_2))}.
\end{aligned}
$$

For $t_0 \geq 1$, we have $t_0^{(n-1)(1-\gamma_1)+(1-\gamma_2)} \leq t_0^{n(1-\gamma_1\wedge\gamma_2)}$ and if $t_0 < 1$, then $t_0^{(n-1)(1-\gamma_1)+(1-\gamma_2)} \leq t_0^{n(1-\gamma_1\vee\gamma_2)}$. While, for n large enough, Γ is an increasing function, therefore

$$\begin{aligned}\Gamma(1+\rho+(n-1)(1-\gamma_1)+(1-\gamma_2)) &\geq \Gamma(1+((1-\gamma_1)\vee(1-\gamma_2))n)\\ &= \Gamma(1+n(1-\gamma_1\wedge\gamma_2))\end{aligned}$$

which concludes the proof. □

9.4 Auxiliary identities and estimates

The proof of the following elementary lemma is straightforward.

Lemma 9.10.

(i) Given $p \in (0,1)$, then for any $c, x \in \mathbb{R}_+$ we have

$$(c+x)^p \leq 2^p(c^p + x^p). \tag{77}$$

(ii) For $b, d \in \mathbb{R}$ and $t \in [0,T]$, there exists some constant $C_{d,b,T}$, such that

$$t^d + t^b \leq C_{d,b,T}\ t^{d\wedge b}. \tag{78}$$

Lemma 9.11. *Given a positive definite matrix M, if $\underline{a}I \leq M \leq \overline{a}I$, then $\frac{1}{\overline{a}}I \leq M^{-1} \leq \frac{1}{\underline{a}}I$*

Lemma 9.12. *For $j > 0$, $c > 0$ and any $x \in \mathbb{R}$,*

$$(c+x)^{-j} = \Gamma(j)^{-1}\int_{\mathbb{R}_+} e^{-s(c+x)}s^{j-1}ds.$$

Proof. By using the time change $s\,(c+x) = k$ the result follows from the definition of the Gamma function. □

Lemma 9.13. *If m is a positive concave increasing function on $\mathbb{R}_+$, then for $\alpha \in (0,1)$, $j > 0$ and constant A independent of s and t, we have*

$$\int_{\mathbb{R}_+} s^{j-1}e^{-s^\alpha m(s)tA}ds \leq C_{\alpha,j,T}\ t^{-\frac{j}{a}}$$

where $C_{\alpha,j,T}$ is independent of t.

Proof. The function m is positive increasing and therefore bounded below on $(1,\infty)$, which gives

$$\int_{\mathbb{R}_+} s^{j-1}e^{-m(s)s^\alpha tA}ds \leq \int_{(0,1]} s^{j-1}ds + \int_{(1,\infty)} s^{j-1}e^{-s^\alpha tC}ds,$$

and to evaluate the second term above, one apply the change of variable $s^\alpha tC = u$, that is $s = \left(\frac{u}{tC}\right)^{1/\alpha}$ to obtain

$$\int_{(0,1]} s^{j-1}ds + \int_{(1,\infty)} s^{j-1}e^{-s^\alpha tC}ds \leq C_{\alpha,j}(1+t^{-\frac{j}{\alpha}}) \leq C_{\alpha,j,T}\ t^{-\frac{j}{\alpha}}.$$

□

9.5 Estimates for the moments of V

In this subsection, we compute the moment estimates of the subordinator V satisfying Hypotheses 3.2.

Lemma 9.14. *For any $j > 0$, we have*

$$\mathbb{E}[V_t^{-j}] \leq C_{\alpha,j} t^{-\frac{j}{\alpha}}$$

Proof. By using Lemma 9.12, we see that

$$\mathbb{E}[V_t^{-j}] = \Gamma(j)^{-1} \int_{\mathbb{R}_+} s^{j-1} \mathbb{E}[e^{-sV_t}] ds \leq C_{\alpha,j,T}\, t^{-\frac{j}{\alpha}}$$

where the second inequality follows from Hypotheses 3.2 (ii) and Lemma 9.13. □

In the following result which is proved by cases on the power parameters λ, δ and γ, for $x \in \mathbb{R}$, we use the notation $x_+ := x \vee 0$.

Lemma 9.15. *Suppose that V satisfies Hypotheses 3.2. Then for any $j > 0$, $s_n \geq 0$ and for any $\lambda > -\alpha$, $\delta > -\alpha$ and $\gamma > -\alpha$ there exist constants $C, M \in \mathbb{R}_+$ such that*

$$\begin{aligned}\int_{\mathbb{R}_+} \mathbb{E}\big[e^{-s_n(V_t - V_s)} \mathbb{W}^{\lambda,\delta,\gamma}_{\eta, V_t - V_s}(c) \,\big|\, \eta\,\big] \widehat{\mu}(dc) &\leq C e^{-M^{\alpha-1} s_n^\alpha m(2s_n) A(t-s)} \left((1-\eta)(t-s)^{-\frac{\widehat{\lambda}_+ \vee \delta_+}{\alpha}} + \eta (t-s)^{-\frac{\gamma_+}{\alpha}} \right) \\ &\leq C e^{-M^{\alpha-1} s_n^\alpha m(2s_n) A(t-s)} (t-s)^{-\frac{\widehat{\lambda}_+ \vee \delta_+ \vee \gamma_+}{\alpha}}.\end{aligned}$$

Here $\widehat{\lambda} = \alpha - w + \lambda$ and the restriction on w is $w \in (\alpha, 1 \wedge (\lambda + \alpha))$. If $\widehat{\lambda} > 0$ then $\frac{\widehat{\lambda}}{\alpha} \in (0,1)$.

Proof. The proof is obtained by estimating each term in (27) in each case. First, we deal with the case $\lambda > 0$, $\delta \geq 0$ and $\gamma \geq 0$. For the first term in (27), we apply Lemma 9.12 to see that, for any $\lambda > 0$,

$$\begin{aligned}\int_{(0,1]} \mathbb{E}\left[e^{-s_n V_t} (V_t + c)^{-\lambda}\right] c\mu(dc) &= \Gamma(\lambda)^{-1} \int_{(0,1]} \int_{\mathbb{R}_+} c e^{-sc} s^{\lambda-1} \mathbb{E}[e^{-(s_n+s)V_t}] ds\, \mu(dc) \\ &\leq \Gamma(\lambda)^{-1} \int_{\mathbb{R}_+} s^{\lambda-1} \mathbb{E}[e^{-(s_n+s)V_t}] \int_{(0,1]} c e^{-sc}\, \mu(dc)\, ds.\end{aligned}$$

Here Fubini's theorem can be used because the integrand is positive. Since $c \in [0,1]$, we have $c \leq c^w$, for any $w \in (\alpha, 1)$. Hypotheses 3.2 (ii) and (iii) can be applied to obtain

$$\Gamma(\lambda)^{-1} \int_{\mathbb{R}_+} s^{\lambda-1} \mathbb{E}[e^{-(s_n+s)V_t}] \int_{(0,1]} c e^{-sc}\, \mu(dc)\, ds \leq C_{\alpha,\lambda,T} \int_{\mathbb{R}_+} s^{\widehat{\lambda}-1} e^{-(s+s_n)^a m(s+s_n) tA} ds.$$

Notice that $m(2s)$ and $s^\alpha m(2s)$ are both positive concave increasing functions in s. Therefore by Jensen's inequality

$$(s + s_n)^a m(s + s_n) \geq M^{\alpha-1} \left(s^\alpha m(2s) + s_n^\alpha m(2s_n)\right),$$

and by using the above and Lemma 9.13 to further bound from above by

$$C_{\alpha,\lambda,T} e^{-M^{\alpha-1} s_n^a m(2s_n) tA} \int_{\mathbb{R}_+} s^{\widehat{\lambda}-1} e^{-M^{\alpha-1} s^a m(2s) tA} ds \leq C_{w,\alpha,\lambda,T} t^{-\frac{\widehat{\lambda}}{\alpha}}\, e^{-M^{\alpha-1} s_n^\alpha m(2s_n) tA}.$$

Similarly, for the second term in (27), if $\delta > 0$,

$$\begin{aligned}
\int_{\mathbb{R}_+} \mathbb{E}\left[e^{-s_n V_t}(V_t+c)^{-\delta}\right] \mathbb{1}_{(1,\infty)}(c)\mu(dc) &\leq \int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \mathbb{E}\left[e^{-s_n V_t} e^{-sc} s^{\delta-1} e^{-sV_t}\right] ds \mathbb{1}_{(1,\infty)}(c)\mu(dc) \\
&\leq \int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \mathbb{E}[e^{-(s_n+s)V_t}] s^{\delta-1} ds\, \mathbb{1}_{(1,\infty)}(c)\mu(dc) \\
&\leq \int_{\mathbb{R}_+} \mathbb{1}_{(1,\infty)}(c)\mu(dc) \int_{\mathbb{R}_+} e^{-(s_n+s)^\alpha m(s_n+s)tA} s^{\delta-1} ds \\
&\leq C_{\alpha,j}\, t^{\frac{\delta}{\alpha}}\, e^{-M^{\alpha-1} s_n^\alpha m(2s_n)tA}.
\end{aligned}$$

where in the last inequality, we have used Hypotheses 3.2 (i). In the case $\delta = 0$ the same inequality is satisfies and the proof follows directly from Hypotheses 3.2 (ii). For the third and last term one uses Lemma 9.14 and finally the conclusion follows from (78).

Let us now show how to deal with the case $\lambda \in (-\alpha, 0]$. The case for $\delta \in (-\alpha, 0)$ is dealt with similarly. By using (77) and Hölder inequality,

$$\begin{aligned}
\int_{\mathbb{R}_+} \mathbb{1}_{(0,1]}(c)\, \mathbb{E}\left[e^{s_n V_t}(V_t+c)^{-\lambda}\right] \widehat{\mu}(dc) &\leq C \int_{\mathbb{R}_+} \mathbb{1}_{(0,1]}(c)\big(c^{-\lambda}\mathbb{E}\left[e^{-s_n V_t}\right] + \mathbb{E}\big[V_t^{-\lambda} e^{-s_n V_t}\big]\big)\widehat{\mu}(dc) \\
&\leq C \int_{\mathbb{R}_+} \mathbb{1}_{(0,1]}(c)\big(c^{-\lambda}\mathbb{E}\left[e^{-s_n V_t}\right] + \mathbb{E}\big[V_T^{-p\lambda}\big]^{\frac{1}{p}} \mathbb{E}\left[e^{-s_n V_t q}\right]^{\frac{1}{q}}\big)\widehat{\mu}(dc)
\end{aligned}$$

where we chose $p > 1$ small enough so that $-p\lambda - \alpha < 0$. Using Hypotheses 3.2 (ii), (iii) and the fact that $\widehat{\mu}$ is a finite measure, one can further bound the above by

$$Ce^{-s_n^\alpha m(s_n)tA} + C_{T,\alpha,p} e^{-tA(qs_n)^\alpha m(qs_n)q^{-1}} \leq C_{T,\alpha,p} e^{-q^{\alpha-1} s_n^\alpha m(s_n)tA}$$

where in the last inequality, we use the fact $m(qs_n) \geq m(s_n)$ and that $q^{\alpha-1} < 1$. □

10 Final Comments

In this article, we have developed the parametrix technique, using subordinated Brownian motions. We have shown that this kind of technique allows on one hand the flexibility of choosing the Lévy measure for the subordinator. On the other hand, it remains to be seen if many properties that are currently being obtained for stable driven stochastic differential equations remain true also for stable-like driven stochastic differential equations.

The analysis is done in separate cases which is useful in order to obtain the results. We have also provided a first glimpse to a stochastic representation for densities of solutions of stochastic differential equations driven by stable-like processes. We believe that such representations maybe useful in order to understand the structure of the problem from a probabilistic point of view.

As examples, one may consider infinite dimensional analysis in the sense of the integration by parts formula. In this case, as an example, one may consider the problem of the integration by parts with respect to N_0. Other possible applications are: explicit bounds for the density (e.g. it is not difficult to obtain upper bounds for the density and its first order derivative in the diagonal case using Theorems 6.1 and 7.2), expansions of the solutions with respect to small parameters and/or with respect to other distances such as the ones given for stability with respect to the coefficients of the equation.

11 Glossary

In this section, for the readers convenience, we compile a list of notations used in the paper. This complements Section 2 on Notations and Definitions.

- B is the Brownian motion,
- V is an α-stable-like subordinator independent of B,
- μ is the Lévy measure of the subordinator V,
- $m(\cdot)$ is a positive concave increasing function.
- $\widehat{\mu}(dc) = c\mathbb{1}_{\{c\leq 1\}}\mu(dc) + \mathbb{1}_{\{c>1\}}\mu(dc)$,
- $\delta_y(dx)$ is the Dirac measure with unit mass at $y \in \mathbb{R}$.
- ψ denotes the Lévy exponent of Z,
- $q(M, x)$ is the Gaussian density with covariance matrix M and $x \in \mathbb{R}^d$,
- φ denotes a regular varying function,
- b is the drift coefficient of X,
- σ is the coefficient of associated with the driving Lévy process $Z := B_V$,
- ζ is the coefficient associated with the diffusion (if X is a jump diffusion process),
- $\mathcal{L}$ is the generator of process X, solution of the stochastic differential equation (1),
- $p_t(x, y)$ is the density of the process X,
- $\widehat{\mathcal{L}}^z$ is the generator of the parametrix or frozen process $\widehat{X}^z$,
- $\widehat{p}^y_t(x, y)$ is the density of the parametrix $\widehat{X}^z$,
- k is the Hölder exponential of $a := \sigma\sigma^T$,
- $\overline{a}$ and $\underline{a}$ are uniform upper and lower bounds for the eigenvalues of $a(\cdot)$,
- k' is the Hölder exponential of $e := \zeta\zeta^T$,
- η and $(\eta_i)_i$ are auxiliary Bernoulli($\frac{1}{2}$) used to link the drift part and the jump part,
- $\widehat{\eta}$ and $(\widehat{\eta}_i)_i$ are auxiliary Bernoulli($\frac{1}{2}$) used to link the small and large jumps,
- $(\tau_j)_{j\in\mathbb{N}}$ are the ordered jump times of an Poisson process.
- $\widehat{\lambda} := \alpha - w + \lambda$, where $w \in (\alpha, 1 \wedge (\alpha + \lambda))$ and $\lambda > 0$,
- $\Delta_{t_{i+1}} V := V_{t_i} - V_{t_{i+1}}$,
- $\Delta_{\tau_{n-i}} V := V_{\tau_{n-i+1}} - V_{\tau_{n-i}}$,
- $x_+ := x \vee 0$ for $x \in \mathbb{R}$.

The following are list of notations for objects that are in the main body of the paper. We omit objects which appears only in proofs.

$$\widehat{\theta}_t(z_2, z_1)\widehat{p}_t^{z_1}(z_2, z_1) := (\mathcal{L}^z - \widehat{\mathcal{L}}^{z_1})(\widehat{p}_t^{z_1}(\cdot, z_1))(z_2)\Big|_{z=z_2}$$
$$\widehat{S}_t^* f(x) := \int f(y)\widehat{\theta}_t(x, y)\widehat{p}_t^y(x, y)dy$$
$$Q_t^* f(x) := \int f(y)\widehat{p}_t^y(x, y)dy$$
$$I_t^n(y, x) := \int_0^t \dots \int_0^{t_{n-1}} dt_n \dots dt_1 \; K(t_n, \dots, t_1, t; x, y)$$
$$K(t_n, \dots, t_1, t; x, y) := \int \widehat{p}_{t_n}^{z_n}(x, z_n) \prod_{i=0}^{n-1} \widehat{\theta}_{t_i - t_{i+1}}(z_{i+1}, z_i)\widehat{p}_{t_i - t_{i+1}}^{z_i}(z_{i+1}, z_i)dz_{i+1},$$
$$D(t; x, y) := b(x)^T \nabla_x \widehat{p}_t^y(x, y)$$
$$J_{(0,1]}(t; x, y) := \int_{(0,1)} \mathbb{E}\big[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\big]\mu(dc)$$
$$J_{(1,\infty)}(t; x, y) := \int_{[1,\infty)} \mathbb{E}\big[q(a(y)V_t + a(x)c, x - y) - q(a(y)V_t + a(y)c, x - y)\big]\mu(dc).$$
$$\mathbb{W}_{\eta,V_t}^{\lambda,\delta,\gamma}(c) := (1 - \eta)\Big[\mathbb{1}_{(0,1]}(c)(c + V_t)^{-\lambda} + \mathbb{1}_{(1,\infty)}(c)(c + V_t)^{-\delta}\Big] + \eta V_t{}^{-\gamma}$$
$$\mathbb{V}_{\eta,V_t}(c, \bar{c}) := \bar{a}(V_t + \eta\bar{c} + (1 - \eta)c)$$
$$G_{V_t}(x) := \int_{\mathbb{R}_+ \times \mathbb{R}} \mathbb{E}\big[q(8\mathbb{V}_{\eta,V_t}(c, \bar{c}), x) \sum_{i=1}^{2} \mathbb{W}_{\eta,V_t}^{\lambda_i,\delta_i,\gamma_i}(c)\big]\widehat{\mu}(dc) \times \delta_0(d\bar{c})$$
$$G_{V_t}^{H_v}(x) := \int_{\mathbb{R}_+ \times \mathbb{R}} \mathbb{E}\big[q(8\mathbb{V}_{\eta,H_v}(c, \bar{c}), x) \sum_{i=1}^{2} \mathbb{W}_{\eta,V_t}^{\lambda_i,\delta_i,\gamma_i}(c)\big]\widehat{\mu}(dc) \times \delta_0(d\bar{c}).$$

References

[1] V. Bally and A. Kohatsu-Higa. A probabilistic interpretation of the parametrix method. *The Annals of Applied Probability*, 25(6):3095–3138, 12 2015.

[2] R. F. Bass, K. Burdzy, and Z.-Q. Chen. Stochastic differential equations driven by stable processes for which pathwise uniqueness fails. *Stochastic Processes and their Applications*, 111(1):1 – 15, 2004.

[3] R. F. Bass, K. Burdzy, and Z.-Q. Chen. Stochastic differential equations driven by stable processes for which pathwise uniqueness fails. *Stochastic Processes and their Applications*, 111(1):1 – 15, 2004.

[4] R. F. Bass and H. Tang. The martingale problem for a class of stable-like processes. *Stochastic Processes and their Applications*, 119(4):1144 – 1167, 2009.

[5] M. Caballero, J. Pardo, and J. Pérez. On the lamperti stable processes. *arXiv preprint arXiv:0802.0851*, 2008.

[6] A. Debussche and N. Fournier. Existence of densities for stable-like driven sdes with hlder continuous coefficients. *Journal of Functional Analysis*, 264(8):1757 – 1778, 2013.

[7] N. Fournier. On pathwise uniqueness for stochastic differential equations driven by stable lévy processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 49(1):138–159, 02 2013.

[8] L. Huang. Density estimates for sdes driven by tempered stable processes. *ArXiv e-prints*, Apr. 2015.

[9] J. Jacod. *Calcul stochastique et problèmes de martingales.* Springer-Verlag, 1979.

[10] V. Knopova and A. Kulik. Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. *ArXiv e-prints*, Dec. 2014.

[11] T. Komatsu. On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations of jump type. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 58(8):353–356, 1982.

[12] V. Konakov and S. Menozzi. Weak error for stable driven stochastic differential equations: Expansion of the densities. *Journal of Theoretical Probability*, 24(2):454–478, 2010.

[13] G. Pagès and H. Luschgy. Moment estimates for lévy processes. *Electronic Communication in Probability*, 13(41):422–434, 2008.

[14] H. Pragarauskas and P. A. Zanzotto. On one-dimensional stochastic differential equations driven by stable processes. *Lithuanian Mathematical Journal*, 40(3):277–295.

[15] K. Sato. *Lévy processes and infinitely divisible distributions.* Cambridge university press, 1999.

[16] I. J. Schoenberg. Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841, 1938.

[17] D. W. Stroock. Diffusion processes associated with lévy generators. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32(3):209–244.

[18] P. Sztonyk. Estimates of tempered stable densities. *Journal of Theoretical Probability*, 23(1):127–147, 2009.

[19] T. Watanabe. Asymptotic estimates of multi-dimensional stable densities and their applications. *Transactions of the American Mathematical Society*, 359(6):2851–2879, 2007.

[20] J. Wendel. Note on the gamma function. *The American Mathematical Monthly*, 55(9):563–564, 1948.

[21] D. R. E. Williams. Path-wise solutions of stochastic differential equations driven by lévy processes. *Revista Matemática Iberoamericana*, 17(2):295–329, 2001.

[22] P. A. Zanzotto. On stochastic differential equations driven by a cauchy process and other stable lévy motions. *The Annals of Probability*, 30(2):802–825, 04 2002.