DENSITY ESTIMATES FOR SDES DRIVEN BY TEMPERED STABLE PROCESSES

L. Huang[1]

Abstract. We study a class of stochastic differential equations driven by a possibly tempered Lévy process, under mild conditions on the coefficients. We prove the well-posedness of the associated martingale problem as well as the existence of the density of the solution. Two sided heat kernel estimates are given as well. Our approach is based on the Parametrix series expansion

1991 Mathematics Subject Classification. 60H15,60H30,47G20.

April 16, 2015.

1. Introduction

This Paper is devoted to the study of Stochastic Differential Equations (SDEs), driven by a class of possibly tempered Lévy processes. Specifically, we show the existence of the density, as well as some associated estimates, under mild assumptions on the coefficients. Weak uniqueness is also derived as a by-product of our approach. More precisely, we study equations with the dynamics:

$$X_t = x + \int_0^t F(u, X_u)du + \int_0^t \sigma(u, X_{u^-})dZ_u, \tag{1.1}$$

where $F : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is measurable bounded, Hölder continuous in space and elliptic, and $(Z_t)_{t\geq 0}$ is a symmetric Lévy process. We will denote by ν its Lévy measure and assume that it satisfies what we call *a tempered stable domination:*

$$\nu(A) \leq \int_{S^{d-1}} \int_0^{+\infty} \mathbf{1}_A(s\theta) \frac{\bar{q}(s)}{s^{1+\alpha}} ds\mu(d\theta), \tag{1.2}$$

where $\bar{q}$ is a non increasing function, and μ is a bounded measure on the sphere S^{d-1}. This is a relatively large class of Lévy processes, that contains in particular the stable processes.

In order to give density estimates on the solution of (1.1), it is first necessary to obtain density estimates for the driving process. Those estimates are clear when $(Z_t)_{t\geq 0}$ is a Brownian motion. However, the Lévy case is much more complicated due to the huge diversity in the class of Lévy processes. Let us mention the papers of Bogdan and Sztonyk [BS07] and Kaleta and Sztonyk [KS13] for density bounds concerning relatively general

Keywords and phrases: Tempered stable process, Parametrix, Density Bounds

[1] LPMA, Université Paris Diderot, 5 rue Thomas Mann 75013 Paris, huang@math.univ-paris-diderot.fr

Lévy processes. In the case of the symmetric stable processes, the Lévy measure writes:

$$\forall A \in \mathcal{B}(\mathbb{R}^d),\ \nu(A) = \int_0^{+\infty} \int_{S^{d-1}} \mathbf{1}_{\{s\theta \in A\}} C_{\alpha,d} \frac{ds}{s^{1+\alpha}} \mu(d\theta), \tag{1.3}$$

for some $\alpha \in (0,2)$. In the above, $C_{\alpha,d}$ is a positive constant that only depends on d and α (see Sato [Sat05] for its exact value), and S^{d-1} stands for the unit sphere of $\mathbb{R}^d$. Also, μ is a symmetric finite measure on the sphere called the spectral measure. When the spectral measure satisfies the non-degeneracy condition:

$$\exists C > 1, \quad \text{s.t.} \quad C^{-1}|p|^\alpha \leq \int_{S^{d-1}} |\langle p, \xi \rangle|^\alpha \mu(d\xi) \leq C|p|^\alpha, \tag{1.4}$$

the driving process Z_t has a density with respect to the Lebesgue measure. In the recent work of Watanabe [Wat07], the author studied asymptotics for the density of a general stable process, and highlighted the importance of the spectral measure on the decay of the densities. Specifically, let us denote by $p_Z(t,\cdot)$ the density of Z_t, and assume that there exists $\gamma > 0$ such that

$$\mu\left(B(\theta, r) \cap S^{d-1}\right) \leq C r^{\gamma - 1},\ \forall \theta \in S^{d-1},\ \forall r \leq 1/2,\ C \geq 1. \tag{1.5}$$

Observe that in the case where the spectral measure has a density with respect to the Lebesgue measure on S^{d-1}, this condition is satisfied with $\gamma = d$. For a general $\gamma \in [1, d]$ such that (1.5) holds, we have for all $x \in \mathbb{R}^d$, $t > 0$:

$$p_Z(t,x) \leq C \frac{t^{-d/\alpha}}{\left(1 + \frac{|x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}}. \tag{1.6}$$

Moreover, a similar lower bound is given for the points $x \in \mathbb{R}^d$ such that two sided estimate hold in (1.5) for $\theta = x/|x|$ (up to a modification of the threshold r). We refer to Theorem 1.1 in Watanabe [Wat07] for a thorough discussion. We would like to point out the difference between assumptions (1.5) and (1.4). The assumption (1.4) alone is enough to show the existence of the density of the driving stable process. However, it turns out that this sole assumption is not enough to get density estimates. Instead, we need to know what we refer to (with a slight abuse of language) as the "concentration properties" of the spectral measure to deduce density bounds. This concentration, reflected by the index γ in (1.5), directly impacts the decay of the density, as shown in the bound (1.6). Observe however that if the concentration index γ is too small with respect to the dimension, namely, $\alpha + \gamma \leq d$, the upper bound (1.6) is not homogeneous to a density, since its integral (over $\mathbb{R}^d$) is not defined. We refer to the work of Watanabe [Wat07] for a detailed presentation of these aspects.

A generalization of this result to the case where the Lévy measure does not factorize as in (1.3), but only satisfies the domination (1.2) has been obtained by Sztonyk [Szt10]. Two sided estimates of the form (1.6) are derived, up to additional multiplicative terms involving the temperation $\bar{q}$ in (1.2), with the same restrictions for the lower bound.

The temperation $\bar{q}$ can be seen as a way to impose finiteness of the moments of Z (see Theorem 25.3 in Sato [Sat05]), and intuitively, the integrability properties of $(Z_t)_{t\geq 0}$ should transfer to $(X_t)_{t\geq 0}$. However, giving a density estimate on the driving process and passing it to the density of the solution of the SDE is not always possible.

In the Brownian setting, if $\sigma\sigma^*$ is uniformly elliptic, bounded and Hölder continuous, and F is Borel bounded, it is known that two sided Gaussian estimates hold for the density of the SDE (1.1), see Friedman [Fri64]. We also mention the approach of Sheu [She91], that also gives estimates on the logarithmic gradient of the density. In the stable non degenerate case, i.e. when the coefficients F, σ are as above, and $\mu(d\xi)$ has a smooth strictly positive density with respect to the Lebesgue measure on the sphere, it can be derived from Kolokoltsov [Kol00], that the density $p(t,s,x,y)$ of (1.1) exists and satisfies the following two sided estimates. Fix $T > 0$, there exists $C > 1$ depending on T, the coefficients and on the non degeneracy conditions, such that for all $x, y \in \mathbb{R}^d$,

$0 < t \leq s \leq T$:

$$C^{-1}\frac{(s-t)^{-d/\alpha}}{\left(1+\frac{|x-y|}{(s-t)^{1/\alpha}}\right)^{\alpha+d}} \leq p(t,s,x,y) \leq C\frac{(s-t)^{-d/\alpha}}{\left(1+\frac{|x-y|}{(s-t)^{1/\alpha}}\right)^{\alpha+d}}. \tag{1.7}$$

This estimate is obtained using a continuity method: the parametrix technique. This approach is well suited to obtain density estimates for the solution of an SDE under mild assumptions on the coefficients, provided that good estimates are available on the driving process and on the so-called *Parametrix kernel*.

We refer to estimates of the form (1.7) as Aronson estimates: two sided bounds that reflect the nature of the noise of the system. In the Gaussian setting, the density of the solution has a Gaussian behavior, and in the stable case, the density of the solution has two sided bounds homogeneous to those of the driving stable process. This work aims at proving Aronson estimates when the driving process is a Lévy process satisfying a tempered domination (in the sense of (1.2)).

Finally, we mention that existence of the density can be investigated via Malliavin calculus. In the Brownian setting, we refer to the works of Kusuoka and Stroock [KS84, KS85, KS87], as well as Norris [Nor86]. The jump case is more difficult, and is treated by various authors. Let us mention Bichteler, Gravereaux and Jacod [BGJ87], and Picard [Pic96]. However, this technique requires regularity on the coefficients. In our approach, the convergence of the Parametrix series will give us the existence of the density as well as weak uniqueness, under relatively mild assumptions on the coefficients.

We will denote by **[H]** the following set of assumptions. These hypotheses ensure the existence of the density, and are those required by Sztonyk [Szt10] in order to have a two sided estimate for the driving process Z.

[H-1] $(Z_t)_{t\geq 0}$ is a symmetric Lévy process. We denote by ν its Lévy measure. There is a non increasing function $\bar{q} : \mathbb{R}_+ \to \mathbb{R}_+$, μ a bounded measure on S^{d-1}, and $\alpha \in (0,2)$, $\gamma \in [1,d]$ such that:

$$\nu(A) \leq \int_{S^{d-1}}\int_0^{+\infty} \mathbf{1}_A(s\theta)\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\theta), \quad \mu\left(B(\theta,r)\cap S^{d-1}\right) \leq Cr^{\gamma-1}, \tag{1.8}$$

with $\gamma + \alpha > d$. Moreover, we assume the following decay for $\bar{q}$: for all $s > 0$, there exists $C > 0$ such that:

$$\bar{q}(s) \leq C\bar{q}(2s), \tag{1.9}$$

$$\forall 0 \leq \delta \leq d, \ s^\delta \bar{q}(s) \leq C\bar{q}(Cs). \tag{1.10}$$

[H-2] For all $p \in \mathbb{R}^d$, there is $\Lambda > 1$ such that

$$\Lambda^{-1}|p|^\alpha \leq \int_{S^{d-1}} |\langle p,\xi\rangle|^\alpha \mu(d\xi) \leq \Lambda|p|^\alpha. \tag{1.11}$$

In particular, denoting φ_Z the Lévy-Kintchine exponent of $(Z_t)_{t\geq 0}$, there is $K > 0$ such that :

$$\mathbb{E}\left(e^{i\langle p,Z_t\rangle}\right) = e^{t\varphi_Z(p)} \leq e^{-Kt|p|^\alpha}. \tag{1.12}$$

[H-3] $F : \mathbb{R}^d \to \mathbb{R}^d$ is uniformly Lipschitz continuous in its second argument, and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded and uniformly η-Hölder continuous in its space.

[H-4] σ is uniformly elliptic. For all $x, \xi \in \mathbb{R}^d$, there exists $\kappa > 1$ such that:

$$\kappa^{-1}|\xi|^2 \leq \langle \xi, \sigma(t,x)\xi\rangle \leq \kappa|\xi|^2. \tag{1.13}$$

[H-5] For all $\forall A \in \mathcal{B}$, Borelian, we define the measure:

$$\nu_t(x,A) = \nu\big(\{z \in \mathbb{R}^d;\ \sigma(t,x)z \in A\}\big). \tag{1.14}$$

We assume these measures to be uniformly Hölder continuous with respect to the first parameter, that is, for all $\forall A \in \mathcal{B}$,

$$|\nu_t(x,A) - \nu_t(x',A)| \leq C|x-x'|^{\eta(\alpha\wedge 1)} \int_{S^{d-1}} \int_0^{+\infty} \mathbf{1}_A(s\theta) \frac{\bar{q}(s)}{s^{1+\alpha}} ds\mu(d\theta).$$

We point out that in the case where $\sigma \in \mathbb{R}$, or when the spherical part of ν is equivalent to the Lebesgue measure on S^{d-1} this is actually a consequence of the Hölder continuity of σ, and the domination **[H-1]**.

[H-LB] There is a non increasing function $\underline{q} : \mathbb{R}_+ \to \mathbb{R}_+$ and $A_{low} \subset \mathbb{R}^d$, such that for all $x \in A_{low}$,

$$\nu\Big(B(x,r)\Big) \geq Cr^\gamma \frac{\underline{q}(|x|)}{|x|^{\alpha+\gamma}}, \ \forall r > 0 \tag{1.15}$$

$$\nu\Big(B(0,r)^c\Big) \leq C\frac{1}{r^\alpha}, \ \forall r \in (0,1). \tag{1.16}$$

In the rest of this paper, we will assume that **[H-1]** to **[H-5]** is in force. Also, we say that **[H]** holds when **[H-1]** to **[H-5]** hold. We point out that **[H-LB]** is needed for the lower bound, and that the upper bound holds independently. Under **[H]**, we are able to prove the following.

Theorem 1.1 (**Weak Uniqueness**). *Assume **[H]** holds. The martingale problem associated with the generator $L_t(x,\nabla_x)$ of the equation* (1.1)*:*

$$L_t(x,\nabla_x)\varphi(x) = \langle F(t,x), \nabla_x\varphi(x)\rangle + \int_{\mathbb{R}^d} \varphi(x+\sigma(t,x)z) - \varphi(x) - \frac{\langle \sigma(t,x)z, \nabla_x\varphi(x)\rangle}{1+|z|^2}\nu(dz),$$

where $\varphi \in \mathcal{C}_0^2(\mathbb{R}^d,\mathbb{R})$ admits a unique solution. That is, for every $x \in \mathbb{R}^d$, there exists a unique probability measure $\mathbb{P}$ on $\Omega = \mathcal{D}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$ the space of càdlàg functions, such that for all $f \in \mathcal{C}_0^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$ (twice continuously differentiable functions with compact support), denoting by $(X_t)_{t\geq 0}$ the canonical process, we have:

$$\mathbb{P}(X_t = x) = 1 \quad \textit{and} \quad f(s,X_s) - \int_t^s (\partial_u + L_u(x,\nabla_x)f(u,X_u)du \quad \textit{is a } \mathbb{P}\textit{- martingale.}$$

Hence, weak uniqueness holds for (1.1).

Also, we have the following density estimate:

Theorem 1.2 (**Density Estimates**). *Under **[H]**, the unique weak solution of* (1.1) *has for every $t > 0$ a density with respect to the Lebesgue measure. Precisely, for all $t > 0$, and $x, y \in \mathbb{R}^d$,*

$$\mathbb{P}(X_s \in dy | X_t = x) = p(t,s,x,y)dy. \tag{1.17}$$

*Also, for a deterministic time horizon $T > 0$. there exists $C_1 \geq 1$ depending on T and **[H]** such that for all $0 \leq t < T$, $(x,y) \in \mathbb{R}^d$,*

$$p(t,T,x,y) \leq C_1 \frac{(T-t)^{-d/\alpha}}{\Big(1 + \frac{|y-\theta_{T,t}(x)|}{(T-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar{q}(|y-\theta_{T,t}(x)|). \tag{1.18}$$

*Moreover, assume **[H-LB]** holds. Then, if there exists $t_0 \in [t,T]$ such that $t_0 - t \geq C(T-t)$, for some positive constant C, and*

$$\forall s \in [t,t_0], \ B\big(\sigma(\theta_{s,t}(x))^{-1}(\theta_{t,T}(y)-x), C(T-t)^{1/\alpha}\big) \subset A_{low}, \tag{1.19}$$

there exists $C_2 > 1$ such that

$$C_2^{-1} \frac{(s-t)^{-d/\alpha}}{\Big(1 + \frac{|y-\theta_{s,t}(x)|}{(s-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \underline{q}(|y-\theta_{s,t}(x)|) \leq p(t,s,x,y). \tag{1.20}$$

Remark 1.1. The condition (1.19) appearing for the lower bound comes from the possibly unbounded feature of the deterministic flow associated with (1.1). Indeed, it states that if a neighborhood at the characteristic time scale of a suitable renormalization of the flow stays in the sets of non degeneracy for ν, then the lower bound holds. Let us mention that the lower should remain valid provided that (1.19) is satisfied for $s \in [\varepsilon_1 t, \varepsilon_2 t]$, $0 \leq \varepsilon_1 < \varepsilon_2 \leq 1$. In this case C_2 should depend on $\varepsilon_2 - \varepsilon_1$ as well. In other words, it should suffice to enter the non degeneracy region for a time interval of order t.

Remark 1.2 (On the constants)**.** We will often use the capital letter C to denote a strictly positive constant that can depend on T and the set of assumptions **[H]** and whose value of C may change from line to line. Similarly, in the temperation, we will often write $\bar{q}(|x|)$ where we actually mean $\bar{q}(C|x|)$. Finally, we will use the symbol $\asymp$ to denote the equivalence:

$$f \asymp g \Leftrightarrow \exists C > 1,\ C^{-1} f(x) \leq g(x) \leq C f(x).$$

Remark 1.3 (About the tempering function)**.** We point out that the additional assumption (1.9) concerning the tempering function $\bar{q}$ is not present in Sztonyk [Szt10]. Formally, it means that some integrability is required in order to perform our techniques (precisely to correct a bad concentration index on the parametrix kernel). Nevertheless, our approach allows to recover the existence and estimates on the density of (1.1) when $(Z_t)_{t\geq 0}$ is a rotationally invariant stable process. See the proof of Proposition 3.3 and Remark 3.2.

Remark 1.4 (Finite time horizon)**.** In the rest, we fix $t \leq T \leq 1$. However, the main results hold for any arbitrary, but finite time. Indeed, Theorem 1.1 is extended to any time by the Markov property, and Theorem 1.2 by convolution arguments (see Lemma 3.4).

The rest of this paper is organized as follows. In Section 2, we set up formally the Parametrix technique, and give the estimates permitting the convergence of the Parametrix series. Section 3 is a technical section and is divided in five subsections. First, in Subsection 3.1 we prove estimates on the Frozen Density. In Subsection 3.2, we investigate the Parametrix Kernel and its smoothing properties. In Subsection 3.3, we tackle the well-posedness of the Martingale Problem, using estimates provided by the two previous subsections. Next, in Subsection 3.4, we prove the estimates giving the convergence of the Parametrix Series. Finally, in Subsection 3.5, we investigate the lower bound (1.20).

2. The Parametrix Setting

We present here a continuity technique known as the Parametrix. Our approach is close to the one of Mc Kean and Singer [MKS67]. The strategy is to approximate the solution of (1.1) by the solution of a simpler equation and control the distance in some sense between the two processes. First of all, let us define the *proxy* we will use. Let $y \in \mathbb{R}^d$ be an arbitrary point. Let $\theta_{t,s}$ be the flow associated with the deterministic differential equation:

$$\frac{d}{dt}\theta_{t,s}(x) = F(t, \theta_{t,s}(x)),\ \theta_{s,s}(x) = x,\ 0 \leq t, s \leq T.$$

We will often refer to $\theta_{t,s}(y)$ as the transport of y by the deterministic part of (1.1). Recall T is the deterministic time horizon. Fix $y \in \mathbb{R}^d$, a terminal point and $t \in [0,T]$, and $x \in \mathbb{R}^d$ initial time and position, we define the *frozen process* $(\tilde{X}_s^{T,y})_{s\in[t,T]}$ as the solution of:

$$\tilde{X}_s^{T,y} = x + \int_t^s F(u, \theta_{u,t}(y))du + \int_t^s \sigma(u, \theta_{u,T}(y))dZ_u. \tag{2.21}$$

We point out that the transport of the terminal point in the drift part comes for the unbounded character of the drift coefficient. Also, in diffusion coefficient σ, the presence of the transport ensures the compatibility between the estimates on the frozen process and the parametrix kernel (see Propositions 3.1 and 3.3). We mention that our approach covers the case of a measurable and bounded drift. In that case, we take as frozen

process $\tilde{X}_t = x + \int_0^t \sigma(u,y)Z_u$. Also, we could restrict ourselves to functions F that are Hölder continuous. In this case, existence of the flow θ is given by the Cauchy Peano theorem. However, the lack of uniqueness poses the problem of the definition of $\theta_{t,s}$, so we decided to assume Lipschitz continuity instead. Anyhow, in the case where F is Hölder continuous, we expect some kind of regularization by the noise, as we recover weak uniqueness, see e.g. Bafico and Baldi [BB82], or Delarue and Flandoli [DF14] for recent developments.

It is clear from the definition of $(\tilde{X}_s^{T,y})_{s\in[t,T]}$ and assumptions **[H-1]** (domination of the Lévy measure) and **[H-2]** (non degeneracy of the spectral measure) that $\tilde{X}^{Tt,y}$ has a density with respect to the Lebesgue measure. We denote the latter:

$$\mathbb{P}(\tilde{X}_s^{T,y} \in dz | \tilde{X}_t = x) = \tilde{p}^{T,y}(t,s,x,z)dz,\ s \in (t,T].$$

To get an explicit representation for it, we proceed by a Fourier inversion. Observe first that the Fourier transform of $\tilde{X}_s^{T,y}$ actually writes:

$$\mathbb{E}(e^{i\langle p, \tilde{X}_s^{T,y}\rangle}) = e^{i\langle p, x + \int_t^s F(u,\theta_{u,T}(y))du\rangle} \exp\left(\int_t^s \varphi_Z(\sigma(u,\theta_{u,T}(y))^* p)\right),$$

where we denoted by $\sigma(u,\theta_{u,T}(y))^*$ the transpose of $\sigma(u,\theta_{u,T}(y))$, and φ_Z is the Lévy-Kintchine exponent of Z. Due to assumptions **[H-1]** and **[H-2]**, this exponent is integrable and the frozen density actually writes:

$$\begin{aligned}\tilde{p}^{T,y}(t,s,x,z) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dp e^{-i\langle p, z - x - \int_t^s F(u,\theta_{u,T}(y))du\rangle} \\ &\quad \times \exp\left(\int_t^s du \int_{\mathbb{R}^d} e^{i\langle p, \sigma(u,\theta_{u,T}(y))\xi\rangle} - 1 - i\frac{\langle p, \sigma(u,\theta_{u,T}(y))\xi\rangle}{1+|\xi|^2}\nu(d\xi)\right). \qquad (2.22)\end{aligned}$$

We will often denote $\tilde{p}(t,T,x,y) = \tilde{p}^{T,y}(t,T,x,y)$, namely, we omit the superscript T,y when the freezing parameters and the points where the density is considered are the same. Observe that in this case, we have

$$y - \int_t^T F(u,\theta_{u,t}(y))du - x = \theta_{t,T}(y) - x. \qquad (2.23)$$

The following proposition illustrates how the estimates on the frozen process transmit to the solution of the SDE.

Proposition 2.1. *Suppose that there exists a unique weak solution $(X_s)_{s\in[t,T]}$ to (1.1) which has a Feller semigroup $(P_{s,t})_{0\le s,t\le T}$. We have the following formal representation. For all $t > 0$, $(x,y) \in (\mathbb{R}^d)^2$ and any bounded measurable $f : \mathbb{R}^d \to \mathbb{R}$:*

$$P_{T,t}f(x) = \mathbb{E}[f(X_T)|X_t = x] = \int_{\mathbb{R}^d}\left(\sum_{r=0}^{+\infty}(\tilde{p}\otimes H^{(r)})(t,T,x,y)\right) f(y)dy, \qquad (2.24)$$

where H is the parametrix kernel:

$$\forall 0 \le t \le s \le T,\ (x,y) \in (\mathbb{R}^d)^2,\ H(t,T,x,y) := (L_t(x,\nabla_x) - L_t(\theta_{t,T}(y),\nabla_x))\tilde{p}^{T,y}(t,T,x,y). \qquad (2.25)$$

The notation $\otimes$ stands for the time space convolution:

$$f\otimes g(t,T,x,y) = \int_t^T du \int_{\mathbb{R}^d} dz f(t,u,x,z)g(u,T,z,y).$$

Besides, $\tilde{p}\otimes H^{(0)} = \tilde{p}$ and $\forall r \in \mathbb{N},\ H^{(r)}(t,T,x,y) = H^{(r-1)}\otimes H(t,T,x,y)$.

Furthermore, when the above representation can be justified, it yields the existence as well as a representation for the density of the initial process. Namely $\mathbb{P}[X_T \in dy | X_t = x] = p(t,T,x,y)dy$ *where :*

$$\forall t > 0,\ (x,y) \in (\mathbb{R}^d)^2,\ p(t,T,x,y) = \sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t,T,x,y). \tag{2.26}$$

Proof. We refer to Huang and Menozzi [HM14] for the proof of this statement. It relies on the Markov properties of the involved process and the Chapman-Kolmogorov equations. For the sake of completeness, we give here a simpler proof assuming first that the density of $(X_s)_{s\in[t,T]}$ exists and is smooth. Since $(\tilde{X}_s^{T,y})_{s\in[t,T]}$ is an approximation of $(X_s)_{s\in[t,T]}$, we can expect that the densities of these processes are close to each other. We quantify the distance with the help of the generators of the solution of (1.1) and (2.21) and the Kolmogorov equations. For $\xi \in \mathbb{R}^d$, we define the integro-differential operator $\forall \varphi \in \mathcal{C}_b^2(\mathbb{R}^d, \mathbb{R})$:

$$\begin{aligned} L_t(\xi, \nabla_x)\varphi(x) &= \langle F(t,\xi), \nabla_x \varphi(x) \rangle \\ &\quad + \int_{\mathbb{R}^d} \varphi(x + \sigma(t,\xi)z) - \varphi(x) - \langle \nabla_x \varphi(x), \sigma(t,\xi)z \rangle \mathbf{1}_{\{|z|\le 1\}} \nu(dz). \end{aligned} \tag{2.27}$$

Observe that when $\xi = x$ the initial position, the operator $L_t(x, \nabla_x)$ is the generator of (1.1), whereas for $\xi = \theta_{t,T}(y)$, the operator $L_t(\theta_{t,T}(y), \nabla_x)$ is the generator of $(\tilde{X}_s^{T,y})_{s\in[t,T]}$. Also, we emphasize with the notations ∇_x the variable on which the operator acts.

Let us denote by $p(t,s,x,y)$ the density of $(X_s)_{s\ge[t,T]}$:

$$\mathbb{P}(X_s \in dy | X_t = x) = p(t,s,x,y)dy.$$

Then, $p(t,s,x,y)$ satisfies the Forward Chapman-Kolmogorov equations:

$$\begin{aligned} &\partial_s p(t,s,x,z) = L_s(x, \nabla_z)^* p(t,s,x,z), \\ &\text{for all } s > t,\ (x,z) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lim_{s\to t} p(t,s,x,\cdot) = \delta_x(\cdot)\ . \end{aligned} \tag{2.28}$$

On the other hand, we have the Backward Chapman-Kolmogorov equations for the frozen density as well:

$$\begin{aligned} &\partial_t \tilde{p}(t,s,x,z) = -L_t(\theta_{t,T}(y), \nabla_x)\tilde{p}(t,s,x,z), \\ &\text{for all } s > t,\ (x,z) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lim_{s\to t} \tilde{p}(t,s,\cdot,z) = \delta_z(\cdot)\ . \end{aligned} \tag{2.29}$$

We deduce from the Dirac convergences (2.28) and (2.29) that:

$$(p - \tilde{p})(t,T,x,y) = \int_t^T du\ \partial_u \left(\int_{\mathbb{R}^d} p(t,u,x,z)\tilde{p}(u,T,z,y)dz \right).$$

Differentiating formally under the integral, leads to:

$$(p - \tilde{p})(t,T,x,y) = \int_t^T du \left(\int_{\mathbb{R}^d} \partial_u p(t,u,x,z)\tilde{p}(u,T,z,y) + p(t,u,x,z)\partial_u \tilde{p}(u,T,z,y)dz \right).$$

Then, using the Kolmogorov Backward equation (2.29) for $\tilde{p}$ and the Forward equation (2.28) for p, we get:

$$\begin{aligned} (p - \tilde{p})(t,T,x,y) = \int_t^s du \int_{\mathbb{R}^d} dz \Big(& L_u(x, \nabla_z)^* p(t,u,x,z)\tilde{p}(u,T,z,y) \\ & - p(t,u,x,z) L_u(\theta_{u,T}(y), \nabla_z)\tilde{p}(u,T,z,y) \Big). \end{aligned}$$

Passing to the adjoint in the last equality yields:

$$\begin{aligned}(p-\tilde{p})(t,T,x,y) &= \int_t^T \int_{\mathbb{R}^d} p(t,u,x,z)\Big(L_u(x,\nabla_z) - L_u(\theta_{u,T}(y),\nabla_z)\Big)\tilde{p}(u,T,z,y)dz \\ &= p \otimes H(t,T,x,y),\end{aligned}$$

with the notation $\otimes$ for the time space convolution:

$$\varphi \otimes \psi(t,T,x,y) = \int_t^T du \int_{\mathbb{R}^d} dz\ \varphi(t,u,x,z)\psi(u,T,z,y),$$

and the Parametrix Kernel:

$$\forall 0 \le t < s,\ (x,y) \in (\mathbb{R}^d)^2,\ H(t,T,x,y) = \Big(L_t(x,\nabla_x) - L_t(\theta_{t,T}(y),\nabla_x)\Big)\tilde{p}(t,T,x,y). \tag{2.30}$$

Thus, we can iterate this identity to get the following formal representation for the density:

$$\forall 0 \le t < T,\ (x,y) \in (\mathbb{R}^d)^2,\ p(t,T,x,y) = \sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t,T,x,y), \tag{2.31}$$

with $\tilde{p} \otimes H^{(0)} = \tilde{p}$ and $\forall r \in \mathbb{N},\ H^{(r)}(t,T,x,y) = H^{(r-1)} \otimes H(t,T,x,y)$.

□

Remark 2.1. The proof relies on the Markov properties of the processes involved, as well as the Chapman-Kolmogorov equations. In the Brownian setting, the series (2.26) is first obtained for the SDE (1.1) with regularized coefficients. Indeed, in that setting, the Hörmander theorem gives existence and smoothness for the density (see Norris [Nor86]). The next step consists in proving estimates independent of the regularization parameter. Finally, the weak uniqueness, obtained through the well posedness of the martingale problem, as exposed in [Men11], allows to pass to the limit and identify the sum of the series (2.26) as the density of the initial equation (1.1). However, in the Lévy setting, there are no general (Hörmander) theorem to ensure the existence of the density even with regular coefficients. Nevertheless, some results exist in the literature concerning existence of the density in the jump case, let us mention Bichteller Gravereaux Jacod [BGJ87], Ishikawa and Kunita [IK06] in the non degenerate case, Cass [Cas09], which can be seen as the most complete extension to the jump case of the Hörmander theorem, but requires some integrability conditions, or the works of Zhang [Zha14] in the weak Hörmander degenerate stable driven framework. Anyhow, in our current operator-based approach, we do not proceed in that manner. Instead, we provide a representation for the semigroup associated with (1.1), and when the series (2.26) converges, it yields a representation of the density of (1.1).

The existence of the density for the solution of (1.1) will follow from the convergence of the parametrix series. In the following, we will denote

$$\bar{p}(t,T,x,y) = \frac{(T-t)^{-d/\alpha}}{\Big(1 + \frac{|y-\theta_{T,t}(x)|}{(T-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar{q}(|y-\theta_{T,t}(x)|). \tag{2.32}$$

This is the upper bound on the Frozen density under **[H]** derived by Sztonyk [Szt10], adapted to our possible unbounded drift case. We prove that this upper bound holds for the frozen density in Section 3.

The following lemma proves the convergence of the series (2.26).

Lemma 2.2 (**Control of the iterated kernels**). *There exist $C_{2.2} > 0$, $\omega \in (0,1]$ s.t. for all $t \in [0,T]$, $(x,y) \in (\mathbb{R}^d)^2$:*

$$|\tilde{p} \otimes H(t,T,x,y)| \leq C_{2.2}\Big((T-t)^{\omega}\bar{p}(t,T,x,y) + \rho(t,T,x,y)\Big), \tag{2.33}$$
$$|\rho \otimes H(t,T,x,y)| \leq C_{2.2}(T-t)^{\omega}\bar{p}(t,T,x,y), \tag{2.34}$$

where we denoted $\rho(t,T,x,y) = \delta \wedge |x - \theta_{t,T}(y)|^{\eta(\alpha\wedge 1)}\bar{p}(t,T,x,y)$. Now for all $k \geq 1$,

$$|\tilde{p} \otimes H^{(2k)}(t,T,x,y)| \leq (4C_{2.2})^{2k}(T-t)^{k\omega}\Big((T-t)^{k\omega}\bar{p}(t,T,x,y) + (\bar{p}+\rho)(t,T,x,y)\Big), \tag{2.35}$$
$$|\tilde{p} \otimes H^{(2k+1)}(t,T,x,y)| \leq (4C_{2.2})^{2k+1}(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}\bar{p} + (T-t)^{\omega}(\bar{p}+\rho) + \rho\Big)(t,T,x,y). \tag{2.36}$$

The above controls allow to derive under the sole assumption [**H**] the convergence of the Parametrix Series (thus, existence of the density for the solution of (1.1)), and the upper bound (1.18) for the sum of the parametrix series (2.26) in small time. To extend the result to any arbitrary (but finite) time, we use the semi-group property satisfied by $\bar{p}(t,T,x,y)$ (see Lemma 3.4). We point out that this procedure yields exponential dependencies in time in the constants. It is possible however to obtain the convergence of the series (2.26) for any time from Lemma 2.2, by estimating separately the two integrals (in time and space) in the time space convolution $\otimes$. This more technical procedure, yielding better, yet still exponentially explosive constants, is developed in Kolokolstov [Kol00].

Proof. We prove the important estimates (2.33) and (2.34) in Section 3, as the proof is technical and relies on sharp estimates on the Frozen Density and on the Parametrix Kernel (see Lemmas 3.6 and 3.7). Assuming estimates (2.33) and (2.34), we prove estimates (2.35) and (2.36) by induction. The bounds may not be very precise, as we will sometimes bound $t^{k\omega} \leq 1$, but they are sufficient to prove the convergence of the Parametrix series (2.26).

Initialization:
Since $t^{\omega}(\bar{p}+\rho) \geq 0$, we clearly have:

$$|\tilde{p} \otimes H(t,T,x,y)| \leq C_{2.2}\Big((T-t)^{\omega}\bar{p} + \rho + (T-t)^{\omega}(\bar{p}+\rho)\Big)(t,T,x,y).$$

Now, using equations (2.33) and (2.34), we have:

$$\begin{aligned}
|\tilde{p} \otimes H^{(2)}(t,T,x,y)| &\leq C_{2.2}\Big((T-t)^{\omega}|\bar{p}\otimes H| + |\rho \otimes H|\Big)(t,T,x,y) \\
&\leq C_{2.2}\Big(C_{2.2}(T-t)^{2\omega}\bar{p} + C_{2.2}(T-t)^{\omega}\rho + C_{2.2}(T-t)^{\omega}\bar{p}\Big)(t,T,x,y) \\
&\leq (2C_{2.2})^2(T-t)^{\omega}\Big((T-t)^{\omega}\bar{p} + (\bar{p}+\rho)\Big)(t,T,x,y).
\end{aligned}$$

Induction:
Suppose that the estimate for $2k$ holds. Let us prove the estimate for $2k+1$.

$$\begin{aligned}
|\tilde{p} \otimes H^{(2k+1)}|(t,T,x,y) &\leq (4C_{2.2})^{2k}(T-t)^{k\omega}\Big((T-t)^{k\omega}|\bar{p}\otimes H|(t,T,x,y) + |(\bar{p}+\rho)\otimes H|(t,T,x,y)\Big) \\
&\leq (4C_{2.2})^{2k}(T-t)^{k\omega}\Big(C_{2.2}(T-t)^{k\omega}((T-t)^{\omega}\bar{p}+\rho)(t,T,x,y) \\
&\quad + C_{2.2}((T-t)^{\omega}\bar{p}+\rho)(t,T,x,y) + C_{2.2}(T-t)^{\omega}\bar{p}_{\alpha}(t,T,x,y)\Big).
\end{aligned}$$

Recalling that $(T-t)\leq 1$, we have $(T-t)^{k\omega}\rho\leq(T-t)^{\omega}\rho$. Thus:

$$\begin{aligned}|\tilde{p}\otimes H^{(2k+1)}|(t,T,x,y) &\leq (4C_{2.2})^{2k}(T-t)^{k\omega}\Big(C_{2.2}(T-t)^{(k+1)\omega}\bar{p}+2C_{2.2}(T-t)^{\omega}(\bar{p}+\rho)+C_{2.2}\rho)\Big)(t,T,x,y)\\ &\leq (4C_{2.2})^{2k}(2C_{2.2})(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}\bar{p}+(T-t)^{\omega}(\bar{p}+\rho)+\rho)\Big)(t,T,x,y),\end{aligned}$$

which gives the announced estimate.
Suppose now that the estimate for $2k+1$ holds. Let us prove the estimate for $2k+2$.

$$\begin{aligned}|\tilde{p}\otimes H^{(2k+2)}(t,T,x,y)| &\leq (4C_{2.2})^{2k+1}(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}|\bar{p}\otimes H|\\ &\quad+(T-t)^{\omega}|(\bar{p}+\rho)\otimes H|+|\rho\otimes H|\Big)(t,T,x,y)\\ &\leq (4C_{2.2})^{2k+1}(T-t)^{k\omega}\Big(C_{2.2}(T-t)^{(k+1)\omega}[(T-t)^{\omega}\bar{p}+\rho]\\ &\quad+C_{2.2}(T-t)^{\omega}[\{(T-t)^{\omega}\bar{p}+\rho\}+C_{2.2}(T-t)^{\omega}\bar{p}]+C_{2.2}(T-t)^{\omega}\bar{p}\Big)(t,T,x,y)\\ &\leq (4C_{2.2})^{2k+2}(T-t)^{(k+1)\omega}\Big((T-t)^{(k+1)\omega}\bar{p}+(\bar{p}+\rho)\Big)(t,T,x,y),\end{aligned}$$

where to get to the last equation, we used the fact that since $t\in[0,T]$ with T small enough, we have $(T-t)^{\omega}\bar{p}\leq\bar{p}$, and $(T-t)^{k\omega}\rho\leq\rho$.

□

3. Proof of the estimates.

In order for the Parametrix technique to be successful, we must obtain some sharp estimates on the quantities involved in the Parametrix expansion (2.26). This is usually done in two parts, first, we give two sided estimates on the density of the frozen process, as well as a similar upper bound on the Parametrix kernel H, up to a time singularity. Then, we prove that those bounds yield a smoothing effect in time for the time space convolution $\tilde{p}\otimes H$ appearing in (2.26).

3.1. **Estimates on the Frozen Density**

We first give the estimates on the frozen density.

Proposition 3.1. *Assume **[H]** is in force. There exists $C>1$ s.t. for all $t\in[0,T]$, $(x,y)\in(\mathbb{R}^d)^2$:*

$$\tilde{p}^{T,y}(t,s,x,z)\leq C\frac{(s-t)^{-d/\alpha}}{\Big(1+\frac{|z-x-\int_t^s F(u,\theta_{u,T}(y))du|}{(s-t)^{1/\alpha}}\Big)^{\alpha+\gamma}}\bar{q}\left(C^{-1}\left|z-x-\int_t^s F(u,\theta_{u,T}(y))du\right|\right). \tag{3.37}$$

*Moreover, when **[H-LB]** holds, for all $z-x-\int_t^s F(u,\theta_{u,T}(y))du\in A_{low}$, the lower bound holds:*

$$\frac{C^{-1}(s-t)^{-d/\alpha}}{\Big(1+\frac{|z-x-\int_t^s F(u,\theta_{u,T}(y))du|}{(s-t)^{1/\alpha}}\Big)^{\alpha+\gamma}}\underline{q}\left(C\left|z-x-\int_t^s F(u,\theta_{u,T}(y))du\right|\right)\leq\tilde{p}^{T,y}(t,s,x,z). \tag{3.38}$$

Proof. We prove these estimates in the lines of Sztonyk [Szt10]. The idea consist in splitting large jumps and small jumps at the characteristic time scale and exploit the Lévy-Itô decomposition. Fix $t,T\in\mathbb{R}_+$ and $y\in\mathbb{R}^d$. Observe that since the drift part in the frozen process is deterministic, it suffices to prove the estimates or:

$$\Lambda_s=\int_t^s\sigma(u,\theta_{u,T}(y))dZ_u.$$

We point out that $\Lambda_s = \Lambda_s(t,T,y)$, where t,T,y are fixed. The Fourier transform of Λ_s writes:

$$\mathbb{E}(e^{i\langle p,\Lambda_s\rangle}) = \exp\left(\int_t^s du \int_{\mathbb{R}^d} e^{i\langle p,\sigma(u,\theta_{u,T}(y))\xi\rangle} - 1 - i\langle p,\sigma(u,\theta_{u,T}(y))\xi\rangle \mathbf{1}_{\{|\xi|\leq t^{1/\alpha}\}}\nu(d\xi)\right).$$

Changing variables in the time integral to $v\in[0,1]$ and setting $\sigma_v = \sigma((s-t)v+t,\theta_{(s-t)v+t,T}(y))$, we obtain:

$$\mathbb{E}(e^{i\langle p,\Lambda_s\rangle}) = \exp\left((s-t)\int_0^1 du \int_{\mathbb{R}^d} e^{i\langle p,\sigma_v\xi\rangle} - 1 - i\langle p,\sigma_v\xi\rangle \mathbf{1}_{\{|\xi|\leq t^{1/\alpha}\}}\nu(d\xi)\right).$$

Now, defining $\nu_{\mathcal{S}}$ to be the image measure of $dv\nu(d\xi)$ by the application $(v,\xi)\mapsto\sigma_v\xi$, we obtain:

$$\mathbb{E}(e^{i\langle p,\Lambda_s\rangle}) = \exp\left((s-t)\int_{\mathbb{R}^d} e^{i\langle p,\eta\rangle} - 1 - i\langle p,\eta\rangle \mathbf{1}_{\{|\eta|\leq t^{1/\alpha}\}}\nu_{\mathcal{S}}(d\xi)\right),$$

which is the Fourier transform of some Lévy process $(\mathcal{S}_u)_{u\geq 0}$, with Lévy measure $\nu_{\mathcal{S}}$, at time $s-t$. In other words, the marginals of $(\Lambda_s)_{s\in[t,T]}$ corresponds to the marginals of $(\mathcal{S}_{s-t})_{s\in[t,T]}$:

$$\forall s\in[t,T],\ \Lambda_s \overset{\text{(law)}}{=} \mathcal{S}_{s-t}. \tag{3.39}$$

The idea is now to work with the process $(\mathcal{S}_u)_{u\geq 0}$, and prove that it satisfies the assumptions **[H]**. Specifically, we prove that **[H-1]** and **[H-2]** holds for $\nu_{\mathcal{S}}$, and that when **[H-LB]** holds for ν, it holds as well for $\nu_{\mathcal{S}}$.

Let $A\in\mathcal{B}(\mathbb{R}^d)$. By definition of $\nu_{\mathcal{S}}$, we have:

$$\nu_{\mathcal{S}}(A) = \int_0^1\int_{\mathbb{R}^d} \mathbf{1}_{\{\sigma_v\xi\in A\}}\nu(d\xi)dv.$$

From the tempered stable domination, we deduce:

$$\nu_{\mathcal{S}}(A) = \int_0^1\int_{S^{d-1}}\int_0^{+\infty} \mathbf{1}_{\{\sigma_v s\varsigma\in A\}}\frac{\bar{q}(s)ds}{s^{1+\alpha}}\mu(d\varsigma)dv.$$

For fixed v,ς, we change the variables in the integral in ds to $\rho = s|\sigma_v\varsigma|$. Observe that from the uniform ellipticity of σ, we have $\bar{q}\left(\frac{\rho}{|\sigma_v\varsigma|}\right)\leq \bar{C}q(\rho)$. It yields:

$$\nu_{\mathcal{S}}(A) \leq \int_0^1\int_{S^{d-1}}\int_0^{+\infty} \mathbf{1}_{\{\rho\frac{\sigma_v\varsigma}{|\sigma_v\varsigma|}\in A\}}\frac{\bar{q}(\rho)d\rho}{\rho^{1+\alpha}}|\sigma_v\varsigma|^\alpha\mu(d\varsigma)dv.$$

We now define $\mu_{\mathcal{S}}(d\varsigma)$ to be the image measure of $|\sigma_v\varsigma|^\alpha\mu(d\varsigma)dv$ (measure on $[0,1]\times S^{d-1}$) by the application $(v,\varsigma)\mapsto\frac{\sigma_v\varsigma}{|\sigma_v\varsigma|}$. We thus obtain:

$$\nu_{\mathcal{S}}(A) \leq \int_{S^{d-1}}\int_0^{+\infty} \mathbf{1}_{\{\rho\zeta\in A\}}\frac{\bar{q}(\rho)d\rho}{\rho^{1+\alpha}}\mu_{\mathcal{S}}(d\zeta).$$

Consequently, **[H-1]** holds for $\nu_{\mathcal{S}}$. Observe that by construction, we have that **[H-2]** holds for $\mu_{\mathcal{S}}$. Therefore, from Stzonyk [Szt10], denoting $p_{\mathcal{S}}(u,\cdot)$ the density of $\mathcal{S}_u$, the following upper bound holds:

$$p_{\mathcal{S}}(u,w) \leq \frac{u^{-d/\alpha}}{\left(1+\frac{|w|}{u^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(|w|).$$

We deduce the estimate for Λ_s, and the upper bound on $\tilde{p}^{T,y}(t,s,x,z)$ then follows.

To get a lower bound on $\tilde{p}^{T,y}(t,s,x,z)$, we investigate a lower bound for $p_{\mathcal{S}}(u,\cdot)$. To that aim, we prove that when **[H-LB]** holds for ν, it does for $\mu_{\mathcal{S}}$. Specifically, assume **[H-LB]** holds for ν. By definition of $\mu_{\mathcal{S}}$, for all $x\in\mathbb{R}^d$, $r>0$, we have:

$$\nu_{\mathcal{S}}\Big(B(x,r)\Big)=\int_0^1\int_{\mathbb{R}^d}\mathbf{1}_{\{|x-\sigma_v\varsigma|\le r\}}dv\nu(d\varsigma).$$

Now, from uniform ellipticity of σ,

$$\{|x-\sigma_v\varsigma|\le r\}\supset\{|\sigma_v^{-1}x-\varsigma|\le Cr\}.$$

Consequently, we have:

$$\nu_{\mathcal{S}}\Big(B(x,r)\Big)=\int_0^1\nu\Big(\sigma_v^{-1}x,Cr\Big)dv\ge\int_0^1 Cr^\gamma\frac{\underline{q}(|\sigma_v^{-1}x|)}{|\sigma_v^{-1}x|^{\alpha+\gamma}}dv\ge r^\gamma\frac{\underline{q}(|x|)}{|x|^{\alpha+\gamma}},$$

where to get the last inequality, we exploited the uniform ellipticity of σ. Besides, for $r\in(0,1)$, we write using the ellipticity of σ:

$$\nu_{\mathcal{S}}\Big(B(0,r)^c\Big)=\int_0^1\int_{\mathbb{R}^d}\mathbf{1}_{\{|\sigma_v\varsigma|\ge r\}}dv\nu(d\varsigma)\le\int_0^1\int_{\mathbb{R}^d}\mathbf{1}_{\{|\varsigma|\ge Cr\}}dv\nu(d\varsigma)\le C\frac{1}{r^\alpha}.$$

Thus, we recovered **[H-LB]** for $\nu_{\mathcal{S}}$ and the lower bound holds for $p_{\mathcal{S}}(u,\cdot)$. Thus, the one for $\tilde{p}^{T,y}(t,s,x,z)$ follows.

□

Remark 3.1. The idea of the proof was to identify the density of the frozen process to the density of some Lévy process and exploit the Lévy structure to derive bounds on the density. The procedure described above require the uniform ellipticity of σ in order to prove that the assumptions **[H]** holds for the new Lévy process $(\mathcal{S}_u)_{u\ge0}$. Intuitively, we can say that a uniform elliptic coefficient does not alter much the nature of the noise in the system. From the identity in law (3.39) that holds for fixed $s\in[t,T]$ and equation (2.23), we deduce that:

$$\tilde{p}^{T,y}(t,T,x,y)=p_{\mathcal{S}}(T-t,\theta_{t,T}(y)-x). \tag{3.40}$$

This identity will be useful when investigating the parametrix kernel H.

Now, we state a Dirac convergence Lemma for the Frozen process when the freezing parameter changes. This convergence will be used in the proof of the well posedness of the martingale problem. The difficulty comes from the fact that when integrating with respect to the freezing parameter (as it is the case in a parametrix procedure), the Dirac convergence does not follow from the Chapman-Kolmogorov equations. However, since we have good estimates on the frozen density, we manage to prove the following lemma:

Lemma 3.2. *For all bounded continuous function $f:\mathbb{R}^d\to\mathbb{R}$, $x\in\mathbb{R}^d$,*

$$\left|\int_{\mathbb{R}^d}f(y)\tilde{p}^{T,y}(t,T,x,y)dy-f(x)\right|\underset{T\downarrow t}{\longrightarrow}0, \tag{3.41}$$

that is, for all $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$, $\tilde{p}^{T,y}(t,T,x,y)dy\Rightarrow\delta_x(dy)$ weakly when $T\downarrow t$.

Proof. We prove this convergence in the lines of [HM14]. Let us write:

$$\begin{aligned}\int_{\mathbb{R}^d}f(y)\tilde{p}^{T,y}(t,T,x,y)dy-f(x) &= \int_{\mathbb{R}^d}f(y)\Big(\tilde{p}^{T,y}(t,T,x,y)-\tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\Big)dy\\ &\quad+\int_{\mathbb{R}^d}f(y)\Big(\tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\Big)dy-f(x).\end{aligned}$$

From the usual Dirac convergence in the Kolmogorov equation (2.29), the second term tends to zero when $T \to t$. We focus on the first term. Define:

$$\Delta = \int_{\mathbb{R}^d} f(y)\Big(\tilde{p}^{T,y}(t,T,x,y) - \tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\Big)dy. \tag{3.42}$$

For a given threshold $K > 0$ and a certain (small) $\beta > 0$ to be specified, we split $\mathbb{R}^d$ into $D_1 \cup D_2$ where:

$$D_1 = \left\{ y \in \mathbb{R}^d; \frac{|\theta_{t,T}(y) - x|}{(T-t)^{1/\alpha}} \leq K(T-t)^{-\beta} \right\},\ D_2 = \left\{ y \in \mathbb{R}^d; \frac{|\theta_{t,T}(y) - x|}{t^{1/\alpha}} > K(T-t)^{-\beta} \right\}.$$

A direct application of Proposition 3.1 yields:

$$\tilde{p}^{T,y}(t,T,x,y) \leq C \frac{(T-t)^{-d/\alpha}}{\Big(1 + \frac{|\theta_{t,T}(y)-x|}{(T-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar{q}(|\theta_{t,T}(y) - x|).$$

Now, observe that

$$y - x - \int_t^T F(u, \theta_{u,T}(\theta_{T,t}(x)))du = y - x - \int_t^T F(u, \theta_{u,t}(x))du = y - \theta_{T,t}(x).$$

Also, from the Lipschitz property of the flow, we have $|\theta_{t,T}(y) - x| \asymp |y - \theta_{T,t}(x)|$. Consequently, we obtain:

$$\tilde{p}^{t,\theta_{T,t}(x)}(t,T,x,y) \leq C \frac{(T-t)^{-d/\alpha}}{\Big(1 + \frac{|\theta_{t,T}(y)-x|}{(T-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar{q}(|\theta_{t,T}(y) - x|),$$

and we have the same upper bound for the two densities in (3.42). The idea is that on D_2, we use the tail estimate, and on D_1, we will explicitly exploit the compatibility between the spectral measures and the Fourier transform in the Fourier representation of the densities. Set for $i \in \{1, 2\}$:

$$\Delta_{D_i} := \int_{D_i} f(y)\Big(\tilde{p}^{T,y}(t,T,x,y) - \tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\Big)dy.$$

For D_2, we bound the two densities as we described above:

$$\begin{aligned} |I_{D_2}| &\leq C|f|_\infty \int_{D_2} \frac{(T-t)^{-d/\alpha}}{\Big(1 + \frac{|\theta_{t,T}(y)-x|}{(T-t)^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar{q}(C^{-1}|\theta_{t,T}(y) - x|)dy \\ &\leq C|f|_\infty \int_{K(T-t)^{-\beta}}^{+\infty} \frac{r^{d-1}}{1 + r^{\alpha+\gamma}} \bar{q}(r)dr \\ &\leq C(T-t)^{\beta(\gamma+\alpha-d)}. \end{aligned}$$

Thus, for $\beta > 0$, $I_{D_2} \underset{T \downarrow t}{\longrightarrow} 0$. On D_1, we will start from the inverse Fourier representation of $\tilde{p}^{T,z}(t,x,y)$, $z = \theta_{T,t}(x), y$. Recall we denoted by φ_Z the Lévy Khintchine exponent of Z, that is $e^{t\varphi_Z(p)} = \mathbb{E}(e^{i\langle p, Z_t\rangle})$, denoting σ^* the transpose of σ, we have:

$$\tilde{p}^{T,z}(t,T,x,y) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} dp e^{-i\langle p, y - \int_t^T F(u,\theta_{u,T}(z))du - x\rangle} \exp\Big(\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(z))^* p)du\Big).$$

Consequently, we have:

$$\begin{aligned}
&\tilde{p}^{T,y}(t,T,x,y)-\tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\\
=\;&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle p,y-\int_t^T F(u,\theta_{u,T}(y))du-x\rangle}e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)du}\\
&-e^{-i\langle p,y-\int_t^T F(u,\theta_{u,t}(x))du-x\rangle}e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,t}(x))^*p)du}dp\\
=\;&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(e^{-i\langle p,y-\int_t^T F(u,\theta_{u,T}(y))du-x\rangle}-e^{-i\langle p,y-\int_t^T F(u,\theta_{u,t}(x))du-x\rangle}\right)e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)du}dp\\
&+\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle p,y-\int_t^T F(u,\theta_{u,t}(x))du-x\rangle}\left(e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)du}-e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,t}(x))^*p)du}\right)dp\\
=\;&\Gamma_1(t,T,x,y)+\Gamma_2(t,T,x,y).
\end{aligned}$$

We split accordingly:

$$\int_{D_1} f(y)\Big(\tilde{p}^{T,y}(t,T,x,y)-\tilde{p}^{T,\theta_{T,t}(x)}(t,T,x,y)\Big)dy=\int_{D_1} f(y)\Gamma_1(t,T,x,y)dy+\int_{D_1} f(y)\Gamma_2(t,T,x,y)dy.$$

Note first that when $\alpha\le 1$, we assumed $F=0$, so that the term $\Gamma_1(t,T,x,y)=0$ in that case. We now treat this term, with $\alpha>1$. Using the mean value theorem, we write:

$$\begin{aligned}
&\Gamma_1(t,T,x,y)\\
=\;&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\int_0^1 d\lambda i\langle p,(I-\theta_{T,t})(\theta_{t,T}(y)-x)\rangle e^{-i\langle p,[\lambda I+(1-\lambda)\theta_{T,t}](\theta_{t,T}(y)-x)\rangle}e^{\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)du}dp,
\end{aligned}$$

where we denoted by I the identity map of $\mathbb{R}^d$. Recall that from the Lipschitz property of the flow and Gronwall's Lemma, there exists $C>0$ such that for all $t\le T$, $z\in\mathbb{R}^d$, $|(I-\theta_{T,t})(z)|\le C(T-t)(1+|z|)$. Thus, since $y\in D_1$, we have for $\beta\le 1/\alpha$,

$$|\Gamma_1(t,T,x,y)| \;\le\; C(T-t)\int_{\mathbb{R}^d}|p|e^{-K(T-t)|p|^\alpha}dp\le C(T-t)^{1-\frac{1}{\alpha}-\frac{d}{\alpha}}.$$

Integrating on D_1, we obtain:

$$\left|\int_{D_1} f(y)\Gamma_1(t,T,x,y)dy\right|\le C|f|_\infty(T-t)^{1-\frac{1}{\alpha}-\beta d}\underset{T\to t}{\longrightarrow}0,$$

when $1/d(1-1/\alpha)>\beta$. For Γ_2, we write:

$$\begin{aligned}
\Gamma_2(t,T,x,y) \;=\;&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dp e^{-i\langle p,y-\theta_{T,t}(x)\rangle}\int_0^1 d\lambda\; e^{\int_t^T \lambda\varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)+(1-\lambda)\varphi_Z(\sigma(u,\theta_{u,t}(x))^*p)du}\\
&\times\int_t^T(\varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)-\varphi_Z(\sigma(\theta_{u,t}(x))^*p))du.
\end{aligned}$$

We know from assumption **[H-2]** that the Lévy-Khintchine exponent is bounded by $-K(T-t)|p|^\alpha$, thus, we obtain independently of $\lambda\in(0,1)$:

$$e^{\int_t^T \lambda\varphi_Z(\sigma(u,\theta_{u,T}(y))^*p)+(1-\lambda)\varphi_Z(\sigma(u,\theta_{u,t}(x))^*p)du}\le e^{-K(T-t)|p|^\alpha}.$$

On the other hand, using the bound on the Lévy-Khintchine exponent and assumption **[H-5]**, we can rewrite the increment:

$$\begin{aligned}
&\left|\int_t^T \varphi_Z(\sigma(u,\theta_{u,T}(y))^* p) - \varphi_Z(\sigma(u,\theta_{u,t}(x))^* p) du\right| \\
= &\left|\int_t^T \int_{\mathbb{R}^d} \cos(\langle \sigma(u,\theta_{u,T}(y))^* p, \xi\rangle) - \cos(\langle \sigma(u,\theta_{u,t}(x))^* p, \xi\rangle)\nu(dz)du\right| \\
\leq \; & K|p|^\alpha \int_t^T |\theta_{u,t}(x) - \theta_{u,T}(y)|^{\eta(\alpha\wedge 1)} \leq K(T-t)|p|^\alpha |x - \theta_{t,T}(y)|^{\eta(\alpha\wedge 1)}.
\end{aligned}$$

To summarize, we obtained:

$$\begin{aligned}
\int_{D_1} f(y)\Gamma_2(t,T,x,y) &\leq |f|_\infty \int_{D_1} dy\, |\Gamma_2(t,T,x,y)| \\
&\leq C\int_{D_1} dy \int_{\mathbb{R}^d} (T-t)|p|^\alpha |x-\theta_{t,T}(y)|^{\eta(\alpha\wedge 1)} e^{-K(T-t)|p|^\alpha} dp.
\end{aligned}$$

Changing variables, and integrating over p yields

$$\begin{aligned}
\int_{D_1} f(y)\Gamma_2(t,T,x,y) &\leq \frac{C}{t^{d/\alpha}}|f|_\infty \int_{D_1} dy |\theta_{t,T}(y) - x|^{\eta(\alpha\wedge 1)} \\
&= \frac{C}{(T-t)^{d/\alpha}}|f|_\infty \int_0^{(T-t)^{-\beta}} dr\; r^{\eta(\alpha\wedge 1)+d-1}(T-t)^{d/\alpha+\eta(1\wedge 1/\alpha)}.
\end{aligned}$$

Choosing now $\frac{\eta(1/\alpha\wedge 1)}{d+\eta(\alpha\wedge 1)} > \beta > 0$ gives that $|I_{D_1}| \underset{T\downarrow t}{\longrightarrow} 0$, which concludes the proof. □

3.2. **The Smoothing Properties of $H(t,x,y)$.**

First, we investigate an upper bound for the Parametrix Kernel. Recall that:

$$\forall t \geq 0,\ (x,y) \in (\mathbb{R}^d)^2,\ H(t,T,x,y) := \Big(L(x,\nabla_x) - L(\theta_{t,T}(y),\nabla_x)\Big)\tilde{p}^{T,y}(t,T,x,y).$$

Proposition 3.3. *Assume **[H]** is in force. There exists $C > 0$ s.t. for all $t \in (0,T]$, $(x,y) \in (\mathbb{R}^d)^2$:*

$$|H(t,T,x,y)| \leq C\left((T-t)^{-1/\alpha} + \frac{\delta \wedge |x-\theta_{t,T}(y)|^{\eta(\alpha\wedge 1)}}{T-t}\right)\bar{p}(t,T,x,y),$$

where we recall that

$$\bar{p}(t,T,x,y) = \frac{(T-t)^{-d/\alpha}}{\left(1+\frac{|\theta_{t,T}(y)-x|}{(T-t)^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(|\theta_{t,T}(y)-x|).$$

Thus, the upper bound on the Kernel H is the same as the upper bound on the Frozen density $\tilde{p}^{T,y}(t,T,x,y)$ up to the additional multiplier $\big(\delta \wedge |x-\theta_{t,T}(y)|^{\eta(\alpha\wedge 1)}\big)(T-t)^{-1}$, that can be seen as the singularity induced by the difference $L(x,\nabla_x) - L(\theta_{t,T}(y),\nabla_x)$ applied to the frozen density. The proof proceeds following the lines of Sztonyk [Szt10], splitting the large jumps and the small jumps. The small jumps are dealt using Fourier analysis techniques, whereas the big jumps are dealt more directly.

Proof. Recall that the density of $\tilde{X}_s^{T,y}$ can be linked to the density of the Lévy process $(\mathcal{S}_u)_{u\geq 0}$ considered at time $s-t$. We now exploit the Lévy structure of $(\mathcal{S}_u)_{u\geq 0}$ to obtain an upper bound on $H(t,T,x,y)$. Specifically, let us introduce the Lévy-Itô decomposition of $(\mathcal{S}_u)_{u\geq 0}$:

$$\mathcal{S}_u = M_u + N_u,$$

where $(M_u)_{u\geq 0}$ is a martingale and $(N_u)_{u\geq 0}$ is a poisson process. We choose to place the cut-off at the characteristic time-scale, namely $(T-t)^{1/\alpha}$. Therefore, the Fourier transform of M_u writes:

$$\mathbb{E}(e^{i\langle p, M_u\rangle}) = \exp\left(u\int_{\mathbb{R}^d}(e^{i\langle p,\eta\rangle} - 1 - i\langle p,\eta\rangle)\mathbf{1}_{\{|z|\leq (T-t)^{1/\alpha}\}}\nu_{\mathcal{S}}(d\eta)\right).$$

This expression is integrable and regular in the variable p (see Sztonyk [Szt10] and the references therein). Thus, the density $p_M(u,\cdot)$ of M_u exists and is the Schwartz's class. Thus, we can say that this term produces the density in the Lévy-Itô decomposition. Also, denoting by $\bar{\nu}_{\mathcal{S}}(dz) = \mathbf{1}_{\{|z|\geq (T-t)^{1/\alpha}\}}\nu_{\mathcal{S}}(dz)$, we have the following decomposition for the law of the Poisson Process N_u:

$$P_{N_u}(dz) = e^{-u\bar{\nu}_{\mathcal{S}}(\mathbb{R}^d)}\sum_{k=0}^{+\infty}\frac{u^k\bar{\nu}_{\mathcal{S}}^{*k}(dz)}{k!}.$$

Now, by independence of $(M_u)_{u\geq 0}$ and $(N_u)_{u\geq 0}$ and exploiting equation (3.40), we get:

$$\tilde{p}^{T,y}(t,T,x,y) = p_{\mathcal{S}}(T-t,\theta_{t,T}(y)-x) = \int_{\mathbb{R}^d} p_M(T-t,\theta_{t,T}(y)-x-\xi)P_{N_{T-t}}(d\xi). \tag{3.43}$$

From the definition of the generators, the operator naturally splits into three parts, for a test function φ:

$$\begin{aligned}\Big(L_t(x,\nabla_x) - L_t(\theta_{t,T}(y),\nabla_x)\Big)\varphi(x) &= \langle\nabla_x\varphi(x), F(t,x) - F(t,\theta_{t,T}(y))\rangle\\ + \int_{\mathbb{R}^d}\Big(\varphi(x+z)-\varphi(x)-\langle\nabla\varphi(x),z\rangle\Big)\mathbf{1}_{\{|z|\leq(T-t)^{1/\alpha}\}}&(\nu_t(x,dz)-\nu_t(\theta_{t,T}(y),dz))\\ + \int_{\mathbb{R}^d}\Big(\varphi(x+z)-\varphi(x)\Big)\mathbf{1}_{\{|z|\geq(T-t)^{1/\alpha}\}}&(\nu_t(x,dz)-\nu(\theta_{t,T}(y),dz)).\end{aligned}$$

Recall that we defined $\nu_t(\xi,A) = \nu\{z\in\mathbb{R}^d; \sigma(t,\xi)z\in A\}$. Also, observe that by symmetry of ν, we changed the cut-off function to exhibit the intrinsic time-scale. Note that the first order term in the operator is present only in the case $\alpha > 1$. Otherwise, we assumed that $F = 0$.

We treat the three terms separately. For the first order term, we see that the gradient acts on the variable x. From the decomposition (3.43), we see that the dependency in x appears in the density of the martingale, which is in the Schwartz class. It is straightforward from [Szt10] that for all $m\geq 1$:

$$|\langle\nabla_x p_M(T-t,\theta_{t,T}(y)-x-\xi), F(t,x)-F(t,\theta_{t,T}(y))\rangle| \leq C_m\frac{|x-\theta_{t,T}(y)|}{(T-t)^{1/\alpha}}(T-t)^{-d/\alpha}\left(1+\frac{|\theta_{t,T}(y)-x-\xi|}{(T-t)^{1/\alpha}}\right)^{-m}.$$

Thus, we recovered Lemma 2 in [Szt10] for the first order term, up to the additional multiplyier $\frac{|x-\theta_{t,T}(y)|}{(T-t)^{1/\alpha}}$. Consequently, we get:

$$|\langle\nabla_x\tilde{p}^{T,y}(t,T,x,y), F(t,x)-F(t,\theta_{t,T}(y))\rangle| \leq \frac{|x-\theta_{t,T}(y)|}{(T-t)^{1/\alpha}}\bar{p}(t,T,x,y).$$

For the small jumps part, once again, we observe that the operator acts on the variable x, and thus can be put on the density of the martingale. Since p_M is given by a Fourier inversion, it is convenient to use the representation of the operator $L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)$ defined by:

$$\begin{aligned} & \Big(L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)\Big)\varphi(x) \\ = & \int_{\mathbb{R}^d} \Big(\varphi(x+z) - \varphi(x) - \langle \nabla_x \varphi(x), z\rangle\Big) \mathbf{1}_{\{|z|\leq (T-t)^{1/\alpha}\}} (\nu(x,dz) - \nu(\theta_{t,T}(y),dz)), \end{aligned}$$

in terms of symbol. Let us denote by:

$$l_t^M(x,p) - l_t^M(\theta_{t,T}(y),p) = \int_{\mathbb{R}^d} e^{i\langle p,z\rangle} - 1 - i\langle p,z\rangle \mathbf{1}_{|z|\leq (T-t)^{1/\alpha}} \big(\nu_t(x,dz) - \nu_t(\theta_{t,T}(y),dz\big),$$

the symbol of the integro-differential operator $L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)$. We have that:

$$\begin{aligned} & \Big(L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)\Big) p_M\Big(T-t, \theta_{t,T}(y) - x - \xi\Big) \\ = & \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle p, \theta_{t,T}(y) - x - \xi\rangle} \Big(l_t^M(x,p) - l_t^M(\theta_{t,T}(y),p)\Big) \mathbb{E}(e^{i\langle p, M_{T-t}\rangle}) dp. \end{aligned}$$

We change variables to $q = (T-t)^{1/\alpha} p$. This yields:

$$\begin{aligned} & \Big(L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)\Big) p_M\Big(T-t, \theta_{t,T}(y) - x - \xi\Big) \\ = & (T-t)^{-d/\alpha} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} dq e^{-i\langle q, \frac{\theta_{t,T}(y) - x - \xi}{(T-t)^{1/\alpha}}\rangle} \Big(l_t^M(x, \frac{q}{(T-t)^{1/\alpha}}) - l_t^M(\theta_{t,T}(y), \frac{q}{(T-t)^{1/\alpha}})\Big) \\ & \times \mathbb{E}(e^{i\langle q, (T-t)^{-1/\alpha} M_{T-t}\rangle}). \end{aligned}$$

For $x, y \in \mathbb{R}^d$ and $0 \leq t < T$ fixed, let us denote h_{T-t} the function whose Fourier transform is given by:

$$\begin{aligned} \hat{h}_{T-t}(q) = & \frac{T-t}{\delta \wedge |\theta_{t,T}(y) - x|^{\eta(\alpha\wedge 1)}} \Big(l_t^M(x, \frac{q}{(T-t)^{1/\alpha}}) - l_t^M(\theta_{t,T}(y), \frac{q}{(T-t)^{1/\alpha}})\Big) \\ & \times \exp\left((T-t)\int_{\mathbb{R}^d} (e^{i\langle \frac{q}{(T-t)^{1/\alpha}}, \eta\rangle} - 1 - i\langle \frac{q}{(T-t)^{1/\alpha}}, \eta\rangle) \mathbf{1}_{\{|z|\leq (T-t)^{1/\alpha}\}} \nu_{\mathcal{S}}(d\eta)\right). \end{aligned}$$

By definition, we have:

$$\begin{aligned} & \Big(L_t^M(x,\nabla_x) - L_t^M(\theta_{t,T}(y),\nabla_x)\Big) p_M\Big(T-t, \theta_{t,T}(y) - x - \xi\Big) \\ = & \frac{\delta \wedge |\theta_{t,T}(y) - x|^{\eta(\alpha\wedge 1)}}{T-t} (T-t)^{-d/\alpha} h\left(\frac{\theta_{t,T}(y) - x - \xi}{(T-t)^{1/\alpha}}\right). \end{aligned}$$

The goal is now to prove that uniformly in $x, y \in \mathbb{R}^d$ and $0 \leq t \leq T$, for all $m \geq 1$, there exists some positive constant $C_m > 0$ such that:

$$h_{T-t}(\zeta) \leq C_m(1 + |\zeta|^m).$$

To that end, we prove that $\hat{h}_{T-t}$ is in the Schwartz class, uniformly for all $x, y \in \mathbb{R}^d$ and $0 \leq t \leq T$.

We see that thanks to the truncation, the function $\hat{h}_{T-t}$ is smooth (see Sztonyk [Szt10] and the references therein). Moreover, we point out that from **[H-5]**, we have:

$$|l_t^M(x,\frac{q}{(T-t)^{1/\alpha}}) - l_t^M(\theta_{t,T}(y),\frac{q}{(T-t)^{1/\alpha}})| \leq C\delta \wedge |\theta_{t,T}(y)-x|^{\eta(\alpha\wedge 1)}\left(\frac{|q|}{(T-t)^{1/\alpha}}\right)^{\alpha}.$$

Besides, from **[H-1]** and **[H-2]** for $\nu_{\mathcal{S}}$:

$$\begin{aligned}&\exp\left((T-t)\int_{\mathbb{R}^d}(e^{i\langle\frac{q}{(T-t)^{1/\alpha}},\eta\rangle}-1-i\langle\frac{q}{(T-t)^{1/\alpha}},\eta\rangle)\mathbf{1}_{\{|z|\leq(T-t)^{1/\alpha}\}}\nu_{\mathcal{S}}(d\eta)\right)\\ \leq\ &\exp(-K|q|^{\alpha})\exp(C(T-t)^{1/\alpha}\nu_{\mathcal{S}}(B(0,(T-t)^{1/\alpha}))),\end{aligned}$$

see also equation (19) in Stzonyk [Szt10]. Now, from assumption **[H-1]** for $\nu_{\mathcal{S}}$, we get

$$\nu_{\mathcal{S}}(B(0,(T-t)^{1/\alpha})) \leq C(T-t)^{-1/\alpha},$$

so that we actually obtain:

$$\exp\left((T-t)\int_{\mathbb{R}^d}(e^{i\langle\frac{q}{(T-t)^{1/\alpha}},\eta\rangle}-1-i\langle\frac{q}{(T-t)^{1/\alpha}},\eta\rangle)\mathbf{1}_{\{|z|\leq(T-t)^{1/\alpha}\}}\nu_{\mathcal{S}}(d\eta)\right) \leq C\exp(-K|q|^{\alpha}).$$

Thus, uniformly in $x,y\in\mathbb{R}^d$ and $0\leq t\leq T$, $\hat{h}_{T-t}(q)$ is in the Schwartz class, which is stable by Fourier transform. We finally obtain:

$$\begin{aligned}&\left|(L_t^M(x,\nabla_x)-L_t^M(\theta_{t,T}(y),\nabla_x))p_M\Big(T-t,\theta_{t,T}(y)-x-\xi\Big)\right|\\ &\leq C\frac{\delta\wedge|\theta_{t,T}(y)-x|^{\eta(\alpha\wedge1)}}{T-t}(T-t)^{-d/\alpha}\left(1+\frac{|\theta_{t,T}(y)-x-\xi|}{(T-t)^{1/\alpha}}\right)^{-m}.\end{aligned}$$

Consequently, we recovered Lemma 2 in [Szt10] for the Parametrix kernel, up to the additional multiplicative term $\big(\delta\wedge|\theta_{t,T}(y)-x|^{\eta(\alpha\wedge1)}\big)(T-t)^{-1}$, which is the expected singularity for the Kernel (see Kolokoltsov [Kol00]). The upper bound follows from this upper bound and the control of the measure of the balls for $P_{N_{T-t}}$ similarly to the derivation of the upper bound for the density, see Corollary 6 in [Szt10] and the proof of Theorem 1 in [Szt10]. The upper bound for the small jumps part of the kernel follows.

Finally, for the large jumps, the measure $\mathbf{1}_{\{|\xi|\geq(T-t)^{1/\alpha}\}}(\nu(x,d\xi)-\nu(\theta_{t,T}(y),d\xi))$ is no more singular. Thus, we can write:

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}\Big(\tilde{p}(t,T,x+\xi,y)-\tilde{p}(t,T,x,y)\Big)\mathbf{1}_{\{|\xi|\geq(T-t)^{1/\alpha}\}}(\nu(x,d\xi)-\nu(\theta_{t,T}(y),d\xi))\right|\\ &\leq\int_{\mathbb{R}^d}\Big|\tilde{p}(t,T,x+\xi,y)-\tilde{p}(t,T,x,y)\Big|\mathbf{1}_{\{|\xi|\geq(T-t)^{1/\alpha}\}}|\nu(x,d\xi)-\nu(\theta_{t,T}(y),d\xi)|\\ &\leq\delta\wedge|x-\theta_{t,T}(y)|^{\eta(\alpha\wedge1)}\bigg(\int_{S^{d-1}}\int_0^{+\infty}\tilde{p}(t,T,x+s\varsigma,y)\mathbf{1}_{\{s\geq(T-t)^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\varsigma)\\ &\quad+\frac{1}{T-t}\tilde{p}(t,T,x,y)\bigg).\end{aligned}$$

For the last inequality, we exploited **[H-5]**. We focus on the remaining integral term above. When the diagonal regime holds, the estimate is straightforward, as we can directly bound:

$$\tilde{p}(t,T,x+s\varsigma,y)\leq C(T-t)^{-d/\alpha}\leq C\bar{p}(t,T,x,y).$$

The integral then yields the singularity $(T-t)^{-1}$. Therefore, we assume that $|\theta_{t,T}(y)-x| \geq (T-t)^{1/\alpha}$. The regime of $\tilde{p}(t,x+s\varsigma,y)$ is given by $|\theta_{t,T}(y)-x-s\varsigma|$. Thus, thanks to the triangle inequality, when:

$$|\theta_{t,T}(y)-x| \leq 1/2s, \text{ or when } s \leq 1/2|\theta_{t,T}(y)-x|,$$

the density $\tilde{p}(t,T,x+s\varsigma,y)$ is off-diagonal with:

$$\tilde{p}(t,T,x+s\varsigma,y) \leq C\bar{p}(t,T,x,y).$$

Consequently, the problematic case is when $s \asymp |\theta_{t,T}(y)-x|$. Indeed, in this case, $\tilde{p}(t,T,x+s\varsigma,y)$ can be in diagonal regime, whereas $\tilde{p}(t,T,x,y)$ is still in the off-diagonal regime. But in this case, we have

$$\frac{\bar{q}(s)}{s^{1+\alpha}} \asymp \frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{1+\alpha}}.$$

Hence, we can take this part out of the integral and integrate a density to get:

$$\begin{aligned}
&\int_{1/2|\theta_{t,T}(y)-x|}^{3/2|\theta_{t,T}(y)-x|} \int_{S^{d-1}} \tilde{p}(t,T,x+s\varsigma,y)\mathbf{1}_{\{s\geq(T-t)^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\varsigma)\\
\leq\ & \frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{1+\alpha}}\int_0^{+\infty}\int_{S^{d-1}}\tilde{p}(t,T,x+s\varsigma,z)\mathbf{1}_{\{s\geq(T-t)^{1/\alpha}\}}ds\mu(d\varsigma)\\
\leq\ & C\frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{1+\alpha}}.
\end{aligned}$$

Rewriting the right hand side to make the time dependencies appear:

$$\begin{aligned}
\frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{1+\alpha}} &= \frac{1}{T-t}\frac{(T-t)^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{t,T}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{t,T}(y)-x|)\times(T-t)^{\frac{d-\gamma}{\alpha}}\\
&\leq C\frac{1}{T-t}\frac{(T-t)^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{t,T}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{t,T}(y)-x|).
\end{aligned}$$

In the last inequality, we recall that $\gamma \leq d$, so that $(T-t)^{\frac{d-\gamma}{\alpha}} \leq 1$. The estimate on the kernel follows from the fact that thanks to equation (1.10) in assumption **[H-1]**, we have:

$$\frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{1+\alpha}} \leq C\frac{\bar{q}(C|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\alpha+\gamma}}.$$

In other words, we can correct the wrong decay thanks to temperation.

□

Remark 3.2. In the above proof, the temperation only serves to compensate the bad concentration in the generator. Also, we see that when the spectral measure μ dominating the Lévy measure ν has a density on the sphere, then, the large jump part of the difference of the generators becomes:

$$\int_{\mathbb{R}^d}\tilde{p}(t,T,x+\xi,y)\mathbf{1}_{\{|\xi|\geq(T-t)^{1/\alpha}\}}\nu(d\xi) \leq C\int_{\mathbb{R}^d}\tilde{p}(t,T,x+\xi,y)\mathbf{1}_{\{|\xi|\geq(T-t)^{1/\alpha}\}}\frac{\bar{q}(|\xi|)}{|\xi|^{d+\alpha}}d\xi.$$

Thus, when $s \asymp |\theta_{t,T}(y)-x|$, as in the last case discussed above, we have directly the good concentration index and the temperation is not needed. In particular, when $\bar{q}=1$, we recovered the results in Kolokolstov [Kol00] in a weaker framework for the coefficients in (1.1).

We have obtained the same type of estimate on the kernel and on the frozen density. Let us observe that the upper bound satisfies a "semigroup" property in the following sense.

Lemma 3.4. *Fix $t \in [0,T]$. Let us denote $\bar{p}_C(t,T,x,z) = \frac{(T-t)^{-d/\alpha}}{\left(1+\frac{|\theta_{t,T}(z)-x|}{(T-t)^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(C|\theta_{t,T}(z)-x|)$. Let $C_1, C_2 > 0$. For all $\tau \in [0,t]$, there exists $C_3 > 0$:*

$$\int_{\mathbb{R}^d} \bar{p}_{C_1}(t,\tau,x,z)\bar{p}_{C_2}(\tau,T,z,y)dz \leq C\bar{p}_{C_3}(t,T,x,y).$$

Proof. The proof follows by an application of the triangle inequality and the Lipschitz property of the flow plus the fact that $\bar{q}$ is non increasing.

□

We exhibit here some smoothing properties in time of the Parametrix Kernel. These properties will become crucial when investigating the convergence of the series (2.26) on the one hand and the lower bound of Theorem 1.2 on the other. The following lemma is a regularizing effect in time of the Parametrix kernel.

Lemma 3.5. *There exists $C > 1$, $\omega > 0$ s.t. for all $\tau \in (t,T)$, $(x,y) \in (\mathbb{R}^d)^2$:*

$$\begin{aligned}\int_{\mathbb{R}^d} \delta \wedge |x-\theta_{\tau,t}(z)|^{\eta(\alpha\wedge 1)}\bar{p}(t,\tau,x,z)dz &\leq C(\tau-t)^{\omega},\\ \int_{\mathbb{R}^d} \delta \wedge |\theta_{\tau,T}(y)-z|^{\eta(\alpha\wedge 1)}\bar{p}(\tau,T,z,y)dz &\leq C(T-\tau)^{\omega}.\end{aligned}$$

As a corollary, we get that

$$\int_t^T d\tau \int_{\mathbb{R}^d} |H(\tau,T,z,y)|dz \leq C(T-t)^{\omega}.$$

Thus, when integrated in time, the parametrix Kernel yields has a smoothing property in time.

Proof. The two estimates are similar, we shall only prove one. Let us denote the quantity of interest:

$$I = \int_{\mathbb{R}^d} dz\ \delta \wedge |x-\theta_{t,\tau}(z)|^{\eta(\alpha\wedge 1)}\frac{(\tau-t)^{-d/\alpha}}{\left(1+\frac{|x-\theta_{t,\tau}(z)|}{(\tau-t)^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(|x-\theta_{t,\tau}(z)|).$$

We split $\mathbb{R}^d = D_1 \cup D_2$, with

$$\begin{aligned}D_1 &= \{z \in \mathbb{R}^d; |x-\theta_{t,\tau}(z)| \leq C(\tau-t)^{1/\alpha}\}\\ D_2 &= \{z \in \mathbb{R}^d; |x-\theta_{t,\tau}(z)| > C(\tau-t)^{1/\alpha}\}.\end{aligned}$$

We write I_{D_i} for the integral over $z \in D_i$. For $z \in D_1$ we have:

$$\frac{(\tau-t)^{-d/\alpha}}{\left(1+\frac{|x-\theta_{t,\tau}(z)|}{(\tau-t)^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(|x-\theta_{t,\tau}(z)|) \leq (\tau-t)^{-d/\alpha},\ |x-\theta_{t,\tau}(z)|^{\eta(\alpha\wedge 1)} \leq (\tau-t)^{\eta(1\wedge 1/\alpha)}.$$

Also, D_1 is a compact and its Lebesgue measure is exactly $(\tau-t)^{d/\alpha}$, thus, we obtain $I_{D_1} \leq (\tau-t)^{\eta(1\wedge 1/\alpha)}$.

When $z \in D_2$, we have:

$$\begin{aligned} I_{D_2} &\leq \int_{D_2} dz \delta \wedge |x - \theta_{t,\tau}(z)|^{\eta(\alpha\wedge 1)} \frac{(\tau - t)^{1+\frac{\gamma-d}{\alpha}}}{|x - \theta_{t,\tau}(z)|^{\alpha+\gamma}} \\ &\leq (\tau - t)^{1+\frac{\gamma-d}{\alpha}} \int_{|z-\theta_{\tau,t}(x)|>C(\tau-t)^{1/\alpha}} dz \frac{\delta \wedge |z - \theta_{\tau,t}(x)|^{\eta(\alpha\wedge 1)}}{|z - \theta_{\tau,t}(x)|^{\alpha+\gamma}}. \end{aligned}$$

Observe that we used the Lipschitz property of the flow to switch from $x - \theta_{t,\tau}(z)$ to $z - \theta_{\tau,t}(x)$. This allows us to change variables and set $X = (z - \theta_{\tau,t}(x))/(\tau - t)^{1/\alpha}$, we get:

$$I_{D_2} \leq (\tau - t)^{1+\frac{\gamma-d}{\alpha}} \int_{|X|>1} \frac{(\tau - t)^{\eta(1\wedge\frac{1}{\alpha})} |X|^{\eta(\alpha\wedge 1)}}{|X|^{\alpha+\gamma}} dX.$$

Thus, the result follows when $\alpha + \gamma - d > \eta(\alpha \wedge 1)$. When it is not the case, we split again:

$$\begin{aligned} \int_{|z-\theta_{\tau,t}(x)|>(\tau-t)^{1/\alpha}} \frac{\delta \wedge |z - \theta_{\tau,t}(x)|^{\eta(\alpha\wedge 1)}}{|z - \theta_{\tau,t}(x)|^{\alpha+\gamma}} dz = &\int_{1\geq|z-\theta_{\tau,t}(x)|>(\tau-t)^{1/\alpha}} \frac{\delta \wedge |z - \theta_{\tau,t}(x)|^{\eta(\alpha\wedge 1)}}{|z - \theta_{\tau,t}(x)|^{\alpha+\gamma}} dz \\ &+ \int_{|z-\theta_{\tau,t}(x)|>1} \frac{\delta \wedge |z - \theta_{\tau,t}(x)|^{\eta(\alpha\wedge 1)}}{|z - \theta_{\tau,t}(x)|^{\alpha+\gamma}} dz. \end{aligned}$$

The second part of the right hand side is clearly a constant, bounding $\delta \wedge |z - \theta_{\tau,t}(x)|^{\eta(\alpha\wedge 1)} \leq \delta$, since $\alpha + \gamma > d$. For the first part, we change variable again to $Y = (z - \theta_{\tau,t}(x))$, which yields when $\alpha + \gamma - d < \eta(\alpha \wedge 1)$:

$$\int_{1>|Y|>(\tau-t)^{1/\alpha}} \frac{|Y|^{\eta(\alpha\wedge 1)}}{|Y|^{\alpha+\gamma}} dY \leq C.$$

On the other hand, when $\alpha + \gamma - d = \eta(\alpha \wedge 1)$

$$\int_{1>|Y|>(\tau-t)^{1/\alpha}} \frac{1}{|Y|^d} dY = [\log(|Y|)]_{(\tau-t)^{1/\alpha}}^1 \leq \frac{1}{\alpha} |\log(\tau - t)|.$$

Thus the proof is complete.

□

3.3. **Proof of Theorem 1.1: Uniqueness to the Martingale Problem**

We are now in position to prove the uniqueness to the martingale problem. Our approach relies on the smoothing properties of the Parametrix kernel H.

Proof. We focus on uniqueness. The existence can be derived from standard compactness arguments (see e.g. Chapter 6 in Stroock and Varadhan [SV79], or Stroock [Str75]). Suppose we are given two solutions $\mathbb{P}^1$ and $\mathbb{P}^2$ of the martingale problem associated with $L(\cdot, \nabla_\cdot)$, starting in x at time 0. We can assume w.l.o.g. that $t \leq T$, the fixed time horizon. Define for a bounded Borel function $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$,

$$S^i f = \mathbb{E}^i \left(\int_t^T f(s, X_s) ds \right), \ i \in \{1, 2\},$$

where $(X_t)_{t\geq 0}$ stands for the canonical process associated with $(\mathbb{P}^i)_{i\in\{1,2\}}$. Let us specify that $S^i f$ is *a priori* only a linear functional and not a function since $\mathbb{P}^i$ does not need to come from a Markov process. We denote:

$$S^\Delta f = S^1 f - S^2 f,$$

and the aim of this section is to prove that $S^{\Delta} f = 0$ for f in a suitable class of test functions.

If $f \in \mathcal{C}_0^{1,2}([0,T) \times \mathbb{R}^d, \mathbb{R})$, since $(\mathbb{P}^i)_{i\in\{1,2\}}$ both solve the martingale problem, we have:

$$f(t,x) + \mathbb{E}^i \left(\int_t^T (\partial_s + L_s(x, \nabla_x) f(s, X_s) ds \right) = 0, \ i \in \{1,2\}. \tag{3.44}$$

As a consequence we thus have that for all $f \in C_0^{1,2}([0,T] \times \mathbb{R}^{nd}, \mathbb{R})$,

$$S^{\Delta}\Big((\partial_s + L_s(x, \nabla_x)) f \Big) = 0. \tag{3.45}$$

We now want to apply (3.45) to a suitable function f. For a fixed point $y \in \mathbb{R}^d$ and a given $\varepsilon \geq 0$, introduce for all $f \in \mathcal{C}_0^{1,2}([0,T) \times \mathbb{R}^d, \mathbb{R})$ the Green kernel:

$$\forall (t,x) \in [0,T) \times \mathbb{R}^d, G^{\varepsilon,y} f(t,x) = \int_t^T ds \int_{\mathbb{R}^d} dz \tilde{p}^{s+\varepsilon,y}(t,s,x,z) f(s,z).$$

We define for all $f \in C_0^{1,2}([0,T) \times \mathbb{R}^d, \mathbb{R})$:

$$M_{t,x}^{\varepsilon,y} f(t,x) = \int_t^T ds \int_{\mathbb{R}^d} dz L_s(\theta_{t,s+\varepsilon}(y), \nabla_x) \tilde{p}^{s+\varepsilon,y}(t,s,x,z) f(s,z).$$

We derive from the Backward Kolmogorov equation for the frozen density that the following equality holds:

$$\partial_t G^{\varepsilon,y} f(t,x) + M_{t,x}^{\varepsilon,y} f(t,x) = -f(t,x), \ \forall (t,x) \in [0,T) \times \mathbb{R}^d. \tag{3.46}$$

Now, let $h \in C_0^{1,2}([0,T) \times \mathbb{R}^{nd}, \mathbb{R})$ be an arbitrary function and define for all $(t,x) \in [0,T) \times \mathbb{R}^{nd}$:

$$\phi^{\varepsilon,y}(t,x) := \tilde{p}^{t+\varepsilon,y}(t, t+\varepsilon, x, y) h(t,y), \Psi_\varepsilon(t,x) := \int_{\mathbb{R}^d} dy G^{\varepsilon,y}(\phi^{\varepsilon,y})(t,x).$$

Then, by semigroup property, we have:

$$\begin{aligned} \Psi_\varepsilon(t,x) &= \int_{\mathbb{R}^d} dy \int_t^T ds \int_{\mathbb{R}^d} dz \tilde{p}^{s+\varepsilon,y}(t,s,x,z) \tilde{p}^{s+\varepsilon,y}(s, s+\varepsilon, z, y) h(s,y) \\ &= \int_{\mathbb{R}^d} dy \int_t^T ds \tilde{p}^{s+\varepsilon,y}(t, s+\varepsilon, x, y) h(s,y). \end{aligned}$$

Hence, we can write:

$$\begin{aligned} \partial_t \Psi_\varepsilon(t,x) + L_t(x, \nabla_x) \Psi_\varepsilon(t,x) &= \int_{\mathbb{R}^d} dy \Big(\partial_t G^{\varepsilon,y} \phi^{\varepsilon,y}(t,x) + M_{t,x}^{\varepsilon,y} \phi^{\varepsilon,y}(t,x) \Big) \\ &\quad + \int_{\mathbb{R}^d} dy \Big(L_t(x, \nabla_x) G^y \phi^{\varepsilon,y}(t,x) - M_{t,x}^{\varepsilon,y} \phi^{\varepsilon,y}(t,x) \Big) \\ &:= I_1^\varepsilon(t,x) + I_2^\varepsilon(t,x). \end{aligned}$$

Observe that from (3.46), we have:

$$I_1^\varepsilon(t,x) = - \int_{\mathbb{R}^d} \tilde{p}^{t+\varepsilon,y}(t, t+\varepsilon, x, y) h(t,y) dy.$$

Now, from Lemma 3.2, when $\varepsilon \to 0$ we have the convergence:

$$\int_{\mathbb{R}^d} \tilde{p}^{t+\varepsilon,y}(t,t+\varepsilon,x,y)h(t,y)dy \underset{\varepsilon\to 0}{\longrightarrow} h(t,x).$$

Consequently, $I_1^\varepsilon(t,x)$ allows us to recover the test function $h(t,x)$ when ε tends to zero, that is:

$$\lim_{\varepsilon\to 0} \left|S^\Delta(I_1^\varepsilon)\right| = |S^\Delta h|.$$

On the other hand,

$$I_2^\varepsilon(t,x) = \int_{\mathbb{R}^d} dy\Big(L_t(x,\nabla_x)G^y\phi^{\varepsilon,y}(t,x) - M_{t,x}^{\varepsilon,y}\phi^{\varepsilon,y}(t,x)\Big)$$

$$= \int_{\mathbb{R}^d} dy \int_t^T ds\Big(L_t(x,\nabla_x) - L_t(\theta_{t,s+\varepsilon}(y),\nabla_x)\Big)\tilde{p}_\alpha^{s+\varepsilon,y}(t,s+\varepsilon,x,y)h(s,y)$$

$$= \int_{\mathbb{R}^d} dy \int_t^T dsH(t,s+\varepsilon,x,y)h(s,y).$$

From the controls of Subsection 3.2, specifically, Lemma 3.3, we have for all $(t,x) \in [0,T]\times\mathbb{R}^d$:

$$|I_2^\varepsilon(t,x)| \le |h|_\infty \int_{\mathbb{R}^d} dy \int_t^T ds|H(t,s+\varepsilon,x,y)| \le C(T+\varepsilon-t)^\omega |h|_\infty.$$

Thus, denoting by $||S^\Delta|| := \sup_{|f|_\infty\le 1}|S^\Delta f|$, we have:

$$\lim_{\varepsilon\to 0}\left|S^\Delta(I_2^\varepsilon)\right| \le ||S^\Delta|| \liminf_{\varepsilon\to 0} |I_2^\varepsilon|_\infty \le C||S^\Delta||(T-t)^\omega|h|_\infty.$$

Now, from (3.45) with $f(t,x) = \Psi_\varepsilon(t,x)$, we have

$$S^\Delta\Big((\partial_\cdot + L(\cdot,\nabla_\cdot))\Psi_\varepsilon\Big) = 0 \Rightarrow |S^\Delta(I_1^\varepsilon)| = |S^\Delta(I_2^\varepsilon)|.$$

Thus, for $T-t$ small enough,

$$|S^\Delta h| = \lim_{\varepsilon\to 0}\left|S^\Delta(I_1^\varepsilon)\right| = \lim_{\varepsilon\to 0}\left|S^\Delta I_2^\varepsilon\right| \le 1/2\|S^\Delta\||h|_\infty.$$

By a monotone class argument, the previous inequality still holds for bounded Borel functions h compactly supported in $[0,T)\times\mathbb{R}^d$. Taking the supremum over $|h|_\infty \le 1$ leads to $\|S^\Delta\| \le 1/2\|S^\Delta\|$. Since $\|S^\Delta\| \le T-t$, we deduce that $\|S^\Delta\| = 0$ which proves the result on $[0,T]$. Regular conditional probabilities allow to extend the result on $\mathbb{R}^+$, see e.g. Theorem 4, Chapter II, paragraph 7, in [Shi96]. □

3.4. **Proof of Lemma 2.2.**

In Subsection 3.1, we have obtained estimates for both the frozen density and the Parametrix Kernel. In this section, we expose how these estimates are used to deduce the convergence of the Parametrix series through the controls of Lemma 2.2.

Lemma 3.6. *Fix $t \in (0,T]$. There exists $C>1$, $\omega>0$ such that for all $(x,y)\in(\mathbb{R}^d)^2$:*

$$|\tilde{p}\otimes H(t,T,x,y)| \le C\Big((T-t)^\omega \bar{p}(t,T,x,y) + \rho(t,T,x,y)\Big),$$

where we recall the notation $\rho(t,T,x,y) = \delta \wedge |\theta_{t,T}(y)-x|^{\eta(\alpha\wedge 1)}\bar{p}(t,T,x,y)$.

Proof. From the upper bound on the Frozen density and the parametrix kernel H, we have:

$$\begin{aligned}|\tilde{p}\otimes H(t,T,x,y)| &\leq C\int_t^T d\tau\int_{\mathbb{R}^d}\bar{p}(t,\tau,x,z)\frac{\delta\wedge|z-\theta_{\tau,T}(y)|^{\eta(\alpha\wedge 1)}}{T-\tau}\bar{p}(\tau,T,z,y)dz\\ &\leq C\int_t^T d\tau\int_{\mathbb{R}^d}\frac{(\tau-t)^{-d/\alpha}}{\Big(1+\frac{|x-\theta_{t,\tau}(z)|}{(\tau-t)^{1/\alpha}}\Big)^{\alpha+\gamma}}\bar{q}(|x-\theta_{t,\tau}(z)|)\\ &\quad\times\frac{\delta\wedge|z-\theta_{\tau,T}(y)|^{\eta(\alpha\wedge 1)}}{T-\tau}\frac{(T-\tau)^{-d/\alpha}}{\Big(1+\frac{|z-\theta_{\tau,T}(y)|}{(T-\tau)^{1/\alpha}}\Big)^{\alpha+\gamma}}\bar{q}(|z-\theta_{\tau,T}(y)|)dz.\end{aligned}$$

Assume first that $|\theta_{t,T}(y)-x|\leq C(T-t)^{1/\alpha}$. Then, we split the time integral in $\int_t^{\frac{T+t}{2}}d\tau+\int_{\frac{T+t}{2}}^T d\tau$, and we use the fact that the Diagonal estimate is global. In the integral over $[\frac{T+t}{2},T]$ we have that $\tau-t\asymp T-t$, so that

$$\bar{p}(t,\tau,x,z)\leq(\tau-t)^{-d/\alpha}\asymp(T-t)^{-d/\alpha}\asymp\bar{p}(t,T,x,y).$$

Consequently, we take $\bar{p}(\tau,x,z)$ out of the integral and use the smoothing property of Lemma 3.5:

$$\begin{aligned}\bar{p}(t,T,x,y)\int_{\frac{T+t}{2}}^T d\tau\int_{\mathbb{R}^d}\frac{\delta\wedge|z-\theta_{\tau,T}(y)|^{\eta(\alpha\wedge1)}}{T-\tau}\frac{(T-\tau)^{-d/\alpha}}{\Big(1+\frac{|z-\theta_{\tau,T}(y)|}{(T-\tau)^{1/\alpha}}\Big)^{\alpha+\gamma}}\bar{q}(|z-\theta_{\tau,T}(y)|)dz\\ \leq C(T-t)^{\omega}\bar{p}(t,T,x,y).\end{aligned}$$

When, $\tau\in[0,\frac{T+t}{2}]$ we have $T-\tau\asymp T-t$, and we have

$$\frac{1}{T-\tau}\bar{p}(\tau,T,z,y)\leq C\frac{(T-\tau)^{-d/\alpha}}{T-\tau}\leq C\frac{(T-t)^{-d/\alpha}}{T-t}\leq C\frac{1}{T-t}\bar{p}(t,T,x,y).$$

Next, we can bound

$$\delta\wedge|z-\theta_{\tau,T}(y)|^{\eta(\alpha\wedge1)}\leq C_T(\delta\wedge|\theta_{t,\tau}z-x|^{\eta(\alpha\wedge1)}+\delta\wedge|x-\theta_{t,T}(y)|^{\eta(\alpha\wedge1)}).$$

Thus, we finally obtain:

$$\begin{aligned}\frac{1}{T-t}\bar{p}(t,T,x,y)\int_0^{\frac{T+t}{2}}d\tau\int_{\mathbb{R}^d}\frac{(\tau-t)^{-d/\alpha}}{\Big(1+\frac{|\theta_{t,\tau}(z)-x|}{(\tau-t)^{1/\alpha}}\Big)^{\alpha+\gamma}}\bar{q}(C|\theta_{t,\tau}(z)-x|)\\ \times\Big(\delta\wedge|\theta_{t,\tau}(z)-x|^{\eta(\alpha\wedge1)}+\delta\wedge|\theta_{t,T}(y)-x|^{\eta(\alpha\wedge1)}\Big)dz\\ \leq C\Big((T-t)^{\omega}+\delta\wedge|\theta_{t,T}(y)-x|^{\eta(\alpha\wedge1)}\Big)\bar{p}(t,T,x,y).\end{aligned}$$

Assume now that $|\theta_{t,T}(y)-x|\geq C(T-t)^{1/\alpha}$. In this case, the off-diagonal estimate holds for $\bar{p}(t,T,x,y)$, that is:

$$\bar{p}(t,T,x,y)\asymp\frac{(T-t)^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{t,T}(y)-x|^{\alpha+\gamma}}\bar{q}(|\theta_{t,T}(y)-x|).$$

On the other hand, we have:

$$|\theta_{t,T}(y)-x|\leq C_T\Big(|\theta_{t,\tau}(z)-x|+|\theta_{\tau,T}(y)-z|\Big)$$

In other words, we have either $|\theta_{\tau,T}(y)-z| \geq C|\theta_{t,T}(y)-x|$, or $|\theta_{t,\tau}(z)-x| \geq C|\theta_{t,T}(y)-x|$. Consequently, we split $\mathbb{R}^d = D_1 \cup D_2$ with

$$\begin{aligned} D_1 &= \{z \in \mathbb{R}^d, |\theta_{\tau,T}(y)-z| \leq |\theta_{t,\tau}(z)-x|\}, \\ D_2 &= \{z \in \mathbb{R}^d, |\theta_{\tau,T}(y)-z| > |\theta_{t,\tau}(z)-x|\}. \end{aligned}$$

Now, when $z \in D_1$, we have that $|\theta_{t,T}(y)-x| \asymp |\theta_{t,\tau}(z)-x|$, thus $\bar{p}(t,\tau,x,z)$ is off-diagonal and we can bound:

$$\begin{aligned} \bar{p}(t,\tau,x,z) &\leq C\frac{(\tau-t)^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{t,\tau}(z)-x|^{\alpha+\gamma}}\bar{q}(|\theta_{t,\tau}(z)-x|) \\ &\leq C\frac{(T-t)^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{t,T}(y)-x|^{\alpha+\gamma}}\bar{q}(|\theta_{t,T}(y)-x|) \asymp \bar{p}(t,T,x,y). \end{aligned}$$

For the last inequality, we used the fact that $\bar{q}$ is non increasing and that $\gamma+\alpha > d$ so that the exponent in $\tau - t$ is positive. Thus, we can take out $\bar{p}(t,\tau,x,z)$ of the integral, and use the smoothing property of H, Lemma 3.5. Denoting by I_{D_1} the convolution $|\tilde{p} \otimes H|$ where the space integration is over D_1, we have:

$$\begin{aligned} I_{D_1} &\leq C\bar{p}(t,T,x,y)\int_t^T d\tau \int_{D_1} \frac{\delta \wedge |\theta_{\tau,t}(y)-z|^{\eta(\alpha\wedge 1)}}{T-\tau}\frac{(T-\tau)^{-d/\alpha}}{\left(1+\frac{|\theta_{\tau,T}(y)-z|}{(T-\tau)^{1/\alpha}}\right)^{\alpha+\gamma}}\bar{q}(|\theta_{\tau,T}(y)-z|)dz \\ &\leq C(T-t)^{\omega}\bar{p}(t,T,x,y). \end{aligned}$$

When $z \in D_2$, we have $|\theta_{\tau,T}(y)-z| \asymp |\theta_{t,T}(y)-x|$. In this case, observe that we have:

$$\frac{1}{T-\tau}\bar{p}(T-\tau,z,y) \asymp \frac{(T-\tau)^{\frac{\gamma-d}{\alpha}}}{|\theta_{\tau,T}(y)-z|^{\alpha+\gamma}}\bar{q}(|\theta_{\tau,T}(y)-z|) \leq (T-\tau)^{\frac{\gamma-d}{\alpha}}\frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\alpha+\gamma}}.$$

Thus, the integral becomes:

$$\begin{aligned} I_{D_2} \leq \frac{\bar{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\alpha+\gamma}}\int_t^T d\tau(T-\tau)^{\frac{\gamma-d}{\alpha}}\int_{D_2} dz\bar{p}(t,\tau,x,z) \\ \times(\delta \wedge |\theta_{\tau,t}(y)-z|^{\eta(\alpha\wedge 1)} + \delta \wedge |\theta_{t,T}(y)-x|^{\eta(\alpha\wedge 1)}) \\ \leq C\Big((T-t)^{\omega} + \delta \wedge |\theta_{t,T}(y)-x|^{\eta(\alpha\wedge 1)}\Big)\bar{p}(t,T,x,y). \end{aligned}$$

To get the last inequality, we used Lemma 3.5 and integrated in time to recover $\bar{p}(t,T,x,y)$. In every case, we obtained the announced bound, thus the proof is complete.

□

The following Lemma controls the second step of the iterated convolutions.

Lemma 3.7. *Fix $t \in (0,T]$. There exists $C > 1$, $\omega > 0$ such that for all $(x,y) \in (\mathbb{R}^d)^2$:*

$$|\rho \otimes H(t,T,x,y)| \leq C(T-t)^{\omega}\bar{p}(t,T,x,y).$$

Proof. The proof is similar to the previous one, but now, due to the presence of $\delta \wedge |\theta_{t,\tau}(z)-x|^{\eta(\alpha\wedge 1)}$ multiplying the first density, we do not use the triangle inequality anymore, because we are always in position to use Lemma 3.5. □

3.5. **Proof of the Lower Bound.**

Observe first, that due to the controls on the Parametrix series, the convergence of the series actually yields a diagonal lower found for the density of $(X_t)_{t>0}$. Indeed, we have $p(t,T,x,y) = \tilde{p}(t,T,x,y) + p \otimes H(t,T,x,y)$. Also, we have the upper bound $p(t,T,x,y) \leq \bar{p}(t,T,x,y)$, which yields

$$\begin{aligned} p \otimes H(t,T,x,y) &\leq \int_t^T d\tau \int_{\mathbb{R}^d} \bar{p}(t,\tau,x,z) \frac{\delta \wedge |\theta_{\tau,T}(y) - z|^{\eta(\alpha \wedge 1)}}{T-\tau} \bar{p}(\tau,T,z,y) dz \\ &\leq \Big((T-t)^{\omega} + \delta \wedge |\theta_{t,T}(y) - x|^{\eta(\alpha \wedge 1)} \Big) \bar{p}(t,T,x,y). \end{aligned}$$

Thus, in diagonal regime: $|\theta_{t,T}(y) - x| \leq C(T-t)^{1/\alpha}$, we have for t small enough $p(t,T,x,y) \geq C(T-t)^{-d/\alpha}$. In other words, we have a diagonal lower bound for the density of (1.1).

We now turn to the off-diagonal regime. The idea to derive a lower bound for the density is to say that in order to go from x to y in time $T-t$, we stay close to the transport of x by the deterministic system, for a certain amount of time, then, a big jump brings us to a neighborhood of the pull back of y by the deterministic system and the process stays in a neighborhood of this curve.

In the off-diagonal regime: $|\theta_{t,T}(y) - x| \geq C(T-t)^{1/\alpha}$, we write from the Chapman-Kolmogorov equation for some $t_0 \in [t,T]$:

$$\begin{aligned} p(t,T,x,y) &= \int_{\mathbb{R}^d} dz p(t,t_0,x,z) p(t_0,Tz,y) \geq \int_{B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})} p(t,t_0,x,z) p(t_0,T,z,y) dz \\ &\geq \mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}) \Big| X_t = x \Big) \inf_{z \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})} p(t_0,T,z,y) \\ &\geq \mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}) \Big| X_t = x \Big) C(T-t_0)^{-d/\alpha}. \end{aligned}$$

Consequently we have to give a lower bound for $\mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}) \Big| X_t = x \Big)$. To this end, we introduce the process $(X_t^{\delta})_{t \geq 0}$ with jumps larger than δ removed, for some δ to be specified. Specifically, $(X_s^{\delta})_{s \geq 0}$ solves the SDE:

$$X_s^{\delta} = x + \int_t^s F(u, X_u^{\delta}) du + \int_t^s \sigma(u, X_u^{\delta}) dZ_u^{\delta},$$

where $(Z_u^{\delta})_{u \geq 0}$ is the process $(Z_u)_{u \geq 0}$ with jumps larger that δ removed. Its Lévy measure is $\mathbf{1}_{\{|z| \leq \delta\}} \nu(dz)$. Now, observe that we can recover the process $(X_s)_{s \geq 0}$ from $(X_s^{\delta})_{s \geq 0}$ by introducing the arrival times of the compound poisson process:

$$N_s = \sum_{t < u \leq s} \Delta Z_u \mathbf{1}_{\{|\Delta Z_u| \geq \delta\}}.$$

Let us denote by $(T_k)_{k \geq 1}$ the arrival times of the process $(N_s)_{s \in [t,T]}$. We know that the variables $T_{k+1} - T_k$ are independent and have exponential distribution of parameter $\nu\big(B(0,\delta)^c\big)$. Then, we have:

$$\begin{aligned} &\forall t \leq s \leq T_1, \ X_s = X_s^{\delta}, \\ &X_{T_1} = X_{T_1^-}^{\delta} + \sigma(X_{T_1^-}^{\delta}) \Delta Z_{T_1}, \\ &\forall T_1 \leq s \leq T_2, \ X_s = X_{T_1} + X_s^{\delta} - X_{T_1^-}^{\delta}, \end{aligned}$$

and so on. We refer to the Theorem 6.2.9 in Applebaum [App09] for a proof of this statement. We now split:

$$\begin{aligned} \mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}) \Big| X_t = x \Big) &= \mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \geq t_0 \Big| X_t = x \Big) \\ &\quad + \mathbb{P}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \leq t_0 \Big| X_t = x \Big). \end{aligned}$$

Using the Markovian notations $\mathbb{P}^{t,x}(\cdot) = \mathbb{P}(\cdot|X_t = x)$, we thus focus on:

$$\begin{aligned}&\mathbb{P}^{t,x}\Big(X_{t_0} \in B\big(\theta_{t_0,t}(y), C(T-t_0)^{1/\alpha}\big); T_1 \le t_0\Big)\\ =\;& \mathbb{E}^{t,x}\left[\mathbb{P}^{t,x}\Big(X_{t_0} \in B\big(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}\big)\Big|\mathcal{F}_{T_1}\Big)\mathbf{1}_{\{T_1\le t_0\}}\right],\end{aligned}$$

where we denoted $\mathcal{F}_{T_1} = \sigma(X^\delta_s; s \le T_1)$, the filtration generated by X^δ_s until time T_1. Now, by the strong Markov property, we have that

$$\begin{aligned}\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t)^{1/\alpha})\Big|\mathcal{F}_{T_1}\Big) &= \mathbb{P}^{T_1,X_{T_1}}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})\Big)\\ &= \int_{B(\theta_{t_0,T}(y),C(T-t_0)^{1/\alpha})} p(T_1,t_0,X_{T_1},z)dz.\end{aligned}$$

Thus:

$$\begin{aligned}&\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t)^{1/\alpha}); T_1 \le t_0\Big)\\ =\;& \mathbb{E}^{t,x}\left[\int_{B(\theta_{t_0,T}(y),C(T-t_0)^{1/\alpha})} p(T_1,t_0,X_{T_1},z)dz\mathbf{1}_{\{T_1\le t_0\}}\right].\end{aligned}$$

Now, since $X_{T_1} = X^\delta_{T_1^-} + \sigma(X^\delta_{T_1^-})\Delta Z_{T_1}$, and since T_1 is the first jump larger that δ, conditionally to $X_{T_1^-}$ we have that $\sigma(X_{T_1^-})\Delta Z_{T_1} + X_{T_1^-}$ is a Poisson process on $\mathbb{R}^d\backslash B(0,\delta)$. Thus, we have for all test function f, given $X_{T_1^-}$, the law of X_{T_1} is:

$$\begin{aligned}\mathbb{E}[f(X_{T_1})|X_{T_1^-}] &= \mathbb{E}[f(\sigma(X_{T_1^-})\Delta Z_{T_1} + X_{T_1^-})|X_{T_1^-}]\\ &= \int_{\{|w|\ge\delta\}} f\big(\sigma(X_{T_1^-})w + X_{T_1^-}\big)\frac{\nu(dw)}{\nu\big(B(0,\delta)^c\big)}.\end{aligned}$$

Consequently, we obtain:

$$\begin{aligned}&\mathbb{E}^{t,x}\left(\int_{B(\theta_{t_0,T}(y),C(T-t_0)^{1/\alpha})} p(T_1,t_0,X_{T_1},z)dz\Big|X_{T_1^-}\right)\\ &= \int_{\mathbb{R}^d}\int_{B(\theta_{t_0,T}(y),C(T-t_0)^{1/\alpha})} dz\; p(T_1,t_0,\sigma(X^\delta_{T_1^-})w + X_{T_1^-}\delta, z)\frac{\nu(dw)}{\nu\big(B(0,\delta)^c\big)}.\end{aligned}$$

Now, we exploit the fact that T_1 in independent and exponentially distributed with parameter $\nu(B(0,\delta)^c)$ to write:

$$\begin{aligned}\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t)^{1/\alpha}); T_1 \le t_0\Big) &= \mathbb{E}^{t,x}\Big[\int_t^{t_0} ds\int_{B(\theta_{t_0,T}(y),C(T-t_0)^{1/\alpha})} dz\int_{\mathbb{R}^d}\frac{\nu(dw)}{\nu\big(B(0,\delta)^c\big)}\\ &\quad\times p(s,t_0,\sigma(X^\delta_s)w + X^\delta_s, z)\nu\big(B(0,\delta)^c\big)e^{-s\nu\left(B(0,\delta)^c\right)}\Big].\end{aligned}$$

Observe that the quantity $\nu\big(B(0,\delta)^c\big)$ gets cancelled. Now, we can give a lower bound by localizing the integral over w so that $\sigma(X^\delta_s)w + X^\delta_s$ is close to $\theta_{s,t_0}(z)$. That is, where the density $p(s,t_0,\sigma(X^\delta_s)w + X^\delta_s, z)$ is in diagonal regime:

$$
\begin{aligned}
&\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \le t_0\Big)\\
\ge\ & \mathbb{E}^{t,x}\Big[\int_t^{t_0} ds \int_{B(\theta_{t_0,T}(y), C(T-t_0)^{\frac{1}{\alpha}})} dz \int_{\{|\sigma(X_s^\delta)w + X_s^\delta - \theta_{s,t_0}(z)| \le C(t_0-s)^{\frac{1}{\alpha}}\}} \nu(dw)\\
&\times p\Big(s, t_0, \sigma(X_s^\delta)w + X_s^\delta, z\Big) e^{-s\nu(B(0,\delta)^c)}\Big]\\
\ge\ & \mathbb{E}^{t,x}\Big[\int_t^{t_0} ds \Big(t_0 - s\Big)^{-d/\alpha} \int_{B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})} dz\\
&\times \nu\Big(B\Big(\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta), C\,(t_0-s)^{1/\alpha}\Big)\Big) e^{-s\nu(B(0,\delta)^c)}\Big].
\end{aligned}
$$

Additionally, we can lower bound the last probability by localizing X_s^δ close to $\theta_{s,t}(x)$:

$$
\begin{aligned}
&\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t)^{1/\alpha}); T_1 \le t_0\Big)\\
\ge\ & \mathbb{E}^{t,x}\Big[\int_t^{t_0} ds \mathbf{1}_{\{|X_s^\delta - \theta_{s,t}(x)| \le C(s-t)^{1/\alpha}\}} \Big(t_0 - s\Big)^{-d/\alpha} \int_{B(\theta_{t_0,t}(y), C(T-t_0)^{1/\alpha})} dz\\
&\times \nu\Big(B\Big(\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta), C\,(t_0-s)^{1/\alpha}\Big)\Big) e^{-s\nu(B(0,\delta)^c)}\Big].
\end{aligned}
$$

Now, from assumption **[H-LB]**, $\nu(B(0,\delta)^c) \le 1/\delta^\alpha$ so that taking $\delta = (T-t)^{1/\alpha}$ yields $e^{-s\nu(B(0,\delta)^c)} \ge C$. Also, since $z \in B(\theta_{t_0,T}(y), C(T-t_0)^{\frac{1}{\alpha}})$, by the Lipschitz property of the flow,

$$
\theta_{s,t_0}(z) \in \theta_{s,t_0}\Big(B(\theta_{t_0,T}(y), C(T-t_0)^{\frac{1}{\alpha}})\Big) \subset B(\theta_{s,T}(y), C(T-t_0)^{\frac{1}{\alpha}}),
$$

up to a modification of C for the last inclusion. On the other hand, $X_s^\delta \in B(\theta_{s,t}(x), s^{1/\alpha})$, thus,

$$
\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta) \in B\big(\sigma(X_s^\delta)^{-1}(\theta_{t,T}(y) - x), C(T-t)^{1/\alpha}\big).
$$

Now, using the Hölder continuity of σ, and the localization $|X_s^\delta - \theta_{s,t}(x)| \le (s-t)^{1/\alpha}$, we have up to a modification of C:

$$
B\big(\sigma(X_s^\delta)^{-1}(\theta_{t,T}(y) - x), C(T-t)^{1/\alpha}\big) \subset B\big(\sigma(\theta_{s,t}(x))^{-1}(\theta_{t,T}(y) - x), C(T-t)^{1/\alpha}\big).
$$

so that we obtain:

$$
\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta) \in B\big(\sigma(\theta_{s,t}(x))^{-1}(\theta_{t,T}(y) - x), C(T-t)^{1/\alpha}\big)
$$

Thus, if we have:

$$
\forall s \in [t, t_0],\ B\big(\sigma(\theta_{s,t}(x))^{-1}(\theta_{t,T}(y) - x), C(T-t)^{1/\alpha}\big) \subset A_{low},
$$

then $\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta) \in A_{low}$. Observe that this is exactly the condition (1.19) of Theorem 1.2.

Consequently, we can use the lower bound in **[H-LB]** to get:

$$
\nu\Big(B\Big(\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta), C\,(t_0-s)^{1/\alpha}\Big)\Big) \ \ge\ C\,(t_0-s)^{\gamma/\alpha}\,\frac{\underline{q}(|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)|)}{|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)|^{\gamma+\alpha}}.
$$

We thus obtain:

$$\begin{aligned}
&\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \leq t_0\Big) \\
\geq\ & C\mathbb{E}^{t,x}\Big[\int_t^{t_0} ds \mathbf{1}_{\{|X_s^\delta - \theta_{s,t}(x)| \leq C(s-t)^{1/\alpha}\}} (t_0-s)^{\frac{\gamma-d}{\alpha}} \\
&\times \int_{B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})} dz \frac{\underline{q}(|\sigma(X_s^\delta)^{-1}(\theta_{t,s}(z) - X_s^\delta)|)}{|\sigma(X_s^\delta)^{-1}(\theta_{t,s}(z) - X_s^\delta)|^{\gamma+\alpha}}\Big].
\end{aligned}$$

Consequently, since the function $u \mapsto \underline{q}(u)|u|^{-\gamma-\alpha}$ is decreasing, the lower bound will follow from the upper bound:

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)| \leq C|y - \theta_{T,t}(x)|.$$

We write from the ellipticity of σ:

$$\begin{aligned}
|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)| &\leq C|\theta_{s,t_0}(z) - X_s^\delta| \\
&\leq C(|\theta_{s,t_0}(z) - \theta_{s,t}(x)| + |\theta_{s,t}(x) - X_s^\delta|).
\end{aligned}$$

Now, in the considered set, $|\theta_{s,t}(x) - X_s^\delta| \leq C(s-t)^{1/\alpha} \leq C(T-t)^{1/\alpha} \leq C|\theta_{t,T}(y) - x|$. Thus, we have:

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)| \leq C\big(|\theta_{s,t_0}(z) - \theta_{s,t}(x)| + C|\theta_{t,T}(y) - x|\big).$$

On the other hand, we can write:

$$|\theta_{s,t_0}(z) - \theta_{s,t}(x)| \leq |\theta_{s,t_0}(z) - \theta_{s,T}(y)| + |\theta_{s,T}(y) - \theta_{s,t}(x)|.$$

Thus, from the Lipschitz property of the flow,

$$|\theta_{s,T}(y) - \theta_{s,t}(x)| \leq C_T|\theta_{t,T}(y) - x|.$$

On the other hand, we have $\theta_{s,T}(y) = \theta_{s,t_0} \circ \theta_{t_0,T}(y)$ so that:

$$|\theta_{s,t_0}(z) - \theta_{s,T}(y)| = |\theta_{s,t_0}(z) - \theta_{s,t_0} \circ \theta_{t_0,T}(y)| \leq C_T|z - \theta_{t_0,T}(y)|,$$

where to the get the last inequality, we once again relied on the Lipschitz property of the flow. We recall that $|z - \theta_{t_0,T}(y)| \leq C(T-t_0)^{1/\alpha} \leq |\theta_{t,T}(y) - x|$, consequently we finally obtain:

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t_0}(z) - X_s^\delta)| \leq C_T|\theta_{t,T}(y) - x|.$$

Using this last inequality to estimate the probability:

$$\begin{aligned}
&\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t,t_0}, C(T-t)^{1/\alpha}); T_1 \leq t_0\Big) \\
\geq\ & C\mathbb{E}^{t,x}\left[\int_t^{t_0} ds \mathbf{1}_{\{|X_s^\delta - \theta_{s,t}(x)| \leq (s-t)^{1/\alpha}\}} (t_0-s)^{\frac{\gamma-d}{\alpha}} \int_{B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})} dz \frac{\underline{q}(|\theta_{t,T}(y) - x|)}{|\theta_{t,T}(y) - x|^{\gamma+\alpha}}\right] \\
\geq\ & C(T-t_0)^{d/\alpha} \frac{\underline{q}(|\theta_{t,T}(y) - x|)}{|\theta_{t,T}(y) - x|^{\gamma+\alpha}} \int_t^{t_0} ds\, (t_0-s)^{\frac{\gamma-d}{\alpha}}\, \mathbb{P}(|X_s^\delta - \theta_{s,t}(x)| \leq (s-t)^{1/\alpha}),
\end{aligned}$$

where $(T-t_0)^{d/\alpha}$ comes from the volume of the ball $B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha})$ obtained from the integral in dz. Using the diagonal lower estimates for the density, we actually see that $\mathbb{P}(|X_s^\delta - \theta_{s,t}(x)| \leq (s-t)^{1/\alpha}) \asymp 1$, therefore:

$$\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \leq t_0\Big) \geq C(T-t_0)^{d/\alpha}(t_0-t)^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\gamma+\alpha}}.$$

Returning to the first estimate on the density yields:

$$\begin{aligned} p(t,T,x,y) &\geq C(T-t_0)^{-d/\alpha}\mathbb{P}^{t,x}\Big(X_{t_0} \in B(\theta_{t_0,T}(y), C(T-t_0)^{1/\alpha}); T_1 \leq t_0\Big) \\ &\geq C(T-t_0)^{-d/\alpha}(T-t_0)^{d/\alpha}(t_0-t)^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\gamma+\alpha}} \\ &= C(t_0-t)^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\gamma+\alpha}}. \end{aligned}$$

Finally, to get the announced bound, we see that we have to choose t_0 such that $t_0 - t \asymp T - t$. This gives:

$$p(t,T,x,y) \geq C(T-t)^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{t,T}(y)-x|)}{|\theta_{t,T}(y)-x|^{\gamma+\alpha}}.$$

which is the off-diagonal lower bound announced.

Remark 3.3. We point out that the assumption **[H-LB]** appears quite naturally in this procedure as it serves here to give a lower bound on the ν-measure of balls. Also, the time t_0 is pretty much arbitrary, but still has to be comparable to $T-t$. Heuristically, in (1.19), this means that the process $(X_s)_{s\in[t,T]}$ has to stay an amount of time comparable to $T-t$ in A_{low} in order to have a lower bound.

References

[App09] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[BB82] R. Bafico and P. Baldi. Small random perturbations of peano phenomena. *Stochastics*, 6(3-4):279–292, 1982.

[BGJ87] K. Bichteler, J. B. Gravereaux, and J. Jacod. *Malliavin calculus for processes with jumps*. Gordon and Breach Science Publishers, January 1987.

[BS07] K. Bogdan and P. Sztonyk. Estimates of the potential kernel and Harnack's inequality for the anisotropic fractional Laplacian. *Studia Math.*, 181(2):101–123, 2007.

[Cas09] Thomas Cass. Smooth densities for solutions to stochastic differential equations with jumps. *Stochastic Process. Appl.*, 119(5):1416–1435, 2009.

[DF14] F. Delarue and F. Flandoli. The transition point in the zero noise limit for a 1d Peano example. *Discrete Contin. Dyn. Syst*, 34:4071–4083, 2014.

[Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.

[HM14] L. Huang and S. Menozzi. Density bounds for some degenerate stable driven sdes. *arXiv:1402.3997*, 2014.

[IK06] Yasushi Ishikawa and Hiroshi Kunita. Malliavin calculus on the Wiener-Poisson space and its application to canonical SDE with jumps. *Stochastic Process. Appl.*, 116(12):1743–1769, 2006.

[Kol00] V. Kolokoltsov. Symmetric stable laws and stable-like diffusion. *Proceedings of the London Mathematical Society*, 80(03):725–768, 2000.

[KS84] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. I. *Stochastic analysis (Katata/Kyoto, 1982)*, North-Holland Math. Library, 32:271–306, 1984.

[KS85] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math*, 32:1–76, 1985.

[KS87] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math*, 34:391–442, 1987.

[KS13] K. Kaleta and P. Sztonyk. Upper estimates of transition densities for stable-dominated semigroups. *J. Evol. Equ.*, 13(3):633–650, 2013.

[Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate kolmogorov equations. *Electronic Communications in Probability*, 16:234–250, 2011.

[MKS67] H. Mc Kean and I. Singer. Curvature and eigen values of the Laplacian. *J. Differential Geometry*, pages 43 – 69, 1967.

[Nor86] James Norris. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 101–130. Springer, Berlin, 1986.

[Pic96] J. Picard. On the existence of smooth densities for jump processes. *Probability Theory and Related Fields*, 105(4):481–511, 1996.

[Sat05] K. Sato. *Lévy processes and Infinitely divisible Distributions*. Cambridge University Press, 2005.

[She91] S. J. Sheu. Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.*, 19–2:538–561, 1991.

[Shi96] A.N. Shiryaev. *Probability, Second Edition.* Graduate Texts in Mathematics, 95. Springer-Verlag, New York., 1996.

[Str75] D. W. Stroock. Diffusion processes associated with Lévy generators. *Probability Theory and Related Fields*, 32(3):209–244, 1975.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.

[Szt10] Paweł Sztonyk. Estimates of tempered stable densities. *J. Theoret. Probab.*, 23(1):127–147, 2010.

[Wat07] T. Watanabe. Asymptotic estimates of multi-dimensional stable densities and their applications. *Transactions of the American Mathematical Society*, 359(6):2851–2879, 2007.

[Zha14] Xicheng Zhang. Densities for SDEs driven by degenerate α-stable processes. *Ann. Probab.*, 42(5):1885–1910, 2014.