Stable-driven SDEs: a Parametrix Approach to Heat Kernel Estimates with an Application to Stochastic Algorithms

Lorick Huang

Université Paris Diderot (Paris 7) Sorbonne Paris Cité

École Doctorale ED386: Sciences Mathématiques de Paris Centre

THÈSE DE DOCTORAT

Discipline: Mathématiques Appliquées

Présentée par

Lorick HUANG

EDS dirigées par des processus stables. Méthode paramétrix pour des estimées de densités et application aux algorithmes stochastiques.

Thèse dirigée par Stéphane MENOZZI

soutenue le 3 Juillet 2015

Jury

François Delarue	Examinateur
Arnaud Gloter	Examinateur
Valentin Konakov	Examinateur
Stéphane Menozzi	Directeur de Thèse
Gilles Pagès	Examinateur
Thomas Simon	Rapporteur
Denis Talay	Examinateur
Peter Tankov	Examinateur

Au vu des rapports de

Vassili Kolokolstov
Thomas Simon

Remerciements

Je prends ces quelques lignes pour remercier toutes les personnes qui m'ont soutenu pendant ces années de thèse.

Je voudrais tout d'abord remercier mon directeur de thèse, Stéphane Menozzi pour m'avoir guidé pendant ces année de theses. de m'avoir introduit au monde de la recherche mathématiques, avec ses hauts et ses bas ! toutes ces nuits à prendre le dernier RER (ou metro ce qui est plus grave) paraissent bien loin maintenant !

Je remercie les rapporteurs de cette thèse, messieurs Thomas Simon et Vassili Kolokoltsov, pour leur relecture minutieuse et leur commentaires pertinents.

Merci à messieurs Valentin Konakov, Arnaud Gloter, Denis Talay, François Delarue, Gilles Pages, et Peter Tankov d'avoir accepté de faire partie de mon jury.

Je tiens également à remercier Noufel Frikha, pour son encadrement au quotidien au LPMA et sa bienveillance envers les doctorants en général.

Merci à Arturo Kohatsu-higa, et à toute son équipe pour son accueil au japon. Merci a Yui, Taguchi, Yuuki et Tran pour ce magnifique séjour. J'espère que nous aurons l'occasion de nous revoir.

Merci à Valentin Konakov pour son accueil à Moscou. Ce fut une semaine enrichissante, je suis très enthousiaste à l'idée de collaborer l'année prochaine.

Merci à Peter Tankov pour le voyage en Argentine, et plus généralement, merci au LPMA, pour ce cadre exceptionnel. Merci Valerie, Nathalie et Florence, et merci à l'UFR de mathématiques en la personne de Pascal Chiettini.

Merci aux personnels enseignants avec qui j'ai pu travailler, Aurélie, Raphael et Gabrielle.

Merci aux doctorants du LPMA pour tous ces moments de convivialité autour d'un café ou d'une bière ! Par ordre d'ancienneté : Christophe, Oriane, Thomas, Maud, Marc-antoine, Bastien, Guillaume, Aser, Jiatu, Shanqiu, Huy, Sophie, Clément, Adrien, Guillaume, Arturo, Anna, Vu-lan, Olga, Sarah, David, Thomas, et tous ceux que j'ai oublié. So long and thanks for all the chocolate!

Je remercie mes parents pour leur soutient précieux tout au long de mes études, et plus généralement pour tout ce qu'ils ont fait pour moi jusqu'ici.

Merci aussi à mes amis de longue date Vai et Mathieu qui sont toujours disponibles quand il s'agit de manger!

Je remercie également la grande famille de France Haidong Gumdo sans qui ces années de thèses auraient été beaucoup plus dures à traverser. merci donc à Maitre Capozzi, à Franck, à toute l'école des cadres et des assistants, à mon club de paris 8, et à tous les pratiquants des autres clubs que j'ai pu rencontrer.

Enfin, un grand merci à Sandrine, pour sa présence au quotidien, son soutien et ses conseils.

Contents

Résumé détaillé

1 Résumé en français

1.1 Estimées de densité pour des EDS dirigées par des processus stables tempérés.

Nous étudions une classe d'équations différentielles stochastiques dirigées par des processus stables (possiblement tempérés), sous des hypothèses de régularité Hölder sur les coefficients. Nous prouvons que le problème de martingale associé est bien posé, établissant ainsi l'unicité faible pour l'EDS. Nous donnons aussi un encadrement de la densité de la solution par celle d'un processus stable (possiblement tempéré). Notre approche est basée sur la méthode parametrix.

1.2 La méthode parametrix pour des EDS dégénérées dirigées par des processus stables.

Nous considérons une équation différentielle stochastique dégénérée dirigée par un processus stable dont les coefficients satisfont une sorte d'hypothèse de Hörmander faible. Sous de relativement faibles hypothèses de régularité et des restrictions dimensionnelles, nous prouvons que le problème de martingale est bien posé. Nous donnons également un majorant de la densité reflétant le caractère multi-échelle du processus sous-jacent dans le cas scalaire du stable tempéré.

1.3 Une extrapolation à pas multiples de Richardson-Romberg pour l'approximation stochastique

Nous obtenons un développement pour l'erreur de discrétisation de la cible d'un algorithme stochastique à la suite de [Fri13]. Ceci nous permet de mettre en place une extrapolation de Richardson-Romberg dans le cadre des algorithmes stochastiques, déjà obtenue pour les estimateurs de Monte Carlo linéaires (introduite par Talay et Tubaro [TT90] et pleinement étudiée dans Pagès [Pag07]). Nous appliquons nos résultats à l'estimation du quantile de la solution d'une EDS dirigée par un processus stable. Les résultats numériques produits à partir de notre méthode montrent une réduction significative de la complexité.

2 Summary in English

2.1 Density Estimates for SDEs Driven by Tempered Stable Processes

We study a class of stochastic differential equations driven by a possibly tempered Lévy process, under mild conditions on the coefficients (Hölder continuity). 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.

2.2 A Parametrix Approach for some Degenerate Stable Driven SDEs

We consider a stable driven degenerate stochastic differential equation, whose coefficients satisfy a kind of weak Hörmander condition. Under mild smoothness assumptions we prove the uniqueness of the martingale problem for the associated generator under some dimension constraints. Also, when the driving noise is scalar and tempered, we establish density bounds reflecting the multi-scale behavior of the process.

2.3 A Multi-step Richardson-Romberg extrapolation method for stochastic approximation

We obtain an expansion of the implicit weak discretization error for the target of stochastic approximation algorithms introduced and studied in [Fri13]. This allows us to extend and develop the Richardson-Romberg extrapolation method for Monte Carlo linear estimator (introduced in [TT90] and deeply studied in [Pag07]) to the framework of stochastic optimization by means of stochastic approximation algorithms. We notably apply the method to the estimation of the quantile of diffusion processes. Numerical results confirm the theoretical analysis and show a significant reduction in the initial computational cost.

Chapter 1

Introduction

1 The Problem

This work deals with density estimates for the solution of some Stochastic Differential Equations (SDEs). Let us consider an SDE of the form:

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

where $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$. Such equations appear in various fields, from Hamiltonian mechanics (see e.g. Talay [Tal02]), to finance (see for instance Jeanblanc *et al* [JYC09]). Also, a strong connection exists between (1.1) and certain partial differential equations. Indeed, the density of (1.1) (when it exists) is the so-called fundamental solution (also known as Heat Kernel) associated with the operator $\partial_t + L_t$, where L_t is the generator of (1.1). For instance, when $(Z_t)_{t\geq 0}$ is an $\mathbb{R}^d$ valued Brownian motion, equation (1.1) can be linked to the solution of a diffusion equation, and more generally, when $(Z_t)_{t\geq 0}$ is a Lévy process, we can link the solution of (1.1) to an integro-differential equation. Thus giving density estimates on the solution of (1.1) is relevant both for the theory and the applications.

Assume first that $(Z_t)_{t\geq 0}$ is an $\mathbb{R}^d$ valued Brownian motion. Consider the following elliptic setting: $\exists C > 1$, $\forall x, \xi \in \mathbb{R}^d$, $\forall t > 0$,

$$C^{-1}|\xi|^2 \leq \langle \xi, \sigma\sigma^*(t,x)\xi \rangle \leq C|\xi|^2, \tag{1.2}$$

where σ^* stands for the transpose of σ. Suppose as well that $\sigma\sigma^*$ is Hölder continuous and b is bounded Borel. It is then known from the works of Friedman [Fri64] that $(X_t)_{t\geq 0}$ has a density, $\mathbb{P}(X_s \in dy | X_t = x) = p(t,s,x,y)dy$. Moreover, two sided Gaussian estimates hold uniformly on compact sets in time. Precisely, for all $T > 0$, $\exists C_1, C_2 \geq 1$, where C_1, C_2 depend on the non degeneracy constant appearing in (1.2) and the drift b, and C_1 depends as well on T, such that, for all $x, y \in \mathbb{R}^d$, $0 \le t < s \le T$:

$$\frac{C_1^{-1}}{(s-t)^{d/2}} \exp\left(-C_2 \frac{|x-y|^2}{s-t}\right) \le p(t,s,x,y) \le \frac{C_1}{(s-t)^{d/2}} \exp\left(-C_2^{-1} \frac{|x-y|^2}{s-t}\right). \tag{1.3}$$

We refer to the last inequality as Aronson estimates: two sided estimates that reflect the nature of the noise in the system (see [Aro59]). In other words, the estimate on the noise is transmitted to the solution of the SDE, and the marginals of the solution of (1.1) have the same asymptotic behavior as the driving Brownian motion. Historically, such results are derived from continuity techniques. Intuitively, when the coefficients have some regularity, we can expect the solution of (1.1) to have the same behavior as the driving process, at least for small times. This is exactly the information in the Aronson estimates.

Observe that the regularity is here crucial. We refer to the well known counter examples by Krylov (see e.g. Section 3 of chapter V in Bass [Bas97]): for $d \ge 2$, the radial part of an SDE of the form (1.1), with σ satisfying (1.2), but non continuous at zero can behave as a Bessel process of low order so that zero can be hit.

A lot of different techniques have been developed to recover the estimates in (1.3). When the coefficients b and σ are smooth and satisfy some non degeneracy conditions (Hörmander setting), possibly weaker than (1.2), Malliavin calculus can be used to prove existence and estimates on the density. This approach is thoroughly presented in the series of papers by Kusuoka and Stroock [KS84, KS85, KS87], which can be seen as a "masterwork" in this field. In particular, Aronson estimates under the strong Hörmander assumption are established in [KS87]. We also mention under (1.2) the control based method successfully used in Sheu [She91] to recover (1.3).

The uniform elliptic assumption intuitively means that the diffusion coefficient σ diffuses the noise in the whole $\mathbb{R}^d$ space. However, this condition is not always fulfilled in some practical cases: consider for instance, in the Hamiltonian setting, the Langevin equations (speed/position dynamics) with a perturbation on the speed component, or in financial mathematics, the Asian option pricing. Nevertheless, when the so-called weak Hörmander condition is satisfied, we can recover similar results. The weak Hörmander condition intuitively states that with the help of the drift coefficient b, the noise fills the space $\mathbb{R}^d$ (see Norris [Nor86] for more details). In the work of Delarue and Menozzi [DM10], the authors consider a chain of n differential equations, and only the first component is affected by a Brownian noise. Under the weak Hörmander assumption, the noise propagates through the system, giving existence for the density. Two sided Gaussian (multi-scale) estimates are derived as well.

Let us now consider the case where the driving process $(Z_t)_{t\ge0}$ in (1.1) is a general Lévy process. In order to obtain a density estimate on the solution of the SDE driven by $(Z_t)_{t\ge0}$, the first step would be getting one for the driving process. However, even establishing the absolute continuity of the law of Z_t can already be a hard task. For infinitely divisible distributions (without Gaussian part), giving general a criterion for absolute continuity in terms of the associated Lévy measures is still an open problem for which no sensible conjecture has been formulated yet, even in the one dimensional case. The difficulty comes from the time dependency of the absolute continuity property. Some Lévy processes even show drastic temporal evolution, from discrete continuous to absolute continuity. The multivariate setting is even more intricate, as the geometry of the space plays an important role. We refer to Watanabe [Wat01] for an overview of these aspects. See also the paper of Yamazato [Yam11] that gives some geometric conditions for the absolute continuity of multivariate Lévy processes.

Nevertheless, existence of the density of an SDE driven by a Lévy process can be derived with Malliavin calculus. A survey of the extension to the jump case of the Malliavin calculus for diffusions (see e.g. [Nor86]) can be found in Bichteller Gravereaux and Jacod [BGJ87]. Therein, some technical restrictions are imposed, such as the non degeneracy and smoothness of the Lévy measure and existence of moments. The key tool still relies on integration by parts formulas that are established through suitable perturbations in the amplitude of the jumps, either following the approach of Bismut based on a Girsanov transform, or following the Malliavin-Stroock operator-based techniques. We also mention the work of Picard [Pic96], who proves absolute continuity results for SDEs when the Lévy measure of the driving process can be singular (but yet non degenerate). In addition, let us indicate the paper of Léandre [Léa88b], who develops a Malliavin calculus for Lévy processes that can be written as subordinated Brownian motions. We insist that in the previous works the dimension of the non degenerate noise corresponds to the that of the underlying space.

When the noise can be in a strict subset, assumptions on the coefficients have to be made so that the noise effectively transmits to the whole space in order to have the existence of the density. Let us indicate the work of Cass [Cas09] which might be seen as the most general, yet incomplete, extension of Hörmander's theorem for SDEs driven by (jump) Lévy processes. See also, Simon [Sim11], who obtains a sharp geometric characterisation for the absolute continuity of degenerate Lévy driven Ornstein-Uhlenbeck processes. In a similar framework, let us also refer to the Fourier-based approach of Priola and Zabczyk [PZ09]. Eventually, the subordinated Malliavin calculus has been extended to prove absolute continuity in the weak Hörmander degenerate setting by Zhang [Zha14b].

As we mentioned above, the sole existence of the density for the solution of the SDE is complicated. Thus, giving an estimate on the density for a general Lévy process seems impossible. In some specific cases, though, density estimates can be derived. When $(Z_t)_{t\geq 0}$ is a symmetric α-stable process, $0 < \alpha < 2$ (the case $\alpha = 2$ corresponds to the Brownian motion that we already mentioned), the Lévy exponent of $(Z_t)_{t\geq 0}$

writes:

$$\mathbb{E}(e^{i\langle p,Z_t\rangle}) = \exp\left(-t\int_0^{+\infty}\int_{S^{d-1}}\left(e^{i\langle p,s\theta\rangle}-1-i\langle p,s\theta\rangle\mathbf{1}_{\{|s|\le 1\}}\right)\frac{C_{d,\alpha}ds}{s^{1+\alpha}}\mu(d\theta)\right). \quad (1.4)$$

In the above equation, $\mu(d\theta)$ is called the *spectral measure*. It is a finite measure on the sphere S^{d-1}, and $C_{d,\alpha}$ is a positive constant (see the exact value in Sato [Sat05]). Equation (1.1) in the context of a stable perturbation has been studied by Kolokoltsov [Kol00b]. The author established under some regularity assumptions on b and σ and the strict positivity of the density of the spectral measure (which can be seen as an additional non degeneracy condition) that for all $T>0$, there exists $C>1$, depending on T and the non degeneracy conditions such that for all $x,y\in\mathbb{R}^d$; $0\le t<s\le T$:

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

We see that in this case as well, two sided Stable estimates hold for the density of (1.1). Thus, the solution of the SDE (1.1) once again has the same global behavior as the driving noise, in this case a Stable process, for its marginal densities.

From here on, one can wonder if the Aronson estimates hold for other types of noise, in the Lévy case. We consider that the driving process $(Z_t)_{t\ge0}$ satisfies the following condition:

$$\exists K>0,\ \forall (t,p)\in\mathbb{R}_+\times\mathbb{R}^d,\ \mathbb{E}(e^{i\langle p,Z_t\rangle})\le e^{-Kt|p|^\alpha},$$

yielding the existence of the density of Z_t, for $t>0$. In this work, we investigate two specific cases.

- The tempered stable driven SDE, that is when the Lévy measure of $(Z_t)_{t\ge0}$ is dominated by the one of a tempered stable process. Specifically, denoting by ν its Lévy measure, we assume that ν is symmetric and that there exists μ a finite measure on S^{d-1}, a positive and non increasing function $\bar q$ on $\mathbb{R}_+$ (possibly vanishing at infinity) such that $\forall A\in\mathcal{B}(\mathbb{R}^d)$,

$$\nu(A)\le\int_{S^{d-1}}\int_0^{+\infty}\mathbf{1}_A(s\theta)\frac{\bar q(s)}{s^{1+\alpha}}ds\mu(d\theta). \quad (1.5)$$

  The temperation can be seen as a way to impose finite moments to the driving process $(Z_t)_{t\ge0}$. It is therefore natural to ask if the solution $(X_t)_{t\ge0}$ of the SDE retains the integrability of the driving process $(Z_t)_{t\ge0}$. One of the main differences with the standard Stable case comes from the fact that in the tempered setting, we lose the Scaling property in the noise (compare the Lévy exponent (1.4) with the domination (1.5)). We state and comment our results in Section 4.

- The degenerate stable process, or chain of stable oscillators:

$$\begin{aligned}
dX_t^1 &= \left(a_t^{1,1}X_t^1 + \cdots + a_t^{1,n}X_t^n\right)dt + \sigma(t, X_{t^-})dZ_t \qquad (1.6)\\
dX_t^2 &= \left(a_t^{2,1}X_t^1 + \cdots + a_t^{2,n}X_t^n\right)dt\\
dX_t^3 &= \left(a_t^{3,2}X_t^2 + \cdots + a_t^{3,n}X_t^n\right)dt\\
&\vdots\\
dX_t^n &= \left(a_t^{n,n-1}X_t^{n-1} + a_t^{n,n}X_t^n\right)dt,\ X_0 = x \in \mathbb{R}^{nd},
\end{aligned}$$

  where $a_t^{i,j}$ are $d \times d$ matrices satisfying a kind of weak Hörmander condition. The degeneracy here comes from the fact that the noise only acts on the d first components of the space (here $\mathbb{R}^{nd}$). Also, we point out the particular form of the drift coefficient, that represents the propagation of the noise through the system. Indeed, to reach the k^{th} component, the noise has to go through the $k-1$ first ones. This equation presents other difficulties with respect to its Brownian counterpart, that has been investigated in [DM10], coming from the lack of integrability of the stable process on the one hand, and on the fact that the generator of the stable process is a non local operator on the other. However, some aspects are preserved, namely, the transport of the initial condition in the renormalized action and the multi-scale property. We refer to section Section 5 for a thorough presentation of this problem.

To tackle those problems, we use a continuity method known as the Parametrix technique. The parametrix approach was first introduced by Levi [Lev07], who required regularity on the coefficients. Another formulation has then been proposed by McKean and Singer [MKS67] which presents the advantage to not require any regularity on the coefficients, besides Hölder continuity (although it was initially presented under $\mathcal{C}^\infty$ assumptions). Additionally, this approach allows us to derive uniqueness to the Martingale problem associated with the generator of (1.1), in the lines of [BP09] and [Men11]. We will make a presentation of this method in Section 2. It turns out that this technique allows also to prove regularity for the solution of the SDE, as it provides an explicit representation of the density in terms of the density of a simpler process, namely the Frozen process and the so-called Parametrix kernel. We discuss how this regularity can be used in the context of stochastic approximation in Section 6.

2 The Parametrix Setting

2.1 The Parametrix Technique for PDEs

The parametrix technique is a perturbation method coming from PDE theory. It was originally formulated in order to give an approximation of the fundamental solution of an elliptic linear differential equation of order $2n$, $n \in \mathbb{N}^*$, with variable coefficients (see Levi [Lev07]). Roughly speaking, the idea consists in isolating a principal part of the fundamental solution and controlling the remainder. A common choice for the principal part consists in considering the solution of the underlying equation with constant coefficients. This technique has then been extended to equations of hyperbolic type by Hadamard (see [Had32, Had64]). In our probabilistic setting, we will focus on parabolic type equations whose fundamental solutions can be linked to the density of suitable SDEs.

Consider a second order differential operator $L_t(x, \nabla_x)$, and the Cauchy problem:

$$\begin{cases} \partial_t u(t,x) + L_t(x,\nabla_x)u(t,x) = 0, & (t,x) \in [0,T) \times \mathbb{R}^d, \\ u(T,x) = f(x), & x \in \mathbb{R}^d, \ f \in \mathcal{C}_b(\mathbb{R}^d). \end{cases} \tag{2.1}$$

We say that a function $p(t,s,x,y)$, $s > t \geq 0$, $(x,y) \in \mathbb{R}^{2d}$, is a fundamental solution of the Cauchy problem (2.1), if given the initial condition f, the solution u to the PDE (2.1) can be written as:

$$u(t,x) = \int_{\mathbb{R}^d} f(z)p(t,T,x,z)dz.$$

Let us denote by $L_t(y, \nabla_x)$ a differential operator with constant coefficients (depending on a fixed point $y \in \mathbb{R}^d$), and $\tilde{p}$ the fundamental solution of

$$\begin{cases} \partial_t u(t,x) + L_t(y,\nabla_x)u(t,x) = 0, & (t,x) \in [0,T) \times \mathbb{R}^d, \\ u(T,x) = f(x), & x \in \mathbb{R}^d, \ f \in \mathcal{C}_b(\mathbb{R}^d). \end{cases}$$

We now look at $\tilde{p}$ as the principal part of p writing:

$$p(t,s,x,y) = \tilde{p}(t,s,x,y) + \int_t^s du \int_{\mathbb{R}^d} \tilde{p}(t,u,x,z)\Phi(u,s,z,y)dz, \tag{2.2}$$

and here Φ is to be determined by the condition that $\big(\partial_t + L_t(x,\nabla_x)\big)p(t,s,x,y) = 0$. Intuitively, $\tilde{p}$ is the principal part of the fundamental solution and Φ has to be seen as a remainder.

Note that the representation (2.2) is not unique, i.e.: various choices of Φ can lead to controllable expansions of the fundamental solution. In the non degenerate diffusive framework, we can for instance mention the approach developed in the works of Friedman [Fri64] who derived Gaussian bounds of type (1.3), when the coefficients of $L_t(x,\nabla_x)$ are Hölder continuous in time and space. On the other hand, the techniques introduced by McKean and Singer [MKS67] allow to obtain the same estimates under the sole assumption of measurability in time and Hölder continuity in space.

Let us mention that the Hölder continuity in space is minimal in order to obtain pointwise estimates on the density. Indeed, in the non degenerate Brownian case, under uniform continuity assumptions on the diffusion coefficient, L^p estimates can be obtained for the density through Harmonic Analysis techniques (see Stroock and Varadhan [SV79]). Under the sole non degeneracy assumption on σ, L^p estimates are derived as well for Itô processes in Krylov [Kry87], based on Alexandroff-Bakelman-Pucci estimates (see e.g. Section 4 Chapter V [Bas97]).

2.2 The Parametrix for Density Estimates

In our approach, we chose a probabilistic point of view, as PDE of the parabolic type can be linked to the solution of an SDE. Again, one can identify the density of the solution of an SDE with the fundamental solution of the Cauchy problem associated with the generator of the SDE. The parametrix now reformulates as a continuity technique that provides a formal representation for the density of an SDE in terms of a series involving the density of another, simpler, Markov process.

Let us now describe how the parametrix method can be formulated. Once again, let us denote by $(X_t)_{t\geq 0}$ the solution of the stable-driven SDE:

$$X_t = x + \int_0^t b(u, X_u)du + \int_0^t \sigma(u, X_{u^-})dZ_u, \tag{2.3}$$

and assume that this SDE admits a unique weak solution $(X_s^{t,x})_{0\leq t\leq s}$ to (2.3) which has a Feller semigroup:

$$P_{t,s}f(x) = \mathbb{E}\Big(f(X_s)|X_t = x\Big).$$

We do not assume a priori that this semigroup is absolutely continuous (that is that the density of X_t exists). Also, we do not assume regularity on the coefficients, as opposed to a Malliavin calculus approach. Instead, we produce a proxy candidate $(\tilde{X}_t^y)_{t\geq 0}$ to approximate $(X_t)_{t\geq 0}$. Let $y \in \mathbb{R}^d$ be an arbitrary terminal point. Intuitively, $y \in \mathbb{R}^d$ is the point where want to approximate the density of (2.3). We freeze the coefficients of (2.3) at the terminal point y:

$$\tilde{X}_t^y = x + \int_0^t b(u, y)du + \int_0^t \sigma(u, y)dZ_u. \tag{2.4}$$

Let us point out that it could seem more natural to freeze the coefficients at the initial position $x \in \mathbb{R}^d$. Besides, it was the approach initially developed by Levi [Lev07]. However, in this case, regularity on the coefficients is needed.

Observe that when σ is uniformly elliptic, then the *frozen process* $(\tilde{X}_t^y)_{t\geq 0}$ has a density whenever $(Z_t)_{t\geq 0}$ does. We denote throughout this document $\tilde{p}^y$ the density of the frozen process. Also, we drop the superscript y when the point considered in the density and the freezing point are the same and write $\tilde{p}(t, x, y) = \tilde{p}^y(t, x, y)$. Consequently, since $\tilde{X}^y$ is an approximation of X, we can expect that the density of

$\tilde{X}^y$ is not too far from the density of X. We quantify the distance with the help of the generators of the solution of (2.3) and (2.4) and the Kolmogorov equations. Let us define for $\xi \in \mathbb{R}^d$, the integro-differential operator:

$$\begin{aligned} L_t(\xi, \nabla_x)\varphi(x) &= \langle \nabla_x \varphi(x), b(t,\xi)\rangle \\ &+ \int_{\mathbb{R}^d} \left(\varphi(x+\sigma(t,\xi)z) - \varphi(x) - \frac{\langle \nabla_x \varphi(x), \sigma(t,\xi)z\rangle}{1+|z|^2}\right)\nu(dz). \end{aligned}$$

Observe that for $\xi = x$, the initial position, $L(x, \nabla_x)$ is the generator of $(X_t)_{t\geq 0}$, whereas for $\xi = y$ the terminal point, $L(y, \nabla_x)$ is the generator of $\tilde{X}^y$. Also, we emphasize with the notations ∇_x the variable on which the operator acts. Assume first that $(X_t)_{t\geq 0}$ has a smooth density and smooth coefficients (so that the adjoint operator is well-defined):

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

Then, this density satisfies the Forward Chapman-Kolmogorov equations:

$$\begin{aligned} &\partial_s p(t,s,x,z) = L_s(x, \nabla_z)^* p(t,s,x,z), \\ &\text{for all } s > t,\ (x,z) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lim_{s\downarrow t} p(t,s,x,\cdot) = \delta_x(\cdot)\ . \end{aligned} \tag{2.5}$$

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

$$\begin{aligned} &\partial_t \tilde{p}(t,s,x,z) = -L_t(y, \nabla_x)\tilde{p}(t,s,x,z), \\ &\text{for all } s > t,\ (x,z) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lim_{t\uparrow s} \tilde{p}(t,s,\cdot,z) = \delta_z(\cdot)\ . \end{aligned} \tag{2.6}$$

We deduce from the Dirac convergences (2.5) and (2.6) that:

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

Differentiating formally under the integral, leads to:

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

Then, using the Kolmogorov Backward equation (2.6) for $\tilde{p}$ and the Forward equation (2.5) for p, we get:

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

Passing to the adjoint in the last equality yields:

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

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

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

and the Parametrix Kernel:

$$\forall 0\le t<s,\ (x,y)\in(\mathbb{R}^d)^2,\ H(t,s,x,y)=\Big(L_t(x,\nabla_x)-L_t(y,\nabla_x)\Big)\tilde{p}(t,s,x,y). \quad (2.7)$$

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

$$\forall 0\le t<T,\ (x,y)\in(\mathbb{R}^d)^2,\ p(t,s,x,y)=\sum_{r=0}^{+\infty}(\tilde{p}\otimes H^{(r)})(t,s,x,y), \quad (2.8)$$

with $\tilde{p}\otimes H^{(0)}=\tilde{p}$ and $\forall r\in\mathbb{N},\ H^{(r)}(t,s,x,y)=H^{(r-1)}\otimes H(t,s,x,y)$. Note that in this formulation, we actually obtain that the function Φ in (2.2) writes:

$$\Phi(t,s,x,y)=\sum_{k=1}^{+\infty}H^{(k)}(t,s,x,y).$$

Observe that to derive the series representation (2.8) we assumed the existence and regularity of the density of $(X_t)_{t\ge 0}$ and the smoothness of the coefficients. However, the series (2.8) can be investigated without strong smoothness assumptions.

To justify the parametrix representation in the Brownian setting with Hölder coefficients, the usual approach is as follows. First we regularize the coefficients and use a theorem ensuring existence and regularity of the solution when the coefficients are smooth (Hörmander theorem). We then obtain estimates on the series (2.8) which are uniform with respect to the regularizing parameters. We eventually pass to the limit thanks to the weak uniqueness, proved through the well posedness of the martingale problem, as exposed in [Men11]. This allows to identify the sum of the series (2.8) with the density of the SDE (2.3), and to transfer the density estimates of the regular case to the limit.

However, in a very general Lévy setting, the existence is not guaranteed, let alone the regularity. Indeed, in the discontinuous case, there are no general (Hörmander) theorem to ensure the existence of the density even with regular coefficients, see anyhow the references page 11.

Therefore, we developed a semigroup approach that instead, provides a formal representation of the semigroup of (2.3) in terms of the series (2.8). The proof relies on the Markov property, but is more complex to give in a semigroup formulation.

Proposition 2.1. *Let* $(P_{t,s})_{0\le t\le s}$ *denote the semigroup associated with* $(X_s^{t,x})_{0\le t\le s}$. *We have the following formal representation. For all* $0 \le t < s$, $(x,y) \in (\mathbb{R}^d)^2$ *and any bounded measurable* $f : \mathbb{R}^d \to \mathbb{R}$*:*

$$P_{t,s}f(x) = \mathbb{E}[f(X_s)|X_t = x] = \int_{\mathbb{R}^d} \left(\sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t,s,x,y) \right) f(y)dy, \tag{2.9}$$

Furthermore, when the sum of the series appearing in (2.9) *is well defined, it yields the existence as well as a representation for the density of the initial process. Namely* $\mathbb{P}[X_s \in dy|X_t = x] = p(t,s,x,y)dy$ *where :*

$$\forall 0 \le t < T,\ (x,y) \in (\mathbb{R}^d)^2,\ p(t,s,x,y) = \sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t,s,x,y). \tag{2.10}$$

The proof of this statement is provided in the text (see [HM14]). It relies on the Markov properties (specifically the Chapman-Kolmogorov equations) of the processes involved. We insist on the fact that this representation is formal. Nevertheless, when the series in (2.9) converges, we deduce the existence of the density of X as well as a representation for it.

2.3 Convergence of the Series

The convergence is usually investigated by giving upper bounds on the frozen process and the parametrix kernel that are homogeneous to a density (up to a singularity for the kernel), and exploiting a smoothing effect in time of the parametrix kernel. Let us mention that in a very general setting, the convergence of the series cannot always be obtained. For instance, in Delarue and Menozzi [DM10], for weak Hörmander-Kolmogorov diffusions, the authors must truncate the series and prove that the rest of the sum is a remainder. However, even when the series does not converge, the parametrix representation is interesting in the sense that it provides a principal term in the expansion of the density of the SDE (2.3).

Let us detail the steps usually used to investigate the convergence of the series. From (2.4), we see that the density of the frozen process can be derived from the density of the driving process. Consequently, the first step in proving the convergence of the parametrix series consists in giving an upper bound for the frozen density of the form $\tilde{p}(t,s,x,y) \le C\bar{p}(s-t,y-x)$, where $\bar{p}(s-t,y-x)$ is homogeneous to the density of the driving process. For instance, in the non degenerate stable case, $\alpha \in (0,2)$, when b is bounded, σ satisfies (1.2), and the spectral measure of $(Z_t)_{t\ge 0}$ has a positive density on the sphere, one can take

$$\bar{p}(s-t,y-x) = C(s-t)^{-d/\alpha}\big(1+\frac{|y-x|}{(s-t)^{\frac{1}{\alpha}}}\big)^{-(d+\alpha)}. \tag{2.11}$$

The second step consists in giving a similar estimate for the kernel, up to a singularity. Formally, the generators of a stable process is a fractional derivative of order α. Since the spatial derivative of a stable density yields a singularity of order $(s-t)^{-1/\alpha}$, the expected upper bound for the parametrix kernel H is:

$$H(t,s,x,y) \le C\frac{\delta \wedge |y-x|^\eta}{s-t}\bar{p}(s-t,y-x), \tag{2.12}$$

where the contribution $\delta \wedge |y-x|^\eta$ comes from the Hölder regularity of the coefficient. We point out that the time singularity obtained is coherent with the Brownian case, for which we apply a non degenerate second order operator in space to a Gaussian density. Now, we say that the parametrix kernel presents a *smoothing property* if there exists $\omega > 0$, such that:

$$\int_{\mathbb{R}^d} |H(t,s,x,y)|dx \le C\int_{\mathbb{R}^d}\frac{\delta \wedge |y-x|^\eta}{s-t}\bar{p}(s-t,y-x)dx \le C(s-t)^\omega. \tag{2.13}$$

In other words, when integrated in space, the singularities are compensated and the parametrix kernel gives a positive power in time. This important property actually yields the convergence of the parametrix series in small time.

Let us now illustrate how this property can be used in the first term of the expansion:

$$|\tilde{p}\otimes H(t,s,x,y)| \le C\int_t^s d\tau \int_{\mathbb{R}^d}\bar{p}(\tau-t,z-x)\frac{\delta\wedge|y-z|^\eta}{s-\tau}\bar{p}(s-\tau,y-z)dz.$$

Roughly speaking, one of the two densities $\bar{p}(\tau-t,z-x)$, $\bar{p}(s-\tau,y-z)$ is homogeneous to $\bar{p}(s-t,y-x)$, and can therefore be taken out of the integral, the remaining one yielding the smoothing effect. When iterating the kernel $\tilde{p}\otimes H^{(k+1)} = \tilde{p}\otimes H^{(k)}\otimes H$, the exponent in time will grow with each iteration, giving the convergence of the series, as well an upper bound for the density:

$$p(t,s,x,y) \le C\bar{p}(s-t,y-x), \text{ for } s-t \text{ small enough.}$$

Observe that the upper bound $\bar{p}$ presents a "semigroup" property in the following sense:

$$\forall u \in (t,s), \int_{\mathbb{R}^d}\bar{p}(u-t,z-x)\bar{p}(s-u,y-z)dz \le C\bar{p}(s-t,y-x).$$

This property is crucial as it allows the propagation of the upper bound obtained on the density of the initial SDE in small time to an arbitrary but finite time. Indeed, let $t, T \in \mathbb{R}_+$, and define $(\tau_i)_{i \in [\![0,n+1]\!]}$ a subdivision of the interval $[t, T]$ with $\tau_0 = t$ and $\tau_{n+1} = T$, whose mesh is small enough so that the upper bound on p holds on each subinterval $[\tau_i, \tau_{i+1}]$. From the Chapman-Kolmogorov equations for p, we can write:

$$p(t, T, x, y) = \int_{\mathbb{R}^d} dz_1 \cdots \int_{\mathbb{R}^d} dz_n \prod_{i=0}^{n} p(\tau_i, \tau_{i+1}, z_i, z_{i+1}),$$

where $z_0 = x$ and $z_{n+1} = y$. Now, for each $i \in [\![0, n]\!]$, the upper bound holds:

$$p(\tau_i, \tau_{i+1}, z_i, z_{i+1}) \leq C\bar{p}(\tau_i, \tau_{i+1}, z_i, z_{i+1}).$$

Consequently, exploiting the semigroup property on $\bar{p}$ yields:

$$p(t, T, x, y) \leq C^n \bar{p}(t, T, x, y).$$

Let us insist on the fact that this procedure yields constants with an exponential dependency in time as n depends on $T - t$.

2.4 Parametrix and Martingale Problem

Another important consequence of the smoothing property (2.13) of the parametrix kernel H is a way to derive weak uniqueness to the martingale problem associated with the generator of (2.3). This approach is presented in Section 3.3 of Chapter 2 in the non degenerate framework, and the same arguments apply for the degenerate case as well (see Section 4 in Chapter 3). This method was first introduced by Bass and Perkins [BP09], and the connection with the parametrix setting has been established in [Men11]. It only relies on the smoothing property of the parametrix kernel. Let us mention that under weaker assumptions on σ, namely continuity and ellipticity, weak uniqueness has been derived by various authors. Let us mention the works of Stroock [Str75], where the driving process presents a Brownian part. Besides, Komatsu [Kom08] considers perturbations of stable-like operators, in the sense that the "stability index" α can depend on the spatial variable in a smooth and non degenerate way. Also, these stable-like operators are assumed to have smooth positive spherical densities. Furthermore, Bass and Tang [BT09] consider generators involving jump measures of the form $\nu(x, dy) = \frac{A(x,y)}{|y|^{d+\alpha}}$, where A is bounded from below and above and continuous[1]. Observe that this particular form of jump measure admits a polar decomposition and that the spherical part is equivalent to the Lebesgue measure on the

[1] actually, the weak uniqueness is proved under a slightly weaker condition, see Assumption 1.1 in [BT09]

sphere. Eventually, Bass and Chen [BC06] consider n independent one-dimensional stable processes as driving process for the SDE.

Anyhow, in our setting, weak uniqueness cannot be derived from those works. Indeed, in the non degenerate setting of Chapter 2, we consider Lévy measures whose spherical parts can be non equivalent or singular with respect to the Lebesgue measure on the sphere. In addition, to the best of our knowledge, no previously established results exist in the degenerate framework of Chapter 3.

2.5 Parametrix and Numerical Probabilities

We conclude this section with an application of the parametrix technique to numerical probabilities. The parametrix technique allows to investigate the weak error expansion for the Euler-Maruyama scheme associated with (2.3), as it can be applied to Markov chains. Indeed, let us introduce the Euler-Maruyama scheme associated with (2.3) with time homogeneous coefficients:

$$X_t^n = x + \int_0^t b(X^n_{\phi(s)})ds + \int_0^t \sigma(X^n_{\phi(s)})dZ_s, \quad \phi(s) = \sup\{t_i :\ t_i \le s\}, \tag{2.14}$$

for a given time step $\Delta = \frac{1}{n}$, $n \in \mathbb{N}^*$, setting for all $i \in \mathbb{N}$, $t_i = i\Delta$. We chose to write the scheme for time homogeneous coefficients, for notational convenience. The Euler-Maruyama scheme enjoys the Markov property, at the discretization times, and a parametrix expansion can be derived for it as well. The corresponding time homogeneous version of (2.3) writes:

$$X_t = x + \int_0^t b(X_u)du + \int_0^t \sigma(X_{u^-})dZ_u. \tag{2.15}$$

We introduce the "frozen Markov chains" $(\tilde{X}^n_{t_k})_{k\in[\![0,n]\!]}$:

$$\tilde{X}^n_{t_k} = x,\ \tilde{X}^n_{t_{k+1}} = \tilde{X}^n_{t_k} + b(y)\Delta + \sigma(y)(Z_{t_{k+1}} - Z_{t_k}). \tag{2.16}$$

Observe that with that definition, the density of the Markov chain (2.16) is $\tilde{p}^y$ considered at discrete times.

We denote the discrete generators by:

$$\begin{aligned}
&L_n f(t_k - t_j, x, y)\\
&= \Delta^{-1}\left(\int p_n(\Delta, x, z) f(t_k - t_{j+1}, z, y)dz - f(t_k - t_{j+1}, x, y)\right), \qquad (2.17)\\
&\tilde{L}_n f(t_k - t_j, x, y)\\
&= \Delta^{-1}\left(\int \tilde{p}^y(\Delta, x, z) f(t_k - t_{j+1}, z, y)dz - f(t_k - t_{j+1}, x, y)\right). \qquad (2.18)
\end{aligned}$$

Observe the shift in the index (t_{j+1} instead of t_j) in (2.17) and (2.18) due to the discrete setting. We obtain a parametrix representation for the density of the Euler-Maruyama scheme using the frozen density and the discrete generators.

Proposition 2.2. *The density $p_n(t_k,x,y)$ of the Euler-Maruyama scheme admits the following representation:*

$$p_n(t_k-t_j,x,y)=\sum_{r=0}^{k-j}\tilde{p}\otimes_n H_n^{(r,n)}(t_k-t_j,x,y), \tag{2.19}$$

where we denoted $H_n(t_k,x,y)=(L_n-\tilde{L}_n)\tilde{p}(t_k,x,y)$, and $\otimes_n$ is the discretized space-time convolution:

$$f\otimes_n g\,(t_k,x,y)=\frac{1}{n}\sum_{i=0}^{k-1}\int_{\mathbb{R}^d} f\,(t_i,x,z)\,g\,(t_k-t_i,z,y)\,dz,$$

and $H_n^{(r,n)}(t_k,x,y)=H_n^{(r-1,n)}\otimes_n H_n(t_k,x,y)$, where $\tilde{p}\otimes H_n^{(0,n)}(t_k,x,y)=\tilde{p}(t_k,x,y)$.

We use the notation $H_n^{(r,n)}(t_k,x,y)$ to emphasize the dependency in the discretization of the convolution. That is, the subscript n refers to the discrete generators, whereas the arguments in the superscript (r,n) refer respectively to the number of steps we iterate the convolution, and the number of discretization dates. Therefore, we have $H_n^{(1,n)}(t_k,x,y)=H_n(t,x,y)$. Using the convention $H_n^{(r,n)}=0$ for $r>k-j$, we can write $p_n(t_k-t_j,x,y)=\sum_{r=0}^{+\infty}\tilde{p}\otimes_n H_n^{(r,n)}(t_k-t_j,x,y)$. We can now use this expansion to investigate the weak error $p-p_n$. We introduce for all $k\in[\![0,n-1]\!]$:

$$\begin{aligned}
p^d(t_k,x,y) &= \sum_{r=0}^{+\infty}\tilde{p}\otimes_n H^{(r,n)}(t_k,x,y), \qquad (2.20)\\
H^{(r,n)}(t_k,x,y) &= H^{(r-1,n)}\otimes_n H(t_k,x,y),\ \text{where } \tilde{p}\otimes H^{(0,n)}(t_k,x,y)=\tilde{p}_\alpha(t_k,x,y),
\end{aligned}$$

Now, we use p^d to compare the expansions, writing:

$$(p-p_n)(t_k,x,y)=(p-p^d)(t_k,x,y)+(p^d-p_n)(t_k,x,y).$$

For $(p-p^d)(t_k,x,y)$, comparing the expansions (2.10) and (2.20), we use the Taylor formula to quantify the difference between the discretized time space convolution and the usual one. For the second part, $(p^d-p_n)(t_k,x,y)$, we compare the expansions (2.19) and (2.20), and once again use the Taylor formula to quantify the distance between the two kernels H and H_n. We thus derive the following expansion for the weak error:

Theorem 2.3. *Assume b and σ are $\mathcal{C}^\infty$, and let $M \in \mathbb{N}^*$. Then, for all $x, y \in \mathbb{R}^d$, we have:*

$$p(1,x,y) - p_n(1,x,y) = \sum_{k=1}^{M-1} \frac{1}{(k+1)!n^k} \left(p \otimes_n \left(L(x,\nabla_x) - L(y,\nabla_x) \right)^{k+1} p^d \right)(1,x,y)$$
$$- \frac{1}{(k+1)!n^k} \left(p^d \otimes_n \left(L(x,\nabla_x) - L(y,\nabla_x) \right)^{k+1} p_n \right)(1,x,y) + \frac{R(x,y)}{n^M}.$$

Moreover, each term and the remainder can be bounded by $C\bar{p}(s-t, y-x)$, defined in (2.11).

We stated the expansion in the case where the coefficients are $\mathcal{C}^\infty$ for simplicity. The reader may consult [KM02] in the Brownian case or [KM10] in the stable case for a proof of this result. Expansions for quantities $\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]$ have been investigated since Talay and Tubaro [TT90]. One of the advantages of the representation in Theorem 2.3 is that the dependences are quite explicit. Therefore, when investigating regularity, we can differentiate each term of the expansion separately to see if it is well defined. Also, we point out another advantage of the parametrix method. In the context of weak error expansion $\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]$, we need low regularity on the coefficients of the *SDE* as well as on the test function f (in Talay and Tubaro [TT90], smoothness of the quantities involved is required). In a parametrix approach, the expansion relies on the regularity of the law of the driving process. Note however that this requires some structure on the coefficients, namely non degeneracy conditions.

We point out that the freezing point is up to now relatively arbitrary. This is actually a forte of the parametrix technique, as we will see in Section 5. Indeed, in the degenerate case of equation (1.6), the process has an intrinsic geometry that we have to take into account when investigating the deviations from the typical behavior. Fix a terminal time $T > 0$ and a final position $y \in \mathbb{R}^d$. We will freeze the "diffusion" coefficient in the process (1.6) along the curve $(R_{s,T}y)_{s\in[t,T]}$, backward flow of the final point y by the deterministic system, to compensate the transport of the initial condition x (see Section 5 for details). Also, since the drift in the degenerate case is linear, we keep it as it is. Indeed, denoting $A_t = (a_t^{i,j})_{(i,j)\in[\![1,n]\!]^2}$, the drift part in (1.6), it yields the term $\langle \nabla_x \varphi(x), A_t x\rangle$ in the generator $L_t\varphi(x)$. This contribution is the same for both the frozen process and the initial one, so that it vanishes when taking the difference of the generators. In the case of a Lipschitz unbounded drift, we could proceed to a linearization as showed in [DM10], see as well Theorem 2.1 in Chapter 3 in the current case.

2.6 Conclusion and Perspectives

In conclusion to this section, we mention that the parametrix technique is a versatile tool that can have multiple applications, from numerical probabilities to heat kernel estimates. However, the computations induced are often quite tricky to handle. Recent development are still made around the parametrix. Let us mention the recent paper of Bally and Kohatsu-Higa [BKH14], providing a probabilistic interpretation of the parametrix method, and an oncoming work of Frikha and Kohatsu-Higa concerning small noise density expansions generalizing the results of Azencott *et al.* [A^+81], based on a parametrix approach coupled with Malliavin calculus. This method also seems to give small time heat kernel expansions in the weak Hörmander setting for diffusions.

3 The stable process

We take a moment here to write about the asymptotics of a multidimensional stable process, since it has been made clear from the last section that such estimates are crucial. The Stable process has been extensively studied, and arises in a huge variety of cases, from statistical mechanics to financial mathematics (see e.g. Borovkova *et al.* [BPP09]). For background on Stable laws, let us mention the books of Sato [Sat05], Samorodnitsky and Taqqu [ST94], and in the one dimensional case, the one of Zolotarev [Zol86]. In this latter framework, the asymptotic behavior of the density is well understood. In Zolotarev [Zol86], the author gives precise asymptotics. However, in the multidimensional case, the question is much more difficult.

As we mentioned earlier, the Lévy measure of a stable process factorizes in polar coordinates, and the spherical part, μ of the Lévy measure is referred to as the spectral measure. Observe that when the spectral measure is non degenerate in the following sense:

$$\forall p \in \mathbb{R}^d,\ \exists C > 1,\quad C^{-1}|p|^\alpha \le \int_{S^{d-1}} |\langle p, \xi\rangle|^\alpha \mu(d\xi) \le C|p|^\alpha, \tag{3.1}$$

then the process $(Z_t)_{t\ge 0}$ has a density for each $t > 0$ with respect to the Lebesgue measure, since the Fourier transform is then integrable. However, it turns out that (3.1) alone is not enough to get a density estimate, except a global diagonal bound. Denoting $p_Z(t,x)$ the density of the stable process, the integrability of the Fourier transform yields the global upper bound $p_Z(t,x) \le Ct^{-d/\alpha}$.

In Kolokoltsov [Kol00b], in order to derive density estimates, it is assumed in addition that the spectral measure has a smooth positive density with respect to the Lebesgue measure. This assumption imposes a "decay rate" on the spectral measure in the following way. For all $r > 0$, we define:

$$\mu^*(r) = \sup_{\xi \in S^{d-1}} \mu\Big(B(\xi, 1/r) \cap S^{d-1}\Big).$$

With a slight abuse in terminology, we will refer to the decay rate of $\mu(r)$ when speaking of the decay rate of μ. Then, when the spectral measure has a smooth positive density with respect to the Lebesgue measure, we actually have: $C^{-1}r^{d-1} \le \mu^*(r) \le Cr^{d-1}$, for some constant $C > 1$, so that the decay rate of μ is of order $d-1$. More generally, in a recent work, Watanabe [Wat07] proved how the decay rate of μ impacts the decay of the density. Indeed, we have $\exists C > 1$, $\forall x \in \mathbb{R}^d$, $\forall t > 0$,

$$p_Z(t,x) \leq C \frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{1+\alpha}} \mu^* \left(1+\frac{|x|}{t^{1/\alpha}}\right). \tag{3.2}$$

To give a similar lower bound, we define $S_\mu \subset S^{d-1}$ to be the support of the spectral measure μ, and

$$C^0_\mu(n) = \{\xi \in S^{d-1};\ \exists c_1, \cdots, c_n > 0,\ \exists \xi_1, \ldots, \xi_n \in S_\mu;\ \xi = \sum_{j=1}^n c_j \xi_j\},$$

the cone of points reachable as a sum of exactly n points of the support.

Then, for all $x \in \mathbb{R}^d$ such that $x/|x|$ is in the interior of $C^0_\mu(d)$, $\forall t > 0$, we have:

$$p_Z(t,x) \geq C^{-1} \frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{1+\alpha}} \mu\left(B\left(\frac{x}{|x|}, 1+\frac{|x|}{t^{1/\alpha}}\right) \cap S^{d-1}\right). \tag{3.3}$$

Let us mention that a more thorough study is established in Watanabe [Wat07]. Anyway, in the above (3.2) and (3.3), we see the impact of the decay rate of μ. Indeed, assume that μ is such that there exists a compact set $K \subset S^{d-1}$with:

$$C^{-1} r^{\gamma-1} \leq \mu\left(B(\theta, r) \cap K\right) \leq C r^{\gamma-1},\ \forall \theta \in K \subset S^{d-1},\ \forall r \leq 1/2. \tag{3.4}$$

Then, the two sided estimate holds for all x such that $x/|x| \in K$:

$$C^{-1} \frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \leq p_Z(t,x) \leq C \frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}}, \tag{3.5}$$

which emphasizes that, locally, every concentration regime is possible between $\alpha+1$ and $\alpha+d$. At time $t=1$, Pruitt and Taylor [PT69], called the order $(1+|x|)^{-1-\alpha}$ the worst possible order. This rate is realized when the spectral measure is a Dirac mass, and corresponds to one dimensional stable process seen in a space of dimension d. In extension to this terminology, Watanabe determines the order $(1+|x|)^{-(1+\alpha)d}$ to be the best possible order.

To illustrate this phenomena, let us consider two independent one dimensional stable processes Z^1_t and Z^2_t, and $Z_t = (Z^1_t, Z^2_t) \in \mathbb{R}^2$. Then, by independence,

$$p_Z\left(1, \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}\right) \asymp \frac{1}{(1+|x^1|)^{1+\alpha}} \frac{1}{(1+|x^2|)^{1+\alpha}}.$$

Now, observe that if $|x^1|$ is small, then:

$$p_Z\left(1, \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}\right) \asymp \frac{1}{(1+|x^2|)^{1+\alpha}},$$

which corresponds to the worst possible order, whereas when $|x^1| \asymp |x^2|$,

$$p_Z\left(1, \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}\right) \asymp \frac{1}{(1+|x^1|)^{2+2\alpha}},$$

corresponds to the best possible order. We point out that the above description is local, and in order to have global two sided estimates, we have to control the spectral measure globally. For instance, in the case of the rotationally invariant stable process, the spectral measure is equivalent to the Lebesgue measure on S^{d-1}, and in this case, we can take $K = S^{d-1}$, and $\gamma = d$ in (3.4) and (3.5).

Observe however that if the decay rate of μ is too low regarding 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.

We point out the ratio $\frac{|x|}{t^{1/\alpha}}$ appearing in the two sided estimates (3.5). Observe that when $|x| \leq Ct^{1/\alpha}$, the estimates become:

$$C^{-1}t^{-d/\alpha} \leq p_Z(t,x) \leq Ct^{-d/\alpha}.$$

We will refer to this regime as the diagonal regime, and to the previous bound as the diagonal estimate. These estimates highlight the auto-similarity index, and can be considered as the behavior of the density on the *typical* sets $\{x \in \mathbb{R}^d; |x| \leq Ct^{1/\alpha}\}$.

When $|x| \geq Ct^{1/\alpha}$, the estimates become for $x/|x| \in K$,

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

We will refer to these estimates as the off-diagonal estimates. They appear in a *large deviation* regime, that emphasizes the heavy tails of the process. We see that when $\alpha + \gamma > d$, not only the bound is integrable in the off-diagonal regime, but it also yields a positive power in time. This crucial fact will give the smoothing property of the parametrix kernel $H(t,s,x,y) = \Big(L_t(x,\nabla_x) - L_t(y,\nabla_x)\Big)\tilde{p}(t,s,x,y)$, and fix dimension constraints in a parametrix approach. Indeed, (2.12) reformulates as follows in the off-diagonal regime:

$$\begin{aligned} H(t,s,x,y) &\leq C\frac{\delta \wedge |y-x|^\eta}{s-t}\frac{(s-t)^{-d/\alpha}}{\left(1+\frac{|y-x|}{(s-t)^{1/\alpha}}\right)^{\alpha+\gamma}} \\ &\leq C\frac{\delta \wedge |y-x|^\eta}{s-t}\frac{(s-t)^{1+\frac{\gamma-d}{\alpha}}}{|y-x|^{\alpha+\gamma}}, \end{aligned}$$

and we see that in the off-diagonal regime, when $\alpha+\gamma > d$, the estimate on the frozen density compensates the singularity of the kernel. Thus the parametrix kernel H has the smoothing effect presented in (2.13).

We conclude this section by saying that the support of the spectral measure plays a key role in the obtention of density bounds for the stable process. The impact of the dimension of the support is reflected in the decay order of the polynomial tails in the off-diagonal regime.

4 The Tempered case.

In Section 3, we discussed the asymptotics of multidimensional stable densities. In this section, we push the discussion a step forward considering driving processes whose Lévy measure satisfies what we call a *tempered stable domination*. Namely, denoting by ν the Lévy measure of the driving process $(Z_t)_{t\geq 0}$, we assume that ν is symmetric and that for all $A \in \mathcal{B}(\mathbb{R}^d)$:

$$\nu(A) \leq \int_{S^{d-1}} \int_0^{+\infty} \mathbf{1}_A(s\theta) \frac{\bar{q}(s)}{s^{1+\alpha}} ds \mu(d\theta). \tag{4.1}$$

The function $\bar{q}$ appearing in the radial part is non increasing, positive and will be referred to as the temperation. With a slight abuse of language, we shall still refer to μ as the spectral measure. We also assume that the measure ν is non degenerate: denoting by φ_Z the Lévy-Khintchine exponent of Z, the following upper bound holds:

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

This assumption ensures the existence of the density of Z_t. Let us emphasize that in (4.1), we do not assume the factorization of the Lévy measure anymore. However, our results allow to recover the known case of the standard symmetric α-stable process. The asymptotics of such Lévy processes are investigated in Sztonyk [Szt10]. Once again, following the approach of [Wat07], the importance of the decay rate of the spectral measure is highlighted. Let us denote by $p_Z(t,x)$ the density of Z_t, the following theorems are proved in Sztonyk [Szt10].

Theorem 4.1. ***(Upper Bound).*** *Assume that the temperation $\bar{q} : \mathbb{R}_+ \to \mathbb{R}_+$, satisfies a doubling condition: $\bar{q}(s) \leq C\bar{q}(2s)$. If there exists $\gamma \in [1,d]$ such that*

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

then we have the following upper bound:

$$p_Z(t,x) \leq C \frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}(|x|). \tag{4.2}$$

(Lower Bound). *Assume moreover that there exists $A_{low} \subset \mathbb{R}^d$ and a non increasing function $\underline{q}$. If there exists $\gamma \in [1, d]$ such that $\forall x \in A_{low}$, $\forall r > 0$:*

$$\nu\Big(B(x,r)\Big) \geq Cr^{\gamma}\frac{\underline{q}(|x|)}{|x|^{\alpha+\gamma}},\ \forall r>0,\ \text{and}\ \nu\Big(B(0,r)^c\Big) \leq Cr^{-\alpha},\ \forall 0<r<1,$$

then the lower bound holds:

$$C^{-1}\frac{t^{-d/\alpha}}{\left(1+\frac{|x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}}\underline{q}(|x|) \leq p_Z(t,x). \tag{4.3}$$

Note that the estimates (4.2) and (4.3) present the same qualitative behavior as those of [Wat07], up to a multiplication by the tempering function. The goal is to transfer these estimates to the solution of an SDE driven by $(Z_t)_{t\geq 0}$. Let us define again:

$$X_t = x + \int_0^t F(X_u)du + \int_0^t \sigma(X_{u^-})dZ_u, \tag{4.4}$$

where $F : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz (or measurable and bounded), and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is Hölder continuous and satisfies the non degeneracy condition (1.2). If $\alpha \leq 1$, we take $F = 0$. This assumption (already present in Kolokolstov [Kol00b]) comes from the fact that when $\alpha \leq 1$, the noise does not "dominate" in the in the dynamics (4.4). Intuitively, the intrinsic time scale of the stable process is $t^{1/\alpha}$, which does not dominate t, the time scale of the drift, in small time when $\alpha \leq 1$. We mention the recent paper of Knopova and Kulik [KK14] who establish a kind of trade-off between the regularity of a non zero drift F and the index of the stable process $\alpha \leq 1$ to derive the well posedness of the martingale problem.

The temperation $\bar{q}$ can be seen as a mean 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 SDE is not always possible. In Kolokoltsov [Kol00b], the author succeeded by using the parametrix technique described in Section 2. In our setting, we manage to prove the convergence of the series (2.8), proving the existence of the density as well as a global upper bound and a diagonal lower bound. We then establish the lower bound from probabilistic techniques.

A big difference between the stable process and the tempered stable one comes from the fact that the temperation disrupts the scaling property. It is then not always possible from the results at time 1 to extrapolate results for arbitrary time. In a global approach like the parametrix technique, we have to consider the density at various times and the lack of scaling prevents us from splitting time and space in the computations.

We somehow manage to recover a time-space separation in our computations by using the Lévy-Itô decomposition at the characteristic time-scale. Indeed, fix $t > 0$, we can split the driving tempered stable process writing for $s \in [0, t]$, $Z_{s\wedge t} = M_{s\wedge t} + N_{s\wedge t}$, where $(M_{s\wedge t})_{s\in[0,t]}$ is a martingale and $(N_{s\wedge t})_{s\in[0,t]}$ is a Poisson process, by discussing if the jumps are bigger than a given arbitrary positive threshold. In particular, we take that threshold to be $t^{1/\alpha}$, the characteristic time scale, in order to recover the diagonal and off-diagonal regimes, when investigating the marginals at time t. See also the techniques developed in Sztonyk [Szt10].

We will denote $\mathbf{[H_{TS}]}$ the following set of assumptions. These hypotheses ensure the existence of the density, and are those required by Sztonyk [Szt10] in order to have two sided estimates on the density of the driving process Z.

$\mathbf{[H_{TS} - 1]}$ $(Z_t)_{t\geq 0}$ 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:

$$\nu(A) \leq \int_{S^{d-1}} \int_0^{+\infty} \mathbf{1}_A(s\theta) \frac{\bar{q}(s)}{s^{1+\alpha}} ds\mu(d\theta). \tag{4.5}$$

We assume one of the following:

$\mathbf{[H_{TS} - 1}a\mathbf{]}$ μ has a density with respect to the Lebesgue measure on the sphere.

$\mathbf{[H_{TS} - 1}b\mathbf{]}$ there exists $\gamma \in [1, d]$ such that $\mu\left(B(\theta, r) \cap S^{d-1}\right) \leq Cr^{\gamma-1}$, with $\gamma + \alpha > d$, and for all $s > 0$, there exists $C > 0$ such that:

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

$\mathbf{[H_{TS} - 2]}$ Denoting by φ_Z the Lévy-Khintchine exponent of $(Z_t)_{t\geq 0}$, there is $C > 0$ such that :

$$\mathbb{E}\left(e^{i\langle p, Z_t\rangle}\right) = e^{t\varphi_Z(p)} \leq e^{-Ct|p|^\alpha}, \ |p| > 1. \tag{4.6}$$

$\mathbf{[H_{TS} - 3]}$ $F : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous or measurable and bounded, if $\alpha > 1$, and $F = 0$ when $\alpha \leq 1$, and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded and η-Hölder continuous, $\eta \in (0, 1)$.

$\mathbf{[H_{TS} - 4]}$ σ is uniformly elliptic. There exists $C > 1$, such that for all $x, \xi \in \mathbb{R}^d$,

$$C^{-1}|\xi|^2 \leq \langle \xi, \sigma(x)\xi\rangle \leq C|\xi|^2. \tag{4.7}$$

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

$$\nu(x, A) = \nu\big(\{z \in \mathbb{R}^d;\ \sigma(x)z \in A\}\big). \tag{4.8}$$

We assume these measures to be Hölder continuous with respect to the first parameter, that is, for all $\forall A \in \mathcal{B}(\mathbb{R}^d)$,

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

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 of S^{d-1}, this is actually a consequence of the Hölder continuity of σ and the domination **[H-1]**.

[$\mathbf{H_{TS}}$ − 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 $x \in A_{low}$,

$$\begin{aligned} \nu\Big(B(x,r)\Big) &\geq Cr^{\gamma}\frac{\underline{q}(|x|)}{|x|^{\alpha+\gamma}},\ \forall r > 0 \qquad (4.9)\\ \nu\Big(B(0,r)^c\Big) &\leq C\frac{1}{r^{\alpha}},\ \forall r \in (0,1). \qquad (4.10)\end{aligned}$$

We say that **[$\mathbf{H_{TS}}$]** holds when **[$\mathbf{H_{TS}}$ − 1]** to **[$\mathbf{H_{TS}}$ − 5]** hold. We point out that **[$\mathbf{H_{TS}}$ − LB]** gives the lower bound, and that the upper bound holds independently. Assumption **[$\mathbf{H_{TS}}$ − 2]** ensures the existence of the density of Z.

Under **[$\mathbf{H_{TS}}$]**, we are able to prove the following.

Theorem 4.2 (**Weak Uniqueness**). *Assume **[$\mathbf{H_{TS}}$]** holds. The martingale problem associated with the generator $L(x, \nabla_x)$ of the solution of:*

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

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 $f \in \mathcal{C}_0^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$, denoting by $(X_t)_{t\geq 0}$ the canonical process, we have:

$$\mathbb{P}(X_0 = x) = 1 \quad \textit{and} \quad f(t, X_t) - \int_0^t (\partial_u + L(x, \nabla_x)f(u, X_u)du \quad \textit{is a } \mathbb{P}\textit{- martingale.}$$

Hence, weak uniqueness for the SDE holds.

Also, we have the following density estimate:

Theorem 4.3 (**Density Estimates**). *Under **[$\mathbf{H_{TS}}$]**, the unique weak solution of* (4.4) *has for every $t > 0$ a density with respect to the Lebesgue measure. Precisely, for all $t > 0$, and $x, y \in \mathbb{R}^d$,*

$$\mathbb{P}(X_t \in dy|X_0 = x) = p(t, x, y)dy. \qquad (4.11)$$

*Assume that the function Q defined below is decreasing, and fix a deterministic time horizon $T > 0$. There exists $C_1 \geq 1$ depending on T and the parameters in **[$\mathbf{H_{TS}}$]**, such that the following density estimates holds:*

$\forall 0 \le t \le T,\ \forall (x,y) \in \mathbb{R}^d$,

$$p(t,x,y) \le C_1 \frac{t^{-d/\alpha}}{\left(1+\frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|y-\theta_{t,0}(x)|), \tag{4.12}$$

where:

- *when the drift F is bounded, θ is the identity map: $\theta_{t,0}(x) = x$, and*
  - *under **[H_{TS} − 1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \bar{q}(s)$,*
  - *under **[H_{TS} − 1b]**, for all $s > 0$, $Q(s) = \min(1, s^{\gamma-1})\bar{q}(s)$,*
- *when the drift F is Lipschitz continuous, $\theta_{s,t}(x)$ denotes the solution to the ordinary differential equation:*

  $$\frac{d}{ds}\theta_{s,t}(x) = F(\theta_{s,t}(x)),\ \theta_{t,t}(x) = x,\ \forall 0 \le t,s \le T,$$

  and
  - *under **[H_{TS} − 1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \min(1,s)\bar{q}(s)$,*
  - *under **[H_{TS} − 1b]**, for all $s > 0$, $Q(s) = \min(1, s, s^{\gamma-1})\bar{q}(s)$.*

*Moreover, assume **[H_{TS} − LB]** holds. Then, if*

$$\forall s \in [0,t/2],\ B\big(\sigma(\theta_{s,0}(x))^{-1}(\theta_{0,t}(y) - x), Ct^{1/\alpha}\big) \subset A_{low}, \tag{4.13}$$

there exists $C_2 > 1$ such that

$$C_2^{-1} \frac{t^{-d/\alpha}}{\left(1+\frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \underline{q}(|y-\theta_{t,0}(x)|) \le p(t,x,y). \tag{4.14}$$

Remark 4.1. The condition (4.13) appearing for the lower bound comes from the possibly unbounded feature of the deterministic flow associated with (4.4). 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 (4.13) is satisfied for $s \in [\varepsilon_1 t, \varepsilon_2 t]$, $0 \le \varepsilon_1 < \varepsilon_2 \le 1$. 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 4.2. We point out that when $[\mathbf{H_{TS}} - 1a]$ holds, the condition

$$\mu\left(B(\theta, r) \cap S^{d-1}\right) \leq C r^{\gamma-1},$$

actually holds with $\gamma = d$. Besides, the function Q appearing in the upper bound (4.12) is decreasing typically when considering tempering function of the form:

$$\bar{q}(s) = \frac{1}{1+s^m},$$

where m is large enough. On the other hand, we highlight the fact that the doubling condition appearing in $[\mathbf{H_{TS}} - 1b]$ is a frequent assumption in the literature. In Sztonyk [Szt10], it is a crucial tool to control the iterated convolutions of the measure associated with the large jumps. We also refer to the paper of Jacob *et al.* [JKLS12] for density estimates related to symmetric Lévy processes which involve a metric associated with the Lévy exponent. The crucial assumption is that $\mathbb{R}^n$ endowed with this metric is a doubling space. This kind of assumptions also usually appears in harmonic analysis to control the measure of balls. Indeed, the Calderón-Zygmund theory for singular integrals naturally extends to doubling spaces (see e.g. Coifman and Weiss [CW71]).

In conclusion, in Theorem 4.3, the density estimates on the driving process transfer to the solution of the SDE as expected. Let us highlight that our results cover the case of the Stable process, that is when the spectral measure satisfies only $[\mathbf{H_{TS}}]$, with $\bar{q} = 1$. In other words, we do not need strong regularity of the spectral measure in order to derive the density bounds for the solution of the SDE. Also, we manage to prove the impact of the decay of μ on the decay of the density of the solution of (4.4), via the presence of the index γ in our estimates. Finally, let us comment that the presence of $\theta_{s,t}$ in the above bounds reflects the possibly unbounded deterministic transport associated with the drift F.

5 The degenerate Case

In this section, we present our results for the degenerate case. The degeneracy in our setting comes from the fact that the noise only affects the first component. Specifically, we focus on equations with dynamics:

$$\begin{aligned}
dX_t^1 &= \left(a_t^{1,1}X_t^1 + \cdots + a_t^{1,n}X_t^n\right)dt + \sigma(t, X_{t^-})dZ_t \\
dX_t^2 &= \left(a_t^{2,1}X_t^1 + \cdots + a_t^{2,n}X_t^n\right)dt \\
dX_t^3 &= \left(a_t^{3,2}X_t^2 + \cdots + a_t^{3,n}X_t^n\right)dt \\
&\ \vdots \\
dX_t^n &= \left(a_t^{n,n-1}X_t^{n-1} + a_t^{n,n}X_t^n\right)dt,\ X_0 = x \in \mathbb{R}^{nd},
\end{aligned} \tag{5.1}$$

where $a^{i,j}:\mathbb{R}_+\to\mathbb{R}^d\otimes\mathbb{R}^d,\ i\in[\![1,n]\!],\ j\in[\![(i-1)\vee 1,n]\!]$ and $(Z_t)_{t\geq 0}$ is an $\mathbb{R}^d$ valued symmetric $\alpha\in(0,2)$ stable process (possibly tempered), $\sigma:\mathbb{R}_+\times\mathbb{R}^{nd}\to\mathbb{R}^d\otimes\mathbb{R}^d$. Observe that with this definition, $(X_t)_{t\geq 0}=(X_t^1,\cdots,X_t^n)_{t\geq 0}$ is $\mathbb{R}^{nd}$ valued.

We point out the particular form of the drift coefficient. Denoting the matrix $A_t\in\mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$ whose entries are the $d\times d$ matrices appearing in the above equation, that is

$$A_t=\begin{pmatrix} a_t^{1,1} & \dots & \dots & \dots & a_t^{1,n}\\ a_t^{2,1} & \ddots & & & a_t^{2,n}\\ 0 & a_t^{3,2} & \ddots & & a_t^{3,n}\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \dots & 0 & a_t^{n,n-1} & a_t^{n,n}\end{pmatrix},$$

we see that A_t is zero in the bottom left corner. This serves to model the propagation of the noise in the system. Indeed, in order for the noise to reach the k^{th} component, the noise must go through the $k-1$ previous components. Consequently, to ensure existence of the solution, in addition to the usual uniform ellipticity on σ, we will assume a kind of weak Hörmander condition on A_t. This can be seen as a way to ensure the transmission of the noise. Indeed, the uniform ellipticity guarantees that σ diffuses the noise in the first d component, whereas the weak Hörmander condition provides the transmission of the noise from a component to another.

Specifically, we make the following assumptions on the coefficients.

[HD-1]: (**Hölder regularity**) $\exists H>0,\ \eta\in(0,1],\ \forall x,y\in\mathbb{R}^{nd}$ and $\forall t\geq 0$,

$$||\sigma(t,x)-\sigma(t,y)||\leq H|x-y|^\eta.$$

[HD-2]: (**Ellipticity**) $\exists\ \overline{c},\ \underline{c}>0$, $\forall\xi\in\mathbb{R}^d$, $\forall z\in\mathbb{R}^{nd}$ and $\forall t\geq 0$,

$$\underline{c}|\xi|^2\leq\langle\xi,\sigma\sigma^*(t,z)\xi\rangle\leq\overline{c}|\xi|^2. \tag{5.2}$$

Also, we assume for all $x\in\mathbb{R}^{nd}$, $t>0$ that $\|\sigma(t,x)\|\leq\overline{c}$.

[HD-3]: (**Hörmander-like condition for** $(A_t)_{t\geq 0}$) $\exists\overline{\alpha},\ \underline{\alpha}\in\mathbb{R}_+^*$, $\forall\xi\in\mathbb{R}^d$ and $\forall t\geq 0$, $\underline{\alpha}|\xi|^2\leq\langle a_t^{i,i-1}\xi,\xi\rangle\leq\overline{\alpha}|\xi|^2,\ \forall i\in[\![2,n-1]\!]$. Also, for all $(i,j)\in[\![1,n]\!]^2,\ \|a_t^{i,j}\|\leq\overline{\alpha}$.

Now, let us turn to the noise. We consider the case where $(Z_t)_{t\geq 0}$ a symmetric Lévy process defined on some filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$. A degenerate equation of the form (5.1) has been investigated in the Brownian setting by Delarue and Menozzi [DM10], with a fully non linear drift. Also, it appears natural to investigate the possible extensions to the stable case. It turns out that additional difficulties with respect to the Brownian setting appear, leading us to temper the noise to derive density estimates. However, some results, concerning the well posedness of the martingale problem, remain true for a stable driving noise under dimension constraints. Therefore, we consider two sets of assumptions, one corresponding to the standard symmetric α stable process, and one for the tempered stable process.

Stable Case: $(Z_t)_{t\geq 0}$ is a symmetric α stable process, that is:

$$\mathbb{E}(e^{i\langle p,Z_t\rangle}) = \exp\left(-t\int_{S^{d-1}} |\langle p,\varsigma\rangle|^\alpha \mu(d\varsigma)\right),\ \forall p\in\mathbb{R}^d.$$

In that case we suppose

[HD-4]: (**Non degeneracy of the spectral measure**) We assume that μ is absolutely continuous w.r.t. to the Lebesgue measure of S^{d-1} with Lipschitz density h and that there exist $\Lambda_1,\Lambda_2\in\mathbb{R}_+^*$, s.t. for all $u\in\mathbb{R}^d$,

$$\Lambda_1|u|^\alpha \leq \int_{S^{d-1}} |\langle u,\varsigma\rangle|^\alpha \mu(d\varsigma) \leq \Lambda_2|u|^\alpha. \tag{5.3}$$

Tempered Case: $(Z_t)_{t\geq 0}$ is a tempered stable process, that is, a Lévy process with generator:

$$L_Z\phi(x) = \int_{\mathbb{R}^d} \{\phi(x+z)-\phi(x)-\frac{\langle\nabla\phi(x),z\rangle}{1+|z|^2}\}g(|z|)\nu(dz),\ \phi\in C_0^2(\mathbb{R}^d,\mathbb{R}), \tag{5.4}$$

where the measure ν is as in the stable case and the tempering function $g:\mathbb{R}^{+*}\to\mathbb{R}^{+*}$ satisfies

[T]: **(Smoothness, Doubling property and Decay associated with the tempering function** g**)** We first assume that there exists $a>0$ s.t. $g\in C^1([0,a],\mathbb{R}^{+*})$ if $\alpha\in(0,1)$ and $g\in C^2([0,a],\mathbb{R}^{+*})$ if $\alpha\in[1,2)$. We also suppose that there exists $c>0$ s.t. for all $r>0, g(r)\leq c\theta(r)$ where $\theta:\mathbb{R}^{+*}\to\mathbb{R}^{+*}$ is a bounded non-increasing function satisfying:

$$\exists D\geq 1,\ \forall r>0,\ \theta(r)\leq D\theta(2r),\ (1+r)\theta(r) := \Theta(r) \underset{r\to+\infty}{\to} 0.$$

Typical examples of tempering functions satisfying **[T]** are for instance $r\to g(r) = \exp(-cr), c>0, g(r)=(1+r)^{-m},\ m\geq 2$.

We say that **[HDS]** (resp. **[HDT]**) holds if conditions **[HD-1]** to **[HD-4]** are fulfilled and the driving noise Z is a symmetric stable process (resp. a tempered stable process satisfying **[T]**). We say that **[HD]** is satisfied if **[HDS]** or **[HDT]** holds, i.e. the results under **[HD]** hold for both the stable and the tempered stable driving process.

Our main results are the following.

Theorem 5.1 (**Weak Uniqueness**). *Under **[HD]**, i.e. in both the stable and the tempered stable case, the martingale problem associated with the generator $(L_t)_{t\geq 0}$, of the degenerate equation* (5.1)*:*

$$dX_t = A_t X_t dt + B\sigma(t, X_{t-})dZ_t,$$

admits a unique solution provided that $d(1-n)+1+\alpha > 0$. That is, for every $x \in \mathbb{R}^{nd}$, there exists a unique probability measure $\mathbb{P}$ on $\Omega = \mathcal{D}(\mathbb{R}_+, \mathbb{R}^{nd})$ the space of càdlàg functions, such that for all $f \in C_0^{1,2}(\mathbb{R}_+ \times \mathbb{R}^{nd}, \mathbb{R})$, denoting by $(X_t)_{t\geq 0}$ the canonical process, we have:

$$\mathbb{P}(X_0 = x) = 1 \quad and \quad f(t, X_t) - \int_0^t (\partial_u + L_u) f(u, X_u) du \quad is\ a\ \mathbb{P}\text{- } martingale.$$

Hence, weak uniqueness holds for (5.1).

The dimension constraints appearing in this theorem are due to a specificity of the degenerate case. Namely, we plan to prove weak uniqueness using a parametrix approach. It turns out that the so-called parametrix kernel does not present the expected smoothing effect (yielding weak uniqueness), unless the dimension constraint $d(1-n)+1+\alpha > 0$ is satisfied. This is due to the fact that in our approach, we are led to consider nd-dimensional stable processes whose spectral measures are either non equivalent or singular with respect to the Lebesgue measure on S^{nd-1}. Also, when $d = 1$ and $n = 2$ in (5.1) we are able to prove the following density estimates in the tempered case.

Theorem 5.2 (**Density Estimates**). *Assume that $d = 1$, $n = 2$. Under **[HDT]** and for $\sigma(t,x) := \sigma(t, x_2)$, i.e. the diffusion coefficient depends on the fast component, provided $1 \geq \eta > \frac{1}{(1\wedge\alpha)(1+\alpha)}$, the unique weak solution of* (5.1) *has for every $s > 0$ a density with respect to the Lebesgue measure. Precisely, for all $0 \leq t < s$ and $x \in \mathbb{R}^2$,*

$$\mathbb{P}(X_s \in dy | X_t = x) = p(t, s, x, y)dy. \tag{5.5}$$

Also, for a deterministic time horizon $T > 0$, and a fixed threshold $K > 0$, there exists $C \geq 1$, s.t. $\forall 0 \leq t < s \leq T$, $\forall (x, y) \in (\mathbb{R}^2)^2$,

$$p(t, s, x, y) \leq C\bar{p}_{\alpha,\Theta}(t, s, x, y) \left(1 + \log(K \vee |(\mathbb{T}^\alpha_{s-t})^{-1}(y - R_{s,t}(x))|\right), \tag{5.6}$$

where for all $u \in \mathbb{R}_+$, $\mathbb{T}^\alpha_u := \mathrm{Diag}\big((u^{1/\alpha}, u^{1+1/\alpha})\big)$, and

$$\bar{p}_{\alpha,\Theta}(t, s, x, y) = \frac{\bar{C}_{\alpha,\Theta}(s-t)^{-\left(1+\frac{2}{\alpha}\right)}}{K + \left(\frac{|(y-R_{s,t}x)^1|}{(s-t)^{\frac{1}{\alpha}}} + \frac{|(y-R_{s,t}x)^2|}{(s-t)^{1+\frac{1}{\alpha}}}\right)^{2+\alpha}} \Theta\left(|(y - R_{s,t}x)^1| + \frac{|(y-R_{s,t}x)^2|}{(s-t)}\right).$$

Here, $R_{s,t}$ stands for the resolvent associated with the deterministic part of (5.1)*, i.e.* $\frac{d}{ds}R_{s,t} = A_s R_{s,t},\ R_{t,t} = I_2$*, and* $\bar{C}_{\alpha,\Theta}$ *is s.t.* $\int_{\mathbb{R}^2} \bar{p}_{\alpha,\Theta}(t,s,x,y)dy = 1$.

Eventually for $0 < T \le T_0 := T_0(\textbf{[HD]}, K)$ *small enough, the following diagonal lower bound holds* $\forall 0 \le t < s \le T,\ \forall (x,y) \in (\mathbb{R}^2)^2$ *s.t. when*

$$|(\mathbb{T}^{\alpha}_{s-t})^{-1}(y - R_{s,t}(x))| \le K,\ p(t,s,x,y) \ge C^{-1}\det(\mathbb{T}^{\alpha}_{s-t})^{-1}. \tag{5.7}$$

In the last Theorem, we actually show that under our dimension constraints, the solution of a degenerate SDE driven by a tempered stable process admits an upper bound homogeneous to a multi-scale tempered stable process, up to a logarithmic correction. The multi-scale behavior can be seen in the fact that each component is normalized by its intrinsic time scale. Also, the diagonal lower bound says that we have the expected behavior (stable estimate) in the characteristic sets.

As we mentioned above, we tackle these problems with a parametrix technique. The strategy is to write the parametrix series, prove the convergence to derive existence and exploit the smoothing effect of the parametrix kernel to derive weak uniqueness in the lines of [BP09], [Men11]. However, the degenerate case is much more subtle, which leads us to the restrictions of the theorem. We present here the various problems we encountered.

A key tool in our approach is the Frozen Process. In the non degenerate case, we simply freeze the coefficients at a terminal point y. However, in the degenerate setting, we have to be more careful because of the transport of the initial condition. Let us consider the following particular case that already illustrates the encountered difficulties:

$$dX_t^1 = dZ_t,\ dX_t^2 = X_t^1 dt,\ \cdots,\ X_t^n = X_t^{n-1}dt, X_0 = (x^1, \dots, x^n).$$

Integrating from line to line yields:

$$\begin{aligned}
X_t^1 &= x^1 + Z_t, \\
X_t^2 &= x^2 + tx^1 + \int_0^t Z_s ds, \\
&\vdots \\
X_t^n &= x^n + \cdots + \frac{t^{n-1}}{(n-1)!}x^1 + \int_0^t ds_1 \dots \int_0^{s_{n-2}} Z_{s_{n-1}} ds_{n-1}.
\end{aligned}$$

Then, the random part in $(X_t)_{t\ge 0}$ comes from the stable process and its iterated integrals in time. Observe that the initial condition $x = (x^1, \dots, x^n)$ is transported through the deterministic system. Define:

$$\left(x^1, x^2 + tx^1, \dots, x^n + \cdots + \frac{t^{n-1}}{(n-1)!}x^1\right)^* = R_t x.$$

Then, $R_t x$ is the solution of $\frac{d}{dt}R_t x = AR_t x$, with $R_0 x = x$, and A is the sub diagonal matrix:

$$A = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

In the general setting, $R_{s,t}$ is the resolvent associated with A_t in (5.1). It is a $nd \times nd$ matrix, and satisfies $\frac{d}{ds}R_{s,t} = A_s R_{s,t}$, with $R_{t,t} = I$, the identity matrix. By symmetry of $(Z_t)_{t\geq 0}$, we can think of the transport $R_t x$ as the "mean" of the process (even though for $\alpha < 1$ the expectation is not defined), or at least the value around which the process X will fluctuate.

Moreover, we point out the different time scales of the system: the first component has the scale of the stable process $t^{1/\alpha}$, the second component is the integral of the stable process and behaves as $t^{1+1/\alpha}$ and so on. Consequently, when turning to density estimates, we investigate the deviations from the typical behavior, renormalized by the typical time scale. In our degenerate framework, the quantity of interest is consequently $|(\mathbb{T}^\alpha_{s-t})^{-1}(y - R_{s,t}x)|$, where:

$$\mathbb{T}^\alpha_{s-t} = \begin{pmatrix} (s-t)^{\frac{1}{\alpha}} I_{d\times d} & & & 0 \\ 0 & (s-t)^{1+\frac{1}{\alpha}} I_{d\times d} & & 0 \\ & & \ddots & \\ 0 & & & (s-t)^{n-1+\frac{1}{\alpha}} I_{d\times d} \end{pmatrix}.$$

Note that this quantity is in fact the quantity appearing in our density estimates of Theorem 2.2. When $|(\mathbb{T}^\alpha_{s-t})^{-1}(y - R_{s,t}x)| \leq C$, we will speak of diagonal regime, whereas when $|(\mathbb{T}^\alpha_{s-t})^{-1}(y - R_{s,t}x)| \geq C$, we will say the off-diagonal regime holds.

Let us specify how the frozen process can be seen as a multi-scale stable process. First, let us define explicitly the frozen process. Fixing a terminal time T and a terminal spatial point $y \in \mathbb{R}^{nd}$, we define:

$$d\tilde{X}^{T,y}_t = A_t \tilde{X}^{T,y}_t dt + B\sigma(t, R_{t,T}y)dZ_t,\ X_0 = x. \tag{5.8}$$

We point out that we actually freeze at point $R_{t,T}y$ in order to compensate the transport of the initial condition as mentioned above. In addition, we did not alter the drift part because since it is linear, it vanishes in the difference of the generators. This equation is a linear SDE and can be solved explicitly using the resolvent $(R_{s,t})_{0\leq s\leq t\leq T}$:

$$\tilde{X}^{t,x,T,y}_s = R_{s,t}x + \int_t^s R_{s,u}B\sigma(u, R_{u,T}y)dZ_u,$$

where we used Markovian notations to highlight the initial time and position in $\tilde{X}_s^{t,x,T,y}$. This identity becomes useful when computing the Fourier transform of the frozen process. Also, this expression gives that $\tilde{X}_s^{t,x,T,y}$ can be seen as a multi-scale stable (possibly tempered) process. Let us specify how this property appears in the case where $n = 2$ and Z is a symmetric stable process. In this case the exponent writes $\forall s \in (t, T]$, $\forall p = (p^1, p^2) \in \mathbb{R}^{2d}$:

$$\mathbb{E}(e^{i\langle p, \tilde{X}_s^{t,x,T,y}\rangle}) = e^{i\langle p, R_{s,t}x\rangle} \exp\left(- \int_t^s du \int_{S^{d-1}} |\langle p^1 + (s-u)p^2, \sigma(u, R_{u,T}y)\varsigma\rangle|^\alpha \mu(d\varsigma) \right).$$

Let us focus on the exponent. Changing variables in the time integral yields:

$$\int_t^s du \int_{S^{d-1}} |\langle p^1 + (s-u)p^2, \sigma(u, R_{u,T}y)\varsigma\rangle|^\alpha \mu(d\varsigma) =$$
$$\int_0^1 dv \int_{S^{d-1}} |\langle (s-t)^{1/\alpha} p^1 + v(s-t)^{1+1/\alpha} p^2, \sigma_v \varsigma\rangle|^\alpha \mu(d\varsigma),$$

where we set $\sigma_v = \sigma(s-(s-t)v, R_{s-(s-t)v,T}y)$. Now, the scalar product in $\mathbb{R}^d$ can be written as a scalar product in $\mathbb{R}^{2d}$:

$$\langle (s-t)^{1/\alpha} p^1 + v(s-t)^{1+1/\alpha} p^2, \sigma_v\varsigma\rangle = \left\langle \mathbb{T}^\alpha_{s-t} p, \begin{pmatrix} \sigma_v\varsigma \\ v\sigma_v\varsigma \end{pmatrix} \right\rangle,$$

recalling $\mathbb{T}^\alpha_{s-t} p = \begin{pmatrix} (s-t)^{1/\alpha} p^1 \\ (s-t)^{1+1/\alpha} p^2 \end{pmatrix}$. Denoting $M_v\varsigma = \begin{pmatrix} \sigma_v\varsigma \\ v\sigma_v\varsigma \end{pmatrix}$, the exponent becomes:

$$\begin{aligned} &\int_t^s du \int_{S^{d-1}} |\langle p^1 + (s-u)p^2, \sigma(u, R_{u,T}y)\varsigma\rangle|^\alpha \mu(d\varsigma) \\ &= \int_0^1 dv \int_{S^{d-1}} \left| \left\langle \mathbb{T}^\alpha_{s-t} p, \frac{1}{|M_v\varsigma|} M_v\varsigma \right\rangle \right|^\alpha |M_v\varsigma|^\alpha \mu(d\varsigma) \\ &= \int_{S^{2d-1}} |\langle \mathbb{T}^\alpha_{s-t} p, \eta\rangle|^\alpha \mu_S(d\eta), \end{aligned}$$

where we defined $\mu_S(d\eta)$ to be the symmetrized image measure of $|M_v\varsigma|^\alpha dv\mu(d\varsigma)$ by the application $(v, \varsigma) \mapsto \frac{1}{|M_v\varsigma|} M_v\varsigma \in S^{2d-1}$. This measure depends on the time horizon $T > 0$, the initial time t and the terminal point y. Now, because of the uniform ellipticity of σ and the scaled form of the resolvent, this measures satisfies **[HD-4]** (non degeneracy of the spectral measure), that is:

$$\bar{\Lambda}_1 |\mathbb{T}^\alpha_{s-t} p|^\alpha \le \int_{S^{2d-1}} \langle \mathbb{T}^\alpha_{s-t} p, \eta\rangle \mu_S(d\eta) \le \bar{\Lambda}_2 |\mathbb{T}^\alpha_{s-t} p|^\alpha.$$

We refer to the text, especially Section 5.2 of Chapter 3 for details on that construction.

Consequently, we obtained the identity:

$$\mathbb{E}(e^{i\langle p, \tilde{X}_s^{t,x,T,y}\rangle}) = e^{i\langle p, R_{s,t}x\rangle} \exp\left(-\int_{S^{2d-1}} |\langle \mathbb{T}_{s-t}^{\alpha} p, \eta\rangle|^{\alpha} \mu_S(d\eta)\right).$$

Now, denoting $(S_u)_{u\geq 0}$ the Lévy process in $\mathbb{R}^{2d}$ whose Lévy exponent writes

$$\int_{S^{2d-1}} |\langle p, \eta\rangle|^{\alpha} \mu_S(d\eta),$$

we have the identity in law at fixed $s, t > 0$:

$$\tilde{X}_s^{t,x,T,y} \overset{\text{(law)}}{=} R_{s,t}x + \mathbb{T}_{s-t}^{\alpha} S_1.$$

Consequently, the marginals of $\tilde{X}_s^{t,x,T,y}$ can be identified with a multi-scale stable process in dimension $2d$, where the various scales are read in the matrix $\mathbb{T}_{s-t}^{\alpha}$, whose spectral measure satisfies a non degeneracy condition ensuring existence of the density. Thus, estimates on the density of $\tilde{X}_s^{t,x,T,y}$ will be obtained from the estimates on the density of $(S_u)_{u\geq 0}$. However, we discussed in Section 3 that the estimates on the density of a stable process are related to the dimension of the support of the spectral measure. In our construction, the support of μ_S will be the image of the support of $|M_v\varsigma|^{\alpha} dv\mu(d\varsigma)$ by the application $(v, \varsigma) \mapsto \frac{1}{|M_v\varsigma|} M_v\varsigma \in S^{2d-1}$. Assuming that the support of the driving process $(Z_t)_{t\geq 0}$ is the sphere S^{d-1}, we see that the dimension of the support of μ_S will be $d - 1 + 1 = d$. Thus, from Watanabe's estimates recalled in Section 3, we deduce that:

$$\tilde{p}^{T,y}(t, s, x, z) \leq C \frac{\det(\mathbb{T}_{s-t}^{\alpha})^{-1}}{\Big(1 + |(\mathbb{T}_{s-t}^{\alpha})^{-1}(y - R_{s,t}x)|\Big)^{d+\alpha+1}}. \tag{5.9}$$

This is what leads us to consider some restrictions on the dimensions in the theorem. We emphasize that even in the tempered case, we cannot get rid of these restrictions, as they come from the stable contribution. Note that this construction is general and can be conducted for any $n, d \in \mathbb{N}$. Viewing the density of the stable process Z and its iterated integrals as the density of an nd-dimensional multi-scale stable process yields to consider a Lévy measure on $\mathbb{R}^{nd}$ for which the support of the spectral measure has dimension $(d - 1) + 1 = d$. In an ambient space of dimension nd, a polynomial tail of order $d + \alpha + 1$ is integrable only when $\alpha > (n - 1)d - 1$. In practice the condition is fulfilled for:

- $d = 1, n = 2$ for $\alpha \in (0, 2)$.
- $d = 1, n = 3$ for $\alpha \in (1, 2)$.

- $d = 2, n = 2$ for $\alpha \in (1, 2)$.

When using a parametrix technique as described in Section 2, we are led to apply the difference of the generators to the density of the frozen process. This brings us to another difficulty in the degenerate setting: the non-local character of the generators. Here, we write according to the notations in Section 2,

$$\begin{aligned} L_t(x, \nabla_x)\varphi(x) &= \langle \nabla_x \varphi(x), A_t x\rangle \\ &+ \int_{\mathbb{R}^d} \varphi(x + B\sigma(t, x)z) - \varphi(x) - \langle \nabla_x \varphi(x), B\sigma(t, x)z\rangle \mathbf{1}_{\{|z|\leq 1\}}\nu(dz), \end{aligned}$$

the generator of the solution of (5.1), and

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

the generator of the Frozen process. We look for an upper bound for $H(t, T, x, y) = \Big(L_t(x, \nabla_x) - L_t(R_{t,T}y, \nabla_x)\Big)\tilde{p}^{T,y}(t, T, x, y)$. First, assume the driving process is a rotationally invariant α stable process in $\mathbb{R}^d$. Then, the Lévy measure becomes: $\nu(dz) = C_{\alpha,d}|z|^{-d-\alpha}dz$. Also, instead of investigating $\Big(L_t(x, \nabla_x) - L_t(R_{t,T}y, \nabla_x)\Big)\tilde{p}^{T,y}(t, T, x, y)$, let us simplify the problem by investigating an upper bound for

$$L\tilde{p}^{T,y}(t, T, x, y) = \int_{\mathbb{R}^d} \Big(\tilde{p}^{T,y}(t, T, x + Bz, y) - \tilde{p}^{T,y}(t, T, x, y)\Big)\mathbf{1}_{\{|z|\geq (T-t)^{1/\alpha}\}}\frac{dz}{|z|^{d+\alpha}},$$

which is the typical quantity to investigate thanks to the smoothness of the coefficients. This corresponds to the large jumps part of the generator of a rotationally invariant stable process in $\mathbb{R}^d$. The small jumps part can be dealt with Fourier arguments. Assume that $\tilde{p}^{T,y}(t, T, x, y)$ is in the off-diagonal regime. Then, when $z \notin B((x - R_{t,T}y)^1, \varepsilon|(x - R_{t,T}y)^1|)$, we have

$$|(\mathbb{T}^{\alpha}_{T-t})^{-1}(x + Bz - R_{t,T}y)| \geq \frac{|z - (R_{t,T}y - x)^1|}{(T - t)^{1/\alpha}} \geq \varepsilon\frac{|(x - R_{t,T}y)^1|}{(T - t)^{1/\alpha}},$$

and $\tilde{p}^{T,y}(t, T, x + Bz, y)$ is off-diagonal with the expected estimate. On the contrary, when $z \in B((x - R_{t,T}y)^1, \varepsilon|(x - R_{t,T}y)^1|)$, we see that the term $\tilde{p}^{T,y}(t, T, x + Bz, y)$ can be in the diagonal regime. Recall that we want to give an upper bound on $L\tilde{p}^{T,y}(t, T, x, y)$ that is homogeneous to the density (up to a singularity). On this set, we can only use the global diagonal bound for $\tilde{p}^{T,y}(t, T, x + Bz, y)$. Since we assumed $\tilde{p}^{T,y}(t, T, x, y)$ to be off-diagonal, this is not the right estimate. We refer to this phenomena as the *rediagonalization*.

Nevertheless, observe that in this situation, we actually have $z \asymp |(x - R_{t,T}y)^1|$. Thus, we can use the Lévy measure to obtain the polynomial decay. We write $\frac{1}{|z|^{d+\alpha}} \asymp \frac{1}{|(x-R_{t,T}y)^1|^{d+\alpha}}$, and we can take this part out of the integral, to be left with the integral of the density. However, we see that we have a gap between the power obtained $d + \alpha$ with this procedure and power of the frozen density decay $d + 1 + \alpha$. In the non degenerate case, this phenomenon already occurs, but the density of the frozen process behaves in large x as $|x|^{-d-\alpha}$, which is exactly the obtained bound. There is a dimension mismatch between the tail behavior of $\tilde{p}^{T,y}(t, T, x, \cdot)$, density of $\tilde{X}_t^{T,y}$, multi-scale stable process of dimension nd, and the one of the jump, stable process of dimension d. Let us mention that this yields an additional diagonal singularity at least of order $(T - t)^{-1/\alpha}$, with respect to the expected one for the kernel $H(t, T, x, y)$ (see Lemma 3.7 and Remark 8.2 in Chapter 3 for details).

Tempering the noise allows to correct the gap between the expected power and the one obtained. Indeed, we can thus correct the concentration index to the expected one. This procedure then yields density estimates through a parametrix continuity technique, under additional dimension constraints. These are due to technical reasons that we detail in the text. We establish the density estimates when $d = 1, n = 2$ (scalar non-degenerate diffusion and associated non-degenerate integral) the expected upper-bound up to an additional logarithmic contribution, when the coefficient $\sigma(t, x) = \sigma(t, x^2)$ depends on the *fast* variable (see Section 5.3 of Chapter 3). This dependence provides a better smoothing property of the parametrix kernel $H(t, T, x, y) = \Big(L_t(x, \nabla_x) - L_t(R_{t,T}y, \nabla_x)\Big)\tilde{p}^{T,y}(t, T, x, y)$. This is due to the fact that the difference of the generators yields a multiplicative term in $|x - R_{t,T}y|^{\eta(\alpha\wedge 1)}$ in the full dependence case, or $|(x-R_{t,T}y)^2|^{\eta(\alpha\wedge 1)}$, in the case $\sigma(t, x) = \sigma(t, x^2)$. To make this contribution homogeneous to $|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|$, which is the quantity appearing in the density estimate (5.9), we have to recover the scale matrix $\mathbb{T}^{\alpha}_{T-t}$. In the full dependence case, we obtain $|x-R_{t,T}y|^{\eta(\alpha\wedge 1)} \le (T-t)^{\eta\left(1\wedge\frac{1}{\alpha}\right)}|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|$, whereas for $\sigma(t, x) = \sigma(t, x^2)$, we have $|(x-R_{t,T}y)^2|^{\eta(\alpha\wedge 1)} \le (T-t)^{\left(1+\frac{1}{\alpha}\right)\eta(\alpha\wedge 1)}|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|$. Observe anyhow that this better smoothing effect is required to compensate the additional singularities brought by the rediagonalization. Also, let us mention that this additional multiplicative term $|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}$ is precisely why we chose freeze at point $R_{t,T}y$. This ensures a compatibility between the estimate on the parametrix kernel H and the frozen process $\tilde{p}$.

6 A Multi-step Richardson-Romberg extrapolation method for stochastic approximation.

Stochastic approximation (SA) algorithms are simulation based procedures to approximate the zeros of a function $h : \mathbb{R}^d \to \mathbb{R}^d$ which writes $h(\theta) = \mathbb{E}[H(\theta, U)]$, for some $\mathbb{R}^d$-valued random vector U. The function, H is assumed to be known by the experimenter, and it is implicitly supposed that the computation of h is more costly (in terms of computational time) than the simulation of U and the computation of H. For simplicity, we assume that h has only one zero θ^*, even though the theory extends to the case of multiple zeros.

Robbins and Monro in [RM51] proposed the following recursive algorithm to approximate θ^*. Let $(U^p)_{p\geq 1}$ be an *i.i.d.* sequence of random variables with the same law as U, and θ_0 independent of the sequence $(U^p)_{p\geq 1}$ with $\mathbb{E}[|\theta_0|^2] < +\infty$. We consider the following recursive scheme:

$$\theta_{p+1} = \theta_p - \gamma_{p+1} H(\theta_p, U^{p+1}),\ p \geq 0. \tag{6.1}$$

Here, $\gamma = (\gamma_p)_{p\geq 1}$ is a deterministic and decreasing sequence of non-negative numbers satisfying the assumptions

$$\sum_{p\geq 1} \gamma_p = +\infty, \quad \text{and} \quad \sum_{p\geq 1} \gamma_p^2 < +\infty. \tag{6.2}$$

Observe that in the deterministic case i.e. $H(\theta, u) = h(\theta)$, when h is the gradient of a convex potential, we recover the classical descent gradient procedure. Concerning the *a.s.* convergence of the scheme (6.1), we have the following theorem.

Theorem 6.1. *Assume that h satisfies the mean-reverting condition*

$$\forall \theta \neq \theta^*,\ \langle \theta - \theta^*, h(\theta)\rangle > 0,$$

and that

$$\forall \theta \in \mathbb{R}^d,\ \mathbb{E}[|H(\theta, U)|^2] \leq C(1 + |\theta - \theta^*|^2).$$

Then the scheme (6.1) *satisfies $\theta_p \to \theta^*$ a.s. when $p \to +\infty$.*

We also mention that the above Theorem admits several extensions. A more general result involving Lyapunov functions can be found in the literature. We refer to Duflo [Duf96] or Kushner and Yin [KY03] for more details. Besides, a rate of convergence is obtained by assuming more regularity.

Proposition 6.2. *Assume that h is twice continuously differentiable in a neighborhood of θ^*, and that the eigenvalues of $Dh(\theta^*)$ have strictly positive real parts. Assume as well that the function H satisfies:*

- $\Gamma : \theta \mapsto \mathbb{E}[H(\theta, U)H(\theta, U)^*]$ *is continuous on* $\mathbb{R}^d$*, and* $\Gamma(\theta^*)$ *is positive definite matrix.*
- $\exists \varepsilon > 0$ *such that* $\theta \mapsto \mathbb{E}[|H(\theta, U)|^{2+\varepsilon}]$ *is locally bounded on* $\mathbb{R}^d$.

Finally, assume that the step sequence γ_p *is given by a function* $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$*, decreasing to zero, that is,* $\forall p \in \mathbb{N},\ \gamma(p) = \gamma_p$ *that satisfies one of the two assumptions below:*

- *there exists* $a \in (1/2, 1)$ *such that for any* $x > 0$, $\lim_{t \to +\infty} \frac{\gamma(tx)}{\gamma(t)} = \frac{1}{x^a}$*. In this case, set* $\zeta = 0$.
- *for* $t \geq 1$, $\gamma(t) = \gamma_0/t$ *and* γ_0 *is such that* $2\mathcal{R}e(\lambda_{min})\gamma_0 > 1$*, where* λ_{min} *is the eigenvalue of* $Dh(\theta^*)$ *with the lowest real part. In this case, we set* $\zeta = 1/(2\gamma_0)$.

Then, there exists a positive definite matrix

$$\Sigma = \int_0^{+\infty} \exp\Big(- s(Dh(\theta^*) - \zeta I_d)\Big)^* \Gamma(\theta^*) \exp\Big(- s(Dh(\theta^*) - \zeta I_d)\Big) ds,$$

such that the convergence holds when p *tends to infinity:*

$$\gamma(p)^{-1/2}(\theta^* - \theta_p) \underset{p \to +\infty}{\Longrightarrow} \mathcal{N}(0, \Sigma).$$

These results are standard, and we refer to Duflo [Duf96] for a quite extensive presentation of the matter. In addition, we mention some recent non-asymptotic concentration bounds in [FM12] and [FF13]. Finally, let us mention that there exist results on stochastic algorithms where innovations are Markovian and satisfies mixing properties. See [Lar11] for a detailed presentation of this topic.

In many applications, notably in mathematical finance, the random variable U is not available through exact simulation, and a first step of approximation is needed. For instance, U may be given by X_T, where $(X_t)_{t \in [0,T]}$ is the solution of an SDE with dynamics:

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

for $(Z_t)_{t \geq 0}$ a symmetric α-stable process, $\alpha \in (0, 2]$, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$. In that case, one has to consider a numerical scheme like the Euler-Maruyama continuous approximation scheme.

We introduce a collection of random variables $(U^n)_{n \in \mathbb{N}}$ such that $U^n \to U$, weakly or strongly, and define accordingly $h^n(\theta) = \mathbb{E}[H(\theta, U^n)]$. Keeping the example of the solution of an SDE, one can take $(U^n)_{n \in \mathbb{N}}$ to be the value at maturity X_T^n of the associated Euler-Maruyama scheme:

$$X_t^n = x + \int_0^t b(\phi(s), X_{\phi(s)}^n) ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^n) dZ_s,$$

with $\phi(s) = \sup\{t_i : t_i \leq s\}$, for a given time step $\Delta = \frac{1}{n}$, $n \in \mathbb{N}^*$, setting for all $i \in \mathbb{N}$, $t_i = i\Delta$. In this case, for instance, when the coefficients b and σ are locally Lipschitz with linear growth, we have: $X^n \xrightarrow{\mathbb{P}} X$ when n tends to infinity (see e.g. Theorem 3.1 in Jacod Protter [JP98] and the references therein).

The goal is to approximate θ^*, the zero of h by $\theta^{*,n}$, the zero of h^n.This induces an implicit discretization error $\mathcal{E}_D(n) = \theta^* - \theta^{*,n}$. Now let us observe that since U^n is easily simulable, we can approximate $\theta^{*,n}$ by M steps in the corresponding stochastic algorithm:

$$\theta^n_{p+1} = \theta^n_p - \gamma_{p+1} H(\theta^n_p, (U^n)^{p+1}), \ p \geq 0, \tag{6.3}$$

This in turn produces a statistical error $\mathcal{E}_S(n, M) = \theta^{*,n} - \theta^n_M$. Consequently, the global error obtained by approximating θ^* by θ^n_M naturally splits in:

$$\begin{aligned} \mathcal{E}_{glob}(n, M) &= \theta^* - \theta^{*,n} + \theta^{*,n} - \theta^n_M \\ &:= \mathcal{E}_D(n) + \mathcal{E}_S(n, M). \end{aligned}$$

Concerning the implicit discretization error, $\mathcal{E}_D(n) = \theta^* - \theta^{*,n}$ the following result is established in [Fri13]:

Proposition 6.3. *For all $n \in \mathbb{N}^*$, assume that h and h^n satisfy the mean reverting assumption:*

$$\forall \theta \neq \theta^*, \ \langle \theta - \theta^*, h(\theta) \rangle > 0 \quad \textit{and} \ \ \forall \theta \neq \theta^{*,n}, \ \langle \theta - \theta^{*,n}, h^n(\theta) \rangle > 0.$$

Moreover, suppose that $(h^n)_{n\geq 1}$ converges locally uniformly towards h, then we have $\theta^{*,n} \underset{n\to+\infty}{\longrightarrow} \theta^*$. Moreover, assume that $\forall \theta \in \mathbb{R}^d$, $n^\alpha (h - h^n)(\theta) \to \Lambda^0_1(\theta)$ when $n \to +\infty$. Then, we have

$$n^\alpha \left(\theta^{*,n} - \theta^*\right) \underset{n\to\infty}{\longrightarrow} Dh(\theta^*)^{-1} \Lambda^0_1(\theta^*).$$

The last result can be seen as an expansion of order one for $\theta^{*,n} - \theta^*$ in power of $n^{-\alpha}$. It is therefore natural to ask if the expansion holds at higher order. Assuming additional regularity on h, h^n allows us to answer positively.

Proposition 6.4. *Under the assumptions of Proposition 6.3, if the following hypotheses hold: for some $R \in \mathbb{N}^*$,*

1. *For all $\theta \in \mathbb{R}^d$,*

$$h(\theta) - h^n(\theta) = \frac{\Lambda^0_1(\theta)}{n^\alpha} + \cdots + \frac{\Lambda^0_R(\theta)}{n^{\alpha R}} + o\left(\frac{1}{n^{\alpha R}}\right). \tag{6.4}$$

2. *$h, h^n \in \mathcal{C}^R(\mathbb{R}^d, \mathbb{R}^d)$ and for all $l \leq R-1$, for all $\theta \in \mathbb{R}^d$,*

$$D^l h^n(\theta) - D^l h(\theta) = \frac{\Lambda_1^l(\theta)}{n^\alpha} + \cdots + \frac{\Lambda_{R-l}^l(\theta)}{n^{\alpha(R-l)}} + o\left(\frac{1}{n^{\alpha(R-l)}}\right) \tag{6.5}$$

   where for all $\theta \in \mathbb{R}^d$, $\Lambda_1^l(\theta), \cdots, \Lambda_{p-l}^l(\theta)$ and $o(n^{-\alpha(p-l)})$ are multilinear maps from $(\mathbb{R}^d)^l$ to $\mathbb{R}^d$.

3. *For all $l \in [\![1, R]\!]$, $(D^l h^n)_{n\geq 1}$ converges locally uniformly towards $D^l h$.*

4. *$Dh(\theta^*)$ is invertible.*

Then, $\theta^{*,n} - \theta^*$ has an expansion up to order R, that is, the following expansion holds:

$$\exists (C_1, \cdots, C_R) \in (\mathbb{R}^d)^R, \quad \theta^{*,n} - \theta^* = \frac{C_1}{n^\alpha} + \cdots + \frac{C_R}{n^{\alpha R}} + o\left(\frac{1}{n^{\alpha R}}\right). \tag{6.6}$$

This result is proved in Theorem 2.3 of Chapter 4, and can be used to produce a Richardson Romberg extrapolation estimator. Let us be more specific on this subject. We introduce a sequence of R random vectors $\{U^{rn}, r \in [\![1, R]\!]\}$, $n \in \mathbb{N}^*$ such that $U^{rn} \xrightarrow{\mathbb{P}} U^r$ as $n \to +\infty$ with $U^r \overset{d}{=} U$, $r \in [\![1, R]\!]$. In the example of an SDE, this sequence comes from the same Brownian motion, considered at various time steps. Under the assumptions of Proposition 6.4, we have for all $r \in [\![1, R]\!]$:

$$\theta^{*,rn} = \theta^* + \sum_{p=1}^{R-1} \frac{C_p}{r^{\alpha p}} \frac{1}{n^{\alpha p}} + \frac{C_R}{r^{\alpha R}} \frac{1}{n^{\alpha R}} (1 + \epsilon_r(n))$$

with $\epsilon_r(n) \to 0$ as $n \to +\infty$. The idea is to find a collection of weights $\mathbf{w}_r$, $r = 1, \ldots, R$ in order to kill the R first terms in the last expansion. Precisely, the last identity writes in matrix form:

$$\begin{pmatrix} \vdots \\ \theta^{*,rn} \\ \vdots \end{pmatrix}_{1\leq r\leq R} = \mathbf{I}\theta^* + V \begin{pmatrix} \vdots \\ \frac{C_r}{n^{\alpha r}} \\ \vdots \end{pmatrix}_{1\leq r\leq R-1} + \begin{pmatrix} \vdots \\ \frac{C_R}{r^{\alpha R}} \frac{1}{n^{\alpha R}} (1 + \epsilon_r(n)) \\ \vdots \end{pmatrix}_{1\leq r\leq R},$$

where $\mathbf{I} = (I_d, \cdots, I_d)^T$ with I_d is the identity matrix of dimension d, and V is the $Rd \times (R-1)d$ Vandermonde matrix $V = [r^{-\alpha p} I_d]_{1\leq r\leq R, 1\leq p\leq R-1}$.

We now look for a weight matrix $\mathbf{w} = (\mathbf{w}_1, \cdots, \mathbf{w}_R)^T$ so that the two conditions are fulfilled:

$$\mathbf{w}^T \mathbf{I} = I_d \quad \text{and} \quad \mathbf{w}^T V = 0_{d\times d(R-1)}. \tag{6.7}$$

We can find $\mathbf{w}$ explicitly using Cramer's rule (see Section 2.2 of Chapter 4):

$$\forall r \in \{1, \cdots, R\}, \quad \mathbf{w}_r = (-1)^{R-r} \frac{r^{\alpha R}}{\prod_{j=0}^{r-1}(r^\alpha - j^\alpha) \prod_{j=r+1}^{R}(j^\alpha - r^\alpha)} I_d, \tag{6.8}$$

For such a $\mathbf{w}$, we obtain that:

$$\sum_{r=1}^{R} \mathbf{w}_r \left(\theta^{*,rn} - \theta^* \right) = \frac{C_R}{n^{\alpha R}} \frac{(-1)^{R-1}}{R!^\alpha} \left(1 + \epsilon_{R+1}(n)\right). \tag{6.9}$$

Now, we approximate each $\theta^{*,nr}$ using M steps in the corresponding stochastic algorithm, that is each $(\theta_p^{rn})_{p\in\mathbb{N}}$ is defined by the corresponding recursive algorithm:

$$\forall r \in [\![1, R]\!], \quad \theta_{p+1}^{rn} = \theta_p^{rn} - \gamma_{p+1} H(\theta_p^{rn}, (U^{rn})^{p+1}), \; p \in [\![0, M-1]\!], \tag{6.10}$$

with $((U^{rn})^p, r = 1, \cdots, R)_{p\in[\![1,M]\!]}$ an i.i.d sequence with the same law as $(U^{rn}, r = 1, \cdots, R)$, θ_0^{rn}, $r = 1 \cdots, R$ the initial conditions independent of the innovation sequence satisfying $\sup_{n\geq 1} \mathbb{E}|\theta_0^n|^2 < +\infty$ and the sequence $(\gamma_p)_{p\geq 1}$ satisfying (6.2). This produces the Richardson Romberg extrapolation estimator:

$$\Theta_M^{n,Rn} = \sum_{r=1}^{R} \mathbf{w}_i \theta_M^{rn}.$$

We now consider the problem of the complexity of the new Richardson Romberg extrapolation estimator. Fixing the tolerance level ε, and denoting $\mathcal{E}_{glob}^{R-R} = \theta^* - \Theta_M^{n,Rn}$, we want to find M and n such that, keeping the global error under ε, the computational cost is minimal. We assume that the cost of a single simulation of U^n is proportional to n and is given by $K \times n$, where K is a generic positive constant independent of n, which is typically the case for the Euler-Maruyama approximation case.

At each step $p \in [\![1, M]\!]$ of our procedure, for every $r \in [\![1, R]\!]$, we have to simulate the random vector $(U^n, U^{2n}, \cdots, U^{Rn})$. The global computational cost is given by

$$\text{Cost(R-R)} := KM \sum_{r=1}^{R} rn = KMn \frac{R(R+1)}{2}.$$

In other words, the problem now states as follows:

$$(n(\varepsilon), M(\varepsilon)) = argmin_{\mathbb{E}|\mathcal{E}_{glob}^{R-R}| \leq \varepsilon} \text{Cost(R-R)}.$$

However, the condition $\mathbb{E}|\mathcal{E}_{glob}^{R-R}| \leq \varepsilon$ is not explicitly tractable and we will consider a suboptimal cost optimization problem.

The new global error splits in:

$$
\begin{aligned}
\mathcal{E}^{R-R}_{glob}(n,M) &= \theta^* - \sum_{r=1}^{p} \mathbf{w}_i \theta^{*,rn} + \sum_{r=1}^{p} \mathbf{w}_i \theta^{*,rn} - \Theta^{n,pn}_M \\
&= \sum_{r=1}^{p} \mathbf{w}_i \Big(\theta^* - \theta^{*,rn}\Big) + \sum_{r=1}^{p} \mathbf{w}_i \Big(\theta^{*,rn} - \theta^{rn}_M\Big) \\
&= \mathcal{E}^{R-R}_D(n) + \mathcal{E}^{R-R}_S(n,M),
\end{aligned}
$$

The key observation is that thanks to (6.9), this new target satisfies the following implicit error:

$$
|\mathcal{E}^{R-R}_D(n)| \leq \frac{|C_R|}{R!^\alpha} n^{-\alpha R} \left(1 + |\epsilon_{R+1}(n)|\right)
$$

Besides, under standard assumptions, the linear combination of stochastic algorithm estimators $\Theta^{n,pn}_M$ converges *a.s.* to the target $\sum_{r=1}^{p} \mathbf{w}_i \theta^{*,rn}$ as the number of steps M goes to infinity. Consequently, to find the optimal $(n(\varepsilon), M(\varepsilon))$, it remains to bound the statistical error $\mathcal{E}^{R-R}_S(n,M)$. Actually, we give an $L^1(\mathbb{P})$ control of the statistical error under the following assumptions.

(HUI) $\exists \delta > 0$, such that $\forall \theta \in \mathbb{R}^d$, $\sup_{n\in\mathbb{N}^*} \mathbb{E}[|H(\theta, U^n)|^{2+\delta}] < +\infty$.

(HC1) $\exists C > 0$ such that $\forall n \in \mathbb{N}^*, \forall \theta \in \mathbb{R}^d$, $\mathbb{E}[|H(\theta,U^n)|^2] \leq C(1 + |\theta - \theta^{*,n}|^2)$.

(HC2) $\forall \theta \in \mathbb{R}^d$, $\mathbb{P}(U \notin \mathcal{C}_\theta) = 0$ with $\mathcal{C}_\theta := \{x \in \mathbb{R}^q : x \mapsto H(\theta,x) \text{ is continuous at } x\}$.

(HRG) There exists $a \in (0,1]$,

$$
\sup_{n\in\mathbb{N}^*, (\theta,\theta')\in(\mathbb{R}^d)^2} \frac{\mathbb{E}|H(\theta,U^n) - H(\theta',U^n)|^2}{|\theta-\theta'|^{2a}} < +\infty.
$$

(HUA) For each $n \in \mathbb{N}^*$, the map $h^n : \theta \in \mathbb{R}^d \mapsto \mathbb{E}[H(\theta,U^n)]$ is continuously differentiable with Dh^n Lipschitz-continuous uniformly in n and there exists $\underline{\lambda} > 0$ s.t. $\inf_{n\in\mathbb{N}^*,\theta\in\mathbb{R}^d} \lambda_{min}\left((Dh^n(\theta) + Dh^n(\theta)^T)/2\right) > \underline{\lambda}$ where $\lambda_{min}(A)$ denotes the lowest eigenvalue of the matrix A. (*Uniform Attractivity*).

(HS) The step sequence is given by $\gamma_p = \gamma(p)$, $p \geq 1$, where γ is a positive function defined on $[0, +\infty[$ decreasing to zero satisfying one of the following assumptions:

- γ varies regularly with exponent $(-\rho)$, $\rho \in (1/2, 1)$, that is, for any $x > 0$, $\lim_{t\to+\infty} \gamma(tx)/\gamma(t) = x^{-\rho}$.

- for $t \geq 1$, $\gamma(t) = \gamma_0/t$ and γ_0 satisfies $2\underline{\lambda}\gamma_0 > 1$.

Then, we have the following estimate for the $L^1(\mathbb{P})$ norm of the statistical error:

Proposition 6.5. *Let $R \in \mathbb{N}^*$. Suppose that for $r \in [\![1, R]\!]$, $U^{rn} \xrightarrow{\mathbb{P}} U^r$ and $\theta_0^n \xrightarrow{\mathbb{P}} \theta_0$, as $n \to +\infty$. Under* ***(H-R)**, **(HUI)**, **(HC1)**, **(HC2)**, **(HRG)**, **(HS)*** *and* ***(HUA)**, one has for some positive constant $C := C(\gamma, \underline{\lambda})$*

$$\mathbb{E}[|\mathcal{E}_S^{R-R}(n,M)|] \leq C\mathbb{E}\left[\left|\sum_{r=1}^{R} \mathbf{w}_r H(\theta^*, U^r)\right|^2\right]^{1/2} \gamma^{1/2}(M)\left(1 + \phi_1^R(n) + \phi_2^R(M)\right)$$

where ϕ_1^R, ϕ_2^R are two positive functions satisfying: $\phi_1^R(n) \to 0$ and $\phi_2^R(M) \to 0$ respectively as $M \to +\infty$, $n \to +\infty$ and ϕ_2^R is non-increasing.

From the two bounds on the discretization error $|\mathcal{E}_D^{R-R}(n)|$ and on the statistical error $\mathbb{E}[|\mathcal{E}_S^{R-R}(n,M)|]$, the global error is bounded by:

$$\begin{aligned} \mathbb{E}|\mathcal{E}_{glob}^{R-R}| &\leq |\mathcal{E}_D^{R-R}(n)| + \mathbb{E}[|\mathcal{E}_S^{R-R}(n,M)|] \\ &\leq \mu_R n^{-\alpha R}\left(1 + |\epsilon_{R+1}(n)|\right) + \nu_R \gamma^{1/2}(M)(1 + \phi_1^R(n) + \phi_2^R(M)), \end{aligned}$$

where $\mu_R = \frac{|C_R|}{R!^\alpha}$ and $\nu_R = C\mathbb{E}\left[\left|\sum_{r=1}^R \mathbf{w}_r H(\theta^*, U^r)\right|^2\right]^{1/2}$. Thus, we are naturally led to consider the following suboptimal computational cost optimization problem

$$(n(\epsilon), M(\epsilon)) = \underset{\mu_R n^{-\alpha R}(1+|\epsilon_{R+1}(n)|)+\nu_R\gamma^{1/2}(M)(1+\phi_1^R(n)+\phi_2^R(M))\leq\varepsilon}{argmin} \mathrm{Cost(R\text{-}R)}, \qquad (6.11)$$

where assuming the step sequence γ is given by: $\gamma(p) = \gamma_0/p^\beta$, $\gamma_0 > 0$, $p > 0$, $\beta \in (1/2, 1]$ admits the following asymptotic cost:

$$\begin{aligned} &\underset{\mu_R n^{-\alpha R}(1+|\epsilon_{R+1}(n)|)+\nu_R\gamma^{1/2}(M)(1+\phi_1^R(n)+\phi_2^R(M))\leq\varepsilon}{argmin} \mathrm{Cost(R\text{-}R)} \\ &\sim K\frac{R(R+1)}{2}\gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\mu_R^{\frac{1}{\alpha R}}\frac{1}{\varepsilon^{\frac{2}{\beta}+\frac{1}{\alpha R}}}\left(1 + \frac{2\alpha R}{\beta}\right)^{\frac{1}{\alpha R}}\left(1 + \frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}, \end{aligned}$$

and the following asymptotic optimal parameters (n, M):

$$n(\varepsilon) \sim \left(\frac{2\alpha R}{\beta} + 1\right)^{\frac{1}{\alpha R}} \mu_R^{\frac{1}{\alpha R}}\varepsilon^{-\frac{1}{\alpha R}} \quad \text{and} \quad M(\varepsilon) \sim \gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1 + \frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}} \quad \text{as } \varepsilon \to 0.$$

Returning to the example of the solution of an SDE, one can ask if the assumptions of Proposition 6.4 are fulfilled in practical cases. Specifically, we see that regularity on the functions h, h^n is required. Since $h(\theta) = \mathbb{E}[H(\theta, U)]$ and $h^n(\theta) = \mathbb{E}[H(\theta, U^n)]$, the regularity can be obtained by two means, either the function H is regular and admits an expansion as specified, or the law of U and U^n are smooth. In particular, when we address the estimation problem of the quantile at level ℓ, of one component of the solution of an SDE, we are typically in the latter case. Indeed, solving $\mathbb{P}(X_T^1 \leq \theta^*) = \ell$ for θ^*, where, $X_T^1 \in \mathbb{R}$ is the first component of $X_T \in \mathbb{R}^d$ can be seen as an equation of the form $h(\theta^*) = \mathbb{E}[H(\theta^*, X_T)] = 0$, where $H(\theta, y) = 1 - \frac{1}{1-\ell}\mathbf{1}_{\{y^1 \geq \theta\}}$, setting $y = (y^1, \cdots, y^d) \in \mathbb{R}^d$. Thus, it is clear that we can derive a stochastic algorithm to approximate $\theta^{*,n}$. However, in this context, the function $H(\theta, y)$ presents low regularity. Thus, we must rely on the regularity of the law of X_T and X_T^n to compensate, as we already mentioned in Section 2. Hopefully, we have that:

$$
\begin{aligned}
h(\theta) - h^n(\theta) &= \mathbb{E}[H(\theta, X_T)] - \mathbb{E}[H(\theta, X_T^n)] \\
&= \frac{1}{1-\ell} \int_{\mathbb{R}^d} \mathbf{1}_{\{y^1 \leq \theta\}} \Big(p(T, x, y) - p_n(T, x, y) \Big) dy \\
&= \frac{1}{1-\ell} \left(\mathbb{P}^x(X_T^1 \leq \theta) - \mathbb{P}^x(X_T^{n,1} \leq \theta) \right),
\end{aligned}
$$

where $p(T, x, \theta)$ is the density of the diffusion, and $p_n(T, x, \theta)$ the density of the Euler-Maruyama scheme at time T, and $X_T^{n,1}$ is the first component of the Euler-Maruyama scheme. Thus, to obtain an expansion for $h(\theta) - h^n(\theta)$, we need an expansion for $(p - p_n)(T, x, y)$. Moreover, denoting by $p^{X_T^1}(T, x, \theta)$ and $p_n^{X_T^{n,1}}(T, x, \theta)$ the marginal densities of X_T^1 and $X_T^{n,1}$, the derivative w.r.t. θ of the previous equality is $\forall k \geq 1$, $\forall (\theta, x) \in \mathbb{R} \times \mathbb{R}^d$:

$$
\frac{d^k}{d\theta^k} h(\theta) - \frac{d^k}{d\theta^k} h^n(\theta) = \frac{1}{1-\ell} \left(\frac{\partial^{k-1}}{\partial \theta^{k-1}} p^{X_T^1}(T, x, \theta) - \frac{\partial^{k-1}}{\partial \theta^{k-1}} p_n^{X_T^{n,1}}(T, x, \theta) \right),
$$

Thus, in order for the assumptions of Proposition 6.4 to hold, we have to get an expansion of the density of the SDE and the density of the Euler-Maruyama scheme and their derivatives. It is known since Talay and Tubaro [TT90] that an expansion for the Euler-Maruyama scheme can be obtained. However, in their approach, regularity on the coefficients and on the test function are needed. We choose a parametrix approach, that holds under mild regularity assumptions. It has the advantage as well to give precise dependence in the variables of the coefficients, which is useful in order to obtain an expansion for the derivatives of $h - h^n$ (see Theorem 2.3 in Section 2). Also, this approach allows us to handle both the Brownian case and the stable case, $\alpha \in (0, 2)$, with similar arguments.

We denote by **[A]** the following set of assumptions. Fix an integer $m \in \mathbb{N}$ referring to the regularity of the coefficients.

[A-1] $b \in \mathcal{C}^m(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in \mathcal{C}^m(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ with bounded derivatives. Also, when $\alpha \leq 1$, we put $b = 0$.

[A-2] There exists $C > 1$ such that for all $x, \xi \in \mathbb{R}^d$, setting $\Sigma(x) = \sigma(x)\sigma(x)^*$:

$$C^{-1}|\xi|^2 \leq \langle \xi, \Sigma(x)\xi \rangle \leq C|\xi|^2.$$

[A-3] When $\alpha < 2$, the spectral measure μ has a positive $\mathcal{C}^m(S^{d-1})$ surface density and satisfies: there exists $C > 1$, such that for all $\xi \in \mathbb{R}^d$:

$$C^{-1}|\xi|^\alpha \leq \int_{S^{d-1}} |\langle \xi, \vartheta \rangle|^\alpha \mu(d\vartheta) \leq C|\xi|^\alpha. \tag{6.12}$$

These are usual assumptions for our setting, see also the papers of Konakov and Mammen [KM02] and Konakov and Menozzi [KM10]. Under the previous set of assumptions, we now prove:

Theorem 6.6. *Assume that **[A]** holds. Let $M \in \mathbb{N}^*$ be such that when $\alpha = 2$, $0 < M \leq m/2$, and when $\alpha < 2$, we assume $m > d+4$ and $0 < M \leq m-(d+4)$. Let $\gamma \in \mathbb{N}^d$, with $|\gamma| \leq M$. Then, for all $x, y \in \mathbb{R}^d$, we have:*

$$\begin{aligned}\partial_y^\gamma p(1,x,y) - \partial_y^\gamma p_n(1,x,y) = \sum_{k=1}^{M-1-|\gamma|} & \frac{1}{(k+1)!n^k}\partial_y^\gamma \left(p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d\right)(1,x,y) \\ & -\frac{1}{(k+1)!n^k}\partial_y^\gamma \left(p^d \otimes_n \left(\tilde{L}_* - \tilde{L}^*\right)^{k+1} p_n\right)(1,x,y) + \frac{\partial_y^\gamma R(x,y)}{n^{M-|\gamma|}}.\end{aligned}$$

*Also, there is a constant $C > 0$ depending on the set of assumptions **[A]**, T, γ, and M such that the following bound holds for each term and the remainders:*

$$\begin{aligned}\sum_{k=1}^{M-|\gamma|-1} \left|\partial_y^\gamma \left(p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d\right)(1,x,y)\right| + \left|\partial_y^\gamma \left(p^d \otimes_n \left(\tilde{L}_* - \tilde{L}^*\right)^{k+1} p_n\right)(1,x,y)\right| & \\ +|\partial_y^\gamma R(x,y)| \leq C\bar{p}_K^\alpha(1,x,y), &\end{aligned}$$

where for a given $K > 0$, we denoted $\bar{p}_K^\alpha(t,x,y)$ the following quantity:

$$\bar{p}_K^\alpha(t,x,y) = \begin{cases} t^{-d/2}\exp\left(-K\frac{|y-x|^2}{t}\right), & \text{if } \alpha = 2, \\ \frac{t^{-d/\alpha}}{\left[K \vee \frac{|y-x|}{t^{\frac{1}{\alpha}}}\right]^{d+\alpha}}, & \text{if } \alpha \in (0,2). \end{cases}$$

Now, we see that **[A]** implies the assumptions of Proposition 6.4 when the regularity m is large enough, so that we can derive an expansion for the difference between the density of an SDE and the density of the Euler-Maruyama scheme. Thus, we obtain the expansion:

$$\theta^{*,n} - \theta^* = \frac{C_1}{n^\alpha} + \cdots + \frac{C_R}{n^{\alpha R}} + o\left(\frac{1}{n^{\alpha R}}\right),$$

for $(C_1, \cdots, C_R) \in (\mathbb{R}^d)^R$, which allows us to use a Richardson Romberg extrapolation in order to reduce the computational cost in the quantile approximation procedure. Numerical illustration is provided in Figure 4.1 of Chapter 4. In particular, the correspondence between the theoretical L^1 controls, used to calibrate the parameters in order to get a given target error bound, and the empirical L^1 error, obtained through multiple runs of the algorithm, appears satisfactory. Let us mention that the problem of the estimation of the quantile of a diffusion has been investigated in [TZ04] as well.

Chapter 2

Density Estimates for SDEs Driven by a Possibly Tempered Stable Processes

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.

1 Introduction

This Chapter is devoted to the study of Stochastic Differential Equations (SDEs), driven by a class of possibly tempered stable process. Specifically, we use a continuity method, known as the parametrix technique. After obtaining preliminary estimates, we are able to prove weak uniqueness to the SDE. Furthermore, we show the existence of the density, as well as some associated estimates, under mild assumptions on the coefficients. More precisely, we study equations with the dynamics:

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

where $F : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded, Hölder continuous 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 is symmetric and 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 probability 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 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 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 probability measure on the sphere called the spectral measure. When the spectral measure satisfies the non-degeneracy condition:

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

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). See Theorem 1.1 in Watanabe [Wat07]. 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 the concentration properties of the spectral measure to deduce density bounds. Also, the concentration of the spectral measure, 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, $\alpha+\gamma \le d$, 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, 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 $(Z_t)_{t\ge 0}$ should transfer to $(X_t)_{t\ge 0}$. 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 σ 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 [Kol00b], that the density $p(t, x, y)$ 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 $x, y \in \mathbb{R}^d$, $t \in (0, T]$:

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

This estimate is obtained using a continuity method: the parametrix technique. This approach is well suited for obtaining density estimates for the solution of an SDE under mild assumptions on the coefficients, provided that good estimates can be obtained 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 and 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 without Gaussian part. 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:

$$\nu(A)\leq\int_{S^{d-1}}\int_0^{+\infty}\mathbf{1}_A(s\theta)\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\theta). \tag{1.8}$$

We assume one of the following:

[H-1a] μ has a density with respect to the Lebesgue measure on the sphere.

[H-1b] there exists $\gamma\in[1,d]$ such that $\mu\left(B(\theta,r)\cap S^{d-1}\right)\leq Cr^{\gamma-1}$, with $\gamma+\alpha>d$, and for all $s>0$, there exists $C>0$ such that:

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

[H-2] Denoting by φ_Z the Lévy-Kintchine exponent of $(Z_t)_{t\geq 0}$, there is $K>0$ such that :

$$\mathbb{E}\left(e^{i\langle p,Z_t\rangle}\right)=e^{t\varphi_Z(p)}\leq e^{-Kt|p|^\alpha},\ |p|>1. \tag{1.10}$$

[H-3] $F:\mathbb{R}^d\to\mathbb{R}^d$ is Lipschitz continuous or measurable and bounded and $\sigma:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ is bounded and η-Hölder continuous $\eta\in(0,1)$.

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

$$\kappa^{-1}|\xi|^2\leq\langle\xi,\sigma(x)\xi\rangle\leq\kappa|\xi|^2. \tag{1.11}$$

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

$$\nu(x,A)=\nu\big(\{z\in\mathbb{R}^d;\ \sigma(x)z\in A\}\big). \tag{1.12}$$

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

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

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 $x \in A_{low}$,

$$\nu\Big(B(x,r)\Big) \geq Cr^\gamma \frac{\underline{q}(|x|)}{|x|^{\alpha+\gamma}},\ \forall r > 0 \tag{1.13}$$

$$\nu\Big(B(0,r)^c\Big) \leq C\frac{1}{r^\alpha},\ \forall r \in (0,1). \tag{1.14}$$

In the rest of this Chapter, 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. Note that assumption **[H-2]** is crucial in order to get the existence of the density. 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 $L(x,\nabla_x)$ of the equation* (1.1)*:*

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

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:

$$\mathbb{P}(X_0 = x) = 1 \ \ \textit{and} \ \ f(t,X_t) - \int_0^t (\partial_u + L(x,\nabla_x)f(u,X_u)du \ \ \textit{is a } \mathbb{P}\textit{- martingale.}$$

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 $x, y \in \mathbb{R}^d$,*

$$\mathbb{P}(X_t \in dy | X_0 = x) = p(t,x,y)dy. \tag{1.15}$$

*Assume that the function Q defined below is decreasing, and fix a deterministic time horizon $T > 0$. There exists $C_1 \geq 1$ depending on T and the parameters in **[H]**, such that the following density estimates holds:*
$\forall 0 \leq t \leq T,\ \forall (x,y) \in \mathbb{R}^d$,

$$p(t,x,y) \leq C_1 \frac{t^{-d/\alpha}}{\left(1 + \frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|y-\theta_{t,0}(x)|), \tag{1.16}$$

where:

- *when the drift F is bounded, θ is the identity map: $\theta_{t,0}(x) = x$, and*
  - *under **[H-1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \bar{q}(s)$,*
  - *under **[H-1b]**, for all $s > 0$, $Q(s) = \min(1, s^{\gamma-1})\bar{q}(s)$,*
- *when the drift F is Lipschitz continuous, $\theta_{s,t}(x)$ denotes the solution to the ordinary differential equation:*

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

  and
  - *under **[H-1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \min(1,s)\bar{q}(s)$,*
  - *under **[H-1a]**, for all $s > 0$, $Q(s) = \min(1, s, s^{\gamma-1})\bar{q}(s)$.*

*Moreover, assume **[H-LB]** holds. Then, if*

$$\forall s \in [0,t/2],\ B\big(\sigma(\theta_{s,0}(x))^{-1}(\theta_{0,t}(y) - x), Ct^{1/\alpha}\big) \subset A_{low}, \tag{1.17}$$

there exists $C_2 > 1$ such that

$$C_2^{-1} \frac{t^{-d/\alpha}}{\left(1 + \frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \underline{q}(|y-\theta_{t,0}(x)|) \leq p(t,x,y). \tag{1.18}$$

Remark 1.1. The condition (1.17) 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 (1.17) 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. We point out that when **[H-1a]** holds, the condition

$$\mu\left(B(\theta, r) \cap S^{d-1}\right) \leq C r^{\gamma-1},$$

actually holds with $\gamma = d$. Besides, the function Q appearing in the upper bound (4.12) is decreasing typically when considering tempering function of the form:

$$\bar{q}(s) = \frac{1}{1+s^m},$$

where m is large enough.

The rest of this Chapter 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 (1.18).

Remark 1.3 (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 $Q(|x|)$ where we actually mean $Q(C|x|)$. 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 1.4 (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).

Remark 1.5 (Time-Homogeneous Coefficients)**.** In this work, we have restricted ourselves to time-homogeneous coefficients. It turns out that our techniques can be used even in the time-dependent case. We choose to stick to the time homogeneous case for notational simplicity.

2 The Parametrix Setting

We present here a continuity technique known as the Parametrix. 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:

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

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

$$\begin{aligned}\tilde{X}_s^{t,y} &= x + \int_0^s F(\theta_{u,t}(y))du + \int_0^s \sigma(\theta_{0,t}(y))dZ_u \\ &= x + \int_0^s F(\theta_{u,t}(y))du + \sigma(\theta_{0,t}(y))Z_s. \end{aligned} \tag{2.19}$$

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 the 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 $\tilde{X}_t = x + \sigma(y)Z_t$. Note that in that case, we do not need the existence of the flow $\theta_{s,t}$ associated with the ODE. Also, we could restrict ourselves to F Hölder continuous, as in this case, the existence of the flow θ is given by the Cauchy Peano theorem. However, the lack of uniqueness raises 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 $\tilde{X}^{t,y}$ and assumptions **[H-2]** (non degeneracy Fourier Transform) and **[H-4]** (ellipticity of σ) that $\tilde{X}^{t,y}$ has a density with respect to the Lebesgue measure. We denote the latter:

$$\tilde{p}^{t,y}(s,x,z)dz = \mathbb{P}(\tilde{X}_s^{t,y} \in dz | \tilde{X}_0 = x),\ s \in (0,t].$$

Recall that we have denoted by $p_Z(t,x)$ the density of the driving process Z. The frozen density relates to the density of Z through the relation:

$$\tilde{p}^{t,y}(t,x,z) = \det \sigma(\theta_{0,t}(y))^{-1} p_Z\left(t, \sigma(\theta_{0,t}(y))^{-1}\left(z - \int_0^t F(\theta_{u,t}(y))du - x\right)\right). \tag{2.20}$$

We will often denote $\tilde{p}(t,x,y) = \tilde{p}^{t,y}(t,x,y)$, 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

$$y - \int_0^t F(\theta_{u,t}(y))du - x = \theta_{0,t}(y) - x.$$

Moreover, for $\xi \in \mathbb{R}^d$, we define the integro-differential operator $\forall \varphi \in \mathcal{C}_b^2(\mathbb{R}^d, \mathbb{R})$:

$$\begin{aligned} L(\xi, \nabla_x)\varphi(x) &= \langle F(\xi), \nabla_x \varphi(x)\rangle \\ &+ \int_{\mathbb{R}^d} \varphi(x+\sigma(\xi)z) - \varphi(x) - \langle \nabla_x \varphi(x), \sigma(\xi)z\rangle \mathbf{1}_{\{|z|\le 1\}} \nu(dz). \end{aligned} \tag{2.21}$$

Observe that when $\xi = x$ the initial position, the operator $L(x, \nabla_x)$ is the generator of (1.1). Also, for a given $(t, y) \in (0, T] \times \mathbb{R}^d$, the operator $L(\theta_{0,t}(y), \nabla_x)$ is the generator of $(\tilde{X}_s^{t,y})_{s\in[0,t]}$ at time $s = 0$. 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 $(X_s^{t,x})_{0\le t\le s}$ to* (1.1) *which has a Feller semigroup $(P_t)_{t\ge 0}$. We have the following formal representation. For all $t > 0$, $(x, y) \in (\mathbb{R}^d)^2$ and any bounded measurable $f : \mathbb{R}^d \to \mathbb{R}$:*

$$P_t f(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int_{\mathbb{R}^d} \left(\sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t, x, y) \right) f(y) dy, \tag{2.22}$$

where H is the parametrix kernel:

$$\forall t \ge 0,\ (x, y) \in (\mathbb{R}^d)^2,\ H(t, x, y) := (L(x, \nabla_x) - L(\theta_{0,t}(y), \nabla_x))\tilde{p}^{t,y}(t, x, y). \tag{2.23}$$

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

$$f \otimes g(t, x, y) = \int_0^t du \int_{\mathbb{R}^d} dz f(u, x, z) g(t-u, z, y).$$

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

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 $\mathbb{P}[X_t \in dy|X_0 = x] = p(t, x, y)dy$ where :

$$\forall t > 0,\ (x, y) \in (\mathbb{R}^d)^2,\ p(t, x, y) = \sum_{r=0}^{+\infty} (\tilde{p} \otimes H^{(r)})(t, x, y). \tag{2.24}$$

Proof. We refer to Huang and Menozzi [HM14] for the proof of this statement. See as well Proposition 3.6 in Chapter 3. □

The proof relies on the Markov properties of the processes involved, as well as the Chapman-Kolmogorov equations. In the Brownian setting, the series (2.24) 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 (2.24) as the density of the initial equation (1.1). However, as we mention in Chapter 1, in the Lévy setting, there are no general (Hörmander) theorem to ensure the existence of the density even with regular coefficients. Nevertheless, in addition to the already mentioned references, we can refer to Ishikawa and Kunita [IK06] in the non degenerate case, and Cass [Cas09], which can be seen as the most complete extension to the jump case of the Hörmander theorem, but requires some integrability conditions.

Also the works of Zhang [Zha14a, Zha14b] 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 (2.24) 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

$$\bar{p}(t,x,y) = \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y)-x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|\theta_{0,t}(y) - x|), \tag{2.25}$$

where

- when the drift F is bounded, θ is the identity map: $\theta_{t,0}(x) = x$, and
  - under **[H-1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \bar{q}(s)$,
  - under **[H-1b]**, for all $s > 0$, $Q(s) = \min(1, s^{\gamma-1})\bar{q}(s)$,
- when the drift F is Lipschitz continuous, $\theta_{s,t}(x)$ denotes the solution to the ordinary differential equation:

  $$\frac{d}{ds}\theta_{s,t}(x) = F(\theta_{s,t}(x)),\ \theta_{t,t}(x) = x,\ \forall 0 \le t, s \le T,$$

  and
  - under **[H-1a]**, $\gamma = d$ and for all $s > 0$, $Q(s) = \min(1, s)\bar{q}(s)$,
  - under **[H-1a]**, for all $s > 0$, $Q(s) = \min(1, s, s^{\gamma-1})\bar{q}(s)$.

We also assume that the function Q is decreasing. Note that $\bar{p}$ is the upper bound on the Frozen density under **[H]** derived by Sztonyk [Szt10], up to the degradation of the tempering function $\bar{q}(s)$. We refer to Section 3 for more details on this estimate.

The following lemma proves the convergence of the series (2.24).

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

$$|\tilde{p} \otimes H(t,x,y)| \leq C_{2.2}\Big(t^\omega \bar{p}(t,x,y) + \rho(t,x,y)\Big), \tag{2.26}$$

$$|\rho \otimes H(t,x,y)| \leq C_{2.2} t^\omega \bar{p}(t,x,y), \tag{2.27}$$

where we denoted $\rho(t,x,y) = \delta \wedge |x - \theta_{0,t}(y)|^{\eta(\alpha \wedge 1)} \bar{p}(t,x,y)$. Now for all $k \geq 1$,

$$|\tilde{p} \otimes H^{(2k)}(t,x,y)| \leq (4C_{2.2})^{2k} t^{k\omega}\Big(t^{k\omega}\bar{p}(t,x,y) + (\bar{p}+\rho)(t,x,y)\Big), \tag{2.28}$$

$$|\tilde{p} \otimes H^{(2k+1)}(t,x,y)| \leq (4C_{2.2})^{2k+1} t^{k\omega}\Big(t^{(k+1)\omega}\bar{p} + t^\omega(\bar{p}+\rho) + \rho\Big)(t,x,y). \tag{2.29}$$

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 (1.16) for the sum of the parametrix series (2.24) in small time. To extend the result to any arbitrary (but finite) time, we use the semi-group property satisfied by $\bar{p}(t,x,y)$ (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 (2.24) for any time from Lemma 3.10, 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 [Kol00b].

Proof. We prove the important estimates (2.26) and (2.27) 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 (2.26) and (2.27), we prove estimates (2.28) and (2.29) 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 (2.24).

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

$$|\tilde{p} \otimes H(t,x,y)| \leq C_{3.10}\Big(t^\omega \bar{p} + \rho + t^\omega(\bar{p}+\rho)\Big)(t,x,y).$$

Now, using equations (2.26) and (2.27), we have:

$$\begin{aligned}
|\tilde{p} \otimes H^{(2)}(t,x,y)| &\leq C_{3.10}\Big(t^\omega |\bar{p} \otimes H| + |\rho \otimes H|\Big)(t,x,y) \\
&\leq C_{3.10}\Big(C_{3.10} t^{2\omega}\bar{p} + C_{3.10} t^\omega \rho + C_{3.10} t^\omega \bar{p}\Big)(t,x,y) \\
&\leq (2C_{3.10})^2 t^\omega \Big(t^\omega \bar{p} + (\bar{p}+\rho)\Big)(t,x,y).
\end{aligned}$$

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

$$\begin{aligned}|\tilde{p}\otimes H^{(2k+1)}|(t,x,y) &\leq (4C_{3.10})^{2k}t^{k\omega}\Big(t^{k\omega}|\bar{p}\otimes H|(t,x,y)+|(\bar{p}+\rho)\otimes H|(t,x,y)\Big)\\ &\leq (4C_{3.10})^{2k}t^{k\omega}\Big(C_{3.10}t^{k\omega}(t^{\omega}\bar{p}+\rho)(t,x,y)\\ &\quad +C_{3.10}(t^{\omega}\bar{p}+\rho)(t,x,y)+C_{3.10}t^{\omega}\bar{p}(t,x,y)\Big).\end{aligned}$$

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

$$\begin{aligned}|\tilde{p}\otimes H^{(2k+1)}|(t,x,y) &\leq (4C_{3.10})^{2k}t^{k\omega}\Big(C_{3.10}t^{(k+1)\omega}\bar{p}+2C_{3.10}t^{\omega}(\bar{p}+\rho)+C_{3.10}\rho)\Big)(t,x,y)\\ &\leq (4C_{3.10})^{2k}(2C_{3.10})t^{k\omega}\Big(t^{(k+1)\omega}\bar{p}+t^{\omega}(\bar{p}+\rho)+\rho)\Big)(t,x,y),\end{aligned}$$

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

$$\begin{aligned}|\tilde{p}\otimes H^{(2k+2)}(t,x,y)| &\leq (4C_{3.10})^{2k+1}t^{k\omega}\Big(t^{(k+1)\omega}|\bar{p}\otimes H|\\ &\quad +t^{\omega}|(\bar{p}+\rho)\otimes H|+|\rho\otimes H|\Big)(t,x,y)\\ &\leq (4C_{3.10})^{2k+1}t^{k\omega}\Big(C_{3.10}t^{(k+1)\omega}[t^{\omega}\bar{p}+\rho]\\ &\quad +C_{3.10}t^{\omega}[\{t^{\omega}\bar{p}+\rho\}+C_{3.10}t^{\omega}\bar{p}]+C_{3.10}t^{\omega}\bar{p}\Big)(t,x,y)\\ &\leq (4C_{3.10})^{2k+2}t^{(k+1)\omega}\Big(t^{(k+1)\omega}\bar{p}+(\bar{p}+\rho)\Big)(t,x,y),\end{aligned}$$

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

□

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 (2.24). 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 (2.24). In the following, we take $\gamma=d$ if **[H-1a]** holds.

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 $t \in [0, T]$, $(x, y) \in (\mathbb{R}^d)^2$:*

$$\tilde{p}^{t,y}(t, x, z) \le C \frac{t^{-d/\alpha}}{\left(1 + \frac{|z-x-\int_0^t F(\theta_{u,t}(y))du|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}\left(C^{-1}\left|z - x - \int_0^t F(\theta_{u,t}(y))du\right|\right). \tag{3.30}$$

Moreover, when ***[H-LB]*** *holds, for all $z - x - \int_0^t F(\theta_{u,t}(y))du \in A_{low}$, the lower bound holds:*

$$\frac{C^{-1}t^{-d/\alpha}}{\left(1 + \frac{|z-x-\int_0^t F(\theta_{u,t}(y))du|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \underline{q}\left(C\left|z - x - \int_0^t F(\theta_{u,t}(y))du\right|\right) \le \tilde{p}^{t,y}(t, x, z). \tag{3.31}$$

Proof. Recall that $\tilde{X}_t^{t,y} = x + \int_0^t F(\theta_{u,t}(y))du + \sigma(\theta_{0,t}(y))Z_t$. From assumption **[H]**, we know that Z_t has a density $p_Z(t, \cdot)$. Also, the density of $(\tilde{X}_t)_{t\ge0}$ can be expressed in terms of $p_Z(t, \cdot)$ (see equation (2.20)). Now, under **[H]**, the conclusion of Theorems 1 and 2 in Sztonyk [Szt10] holds, namely:

$$p_Z(t, u) \le C \frac{t^{-d/\alpha}}{\left(1 + \frac{|u|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}(C^{-1}|u|),$$

and when **[H-LB]** holds, for all $u \in A_{low}$,

$$C^{-1} \frac{t^{-d/\alpha}}{\left(1 + \frac{|u|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \underline{q}(C|u|) \le p_Z(t, u).$$

Thus, replacing $u = \sigma(\theta_{0,t}(y))^{-1}(z - x - \int_0^t F(\theta_{u,t}(y))du)$, and recalling that σ is elliptic and bounded yields the announced result. □

Remark 3.1. Observe that the upper bound on the frozen density is exactly the one given by Sztonyk in [Szt10]. However, in Theorem 1.2, the upper bound is different. This is due to the fact that in our approach, the Parametrix Kernel presents a different concentration index. In order to correct this concentration, we have to deteriorate the tempering function, replacing $\bar{q}$ by Q. See also Proposition 3.3. Anyhow, since we have in general $\bar{q}(s) \le Q(s)$, we trivially have:

$$\tilde{p}^{t,y}(t, x, z) \le C \frac{t^{-d/\alpha}}{\left(1 + \frac{|z-x-\int_0^t F(\theta_{u,t}(y))du|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q\left(C^{-1}\left|z - x - \int_0^t F(\theta_{u,t}(y))du\right|\right).$$

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 follows 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 $f : \mathbb{R}^d \to \mathbb{R}, x \in \mathbb{R}^d$ and a given $t_0 \in [0,T]$:*

$$\left| \int_{\mathbb{R}^d} f(y)\tilde{p}^{t_0+t,y}(t,x,y)dy - f(x) \right| \underset{t\downarrow 0}{\longrightarrow} 0, \tag{3.32}$$

that is, for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, $\tilde{p}^{t,y}(t,x,y)dy \Rightarrow \delta_x(dy)$ weakly when $t \to 0$.

Proof. For notational convenience, we take $t_0 = 0$. The general case can be proved similarly. Let us write:

$$\begin{aligned} \int_{\mathbb{R}^d} f(y)\tilde{p}^{t,y}(t,x,y)dy - f(x) &= \int_{\mathbb{R}^d} f(y)\Big(\tilde{p}^{t,y}(t,x,y) - \tilde{p}^{t,\theta_{t,0}(x)}(t,x,y)\Big)dy \\ &\quad + \int_{\mathbb{R}^d} f(y)\Big(\tilde{p}^{t,\theta_{t,0}(x)}(t,x,y)\Big)dy - f(x). \end{aligned}$$

From the usual Dirac convergence in the Kolmogorov equations (2.6) in Chapter 1, the second term tends to zero when $t \to 0$. Let us discuss the first term. Define:

$$I = \int_{\mathbb{R}^d} f(y)\Big(\tilde{p}^{t,y}(t,x,y) - \tilde{p}^{t,\theta_{t,0}(x)}(t,x,y)\Big)dy. \tag{3.33}$$

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:

$$D_1 = \left\{ y \in \mathbb{R}^d; \frac{|\theta_{0,t}(y) - x|}{t^{1/\alpha}} \le Kt^{-\beta} \right\},\ D_2 = \left\{ y \in \mathbb{R}^d; \frac{|\theta_{0,t}(y) - x|}{t^{1/\alpha}} > Kt^{-\beta} \right\}.$$

We have from Proposition 3.1,

$$\tilde{p}^{t,y}(t,x,y) \le \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y)-x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}(C^{-1}|\theta_{0,t}(y) - x|),$$

and

$$\tilde{p}^{t,\theta_{t,0}(x)}(t,x,y) \le \frac{t^{-d/\alpha}}{\left(1 + \frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}(C^{-1}|y - \theta_{t,0}(x)|).$$

Observe that we used the fact that:

$$y-x-\int_0^t F(\theta_{u,t}(\theta_{t,0}(x)))du = y-x-\int_0^t F(\theta_{u,0}(x))du = y-\theta_{t,0}(x).$$

Now, from the Lipschitz property of the flow, we have $|\theta_{0,t}(y)-x| \asymp |y-\theta_{t,0}(x)|$. Consequently, the same upper bound for the two densities in (3.33).

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 $i\in\{1,2\}$, $I_{D_i} := \int_{D_i} f(y)\Big(\tilde p^{t,y}(t,x,y) - \tilde p^{t,\theta_{t,0}(x)}(t,x,y)\Big)dy$. We derive:

$$\begin{aligned}
|I_{D_2}| &\le C|f|_\infty \int_{D_2} \frac{t^{-d/\alpha}}{\Big(1+\frac{|y-\theta_{t,0}(x)|}{t^{1/\alpha}}\Big)^{\alpha+\gamma}} \bar q(C^{-1}|y-\theta_{t,0}(x)|)dy\\
&\le C|f|_\infty \int_{Kt^{-\beta}}^{+\infty} \frac{r^{d-1}}{1+r^{\alpha+\gamma}} \bar q(rt^{1/\alpha})dr\\
&\le Ct^{\beta(\gamma+\alpha-d)}.
\end{aligned}$$

To get to the second inequality, we changed variables to $z=\frac{y-\theta_{t,0}(x)}{t^{1/\alpha}}$, then pass to the polar coordinates $z=r\varsigma$, for $r\in\mathbb{R}_+$ and $\varsigma\in S^{d-1}$. Thus, since $\gamma+\alpha>d$, for $\beta>0$, $I_{D_2} \underset{t\downarrow 0}{\longrightarrow} 0$. On D_1, we will start from the inverse Fourier representation of $\tilde p^{t,z}(t,x,y)$, $z=\theta_{t,0}(x), y$. Recall we denoted 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:

$$\tilde p^{t,z}(t,x,y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dp e^{-i\langle p, y-\int_0^t F(\theta_{u,t}(z))du - x\rangle} e^{t\varphi_Z(\sigma(z)^*p)}.$$

Consequently, we have:

$$\begin{aligned}
&\tilde p^{t,y}(t,x,y) - \tilde p^{t,\theta_{t,0}(x)}(t,x,y)\\
=\ & \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle p, y-\int_0^t F(\theta_{u,t}(y))du - x\rangle} e^{t\varphi_Z(\sigma(\theta_{0,t}(y))^*p)}\\
& - e^{-i\langle p, y-\int_0^t F(\theta_{t,0}(x))du - x\rangle} e^{t\varphi_Z(\sigma(\theta_{t,0}(x))^*p)} dp\\
=\ & \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \Big(e^{-i\langle p, y-\int_0^t F(\theta_{u,t}(y))du - x\rangle} - e^{-i\langle p, y-\int_0^t F(\theta_{t,0}(x))du - x\rangle}\Big) e^{t\varphi_Z(\sigma(\theta_{0,t}(y))^*p)} dp\\
& + \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle p, y-\int_0^t F(\theta_{t,0}(x))du - x\rangle}\Big(e^{t\varphi_Z(\sigma(\theta_{0,t}(y))^*p)} - e^{t\varphi_Z(\sigma(\theta_{t,0}(x))^*p)}\Big) dp\\
=\ & \Gamma_1(t,x,y) + \Gamma_2(t,x,y).
\end{aligned}$$

Thus, we have:

$$\int_{D_1} f(y)\Big(\tilde{p}^{t,y}(t,x,y)-\tilde{p}^{t,\theta_{t,0}(x)}(t,x,y)\Big)dy = \int_{D_1} f(y)\Gamma_1(t,x,y)dy+\int_{D_1} f(y)\Gamma_2(t,x,y)dy.$$

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

$$\begin{aligned}&\Gamma_1(t,x,y)\\ =\ & \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\int_0^1 d\lambda i\langle p,(I-\theta_{t,0})(\theta_{0,t}(y)-x)\rangle e^{-i\langle p,[\lambda I+(1-\lambda)\theta_{t,0}](\theta_{0,t}(y)-x)\rangle}e^{t\varphi_Z(\sigma(y)^*p)}dp,\end{aligned}$$

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 $t \le T$, $z \in \mathbb{R}^d$, $|(I-\theta_{t,0})(z)| \le Ct(1+|z|)$. Thus, since $y \in D_1$, we have for $\beta \le 1/\alpha$,

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

Integrating on D_1, we obtain:

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

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

$$\begin{aligned}\Gamma_2(t,x,y) \ &=\ \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dp e^{-i\langle p,y-\theta_{t,0}(x)\rangle}\int_0^1 d\lambda e^{\lambda t\varphi_Z(\sigma(\theta_{0,t}(y))^*p)+(1-\lambda)t\varphi_Z(\sigma(\theta_{t,0}(x))^*p)}\\ &\quad \times t(\varphi_Z(\sigma(\theta_{0,t}(y))^*p)-\varphi_Z(\sigma(\theta_{t,0}(x))^*p)).\end{aligned}$$

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

$$e^{\lambda t\varphi_Z(\sigma(\theta_{0,t}(y))^*p)+(1-\lambda)t\varphi_Z(\sigma(\theta_{t,0}(x))^*p)} \le e^{-Kt|p|^\alpha}.$$

On the other hand, using the bound on the Lévy-Khintchine exponent and the Hölder continuity of σ, we can rewrite the increment:

$$\begin{aligned}&t|\varphi_Z(\sigma(\theta_{0,t}(y))^*p)-\varphi_Z(\sigma(\theta_{t,0}(x))^*p)|\\ =\ & t\left|\int_{\mathbb{R}^d}\cos(\langle\sigma(\theta_{0,t}(y))^*p,\xi\rangle)-\cos(\langle\sigma(\theta_{t,0}(x))^*p,\xi\rangle)\nu(dz)\right|\\ \le\ & Ct|p|^\alpha|x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}.\end{aligned}$$

To summarize, we obtained:

$$\begin{aligned}\int_{D_1} f(y)\Gamma_2(t,x,y) &\leq |f|_\infty \int_{D_1} dy\,|\Gamma_2(t,x,y)| \\ &\leq C\int_{D_1} dy \int_{\mathbb{R}^d} t|p|^\alpha |x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)} e^{-Kt|p|^\alpha} dp.\end{aligned}$$

Changing variables, and integrating over p yields

$$\begin{aligned}\int_{D_1} f(y)\Gamma_2(t,x,y) &\leq \frac{C}{t^{d/\alpha}}|f|_\infty \int_{D_1} dy|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)} \\ &= \frac{C}{t^{d/\alpha}}|f|_\infty \int_0^{t^{-\beta}} drr^{\eta(\alpha\wedge 1)+d-1} t^{d/\alpha+\eta(1\wedge 1/\alpha)}.\end{aligned}$$

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

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

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

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

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

$$|H(t,x,y)| \leq C\left(t^{-1/\alpha}\mathbf{1}_{\{\alpha>1\}} + \frac{\delta\wedge|x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}}{t}\right)\bar{p}(t,x,y),$$

where we recall that

$$\bar{p}(t,x,y) = \frac{t^{-d/\alpha}}{\left(1+\frac{|\theta_{0,t}(y)-x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|\theta_{0,t}(y)-x|), \tag{3.34}$$

where

- *when the drift F is bounded, θ is the identity map: $\theta_{t,0}(x)=x$, and*
  - *under **[H-1a]**, $\gamma=d$ and for all $s>0$, $Q(s)=\bar{q}(s)$,*
  - *under **[H-1b]**, for all $s>0$, $Q(s)=\min(1,s^{\gamma-1})\bar{q}(s)$,*

- *when the drift F is Lipschitz continuous, $\theta_{s,t}(x)$ denotes the solution to the ordinary differential equation:*

$$\frac{d}{ds}\theta_{s,t}(x) = F(\theta_{s,t}(x)),\ \theta_{t,t}(x) = x,\ \forall 0 \le t, s \le T,$$

and

  - *under* ***[H-1a]***, *$\gamma = d$ and for all $s > 0$, $Q(s) = \min(1, s)\bar{q}(s)$,*
  - *under* ***[H-1a]***, *for all $s > 0$, $Q(s) = \min(1, s, s^{\gamma-1})\bar{q}(s)$.*

Thus, the upper bound on the Kernel H is the same as the upper bound on the Frozen density $\tilde{p}^{t,y}(t, x, y)$ up to the additional multiplier $\big(\delta \wedge |x - \theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}\big)t^{-1}$, that can be seen as the singularity induced by the difference $L(x, \nabla_x) - L(\theta_{0,t}(y), \nabla_x)$ applied to the frozen density. The proof is 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. From the definition of the generators, the operator naturally splits into three parts. Let φ be a test function,

$$\begin{aligned}\Big(L(x, \nabla_x) - L(\theta_{0,t}(y), \nabla_x)\Big)\varphi(x) &= \langle \nabla\varphi(x), F(x) - F(\theta_{0,t}(y))\rangle \\ + \int_{\mathbb{R}^d} \Big(\varphi(x+z) - \varphi(x) - \langle \nabla\varphi(x), z\rangle\Big)&\mathbf{1}_{\{|z|\le t^{1/\alpha}\}}(\nu(x, dz) - \nu(\theta_{0,t}(y), dz)) \\ + \int_{\mathbb{R}^d} \Big(\varphi(x+z) - \varphi(x)\Big)&\mathbf{1}_{\{|z|\ge t^{1/\alpha}\}}(\nu(x, dz) - \nu(\theta_{0,t}(y), dz)).\end{aligned}$$

Recall that we defined $\nu(\xi, A) = \nu\{z \in \mathbb{R}^d; \sigma(\xi)z \in A\}$. 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 proceed as Sztonyk in [Szt10]. We split at the characteristic time scale the process $Z_t = M_t + N_t$ with its Lévy-Itô decomposition, where M is a martingale and N is a poisson process. Specifically,

$$\mathbb{E}(e^{i\langle p, M_t\rangle}) = \exp\left(t\int_{\mathbb{R}^d}(e^{i\langle p,\eta\rangle} - 1 - i\langle p, \eta\rangle)\mathbf{1}_{\{|z|\le t^{1/\alpha}\}}\nu(d\eta)\right),$$

and since this Fourier transform is integrable and regular (see Sztonyk [Szt10] and the references therein), we can say that this term produces the density in the Lévy-Itô decomposition. Also, we have the following decomposition for the law of the Poisson Process N_t:

$$P_{N_t}(dz) = e^{-t\bar{\nu}(\mathbb{R}^d)}\sum_{k=0}^{+\infty}\frac{t^k\bar{\nu}^{*k}(dz)}{k!},\ \bar{\nu}(dz) = \mathbf{1}_{\{|z|\ge t^{1/\alpha}\}}\nu(dz),$$

that will give the heavy-tailed behavior when we make the convolution between the laws of M_t and N_t. Specifically, denoting by p_M the density of the martingale and P_{N_t} the law of N_t, the density of Z can be written as:

$$p_Z(t,u) = \int_{\mathbb{R}^d} p_M(t, u-\xi) P_{N_t}(d\xi).$$

Recall that $\tilde{p}(t,x,y) = \det(\sigma(\theta_{0,t}(y)))^{-1} p_Z(t, \sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y) - x))$. Thus, a derivative along x of $\tilde{p}(t,x,y)$ acts in fact on the density of the martingale, and we have to control $\nabla_x p_M(t, \sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y) - x) - \xi)$. Borrowing the notations of the proof of Lemma 2 in [Szt10], we have:

$$p_M(t,x) = t^{-d/\alpha} g_t(t^{-1/\alpha} x).$$

Formally, since we chose to split at the characteristic time-scale $t^{1/\alpha}$, the density of the martingale presents a time space separation, and defining g_t as above allows to have estimates independent of t. Thus, uniformly for all $t > 0$, $g_t(y)$ is in Schwartz's class. Therefore, we have for all $m \geq 1$:

$$|\nabla_x p_M(t,x)| \leq \frac{1}{t^{1/\alpha}} C_m t^{-d/\alpha} \left(1 + \frac{|x|}{t^{1/\alpha}}\right)^{-m},$$

and we recovered Lemma 2 in [Szt10]. Thus, a derivative on the density yields a singularity in $t^{-1/\alpha}$ which is integrable when $\alpha > 1$. See also Lemma 8.1 in Chapter 3.

Specifically, when the drift F is bounded and $\alpha > 1$, we write:

$$\begin{aligned} |\langle \nabla_x p(t,x,y), F(x) - F(\theta_{0,t}(y))\rangle| &\leq C2|F|_\infty |\nabla_x p(t,x,y)| \\ &\leq Ct^{-1/\alpha} \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y) - x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} \bar{q}(|\theta_{0,t}(y) - x|). \end{aligned}$$

On the other hand, when F is unbounded, we have to deteriorate the tempering function:

$$\begin{aligned} |\langle \nabla_x p(t,x,y), F(x) - F(\theta_{0,t}(y))\rangle| &\leq C|x - \theta_{0,t}(y)| |\nabla_x p(t,x,y)| \\ &\leq Ct^{-1/\alpha} \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y) - x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} |x - \theta_{0,t}(y)| \bar{q}(|\theta_{0,t}(y) - x|) \\ &\leq Ct^{-1/\alpha} \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y) - x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|\theta_{0,t}(y) - x|) \end{aligned}$$

Consider now the integro-differential part of the kernel. 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. We use the representation in terms of symbols, denoting by $\phi_t(x,p)$ the symbol of an integro-differential operator $\Phi_t(x,\nabla_x)$:

$$\begin{aligned}
&\Phi_t(x,\nabla_x)p_M\Big(t,\sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y)-x)-\xi\Big)\\
=\ &\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle p,\sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y)-x)-\xi\rangle}\phi_t(x,-(\sigma(\theta_{0,t}(y))^{-1})^*pt^{-1/\alpha})\hat{g}_t(p)dp.
\end{aligned}$$

Now, when $\Phi_t(x,\nabla_x)=L^M(x,\nabla_x)-L^M(\theta_{0,t}(y),\nabla_x)$, the small jump part of the difference of the generators, that is:

$$\begin{aligned}
&\Big(L^M(x,\nabla_x)-L^M(\theta_{0,t}(y),\nabla_x)\Big)\varphi(x)\\
=\ &\int_{\mathbb{R}^d}\Big(\varphi(x+z)-\varphi(x)-\langle\nabla\varphi(x),z\rangle\Big)\mathbf{1}_{\{|z|\le t^{1/\alpha}\}}(\nu(x,dz)-\nu(\theta_{0,t}(y),dz)),
\end{aligned}$$

denoting by $l^M(x,p)-l^M(\theta_{0,t}(y),p)$ the corresponding symbol, we have that:

$$|l^M(x,p)-l^M(\theta_{0,t}(y),p)|\le C\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}|p|^\alpha.$$

Moreover, this quantity $\phi_t(x,-(\sigma(\theta_{0,t}(y))^{-1})^*pt^{-1/\alpha})\hat{g}_t(p)$ is smooth (in its p argument) because of the truncation (see Sztonyk [Szt10] and the references therein). Consequently,

$$\frac{t}{\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}}\phi_t(x,-(\sigma(\theta_{0,t}(y))^{-1})^*pt^{-1/\alpha})\hat{g}_t(p)$$

is infinitely differentiable as a function of p and uniformly bounded with all its derivatives. Therefore, it is in Schwartz's space as well as its Fourier inverse. We have $\forall m>1$:

$$\begin{aligned}
&\Big|(L^M(x,\nabla_x)-L^M(\theta_{0,t}(y),\nabla_x))p_M\Big(t,\sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y)-x)-\xi\Big)\Big|\\
&\le C\frac{\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}}{t}t^{-d/\alpha}\left(1+\frac{|\sigma(\theta_{0,t}(y))^{-1}(\theta_{0,t}(y)-x)-\xi|}{t^{1/\alpha}}\right)^{-m}.
\end{aligned}$$

Consequently, we recovered Lemma 2 in [Szt10] for the Parametrix kernel, up to the additional multiplicative term $\big(\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}\big)t^{-1}$, which is the expected singularity for the Kernel (see Kolokoltsov [Kol00b]). The upper bound follows from this upper bound and the control of the measure of the balls for P_{N_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]. See also Section 8 in Chapter 3. The upper bound for the small jumps part of the kernel follows.

Finally, for large jumps, we see that the measure $\mathbf{1}_{\{|\xi|\geq t^{1/\alpha}\}}(\nu(x,d\xi)-\nu(\theta_{0,t}(y),d\xi))$ is no more singular. Thus, we can write:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d}\Big(\tilde{p}(t,x+\xi,y)-\tilde{p}(t,x,y)\Big)\mathbf{1}_{\{|\xi|\geq t^{1/\alpha}\}}(\nu(x,d\xi)-\nu(\theta_{0,t}(y),d\xi))\right|\\
&\leq \int_{\mathbb{R}^d}\Big|\tilde{p}(t,x+\xi,y)-\tilde{p}(t,x,y)\Big|\mathbf{1}_{\{|\xi|\geq t^{1/\alpha}\}}|\nu(x,d\xi)-\nu(\theta_{0,t}(y),d\xi)|\\
&\leq \delta\wedge|x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}\bigg(\int_{S^{d-1}}\int_0^{+\infty}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\varsigma)\\
&\qquad+\frac{1}{t}\bar{p}(t,x,y)\bigg).
\end{aligned}$$

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 $\tilde{p}(t,x+s\varsigma,y)\leq Ct^{-d/\alpha}\leq C\bar{p}(t,x,y)$. The integral then yields the singularity t^{-1}. Therefore, we assume that $|\theta_{0,t}(y)-x|\geq t^{1/\alpha}$. The regime of $\tilde{p}(t,x+s\varsigma,y)$ is given by $|\theta_{0,t}(y)-x-s\varsigma|$. Thus, thanks to the triangle inequality, when $|\theta_{0,t}(y)-x|\leq 1/2s$, or when $s\leq 1/2|\theta_{0,t}(y)-x|$, the density $\tilde{p}(t,x+s\varsigma,y)$ is off-diagonal with $\tilde{p}(t,x+s\varsigma,y)\leq C\bar{p}(t,x,y)$.

Consequently, the problematic case is when $s\asymp|\theta_{0,t}(y)-x|$. Indeed, in this case, $\tilde{p}(t,x+s\varsigma,y)$ can be in diagonal regime, whereas $\tilde{p}(t,x,y)$ is still in the off-diagonal regime.

Assume first that **[H-1-a]** holds, and let us simply denote $\frac{d\mu}{d\varsigma}(\varsigma)$ the density of μ on the sphere. Then, we have:

$$\begin{aligned}
&\int_{1/2|\theta_{0,t}(y)-x|}^{3/2|\theta_{0,t}(y)-x|}\int_{S^{d-1}}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\frac{d\mu}{d\varsigma}(\varsigma)d\varsigma\\
=&\int_{1/2|\theta_{0,t}(y)-x|}^{3/2|\theta_{0,t}(y)-x|}\int_{S^{d-1}}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{\alpha+d}}\frac{d\mu}{d\varsigma}(\varsigma)s^{d-1}dsd\varsigma.
\end{aligned}$$

Now, since $s\asymp|\theta_{0,t}(y)-x|$, we can take $\frac{\bar{q}(s)}{s^{\alpha+d}}$ out of the integral. Also, the density $\frac{d\mu}{d\varsigma}(\varsigma)$ is bounded, so that we obtain:

$$\begin{aligned}
&\int_{1/2|\theta_{0,t}(y)-x|}^{3/2|\theta_{0,t}(y)-x|}\int_{S^{d-1}}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\frac{d\mu}{d\varsigma}(\varsigma)d\varsigma\\
\leq\ & C\frac{\bar{q}(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{\alpha+d}}\int_0^{+\infty}\int_{S^{d-1}}\tilde{p}^y(t,x+s\varsigma,z)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}s^{d-1}dsd\varsigma.
\end{aligned}$$

Finally, the remaining integral can be bounded by some constant as the integral of the density . Consequently, we obtained:

$$\begin{aligned}&\int_{1/2|\theta_{0,t}(y)-x|}^{3/2|\theta_{0,t}(y)-x|}\int_{S^{d-1}}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\frac{d\mu}{d\varsigma}(\varsigma)d\varsigma\\ &\leq C\frac{\bar{q}(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{\alpha+d}}=C\frac{1}{t}\frac{t}{|\theta_{0,t}(y)-x|^{\alpha+d}}\bar{q}(|\theta_{0,t}(y)-x|),\end{aligned}$$

which is the off diagonal estimate for $\bar{p}$ when **[H-1a]** holds, up to the singularity $1/t$.

Now, assume that **[H-1-b]** holds. In this case, we can take out $\frac{\bar{q}(s)}{s^{1+\alpha}}$ and integrate a density to get:

$$\begin{aligned}&\int_{1/2|\theta_{0,t}(y)-x|}^{3/2|\theta_{0,t}(y)-x|}\int_{S^{d-1}}\tilde{p}(t,x+s\varsigma,y)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}\frac{\bar{q}(s)}{s^{1+\alpha}}ds\mu(d\varsigma)\\ \leq\ &\frac{\bar{q}(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{1+\alpha}}\int_0^{+\infty}\int_{S^{d-1}}\tilde{p}^y(t,x+s\varsigma,z)\mathbf{1}_{\{s\geq t^{1/\alpha}\}}ds\mu(d\varsigma)\\ &\leq C\frac{\bar{q}(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{1+\alpha}}.\end{aligned}$$

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

$$\begin{aligned}\frac{\bar{q}(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{1+\alpha}}&=\frac{1}{t}\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{0,t}(y)-x|)\times t^{\frac{d-\gamma}{\alpha}}\\ &\leq C\frac{1}{t}\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{0,t}(y)-x|).\end{aligned}$$

In the last inequality, we recall that $\gamma\leq d$, so that $t^{\frac{d-\gamma}{\alpha}}\leq 1$. Now, we write:

$$\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{0,t}(y)-x|)=\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}}\times|\theta_{0,t}(y)-x|^{\gamma-1}\bar{q}(|\theta_{0,t}(y)-x|).$$

Recalling we denoted by Q:

$$Q(|\theta_{0,t}(y)-x|)=\max(1,|\theta_{0,t}(y)-x|,|\theta_{0,t}(y)-x|^{\gamma-1})\bar{q}(|\theta_{0,t}(y)-x|),$$

we finally obtain:

$$C\frac{1}{t}\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{1+\alpha}}\bar{q}(|\theta_{0,t}(y)-x|)\leq C\frac{1}{t}\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}}Q(|\theta_{0,t}(y)-x|)$$

In other words, we can correct the wrong decay by deteriorating the temperation. Consequently, the global upper bound for the kernel is the one announced.

To sum up, we deteriorated the tempering function in the following cases:

- when the drift is bounded and **[H-1b]** holds. In this case, we replaced $\bar{q}$ by $Q(s) = \max(1, s^{\gamma-1})\bar{q}(s)$.
- when the drift is unbounded and **[H-1a]** holds. In this case, we replaced $\bar{q}$ by $Q(s) = \max(1, s)\bar{q}(s)$.
- when the drift is unbounded and **[H-1b]** holds. In this case, we replaced $\bar{q}$ by $Q(s) = \max(1, s, s^{\gamma-1})\bar{q}(s)$.

Note that when the drift is bounded and **[H-1a]** holds, we do not need to deteriorate the tempering function.

□

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:

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

Thus, when $s \asymp |\theta_{0,t}(y) - x|$, 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 results in Kolokolstov [Kol00b].

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 "semi group" property in the following sense.

Lemma 3.4. *Fix $t \in [0, T]$. Let us denote*

$$\bar{p}_C(t, x, y) = \frac{t^{-d/\alpha}}{\left(1 + \frac{|\theta_{0,t}(y)-x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(C|\theta_{0,t}(y) - x|).$$

Let $C_1, C_2 > 0$. For all $\tau \in [0, t]$, there exists $C_3 > 0$:

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

Proof. The proof follows by an application of the triangle inequality and the Lipschitz property of the flow plus the fact that Q are 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 (2.24) 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 $t \geq \tau > 0$, $(x, y) \in (\mathbb{R}^d)^2$:*

$$\begin{aligned}
\int_{\mathbb{R}^d} \delta \wedge |x - \theta_{0,\tau}(z)|^{\eta(\alpha\wedge 1)} \bar{p}(\tau, x, z) dz &\leq C t^{\omega}, \\
\int_{\mathbb{R}^d} \delta \wedge |\theta_{\tau,t}(y) - z|^{\eta(\alpha\wedge 1)} \bar{p}(t - \tau, z, y) dz &\leq C(t - \tau)^{\omega}.
\end{aligned}$$

As a corollary, we get that

$$\int_0^t \int_{\mathbb{R}^d} |H(t - \tau, z, y)| \leq C t^{\omega}.$$

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. Besides, it is enough to prove the property for $Q(s) = \max(1, s, s^{\gamma-1})\bar{q}(s)$. Also, under **[H-1a]**, we set $\gamma = d$. Let us denote by I the integral:

$$I = \int_{\mathbb{R}^d} dz \delta \wedge |x - \theta_{0,\tau}(z)|^{\eta(\alpha\wedge 1)} \frac{\tau^{-d/\alpha}}{\left(1 + \frac{|x-\theta_{0,\tau}(z)|}{\tau^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|x - \theta_{0,\tau}(z)|).$$

We split $\mathbb{R}^d = \{z \in \mathbb{R}^d; |x - \theta_{0,\tau}(z)| \leq \tau^{1/\alpha}\} \cup \{z \in \mathbb{R}^d; |x - \theta_{0,\tau}(z)| > \tau^{1/\alpha}\} = D_1 \cup D_2$. We write I_{D_i} for the integral over $z \in D_i$. For $z \in D_1$ we have:

$$\frac{\tau^{-d/\alpha}}{\left(1 + \frac{|x-\theta_{0,\tau}(z)|}{\tau^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|x - \theta_{0,\tau}(z)|) \leq \tau^{-d/\alpha}, \ |x - \theta_{0,\tau}(z)|^{\eta(\alpha\wedge 1)} \leq \tau^{\eta(1\wedge 1/\alpha)}.$$

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

When $z \in D_2$, we have:

$$\begin{aligned}
I_{D_2} &\leq \int_{D_2} dz \delta \wedge |x - \theta_{0,\tau}(z)|^{\eta(\alpha\wedge 1)} \frac{\tau^{1+\frac{\gamma-d}{\alpha}}}{|x - \theta_{0,\tau}(z)|^{\alpha+\gamma}} \\
&\leq \tau^{1+\frac{\gamma-d}{\alpha}} \int_{|z-\theta_{\tau,0}(x)|>C\tau^{1/\alpha}} dz \frac{\delta \wedge |z - \theta_{\tau,0}(x)|^{\eta(\alpha\wedge 1)}}{|z - \theta_{\tau,0}(x)|^{\alpha+\gamma}}.
\end{aligned}$$

Observe that we used the Lipschitz property of the flow to switch from $x-\theta_{0,\tau}(z)$ to $z-\theta_{\tau,0}(x)$. This allows us to change variables and set $X=(z-\theta_{\tau,0}(x))\tau^{-1/\alpha}$, we get:

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

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

$$\begin{aligned}\int_{|z-\theta_{\tau,0}(x)|>\tau^{1/\alpha}} \frac{\delta\wedge|z-\theta_{\tau,0}(x)|^{\eta(\alpha\wedge 1)}}{|z-\theta_{\tau,0}(x)|^{\alpha+\gamma}}dz &= \int_{1\ge|z-\theta_{\tau,0}(x)|>\tau^{1/\alpha}} \frac{\delta\wedge|z-\theta_{\tau,0}(x)|^{\eta(\alpha\wedge 1)}}{|z-\theta_{\tau,0}(x)|^{\alpha+\gamma}}dz\\ &+\int_{|z-\theta_{\tau,0}(x)|>1} \frac{\delta\wedge|z-\theta_{\tau,0}(x)|^{\eta(\alpha\wedge 1)}}{|z-\theta_{\tau,0}(x)|^{\alpha+\gamma}}dz.\end{aligned}$$

The second part of the right hand side is clearly a constant, bounding $\delta\wedge|z-\theta_{\tau,0}(x)|^{\eta(\alpha\wedge 1)}\le\delta$, since $\alpha+\gamma>d$. For the first part, we change variable again to $Y=(z-\theta_{\tau,0}(x))$, which yields when $\alpha+\gamma-d<\eta(\alpha\wedge 1)$:

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

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

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

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 is largely inspired by [Men11]. It relies on the smoothing properties, of the Parametrix kernel H.

Proof. We focus on uniqueness. Indeed, the existence stems from Theorem 2.1 in Bass [Bas88]. It relies on compactness arguments, in the lines of those developed in the diffusive case in Chapter 6 in Stroock and Varadhan [SV79], or Stroock [Str75] for a Lévy process with a Brownian part. The main idea consists in proving that the measures $(P_n)_{n\in\mathbb{N}^*}$ induced by the Euler-Maruyama schemes are tight. We also mention the semigroup-based approach to existence of Komatsu [Kom84].

Suppose we are given two solutions $\mathbb{P}^1$ and $\mathbb{P}^2$ of the martingale problem associated with $L(\cdot, \nabla.)$, starting in x at time 0. We can assume w.l.o.g. that $t \le T$, 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 $(X_t)_{t\ge 0}$ 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:

$$f(t,x) + \mathbb{E}^i \left(\int_t^T (\partial_s + L(x, \nabla_x)) f(s, X_s) ds \right) = 0, \ i \in \{1,2\}. \tag{3.35}$$

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

$$S^\Delta \Big((\partial_s + L(x, \nabla_x)) f \Big) = 0. \tag{3.36}$$

We now want to apply (3.36) to a suitable function f. For a fixed point $y \in \mathbb{R}^d$ and a given $\varepsilon \ge 0$, introduce for all $f \in \mathcal{C}_0^{1,2}([0,T) \times \mathbb{R}^d, \mathbb{R})$ the Green kernel:

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

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

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

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

$$\partial_t G^{\varepsilon,y} f(t,x) + M_{t,x}^{\varepsilon,y} f(t,x) = -f(t,x), \ \forall (t,x) \in [0,T) \times \mathbb{R}^d. \tag{3.37}$$

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

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

Then, by semigroup property, we have:

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

Hence, we can write:

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

Observe that from (3.37), we have:

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

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

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

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

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

On the other hand,

$$\begin{aligned}I_2^\varepsilon(t,x) = \int_{\mathbb{R}^d} dy \Big(L(x,\nabla_x)G^y\phi^{\varepsilon,y}(t,x) - M_{t,x}^{\varepsilon,y}\phi^{\varepsilon,y}(t,x)\Big) \\ = \int_{\mathbb{R}^d} dy \int_t^T ds \Big(L(x,\nabla_x) - L(\theta_{t,s+\varepsilon}(y),\nabla_x)\Big)\tilde{p}_\alpha^{s+\varepsilon,y}(s+\varepsilon-t,x,y)h(s,y) \\ = \int_{\mathbb{R}^d} dy \int_t^T ds H(s+\varepsilon-t,x,y)h(s,y).\end{aligned}$$

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

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

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 (3.36) 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 3.10.

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 3.10.

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

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

where we recall the notation $\rho(t,x,y) = \delta \wedge |\theta_{0,t}(y) - x|^{\eta(\alpha\wedge 1)}\bar{p}(t,x,y)$.

Proof. We follow the ideas of Kolokolstov [Kol00b]. Recall that

$$|H(t,x,y)| \le C\left(t^{-1/\alpha}\mathbf{1}_{\{\alpha>1\}} + \frac{\delta\wedge|x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}}{t}\right)\bar{p}(t,x,y),$$

where:

$$\bar{p}(t,x,y) = \frac{t^{-d/\alpha}}{\left(1+\frac{|\theta_{0,t}(y)-x|}{t^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|\theta_{0,t}(y)-x|),$$

and Q is non increasing. Under **[H-1a]**, we take $\gamma = d$. Observe that the following upper bound $\tilde{p}(t,x,y) \le C\bar{p}(t,x,y)$ holds trivially by definition of Q. Also, when $\alpha > 1$, the singularity $t^{-1/\alpha}$ is integrable, and from Proposition 3.4, this term yields the contribution $t^\omega\bar{p}(t,x,y)$, with $\omega = 1 - 1/\alpha > 0$.

When $\alpha \le 1$, we write:

$$\begin{aligned}
|\tilde{p} \otimes H(t,x,y)| &\le C \int_0^t d\tau \int_{\mathbb{R}^d} \bar{p}(\tau,x,z) \frac{\delta \wedge |z-\theta_{\tau,t}(y)|^{\eta(\alpha\wedge 1)}}{t-\tau} \bar{p}(t-\tau,z,y)dz \\
&\le C \int_0^t d\tau \int_{\mathbb{R}^d} \frac{\tau^{-d/\alpha}}{\left(1+\frac{|x-\theta_{0,\tau}(z)|}{\tau^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|x-\theta_{0,\tau}(z)|) \\
&\quad \times \frac{\delta \wedge |z-\theta_{\tau,t}(y)|^{\eta(\alpha\wedge 1)}}{t-\tau} \frac{(t-\tau)^{-d/\alpha}}{\left(1+\frac{|z-\theta_{\tau,t}(y)|}{(t-\tau)^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|z-\theta_{\tau,t}(y)|)dz.
\end{aligned}$$

Assume first that $|\theta_{0,t}(y)-x| \le Ct^{1/\alpha}$. Then, we split the time integral in $\int_0^{t/2} d\tau + \int_{t/2}^t d\tau$, and we use the fact that the Diagonal estimate is global. In the integral over $[t/2,t]$ we have that $\tau \asymp t$, so that

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

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

$$\begin{aligned}
\bar{p}(t,x,y) \int_{t/2}^t d\tau \int_{\mathbb{R}^d} \frac{\delta \wedge |z-\theta_{\tau,t}(y)|^{\eta(\alpha\wedge 1)}}{t-\tau} \frac{(t-\tau)^{-d/\alpha}}{\left(1+\frac{|z-\theta_{\tau,t}(y)|}{(t-\tau)^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|z-\theta_{\tau,t}(y)|)dz \\
\le Ct^{\omega}\bar{p}(t,x,y).
\end{aligned}$$

When, $\tau \in [0,t/2]$ we have $t-\tau \asymp t$, and we have

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

Next, we can bound

$$\delta \wedge |z-\theta_{\tau,t}(y)|^{\eta(\alpha\wedge 1)} \le C_t(\delta \wedge |\theta_{0,\tau}z - x|^{\eta(\alpha\wedge 1)} + \delta \wedge |x-\theta_{0,t}(y)|^{\eta(\alpha\wedge 1)}).$$

Thus, we finally obtain:

$$\begin{aligned}
\frac{1}{t}\bar{p}(t,x,y) \int_0^{t/2} d\tau \int_{\mathbb{R}^d} \frac{\tau^{-d/\alpha}}{\left(1+\frac{|\theta_{0,\tau}(z)-x|}{\tau^{1/\alpha}}\right)^{\alpha+\gamma}} Q(C|\theta_{0,\tau}(z)-x|) \\
\times\Big(\delta \wedge |\theta_{0,\tau}(z)-x|^{\eta(\alpha\wedge 1)} + \delta \wedge |\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}\Big)dz \\
\le C\Big(t^{\omega} + \delta \wedge |\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}\Big)\bar{p}(t,x,y).
\end{aligned}$$

Assume now that $|\theta_{0,t}(y)-x| \geq Ct^{1/\alpha}$. In this case, the off-diagonal estimate holds for $\bar{p}(t,x,y)$, that is:

$$\bar{p}(t,x,y) \asymp \frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}} Q(|\theta_{0,t}(y)-x|).$$

On the other hand, we have:

$$|\theta_{0,t}(y)-x| \leq C_t\Big(|\theta_{0,\tau}(z)-x| + |\theta_{\tau,t}(y)-z|\Big)$$

In other words, we have either $|\theta_{\tau,t}(y)-z| \geq C|\theta_{0,t}(y)-x|$, or $|\theta_{0,\tau}(z)-x| \geq C|\theta_{0,t}(y)-x|$. Consequently, we split $\mathbb{R}^d = D_1 \cup D_2$ with

$$\begin{aligned} D_1 &= \{z \in \mathbb{R}^d, |\theta_{\tau,t}(y)-z| \leq |\theta_{0,\tau}(z)-x|\}, \\ D_2 &= \{z \in \mathbb{R}^d, |\theta_{\tau,t}(y)-z| > |\theta_{0,\tau}(z)-x|\}. \end{aligned}$$

Now, when $z \in D_1$, we have that $|\theta_{0,t}(y)-x| \asymp |\theta_{0,\tau}(z)-x|$, thus $\bar{p}(\tau,x,z)$ is off-diagonal and we can bound:

$$\begin{aligned} \bar{p}(\tau,x,z) &\leq C\frac{\tau^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,\tau}(z)-x|^{\alpha+\gamma}} Q(|\theta_{0,\tau}(z)-x|) \\ &\leq C\frac{t^{1+\frac{\gamma-d}{\alpha}}}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}} Q(|\theta_{0,t}(y)-x|) \asymp \bar{p}(t,x,y). \end{aligned}$$

For the last inequality, we used the fact that Q is non increasing and that $\gamma+\alpha > d$ so that the exponent in τ is positive. Thus, we can take out $\bar{p}(\tau,x,z)$ 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:

$$\begin{aligned} I_{D_1} &\leq C\bar{p}(t,x,y)\int_0^t d\tau \int_{D_1} \frac{\delta \wedge |\theta_{\tau,t}(y)-z|^{\eta(\alpha\wedge 1)}}{t-\tau} \frac{(t-\tau)^{-d/\alpha}}{\left(1+\frac{|\theta_{\tau,t}(y)-z|}{(t-\tau)^{1/\alpha}}\right)^{\alpha+\gamma}} Q(|\theta_{\tau,t}(y)-z|)dz \\ &\leq Ct^{\omega}\bar{p}(t,x,y). \end{aligned}$$

When $z \in D_2$, we have $|\theta_{\tau,t}(y)-z| \asymp |\theta_{0,t}(y)-x|$. In this case, observe that we have:

$$\frac{1}{t-\tau}\bar{p}(t-\tau,z,y) \asymp \frac{(t-\tau)^{\frac{\gamma-d}{\alpha}}}{|\theta_{\tau,t}(y)-z|^{\alpha+\gamma}} Q(|\theta_{\tau,t}(y)-z|) \leq (t-\tau)^{\frac{\gamma-d}{\alpha}} \frac{Q(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}}.$$

Thus, the integral becomes:

$$\begin{aligned} I_{D_2} &\leq \frac{Q(|\theta_{0,t}(y)-x|)}{|\theta_{0,t}(y)-x|^{\alpha+\gamma}} \int_0^t (t-\tau)^{\frac{\gamma-d}{\alpha}} \int_{D_2} \bar{p}(\tau,x,z) \\ &\times(\delta\wedge|\theta_{\tau,t}(y)-z|^{\eta(\alpha\wedge 1)} + \delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}) \\ &\leq C\Big(t^{\omega}+\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}\Big)\bar{p}(t,x,y). \end{aligned}$$

To get the last inequality, we used Lemma 3.5 and integrated in time to recover $\bar{p}(t,x,y)$. 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 $(x,y)\in(\mathbb{R}^d)^2$:*

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

Proof. The proof is similar to the previous one, but now, due to the presence of $\delta\wedge|\theta_{0,\tau}(z)-x|^{\eta(\alpha\wedge 1)}$ 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 $(X_t)_{t\geq 0}$. Indeed, we have $p(t,x,y)=\tilde{p}(t,x,y)+p\otimes H(t,x,y)$. Also, we have the upper bound $p(t,x,y)\leq \bar{p}(t,x,y)$, which yields

$$\begin{aligned} p\otimes H(t,x,y) &\leq \int_0^t du \int_{\mathbb{R}^d} \bar{p}(u,x,z)\frac{\delta\wedge|\theta_{\tau,t}(y)-z|^{\eta(\alpha\wedge 1)}}{t-u}\bar{p}(t-u,z,y)dz \\ &\leq \Big(t^{\omega}+\delta\wedge|\theta_{0,t}(y)-x|^{\eta(\alpha\wedge 1)}\Big)\bar{p}(t,x,y). \end{aligned}$$

Thus, in diagonal regime, we have for t small enough $p(t,x,y)\geq Ct^{-d/\alpha}$. 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 for giving a lower bound for the density is to say that in order to go from x to y in time 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.

For large $|\theta_{0,t}(y) - x| \geq t^{1/\alpha}$, we write from the Chapman-Kolmogorov equation:

$$\begin{aligned} p(t,x,y) &= \int_{\mathbb{R}^d} dz p(t/2,x,z)p(t/2,z,y) \geq \int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} p(t/2,x,z)p(t/2,z,y)dz \\ &\geq \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha})\Big) \inf_{z\in B(\theta_{t/2,t}(y),Ct^{1/\alpha})} p(t/2,z,y) \\ &\geq \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha})\Big) C t^{-d/\alpha}. \end{aligned}$$

Consequently we have to give a lower bound for $\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha})\Big)$. To this end, we introduce the process $(X_t^\delta)_{t\geq 0}$ with jumps larger than δ removed. Specifically, $(X_t^\delta)_{t\geq 0}$ solves the SDE:

$$X_t^\delta = x + \int_0^t b(X_s^\delta)ds + \int_0^t \sigma(X_s^\delta)dZ_s^\delta,$$

where (Z_t^δ) is the process $(Z_t)_{t\geq 0}$ 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 $(X_t)_{t\geq 0}$ from $(X_t^\delta)_{t\geq 0}$ by introducing the arrival times of the compound poisson process:

$$N_t = \sum_{0<s\leq t} \Delta Z_s \mathbf{1}_{\{|\Delta Z_s|\geq\delta\}}.$$

Let us denote by $(T_k)_{k\geq 1}$ the arrival times of the process N_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:

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

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

$$\begin{aligned} \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha})\Big) &= \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha}); T_1 \geq t/2\Big) \\ &\quad + \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha}); T_1 \leq t/2\Big) \end{aligned}$$

We thus focus on:

$$\begin{aligned} &\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha}); T_1 \leq t/2\Big) \\ = \ &\mathbb{E}^x\left[\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y),Ct^{1/\alpha})\Big|\mathcal{F}_{T_1}\Big)\mathbf{1}_{\{T_1\leq t/2\}}\right], \end{aligned}$$

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

$$\begin{aligned}\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha})\Big|\mathcal{F}_{T_1}\Big) &= \mathbb{P}^{X_{T_1}}\Big(X_{t/2-T_1} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha})\Big)\\ &= \int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} p(t/2-T_1, X_{T_1}, z)dz.\end{aligned}$$

Thus, we have:

$$\begin{aligned}&\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha}); T_1 \le t/2\Big)\\ =\ & \mathbb{E}^x\left[\int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} p(t/2-T_1, X_{T_1}, z)dz\mathbf{1}_{\{T_1\le t/2\}}\right]\end{aligned}$$

Now, since we have $X_{T_1} = X_{T_1^-}^\delta + \sigma(X_{T_1^-}^\delta)\Delta Z_{T_1}$, 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:

$$\begin{aligned}&\mathbb{E}^x\left(\int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} p(t/2-T_1, X_{T_1}, z)dz\middle|X_{T_1^-}\right)\\ =& \int_{\mathbb{R}^d}\int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} dz\ p(t/2-T_1, \sigma(X_{T_1^-}^\delta)w + X_{T_1^-}\delta, z)\frac{\nu(dw)}{\nu\big(B(0,\delta)^c\big)}.\end{aligned}$$

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

$$\begin{aligned}&\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha}); T_1 \le t/2\Big)\\ =\ & \mathbb{E}^x\bigg[\int_0^{t/2} ds\int_{B(\theta_{t/2,t}(y),Ct^{1/\alpha})} dz\int_{\mathbb{R}^d}\frac{\nu(dw)}{\nu\big(B(0,\delta)^c\big)}\\ &\times p(t/2-s, \sigma(X_s^\delta)w + X_s^\delta, z)\nu\big(B(0,\delta)^c\big)e^{-s\nu\left(B(0,\delta)^c\right)}\bigg].\end{aligned}$$

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 $\sigma(X_s^\delta)w + X_s^\delta$ is close to $\theta_{s,t/2}(z)$. That is, where the density $p(t/2-s, \sigma(X_s^\delta)w + X_s^\delta, z)$ is in diagonal regime:

$$
\begin{aligned}
&\mathbb{P}^x\Big(X_{t/2}\in B(\theta_{t/2,t}(y),Ct^{1/\alpha});T_1\le t/2\Big)\\
\ge\ &\mathbb{E}^x\Big[\int_0^{\frac{t}{2}}ds\int_{B(\theta_{\frac{t}{2},t}(y),Ct^{\frac{1}{\alpha}})}dz\int_{\{|\sigma(X_s^\delta)w+X_s^\delta-\theta_{s,\frac{t}{2}}(z)|\le C(\frac{t}{2}-s)^{\frac{1}{\alpha}}\}}\nu(dw)\\
&\times p\Big(\frac{t}{2}-s,\sigma(X_s^\delta)w+X_s^\delta,z\Big)e^{-s\nu(B(0,\delta)^c)}\Big]\\
\ge\ &\mathbb{E}^x\Big[\int_0^{\frac{t}{2}}ds\Big(\frac{t}{2}-s\Big)^{-d/\alpha}\int_{B(\theta_{\frac{t}{2},t}(y),Ct^{1/\alpha})}dz\\
&\times\nu\left(B\left(\sigma(X_s^\delta)^{-1}(\theta_{s,\frac{t}{2}}(z)-X_s^\delta),C\left(\frac{t}{2}-s\right)^{1/\alpha}\right)\right)e^{-s\nu(B(0,\delta)^c)}\Big].
\end{aligned}
$$

Additionally, we can lower bound the last probability by localizing X_s^δ close to $\theta_{s,0}(x)$:

$$
\begin{aligned}
&\mathbb{P}^x\Big(X_{t/2}\in B(\theta_{t/2,t}(y),Ct^{1/\alpha});T_1\le t/2\Big)\\
\ge\ &\mathbb{E}^x\Big[\int_0^{\frac{t}{2}}ds\mathbf{1}_{\{|X_s^\delta-\theta_{s,0}(x)|\le Cs^{1/\alpha}\}}\Big(\frac{t}{2}-s\Big)^{-d/\alpha}\int_{B(\theta_{\frac{t}{2},t}(y),Ct^{1/\alpha})}dz\\
&\times\nu\left(B\left(\sigma(X_s^\delta)^{-1}(\theta_{s,\frac{t}{2}}(z)-X_s^\delta),C\left(\frac{t}{2}-s\right)^{1/\alpha}\right)\right)e^{-s\nu(B(0,\delta)^c)}\Big].
\end{aligned}
$$

Now, from assumption **[H-LB]**, we have that $\nu(B(0,\delta)^c)\le 1/\delta^\alpha$ so that taking $\delta=t^{1/\alpha}$ yields $e^{-s\nu(B(0,\delta)^c)}\ge C$. Also, since $z\in B(\theta_{0,\frac{t}{2}}(y),Ct^{\frac{1}{\alpha}})$, by the Lipschitz property of the flow,

$$\theta_{s,t/2}(z)\in\theta_{s,t/2}\Big(B(\theta_{\frac{t}{2},t}(y),Ct^{\frac{1}{\alpha}})\Big)\subset B(\theta_{s,t}(y),Ct^{\frac{1}{\alpha}}).$$

On the other hand, $X_s^\delta\in B(\theta_{s,0}(x),s^{1/\alpha})$, thus,

$$\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z)-X_s^\delta)\in B\big(\sigma(X_s^\delta)^{-1}(\theta_{0,t}(y)-x),Ct^{1/\alpha}\big).$$

Note that the constant C radius $Ct^{1/\alpha}$ have changed. Now, using the Hölder continuity of σ, since $|X_s^\delta-\theta_{s,0}(x)|\le s^{1/\alpha}$, we have up to a modification of C:

$$B\big(\sigma(X_s^\delta)^{-1}(\theta_{0,t}(y)-x),Ct^{1/\alpha}\big)\subset B\big(\sigma(\theta_{s,0}(x))^{-1}(\theta_{0,t}(y)-x),Ct^{1/\alpha}\big).$$

so that we obtain:

$$\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z)-X_s^\delta)\in B\big(\sigma(\theta_{s,0}(x))^{-1}(\theta_{0,t}(y)-x),Ct^{1/\alpha}\big)$$

Thus, if we have:

$$\forall s \in [0, t/2],\ B\big(\sigma(\theta_{s,0}(x))^{-1}(\theta_{0,t}(y) - x), Ct^{1/\alpha}\big) \subset A_{low},$$

then $\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta) \in A_{low}$. Consequently, we can use the lower bound in **[H-LB]** to get:

$$\begin{aligned} & \nu\left(B\Big(\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta), C\left(\frac{t}{2} - s\right)^{1/\alpha}\Big)\right) \\ \geq\ & C\left(\frac{t}{2} - s\right)^{\gamma/\alpha} \frac{\underline{q}(|\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)|)}{|\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)|^{\gamma+\alpha}}. \end{aligned}$$

We thus obtain:

$$\begin{aligned} & \mathbb{P}^x\Big(X_{t/2} \in B(y, Ct^{1/\alpha}); T_1 \leq t/2\Big) \\ \geq\ & C\mathbb{E}^x\Bigg[\int_0^{t/2} ds \mathbf{1}_{\{|X_s^\delta - \theta_{s,0}(x)| \leq Cs^{1/\alpha}\}} \left(\frac{t}{2} - s\right)^{\frac{\gamma-d}{\alpha}} \\ & \times \int_{B(\theta_{t/2,t}(y), Ct^{1/\alpha})} dz \frac{\underline{q}(|\sigma(X_s^\delta)^{-1}(\theta_{t,s}(z) - X_s^\delta)|)}{|\sigma(X_s^\delta)^{-1}(\theta_{t,s}(z) - X_s^\delta)|^{\gamma+\alpha}}\Bigg]. \end{aligned}$$

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

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)| \leq C|y - \theta_{t,0}(x)|.$$

We write from the ellipticity of σ:

$$\begin{aligned} |\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)| &\leq C|\theta_{s,t/2}(z) - X_s^\delta| \\ &\leq C(|\theta_{s,t/2}(z) - \theta_{s,0}(x)| + |\theta_{s,0}(x) - X_s^\delta|). \end{aligned}$$

Now, in the considered set, $|\theta_{s,0}(x) - X_s^\delta| \leq Cs^{1/\alpha} \leq Ct^{1/\alpha} \leq C|\theta_{0,t}(y) - x|$. Thus, we have:

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)| \leq C\big(|\theta_{s,t/2}(z) - \theta_{s,0}(x)| + C|\theta_{0,t}(y) - x|\big).$$

On the other hand, we can write:

$$|\theta_{s,t/2}(z) - \theta_{s,0}(x)| \ \leq\ |\theta_{s,t/2}(z) - \theta_{s,t}(y)| + |\theta_{s,t}(y) - \theta_{s,0}(x)|.$$

Thus, from the Lipschitz property of the flow,

$$|\theta_{s,t}(y) - \theta_{s,0}(x)| \leq C_T|\theta_{0,t}(y) - x|.$$

On the other hand, from the homogeneity of the system, we have $\theta_{s,t}(y) = \theta_{s,t/2} \circ \theta_{t/2,t}(y)$ so that:

$$|\theta_{s,t/2}(z) - \theta_{s,t}(y)| = |\theta_{s,t/2}(z) - \theta_{s,t/2} \circ \theta_{t/2,t}(y)| \leq C_T|z - \theta_{t/2,t}(y)|,$$

where to the get the last inequality, we once again relied on the Lipschitz property of the flow. We recall that $|z - \theta_{t/2,t}(y)| \leq Ct^{1/\alpha} \leq |\theta_{0,t}(y) - x|$, consequently we finally obtain:

$$|\sigma(X_s^\delta)^{-1}(\theta_{s,t/2}(z) - X_s^\delta)| \leq C_T|\theta_{0,t}(y) - x|.$$

Using this last inequality to estimate the probability:

$$\begin{aligned}
& \mathbb{P}^x\Big(X_{t/2} \in B(\theta_{0,t/2}, Ct^{1/\alpha}); T_1 \leq t/2\Big) \\
\geq \; & C\mathbb{E}^x\left[\int_0^{t/2} ds \mathbf{1}_{\{|X_s^\delta - \theta_{s,0}(x)| \leq s^{1/\alpha}\}} \left(\frac{t}{2} - s\right)^{\frac{\gamma-d}{\alpha}} \int_{B(\theta_{t/2,t}(y), Ct^{1/\alpha})} dz \frac{\underline{q}(|\theta_{0,t}(y) - x|)}{|\theta_{0,t}(y) - x|^{\gamma+\alpha}}\right] \\
\geq \; & Ct^{d/\alpha} \frac{\underline{q}(|\theta_{0,t}(y) - x|)}{|\theta_{0,t}(y) - x|^{\gamma+\alpha}} \int_0^{t/2} ds \left(\frac{t}{2} - s\right)^{\frac{\gamma-d}{\alpha}} \mathbb{P}(|X_s^\delta - \theta_{s,0}(x)| \leq s^{1/\alpha}),
\end{aligned}$$

where $t^{d/\alpha}$ comes from the volume of the ball $B(\theta_{t/2,t}(y), Ct^{1/\alpha})$ obtained from the integral in dz. Using the diagonal lower estimates for the density, we actually see that $\mathbb{P}(|X_s^\delta - \theta_{s,0}(x)| \leq s^{1/\alpha}) \asymp 1$, therefore:

$$\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha}); T_1 \leq t\Big) \geq Ct^{d/\alpha} t^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{0,t}(y) - x|)}{|\theta_{0,t}(y) - x|^{\gamma+\alpha}}.$$

Returning to the first estimate on the density yields:

$$\begin{aligned}
p(t,x,y) \;&\geq\; Ct^{-d/\alpha}\mathbb{P}^x\Big(X_{t/2} \in B(\theta_{t/2,t}(y), Ct^{1/\alpha}); T_1 \leq t\Big) \\
&\geq\; Ct^{-d/\alpha} t^{d/\alpha} t^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{0,t}(y) - x|)}{|\theta_{0,t}(y) - x|^{\gamma+\alpha}} \\
&=\; Ct^{1+\frac{\gamma-d}{\alpha}} \frac{\underline{q}(|\theta_{0,t}(y) - x|)}{|\theta_{0,t}(y) - x|^{\gamma+\alpha}},
\end{aligned}$$

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.

Chapter 3

A Parametrix Approach for some Degenerate Stable Driven SDEs

We consider a stable driven degenerate stochastic differential equation, whose coefficients satisfy a kind of weak Hörmander condition. Under mild smoothness assumptions we prove the uniqueness of the martingale problem for the associated generator under some dimension constraints. Also, when the driving noise is scalar and tempered, we establish density bounds reflecting the multi-scale behavior of the process.

1 Introduction.

The aim of this Chapter is to study degenerate stable driven stochastic differential equations of the following form:

$$\begin{aligned}
dX_t^1 &= \left(a_t^{1,1}X_t^1 + \cdots + a_t^{1,n}X_t^n\right)dt + \sigma(t, X_{t^-})dZ_t \\
dX_t^2 &= \left(a_t^{2,1}X_t^1 + \cdots + a_t^{2,n}X_t^n\right)dt \\
dX_t^3 &= \left(a_t^{3,2}X_t^2 + \cdots + a_t^{3,n}X_t^n\right)dt \\
&\vdots \\
dX_t^n &= \left(a_t^{n,n-1}X_t^{n-1} + a_t^{n,n}X_t^n\right)dt,\ X_0 = x \in \mathbb{R}^{nd},
\end{aligned} \tag{1.1}$$

where Z is an $\mathbb{R}^d$ valued symmetric α stable process (possibly tempered and with $\alpha \in (0,2)$), $\sigma : \mathbb{R}^+ \times \mathbb{R}^{nd} \to \mathbb{R}^d \otimes \mathbb{R}^d$, $a^{i,j} : \mathbb{R}^+ \to \mathbb{R}^d \otimes \mathbb{R}^d$, $i \in [\![1, n]\!]$, $j \in [\![(i-1) \vee 1, n]\!]$. Observe that $(X_t)_{t\geq 0} = (X_t^1, \cdots, X_t^n)_{t\geq 0}$ is $\mathbb{R}^{nd}$ valued. We will often use the shortened form:

$$dX_t = A_t X_t dt + B\sigma(t, X_{t^-})dZ_t,\ X_0 = x, \tag{1.2}$$

where $B = \big(I_{d\times d}\ 0_{(n-1)d\times d}\big)^*$ denotes the injection matrix from $\mathbb{R}^d$ into $\mathbb{R}^{nd}$, with $*$ standing for the transposition, and A_t is the matrix :

$$A_t = \begin{pmatrix} a_t^{1,1} & \dots & \dots & \dots & a_t^{1,n} \\ a_t^{2,1} & \ddots & & & a_t^{2,n} \\ 0 & a_t^{3,2} & \ddots & & a_t^{3,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_t^{n,n-1} & a_t^{n,n} \end{pmatrix}.$$

The previous system appears in many applicative fields. It is for instance related for $n = 2$ to the pricing of Asian options in jump diffusions models (see e.g. Jeanblanc *et al* [JYC09] or Barucci *et al* [BPV01] in the Brownian case). The Hamiltonian formulation in mechanics can lead to systems corresponding to the drift part of (1.1) (still with $n = 2$). The associated Brownian perturbation has been thoroughly studied, see e.g. Talay [Tal02] or Stuart *et al.* [MSH02] for the convergence of approximation schemes to equilibrium, but to the best of our knowledge other perturbations, like the current stable one, have not yet been much considered. For a general n, equation (1.1) can be seen as the linear dynamics of n coupled oscillators in dimension d perturbed by a stable anisotropic noise. Observe also that in the diffusive case these oscillator chains naturally appear in statistical mechanics, see e.g Eckman *et al.* [EPRB99].

Equation (1.1) is degenerate in the sense that the noise only acts on the first component of the system. Additionally to the non-degeneracy of the *volatility* σ, we will assume a kind of weak Hörmander condition on the drift component in order to allow the noise propagation into the system.

A huge literature exists on degenerate Brownian diffusions under the *strong* Hörmander condition, i.e. when the underlying space is spanned by the diffusive vector fields and their iterated Lie brackets. The major works in that framework have been obtained in a series of papers by Kusuoka ans Stroock, [KS84], [KS85], [KS87], using a Malliavin calculus approach.

For the weak Hörmander case, many questions are still open even in the Brownian setting. Let us mention in this framework the papers [DM10], [Men11] and [KMM10] dealing respectively with density estimates, martingale problems and random walk approximations for systems of type (1.1) or that can be linearized around such systems. In those works a global multi-scale Gaussian regime holds. For highly non-linear first order vector fields, Franchi [Fra14] and Cinti *et al.* [CMP15] address issues for which there is not a single regime anymore. A specificity of the weak Hörmander condition is the unbounded first order term which does not lead to a time-space separation in the off-diagonal bounds for the density estimates as in the sub-Riemannian setting, see e.g. [KS87], Ben Arous and Léandre [BAL91] and references therein. The energy of the associated deterministic control problem has to be considered instead, see e.g. [DM10]. We have a similar feature in our current stable setting.

In this work we are first interested in proving the uniqueness of the martingale problem associated with the generator $(L_t)_{t\geq}$ of (1.1), i.e. for all $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$ (twice continuously differentiable functions with compact support)

$$\forall x \in \mathbb{R}^{nd},\ L_t\varphi(x) = \langle A_t x, \nabla\varphi(x)\rangle + \int_{\mathbb{R}^d}\left(\varphi(x + B\sigma(t,x)z) - \varphi(x) - \frac{\langle \nabla\varphi(x), B\sigma(t,x)z\rangle}{1+|z|^2}\right) g(|z|)\nu(dz), \tag{1.3}$$

under some mild assumptions on the volatility σ, the Lévy measure ν of a symmetric α stable process and the tempering function g (which is set to 1 in the stable case). To this end, the key tool consists in exploiting some properties of the joint densities of (possibly tempered) stable processes and their iterated integrals, corresponding to the proxy model in a *parametrix* continuity technique (see e.g. Friedman [Fri64] or McKean and Singer [MKS67]). Following the strategies developed in [BP09], [Men11] we then derive uniqueness exploiting the smoothing properties of the parametrix kernel. For this approach to work, we anyhow consider some restrictions on the dimensions n, d. Let us indeed emphasize that the density of a d-dimensional α-stable process and its $n-1$ iterated integrals behaves as the density of an α stable process in dimension nd with a modified Lévy measure and different time-scales. The first point can be checked through Fourier arguments (see Proposition 5.3 and Remark 5.2). Also, the typical time scale of the initial stable process is $t^{1/\alpha}$ and $t^{(i-1)+1/\alpha}$ for the associated $(i-1)^{\text{th}}$ integral. One of the difficulties is now that the associated Lévy measures (on $\mathbb{R}^{nd}$) have spectral parts that are either not equivalent or singular with respect to the Lebesgue measure of S^{nd-1}. The link between the behavior of the stable density and the corresponding spectral measure is discussed intensively in Watanabe [Wat07] and can lead to rather subtle phenomena. Roughly speaking, the lower is the dimension of the support of the spectral measure, the heavier the tail. This is what leads us to consider some restrictions on the dimensions. Also, using the resolvent[1] associated with the ordinary differential equation obtained from (1.1) setting $\sigma = 0$, i.e. $\frac{d}{dt}R_t = A_t R_t,\ R_0 = I_{nd\times nd}$, the mean of the process is $R_t x$ at time t (transport of the initial condition by the resolvent). The process will deviate from its *mean* accordingly to the associated component wise time scales.

When turning to density estimates, one of the dramatic differences with the Gaussian case is the lack of integrability of the driving process. For non-degenerate stable driven SDEs, this difficulty can be bypassed to derive two-sided pointwise bounds for the SDE that are homogeneous to the density of the driving process. Kolokoltsov [Kol00b] establishes in the stable case the analogue of the *Aronson* bounds for diffusions, see e.g. Sheu [She91] or [Aro67]. For approximation schemes of non-degenerate stable-driven SDEs we also mention [KM10]. In our current degenerate framework,

[1] We carefully mention that we use the term *resolvent* in the sense of ordinary differential equations.

working under somehow minimal assumptions to derive pointwise density bounds, that is Hölder continuity of the coefficients, we did not succeed to get rid of those integrability problems. We are also faced with a new difficulty due to the degeneracy and the non-local character of L. Namely, we have a disturbing *rediagonalization phenomenon*: when the density $\tilde{p}(t,x,\cdot)$ of $\tilde{X}_t = x + \int_0^t A_s\tilde{X}_s ds + BZ_t$ is in a *large deviation regime*, estimating the non local part of $L_t\tilde{p}$ (which is the crucial quantity to control in a parametrix approach), the very large jumps can lead to integrate $\tilde{p}$ on a set where it is in its typical regime. This phenomenon already appears in the non-degenerate case, but the difficulty here is that there is a dimension mismatch between the tail behavior of $\tilde{p}(t,x,\cdot)$, density of $\tilde{X}_t$, multi-scale stable process of dimension nd, and the one of the jump, stable process of dimension d.

This quite tricky phenomenon leads us to temper the driving noise in order to obtain density estimates through a parametrix continuity technique. For technical reasons that will appear later on, we establish when $d = 1, n = 2$ (scalar non-degenerate diffusion and associated non-degenerate integral) the expected upper-bound up to an additional logarithmic contribution, when the coefficient $\sigma(t,x) = \sigma(t,x^2)$ depends on the *fast* variable. This constraint appears in order to compensate an additional time singularity deriving from the rediagonalization. Roughly speaking, the dependence on the fast variable only gives a better smoothing effect for the parametrix kernel. Eventually, we derive the expected diagonal lower bound, see Theorem 2.2. To this end we use a parametrix approach similar to the one of Mc Kean and Singer [MKS67]. Working with smoother coefficients would have allowed to consider Malliavin calculus type techniques. In the jump case, this approach has been investigated to establish existence/smoothness of the density for SDEs by Bichteler *et al.* in the non-degenerate case [BGJ87], and Léandre in the degenerate one, see [Léa85],[Léa88a]. Also, we mention the recent work of Zhang [Zha14a] who obtained existence and smoothness results for the density of equations of type (1.1) in arbitrary dimension for smooth coefficients, and a possibly non linear drift, still satisfying a weak Hörmander condition. His approach relies on the *subordinated* Malliavin calculus, which consists in applying the *usual* Malliavin calculus techniques on a Brownian motion observed along the path of an α-stable subordinator.

Let us eventually mention some related works. Priola and Zabczyk establish in [PZ09] existence of the density for processes of type (1.1), under the same kind of weak Hörmander assumption and when σ is constant, for a general driving Lévy process Z provided its Lévy measure is infinite and has itself a density on compact sets. Also, Picard, [Pic96] investigates similar problems for singular Lévy measures. Other results concerning the smoothness of the density of Lévy driven SDEs have been obtained by Ishikawa and Kunita [IK06] in the non-degenerate case but with mild conditions on the Lévy measure and by Cass [Cas09] who gets smoothness in the weak Hörmander framework under technical restrictions. Also, we refer to the work of Watanabe [Wat07] for two-sided heat-kernel estimates for stable processes with very general spectral measures. Those estimates have been extended to the tempered stable case by Sztonyk [Szt10].

The article is organized as follows. We state our main results in Section 2. In Section 3, we explain the procedure to derive those results and also state the density estimates on the process in (1.1) when $\sigma(t,x) = \sigma(t)$ (*frozen process*). We then prove the uniqueness of the martingale problem in Section 4. A non linear extension is discussed in Appendix 9. Sections 5 and 6 are the technical core of the paper. In particular, we prove there the existence of the density and the associated estimates for the frozen process and establish the smoothing properties of the parametrix kernel. Appendices 7 and 8 are dedicated to the derivation of stable density bounds and kernels combining the approaches of [Kol00b] and [Wat07], [Szt10] in our current degenerate setting. We emphasize that the tempering procedure allows to get rid of the integrability problems but does not prevent from the rediagonalization phenomenon. This difficulty would occur even in the truncated case, thoroughly studied in the non-degenerate case by Chen *et al.* [CKK08]. The truncation would certainly *relocalize* the operator but the rediagonalization would still perturb the parametrix iteration in the stable regime.

2 Assumptions, and Main Result.

We will make the following assumptions:

About the Coefficients: The coefficients are assumed to be bounded and measurable in time and also to satisfy the conditions below.

[H-1]: (**Hölder regularity in space**) $\exists H > 0,\ \eta \in (0,1],\ \forall x, y \in \mathbb{R}^{nd}$ and $\forall t \geq 0$,

$$||\sigma(t,x) - \sigma(t,y)|| \leq H|x-y|^{\eta}.$$

[H-2]: (**Ellipticity**) $\exists\ \kappa \geq 1$, $\forall \xi \in \mathbb{R}^d$, $\forall z \in \mathbb{R}^{nd}$ and $\forall t \geq 0$,

$$\kappa^{-1}|\xi|^2 \leq \langle \xi, \sigma\sigma^*(t,z)\xi\rangle \leq \kappa|\xi|^2. \tag{2.1}$$

[H-3]: (**Hörmander-like condition for** $(A_t)_{t\geq 0}$) $\exists \overline{\alpha},\ \underline{\alpha} \in \mathbb{R}^{+*}$, $\forall \xi \in \mathbb{R}^{nd}$ and $\forall t \geq 0$, $\underline{\alpha}|\xi|^2 \leq \langle a_t^{i,i-1}\xi, \xi\rangle \leq \overline{\alpha}|\xi|^2,\ \forall i \in [\![2, n-1]\!]$. Also, for all $(i,j) \in [\![1,n]\!]^2,\ \|a_t^{i,j}\| \leq \overline{\alpha}$.

About the Driving Noise.

Stable Case: Let us first consider $(Z_t)_{t\geq 0}$ to be an α stable symmetric process, defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, that is a Lévy process with Fourier exponent:

$$\mathbb{E}e^{i\langle p, Z_t\rangle} = \exp\left(-t\int_{S^{d-1}} |\langle p, \varsigma\rangle|^{\alpha}\mu(d\varsigma)\right),\ \forall p \in \mathbb{R}^d.$$

In the above expression, we denote by S^{d-1} the unit sphere in $\mathbb{R}^d$, and by μ the spectral measure of Z. This measure is related to the Lévy measure of Z as follows. If ν is the Lévy measure of Z, its decomposition in polar coordinates writes:

$$\nu(dz) = \frac{d\rho}{\rho^{1+\alpha}}\tilde{\mu}(d\varsigma),\ z = \rho\varsigma,\ (\rho,\varsigma) \in \mathbb{R}^+ \times S^{d-1}. \tag{2.2}$$

Then, $\mu = C_{\alpha,d}\tilde{\mu}$ (see Sato [Sat05] for the exact value of $C_{\alpha,d}$). In that case we suppose

[H-4]: (**Non degeneracy of the spectral measure**) We assume that μ is absolutely continuous w.r.t. to the Lebesgue measure of S^{d-1} with Lipschitz density h and that there exists $\lambda \geq 1$, s.t. for all $u \in \mathbb{R}^d$,

$$\lambda^{-1}|u|^\alpha \leq \int_{S^{d-1}} |\langle u, \varsigma\rangle|^\alpha \mu(d\varsigma) \leq \lambda|u|^\alpha. \tag{2.3}$$

Tempered Case: In the tempered case we simply assume that $(Z_t)_{t\geq 0}$ has generator:

$$L_Z\phi(x) = \int_{\mathbb{R}^d} \{\phi(x+z) - \phi(x) - \frac{\langle \nabla\phi(x), z\rangle}{1+|z|^2}\}g(|z|)\nu(dz),\ \phi \in C_0^2(\mathbb{R}^d, \mathbb{R}), \tag{2.4}$$

where the measure ν is as in the stable case and the tempering function $g : \mathbb{R}^{+*} \to \mathbb{R}^{+*}$ satisfies

[T]: **(Smoothness, Doubling property and Decay associated with the tempering function** g**)** We first assume that $g \in C^1(\mathbb{R}^+, \mathbb{R}^{+*})$ and that there exists $a > 0$ s.t. $g \in C^2([0, a], \mathbb{R}^{+*})$ if $\alpha \in [1, 2)$. We also suppose that there exists $c > 0$ s.t. for all $r > 0, g(r) + r\sup_{u\in[\kappa^{-1},\kappa]} g'(ur) \leq c\theta(r)$ for κ as in **[H-2]** and where $\theta : \mathbb{R}^{+*} \to \mathbb{R}^{+*}$ is a bounded non-increasing function satisfying:

$$\exists D \geq 1,\ \forall r > 0,\ \theta(r) \leq D\theta(2r),\ (1+r)\theta(r) := \Theta(r) \underset{r\to+\infty}{\to} 0.$$

Typical examples of tempering functions satisfying **[T]** are for instance $r \to g(r) = \exp(-cr), c > 0, g(r) = (1+r)^{-m},\ m \geq 2$.

We say that **[HS]** (resp. **[HT]**) holds if conditions **[H-1]** to **[H-4]** are fulfilled and the driving noise Z is a symmetric stable process (resp. a tempered stable process satisfying **[T]**). We say that **[H]** is satisfied if **[HS]** or **[HT]** holds, i.e. the results under **[H]** hold for both the stable and the tempered stable driving process.

Our main results are the following.

Theorem 2.1 (**Weak Uniqueness**). *Under [H], i.e. in both the stable and the tempered stable case, the martingale problem associated with the generator* $(L_t)_{t\geq 0}$*, defined in* (1.3)*, of the degenerate equation* (1.1)*:*

$$dX_t = A_tX_tdt + B\sigma(t, X_{t-})dZ_t,$$

admits a unique solution provided $d(1-n)+1+\alpha>0$*. That is, for every* $x\in\mathbb{R}^{nd}$*, there exists a unique probability measure* $\mathbb{P}$ *on* $\Omega=\mathcal{D}(\mathbb{R}^+,\mathbb{R}^{nd})$ *the space of càdlàg functions, such that for all* $f\in C_0^{1,2}(\mathbb{R}^+\times\mathbb{R}^{nd},\mathbb{R})$*, denoting by* $(X_t)_{t\geq 0}$ *the canonical process, we have:*

$$\mathbb{P}(X_0=x)=1 \quad and \quad f(t,X_t)-\int_0^t(\partial_u+L_u)f(u,X_u)du \quad is\ a\ \mathbb{P}\text{- } martingale.$$

Hence, weak uniqueness holds for (1.1)*.*

Now, if $d=1, n=2$ *and* $\alpha>1$*, the well-posedness of the martingale problem extends to the case of a non-linear Lipschitz drift satisfying a Hörmander-like non-degeneracy condition. Namely, weak uniqueness holds for:*

$$\begin{aligned} X_t^1 &= x^1+\int_0^t F_1(s,X_s)ds+\int_0^t\sigma(s,X_{s^-})dZ_s,\\ X_t^2 &= x^2+\int_0^t F_2(s,X_s)ds, \end{aligned} \tag{2.5}$$

provided $F=(F_1,F_2)^*:\mathbb{R}^+\times\mathbb{R}^2\to\mathbb{R}^2$ *is measurable and bounded in time, uniformly Lipschitz in space and such that* $\partial_{x^1}F_2\in[c_0,c_0^{-1}], c_0\in(0,1]$ *and* $\partial_{x^1}F_2$ *is* η*-Hölder continuous,* $\eta\in(0,1]$*.*

Remark 2.1. The dimension constraint comes from the *worst* asymptotic behavior of the stable densities in our current case. Viewing the density of the stable process Z and its iterated integrals as the density of an nd-dimensional multi-scale stable process yields to consider a Lévy measure on $\mathbb{R}^{nd}$ for which the support of the spectral measure has dimension $(d-1)+1=d$. Thus, from Theorem 1.1 in Watanabe [Wat07] we have that, at time 1 (to get rid of the multi-scale feature), the tails will behave at least as $|x|^{-(d+1+\alpha)}$ for large values of $|x|, x\in\mathbb{R}^{nd}$. The condition in the previous Theorem is imposed in order to have the integrability of the worst bound in $\mathbb{R}^{nd}$. We refer to Section 5.2 for details. In practice the condition is fulfilled for:

- $d=1, n=2$ for $\alpha\in(0,2)$.
- $d=1, n=3$ for $\alpha\in(1,2)$.
- $d=2, n=2$ for $\alpha\in(1,2)$.

Remark 2.2. We point out that the tempering function g does not play any role in the proof of the previous theorem. Furthermore, it cannot be used to weaken the previous dimension constraints. Indeed, it can be seen from the estimates in Proposition 3.4 that the additional multiplicative term in θ makes the worst bound integrable but also yields an explosive contribution in small time.

Also, when $d = 1$ and $n = 2$ in (1.1) we are able to prove the following density estimates in the tempered case.

Theorem 2.2 (**Density Estimates**). *Assume that $d = 1$, $n = 2$. Under **[HT]** and for $\sigma(t,x) := \sigma(t,x^2)$, i.e. the diffusion coefficient depends on the fast component, provided $1 \ge \eta > \frac{1}{(1\wedge\alpha)(1+\alpha)}$, the unique weak solution of (1.1) has for every $s > 0$ a density with respect to the Lebesgue measure. Precisely, for all $0 \le t < s$ and $x \in \mathbb{R}^2$,*

$$\mathbb{P}(X_s \in dy | X_t = x) = p(t,s,x,y)dy. \tag{2.6}$$

Also, for a deterministic time horizon $T > 0$, and a fixed threshold $K > 0$, there exists $C_{2.2} := C_{2.2}(\textbf{[HT]}, T, K) \ge 1$, s.t. $\forall 0 \le t < s \le T$, $\forall (x,y) \in (\mathbb{R}^2)^2$,

$$p(t,s,x,y) \le C_{2.2}\bar{p}_{\alpha,\Theta}(t,s,x,y)\left(1 + \log(K \vee |(\mathbb{T}^{\alpha}_{s-t})^{-1}(y - R_{s,t}x)|\right), \tag{2.7}$$

where for all $u \in \mathbb{R}^+$, $\mathbb{T}^{\alpha}_u := \mathrm{Diag}\big((u^{1/\alpha}, u^{1+1/\alpha})\big)$, $\mathbb{M}_u := \mathrm{Diag}\big(1, u\big)$ and

$$\bar{p}_{\alpha,\Theta}(t,s,x,y) = \bar{C}_{\alpha,\Theta}\frac{\det(\mathbb{T}^{\alpha}_{s-t})^{-1}}{\{K \vee |(\mathbb{T}^{\alpha}_{s-t})^{-1}(y - R_{s,t}x)|\}^{2+\alpha}}\Theta(|\mathbb{M}^{-1}_{s-t}(y - R_{s,t}x)|).$$

Here, $R_{s,t}$ stands for the resolvent associated with the deterministic part of (1.1), i.e. $\frac{d}{ds}R_{s,t} = A_s R_{s,t}$, $R_{t,t} = I_{2\times 2}$, and $\bar{C}_{\alpha,\Theta}$ is s.t. $\int_{\mathbb{R}^2}\bar{p}_{\alpha,\Theta}(t,s,x,y)dy = 1$.

Eventually for $0 < T \le T_0 := T_0(\textbf{[HT]}, K)$ small enough, the following diagonal lower bound holds $\forall 0 \le t < s \le T$, $\forall (x,y) \in (\mathbb{R}^2)^2$ s.t.:

$$|(\mathbb{T}^{\alpha}_{s-t})^{-1}(y - R_{s,t}x)| \le K,\ p(t,s,x,y) \ge C^{-1}_{2.2}\det(\mathbb{T}^{\alpha}_{s-t})^{-1}. \tag{2.8}$$

Under the current assumptions, Theorem 2.1 is proved following the lines of [BP09] and [Men11]. In the Gaussian framework, those assumptions are sufficient to derive homogeneous two-sided multi-scale Gaussian bounds, see [DM10]. However, in the current context, we only managed to obtain the expected upper bound up to a logarithmic factor and a diagonal lower bound for $d = 1$ and $n = 2$ for a *tempered* driving noise and $\sigma(t,x) = \sigma(t,x^2)$. This is mainly due to a lack of integrability of the stable process and the rediagonalization phenomenon which becomes really delicate to handle in the degenerate case. Precisely, the parametrix technique consists in applying the difference of two non-local generators of the form (1.3) to the density of some process which is meant to locally behave as (1.1) and for which estimates are available. Such a process is known as the *parametrix* or *proxy*. The density of the stable nd-dimensional process we will use in the degenerate setting as *parametrix* will have decays of order $d+1+\alpha$ in the *large deviation* regime. It is indeed delicate to use other bounds than the worst one in a global approach like the parametrix. Let us mention that this is not the decay of a rotationally invariant stable process in dimension nd (which would be $nd+\alpha$) except if $n = 2, d = 1$. Observe now from (1.3), (2.2) that we have a dimension mismatch between the decays of the densities of the parametrix and those of the jump measure ν, which are in $d+\alpha$. We recall that the large jumps can lead to integrate the density on a set on which it is in its diagonal regime, when applying the non-local generator to the density. This is what we actually call *rediagonalization* and leads in our degenerate framework to additional time-singularities in the parametrix kernel. We manage to handle those singularities when σ depends on the fast component, yielding a better smoothing property in time, see Section 6.

Observe that, in the non-degenerate context, the decays of the rotationally invariant stable densities and the jump measure in (2.2) correspond. This allows Kolokoltsov [Kol00b] to successfully give two sided bounds for the density of the SDE which are homogeneous to those of the rotationally stable case provided the density of the spectral measure is positive. The technical reasons leading to the restriction of Theorem 2.2 will be discussed thoroughly in the dedicated sections (see Sections 3.3 and 6). Let us mention that the above results could be extended to the case of a d-dimensional non-degenerate SDE driven by a tempered stable process and the integral of **one** of its components. We emphasize as well, that our estimates still hold if we had a non-linear bounded drift in the dynamics of X^1 if $\alpha > 1$ (see Remark 5.5). We conclude this paragraph saying that the uniqueness of the martingale problem and the estimates of Section 6 allow to extend in the non-degenerate case, the stable two-sided *Aronson like* estimates of [Kol00b] for Hölder coefficients.

Constants and usual notations:

- The capital letter C will denote a constant whose value may change from line to line, and can depend on the hypotheses **[H]**. Other dependencies (in particular in time), will be specified, using explicit under scripts.

- We will often use the notation $\asymp$ to express equivalence between functions. If f and g are two real valued nonnegative functions, we denote $f(x) \asymp g(x),\ x \in I \subset \mathbb{R}^p,\ p \in \mathbb{N}$, when there exists a constant $C \geq 1$, possibly depending on **[H]**, I s.t. $C^{-1}f(x) \leq g(x) \leq Cf(x),\ \forall x \in I$.

- For $x = (x_1, \cdots, x_{nd}) \in \mathbb{R}^{nd}$ and for all $k \in [\![1, n]\!]$, we use the notations: $x^k := (x_{(k-1)d+1}, \cdots, x_{kd}) \in \mathbb{R}^d$. Accordingly, $x = (x^1, \cdots, x^n)$. Also $x^{2:n} := (x^2, \cdots, x^n)$.

From now on, we assume **[H]** to be in force, specifying when needed, which results are valid under **[HT]** only.

3 Continuity techniques : the Frozen equation and the parametrix series.

For density estimates, a continuity technique consists in considering a simpler equation as *proxy* model for the initial equation. The *proxy* will be significant if it achieves two properties:

- It admits an explicit density or a density that is well estimated.
- The difference between the density of the initial SDE and the one of the proxy can be well controlled.

For the last point a usual strategy consists in expressing the difference of the densities through the difference of the generators of the two SDEs, using Kolmogorov's equations. This approach is known as the *parametrix* method. In the current work, we will use the procedure developed by Mc Kean and Singer [MKS67], which turns out to be well-suited to handle coefficients with mild smoothness properties.

We first introduce the *proxy* model in Section 3.1, and give some associated density bounds. We then analyze in Sections 3.2, 3.3 how this choice can formally lead through a parametrix expansion to a density estimate, exploiting some suitable regularization properties in time. These arguments can be made rigorous provided that the initial SDE admits a Feller transition function. The uniqueness of the martingale problem will actually give this property.

3.1 The Frozen Process.

In this section, we give results that hold in any dimension d, and for any fixed number of oscillators n. Let $T > 0$ (arbitrary deterministic time) and $y \in \mathbb{R}^{nd}$ (final freezing point) be given. Heuristically, y is the point where we want to estimate the density of (1.1) at time T provided it exists. We introduce the *frozen* process as follows:

$$d\tilde{X}_s^{T,y} = A_s \tilde{X}_s^{T,y} ds + B\sigma(s, R_{s,T}y)dZ_s. \tag{3.1}$$

In this equation, $R_{s,T}y$ is the resolvent of the associated deterministic equation, i.e. it satisfies $\frac{d}{ds}R_{s,T} = A_s R_{s,T}$, with $R_{T,T} = I_{nd\times nd}$ in $\mathbb{R}^{nd} \otimes \mathbb{R}^{nd}$. Let us emphasize that the previous choice can seem awkward at first sight. Indeed, a very natural approach for a proxy model would consist in freezing the diffusion coefficient at the terminal point, see e.g. Kolokoltsov [Kol00b]. In our current weak Hörmander setting we need to take into account the backward transport of the final point by the deterministic differential system. This particular choice is actually imposed by the natural metric appearing in the density of the frozen process, see Proposition 3.3. This allows the comparison of the singular parts of the generators of (1.1) and (3.1) applied to the frozen density, see Proposition 3.6 and Lemma 3.10.

Proposition 3.1. *Fix $(t,x) \in [0,T]\times\mathbb{R}^{nd}$. The unique solution of (3.1) starting from x at time t writes:*

$$\tilde{X}_s^{t,x,T,y} = R_{s,t}x + \int_t^s R_{s,u}B\sigma(u, R_{u,T}y)dZ_u. \tag{3.2}$$

Proof. Equation (3.1) is a linear SDE, with deterministic diffusion coefficient. As such, it admits a unique strong solution. The representation (3.2) follows from Itô's formula. □

Introduce for all $u \in \mathbb{R}^+$, the diagonal time scale matrixes:

$$\mathbb{T}_u^\alpha = \begin{pmatrix} u^{\frac{1}{\alpha}}I_{d\times d} & & & 0 \\ 0 & u^{1+\frac{1}{\alpha}}I_{d\times d} & & 0 \\ & & \ddots & \\ 0 & & & u^{n-1+\frac{1}{\alpha}}I_{d\times d} \end{pmatrix}, \tag{3.3}$$

$$\mathbb{M}_u = u^{-\frac{1}{\alpha}}\mathbb{T}_u^\alpha = \begin{pmatrix} I_{d\times d} & & & 0 \\ 0 & uI_{d\times d} & & 0 \\ & & \ddots & \\ 0 & & & u^{n-1}I_{d\times d} \end{pmatrix}.$$

These extend the definitions of Theorem 2.2 for $n = 2$. The entries of the matrix $\mathbb{T}_u^\alpha$ correspond to the *intrinsic* time scales of the iterated integrals of a stable process with index α observed at time u. They reflect the multi-scale behavior of our system. The matrix $\mathbb{M}_u$ appears in the tempered case. We first give an expression of the density of $\tilde{X}_s^{t,x,T,y}$ in terms of its inverse Fourier transform. We refer to Section 5.2 for the proof of this result.

Proposition 3.2. *The frozen process $(\tilde{X}_s^{t,x,T,y})_{s\geq t}$ has for all $s > t$ a density w.r.t. the Lebesgue measure, that is:*

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

For $0 < T - t \leq T_0 := T_0(\textbf{[H]}) \leq 1$ we have:

$$\begin{aligned}\tilde{p}_\alpha^{T,y}(t,s,x,z) &= \frac{\det(\mathbb{M}_{s-t})^{-1}}{(2\pi)^{nd}} \\ &\times \int_{\mathbb{R}^{nd}} e^{-i\langle q, (\mathbb{M}_{s-t}^\alpha)^{-1}(z - R_{s,t}x)\rangle} \exp\left(-(s-t)\int_{\mathbb{R}^{nd}}\Big(1-\cos\langle q,\xi\rangle\Big)\nu_S(d\xi)\right)dq,\end{aligned} \tag{3.4}$$

where $\nu_S := \nu_S(t,T,s,y)$ is a symmetric measure on $\mathbb{R}^{nd}$ s.t. uniformly in $s \in (t, t+T_0]$ for all $A \subset \mathbb{R}^{nd}$:

$$\nu_S(A) \leq \int_{\mathbb{R}^+}\frac{d\rho}{\rho^{1+\alpha}}\int_{S^{nd-1}} \mathbf{1}_A(s\xi)g(c\rho)\bar{\mu}(d\xi),$$

*with $\bar{\mu}$ satisfying **[H-4]** and* $\dim(\mathrm{supp}(\bar{\mu})) = d$. *In the stable case, i.e.* $g = 1$, *we have the equality in the above equation, so that ν_S indeed corresponds to a stable Lévy measure.*

Remark 3.1. The above proposition is important in that it shows in the stable case **[HS]** why the density of a d-dimensional stable process with index $\alpha \in (0, 2)$ and its $n-1$ iterated integrals actually behaves as the density of an nd-dimensional *multi-scale* stable process, where the various scales are read through the matrix $\mathbb{T}^\alpha$. Also, the fact that the associated spectral measure is either non equivalent or singular w.r.t. the Lebesgue measure of S^{nd-1} leads to consider delicate asymptotics for the tails of the density which yields the dimension constraints in Theorem 2.1 and the restrictions of Theorem 2.2.

From the previous remark and the dimension of the support of $\bar{\mu}$ in Proposition 3.2 we derive from points *i)* and *iii)* in Theorem 1.1 in Watanabe [Wat07] the following estimate in the stable case.

Proposition 3.3 (Density Estimates for the frozen process under **[HS]**). *Fix $T > 0$, a threshold $K > 0$ and $y \in \mathbb{R}^{nd}$. For all $(t, x) \in [0, T) \times \mathbb{R}^{nd}$, the density $\tilde{p}_\alpha^{T,y}(t, s, x, z)$ of the frozen process $(\tilde{X}_s^{t,x,T,y})_{s\in(t,T]}$ in* (3.2) *satisfies the following estimates. There exists $C_{3.3} := C_{3.3}$ (**H-2,H-3,H-4**,K)≥ 1, s.t. for all $0 \leq t < s \leq T$, $(x, z) \in (\mathbb{R}^{nd})^2$:*

$$C_{3.3}^{-1}\underline{p}_\alpha(t, s, x, z) \leq \tilde{p}_\alpha^{T,y}(t, s, x, z) \leq C_{3.3}\bar{p}_\alpha(t, s, x, z), \tag{3.5}$$

where we write:

$$\bar{p}_\alpha(t, s, x, z) = C_\alpha \frac{\det(\mathbb{T}_{s-t}^\alpha)^{-1}}{\{K \vee |(\mathbb{T}_{s-t}^\alpha)^{-1}(z - R_{s,t}x)|\}^{d+1+\alpha}}, \tag{3.6}$$

and also

$$\underline{p}_\alpha(t, s, x, z) = C_\alpha^{-1} \frac{\det(\mathbb{T}_{s-t}^\alpha)^{-1}}{\{K \vee |(\mathbb{T}_{s-t}^\alpha)^{-1}(z - R_{s,t}x)|\}^{nd(1+\alpha)}},\ C_\alpha := C_\alpha(\textbf{[H]}) \geq 1.$$

We refer to Section 5.2 for the proof of this result. Observe that $\bar{p}_\alpha(t, s, x, .)$ can be identified with a probability density only under the condition $d(1 - n) + 1 + \alpha > 0$ appearing in Theorem 2.1. Roughly speaking the upper-bound in (3.5) is the worst possible considering the underlying dimension of the support of the spectral measure in S^{nd-1}, which is here d. On the other hand, the lower bound corresponds to the highest possible concentration of a spectral measure on S^{nd-1} satisfying **[H-4]**, see again Section 5.2. This control would for instance correspond to the product at a given point of the one-dimensional stable asymptotics in each direction.

From Proposition 3.2 and Theorem 1 in Sztonyk [Szt10] we also derive the following result in the tempered case.

Proposition 3.4 (Density Estimates for the frozen process under **[HT]**). *Fix $T > 0$, a threshold $K > 0$ and $y \in \mathbb{R}^{nd}$. For all $(t, x) \in [0, T) \times \mathbb{R}^{nd}$, the density $\tilde{p}_\alpha^{T,y}(t, s, x, z)$ of the frozen process $(\tilde{X}_s^{t,x,T,y})_{s\in(t,T]}$ in (3.2) satisfies the following estimates. There exists $C_{3.4} := C_{3.4}$ (**H-2,H-3,H-4**,K)≥ 1, s.t. for all $0 \leq t < s \leq T$, $(x, z) \in (\mathbb{R}^{nd})^2$:*

$$\tilde{p}_\alpha^{T,y}(t, s, x, z) \leq C_{3.4}\bar{p}_\alpha(t, s, x, z), \tag{3.7}$$

where:

$$\bar{p}_\alpha(t, s, x, y) = C_\alpha \frac{\det(\mathbb{T}_{s-t}^\alpha)^{-1}}{\{K \vee |(\mathbb{T}_{s-t}^\alpha)^{-1}(y - R_{s,t}x)|\}^{d+1+\alpha}} \theta(|(\mathbb{M}_{s-t})^{-1}(y - R_{s,t}x)|). \tag{3.8}$$

As a corollary, we have the following important property.

Corollary 3.5 (**“Semigroup" property**). *Under **[H]**, when $d = 1, n = 2$, which is the only case for which $nd+\alpha = d+1+\alpha$ so that $\bar{p}_\alpha$ can be identified with the density of a, possibly tempered, multi-scale stable process in dimension nd whose spectral measure is absolutely continuous with respect to the Lebesgue measure of S^{nd-1}, there exists $C_{3.5} := C_{3.5}$ (**H-2,H-3,H-4**,K) ≥ 1 s.t. for all $0 \leq t < \tau < s$, $(x, y) \in (\mathbb{R}^{nd})^2$:*

$$\int_{\mathbb{R}^{nd}} \bar{p}_\alpha(t, \tau, x, z)\bar{p}_\alpha(\tau, s, z, y)dz \leq C_{3.5}\bar{p}_\alpha(t, s, x, y).$$

The above control is important since it allows to give estimates on the convolution of the frozen densities with possible different freezing points. Namely, for all $T_1, T_2 > 0$, $y_1, y_2 \in \mathbb{R}^{nd}$, for all $t < \tau < s$ and $x, y \in \mathbb{R}^{nd}$:

$$\int_{\mathbb{R}^{nd}} \tilde{p}_\alpha^{T_1,y_1}(t, \tau, x, z)\tilde{p}_\alpha^{T_2,y_2}(\tau, s, z, y)dz \leq C_{3.5}\bar{p}_\alpha(t, s, x, y). \tag{3.9}$$

3.2 The Parametrix Series.

We assume here that the generator $(L_t)_{t\geq 0}$ of (1.1) generates a two-parameter Feller semigroup $(P_{t,s})_{0\leq t\leq s}$. Using the Chapman-Kolmogorov equations satisfied by the semigroup and the pointwise Kolmogorov equations for the *proxy* model, we derive a formal representation of the semigroup in terms of a series, involving the difference of the generators of the initial and frozen processes. Let L_t (already defined in (1.3)) and $\tilde{L}_t^{T,y}$ denote the generators of $X^{t,x}$ and $\tilde{X}^{t,x,T,y}$ at time t respectively. For $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$, from (2.4) (setting $g = 1$ in the stable case), we have for all $x \in \mathbb{R}^{nd}$:

$$\begin{aligned} L_t\varphi(x) &= \langle \nabla\varphi(x), A_t x\rangle \qquad (3.10)\\ &+ \int_{\mathbb{R}^d} \left(\varphi(x + B\sigma(t, x)z) - \varphi(x) - \frac{\langle \nabla\varphi(x), B\sigma(t, x)z\rangle}{1 + |z|^2}\right) g(|z|)\nu(dz),\\ \tilde{L}_t^{T,y}\varphi(x) &= \langle \nabla\varphi(x), A_t x\rangle \qquad (3.11)\\ &+ \int_{\mathbb{R}^d} \left(\varphi(x + B\sigma(t, R_{t,T}y)z) - \varphi(x) - \frac{\langle \nabla\varphi(x), B\sigma(t, R_{t,T}y)z\rangle}{1 + |z|^2}\right) g(|z|)\nu(dz). \end{aligned}$$

Observe that for $\tilde{X}_s^{t,x,T,y}$ defined in (3.2), its density $\tilde{p}_\alpha^{T,y}(t,s,x,\cdot)$ exists and is smooth under **[H]** for $s>t$ (see Proposition 3.2 above).

Proposition 3.6. *Suppose that there exists a unique weak solution $(X_s^{t,x})_{0\le t\le s}$ to (1.1) which has a two-parameter Feller semigroup $(P_{t,s})_{0\le t\le s}$. We have the following formal representation. For all $0\le t<T$, $(x,y)\in(\mathbb{R}^{nd})^2$ and any bounded measurable $f:\mathbb{R}^{nd}\to\mathbb{R}$:*

$$P_{t,T}f(x)=\mathbb{E}[f(X_T)|X_t=x]=\int_{\mathbb{R}^{nd}}\left(\sum_{r=0}^{+\infty}(\tilde{p}_\alpha\otimes H^{(r)})(t,T,x,y)\right)f(y)dy, \qquad (3.12)$$

where H is the parametrix kernel:

$$\forall 0\le t<T,\ (x,y)\in(\mathbb{R}^{nd})^2,\ H(t,T,x,y):=(L_t-\tilde{L}_t^{T,y})\tilde{p}_\alpha(t,T,x,y). \qquad (3.13)$$

In equations (3.12), (3.13), we denote for all $0\le t<u\le T$, $(x,z)\in(\mathbb{R}^{nd})^2$, $\tilde{p}_\alpha(t,u,x,z):=\tilde{p}_\alpha^{u,z}(t,u,x,z)$, i.e. we omit the superscript when the freezing terminal time and point are those where the density is considered. Also, the notation $\otimes$ stands for the time space convolution:

$$f\otimes h(t,T,x,y)=\int_t^T du\int_{\mathbb{R}^{nd}}dzf(t,u,x,z)h(u,T,z,y).$$

Besides, $H^{(0)}=I$ 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 :

$$\forall 0\le t<T,\ (x,y)\in(\mathbb{R}^{nd})^2,\ p(t,T,x,y)=\sum_{r=0}^{+\infty}(\tilde{p}_\alpha\otimes H^{(r)})(t,T,x,y). \qquad (3.14)$$

Proof. Let us first emphasize that the density $\tilde{p}_\alpha^{T,y}(t,s,x,z)$ of $\tilde{X}_s^{t,x,T,y}$ at point z solves the Kolmogorov *backward* equation:

$$\frac{\partial\tilde{p}_\alpha^{T,y}}{\partial t}(t,s,x,z)=-\tilde{L}_t^{T,y}\tilde{p}_\alpha^{T,y}(t,s,x,z),$$
$$\text{for all } t<s,\ (x,z)\in\mathbb{R}^{nd}\times\mathbb{R}^{nd},\ \lim_{t\uparrow s}\tilde{p}_\alpha^{T,y}(t,s,\cdot,z)=\delta_z(\cdot). \qquad (3.15)$$

Here, $\tilde{L}_t^{T,y}$ acts on the variable x. Let us now introduce the family of operators $(\tilde{P}_{t,s})_{0\le t\le s}$. For $0\le t\le s$ and any bounded measurable function $f:\mathbb{R}^{nd}\to\mathbb{R}$:

$$\tilde{P}_{t,s}f(x):=\int_{\mathbb{R}^{nd}}\tilde{p}_\alpha(t,T,x,y)f(y)dy:=\int_{\mathbb{R}^{nd}}\tilde{p}_\alpha^{T,y}(t,T,x,y)f(y)dy. \qquad (3.16)$$

Observe that the family $(\tilde{P}_{t,s})_{0\le t\le s}$ is not a two-parameter semigroup. Anyhow, we can still establish, see Lemma 4.1, that for a continuous f:

$$\lim_{s\to t} \tilde{P}_{s,t} f(x) = f(x). \tag{3.17}$$

This convergence is not a direct consequence of the bounded convergence theorem since the freezing parameter is also the integration variable.

The boundary condition (3.17) and the Feller property yield:

$$(P_{t,T} - \tilde{P}_{t,T})f(x) = \int_t^T du \frac{\partial}{\partial u}\Big\{ P_{t,u}(\tilde{P}_{u,T} f(x))\Big\}.$$

Computing the derivative under the integral leads to:

$$(P_{t,T} - \tilde{P}_{t,T})f(x) = \int_t^T du \Big\{ \partial_u P_{t,u}(\tilde{P}_{u,T} f(x)) + P_{t,u}(\partial_u(\tilde{P}_{u,T} f(x)))\Big\}.$$

Using the Kolmogorov equation (3.15) and the Chapman-Kolmogorov relation $\partial_u P_{t,u}\varphi(x) = P_{t,u}(L_u\varphi(x)),\ \forall \varphi \in C_b^2(\mathbb{R}^{nd},\mathbb{R})$ we get:

$$(P_{t,T} - \tilde{P}_{t,T})f(x) = \int_t^T du P_{t,u}\Big(L_u \tilde{P}_{u,T} f\Big)(x) - P_{t,u}\left(\int_{\mathbb{R}^{nd}} f(y)\tilde{L}_u^{T,y}\tilde{p}_\alpha(u,T,\cdot,y)dy\right)(x).$$

Define now the operator:

$$\mathcal{H}_{u,T}\varphi(z) := \int_{\mathbb{R}^{nd}} \varphi(y)(L_u - \tilde{L}_u^{T,y})\tilde{p}_\alpha(u,T,z,y)dy = \int_{\mathbb{R}^{nd}} \varphi(y) H(u,T,z,y)dy. \tag{3.18}$$

We can thus rewrite:

$$P_{t,T} f(x) = \tilde{P}_{t,T} f(x) + \int_t^T P_{t,u}\big(\mathcal{H}_{u,T}(f)\big)(x)du.$$

The idea is now to reproduce this procedure for $P_{t,u}$ applied to $\mathcal{H}_{u,T}(f)$. This recursively yields the formal representation:

$$P_{t,T} f(x) = \tilde{P}_{t,T} f(x) + \sum_{r\ge 1}\int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-1}} du_r \tilde{P}_{t,u_r}\big(\mathcal{H}_{u_r,u_{r-1}}\circ\dots\circ\mathcal{H}_{u_1,T}\big)(f)(x).$$

Equation (3.12) then formally follows from the following identification. For all $r \in \mathbb{N}^*$:

$$\begin{aligned}\int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-1}} du_r \tilde{P}_{t,u_r}\big(\mathcal{H}_{u_r,u_{r-1}}\circ\dots\circ\mathcal{H}_{u_1,T}\big)(f)(x)du \\ = \int_{\mathbb{R}^{nd}} f(y)\tilde{p}_\alpha \otimes H^{(r)}(t,T,x,y)dy.\end{aligned}$$

We can proceed by immediate induction:

$$
\begin{aligned}
& \int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-1}} du_r \tilde{P}_{t,u_r}\big(\mathcal{H}_{u_r,u_{r-1}} \circ \dots \circ \mathcal{H}_{u_1,T}\big)(f)(x)du \\
= & \int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-1}} du_r \int_{\mathbb{R}^{nd}} dz \\
& \quad \mathcal{H}_{u_r,u_{r-1}} \circ \dots \circ \mathcal{H}_{u_1,T}(f)(z)\tilde{p}_\alpha(t,u_r,x,z) \qquad (3.19) \\
\overset{(3.18)}{=} & \int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-1}} du_r \int_{\mathbb{R}^{nd}} dz \int_{\mathbb{R}^{nd}} dy \\
& \quad \mathcal{H}_{u_{r-1},u_{r-2}} \circ \dots \circ \mathcal{H}_{u_1,T}(f)(y)H(u_r,u_{r-1},z,y)\tilde{p}_\alpha(t,u_r,x,z) \\
= & \int_t^T du_1 \int_t^{u_1} du_2 \dots \int_t^{u_{r-2}} du_{r-1} \int_{\mathbb{R}^{nd}} dy \\
& \quad \mathcal{H}_{u_{r-1},u_{r-2}} \circ \dots \circ \mathcal{H}_{u_1,T}(f)(y)\tilde{p}_\alpha \otimes H(t,u_{r-1},x,y).
\end{aligned}
$$

Thus, we can iterate the procedure from (3.19) with $\tilde{p}_\alpha \otimes H$ instead of $\tilde{p}_\alpha$.

□

Observe that in order to make the identification above, we have exchanged various integrals. Hence, so far the representation (3.14) is *formal*. It will become rigorous provided that we manage to show the convergence of the series and get integrable bounds on its sum.To achieve these points, one needs to give precise bounds on the iterated time-space convolutions appearing in the series. Such controls are stated in Section 3.3 and proved in Section 6 below.

3.3 Controls on the iterated kernels.

From now on, we assume w.l.o.g. that $0 < T \le T_0 := T_0(\textbf{[H]}) \le 1$. The choice of T_0 depends on the constants appearing in **[H]** and will be clear from the proof of Lemma 5.1. Theorems 2.1 and 2.2 can anyhow be obtained for an arbitrary fixed finite $T > 0$, from the results for T sufficiently small. Indeed, the uniqueness of the martingale problem simply follows from the Markov property whereas the upper density estimate stems from the *semigroup* property of $\bar{p}_\alpha$ (see Corollary 3.5 and Lemma 3.12 for the convolutions involving the logarithmic correction). From now on, we consider that the threshold $K > 0$ appearing in Proposition 3.3 is fixed.

We first give pointwise results on the convolution kernel, that hold in any dimension d, and for any number of oscillators n.

Lemma 3.7 (**Control of the kernel**). *Fix $K, \delta > 0$, $\exists C_{3.7} := C_{3.7}(\textbf{[H]}, K, \delta) > 0$ s.t. for all $T \in (0, T_0]$ and $(t,x,y) \in [0,T) \times (\mathbb{R}^{nd})^2$:*

$$
|H(t,T,x,y)| \le C_{3.7} \frac{\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}}{T-t} \left\{\bar{p}_\alpha(t,T,x,y) + \check{p}_\alpha(t,T,x,y)\right\}, \qquad (3.20)
$$

where $\bar p_\alpha$ is as in (3.6) *in the stable case **[HS]** and as in* (3.8) *in the tempered case **[HT]**. Also,*

$$\breve p_\alpha(t,T,x,y) = \frac{\mathbf{1}_{|(x-R_{t,T}y)^1|/(T-t)^{1/\alpha}\asymp|(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)|\ge K}}{(T-t)^{d/\alpha}(1+\frac{|(x-R_{t,T}y)^1|}{(T-t)^{1/\alpha}})^{d+\alpha}}$$

$$\times\frac{1}{(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}(1+|((\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)^{2:n}|)^{1+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(x-R_{t,T}y)|),$$

*recalling that under **[HS]**, $g(r)=1, r>0$.*

The contribution in $\breve p_\alpha$ comes from the *rediagonalization* phenomenon which is specific to the degenerate, non-local case and only appears when the rescaled first (slow) component is equivalent to the energy $|(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)|$. Observe that if $|(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)|\le K$, diagonal regime, both contributions $\bar p$ and $\breve p$ can be upper-bounded by $(T-t)^{-nd/\alpha+n(n-1)d/2}$. In the off-diagonal case, we also have that if there exists $i\in[\![2,n]\!]$ s.t. $|((\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y))^i|\asymp|((\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y))^1|$ then $\breve p_\alpha(t,T,x,y)\le\bar p_\alpha(t,T,x,y)$.

Once integrated in space, under the dimension constraints of Theorem 2.1, this pointwise estimate yields the following *smoothing* property in time.

Lemma 3.8. *Assume that $d(1-n)+1+\alpha>0$. Then, there exists $C_{3.8}:=C_{3.8}($**[H]**$,K)$ and $\omega:=\omega(d,n,\alpha)>0$ s.t. for all $T\in(0,T_0],(x,y)\in(\mathbb{R}^{nd})^2$, $\tau\in[t,T)$, we have the estimate*

$$\begin{aligned}\int_{\mathbb{R}^{nd}}\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge1)}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz &\le C_{3.8}(T-\tau)^\omega, &(3.21)\\ \int_{\mathbb{R}^{nd}}\delta\wedge|z-R_{\tau,t}x|^{\eta(\alpha\wedge1)}\bar p_\alpha(t,\tau,x,z)dz &\le C_{3.8}(\tau-t)^\omega. &(3.22)\end{aligned}$$

Also, when $d=1,n=2$ one has the following better smoothing property for the fast variable:

$$\int_{\mathbb{R}^{nd}}\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz \le C_{3.8}(T-\tau)^{\tilde\omega}, \qquad (3.23)$$

$$\int_{\mathbb{R}^{nd}}\delta\wedge\Big((\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|\Big)^{\eta(\alpha\wedge1)}\bar p_\alpha(t,\tau,x,z)dz \le C_{3.8}(\tau-t)^{\tilde\omega}, \qquad (3.24)$$

with $\tilde\omega=(1+1/\alpha)\eta(\alpha\wedge1)$.

The proof of these results will be given in Section 5.3 and Appendix 8.

Remark 3.2. We can now justify from this Lemma our previous choice for the *proxy* model. Indeed, the contributions $|z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)}$, $|z - R_{\tau,t}x|^{\eta(\alpha\wedge 1)}$ come from the difference of the generators and turn out to be compatible, up to using the Lipschitz property of the flow, with the bounds appearing in Proposition 3.3 for the frozen density. This is what gives this smoothing property and thus allows to get rid of the diagonal singularities coming from the bound (3.20).

Remark 3.3. The l.h.s. of equations (3.24), (3.24) naturally appear in the case $\sigma(t,x) = \sigma(t,x^2)$ which is the one considered for the density estimates in Theorem 2.2. Intuitively, the higher smoothing effect in this case permits to compensate the difficulties arising from the rediagonalization in the degenerate case.

When dealing with convolutions of the kernel and the frozen density we restrict to the case $d = 1, n = 2$ for which we have the *semigroup* property, which is important to handle the off-diagonal regimes. In this framework, the technical computations in Section 6, based on the previous controls on the kernel H, yield the following bound for the first step of the parametrix procedure.

Lemma 3.9. *There exist $C_{3.9} := C_{3.9}(\textbf{[H]}, K)$, $\omega := \omega(\textbf{[H]}) \in (0,1]$ s.t. for all $T \in (0, T_0]$ and $(t,x,y) \in [0,T) \times (\mathbb{R}^{nd})^2$:*

$$\begin{aligned}
|\tilde{p}_\alpha \otimes H(t,T,x,y)| \ \le\ & C_{3.9}\Big(\bar{p}_\alpha(t,T,x,y)\Big((T-t)^\omega \\
& +\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\big(1 + \log[K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|]\big)\Big) \\
& +\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\check{p}(t,T,x,y)\}\Big),
\end{aligned}$$

where

$$\begin{aligned}
&\check{p}(t,T,x,y) := \\
&\inf_{\tau\in[t,T]} \frac{\mathbf{1}_{\frac{(R_{\tau,t}x - R_{\tau,T}y)^1}{(T-t)^{1/\alpha}} \asymp |(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x - R_{\tau,T}y)|}}{(T-t)^{1/\alpha}(1 + |(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x - R_{\tau,T}y)|)^{1+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x - R_{\tau,T}y)|) \\
&\times \frac{1}{(T-t)^{1+1/\alpha}(1 + \frac{|(R_{\tau,t}x - R_{\tau,T}y)^2|}{(T-t)^{1+\frac{1}{\alpha}}})^{1+\alpha}}.
\end{aligned} \tag{3.25}$$

The contribution in $\check{p}$, comes from the bad rediagonalization which is intrinsic to the degenerate case. It first generates a loss of concentration in the estimate, which leads us to temper the driving noise. It also turns out to be very difficult to handle in the iterated convolutions of the kernel. Up to the end of section we thus restrict under **[HT]** to the case $d = 1, n = 2$ and $\sigma(t,x) := \sigma(t,x^2)$, for which we have been able to refine the above results and to derive the convergence of (3.14). This restriction will be discussed thoroughly in Section 6.

Lemma 3.10 (**Control of the iterated kernels**). *Assume under* ***[HT]*** *that* $d = 1, n = 2, \sigma(t,x) = \sigma(t,x^2)$ *and* $1 \geq \eta > ((\alpha \wedge 1)(1+\alpha))^{-1}$. *Then there exist* $C_{3.10} := C_{3.10}($***[HT]***$, K)$, $\omega := \omega($***[HT]***$) \in (0,1]$ *s.t. for all* $T \leq T_0$ *and* $(t,x,y) \in [0,T) \times (\mathbb{R}^2)^2$*:*

$$
\begin{aligned}
|\tilde{p}_\alpha \otimes H(t,T,x,y)| &\leq C_{3.10}\Big((T-t)^\omega \bar{p}_{\alpha,\Theta}(t,T,x,y) + \bar{q}_{\alpha,\Theta}(t,T,x,y)\Big),\\
|\bar{q}_\alpha \otimes H(t,T,x,y)| &\leq C_{3.10}(T-t)^\omega\Big(\bar{p}_{\alpha,\Theta}(t,T,x,y) + \bar{q}_{\alpha,\Theta}(t,T,x,y)\Big),
\end{aligned}
$$

where we denoted

$$
\begin{aligned}
\bar{q}_{\alpha,\Theta}(t,T,x,y) &= \delta \wedge \{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)}\}\\
&\quad \times\Big\{\bar{p}_{\alpha,\Theta}(t,T,x,y)\Big(1+\log(K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|)\ \Big)\Big\}.
\end{aligned}
$$

Now for all $k \geq 1$,

$$
\begin{aligned}
|\tilde{p}_\alpha \otimes H^{(2k)}(t,T,x,y)| &\leq (4C_{3.10})^{2k}(T-t)^{k\omega}\Big((T-t)^{k\omega}\bar{p}_{\alpha,\Theta}(t,T,x,y)\\
&\quad +(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})(t,T,x,y)\Big),\\
|\tilde{p}_\alpha \otimes H^{(2k+1)}(t,T,x,y)| &\leq (4C_{3.10})^{2k+1}(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}\bar{p}_{\alpha,\Theta}\\
&\quad +(T-t)^\omega(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})+\bar{q}_{\alpha,\Theta}\Big)(t,T,x,y).
\end{aligned}
$$

The above controls allow to derive under the sole assumption **[HT]** an upper bound for the sum of the parametrix series (3.14) in small time.

Proposition 3.11 (**Sum of the parametrix series**). *Under the assumptions of Lemma 3.10, for* T_0 *small enough, there exists* $C_{3.11} := C_{3.11}($***[HT]***$, K, T_0)$ *s.t. for all* $T \in (0,T_0]$ *and* $(t,x,y) \in [0,T) \times (\mathbb{R}^2)^2$*:*

$$
\begin{aligned}
\sum_{r\geq 0}|\tilde{p}_\alpha \otimes H^{(r)}(t,T,x,y)| &\leq C_{3.11}\Big(\bar{p}_{\alpha,\Theta}(t,T,x,y)+\bar{q}_{\alpha,\Theta}(t,T,x,y)\Big),\\
C_{3.11}\det(\mathbb{T}^\alpha_{T-t})^{-1} &\leq \sum_{r\geq 0}\tilde{p}_\alpha \otimes H^{(r)}(t,T,x,y),\ \text{for } |(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)| \leq K.
\end{aligned}
$$

The proofs of Lemmas 3.7 and 3.10 are postponed to Section 6.4. Using those controls on the iterated convolutions, we can prove Proposition 3.11.

Proof. The upper-bound can be readily derived from Lemma 3.10 for T_0 small enough (sum of a geometric series). To get the diagonal lower bound, we first write:

$$
\sum_{k\geq 0}\tilde{p}_\alpha \otimes H^{(k)}(t,T,x,y) = \tilde{p}_\alpha(t,T,x,y) + \left(\sum_{k\geq 0}\tilde{p}_\alpha \otimes H^{(k)}\right)\otimes H(t,T,x,y).
$$

Now, since

$$\sum_{k\geq 0} |\tilde{p}_\alpha \otimes H^{(k)}(t,T,x,y)| \leq C(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})(t,T,x,y),$$

we derive:

$$\left| \left(\sum_{k\geq 0} \tilde{p}_\alpha \otimes H^{(k)} \right) \otimes H(t,T,x,y) \right| \leq C|(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta}) \otimes H(t,T,x,y)|.$$

Using once again the first part of Lemma 3.10, we thus get that

$$\begin{aligned} \left| \left(\sum_{k\geq 0} \bar{p}_\alpha \otimes H^{(k)} \right) \otimes H(t,T,x,y) \right| &\leq C\big\{(T-t)^\omega \bar{p}_{\alpha,\Theta}(t,T,x,y) + \bar{q}_{\alpha,\Theta}(t,T,x,y) \\ &\qquad + (T-t)^\omega (\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})(t,T,x,y)\big\}. \end{aligned}$$

Now, if the global regime is diagonal, i.e. $|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| \leq K$, the logarithm contribution vanishes in $\bar{q}_{\alpha,\Theta}$. Observe also that

$$\begin{aligned} \delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)} &\leq C^{\eta(\alpha\wedge 1)}|R_{T,t}x - y|^{\eta(\alpha\wedge 1)} \\ &\leq C^{\eta(\alpha\wedge 1)}(T-t)^{\eta(1/\alpha\wedge 1)}|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x - y)|^{\eta(\alpha\wedge 1)} \\ &\leq (CK)^{\eta(\alpha\wedge 1)}(T-t)^{\eta(1/\alpha\wedge 1)}. \end{aligned}$$

Hence $\left|\left(\sum_{k\geq 0} \tilde{p}_\alpha \otimes H^{(k)}\right) \otimes H(t,T,x,y)\right| \leq C(T-t)^\omega \det(\mathbb{T}^\alpha_{T-t})^{-1}$. Taking $T-t$ small enough yields the announced bound. □

We conclude anyhow the section stating a Lemma that allows to extend the upper bound in Theorem 2.2 to an arbitrary given fixed time. The arguments for its proof would be similar to those of Lemma 6.3.

Lemma 3.12 (**Semigroup property for** $\bar{q}_{\alpha,\Theta}$)**.** *With the notations of Proposition 3.11, for any* $T \in [0, \bar{T}_0)$*, we have that there exists* $C_{3.12} := C_{3.12}([\mathbf{HT}], \bar{T}_0) \geq 1$ *s.t.* $\forall (x,y) \in \mathbb{R}^{nd}$*,* $\forall n \in \mathbb{N}$*:*

$$\int_{\mathbb{R}^{nd}} \bar{q}_{\alpha,\Theta}(0,nT,x,z)\bar{q}_{\alpha,\Theta}(nT,(n+1)T,z,y)dz \leq C_{3.12}^{n+2}\bar{q}_{\alpha,\Theta}(0,(n+1)T,x,y).$$

Observe now that Theorem 2.1 yields that $(X_t)_{t\geq 0}$, the canonical process of $\mathbb{P}$, admits a Feller transition function. On the other hand, when $d = 1, n = 2$ we have from Proposition 3.11 that the series appearing in equation (3.12) of Proposition 3.6 is absolutely convergent. This allows to derive that the Feller transition is absolutely continuous, which in particular means that the process $(X_t)_{t\geq 0}$ admits for all $t > 0$ a density, satisfying the bounds of Proposition 3.11.

4 Proof of the uniqueness of the Martingale Problem associated with (1.1).

In this section, d and n satisfy the conditions $d(1-n)+1+\alpha>0$. As a corollary to the bounds of Section 3.3, specifically Lemmas 3.7 and 3.8 (controls on the kernel and associated smoothing effect), we prove here Theorem 2.1. The existence of a solution to the martingale problem can be derived by compactness arguments adapting the proof of Theorem 2.2 from [Str75], even though our coefficients are not bounded.

Uniqueness of the Martingale Problem associated with (1.3). Suppose we are given two solutions $\mathbb{P}^1$ and $\mathbb{P}^2$ of the martingale problem associated with $(L_s)_{s\in[t,T]}$, starting in x at time t. We can assume w.l.o.g. that $T\le T_0:=T_0(\mathbf{[H]})$. Define for a bounded Borel function $f:[0,T]\times\mathbb{R}^{nd}\to\mathbb{R}$,

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

where $(X_s)_{s\in[t,T]}$ 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:

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

If $f\in C_0^{1,2}([0,T)\times\mathbb{R}^{nd},\mathbb{R})$, since $(\mathbb{P}^i)_{i\in\{1,2\}}$ both solve the martingale problem, we have:

$$f(t,x)+\mathbb{E}^i\left(\int_t^T(\partial_s+L_s)f(s,X_s)ds\right)=0,\ i\in\{1,2\}. \tag{4.1}$$

For a fixed point $y\in\mathbb{R}^{nd}$ and a given $\varepsilon\ge0$, introduce now for all $f\in C_0^{1,2}([0,T)\times\mathbb{R}^{nd},\mathbb{R})$ the Green function:

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

We recall here that $\tilde p_\alpha^{s+\varepsilon,y}(t,s,x,z)$ stands for the density at time s and point z of the process $\tilde X^{s+\varepsilon,y}$ defined in (3.2) starting from x at time t. In particular, ε can be equal to zero in the previous definition. One now easily checks that:

$$\forall(t,x,z)\in[0,s)\times(\mathbb{R}^{nd})^2,\left(\partial_t+\tilde L_t^{s+\varepsilon,y}\right)\tilde p_\alpha^{s+\varepsilon,y}(t,s,x,z)=0,\lim_{s\downarrow t}\tilde p_\alpha^{s+\varepsilon,y}(t,s,x,\cdot)=\delta_x(.). \tag{4.2}$$

Introducing for all $f\in C_0^{1,2}([0,T)\times\mathbb{R}^{nd},\mathbb{R})$ the quantity:

$$M_{t,x}^{\varepsilon,y}f(t,x)=\int_t^T ds\int_{\mathbb{R}^{nd}}dz\tilde L_t^{s+\varepsilon,y}\tilde p_\alpha^{s+\varepsilon,y}(t,s,x,z)f(s,z), \tag{4.3}$$

we derive from (4.2) and the definition of $G^{\varepsilon,y}$ that the following equality holds:

$$\partial_t G^{\varepsilon,y} f(t,x) + M^{\varepsilon,y}_{t,x} f(t,x) = -f(t,x),\ \forall (t,x) \in [0,T) \times \mathbb{R}^{nd}. \tag{4.4}$$

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

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

Then, by semigroup property, we have:

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

Hence,

$$\begin{aligned}
(\partial_t + L_t)\Psi_\varepsilon(t,x) &= \int_{\mathbb{R}^{nd}} dy (\partial_t + L_t)(G^{\varepsilon,y}\phi^{\varepsilon,y})(t,x) \\
&= \int_{\mathbb{R}^{nd}} dy \{\partial_t G^{\varepsilon,y}\phi^{\varepsilon,y}(t,x) + M^{\varepsilon,y}_{t,x}\phi^{\varepsilon,y}(t,x)\} \\
&\quad + \int_{\mathbb{R}^{nd}} dy \{L_t G^{\varepsilon,y}\phi^{\varepsilon,y}(t,x) - M^{\varepsilon,y}_{t,x}\phi^{\varepsilon,y}(t,x)\} \\
&\overset{(4.4)}{=} -\int_{\mathbb{R}^{nd}} dy \phi^{\varepsilon,y}(t,x) + \int_{\mathbb{R}^{nd}} dy \{L_t G^{\varepsilon,y}\phi^{\varepsilon,y}(t,x) - M^{\varepsilon,y}_{t,x}\phi^{\varepsilon,y}(t,x)\} \\
&= I_1^\varepsilon + I_2^\varepsilon.
\end{aligned}$$

We now need the following lemma whose proof is postponed to the end of Section 5.2.

Lemma 4.1. *For all bounded continuous function $f : \mathbb{R}^{nd} \to \mathbb{R}, x \in \mathbb{R}^{nd}$:*

$$\left| \int_{\mathbb{R}^{nd}} f(y)\tilde{p}_\alpha^{T,y}(t,T,x,y)dy - f(x) \right| \underset{T\downarrow t}{\longrightarrow} 0. \tag{4.5}$$

We emphasize that the above lemma is not a direct consequence of the convergence of the law of the frozen process towards the Dirac mass when $T \downarrow t$. Indeed, the integration parameter is also the freezing parameter which makes things more subtle. Lemma 4.1 yields $I_1^\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} -h(t,x)$. On the other hand, we have the following identity:

$$\begin{aligned}
I_2^\varepsilon &= \int_t^T ds \int_{\mathbb{R}^{nd}} dy (L_t - \tilde{L}_t^{s+\varepsilon})\tilde{p}_\alpha^{s+\varepsilon,y}(t,s+\varepsilon,x,y)h(s,y) \\
&= \int_t^T ds \int_{\mathbb{R}^{nd}} dy H(t,s+\varepsilon,x,y)h(s,y).
\end{aligned}$$

The bound of Lemmas 3.7 and 3.8 now yield:

$$\begin{aligned} |I_2^\varepsilon| &\le C\int_t^T ds\int_{\mathbb{R}^{nd}} dy \frac{\delta\wedge|x-R_{t,s+\varepsilon}y|^{\eta(\alpha\wedge 1)}}{s+\varepsilon-t}(\bar p_\alpha+\breve p_\alpha)(t,s+\varepsilon,x,y)|h(s,y)| \\ &\le C|h|_\infty\int_t^T (s+\varepsilon-t)^{\eta(\frac{1}{\alpha}\wedge 1)-1}ds \le C|h|_\infty[(T-t)\vee\varepsilon]^{\eta(\frac{1}{\alpha}\wedge 1)}. \end{aligned}$$

Hence, we may choose T and ε small enough to obtain

$$|I_2^\varepsilon| \le 1/2|h|_\infty. \tag{4.6}$$

Observe now that (4.1) gives $S^\Delta\Big((\partial_\cdot + L_\cdot)\Psi_\varepsilon\Big)=0$ so that $|S^\Delta(I_1^\varepsilon)|=|S^\Delta(I_2^\varepsilon)|$. From Lemma 4.1 and (4.6), defining $\|S^\Delta\| := \sup_{|f|_\infty\le 1}|S^\Delta f|$, we derive:

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

By a monotone class argument, the previous inequality still holds for bounded Borel functions h compactly supported in $[0,T)\times\mathbb{R}^{nd}$. 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, §7, in [Shi96], see also Chapter 6 in [SV79] and [Str75].

□

5 Proof of the results involving the Frozen process.

Introduce for a given $t>0$ and all $s\ge t$ the process:

$$\Lambda_s := \int_t^s R_{s,u}B\sigma_u dZ_u, \tag{5.1}$$

solving $d\Lambda_s = A_s\Lambda_s ds + B\sigma_s dZ_s,\ Z_t=0$, i.e. Λ_s can be viewed as the process of the iterated integrals of Z weighted by the entries of the resolvent. In (5.1), $(\sigma_u)_{u\ge t}$ is a deterministic $\mathbb{R}^d\otimes\mathbb{R}^d$-valued function s.t. $(\sigma_u\sigma_u^*)_{u\ge t}$ satisfies **[H-2]** (uniform ellipticity). It can be seen from Proposition 3.1 that the frozen process will have a density if and only if Λ does for $s>t$. This is what we establish through Fourier inversion. The structure of the resolvent is crucial: it gives the multi-scale behaviour of the frozen process and allows to prove in Proposition 5.3 that the Fourier transform is integrable. Recalling as well that B stands for the embedding matrix from $\mathbb{R}^d$ into $\mathbb{R}^{nd}$, we observe that only the first d columns of the resolvent are taken into account in (5.1). Reasoning by blocks we rewrite: $R_{s,t} = \begin{pmatrix} R_{s,t}^{1,1} & \cdots & R_{s,t}^{1,n} \\ \vdots & \ddots & \vdots \\ R_{s,t}^{n,1} & \cdots & R_{s,t}^{n,n}\end{pmatrix}$, where the entries $(R_{s,t}^{i,j})_{(i,j)\in[\![1,n]\!]^2}$ belong to $\mathbb{R}^d\otimes\mathbb{R}^d$.

5.1 Analysis of the Resolvent.

Lemma 5.1 (Form of the Resolvent). *Let $0 \le t \le s \le T \le T_0 := T_0(\textbf{[H]}) \le 1$. We can write the first column of the resolvent in the following way:*

$$R_{s,t}^{\cdot,1} = \begin{pmatrix} \bar{R}_{s,t}^1 \\ (s-t)\bar{R}_{s,t}^2 \\ \vdots \\ \frac{(s-t)^{n-1}}{(n-1)!}\bar{R}_{s,t}^n \end{pmatrix}, \tag{5.2}$$

where the $(\bar{R}_{s,t}^i)_{i\in[\![1,n]\!]}$ are non-degenerate and bounded matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$, i.e. $\exists C := C(\textbf{[H]}, T_0)$ s.t. for all $\xi \in S^{d-1}$, $C^{-1} \le |\bar{R}_{s,t}^i \xi| \le C$.

Proof. We are going to prove the result by induction. Let us first consider the case $n = 2$. We have, for $i \in \{1,2\}$:

$$\frac{d}{ds}R_{s,t}^{1,1} = a_s^{1,1}R_{s,t}^{1,1} + a_s^{1,2}R_{s,t}^{2,1},\ \frac{d}{ds}R_{s,t}^{2,1} = a_s^{2,1}R_{s,t}^{1,1} + a_s^{2,2}R_{s,t}^{2,1}.$$

In order to obtain, for $i \in \{1,2\}$, a semi-integrated representation of the entry $R_{s,t}^{i,1}$, we use the resolvent $\Gamma_{u,v}^i$ satisfying $\frac{d}{du}\Gamma_{u,v}^i = a_u^{i,i}\Gamma_{u,v}^i$, $\Gamma_{v,v}^i = I_{d\times d}$. This yields:

$$R_{s,t}^{1,1} = \Gamma_{s,t}^1 + \int_t^s \Gamma_{s,u}^1 a_u^{1,2} R_{u,t}^{2,1} du,\ R_{s,t}^{2,1} = \int_t^s \Gamma_{s,u}^2 \left\{ a_u^{2,1} R_{u,t}^{1,1} \right\} du.$$

Hence for all $0 \le t \le s \le T$:

$$\begin{aligned} R_{s,t}^{1,1} &= \Gamma_{s,t}^1 + \int_t^s \Gamma_{s,u}^1 a_u^{1,2} \left\{ \int_t^u \Gamma_{u,v}^2 \left\{ a_v^{2,1} R_{v,t}^{1,1} \right\} dv \right\} du, \\ |R_{s,t}^{1,1}| &\le C_T(1 + \int_t^s |R_{v,t}^{1,1}|(s-t)dv) \le C_T,\ |R_{s,t}^{2,1}| \le C_T(s-t), \end{aligned}$$

using Gronwall's lemma for the last but one inequality. This in particular yields

$$R_{s,t}^{2,1} = \int_t^s \Gamma_{s,u}^2 a_u^{2,1}(\Gamma_{u,t}^1 + O((u-t)^2))du.$$

From the non-degeneracy of $a^{2,1}$ (Hörmander like assumption **[H-3]**) and the resolvents on a compact set we derive that for T small enough $R_{s,t}^{2,1} = (t-s)\bar{R}_{s,t}^2$ where $\bar{R}_{s,t}^2$ is non-degenerate and bounded. Rewriting $R_{s,t}^{1,1} = \Gamma_{s,t}^1 + O((s-t)^2)$ we derive similarly that $R_{s,t}^{1,1} = \bar{R}_{s,t}^1$, $\bar{R}_{s,t}^1$ being non-degenerate and bounded. This proves (5.2) for $n = 2$. Let us now assume that (5.2) holds for a given $n \ge 2$ and let us prove it for $n+1$.

We first need to introduce some notations to keep track of the induction hypothesis. To this end, we denote by $A_t^{n+1} := A_t$ and $R_{s,t}^{n+1} := R_{s,t}$ the matrices in $\mathbb{R}^{(n+1)d} \otimes \mathbb{R}^{(n+1)d}$

associated with the linear system $\frac{d}{ds}R_{s,t} = A_t R_{s,t},\ R_{t,t} = I_{(n+1)d\times(n+1)d}$. Observe now that:

$$A_t^{n+1} = \left(\begin{array}{c|ccc} a_t^{1,1} & \cdots & \cdots & a_t^{1,n+1} \\ \hline a_t^{2,1} & & & \\ 0 & & & \\ \vdots & & A_t^n & \\ 0 & & & \end{array}\right),$$

where A_t^n is an $\mathbb{R}^{nd}\otimes\mathbb{R}^{nd}$ matrix satisfying **[H-3]**. Hence, denoting by $R_{s,t}^n$ the associated resolvent, i.e. $\frac{d}{ds}R_{s,t}^n = A_s^n R_{s,t}^n,\ R_{t,t}^n = I_{nd\times nd}$, $R_{s,t}^n$ satisfies (5.2) from the induction hypothesis, so that

$$\forall i\in[\![1,n]\!],\ \forall 0\le t\le s\le T,\ (R_{s,t}^n)^{i,1} = \frac{(s-t)^{i-1}}{(i-1)!}\bar{R}_{s,t}^{i,n},$$

where the $(\bar{R}_{s,t}^{i,n})_{i\in[\![1,n]\!]}$ are non-degenerate and bounded. Let us now observe that the differential dynamics of $(R_{s,t}^{n+1})^{2:n+1,1} := \big((R_{s,t}^{n+1})^{2,1},\cdots,(R_{s,t}^{n+1})^{n+1,1}\big)^*$ writes:

$$\frac{d}{ds}(R_{s,t}^{n+1})^{2:n+1,1} = A_s^n(R_{s,t}^{n+1})^{2:n+1,1} + G_{s,t}^{n+1},\ G_{s,t}^{n+1} := \big(a_s^{2,1}(R_{s,t}^{n+1})^{1,1}\ 0_{n\times n}\ \cdots\ 0_{n\times n}\big)^*,$$

where

$$(R_{s,t}^{n+1})^{1,1} = \Gamma_{s,t}^{n+1,1} + \int_t^s \Gamma_{s,u}^{n+1,1}\left\{\sum_{j=2}^{n+1} a_u^{1,j}(R_{u,t}^{n+1})^{j,1}\right\}du, \tag{5.3}$$

$\Gamma^{n+1,1}$ standing for the resolvent associated with $a^{1,1}$. Using now the resolvent $R_{s,t}^n$, the above equation can be integrated. We get:

$$\big(R_{s,t}^{n+1}\big)^{2:n+1,1} = \int_t^s R_{s,u}^n G_{u,t}^{n+1}du. \tag{5.4}$$

From the above representation, using the induction assumption, (5.3) and Gronwall's lemma we derive:

$$|(R_{s,t}^{n+1})^{n+1,1}| \le C_T\int_t^s \frac{(s-u)^{n-1}}{(n-1)!}\left\{1+\int_t^u\sum_{j=2}^n |(R_{v,t}^{n+1})^{j,1}|dv\right\}du.$$

By induction one also derives for all $i\in[\![2,n+1]\!]$:

$$|(R_{s,t}^{n+1})^{i,1}| \le C_T\int_t^s \frac{(s-u)^{i-2}}{(i-2)!}\left\{1+\int_t^u\sum_{j=2}^{i-1} |(R_{v,t}^{n+1})^{j,1}|dv\right\}du,$$

up to modifications of C_T at each step. These controls yield that for all $i \in [\![2,n]\!]$, $0 \le t \le s \le T$:

$$|(R^{n+1}_{s,t})^{i,1}| = O((s-t)^{i-1}). \tag{5.5}$$

Now from (5.4), (5.3) and the induction assumption, we obtain, for all $i \in [\![2,n]\!]$, $0 \le t \le s \le T$:

$$(R^{n+1}_{s,t})^{i,1} = \int_t^s \frac{(s-u)^{i-2}}{(i-2)!} \bar{R}^{i-1,n}_{s,u} a^{2,1}_u \{\Gamma^{n+1,1}_{u,t} + \int_t^u \Gamma^{n+1,1}_{u,v} \{\sum_{j=2}^{n+1} a^{1,j}_v (R^{n+1}_{v,t})^{j,1} dv\}\} du.$$

From the non degeneracy of $\bar{R}^{i-1,n}, a^{2,1}, \Gamma^{n+1,1}$ and (5.5), we can conclude as for the case $n = 2$. □

We can also mention some related analysis, emphasizing various specific time-scales, in Chaleyat-Maurel and Elie p. 255-279 in [A$^+$81], Kolokoltsov [Kol00a] and [DM10]. These procedures were performed to derive small time asymptotics of the covariance matrix of, possibly perturbed, Gaussian hypoelliptic diffusions.

To conclude our analysis of the resolvent $R_{s,t}$, we give here a technical lemma that will be useful for the controls of Section 6.

Lemma 5.2 (**Scaling Lemma**). *Under* ***[H-3]****, the resolvent* $(R_{s,T})_{s\in[t,T]}$*, for* $0 \le t < T$ *associated with the linear system* $\frac{d}{ds}R_{s,T} = A_s R_{s,T}$, $R_{T,T} = I_{nd\times nd}$ *can be written as*

$$R_{s,T} = \mathbb{T}^\alpha_{T-t} \hat{R}^{t,T}_{\frac{s-t}{T-t}} (\mathbb{T}^\alpha_{T-t})^{-1},$$

where $\hat{R}^{t,T}_{\frac{s-t}{T-t}}$ *is non-degenerate and bounded uniformly on* $s \in [t,T]$ *with constants depending on* T.

Proof. The proof of the above statement follows from the structure of the matrix A_t (Assumption **[H-3]**), setting for all $u \in [0,1]$, $\hat{R}^{t,T}_u := (\mathbb{T}^\alpha_{T-t})^{-1} R_{t+u(T-t),T} \mathbb{T}^\alpha_{T-t}$ and differentiating:

$$\begin{aligned} \partial_u \hat{R}^{t,T}_u &= (T-t)(\mathbb{T}^\alpha_{T-t})^{-1} A_{t+u(T-t)} R_{t+u(T-t),T} \mathbb{T}^\alpha_{T-t} \\ &= \left((T-t)(\mathbb{T}^\alpha_{T-t})^{-1} A_{t+u(T-t)} \mathbb{T}^\alpha_{T-t}\right) \hat{R}^{t,T}_u := A^{t,T}_u \hat{R}^{t,T}_u. \end{aligned}$$

□

Remark 5.1. Let us observe that the scaling Lemma already gives the right orders for the entries $(R^{i,1}_{t,s})_{i\in[\![1,n]\!]}$ of the resolvent. However for the analysis of the Fourier transform of Λ, we explicitly need that those entries write in the form of equation (5.2).

5.2 Estimates on the frozen density.

Existence and first estimates.

The main result of this section is the following.

Proposition 5.3 (Existence of the density). *Let $T_0 := T_0(\boldsymbol{[H]})$ be as in Lemma 5.1. The process $(\Lambda_s)_{s\in[t,t+T_0]}$, $t \geq 0$, defined in* (5.1) *has for all $s \in (t, t+T_0]$ a density p_{Λ_s} given for all $z \in \mathbb{R}^{nd}$ by:*

$$p_{\Lambda_s}(z) = \frac{\det(\mathbb{M}_{s-t}^{-1})}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} e^{-i\langle q, \mathbb{M}_{s-t}^{-1} z\rangle} \exp\left(-(s-t)\int_{\mathbb{R}^{nd}} \{1-\cos\langle q,\xi\rangle\}\nu_S(d\xi)\right) dq,$$

where $\nu_S := \nu_S(t, T, s, \sigma)$ is a symmetric measure on S^{nd-1} s.t. uniformly in $s \in (t, t+T_0]$ for all $A \subset \mathbb{R}^{nd}$:

$$\nu_S(A) \leq \int_{\mathbb{R}^+} \frac{d\rho}{\rho^{1+\alpha}} \int_{S^{nd-1}} \boldsymbol{1}_A(\rho\eta) g(c\rho)\bar{\mu}(d\eta), \tag{5.6}$$

*where $\bar{\mu}$ satisfies **[H-4]** and $\dim(\mathrm{supp}(\bar{\mu})) = d$. As a consequence of this representation, we get the following global (diagonal) estimate:*

$$\exists C := C(\boldsymbol{[H]}, T_0),\ \forall s \in (t, t+T_0],\ \forall z \in \mathbb{R}^{nd},\ p_{\Lambda_s}(z) \leq C \det(\mathbb{T}_{s-t}^{\alpha})^{-1}. \tag{5.7}$$

Remark 5.2. The previous result emphasizes that the process $(\Lambda_s)_{s\in[t,t+T_0]}$ can actually be seen as a possibly tempered α-stable symmetric process in dimension nd, with non-degenerate spectral measure, (left) multiplied by the intrinsic scale factor $(\mathbb{M}_{s-t})_{s\in[t,t+T_0]}$.

Proof. The proof is divided into two steps:

- The first step is to compute the Fourier transform.

Starting from the representation (5.1), we write the integral as a limit of its increments. Let $\tau_n := \{(t_i)_{i\in[\![0,n]\!]}; t = t_0 < t_1 < \cdots < t_n = s\}$ be a subdivision of $[t, s]$, whose mesh $|\tau_n| := \max_{i\in[\![0,n-1]\!]} |t_{i+1} - t_i|$ tends to zero when $n \to \infty$. Write now for all $p \in \mathbb{R}^{nd}$:

$$\langle p, \Lambda_s\rangle \quad = \quad \lim_{|\tau_n|\to 0} \sum_{i=0}^{n-1} \langle p, R_{s,t_i} B\sigma_{t_i}(Z_{t_{i+1}} - Z_{t_i})\rangle = \lim_{|\tau_n|\to 0} \sum_{i=0}^{n-1} \langle \sigma_{t_i}^* B^* R_{s,t_i}^* p, (Z_{t_{i+1}} - Z_{t_i})\rangle.$$

Since Z has independent increments, we get from (2.4) and the bounded convergence theorem that $\forall p \in \mathbb{R}^{nd}$:

$$\varphi_{\Lambda_s}(p) := \mathbb{E}(e^{i\langle p,\Lambda_s\rangle}) = \exp\left(\int_t^s \int_{\mathbb{R}^d} \{\cos(\langle p, R_{s,u}B\sigma_u z\rangle) - 1\} g(|z|)\nu(dz)du\right). \tag{5.8}$$

- The second one is to prove its integrability.

Setting $v = (s-u)/(s-t)$ and denoting $u(v) := s - v(s-t)$, the exponent in (5.8) writes:

$$\int_t^s \int_{\mathbb{R}^d} \{\cos(\langle p, R_{s,u} B\sigma_u z\rangle) - 1\} g(|z|)\nu(dz)du =$$
$$(t-s)\int_0^1 \int_{\mathbb{R}^d} \{\cos(\langle p, R^{\cdot,1}_{s,u(v)}\sigma_{u(v)} z\rangle) - 1\} g(|z|)\nu(dz)dv.$$

Now, from Lemma 5.1, we have the identity

$$R^{\cdot,1}_{s,u(v)} = \mathbb{M}_{s-t}\bar{R}_v,$$

setting with a slight abuse of notation $\bar{R}_v = \begin{pmatrix} \bar{R}^1_v \\ v\bar{R}^2_v \\ \vdots \\ \frac{v^{n-1}}{(n-1)!}\bar{R}^n_v \end{pmatrix}$, where the $(\bar{R}^k_v)_{k\in[\![1,n]\!]} \in \mathbb{R}^d \otimes \mathbb{R}^d$ are non-degenerate and bounded. The exponent in (5.8) thus rewrites:

$$\begin{aligned} & \int_t^s \int_{\mathbb{R}^d} \{\cos(\langle p, R_{s,u} B\sigma_u z\rangle) - 1\} g(|z|)\nu(dz)du \qquad (5.9) \\ = \ & (t-s)\int_0^1 \int_{\mathbb{R}^d} \{\cos(\langle \mathbb{M}_{s-t} p, \bar{R}_v\sigma_{u(v)} z\rangle) - 1\} g(|z|)\nu(dz)dv \\ = \ & (t-s)\int_0^1 \int_{\mathbb{R}^d} \{\cos(\langle \sigma^*_{u(v)}\bar{R}^*_v\mathbb{M}_{s-t} p, z\rangle) - 1\} g(|z|)\nu(dz)dv. \end{aligned}$$

Observe now from **[H-4]** and **[T]** (recall that g is C^1 for $\alpha \in (0,1)$ or C^2 for $\alpha \in [1,2)$, in a neighborhood of 0) that:

$$\begin{aligned} & \int_t^s \int_{\mathbb{R}^d} \{\cos(\langle p, R_{s,u} B\sigma_u z\rangle) - 1\} g(|z|)\nu(dz)du \\ = \ & (t-s)\int_0^1 \int_{\mathbb{R}^d} \{\cos(\langle \sigma^*_{u(v)}\bar{R}^*_v\mathbb{M}_{s-t} p, z\rangle) - 1\} g(0)\nu(dz)dv \\ & +(t-s)\int_0^1 \int_{\mathbb{R}^d} \{\cos(\langle \sigma^*_{u(v)}\bar{R}^*_v\mathbb{M}_{s-t} p, z\rangle) - 1\} (g(|z|) - g(0))\nu(dz)dv \\ \le \ & c(t-s)\{-\int_0^1 |\sigma^*_{u(v)}\bar{R}^*_v\mathbb{M}_{s-t} p|^\alpha dv + 1\} \\ \le \ & c(t-s)\{-\int_0^1 |\bar{R}^*_v\mathbb{M}_{s-t} p|^\alpha dv + 1\}, \ c := c(\mathbf{[H]}), \qquad (5.10) \end{aligned}$$

using the uniform ellipticity of σ in assumptions **[H]** for the last inequality. In the above computations, introducing $g(0)$ allows to exploit the explicit expression for the integral of the Fourier exponent of the stable Lévy measure ν and to do Taylor expansions in a neighborhood of 0 for the term $g(|z|) - g(0)$ thanks to the smoothness of g. Now, the lower bound of the following lemma, whose proof is postponed to Subsection 5.2, gives that $\varphi_{\Lambda_s} \in L^1(\mathbb{R}^{nd})$ and therefore yields the existence of the density.

Lemma 5.4. *There exists a constant $C_{5.4} := C_{5.4}(\textbf{[H]}, T_0) > 0$, such that for all $s \in [t, t+T_0]$:*

$$\int_0^1 |\bar{R}_v^* \mathbb{M}_{s-t} p|^\alpha dv \geq C_{5.4} |\mathbb{M}_{s-t} p|^\alpha. \tag{5.11}$$

Since φ_{Λ_s} is integrable, we can write by (5.10) and Fourier inversion that for all $z \in \mathbb{R}^{nd}$:

$$\begin{aligned} p_{\Lambda_s}(z) &= \frac{1}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp e^{-i\langle p, z\rangle} \\ &\quad \times \exp\left(-(t-s) \int_0^1 \int_{\mathbb{R}^d} \{1 - \cos(\langle \mathbb{M}_{s-t} p, \bar{R}_v \sigma_{u(v)} z\rangle)\} g(|z|) \nu(dz) dv \right) \\ &\leq \frac{1}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp \exp\Big(-c(t-s)\big(C_{5.11}|\mathbb{M}_{s-t}p|^\alpha + 1\big)\Big), \end{aligned}$$

using (5.10) and (5.11) for the last inequality. This readily gives the global (diagonal) upper bound for the density. Now, let us also write from (2.2) and (2.4)

$$\begin{aligned} p_{\Lambda_s}(z) &= \frac{1}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp e^{-i\langle p, z\rangle} \\ \times &\ \exp\left(-(t-s) \int_{\mathbb{R}^+} \frac{d\rho}{\rho^{1+\alpha}} \int_0^1 \int_{S^{d-1}} \{1 - \cos\langle \mathbb{M}_{s-t} p, \bar{R}_v \sigma_{u(v)} \rho \varsigma\rangle\} g(\rho) \tilde{\mu}(d\varsigma) dv \right) \\ &= \frac{1}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp e^{-i\langle p, z\rangle} \\ &\quad \times \exp\left(-(t-s) \int_{\mathbb{R}^+} \frac{d\tilde{\rho}}{\tilde{\rho}^{1+\alpha}} \int_0^1 \int_{S^{d-1}} \{1 - \cos(\langle \mathbb{M}_{s-t} p, \frac{\bar{R}_v \sigma_{u(v)} \varsigma}{|\bar{R}_v \sigma_{u(v)} \varsigma|} \tilde{\rho}\rangle)\} \right. \\ &\quad \left. \times\, g\left(\frac{\tilde{\rho}}{|\bar{R}_v \sigma_{u(v)} \varsigma|} \right) |\bar{R}_v \sigma_{u(v)} \varsigma|^\alpha \tilde{\mu}(d\varsigma) dv \right). \end{aligned} \tag{5.12}$$

We now define the function

$$\begin{array}{rccc} f: & [0,1] \times S^{d-1} & \longrightarrow & S^{nd-1} \\ & (v, \varsigma) & \longmapsto & \frac{\bar{R}_v \sigma_{u(v)} \varsigma}{|\bar{R}_v \sigma_{u(v)} \varsigma|}, \end{array}$$

and on $[0,1]\times S^{d-1}$ the measure:

$$m_{\alpha,\tilde\rho}(dv,d\varsigma)=g\left(\frac{\tilde\rho}{|\bar R_v\sigma_{u(v)}\varsigma|}\right)|\bar R_v\sigma_{u(v)}\varsigma|^\alpha\tilde\mu(d\varsigma)dv.$$

The exponent in (5.12) thus rewrites:

$$\begin{aligned}&\int_0^1\int_{\mathbb{R}^d}\{1-\cos(\langle\mathbb{M}_{s-t}p,\bar R_v\sigma_{u(v)}z\rangle)\}g(|z|)\nu(dz)dv\\ =&\int_{\mathbb{R}^+}\frac{d\tilde\rho}{\tilde\rho^{1+\alpha}}\int_0^1\int_{S^{d-1}}\{1-\cos(\langle\mathbb{M}_{s-t}p,f(v,\varsigma)\tilde\rho\rangle)\}m_{\alpha,\tilde\rho}(dv,d\varsigma)\\ =&\int_{\mathbb{R}^+}\frac{d\tilde\rho}{\tilde\rho^{1+\alpha}}\int_{S^{nd-1}}\{1-\cos(\langle\mathbb{M}_{s-t}p,\eta\tilde\rho\rangle)\}\mu^*_{\tilde\rho}(d\eta),\end{aligned}$$

denoting by $\mu^*_{\tilde\rho}$ the image measure of $m_{\alpha,\tilde\rho}$ by f (which is a measure on S^{nd-1}). Symmetrizing $\mu^*_{\tilde\rho}$ introducing $\mu^*_{S,\tilde\rho}(A)=\frac{\mu^*_{\tilde\rho}(A)+\mu^*_{\tilde\rho}(-A)}{2}$, by parity of the cosine, we can write the exponent as:

$$\begin{aligned}&\int_0^1\int_{\mathbb{R}^d}\{1-\cos(\langle\mathbb{M}_{s-t}p,\bar R_v\sigma_{u(v)}z\rangle)\}g(|z|)\nu(dz)dv\\ &=\int_{\mathbb{R}^+}\frac{d\tilde\rho}{\tilde\rho^{1+\alpha}}\int_{S^{nd-1}}\{1-\cos(\langle\mathbb{M}_{s-t}p,\eta\tilde\rho\rangle)\}\mu^*_{S,\tilde\rho}(d\eta).\end{aligned}$$

We eventually derive:

$$\begin{aligned}p_{\Lambda_s}(z)&=\frac{1}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}}dpe^{-i\langle p,z\rangle}\\ &\quad\times\exp\left(-(t-s)\int_{\mathbb{R}^+}\frac{d\tilde\rho}{\tilde\rho^{1+\alpha}}\int_{S^{nd-1}}\{1-\cos(\langle\mathbb{M}_{s-t}p,\eta\tilde\rho\rangle)\}\mu^*_{S,\tilde\rho}(d\eta)\right)\\ &=\frac{\det(\mathbb{M}_{s-t}^{-1})}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}}dpe^{-i\langle p,\mathbb{M}_{s-t}^{-1}z\rangle}\\ &\quad\times\exp\big((t-s)\int_{\mathbb{R}^{nd}}\{\cos(\langle p,\xi\rangle)-1\}\nu_S(d\xi)\big),\end{aligned}\tag{5.13}$$

where ν_S is a symmetric measure on $\mathbb{R}^{nd}$. Also, from (5.12) and Lemma 5.4 we get that there exists a symmetric bounded measure $\bar\mu$ on S^{nd-1} and a constant $c>0$ s.t. for all $A\subset\mathbb{R}^{nd}$:

$$\nu_S(A)\le\int_{\mathbb{R}^+}\frac{d\rho}{\rho^{1+\alpha}}\int_{S^{nd-1}}\mathbf{1}_A(s\xi)g(c\rho)\bar\mu(d\xi),\tag{5.14}$$

where $\bar\mu$ satisfies **[H-4]** and $\dim(\mathrm{supp}(\bar\mu))=d$, recalling for this last property that $\tilde\mu$ is absolutely continuous w.r.t. the Lebesgue measure of S^{d-1}. In the stable case, corresponding to $g = 1$ the equality holds in (5.14), and $\bar{\mu}$ is the spherical part of ν_S. In that case $\bar{\mu} = \mu^*_{S,\tilde{\rho}} := \mu_S$ since the measure $\mu^*_{S,\tilde{\rho}}$ introduced above would not depend on ρ. In the general case, the domination in (5.14) can be simply derived from the fact that in (5.12) one has $g\left(\frac{\tilde{\rho}}{|\bar{R}_v \sigma_{u(v)} \varsigma|}\right) \geq g\left(c\tilde{\rho}\right),\ \forall (v, \varsigma) \in [0,1] \times S^{d-1}$.

□

Final derivation of the density bounds.

Diagonal controls. We first consider the case $|(\mathbb{T}^\alpha_{s-t})^{-1}(R_{s,t}x - y)| \leq K$. The upper-bound in (3.5) has already been proven. To obtain the lower-bound we perform computations rather similar to the ones in [Kol00b] which are recalled in Appendix 7.

Off-diagonal controls. We now consider the case $|(\mathbb{T}^\alpha_{s-t})^{-1}(R_{s,t}x - y)| > K$. We begin this paragraph recalling some results of Watanabe [Wat07]. The striking and subtle thing with multi-dimensional stable processes is that their large scale asymptotics highly depend on the spectral measure. Namely, for a given symmetric spectral measure $\bar{\mu}$ on S^{nd-1} satisfying **[H-4]**, implying that the associated symmetric stable process $(\bar{S}_t)_{t\geq 0}$ has a density on $\mathbb{R}^{nd}$ for $t > 0$, the tail asymptotics of $\bar{S}_1$ can behave, when $|x| \to +\infty$, as $p_{\bar{S}}(1, x) \asymp |x|^{-b}$ for $b \in [(1 + \alpha), nd(1 + \alpha)]$. Indeed, the behavior in $|x|^{-(1+\alpha)}$ would correspond to the decay of a scalar stable process and can appear if $\bar{\mu} = \sum_{i=1}^{nd} c_i(\delta_{e_i} + \delta_{-e_i})$, where the $(c_i)_{i\in[\![1,nd]\!]}$ are positive and $(e_i)_{i\in[\![1,nd]\!]}$ stand for the vectors of the canonical basis of $\mathbb{R}^{nd}$, when considering the asymptotics along **one** direction. On the other hand, the fastest possible decay of $|x|^{-nd(1+\alpha)}$ is also associated with this kind of spectral measure when investigating the large asymptotics for **all** the directions. Generally speaking, in the current framework, if $\bar{\mu}$ has support of dimension $k \in [\![0, nd-1]\!]$ the asymptotics of $\bar{S}_1$ satisfy that there exists $\bar{C} \geq 1$ s.t.:

$$\frac{\bar{C}^{-1}}{|x|^{nd(1+\alpha)}} \leq p_{\bar{S}}(1, x) \leq \frac{\bar{C}}{|x|^{k+1+\alpha}}. \tag{5.15}$$

We refer to Theorem 1.1 points *i)* and *iii)* in [Wat07] for the proof of these results. The strategy to derive those bounds consists in carefully splitting the small and large jumps. This approach turns out to be very useful for us to investigate the kernel H and is thoroughly exploited in Appendix 8.

From the representation (3.4) of the density of $\tilde{X}^{t,T,x,y}_s$ and (5.15) we readily get the indicated controls in the stable case. We refer to Appendix 8 for a thorough discussion on the general case.

Proof of Lemma 5.4.

It is enough to show that there exists $C_{5.4} := C_{5.4}(\mathbf{[H]}, T_0)$, s.t. for any $\theta \in S^{nd-1}$, $\int_0^1 |\bar{R}_v^* \theta|^\alpha dv \geq C_{5.4}$. We define

$$\bar{C} := \inf_{\theta \in S^{nd-1}} \int_0^1 |\bar{R}_v^* \theta|^\alpha dv.$$

By continuity of the involved functions and compactness of S^{nd-1}, the infimum is actually a minimum. We need to show that this quantity is not zero. We proceed by contradiction. Assume that $\bar{C} = 0$. Then, there exists $\theta_0 \in S^{nd-1}$ such that for almost all $v \in [0,1]$, $|\bar{R}_v^* \theta_0| = 0$. But since $\bar{R}_v^*$ is a continuous function in v, the previous statement holds for all $v \in [0,1]$, i.e. $\exists \theta_0 \in S^{nd-1}, \forall v \in [0,1], |\bar{R}_v^* \theta_0| = 0$, or equivalently, that $\exists \theta_0 \in S^{nd-1}, \forall v \in [0,1], \theta_0 \in Ker(\bar{R}_v^*)$. Take now arbitrary $(v_i)_{i \in [\![1,n]\!]}$ in $[0,1]$. We have for each $i \in [\![1,n]\!]$:

$$\begin{pmatrix} (\bar{R}^1_{v_i})^* & v_i(\bar{R}^2_{v_i})^* & \cdots & \frac{v_i^n}{(n-1)!}(\bar{R}^n_{v_i})^* \end{pmatrix} \begin{pmatrix} \theta_0^1 \\ \vdots \\ \theta_0^n \end{pmatrix} = 0_{\mathbb{R}^d}.$$

This equivalently writes in matrix form:

$$\begin{pmatrix} (\bar{R}^1_{v_1})^* & v_1(\bar{R}^2_{v_1})^* & \cdots & \frac{v_1^n}{(n-1)!}(\bar{R}^n_{v_1})^* \\ \vdots & \vdots & & \vdots \\ (\bar{R}^1_{v_n})^* & v_n(\bar{R}^2_{v_n})^* & \cdots & \frac{v_n^n}{(n-1)!}(\bar{R}^n_{v_n})^* \end{pmatrix} \begin{pmatrix} \theta_0^1 \\ \vdots \\ \theta_0^n \end{pmatrix} = 0_{\mathbb{R}^{nd}}.$$

Now, taking $v_1 \to 0$ in the first line yields $(\bar{R}^1_{v_1})^* \theta_0^1 = 0_{\mathbb{R}^d}$. Since the $(\bar{R}^i_v)_{i \in [\![1,n]\!]}$ are from Lemma 5.1 non degenerate, we have that $\theta_0^1 = 0_{\mathbb{R}^d}$. Hence, the second line becomes:

$$v_2(\bar{R}^2_{v_2})^* \theta_0^2 + \cdots + \frac{v_2^n}{(n-1)!}(\bar{R}^n_{v_2})^* \theta_0^n = 0_{\mathbb{R}^d}.$$

Dividing by v_2, and taking $v_2 \to 0$, we get $(\bar{R}^2_{v_2})^* \theta_0^2 = 0_{\mathbb{R}^d}$. Hence, $\theta_0^2 = 0_{\mathbb{R}^d}$. By induction, we have that all components $\theta_0^i = 0_{\mathbb{R}^d}$, but this contradicts $\theta_0 \in S^{nd-1}$. This yields $\bar{C} := C_{5.4} > 0$, which concludes the proof.

□

Remark 5.3. In the previous argument, the fact that the powers are increasing plays a key-role. Indeed, we rely on the multi-scale property reflected by the scale matrix $\mathbb{T}^\alpha$.

Proof of Lemma 4.1.

Let us write:

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

From Proposition 3.2, the second term tends to zero as T tends to t. Let us discuss the first term. Define:

$$\Delta = \int_{\mathbb{R}^{nd}} f(y)\Big(\tilde{p}_\alpha^{T,y}(t,T,x,y) - \tilde{p}_\alpha^{T,R_{T,t}x}(t,T,x,y)\Big)dy. \tag{5.16}$$

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

$$D_1 = \{y \in \mathbb{R}^{nd}; |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| \le K(T-t)^{-\beta}\},$$

$$D_2 = \{y \in \mathbb{R}^{nd}; |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| > K(T-t)^{-\beta}\}.$$

From Propositions 3.3, 3.4, the two densities in (5.16) are upper-bounded by

$$\frac{C\det(\mathbb{T}^\alpha_{T-t})^{-1}}{K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|^{d+1+\alpha}}.$$

The idea is that on D_2 they are both in the *off-diagonal* regime so that tail estimates can be used. On the other hand, we will explicitly exploit the compatibility between the spectral measures and the Fourier transform on D_1. Set for $i \in \{1,2\}$, $\Delta_{D_i} := \int_{D_i} f(y)\Big(\tilde{p}_\alpha^{T,y}(t,T,x,y) - \tilde{p}_\alpha^{T,R_{T,t}x}(t,T,x,y)\Big)dy$. We derive:

$$\begin{aligned}|\Delta_{D_2}| &\le C|f|_\infty \int_{D_2} \frac{\det(\mathbb{T}^\alpha_{T-t})^{-1}}{K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|^{d+1+\alpha}}dy \\ &= C|f|_\infty \int_{K(T-t)^{-\beta}}^{+\infty} dr \frac{r^{nd-1}}{K \vee r^{d+1+\alpha}} \\ &\le C(T-t)^{\beta((1-n)d+1+\alpha)}.\end{aligned}$$

Thus, for $\beta > 0$, $\Delta_{D_2} \underset{T\downarrow t}{\longrightarrow} 0$. On D_1, we will start from the inverse Fourier representation of $\tilde{p}_\alpha^{T,w}$ deriving from (5.12), for $w = y$ or $R_{T,t}x$. Namely,

$$\tilde{p}_\alpha^{T,w}(t,T,x,y) = \frac{1}{\det(\mathbb{M}_{T-t})(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp e^{-i\langle p, \mathbb{M}_{T-t}^{-1}(y - R_{T,t}x)\rangle} \exp\left(F_{T-t}(p,w)\right),$$

where the Fourier exponent writes $\forall (p,w)\in(\mathbb{R}^{nd})^2$:

$$F_{T-t}(p,w) = -(T-t)\int_0^1\int_{\mathbb{R}^d}\{1-\cos(\langle p,\bar R_v\sigma(u(v),R_{u(v),T}w)z\rangle)g(|z|)\nu(dz)\}.$$

We thus rewrite:

$$\begin{aligned}\Big(\tilde p_\alpha^{T,y}-\tilde p_\alpha^{T,R_{T,t}x}\Big)(t,T,x,y) &= \frac{1}{\det(\mathbb{M}_{s-t})(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} dp e^{-i\langle p,\mathbb{M}_{s-t}^{-1}(y-R_{T,t}x)\rangle}\\ &\int_0^1 d\lambda\Big(F_{T-t}(p,y)-F_{T-t}(p,R_{T,t}x)\Big)e^{(\lambda F_{T-t}(p,y)+(1-\lambda)F_{T-t}(p,R_{T,t}x))}.\end{aligned}$$

The key point is now to observe that from **[H-2]** the proof of Proposition 5.3 and the bound of Lemma 5.4, we have:

$$\forall (p,w)\in(\mathbb{R}^{nd})^2,\ F_{T-t}(p,w)\le C_{5.4}(T-t)(-|p|^\alpha+1).$$

Hence, $\exp(\lambda F_{T-t}(p,y)+(1-\lambda)F_{T-t}(p,R_{T,t}x))\le \exp(C_{5.4}(T-t)\{-|p|^\alpha+1\})$, independently on $\lambda\in[0,1]$. Now, the smoothness of the tempering function g in **[T]** yields:

$$\begin{aligned}|F_{T-t}(p,y)-F_{T-t}(p,R_{T,t}x)| &\le (T-t)\int_0^1\Big|\int_{\mathbb{R}^d}\cos(\langle\sigma(u(v),R_{u(v),T}y)^*\bar R_v^*p,z\rangle)\\ &\quad -\cos(\langle\sigma(u(v),R_{u(v),t}x)^*\bar R_v^*p,z\rangle)g(|z|)\nu(dz)\Big|dv\\ &\le c(T-t)\Big\{\int_0^1\int_{S^{d-1}}\Big||\langle p,\bar R_v\sigma(u(v),R_{u(v),T}y)\varsigma\rangle|^\alpha\\ &\quad -|\langle p,\bar R_v\sigma(u(v),R_{u(v),t}x)\varsigma\rangle|^\alpha\Big|\mu(d\varsigma)dv+1\Big\},\end{aligned}$$

using the notations of the proof of Proposition 5.3. On the other hand, since σ is η-Hölder continuous in its second variable (see **[H-1]**), we have:

$$\begin{aligned}|F(p,y)-F(p,R_{T,t}x)| &\le c(T-t)\{\int_0^1|p|^\alpha|R_{u(v),T}y-R_{u(v),t}x|^{\eta(\alpha\wedge1)}dv+1\}\\ &\le C(T-t)\{|p|^\alpha|y-R_{T,t}x|^{\eta(\alpha\wedge1)}+1\},\end{aligned}$$

using the Lipschitz property of the flow for the last inequality.

To summarize, we get in all cases:

$$\begin{aligned}|\Delta_{D_1}| &\le |f|_\infty\int_{D_1}dy\left|\tilde p_\alpha^{T,y}(t,T,x,y)-\tilde p_\alpha^{T,x}(t,T,x,y)\right|\\ &\le C\frac{1}{\det(\mathbb{M}_{T-t})}\int_{D_1}dy\int_{\mathbb{R}^{nd}}dp(T-t)\{|p|^\alpha|y-R_{T,t}x|^{\eta(\alpha\wedge1)}+1\}e^{-C_{5.4}(T-t)|p|^\alpha}.\end{aligned}$$

Changing variables, and integrating over p yields

$$\begin{aligned}
|\Delta_{D_1}| &\le \frac{C}{\det(\mathbb{T}^\alpha_{T-t})}\int_{\{|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\le K(T-t)^{-\beta}\}} dy\{|y-R_{T,t}x|^{\eta(\alpha\wedge 1)}+(T-t)\}\\
&\le C\int_{\{|Y|\le K(T-t)^{-\beta}\}} dY\{|\mathbb{T}^\alpha_{T-t}Y|^{\eta(\alpha\wedge 1)}+(T-t)\}\le C(T-t)^{\eta(\frac{1}{\alpha}\wedge 1)-\beta(nd+\eta(\alpha\wedge 1))}.
\end{aligned}$$

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

5.3 Estimates on the convolution kernel H.

In order to derive pointwise bounds on the kernel $H(t,T,x,y):=(L_t-\tilde L_t^{T,y})\tilde p_\alpha^{T,y}(t,T,x,y)$, it is convenient, since $\tilde p_\alpha^{T,y}$ is given in terms of Fourier inversion, to compute the symbols of the operators $L_t,\tilde L_t^{T,y}$. Precisely, we denote by $l_t(p,x)$ (resp. $\tilde l_t^{T,y}(p,x)$) the functions of $(p,x)\in(\mathbb{R}^{nd})^2$ s.t.

$$\begin{aligned}
\forall\varphi\in C_0^2(\mathbb{R}^{nd}),\ \forall x\in\mathbb{R}^{nd},\ L_t\varphi(x) &= \frac{1}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} dp\exp(-i\langle p,x\rangle)l_t(p,x)\hat\varphi(p),\\
\tilde L_t^{T,y}\varphi(x) &= \frac{1}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} dp\exp(-i\langle p,x\rangle)\tilde l_t^{T,y}(p,x)\hat\varphi(p).
\end{aligned}$$

We refer to Jacob [Jac96] for further properties of the symbols associated with an integro-differential operator. From usual properties of the (inverse) Fourier transform, we derive the following expressions.

Lemma 5.5. *Let $(p,x)\in(\mathbb{R}^{nd})^2$ be given. Recalling that B stands for the injection matrix of $\mathbb{R}^d$ into $\mathbb{R}^{nd}$, we have:*

$$\begin{aligned}
l_t(p,x) &= \langle p,A_t x\rangle+\int_{\mathbb{R}^d}\{\cos(\langle p,B\sigma(t,x)z\rangle)-1\}g(|z|)\nu(dz),\\
\tilde l_t^{T,y}(p,x) &= \langle p,A_t x\rangle+\int_{\mathbb{R}^d}\{\cos(\langle p,B\sigma(t,R_{t,T}y)z\rangle)-1\}g(|z|)\nu(dz).
\end{aligned}$$

From Lemma 5.5 we rewrite:

$$\begin{aligned}
H(t,T,x,y) &= \frac{1}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} dpe^{-i\langle p,y-R_{T,t}x\rangle}\\
&\quad\times\left\{\int_{\mathbb{R}^d}\{\cos(\langle p,B\sigma(t,x)z\rangle)-\cos(\langle p,B\sigma(t,R_{t,T}y)z\rangle)\}g(|z|)\nu(dz)\right\}\\
&\quad\times\exp\left(-\int_t^T du\int_{\mathbb{R}^d}\{1-\cos(\langle p,R^{1,\cdot}_{T,u}\sigma(u,R_{u,T}(y)\tilde z\rangle)\}g(|\tilde z|)\nu(d\tilde z)\right).
\end{aligned}$$

Remark 5.4. Observe the interesting fact that since the drift is linear, it disappears in the difference of the generators.

Let us now derive the diagonal bounds on the kernel, i.e. when $|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| \le K$. Observe first from the proof of Proposition 5.3 that we can write:

$$|H(t,T,x,y)| \le C\int_{\mathbb{R}^{nd}} dp \left| \int_{\mathbb{R}^d} \Big(\cos\langle p, B\sigma(t,x)z\rangle - \cos\langle p, B\sigma(t,R_{t,T}y)z\rangle \Big) g(|z|)\nu(dz) \right| \times \exp\left(-c|\mathbb{T}^\alpha_{T-t}p|^\alpha\right).$$

Assume first that $\alpha \in (0,1)$. We then perform a first order Taylor expansion in the variable $z = \rho\varsigma$ associated with a radial cut-off at threshold :

$$1/\{|p^1|\Delta\sigma(t,x,R_{t,T}y)\},\ \Delta\sigma(t,x,R_{t,T}y) := |\sigma(t,x) - \sigma(t,R_{t,T}y)|.$$

Recalling that σ is η-Hölder continuous, we obtain:

$$\begin{aligned}
|H(t,T,x,y)| &\le C\int_{\mathbb{R}^{nd}} dp \left\{ \int_{|z|\le 1/\{|p^1|\Delta\sigma(t,x,R_{t,T}y)\}} |p^1|\Delta\sigma(t,x,R_{t,T}y)\rho\tilde{\mu}(d\varsigma)\frac{d\rho}{\rho^{1+\alpha}} \right.\\
&\quad \left. + 2\int_{\rho > 1/\{|p^1|\Delta\sigma(t,x,R_{t,T}y)\}} \frac{d\rho}{\rho^{1+\alpha}} \right\} \exp\left(-c|\mathbb{T}^\alpha_{T-t}p|^\alpha\right)\\
&\le C\int_{\mathbb{R}^{nd}} dp|p^1|^\alpha\{\Delta\sigma(t,x,R_{t,T}y)\}^\alpha \exp\left(-c|\mathbb{T}^\alpha_{T-t}p|^\alpha\right)\\
&\le C\frac{\delta\wedge|x-R_{t,T}y|^{\alpha\eta}}{T-t}\int_{\mathbb{R}^{nd}} dp(T-t)|p^1|^\alpha \exp\left(-c|\mathbb{T}^\alpha_{T-t}p|^\alpha\right)\\
&\le C\frac{\delta\wedge|x-R_{t,T}y|^{\alpha\eta}}{T-t}\det(\mathbb{T}^\alpha_{T-t})^{-1}\\
&= C\frac{\delta\wedge|x-R_{t,T}y|^{\alpha\eta}}{T-t}\bar{p}_\alpha(t,T,x,y).
\end{aligned}$$

The case $\alpha \in (1,2)$ can be handled as above performing a Taylor expansion at order 2 for the small jumps and 1 for the large ones if for the threshold $1/|p^1|$. The case $\alpha = 1$ is direct in the stable case and can be extended to the tempered one performing a first order Taylor expansion for the small jumps using the smoothness of g around the origin.

This gives the claim of Lemma 3.7 in the diagonal regime. The off-diagonal case is much more involved and leads to consider a quite tricky phenomenon of *rediagonalization*. These aspects are considered in Appendix 8.

Remark 5.5. We emphasize here that we could also consider an additional bounded drift term in the first d components when $\alpha > 1$. Denoting this term by $b : \mathbb{R}^+ \times \mathbb{R}^{nd} \to$

$\mathbb{R}^d$, we could still use the previous frozen process as proxy. Exploiting the above symbol representation, the additional term coming from the difference of the generators would write

$$\langle b(t,x), \nabla_{x^1}\tilde{p}_\alpha(t,T,x,y)\rangle = \frac{1}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} dp e^{-i\langle p, y-R_{T,t}x\rangle}\langle b(t,x), p^1\rangle$$
$$\times \exp\left(-\int_t^T du \int_{\mathbb{R}^{nd}}\{1-\cos(\langle p, R^{1,\cdot}_{T,u}\sigma(u,R_{u,T}y)z\rangle)\}g(|z|)\nu(dz)\right),$$

where ∇_{x^1} stands for the derivative w.r.t. to the first d components. Observe that $|p^1|(T-t)^{1/\alpha}$ is homogeneous to the the contributions associated with p^1 in the exponential. This actually yields:

$$|\langle b(t,x), \nabla_{x^1}\tilde{p}_\alpha(t,T,x,y)\rangle| \le \frac{|b|_\infty}{(T-t)^{1/\alpha}}\bar{p}_\alpha(t,T,x,y),$$

on the diagonal which for $\alpha > 1$ gives an integrable singularity in time. The off-diagonal case can be handled as in Appendix 8.

6 Controls of the convolutions.

In this section we assume w.l.o.g. that $T \le T_0 = T_0(\mathbf{[H]}) \le 1$, as in Lemma 5.1. We first prove Lemma 3.8 that emphasizes how the spatial contribution in the r.h.s. of (3.20) yields, once integrated, a regularizing effect in time.

6.1 Proof of Lemma 3.8.

We prove the first estimate only, the other one is obtained similarly. Let us naturally split the space according to the regimes of $\bar{p}_\alpha$ and $\breve{p}_\alpha$. With the notations of Proposition 3.3 we introduce the partition:

$$D_1 = \{z \in \mathbb{R}^{nd}; |(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)| \le K\},$$
$$D_2 = \{z \in \mathbb{R}^{nd}; |(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)| > K\}.$$

On D_1, the diagonal control holds for $\bar{p}_\alpha + \breve{p}_\alpha$, that is, for $z \in D_1$ and recalling the definition of $\mathbb{T}^\alpha_{T-\tau}$ in Theorem 2.2:

$$(\bar{p}_\alpha + \breve{p}_\alpha)(\tau,T,z,y) \le C_{3.3}\det(\mathbb{T}^\alpha_{T-\tau})^{-1} = C_{3.3}(T-\tau)^{-d(\frac{n}{\alpha}+\frac{n(n-1)}{2})}.$$

On the other hand, denoting by $\|\cdot\|$ the matricial norm, we have from the scaling Lemma 5.2:

$$\begin{aligned}|z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)} &\le \|R_{\tau,T}\|^{\eta(\alpha\wedge1)}\|\mathbb{T}^\alpha_{T-\tau}\|^{\eta(\alpha\wedge1)}|(\mathbb{T}^\alpha_{T-\tau})^{-1}(y-R_{T,\tau}z)|^{\eta(\alpha\wedge1)}\\ &\le C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)},\end{aligned}$$

where the last inequality follows from the boundedness of the resolvent on compact sets and the definition of $\mathbb{T}^{\alpha}_{T-\tau}$.

Besides, the Lebesgue measure of the set D_1 is bounded by $C\det(\mathbb{T}^{\alpha}_{T-\tau})$, compensating exactly the time singularity appearing in the bound of $\tilde p_\alpha + \breve p_\alpha$. In conclusion, we obtained on D_1:

$$\int_{D_1} \delta \wedge |z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz \le C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}.$$

Similarly, for $z\in D_2$, the off-diagonal bound holds for $\bar p_\alpha$ and $\breve p_\alpha$, i.e.:

$$\begin{aligned}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y) \quad \le \quad & C\Bigg\{\frac{\det(\mathbb{T}^{\alpha}_{T-\tau})^{-1}}{|(\mathbb{T}^{\alpha}_{T-\tau})^{-1}(y-R_{T,\tau}z)|^{d+1+\alpha}} \\ & +\frac{\mathbf{1}_{|(z-R_{\tau,T}y)^1|/(T-\tau)^{1/\alpha}|\asymp|\mathbb{T}^{-\alpha}_{(T-\tau)}(z-R_{\tau,T}y)|}}{(T-\tau)^{d/\alpha}(1+\frac{|(z-R_{\tau,T}y)^1|}{(T-\tau)^{1/\alpha}})^{d+\alpha}} \\ & \times\frac{1}{(T-\tau)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}(1+|(\mathbb{T}^{-\alpha}_{T-\tau}(z-R_{\tau,T}y)^{2:n}|)^{1+\alpha}}\Bigg\}.\end{aligned}$$

From the scaling Lemma 5.2 we derive:

$$|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)} \le C|y-R_{T,\tau}z|^{\eta(\alpha\wedge1)} \le C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|(\mathbb{T}^{\alpha}_{T-\tau})^{-1}(y-R_{T,\tau}z)|^{\eta(\alpha\wedge 1)}.$$

Hence setting $\xi := |(\mathbb{T}^{\alpha}_{T-\tau})^{-1}(y-R_{T,\tau}z)|$ we first get

$$\begin{aligned}&\int_{D_2}\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}\bar p_\alpha(\tau,T,z,y)dz \\ \le\ & C\int_{\xi>K}\big(\delta\wedge[(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\xi^{\eta(\alpha\wedge1)}]\big)\xi^{nd-1}\frac{d\xi}{\xi^{d+1+\alpha}}. \qquad (6.1)\end{aligned}$$

Now if $\beta := (1-n)d+2+\alpha-\eta(\alpha\wedge1)>1$, we directly get:

$$\int_{\xi>K}\big(\delta\wedge[(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\xi^{\eta(\alpha\wedge1)}]\big)\frac{d\xi}{\xi^{(1-n)d+2+\alpha}} \le (T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\int_{\xi>K}\frac{d\xi}{\xi^\beta}\le C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}.$$

When $\beta\le 1$ we have to be more subtle. We refine the partition introducing:

$$D_{2,1}=\{\xi\in\mathbb{R}; K\le\xi\le K(T-\tau)^{-1/\alpha}\},\ D_{2,2}=\{\xi\in\mathbb{R};\xi>K(T-\tau)^{-1/\alpha}\}.$$

On $D_{2,1}$, writing $\delta\wedge[(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\xi^{\eta(\alpha\wedge1)}]\le[(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\xi^{\eta(\alpha\wedge1)}]$ we get:

$$(T-\tau)^{\eta(\frac{1}{\alpha}\wedge1)}\int_{\xi\in D_{2,1}}\frac{d\xi}{\xi^\beta}\le C\Big((T-\tau)^{\frac{(1-n)d+1+\alpha}{\alpha}}\mathbf{1}_{\beta<1}+(T-t)^{\eta(\frac{1}{\alpha}\wedge1)}|\log(T-\tau)|\mathbf{1}_{\beta=1}\Big).$$

On $D_{2,2}$, using $\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}\xi^{\eta(\alpha\wedge 1)}] \le \delta$ we derive $\int_{\xi\in D_{2,2}} \frac{d\xi}{\xi^{(1-n)d+2+\alpha}} \le C_\delta (T-\tau)^{((1-n)d+1+\alpha)/\alpha}$. Plugging the above controls in (6.1) yields the stated control. Let us now turn to:

$$\begin{aligned}
&\int_{D_2} \delta \wedge |z - R_{\tau,T} y|^{\eta(\alpha\wedge 1)} \breve p_\alpha(\tau,T,z,y)dz \\
\le\ & C\int_{|\zeta|>K} \big(\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\zeta|^{\eta(\alpha\wedge 1)}]\big)\frac{\mathbf{1}_{|\zeta^1|\asymp|\zeta|}}{(1+|\zeta^1|)^{d+\alpha}}\frac{d\zeta}{(1+|\zeta^{2:n}|)^{1+\alpha}},
\end{aligned}$$

where we have set $\zeta := (\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)$. We can now somehow *tensorize* the two contributions. We obtain on the considered events:

$$\begin{aligned}
&\int_{D_2} \delta \wedge |z - R_{\tau,T} y|^{\eta(\alpha\wedge 1)} \breve p_\alpha(\tau,T,z,y)dz \\
\le\ & C\Big\{\int_{|\zeta^1|>cK} \big(\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\zeta^1|^{\eta(\alpha\wedge 1)}]\big)\frac{d\zeta^1}{|\zeta^1|^{d+\alpha}} \\
&+\int_{|\zeta|>K} \big(\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\zeta^{2:n}|^{\eta(\alpha\wedge 1)}]\big)\frac{\mathbf{1}_{|\zeta^1|\asymp|\zeta|}}{(1+|\zeta^1|)^{d+\alpha}}\frac{d\zeta}{(1+|\zeta^{2:n}|)^{1+\alpha}}\Big\} \\
:=\ & \breve T_1 + \breve T_2.
\end{aligned}$$

For the term $\breve T_1$, we directly have $\breve T_1 \le C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}$ provided $\alpha > \eta(\alpha\wedge 1)$. Otherwise, i.e. the only possible case is $\alpha = \eta(\alpha \wedge 1)$, considering the partition $|\zeta^1| \in D_{2,1}\cup D_{2,2}$ as above replacing K by cK, one can reproduce the previous arguments. Namely, on $D_{2,1}$,

$$(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}\int_{D_{2,1}} r^{-(1+\alpha)+\eta(\alpha\wedge 1)}dr \le C\{(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\log(T-\tau)|\}.$$

On the other hand, on $D_{2,2}$,

$$\int_{D_{2,2}} (\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\zeta^1|^{\eta(\alpha\wedge 1)}])\frac{d\zeta^1}{|\zeta^1|^{d+\alpha}} \le \delta\int_{r>(T-\tau)^{-1/\alpha}Kc}\frac{dr}{r^{1+\alpha}} \le C(T-\tau).$$

For $\breve T_2$, on $\{|\zeta^{2:n}| \le K\}$ we directly get the estimate. Now, for $\{|\zeta|^{2:n} > K\}$ we get:

$$\begin{aligned}
&\int_{|\zeta^{2:n}|>K\cap|\zeta|>K} \big(\delta \wedge [(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}|\zeta^{2:n}|^{\eta(\alpha\wedge 1)}]\big)\frac{\mathbf{1}_{|\zeta^1|\asymp|\zeta|}}{(1+|\zeta^1|)^{d+\alpha}}\frac{d\zeta}{(1+|\zeta^{2:n}|)^{1+\alpha}}\Big\} \\
\le\ & (T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}\int_{|\zeta^1|>\bar cK}\frac{d\zeta^1}{(1+|\zeta^1|)^{d+\alpha}}\int_{c|\zeta^1|\ge|\zeta^{2:n}|\ge K}|\zeta^{2:n}|^{\eta(\alpha\wedge 1)}\frac{d\zeta^{2:n}}{(1+|\zeta^{2:n}|)^{1+\alpha}} \\
\le\ & C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}\int_{|\zeta^1|>\bar cK}\frac{d\zeta^1}{(1+|\zeta^1|)^{d+\alpha}}\{|\zeta^1|^{\eta(\alpha\wedge 1)+(n-1)d-1-\alpha}\mathbf{1}_{\beta<1} + \log(|\zeta^1|)\mathbf{1}_{\beta=1}\},
\end{aligned}$$

for β as above. Thus

$$\begin{aligned}\breve{T}_2 &\leq C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)}\{1+\int_{r>\bar{c}K} dr\{r^{-(d(1-n)+2+2\alpha-\eta(\alpha\wedge 1))}\mathbf{1}_{\beta<1}+r^{-(1+\alpha)}\log(r)\mathbf{1}_{\beta=1}\}\\ &\leq C(T-\tau)^{\eta(\frac{1}{\alpha}\wedge 1)},\end{aligned}$$

using again the condition $d(1-n)+1+\alpha>0$ for the last inequality. The smoothing bounds of equations (3.24), (3.24) for $d=1, n=2$ when the fast component is considered can be derived similarly. □

A useful extension of the previous result is the following lemma involving an additional logarithmic contribution which is *explosive* in the off-diagonal regime. This anyhow does not affect *much* the smoothing effect.

Lemma 6.1. *There exists $C_{6.1} := C_{6.1}(\textbf{[H]}, T_0) > 0$ s.t. for all $T \in (0,T_0], (x,y) \in (\mathbb{R}^{nd})^2$, $\tau \in (t,T)$:*

$$\begin{aligned}C_{6.1}(T-\tau)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)} &\geq \int_{\mathbb{R}^{nd}} \log(K\vee|(\mathbb{T}^{\alpha}_{T-\tau})^{-1}(y-R_{T,\tau}z)|)\\ &\quad\times\big\{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}\big\}(\bar{p}_\alpha+\breve{p}_\alpha)(\tau,T,z,y)dz,\\ C_{6.1}(\tau-t)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)} &\geq \int_{\mathbb{R}^{nd}} \log(K\vee|(\mathbb{T}^{\alpha}_{\tau-t})^{-1}(z-R_{\tau,t}x)|)\\ &\quad\times\big\{\delta\wedge[(\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|]^{\eta(\alpha\wedge 1)}\big\}\\ &\quad\times\bar{p}_{\alpha,\Theta}(t,\tau,x,z)dz\end{aligned}$$

Proof. The proof does not change much from the previous one. Observe first that, from the supremum in the logarithm, the only difference arises for *off-diagonal* regimes, that is, for $z\in D_2$ referring to the partition in the previous proof. The argument in the logarithm is however the same as the denominator of the *off-diagonal estimate*. After changing variables to ξ or ζ with the notations of the previous proof, it suffices to observe that for any $\varepsilon\in(0,\alpha)$, there exists $C_\varepsilon>0$ s.t. for all $r>K$: $\log(K\vee r)\leq C_\varepsilon r^\varepsilon$. Taking $\varepsilon>0$ s.t. $d(1-n)+1+\alpha-\varepsilon>0$ allows to proceed as in the proof of Lemma 3.8. □

We now state a key lemma for our analysis. It gives a control for the first convolution between the frozen density $\tilde{p}_\alpha$ and the parametrix kernel H. The result differs here from the expected one: we get an additional logarithmic factor, w.r.t. the bounds established for this quantity in [DM10] for the Gaussian degenerate case, or [Kol00b] for the stable non-degenerate case, as well as another contribution coming from the *rediagonalization* phenomenon.

Lemma 6.2 (**First Step Convolution**.). *Assume $d=1, n=2$. There exist $C_{6.2} := C_{6.2}($**[H]**$) > 0,\ \omega := \omega($**[H]**$) \in (0,1]$ s.t. for all $T \in (0,T_0],\ T_0 := T_0($**[H]**$) \le 1, (x,y) \in (\mathbb{R}^{nd})^2,\ t \in [0,T)$,*

$$
\begin{aligned}
|\tilde p_\alpha \otimes H|(t,T,x,y) \;\le\; & C_{6.2}\Big(\bar p_\alpha(t,T,x,y)\Big((T-t)^\omega \\
&+\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\big(1+\log[K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|]\big)\Big) \\
&+[\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}]\check p(t,T,x,y)\Big),
\end{aligned}
$$

with $\check p$ as in (3.25)*. Suppose now that* **[HT]** *holds, that $\sigma(t,x) = \sigma(t,x^2)$ and $\eta > 1/[(\alpha \wedge 1)(1+\alpha)]$. We can then improve the previous bound and derive:*

$$
|\tilde p_\alpha \otimes H|(t,T,x,y) \;\le\; C_{6.2}\Big((T-t)^\omega \bar p_{\alpha,\Theta}(t,T,x,y) + \bar q_{\alpha,\Theta}(t,T,x,y)\Big), \qquad (6.2)
$$

where we denote:

$$
\begin{aligned}
\bar q_{\alpha,\Theta}(t,T,x,y) \;=\; & \delta \wedge \{(T-t)|(x - R_{t,T}y)^1| + |(x - R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)} \\
& \times\big[\bar p_{\alpha,\Theta}(t,T,x,y)\left(1 + \log\left[K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|\right]\right)\big].
\end{aligned}
$$

Remark 6.1. The first part of the Lemma gives the bound of Lemma 3.9. Let us emphasize that this bound is not sufficient to derive the convergence of the parametrix series (3.14). The difficulty comes from the term in $\check p$ deriving from the rediagonalization phenomenon that induces a possible loss of concentration in the stable case and also prevents from a regularizing property in the tempered one if σ depends on both variables. Namely, the additional time singularity in $\check p$ can be compensated if σ only depends on the fast variable, which gives a higher order smoothing effect, but does not seem to be easily handleable in the general setting. The control (6.2) is actually sufficient to imply the convergence of the parametrix series when $d=1, n=2,\ \sigma(t,x) = \sigma(t,x^2)$ under the indicated condition on η. It gives the first statement in Lemma 3.10.

Proof. To perform the analysis, we first bound H using (3.20). We thus obtain:

$$
\begin{aligned}
&|\tilde p_\alpha \otimes H|(t,T,x,y) \le \\
&C\int_t^T d\tau \int_{\mathbb{R}^{nd}} \bar p_\alpha(t,\tau,x,z)\frac{\delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau}(\bar p_\alpha + \check p_\alpha)(\tau,T,z,y)dz. \qquad (6.3)
\end{aligned}
$$

For the proof it will be convenient to split the time interval $[t,T]$ into two subintervals $I_1 := [t, \frac{t+T}{2}], I_2 := [\frac{t+T}{2}, T]$. We observe that for $\tau \in I_1,\ T-\tau \asymp T-t$ whereas for

$\tau \in I_2,\ \tau - t \asymp T - t.$

The leading idea for the proof is to partition the space in order to say that one of the densities involved in (6.3) is homogeneous to the *global one* $\bar{p}_\alpha(t, T, x, y)$, and to get some *regularization* from the other contribution, using thoroughly Lemma 3.8.

Diagonal Estimates. When the global diagonal regime holds, i.e. $|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x - y)| \le K$, we will prove the following global diagonal estimate:

$$|\tilde{p}_\alpha \otimes H|(t, T, x, y) \le C\Big((T-t)^\omega + \delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\Big)\bar{p}_\alpha(t, T, x, y). \tag{6.4}$$

Indeed, on I_1, if $|(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)| \le K$, from Proposition 3.3 the diagonal estimate holds for $\bar{p}_\alpha(\tau, T, z, y)$. Since $T - \tau \asymp T - t$, we have:

$$\bar{p}_\alpha(\tau, T, z, y) \le C \det(\mathbb{T}^\alpha_{T-\tau})^{-1} \le C \det(\mathbb{T}^\alpha_{T-t})^{-1} \le C\bar{p}_\alpha(t, T, x, y).$$

On the other hand, if $|(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)| > K$, the off-diagonal expansion holds for $\bar{p}_\alpha(\tau, T, z, y)$ and from Proposition 3.3:

$$\begin{aligned}
\bar{p}_\alpha(\tau, T, z, y) &\le C\frac{\det(\mathbb{T}^\alpha_{T-\tau})^{-1}}{|(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)|^{d+1+\alpha}} \\
&\le C\det(\mathbb{T}^\alpha_{T-\tau})^{-1} \\
&\le C\det(\mathbb{T}^\alpha_{T-t})^{-1} \le C\bar{p}_\alpha(t, T, x, y).
\end{aligned}$$

Observe that we could have used here that the diagonal control is a global bound. We introduced the dichotomy on the regime to emphasize that it is a crucial argument in this section Additionally, the boundedness of the resolvent yields:

$$|z - R_{\tau,T}y| \le |z - R_{\tau,t}x| + |R_{\tau,t}x - R_{\tau,T}y| \le C\Big(|z - R_{\tau,t}x| + |x - R_{t,T}y|\Big). \tag{6.5}$$

On the other hand, on I_1:

$$\breve{p}(\tau, T, z, y) \le C\det(\mathbb{T}^\alpha_{T-t})^{-1} \le C\bar{p}_\alpha(t, T, x, y). \tag{6.6}$$

Denoting by $\otimes_{|I_1}$ the time-space convolution, where the time parameter is restricted to the interval I_1, we have from (6.3), (6.5), (6.6) and Lemma 3.8:

$$\begin{aligned}
|\tilde{p}_\alpha \otimes_{|I_1} H|(t, T, x, y) &\le C\bar{p}_\alpha(t, T, x, y)\int_{I_1} d\tau \int_{\mathbb{R}^{nd}} \bar{p}_\alpha(t, \tau, x, z)\left(\frac{\delta \wedge |z - R_{\tau,t}x|^{\eta(\alpha\wedge 1)}}{\tau - t}\right. \\
&\quad \left. + \frac{\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}}{T-t} + 1\right) dz \\
&\le C\bar{p}_\alpha(t, T, x, y) \\
&\quad \times \int_{I_1} d\tau \left((\tau - t)^{\omega - 1} + \frac{\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}}{T-t} + 1\right) \\
&\le C\bar{p}_\alpha(t, T, x, y)\Big((T-t)^\omega + \delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\Big).
\end{aligned} \tag{6.7}$$

Now, when $\tau \in I_2$, we have $\bar{p}_\alpha(t,\tau,x,z) \le \bar{p}_\alpha(t,T,x,y)$, so that from Lemma 3.8:

$$\begin{aligned}
|\tilde{p}_\alpha \otimes_{|I_2} H|(t,T,x,y) &\le C\bar{p}_\alpha(t,T,x,y) \\
&\quad \times \int_{I_2} d\tau \int_{\mathbb{R}^{nd}} \frac{\delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau} (\bar{p}_\alpha + \check{p}_\alpha)(\tau,T,z,y)dz \\
&\le C\bar{p}_\alpha(t,T,x,y) \int_{I_2} d\tau (T-\tau)^{\omega-1} \le C(T-t)^\omega \bar{p}_\alpha(t,T,x,y).
\end{aligned}$$

Off-Diagonal Estimates. We consider here the case $|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| \ge K$. Since we will need in the proof to exploit the *semigroup* property of Corollary 3.5 we restrict for the off-diagonal estimates to the case $d = 1, n = 2$.

Contributions involving $\bar{p}_\alpha(t,T,x,y)$.

We first consider the contributions involving $\bar{p}_\alpha(t,T,x,y)$ which is in the *off-diagonal* regime. In our current degenerate setting, several scales are involved in the term $|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|$. The *slow* time scale, associated with the first component of the process, induces in the off-diagonal regime additional time singularities in the density w.r.t. to the non-degenerate case. We thus need to be very careful when comparing the two contributions in $\bar{p}_\alpha$ appearing in the convolution $\tilde{p}_\alpha \otimes H$. Observe anyhow from the scaling Lemma 5.2 that:

$$\begin{aligned}
|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| &\le |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,\tau}z)| \qquad (6.8)\\
&\quad + |(\mathbb{T}^\alpha_{T-t})^{-1}(\mathbb{T}^\alpha_{T-t}\hat{R}^{t,T}_{\frac{\tau-t}{T-t}}(\mathbb{T}^\alpha_{T-t})^{-1}\{z - R_{\tau,t}x\})| \\
&\le |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,\tau}z)| + C|(\mathbb{T}^\alpha_{T-t})^{-1}(z - R_{\tau,t}x)| \\
&\le |(\mathbb{T}^\alpha_{T-\tau})^{-1}(y - R_{T,\tau}z)| + C|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z - R_{\tau,t}x)|, \qquad (6.9)
\end{aligned}$$

where $C := C([\mathbf{H}], T_0)$. Hence, at least one of the two densities involved in the convolution is off-diagonal. As emphasized below, the main difficulty w.r.t. the non degenerate case consists in suitably controlling the multi-scale effects that prevent from handling directly the time singularity of H in the convolution $\tilde{p}_\alpha \otimes H$, see e.g. Proposition 3.2 in Kolokoltsov [Kol00b]. Assume now that the component number $k \in \{1,2\}$ dominates in $\bar{p}_\alpha(t,T,x,y)$ when considering the flow at the current time τ of the convolution, the off-diagonal estimate becomes:

$$\begin{aligned}
\bar{p}_\alpha(t,T,x,y) &\le C\frac{\det(\mathbb{T}^\alpha_{T-t})^{-1}}{|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x - R_{\tau,T}y)|^{2+\alpha}} \theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x - R_{\tau,T}y)|) \\
&\le C\frac{(T-t)^{-\zeta(k)}}{|R^k_{\tau,t}x - R^k_{\tau,T}y|^{2+\alpha}} \theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x - R_{\tau,T}y)|), \\
\text{setting: } \zeta(k) &= (\frac{2}{\alpha} + 1) - ((k-1) + \frac{1}{\alpha})(2+\alpha).
\end{aligned}$$

According to the sign of the power of $T-t$, two cases arise. Set for $k \in \{1,2\}$, $\gamma(k) := \zeta(k)-1$. For the second, or *fast*, component, the exponent $\gamma(2)=\zeta(2)-1=1+\alpha$ is non negative. For the first, *slow* component $\gamma(1)=-1$. This is the aforementioned *slow/fast* dichotomy.

- When the fast component dominates, as the off-diagonal estimates are not singular in time anymore, no major problem arises. We refine (6.9) in the following sense:

$$K(T-t)^{1+\frac{1}{\alpha}} \le |R^2_{\tau,T}y - R^2_{\tau,t}x| \le |R^2_{\tau,T}y - z^2| + |z^2 - R^2_{\tau,t}x|.$$

Thus, at least one of the two densities in (6.3) is off-diagonal through a fast component. On the one hand, if $1/2|R^2_{\tau,T}y - R^2_{\tau,t}x| \le |z^2 - R^2_{\tau,t}x|$,

$$\begin{aligned}
\bar{p}_\alpha(t,\tau,x,z) &\le C\frac{\det(\mathbb{T}^\alpha_{\tau-t})^{-1}}{|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{\tau-t}(z-R_{\tau,t}x)|)\\
&\le C\frac{(\tau-t)^{\gamma(2)+1}}{|z^k - R^k_{\tau,t}x|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{\tau-t}(z-R_{\tau,t}x)|)\\
&\le C\frac{(T-t)^{\gamma(2)+1}}{|R^2_{\tau,t}x - R^2_{\tau,T}y|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x)-R_{\tau,T}y|).
\end{aligned}$$

On the other hand, if $1/2|R^2_{\tau,T}y - R^2_{\tau,t}x| \le |z^2 - R^2_{\tau,T}y|$,

$$\begin{aligned}
\frac{1}{T-\tau}\bar{p}_\alpha(\tau,T,z,y) &\le C\frac{(T-\tau)^{\gamma(2)}}{|z^2 - R^2_{\tau,T}y|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{T-\tau}(z-R_{\tau,T}y)|)\\
&\le \frac{C}{T-t}\frac{(T-t)^{\gamma(2)+1}}{|R^2_{\tau,T}y - R^2_{\tau,t}x|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,T}y-R_{\tau,t}x)|).
\end{aligned}$$

In both cases, we are in position to apply Lemma 3.8, directly in the first case, or similarly to (6.7) in the second one. The proof is then the same as in Kolokoltsov [Kol00b]. Observe that in the second case, we have compensated the singularity associated with the contribution $\bar{p}_\alpha$ in the kernel H, independently of the position of the time parameter τ.

- We now focus on the second case, that is when the slow component dominates so that $\gamma(1)$ is negative. We consider the partition $[t,T]=I_1\cup I_2$ and start with $\tau \in I_2$. In this case, we have $T-t \asymp \tau - t$. In other words, this is the case where the singularity induced by the kernel H is the worst.

We split $\mathbb{R}^2$ into

$$\begin{aligned}
D_1 &:= \{z\in\mathbb{R}^2; (T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)| \le |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\},\\
D_2 &:= \{z\in\mathbb{R}^2; (T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)| > |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\}, \quad (6.10)
\end{aligned}$$

for a parameter $\beta > 0$ to be specified later on. We define accordingly, for $i \in \{1,2\}$:

$$\bar{A}_{\alpha,I_2,D_i}(t,T,x,y) := \int_{I_2} d\tau \int_{D_i} \bar{p}_\alpha(t,\tau,x,z)\frac{\delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau}\bar{p}_\alpha(\tau,T,z,y)dz. \quad (6.11)$$

Observe first that since:

$$|R^1_{\tau,t}x - R^1_{\tau,T}y| \le |R^1_{\tau,t}x - z^1| + |z^1 - R^1_{\tau,T}y|,$$

and since the first and slow component dominates, we have that the tempering term for the convolution can be obtained taking out of the integral one of the tempering functions appearing in the densities. Precisely, we have either

$$\theta(|\mathbb{M}^{-1}_{\tau-t}(R_{\tau,t}x - z)|) \le C\theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x - R_{\tau,t}y)|),$$

or

$$\theta(|\mathbb{M}^{-1}_{T-\tau}(R_{\tau,T}y - z)|) \le C\theta(|\mathbb{M}^{-1}_{T-t}(R_{\tau,t}x - R_{\tau,t}y)|).$$

Let us first deal with $z \in D_1$ assuming w.l.o.g. that the first above condition holds, since otherwise the other tempering function in the bound of $\bar{p}_\alpha(\tau,T,z,y)$ can be taken out of the integral without altering the smoothing effect of the kernel. Since $\tau \in I_2$, we have:

$$\begin{aligned}
\bar{p}_\alpha(t,\tau,x,z) &\le C\frac{\det(\mathbb{T}^\alpha_{\tau-t})^{-1}}{|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z - R_{\tau,t}x)|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{\tau-t}(R_{\tau,t}x - z)|)\\
&\le C\frac{\det(\mathbb{T}^\alpha_{T-t})^{-1}}{(T-\tau)^{\beta(2+\alpha)}|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(y - R_{T,t}x)|).
\end{aligned}$$

Hence, as we did in the first part of the proof, we take out $\bar{p}_\alpha(t,\tau,x,z)$ off the integral (6.11). This is done here up to the additional singular coefficient $(T-\tau)^{-\beta(2+\alpha)}$. Still from Lemma 3.8, we get:

$$\bar{A}_{\alpha,I_2,D_1}(t,T,x,y) \le C\bar{p}_\alpha(t,T,x,y)\int_{I_2} d\tau (T-\tau)^{\omega-\beta(2+\alpha)-1}.$$

Then, in order to get an integrable bound, we must take:

$$0 < \beta < \frac{\omega}{2+\alpha}. \quad (6.12)$$

On D_2, we have to be more subtle. From the previous partition, the idea is to say that if $\tau \in [\tau_0, T]$ for τ_0 *close enough* to T, then the diagonal bound holds for the first density on D_2. In such cases we manage to get the global expected bound in the convolution. However, the previous τ_0 will highly depend on the global off-diagonal estimate $|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x - y)|$, and for $\tau \in I_2$, $\tau \le \tau_0$, we did not succeed to do better than integrating the singularity in $(T-\tau)^{-1}$ yielding the logarithmic contribution.

- Let us fix $\delta_0 \in (0, K)$. Observe that for fixed (t, T, x, y), if :

$$\tau \geq \tau_0 := T - \left(\frac{\delta_0}{|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|} \right)^{\frac{1}{\beta}}$$

then $\delta_0 \geq (T-\tau)^\beta |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|$. Then, since $z \in D_2$, we have $\delta_0 \geq |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z - R_{\tau,t}x)|$, and the diagonal estimate holds for $\bar{p}_\alpha(t, \tau, x, z)$. We write:

$$\begin{aligned}
& \bar{A}_{\alpha, I_2 \cap \{\tau \geq \tau_0\}, D_2}(t, T, x, y) \\
\leq \quad & C \int_{I_2 \cap \{\tau \geq \tau_0\}} d\tau \det(\mathbb{T}^\alpha_{\tau-t})^{-1} \int_{D_2} \frac{\delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau} \bar{p}_\alpha(\tau, T, z, y) dz \\
\overset{Lemma\ 3.8}{\leq} \quad & C\theta(|\mathbb{M}^{-1}_{T-t}(y - R_{T,t}x)|) \int_{I_2 \cap \{\tau \geq \tau_0\}} d\tau \det(\mathbb{T}^\alpha_{\tau-t})^{-1} (T-\tau)^{\omega - 1}.
\end{aligned}$$

Now $\delta_0^{2+\alpha} \geq (T-\tau)^{\beta(2+\alpha)} |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|^{2+\alpha}$, so that:

$$\begin{aligned}
\bar{A}_{\alpha, I_2 \cap \{\tau \geq \tau_0\}, D_2}(t, T, x, y) \quad \leq \quad & \int_{I_2} d\tau \det(\mathbb{T}^\alpha_{T-t})^{-1} (T-\tau)^{\omega - \beta(2+\alpha) - 1} \\
& \times \frac{\delta_0^{2+\alpha} \theta(|\mathbb{M}^{-1}_{T-t}(y - R_{T,t}x)|)}{|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|^{2+\alpha}}.
\end{aligned}$$

Thus, as long as β satisfies (6.12), $\bar{A}_{\alpha, I_2 \cap \{\tau \geq \tau_0\}, D_2}(t, T, x, y) \leq (T-t)^{\bar{\omega}} \bar{p}_\alpha(t, T, x, y)$, $\bar{\omega} := \omega - \beta(2+\alpha)$.

- Assume now that $\tau < \tau_0 = T - \left(\frac{\delta_0}{|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|} \right)^{\frac{1}{\beta}}$. The singularity induced by H is then integrable, and yields the logarithmic contribution. Specifically:

$$\begin{aligned}
& \bar{A}_{\alpha, I_2 \cap \{\tau < \tau_0\}, D_2}(t, T, x, y) \\
\leq \quad & C \int_{I_2} d\tau \mathbf{1}_{\tau \leq \tau_0} \int_{D_2} \bar{p}_\alpha(t, \tau, x, z) \frac{\delta \wedge |z - R_{\tau,T}(y)|^{\eta(\alpha\wedge 1)}}{T-\tau} \bar{p}_\alpha(\tau, T, z, y) dz.
\end{aligned}$$

Now, the key-point to get a smoothing effect is to keep the $\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}$ part in the control of the convolution. In order to keep track of this term, we need to determine which component dominates in $|x - R_{t,T}y|$. This can be rather intricate in the multi-scale setting. In the case $n = 2$, the only *slow* component is the first one. Saying that it dominates at a given integration time τ is asking:

$$|R^2_{\tau,T}y - R^2_{\tau,t}x| \leq (T-t)|R^1_{\tau,T}y - R^1_{\tau,t}x|. \tag{6.13}$$

Furthermore, we can write:

$$|R^1_{T,t}x - y^1| \quad \geq \quad |R^1_{\tau,t}x - R^1_{\tau,T}y| - \|R_{T,\tau} - I\| |R_{\tau,t}x - R_{\tau,T}y|.$$

From Lemma 5.1, and observing from its proof that we could also establish that $\sum_{j=1}^{2} \|(R_{T,\tau}-I)^{j,2}\| + \|(R_{T,\tau}-I)^{1,1}\| \le C(T-\tau),\ C := C([\mathbf{H}], T_0),\ T_0 \le 1$ we get using (6.13):

$$|R^1_{T,t}x - y^1| \ \ge \ |R^1_{\tau,t}x - R^1_{\tau,T}y|(1 - C(T-\tau)).$$

Thus, for T small enough we get:

$$(T-t)|R^1_{T,t}x - y^1| \ge \frac{T-t}{2}|R^1_{\tau,t}x - R^1_{\tau,T}y| \overset{(6.13)}{\ge} \frac{1}{2}|R^2_{\tau,t}x - R^2_{\tau,T}y|.$$

We then derive similarly that:

$$\begin{aligned} |R^2_{\tau,t}x - R^2_{\tau,T}y| &\ge |R^2_{T,t}x - y^2| - \|R_{\tau,T} - I\||R_{T,t}x - y| \\ &\ge \frac{|R^2_{T,t}x - y^2|}{2} - C(T-\tau)|R^1_{T,t}x - y^1|. \end{aligned}$$

This finally yields that

$$(T-t)|R^1_{T,t}x - y^1| \ge \frac{|R^2_{T,t}x - y^2|}{4(1+C)}, \tag{6.14}$$

that is, the first component dominates in the contribution $|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x - y)|$ appearing in D_2. Write now:

$$|z - R_{\tau,T}y| \le |z^1 - R^1_{\tau,t}x| + |z^2 - R^2_{\tau,t}x| + |R_{\tau,t}x - R_{\tau,T}y|. \tag{6.15}$$

◇ Suppose first that $(\tau - t)|z^1 - R^1_{\tau,t}x| \le |z^2 - R^2_{\tau,t}x|$. Since $z \in D_2$, we have from (6.14):

$$|z^2 - R^2_{\tau,t}x| \le C(\tau - t)(T-\tau)^\beta |R^1_{T,t}x - y^1|.$$

Consequently, plugging the last two inequalities into (6.15), we get:

$$\begin{aligned} |z - R_{\tau,T}y| &\le \left(\frac{1}{\tau - t} + 1\right)|z^2 - R^2_{\tau,t}x| + |R_{\tau,t}x - R_{\tau,T}y| \\ &\le \Big(1 + (\tau - t)\Big)(T-\tau)^\beta |R^1_{T,t}x - y^1| + |R_{\tau,t}x - R_{\tau,T}y| \\ &\le C|x - R_{t,T}y|, \end{aligned}$$

using the Lipschitz property of the flow for the last inequality.

◇ Assume now that $|z^2 - R^2_{\tau,t}x| \le (\tau - t)|z^1 - R^1_{\tau,t}x| \le |z^1 - R^1_{\tau,t}x|$. We exploit that $z \in D_2$ and (6.14) to write:

$$|z^1 - R^1_{\tau,t}x| \le C(T-\tau)^\beta |R^1_{T,t}x - y^1|.$$

Plugging the last two inequalities into (6.15) yields:

$$\begin{aligned}
|z - R_{\tau,T}y| &\le 2|z^1 - R^1_{\tau,t}x| + |R_{\tau,t}x - R_{\tau,T}y| \\
&\le 2C(T-\tau)^\beta |R^1_{T,t}x - y^1| + |R_{\tau,t}x - R_{\tau,T}y| \\
&\le C|x - R_{t,T}y|,
\end{aligned}$$

using again the Lipschitz property of the flow for the last inequality.

Thus, in both cases,

$$|z - R_{\tau,T}y| \le C|x - R_{t,T}y| \Rightarrow \delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)} \le C\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}. \quad (6.16)$$

It could similarly be shown that when $\sigma(t,x) := \sigma(t,x^2)$:

$$\begin{aligned}
& \delta \wedge |(z - R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)} \quad (6.17)\\
\le\ & C\delta \wedge \{(T-t)|(x - R_{t,T}y)^1| + |(x - R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)} \\
\le\ & C\delta \wedge \{(T-t)|(R_{T,t}x - y)^1| + |(R_{T,t}x - y)^2|\}^{\eta(\alpha\wedge 1)},
\end{aligned}$$

using a direct modification of Lemma 5.2 for the last inequality. Taking out this contribution from the spatial integral we get:

$$\begin{aligned}
\bar{A}_{\alpha, I_2 \cap \{\tau \le \tau_0\}, D_2}(t,T,x,y) &\le C \int_{I_2} d\tau \frac{\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau} \\
&\quad \times \mathbf{1}_{\tau \le \tau_0} \int \bar{p}_\alpha(t,\tau,x,z)\bar{p}_\alpha(\tau,T,z,y)dz \\
&\le C\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)} \log\left(K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|\right) \\
&\quad \times \bar{p}_\alpha(t,T,x,y),
\end{aligned}$$

using the semigroup property of Corollary 3.5 for the last inequality.

To complete the analysis for this contribution, it remains to consider the case $\tau \in I_1$. In this case, $T - t \asymp T - \tau$, and we have by triangle inequality:

$$\delta \wedge |z - R_{\tau,T}y|^{\eta(\alpha\wedge 1)} \le C\left(\delta \wedge |z - R_{\tau,t}x|^{\eta(\alpha\wedge 1)} + \delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}\right).$$

Recalling that $T - \tau$ is not singular and splitting the integrals accordingly yields:

$$\begin{aligned}
\bar{A}_{\alpha, I_1}(t,T,x,y) &\le C \int_{I_1} d\tau \int_{\mathbb{R}^{nd}} dz \bar{p}_\alpha(t,\tau,x,z) \frac{\delta \wedge |x - R_{t,\tau}z|^{\eta(\alpha\wedge 1)}}{\tau - t} \bar{p}_\alpha(\tau,T,z,y) \\
&\quad + C\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)} \bar{p}_\alpha(t,T,x,y),
\end{aligned}$$

where we used the semigroup property of Corollary 3.5 for the last term in the r.h.s. Now, for the first term in the above r.h.s., the previous arguments apply. Similarly to (6.9) one of the two terms $|(\mathbb{T}^\alpha_{\tau-t})^{-1}(R_{\tau,t}x - z)|$, $|(\mathbb{T}^\alpha_{T-\tau})^{-1}(R_{T,\tau}z - y)|$ is in the *off-diagonal* regime. If it is the second one, then $\bar p_\alpha(\tau,T,z,y) \le C\bar p_\alpha(t,T,x,y)$ and we conclude using Lemma 3.8. If it is the first term, then we can still perform the previous dichotomy along the dominating component in $|(\mathbb{T}^\alpha_{\tau-t})^{-1}(R_{\tau,t}x - z)|$. If the fast component dominates, the density is not singular. When the first component dominates, we modify the previous partition $(D_i)_{i\in\{1,2\}}$, considering:

$$D_1 = \{z \in \mathbb{R}^{nd}; (\tau-t)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| \le |(\mathbb{T}^\alpha_{T-\tau})^{-1}(z - R_{\tau,T}y)|\},$$

$$D_2 = \{z \in \mathbb{R}^{nd}; (\tau-t)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)| > |(\mathbb{T}^\alpha_{T-\tau})^{-1}(z - R_{\tau,T}y)|\}.$$

From this point on, the proof is similar: on D_1, we compensate the singularity, as long as β is like in (6.12). When $z \in D_2$, we subdivide along $\delta_0 \le$ or $> (\tau-t)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y - R_{T,t}x)|$. The first case is dealt as above. In the second case, we can integrate the time singularity.

Contributions involving $\breve p_\alpha(t,T,x,y)$. We first focus on the contribution

$$\begin{aligned}
\breve A_{\alpha,I_2} &:= \int_{I_2} d\tau \int \bar p_\alpha(t,\tau,x,z)\frac{\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau}\breve p_\alpha(\tau,T,z,y)dz\\
&\le \int_{I_2} d\tau \int \bar p_\alpha(t,\tau,x,z)\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}\mathbf{1}_{\frac{|(z-R_{\tau,T}y)^1|}{(T-\tau)^{1/\alpha}}\asymp|(\mathbb{T}^\alpha_{T-\tau})^{-1}(z-R_{\tau,T}y)|\ge K}\\
&\quad\times\frac{1}{|(z-R_{\tau,T}y)^1|^{1+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(z-R_{\tau,T}y)|)\frac{dz}{(T-\tau)^{1+\frac{1}{\alpha}}(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}.
\end{aligned}$$

Using again the partition in equation (6.10), we readily get from Lemma 3.8, similarly to the previous paragraph, that:

$$\begin{aligned}
\breve A_{\alpha,I_2,D_1} &:= \int_{I_2} d\tau \int_{D_1} \bar p_\alpha(t,\tau,x,z)\frac{\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau}\breve p_\alpha(\tau,T,z,y)dz\\
&\le C(T-\tau)^\omega\bar p_\alpha(t,T,x,y).
\end{aligned}$$

On D_2 the previous arguments also apply for $\{\tau \ge \tau_0\}$, with the same definition of τ_0. Hence:

$$\breve A_{\alpha,I_2\cap\{\tau>\tau_0\},D_2} \le C(T-\tau)^\omega\bar p_\alpha(t,T,x,y).$$

The only remaining case to handle is when the slow component dominates at the current time τ, i.e. $|(R_{\tau,t}x - R_{\tau,T}y)^1| \ge c_0(T-t)|(R_{\tau,t}x - R_{\tau,T}y)^2|$.

On the considered set, it has previously been proven on D_2 (see (6.16)) that $\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge1)} \le C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge1)}$ which can be taken out of the integral. Thus on the considered set, recalling that $|(z-R_{\tau,T}y)^1| \ge c|(R_{\tau,t}x - R_{\tau,T}y)^1|$:

$$
\begin{aligned}
&\breve{A}_{\alpha,I_2\cap\{\tau\le\tau_0\},D_2} \le \\
&C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\int_{I_2}d\tau\mathbf{1}_{\tau\le\tau_0}\int_{D_2}\bar{p}_\alpha(t,\tau,x,z)\times\mathbf{1}_{\frac{|(z-R_{\tau,T}y)^1|}{(T-\tau)^{1/\alpha}}\asymp|(\mathbb{T}^\alpha_{T-\tau})^{-1}(z-R_{\tau,T}y)|\ge K}\\
&\times\frac{1}{|(z-R_{\tau,T}y)^1|^{1+\alpha}}\frac{dz}{(T-\tau)^{1+\frac{1}{\alpha}}(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}\theta(|\mathbb{M}^{-1}_{T-\tau}(R_{\tau,T}y-z)|)\\
&\le C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\theta(|\mathbb{M}^{-1}_{T-t}(R_{t,T}y-x)|)\int_{I_2}d\tau\frac{\mathbf{1}_{\tau\le\tau_0}}{|(R_{\tau,t}x-R_{\tau,T}y)^1|^{1+\alpha}}\\
&\times\int_{D_2}\frac{dz}{(\tau-t)^{\frac{2}{\alpha}+1}(1+\frac{|(R_{\tau,t}x-z)^1|}{(\tau-t)^{1/\alpha}}+\frac{|(R_{\tau,t}x-z)^2|}{(\tau-t)^{1+1/\alpha}})^{2+\alpha}}\times\frac{\mathbf{1}_{\frac{|(z-R_{\tau,T}y)^1|}{(T-\tau)^{1/\alpha}}\asymp|(\mathbb{T}^\alpha_{T-\tau})^{-1}(z-R_{\tau,T}y)|\ge K}}{(T-\tau)^{1+\frac{1}{\alpha}}(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}\\
&\le C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\frac{\theta(|\mathbb{M}^{-1}_{T-t}(R_{t,T}y-x)|)}{(T-t)^{1/\alpha}(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|)^{1+\alpha}}\\
&\int_{I_2}d\tau\frac{\mathbf{1}_{\tau\le\tau_0}}{T-t}\int\frac{dz_2}{(\tau-t)^{\frac{1}{\alpha}+1}(1+\frac{|(R_{\tau,t}x-z)^2|}{(\tau-t)^{1+1/\alpha}})^{1+\alpha}}\frac{1}{(T-\tau)^{1+\frac{1}{\alpha}}(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}\\
&\le C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\frac{\theta(|\mathbb{M}^{-1}_{T-t}(R_{t,T}y-x)|)}{(T-t)^{1/\alpha}(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|)^{1+\alpha}}\\
&\int_{I_2}d\tau\frac{\mathbf{1}_{\tau\le\tau_0}}{T-t}\frac{1}{(T-t)^{\frac{1}{\alpha}+1}(1+\frac{|(R_{\tau,t}x-R_{\tau,T}y)^2|}{(T-t)^{1+1/\alpha}})^{1+\alpha}}.
\end{aligned}
$$

From this last inequality we deduce that if

$$|(R_{\tau,t}x-R_{\tau,T}y)^2|\ge c_1(T-t)|(R_{\tau,t}x-R_{\tau,T}y)^1|,$$

i.e. the components are equivalent, we get the expected control, which could have already been deduced from the fact that the fast component is equivalent to the global energy. If such an equivalence does not hold, the natural control is:

$$
\begin{aligned}
\breve{A}_{\alpha,I_2\cap\{\tau\le\tau_0\},D_2} \le\ & C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\frac{\theta(|\mathbb{M}^{-1}_{T-t}(R_{t,T}y-x)|)}{(T-t)^{1/\alpha}(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|)^{1+\alpha}}\\
&\times\frac{1}{(T-t)^{\frac{1}{\alpha}+1}(1+\inf_{\tau\in[t,T]}\frac{|(R_{\tau,t}x-R_{\tau,T}y)^2|}{(T-t)^{1+1/\alpha}})^{1+\alpha}}.
\end{aligned}
$$

Now in the stable case **[HS]**, we obtain:

$$
\begin{aligned}
\breve{A}_{\alpha,I_2\cap\{\tau\le\tau_0\},D_2} \le\ & \frac{C\delta\wedge|R_{\tau^*,t}x-R_{\tau^*,T}y|^{\eta(\alpha\wedge 1)}}{(T-t)^{1/\alpha}(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau^*,t}x-R_{\tau^*,T}y)|)^{1+\alpha}}\\
&\times\frac{1}{(T-t)^{\frac{1}{\alpha}+1}(1+\frac{|(R_{\tau^*,t}x-R_{\tau^*,T}y)^2|}{(T-t)^{1+1/\alpha}})^{1+\alpha}},
\end{aligned}
$$

where τ^* achieves the minimum. In the tempered case **[HT]**, the control reads:

$$\begin{aligned}\breve{A}_{\alpha,I_2\cap\{\tau\le\tau_0\},D_2} &\le C\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}\bar{p}_{\alpha,\Theta}(t,T,x,y)\\ &\times\frac{1}{(T-t)^{\frac{1}{\alpha}}(1+\frac{|(R_{\tau^*,t}x-R_{\tau^*,T}y)^2|}{(T-t)^{1+1/\alpha}})^{1+\alpha}}.\end{aligned}$$

Similar controls could be established by symmetry for $\breve{A}_{\alpha,I_1}$. These bounds thus yield in both cases an additional time-singularity in $(T-t)^{-1/\alpha}$ if $|(R_{\tau^*,t}x-R_{\tau^*,T}y)^2|\le K(T-t)^{1+1/\alpha}$ and a possible loss of concentration in the stable case. They also turn out to be difficult to exploit in order to iterate in the series to establish the existence of the density and related bounds.

Now if $\sigma(t,x):=\sigma(t,x_2)$ we can get rid of the additional singularity in the tempered case, writing:

$$\begin{aligned}&\breve{A}_{\alpha,\{\tau\le\tau_0\},D_2}\le\\ &C\int_t^{T-\tau_0}d\tau\int_{D_2}\bar{p}_\alpha(t,\tau,x,z)\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}}{|(z-R_{\tau,T}y)^1|^{1+\alpha}}\theta(|\mathbb{M}_{T-\tau}^{-1}(z-R_{\tau,T}y)|)\\ &\times\frac{1}{(T-\tau)^{1+\frac{1}{\alpha}}}\frac{dz}{(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}\\ &\le C\frac{\theta(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|)}{|(R_{T,t}x-y)^1|^{1+\alpha}}\times C\int_t^{T-\tau_0}d\tau\int\frac{1}{(\tau-t)^{1+\frac{1}{\alpha}}(1+\frac{|(R_{\tau,t}x-z)^2|}{(\tau-t)^{1+\frac{1}{\alpha}}})^{1+\alpha}}\\ &\times\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}\frac{1}{(T-\tau)^{1+\frac{1}{\alpha}}}\frac{dz^2}{(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{1+\alpha}}.\end{aligned}$$

Observe now from **[HT]** that we have the control:

$$\begin{aligned}\frac{\theta(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|)}{|R_{T,t}x-y|^{1+\alpha}} &\le \frac{|(R_{T,t}x-y)^1|\theta(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|)}{|(R_{T,t}x-y)^1|^{2+\alpha}}\\ &\le \frac{\Theta(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|)}{|R_{T,t}x-y|^{1+\alpha}}\\ &= \bar{p}_{\alpha,\Theta}(t,T,x,y),\end{aligned}$$

on the considered case (i.e. the first component dominates in the off-diagonal regime).

Hence:

$$
\begin{aligned}
&\breve{A}_{\alpha,\{\tau\le\tau_0\},D_2} \le \\
&C\bar p_{\alpha,\Theta}(t,T,x,y)\int_t^{T-\tau_0} d\tau \left\{(\tau-t)^{-(1+\frac{1}{\alpha})(q-1)}\int \frac{(\tau-t)^{-\left(1+\frac{1}{\alpha}\right)}dz^2}{(1+\frac{|(z-R_{\tau,t}x)^2|}{(\tau-t)^{1+\frac{1}{\alpha}}})^{(1+\alpha)q}}\right\}^{1/q}\\
&\times\left\{(T-\tau)^{-\{(1+\frac{1}{\alpha})(p-1)\}}\int [\delta\wedge |(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}]^p\frac{(T-\tau)^{-\left(1+\frac{1}{\alpha}\right)}dz^2}{(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac{1}{\alpha}}})^{(1+\alpha)p}}\right\}^{1/p} \qquad (6.18)\\
&\le C\bar p_{\alpha,\Theta}(t,T,x,y)\int_t^{T-\tau_0} d\tau(\tau-t)^{-(1+\frac{1}{\alpha})\frac{1}{p}}(T-\tau)^{-\{(1+\frac{1}{\alpha})\frac{1}{q}\}}\{(T-t)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)}\},
\end{aligned}
$$

where $p,q>1, p^{-1}+q^{-1}=1$ and s.t. $p>1+\frac{1}{\alpha}$ for $\tau\in[t,\frac{t+T}{2}]$ and $q>1+\frac{1}{\alpha}$ for $\tau\in[\frac{t+T}{2},T]$. Also, the regularizing term $(T-t)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)}$ in the last control can be derived following the proof of Lemma 3.8. We thus derive:

$$
\begin{aligned}
\breve{A}_{\alpha,\{\tau\le\tau_0\},D_2} &\le C\bar p_{\alpha,\Theta}(t,T,x,y)(T-t)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)}(T-t)^{1-(1+\frac{1}{\alpha})}\\
&\le C\bar p_{\alpha,\Theta}(t,T,x,y)(T-t)^{1+(1+\frac{1}{\alpha})(\eta(\alpha\wedge 1)-1)}.
\end{aligned}
$$

Therefore, the last contribution gives a smoothing effect provided:

$$
(1+\frac{1}{\alpha})(\eta(\alpha\wedge 1)-1)>-1 \iff \eta>\frac{1}{(\alpha\wedge 1)(1+\alpha)}.
$$

The controls associated with $\bar p_\alpha$, when $\sigma(t,x)=\sigma(t,x^2)$, yielding the contribution in $\bar q_{\alpha,\Theta}$ in the Lemma, could be easily deduced in the current case from the previous analysis, exploiting (6.17) instead of (6.16).

□

The convergence of the parametrix series (3.14) will now follow from controls involving the convolutions of H with the last term $\bar q_{\alpha,\Theta}(t,T,x,y)$. The following lemma completes the proof of Lemma 3.10.

Lemma 6.3. *Assume $d=1,n=2,\sigma(t,x)=\sigma(t,x_2)$ and $\eta>\frac{1}{(\alpha\wedge 1)(1+\alpha)}$. There exist $C_{6.3}:=C_{6.3}(\textbf{[H]})>0,\ \omega:=\omega(\textbf{[H]})\in(0,1]$ s.t. for all $T\in(0,T_0]$, $T_0:=T_0(\textbf{[H]})\le 1$, $(x,y)\in(\mathbb{R}^{nd})^2$, $t\in[0,T)$,*

$$
\begin{aligned}
|\bar q_{\alpha,\Theta}\otimes H|(t,T,x,y) \le\ & C(T-t)^{\omega}\Big[\bar p_{\alpha,\Theta}(t,T,x,y)\\
&+\delta\wedge\Big((T-t)|(x-R_{t,T}y)^1|+|(x-R_{t,T}y)^2|\Big)^{\eta(\alpha\wedge 1)}\\
&\times\log\Big(K\vee|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|\Big)\bar p_{\alpha,\Theta}(t,T,x,y)\Big].
\end{aligned}
$$

Proof. Recall that $\bar{q}_{\alpha,\Theta}(t,T,x,y)$ writes as the sum of

$$q_{\alpha,\Theta}(t,T,x,y) := \delta \wedge \Big((T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\Big)^{\eta(\alpha\wedge 1)} \bar{p}_{\alpha,\Theta}(t,T,x,y)$$

and

$$\begin{aligned}\rho_{\alpha,\Theta}(t,T,x,y) &:= \delta \wedge \Big((T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\Big)^{\eta(\alpha\wedge 1)} \\ &\quad \times \log\Big(K \vee |(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|\Big)\bar{p}_{\alpha,\Theta}(t,T,x,y).\end{aligned}$$

Though the lines of the proof are similar to those of Lemma 6.2, we treat the two convolutions separately, to emphasize the difficulties induced by the rediagonalization and the logarithmic factor. First, for $|q_{\alpha,\Theta}\otimes H|(t,T,x,y)$, we bound $|H|$ using Lemma 3.7, to get:

$$\begin{aligned}|q_{\alpha,\Theta}\otimes H|(t,T,x,y) \le C\int_t^T d\tau \int_{\mathbb{R}^{nd}} &\delta \wedge \Big((\tau-t)|(z-R_{\tau,t}x)^1| + |(z-R_{\tau,t}x)^2|\Big)^{\eta(\alpha\wedge 1)} \\ &\times \bar{p}_{\alpha,\Theta}(t,\tau,x,z)\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}}{T-\tau}(\bar{p}_\alpha+\check{p}_\alpha)(\tau,T,z,y).\end{aligned}$$

The above contribution can be handled as in Lemma 6.2, in the *diagonal* case $|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)| \le K$, or in the *off-diagonal* case $|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)| > K$ when for a given integration time $\tau\in[t,T]$ the *fast component* dominates, i.e.

$$|R^2_{\tau,T}y - R^2_{\tau,t}x| \ge (T-t)|R^1_{\tau,T}y - R^1_{\tau,t}x|.$$

The only difference is that we do not need to use the triangle inequality in order to apply Lemma 3.8. Indeed, regularizing terms $\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}$, and $\delta\wedge\{(\tau-t)|(z-R_{\tau,t}x)^1| + |(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge 1)}$ already appear for both densities.

When $|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)| > K$ and

$$|R^2_{\tau,T}y - R^2_{\tau,t}x| \le (T-t)|R^1_{\tau,T}y - R^1_{\tau,t}x|,$$

we split as in the previous proof the time interval into $I_1\cup I_2 := [t,\frac{T+t}{2}]\cup[\frac{T+t}{2},T]$. Suppose $\tau\in I_2$. We consider the spatial partition introduced in (6.10).

For $z\in D_1$, we have $\bar{p}_{\alpha,\Theta}(t,\tau,x,z)\le C(T-\tau)^{-\beta(2+\alpha)}\bar{p}_{\alpha,\Theta}(t,T,x,y)$. This yields a regularization property from Lemma 3.8 when β satisfies (6.12). For $z\in D_2$ and a given $\delta_0>0$, we use again the partition $(T-\tau)^\beta|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)| \ge$ or $<\delta_0$. The case $(T-\tau)^\beta|(\mathbb{T}^{\alpha}_{T-t})^{-1}(y-R_{T,t}x)|\le\delta_0$ yields a regularization in time similarly to the previous proof.

In order for $(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|$ to exceed δ_0, we see that τ must be lower than $\tau_0 := T - \Big(\frac{\delta_0}{|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|}\Big)^{\frac{1}{\beta}}$. In that case, the time singularity is still logarithmically explosive but integrable. We are led to consider:

$$\begin{aligned}\Gamma := &\int_{I_2} d\tau \frac{1}{T-\tau}\mathbf{1}_{\tau\le\tau_0}\int_{D_2} \delta\wedge\{(\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge 1)}\\ &\times \bar p_{\alpha,\Theta}(t,\tau,x,z)\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz.\end{aligned}\tag{6.19}$$

Using iteratively the scaling Lemma 5.2 we derive:

$$\begin{aligned}&\frac{|y^1-R^1_{T,\tau}z|}{(T-t)^{\frac{1}{\alpha}}}+\frac{|y^2-R^2_{T,\tau}z|}{(T-t)^{1+\frac{1}{\alpha}}}\\ \ge\ & c_2|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,\tau}z)|\\ \ge\ & c_2C^{-1}\Big(|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x-R_{\tau,T}y)|-|(\mathbb{T}^\alpha_{T-t})^{-1}(z-R_{\tau,t}x)|\Big)\\ \ge\ & c_2\Big(C^{-1}|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x-R_{\tau,T}y)|-C^{-1}(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|\Big)\\ \ge\ & c_2\Big(C^{-1}-(T-\tau)^\beta\Big)|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{\tau,t}x-R_{\tau,T}y)|,\end{aligned}$$

where $c_2>0, C:=C(T)\ge 1$ and recalling that $z\in D_2$ for the last but one inequality. Thus, for T small enough and up to a modification of C, we have either $|y^1-R^1_{T,\tau}z|\ge C|R^1_{\tau,t}x-R^1_{\tau,T}y|$, or $|y^2-R^2_{T,\tau}z|\ge C(T-t)|R^1_{\tau,t}x-R^1_{\tau,T}y|$. In both cases, $\bar p_\alpha(\tau,T,z,y)\le \frac{C}{|R^1_{\tau,t}x-R^1_{\tau,T}y|^{2+\alpha}}\theta(|\mathbb{M}^{-1}_{T-t}(R^1_{\tau,t}x-R^1_{\tau,T}y)|)$. This yields from Proposition 3.3 $\bar p_\alpha(\tau,T,z,y)\le C\bar p_\alpha(t,T,x,y)$. In our current case, we then derive from (6.17) that:

$$\begin{aligned}&\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}\bar p_\alpha(\tau,T,z,y)\\ \le\ &\delta\wedge\Big((T-t)|(x-R_{t,T}y)^1|+|(x-R_{t,T}y)^2|\Big)^{\eta(\alpha\wedge1)}\bar p_\alpha(t,T,x,y).\end{aligned}$$

Consequently, we can bound (6.19) by:

$$\begin{aligned}\Gamma\ \le\ & C(T-t)^\omega\bar p_\alpha(t,T,x,y)\\ &+\int_{I_2} d\tau\frac{1}{T-\tau}\mathbf{1}_{\tau\le\tau_0}\int_{D_2}\delta\wedge\{(\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge1)}\\ &\times\bar p_{\alpha,\Theta}(t,\tau,x,z)\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}\breve p_\alpha(\tau,T,z,y)dz\\ :=\ &\Gamma_1+\Gamma_2.\end{aligned}$$

It thus remains to handle Γ_2 which derives from the rediagonalization. We write:

$$\begin{aligned}
\Gamma_2 &\le C\bar p_{\alpha,\Theta}(t,T,x,y)\int_t^T d\tau \int_{\mathbb{R}^2} \bar p_{\alpha,\Theta}(t,\tau,x,z)\\
&\quad \times\Big(\delta\wedge\Big((\tau-t)|(z-R_{\tau,t}x)^1|\Big)^{\eta(\alpha\wedge 1)} + \delta\wedge|(z-R_{\tau,t}x)^2|^{\eta(\alpha\wedge1)}\Big)\\
&\quad \times\frac{1}{(T-\tau)^{1+\frac1\alpha}}\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}}{(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac1\alpha}})^{1+\frac1\alpha}}dz\\
&\le C\bar p_{\alpha,\Theta}(t,T,x,y)\int_t^T d\tau\int dz_2\frac{1}{(T-\tau)^{1+\frac1\alpha}}\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge1)}}{(1+\frac{|(z-R_{\tau,T}y)^2|}{(T-\tau)^{1+\frac1\alpha}})^{1+\frac1\alpha}}\\
&\quad \times\Big((\tau-t)^{(1+\frac1\alpha)[\eta(\alpha\wedge1)-1]}+\frac{\delta\wedge|(R_{\tau,t}x-z)^2|^{\eta(\alpha\wedge1)}}{(\tau-t)^{1+\frac1\alpha}(1+\frac{|(z-R_{\tau,t}x)^2|}{(\tau-t)^{1+\frac1\alpha}})^{1+\alpha}}\Big)
\end{aligned}$$

proceeding as in (6.18). Now, from Lemma 3.8, we obtain:

$$\begin{aligned}
\Gamma_2 &\le C\bar p_{\alpha,\Theta}(t,T,x,y)\int_t^T d\tau\Big((T-\tau)^{(1+\frac1\alpha)\eta(\alpha\wedge1)}(\tau-t)^{(1+\frac1\alpha)[\eta(\alpha\wedge1)-1]}\\
&\quad +(\tau-t)^{(1+\frac1\alpha)\{\eta(\alpha\wedge1)-1/2\}}(T-\tau)^{(1+\frac1\alpha)\{\eta(\alpha\wedge1)-1/2\}}\Big)\\
&\le C\bar p_{\alpha,\Theta}(t,T,x,y)(T-t)^{2(1+\frac1\alpha)\eta(\alpha\wedge1)-\frac1\alpha},
\end{aligned}$$

This indeed gives a regularizing effect recalling that we have assumed $1\ge\eta>\frac{1}{(\alpha\wedge1)(1+\alpha)}$. Note that the case $\tau\in I_1$ could be handled similarly, see Lemma 6.2. The controls become:

$$\begin{aligned}
q_{\alpha,\Theta}\otimes|H|(t,T,x,y) &\le C(T-t)^{\omega}\Big(\bar p_{\alpha,\Theta}(t,T,x,y) \qquad (6.20)\\
&\quad +\log\Big(K\vee|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\Big)q_{\alpha,\Theta}(t,T,x,y)\Big).
\end{aligned}$$

We point out that the important contribution in the above equation is the factor $(T-t)^\omega$, whose power will grow at each iteration. This key feature gives the convergence of the series (3.14).

Now, for $\rho_{\alpha,\Theta}\otimes|H|(t,T,x,y)$, we still bound $|H|$ using Lemma 3.7:

$$\begin{aligned}
\rho_{\alpha,\Theta}\otimes|H|(t,T,x,y) &\le C\int_t^T d\tau\int_{\mathbb{R}^2}dz\log\Big(K\vee|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\Big)\\
&\quad \times\delta\wedge\Big((\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|\Big)^{\eta(\alpha\wedge1)}\\
&\quad \times\bar p_{\alpha,\Theta}(t,\tau,x,z)\frac{\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge1)}}{T-\tau}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y).
\end{aligned}$$

W.r.t. the previous contribution, the main difference comes from the logarithm. However, the lines of the proof remain the same. Suppose first that $|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)| \le K$. Depending on the time parameter τ, we can show that we always have either $\bar p_{\alpha,\Theta}(t,\tau,x,z) \le C\bar p_{\alpha,\Theta}(t,T,x,y)$ or $\bar p_\alpha(\tau,T,z,y) \le C\bar p_\alpha(t,T,x,y) \le C\bar p_{\alpha,\Theta}(t,T,x,y)$. The second case occurs when $\tau \in I_1$. Using the notations of the previous proof, this yields:

$$\begin{aligned}\rho_{\alpha,\Theta} \otimes_{|I_1} |H|(t,T,x,y) \le C\bar p_{\alpha,\Theta}(t,T,x,y) \int_{I_1} d\tau \int_{\mathbb{R}^2} \log\Big(K \vee |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\Big)\\ \times \frac{\delta \wedge \{(\tau-t)|(z-R_{\tau,t}x)^1| + |(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge 1)}}{\tau-t} \bar p_{\alpha,\Theta}(t,\tau,x,z)dz,\end{aligned}$$

and we conclude by Lemma 6.1. In the case when $\bar p_{\alpha,\Theta}(t,\tau,x,z) \le C\bar p_{\alpha,\Theta}(t,T,x,y)$, which happens for $\tau \in I_2$, we have:

$$\begin{aligned}|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| &\le C(|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,t}x)| + |(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|)\\ &\le C(K + |(\mathbb{T}^\alpha_{T-\tau})^{-1}(y-R_{T,\tau}z)|).\end{aligned}$$

Plugging this inequality into the logarithm and taking out the first density, we can bound:

$$\begin{aligned}\rho_{\alpha,\Theta} \otimes_{|I_2} |H|(t,T,x,y) &\le C\bar p_{\alpha,\Theta}(t,T,x,y) \int_{I_2} d\tau \int_{\mathbb{R}^2} \log\Big(K \vee |(\mathbb{T}^\alpha_{T-\tau})^{-1}(y-R_{T,\tau}z)|\Big)\\ &\quad \times \frac{\delta \wedge \{(T-\tau)|(z-R_{\tau,T}y)^1| + |(z-R_{\tau,T}y)^2|\}^{\eta(\alpha\wedge 1)}}{T-\tau}\\ &\quad \times (\bar p_{\alpha,\Theta} + \breve p_\alpha)(\tau,T,z,y)dz,\end{aligned}$$

and once again, we conclude by Lemma 6.1. Thus, we have so far managed to show that in the global diagonal regime,

$$\rho_{\alpha,\Theta} \otimes |H|(t,T,x,y) \le C(T-t)^\omega \bar p_{\alpha,\Theta}(t,T,x,y).$$

It remains to deal with the case when $|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)| \ge K$. Suppose first that $\tau \in I_2$, and that the first component dominates in the global action $|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|$, i.e.

$$|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)| \asymp \frac{|y^1 - R^1_{T,t}x|}{(T-t)^{1/\alpha}}.$$

We still consider the partition in (6.10). When $z \in D_1$, we can bound

$$\bar p_{\alpha,\Theta}(t,\tau,x,z) \le C(T-\tau)^{-\beta(2+\alpha)} \bar p_{\alpha,\Theta}(t,T,x,y). \tag{6.21}$$

On the other hand, the triangle inequality and the scaling Lemma 5.2 yield:

$$|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| \le C\Big(|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)| + |(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,t}x)|\Big).$$

Consequently, up to a modification of C, we have either:

$$|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| \le C|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,t}x)|$$

or

$$|(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| \le C|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|.$$

Define accordingly,

$$D_{1,1} = \{z \in D_1; |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| \le C|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,t}x)|\},$$

$$D_{1,2} = \{z \in D_1; |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)| \le C|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|\}.$$

Observe that with this definition, $D_{1,1}$ and $D_{1,2}$ is not a partition of D_1. However, $D_1 \subset D_{1,1} \cup D_{1,2}$.

When $z \in D_{1,1}$, we can bound

$$\log\Big(K \vee |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\Big) \le \log\Big(K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\Big) + C.$$

On the other hand, for $\tau \in I_2$, we get from the definition of $D_{1,1}$:

$$\begin{aligned}&\delta \wedge \{(\tau-t)|(z-R_{\tau,t}x)^1| + |(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge 1)}\\ &\le C(\delta \wedge \{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)}).\end{aligned}$$

From (6.21), we thus have:

$$\begin{aligned}
& \rho_{\alpha,\Theta} \otimes_{|I_2,D_{1,1}} |H|(t,T,x,y)\\
\le\ & C\Big(\log\Big(K \vee |(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\Big) + 1\Big)\\
& \delta \wedge \{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)}\\
& \times \int_{I_2} d\tau \int_{D_{1,1}} \bar p_{\alpha,\Theta}(t,\tau,x,z)\frac{\delta \wedge |(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}}{T-\tau}(\bar p_\alpha + \breve p_\alpha)(\tau,T,z,y)dz\\
\le\ & (T-t)^\omega(\rho_\alpha + q_\alpha)(t,T,x,y),
\end{aligned}$$

choosing β satisfying (6.12).

When $z \in D_{1,2}$, we can bound:

$$\log\Big(K \vee |(\mathbb{T}^\alpha_{\tau-t})^{-1}(z-R_{\tau,t}x)|\Big) \le \log\Big(K \vee |(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|\Big) + C.$$

Bounding also roughly $\delta \wedge \{(\tau-t)|(z-R_{\tau,t}x)^1|+|(z-R_{\tau,t}x)^2|\}^{\eta(\alpha\wedge 1)} \le \delta$, and using the bound (6.21), we can write:

$$
\begin{aligned}
&\rho_{\alpha,\Theta}\otimes_{|I_2,D_{1,2}}|H|(t,T,x,y)\\
\le\ & C\int_{I_2} d\tau \int_{D_{1,2}} \bar p_{\alpha,\Theta}(t,\tau,x,z)\left(\log\left(K\vee|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|\right)+1\right)\\
&\times\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}}{T-\tau}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz\\
\le\ & C\bar p_{\alpha,\Theta}(t,T,x,y)\int_{I_2} d\tau(T-\tau)^{-\beta(2+\alpha)}\\
&\times\int_{D_{1,2}}\left(\log\left(K\vee|(\mathbb{T}^\alpha_{\tau-t})^{-1}(y-R_{T,\tau}z)|\right)+1\right)\\
&\times\frac{\delta\wedge|z-R_{\tau,T}y|^{\eta(\alpha\wedge 1)}}{T-\tau}(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz.
\end{aligned}
$$

Thus, using Lemma 6.1, we have

$$
\rho_{\alpha,\Theta}\otimes_{|I_2,D_{1,2}}|H|(t,T,x,y)\le (T-t)^\omega \bar p_\alpha(t,T,x,y).
$$

We have to deal with $z\in D_2$. In this case, **and because** $d=1$, we have

$$
\bar p_\alpha(\tau,T,z,y)\le C\bar p_\alpha(t,T,x,y).
$$

As above, we split for a given $\delta_0>0$, the time interval I_2 in $(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\ge\delta_0$ and $(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|<\delta_0$.

Assume first that $(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\le\delta_0$. Then, taking $\delta_0\le K$ gives that the first density is diagonal. Hence, the logarithm part disappears, and we have to deal with:

$$
\begin{aligned}
\rho_{\alpha,\Theta}\otimes_{I_2,D_2}|H|(t,T,x,y) &\le C\int_{\tau_0}^T d\tau\int_{D_2}\frac{1}{(T-t)^{\frac{2}{\alpha}+1}}\frac{\delta\wedge|(z-R_{\tau,T}y)^2|^{\eta(\alpha\wedge 1)}}{T-\tau}\\
&\quad\times(\bar p_\alpha+\breve p_\alpha)(\tau,T,z,y)dz\\
&\overset{Lemma\ 3.8}{\le}\frac{\delta_0^{2+\alpha}}{(T-t)^{\frac{2}{\alpha}+1}|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|^{2+\alpha}}\\
&\quad\times\int_{\tau_0}^T d\tau(T-\tau)^{(1+\frac{1}{\alpha})\eta(\alpha\wedge 1)-1-\beta(2+\alpha)}\\
&\le (T-t)^\omega\bar p_\alpha(t,T,x,y).
\end{aligned}
$$

Finally, we have to deal with the case $(T-\tau)^\beta|(\mathbb{T}^\alpha_{T-t})^{-1}(y-R_{T,t}x)|\ge\delta_0$. Observe that, on I_2, this imposes that $\tau\in[\frac{T+t}{2},\tau_0]$, with τ_0 defined above. In the considered set, we have from (6.16):

$$\begin{aligned}|(z-R_{\tau,T}y)^2| &\le |(z-R_{\tau,t}x)^2| + C\{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}\\ &\le C(1+(T-\tau)^\beta)\{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}.\end{aligned}$$

Plugging this estimate into the convolution and recalling for $z \in D_2$, $\bar p_\alpha(\tau, T, z, y) \le C\bar p_\alpha(t, T, x, y)$, we obtain from Lemma 6.1 and the previous controls for the contribution in $\breve p_\alpha$:

$$\begin{aligned}&\rho_{\alpha,\Theta} \otimes_{I_2,D_2} |H|(t,T,x,y) \le C\bar p_{\alpha,\Theta}(t,T,x,y)\\ &\Bigg(\big(\delta \wedge \{(T-t)|(x-R_{t,T}y)^1| + |(x-R_{t,T}y)^2|\}^{\eta(\alpha\wedge 1)}\big)\\ &\qquad \times \int_{\frac{T+t}{2}}^{\tau_0} d\tau \frac{1}{T-\tau}(\tau-t)^\omega + (T-t)^\omega\Bigg).\end{aligned}$$

Hence, integrating over τ yields the logarithmic contribution:

$$\rho_{\alpha,\Theta} \otimes_{I_2,D_2} |H|(t,T,x,y) \le C(T-t)^\omega \rho_{\alpha,\Theta}(t,T,x,y).$$

In order to complete the proof, we have to specify how to proceed in the remaining cases, that is when $\tau \in I_2$ and the second component dominates or when $\tau \in I_1$. When a fast component dominates, as we have seen in the previous proof, we can compensate the singularities brought by the kernel H, and conclude directly with Lemmas 3.8 and 6.1. When $\tau \in I_1$, we can adapt the previous strategy following the procedure described in Lemma 6.2. □

Using the previous lemmas, we get the following result.

Corollary 6.4. *Under the Assumptions of Lemma 6.3, there exists $C_{6.4} := 4C_{3.10} > 0$, s.t. for all $T \in (0, T_0]$, $T_0 = T_0(\textbf{[H]}) \le 1$, $(x, y) \in (\mathbb{R}^2)^2$, $t \in [0, T)$, $\forall k \in \mathbb{N}$:*

$$\begin{aligned}|\tilde p_\alpha \otimes H^{(2k)}(t,T,x,y)| &\le C_{6.4}^{2k}(T-t)^{k\omega}\Big((T-t)^{k\omega}\bar p_{\alpha,\Theta}(t,T,x,y)\\ &\quad +(\bar p_{\alpha,\Theta} + \bar q_{\alpha,\Theta})(t,T,x,y)\Big)\\ |\tilde p_\alpha \otimes H^{(2k+1)}(t,T,x,y)| &\le C_{6.4}^{2k+1}(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}\bar p_{\alpha,\Theta}\\ &\quad +(T-t)^\omega(\bar p_{\alpha,\Theta} + \bar q_{\alpha,\Theta}) + \bar q_{\alpha,\Theta}\Big)(t,T,x,y).\end{aligned}$$

Proof. We prove the estimate by induction. The idea is to use the controls of Lemmas 6.2 and 6.3 gathered in Lemma 3.10 to get from an estimate to the following one. The bounds may not be very precise, as we will sometimes bound $(T-t)^{k\omega} \le 1$, but they are sufficient to prove the convergence of the Parametrix series (3.14).

Initialization:
Since $(T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta}) \ge 0$, we clearly have:

$$|\tilde{p}_\alpha \otimes H(t,T,x,y)| \le C_{3.10}\Big((T-t)^{\omega}\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta} + (T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})\Big)(t,T,x,y).$$

Now, using Lemmas 6.2 and 6.3, we have:

$$\begin{aligned}
& |\tilde{p}_\alpha \otimes H^{(2)}(t,T,x,y)| \\
\le\ & C_{3.10}\Big((T-t)^{\omega}|\bar{p}_{\alpha,\Theta} \otimes H| + |\bar{q}_{\alpha,\Theta} \otimes H|\Big)(t,T,x,y) \\
\le\ & C_{3.10}\Big(C_{3.10}(T-t)^{2\omega}\bar{p}_{\alpha,\Theta} + C_{3.10}(T-t)^{\omega}\bar{q}_{\alpha,\Theta} + C_{3.10}(T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})\Big)(t,T,x,y) \\
\le\ & (2C_{3.10})^2(T-t)^{\omega}\Big((T-t)^{\omega}\bar{p}_{\alpha,\Theta} + (\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})\Big)(t,T,x,y).
\end{aligned}$$

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

$$\begin{aligned}
& |\tilde{p}_\alpha \otimes H^{(2k+1)}|(t,T,x,y) \\
\le\ & (4C_{3.10})^{2k}(T-t)^{k\omega}\Big((T-t)^{k\omega}|\bar{p}_{\alpha,\Theta} \otimes H|(t,T,x,y) + |(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta}) \otimes H|(t,T,x,y)\Big) \\
\le\ & (4C_{3.10})^{2k}(T-t)^{k\omega}\Big(C_{3.10}(T-t)^{k\omega}((T-t)^{\omega}\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})(t,T,x,y) \\
& + C_{3.10}((T-t)^{\omega}\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})(t,T,x,y) + C_{3.10}(T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta})(t,T,x,y)\Big).
\end{aligned}$$

Recalling that $T-t \le 1$, we have $(T-t)^{k\omega}\bar{q}_{\alpha,\Theta} \le (T-t)^{\omega}\bar{q}_{\alpha,\Theta}$. Thus:

$$\begin{aligned}
|\tilde{p}_\alpha \otimes H^{(2k+1)}|(t,T,x,y) \ \le\ & (4C_{3.10})^{2k}(T-t)^{k\omega}\Big(C_{3.10}(T-t)^{(k+1)\omega}\bar{p}_{\alpha,\Theta} \\
& + 2C_{3.10}(T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta}) + C_{3.10}\bar{q}_{\alpha,\Theta})\Big)(t,T,x,y) \\
\le\ & (4C_{3.10})^{2k}(2C_{3.10})(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}\bar{p}_{\alpha,\Theta} \\
& + (T-t)^{\omega}(\bar{p}_{\alpha,\Theta} + \bar{q}_{\alpha,\Theta}) + \bar{q}_{\alpha,\Theta})\Big)(t,T,x,y),
\end{aligned}$$

which gives the announced estimate.

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

$$\begin{aligned}
&|\tilde{p}_\alpha \otimes H^{(2k+2)}(t,T,x,y)| \\
\leq\ & (4C_{3.10})^{2k+1}(T-t)^{k\omega}\Big((T-t)^{(k+1)\omega}|\bar{p}_{\alpha,\Theta}\otimes H| \\
&+(T-t)^{\omega}|(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})\otimes H| + |\bar{q}_{\alpha,\Theta}\otimes H|\Big)(t,T,x,y) \\
\leq\ & (4C_{3.10})^{2k+1}(T-t)^{k\omega}\Big(C_{3.10}(T-t)^{(k+1)\omega}[(T-t)^{\omega}\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta}] \\
&+C_{3.10}(T-t)^{\omega}[\{(T-t)^{\omega}\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta}\}+C_{3.10}(T-t)^{\omega}(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})] \\
&+C_{3.10}(T-t)^{\omega}(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})\Big)(t,T,x,y) \\
\leq\ & (4C_{3.10})^{2k+2}(T-t)^{(k+1)\omega}\Big((T-t)^{(k+1)\omega}\bar{p}_{\alpha,\Theta}+(\bar{p}_{\alpha,\Theta}+\bar{q}_{\alpha,\Theta})\Big)(t,T,x,y),
\end{aligned}$$

where to get to the last equation, we used the fact that $(T-t)^{\omega}\bar{p}_{\alpha,\Theta} \leq \bar{p}_{\alpha,\Theta}$, and $(T-t)^{k\omega}\bar{q}_{\alpha,\Theta} \leq \bar{q}_{\alpha,\Theta}$. □

7 Proof of the diagonal lower bound for the frozen density.

In this section we prove the diagonal lower bound for the frozen density. Recall from Proposition 5.3, that the frozen density p_{Λ_s} writes for all $z \in \mathbb{R}^{nd}$ as:

$$\begin{aligned}
p_{\Lambda_s}(z) &= \frac{\det(\mathbb{M}_{s-t})^{-1}}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} e^{-i\langle q,(\mathbb{M}_{s-t})^{-1}z\rangle} \\
&\quad\times\exp\left(-(s-t)\int_{\mathbb{R}^{nd}}\{1-\cos(\langle q,\xi\rangle)\}\nu_S(d\xi)\right)dq \\
&= \frac{\det(\mathbb{T}^{\alpha}_{s-t})^{-1}}{(2\pi)^{nd}}\int_{\mathbb{R}^{nd}} e^{-i\langle q,(\mathbb{T}^{\alpha}_{s-t})^{-1}z\rangle} \\
&\quad\times\exp\left(-(s-t)\int_{\mathbb{R}^{nd}}\{1-\cos(\langle\frac{q}{(s-t)^{\frac{1}{\alpha}}},\xi\rangle)\}\nu_S(d\xi)\right)dq.
\end{aligned}$$

The complex exponential can be written as a cosine. Denoting $\bar{x}$ the projection of $x \in \mathbb{R}^{nd}$ on the sphere, we change variable to the polar coordinates by setting $q = |q|\bar{q}$, where $(|q|,\bar{q}) \in \mathbb{R}^+ \times S^{nd-1}$. Also, we take a parametrization of the sphere by setting $\bar{q} = (\theta,\phi) \in [0,\pi] \times S^{nd-2}$, along the axis defined by $(\mathbb{T}^{\alpha}_{s-t})^{-1}z$. Set finally $\tau = \cos(\theta)$, the density writes:

$$\begin{aligned} p_{\Lambda_s}(z) &= \frac{\det(\mathbb{T}^\alpha_{s-t})^{-1}}{(2\pi)^{nd}} \int_0^{+\infty} d|q||q|^{nd-1} \int_{-1}^1 d\tau(1-\tau^2)^{\frac{nd-3}{2}} \int_{S^{nd-2}} d\phi \\ &\times \cos(|q||(\mathbb{T}^\alpha_{s-t})^{-1}z|\tau) \\ &\times \exp\left(-(s-t)\int_{\mathbb{R}^{nd}} \{1-\cos(\langle \frac{\bar{q}|q|}{(s-t)^{1/\alpha}},\xi\rangle)\}\nu_S(d\xi)\right). \end{aligned} \tag{7.1}$$

The idea is the following: since $|(\mathbb{T}^\alpha_{s-t})^{-1}z|$ is small, we can expand the cosine and show that the first term is positive, giving the two-sided diagonal estimate. We focus on the diagonal lower bound.

Proposition 7.1 (**Diagonal Lower bound**). *For K sufficiently small, there exists C_K s.t. for all $z \in \mathbb{R}^{nd}$, $|(\mathbb{T}^\alpha_{s-t})^{-1}z| \le K$:*

$$p_{\Lambda_s}(z) \ge C_K \det(\mathbb{T}^\alpha_{s-t})^{-1}.$$

Proof. There is no difference with the non degenerate case for the diagonal expansion, see [Kol00b]. For small $|(\mathbb{T}^\alpha_{s-t})^{-1}z|$, we expand $\cos(|q||(\mathbb{T}^\alpha_{s-t})^{-1}z|\tau)$ using Taylor's formula in equation (7.1):

$$\begin{aligned} p_{\Lambda_s}(z) &= \frac{\det(\mathbb{T}^\alpha_{s-t})^{-1}}{(2\pi)^{nd}} \int_0^{+\infty} d|q||q|^{nd-1} \int_{-1}^1 d\tau(1-\tau^2)^{\frac{nd-3}{2}} \qquad (7.2) \\ &\times \left(\sum_{k=0}^N \frac{(-1)^k}{(2k)!}|q|^{2k}|(\mathbb{T}^\alpha_{s-t})^{-1}z|^{2k}\tau^{2k} + \tilde{R}_N(|(\mathbb{T}^\alpha_{s-t})^{-1}z|)\right) \\ &\times \int_{S^{nd-2}} d\phi \exp\left(-(s-t)\int_{\mathbb{R}^{nd}} \{1-\cos(\langle \frac{\bar{q}|q|}{(s-t)^{1/\alpha}},\xi\rangle)\}\nu_S(d\xi)\right) \\ &= \frac{\det(\mathbb{T}^\alpha_{s-t})^{-1}}{(2\pi)^{nd}} \sum_{k=0}^N \frac{(-1)^k}{(2k)!}|(\mathbb{T}^\alpha_{s-t})^{-1}z|^{2k} \int_0^{+\infty} d|q||q|^{2k+nd-1}\int_{-1}^1 d\tau(1-\tau^2)^{\frac{nd-3}{2}}\tau^{2k} \\ &\times \int_{S^{nd-2}} d\phi \exp\left(-(s-t)\int_{\mathbb{R}^{nd}} \{1-\cos(\langle \frac{\bar{q}|q|}{(s-t)^{1/\alpha}},\xi\rangle)\}\nu_S(d\xi)\right) \\ &\quad + R_N(|(\mathbb{T}^\alpha_{s-t})^{-1}z|). \end{aligned}$$

The estimate on the coefficient also serves to estimate $R_N(|(\mathbb{T}^\alpha_{s-t})^{-1}z|)$. To bound the coefficient, we use the domination condition in (5.6) and the property that g is non-increasing:

$$\begin{aligned}
&\exp\left(-(s-t)\int_{\mathbb{R}^{nd}}\{1-\cos(\langle\frac{\bar q|q|}{(s-t)^{1/\alpha}},\xi\rangle)\}\nu_S(d\xi)\right)\\
\geq\ &\exp\left(-(s-t)g(0)\int_{\mathbb{R}^+}\frac{d\rho}{\rho^{1+\alpha}}\int_{S^{nd-1}}\{1-\cos(\langle\frac{\bar q|q|}{(s-t)^{1/\alpha}},\rho\eta\rangle)\bar\mu(d\eta)\}\right)\\
=\ &\exp\left(-g(0)\int_{\mathbb{R}^+}\frac{d\rho}{\rho^{1+\alpha}}\int_{S^{nd-1}}\{1-\cos(\langle\bar q|q|,\rho\eta\rangle)\}\bar\mu(d\eta)\}\right)\\
=\ &\exp\left(-c_\alpha g(0)|q|^\alpha\int_{S^{nd-1}}|\langle\bar q,\eta\rangle|^\alpha\bar\mu(d\eta)\right)\\
\geq\ &\exp\left(-\bar c|q|^\alpha\right),\bar c:=\bar c(\alpha,\mathbf{[H]})\geq 1,
\end{aligned}$$

using that $\bar\mu$ satisfies **[H-4]** for the last inequality. The above control can be used to give a lower bound for the even terms in the previous expansion (7.2). On the other hand, similarly to the proof of Proposition 5.3, we get

$$\exp\left(-(s-t)\int_{\mathbb{R}^{nd}}\{1-\cos(\langle\frac{\bar q|q|}{(s-t)^{1/\alpha}},\xi\rangle)\}\nu_S(d\xi)\right)\leq\exp\left(-\bar c^{-1}|q|^\alpha\right),$$

for $|q|>1$, which can be used to derive lower bound for the odd terms of the expansion (7.2). Note that the coefficient $a_k(\overline{(\mathbb{T}^\alpha_{s-t})^{-1}z})$ depends on $\overline{(\mathbb{T}^\alpha_{s-t})^{-1}z}$ because of the choice of the parametrization of the sphere S^{nd-2}.

□

8 Off-diagonal Estimates on the Kernel H.

We thoroughly exploit the decomposition of the density used by Watanabe [Wat07] in the stable case followed by Sztonyk [Szt10] in the tempered one. From the identity (5.13), we have:

$$\forall z\in\mathbb{R}^{nd},\ p_{\Lambda_T}(z)=\det(\mathbb{M}^\alpha_{T-t})^{-1}p_S(T-t,(\mathbb{M}^\alpha_{s-t})^{-1}z),\tag{8.1}$$

where $(S_u)_{u\geq 0}$ has Lévy measure ν_S.

For a fixed $T-t$ we can write $S_{T-t}=M_{T-t}+N_{T-t}$ where $(M_u)_{u\geq 0}$ and $(N_u)_{u\geq 0}$ are two independent processes with respective generators:

$$\begin{aligned}
\mathcal{L}^M\varphi(z)&=\int_{\mathbb{R}^{nd}}(\varphi(z+\xi)-\varphi(z)-\frac{\langle\nabla\varphi(z),\xi\rangle}{1+|\xi|^2})\mathbf{1}_{|\xi|\leq(T-t)^{1/\alpha}}\nu_S(d\xi),\\
\mathcal{L}^N\varphi(z)&=\int_{\mathbb{R}^{nd}}(\varphi(z+\xi)-\varphi(z)-\frac{\langle\nabla\varphi(z),\xi\rangle}{1+|\xi|^2})\mathbf{1}_{|\xi|>(T-t)^{1/\alpha}}\nu_S(d\xi),
\end{aligned}$$

for all $z \in \mathbb{R}^{nd}$ and $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$. We have separated the jumps that are at the typical scales, i.e. $(T-t)^{1/\alpha}$, from the big ones which induce a compound Poisson process. It can be proved similarly to Proposition 5.3 that M_{T-t} has a density, intuitively the *small* jumps generate the density. We therefore disintegrate $p_S(T-t,.)$ in the following way:

$$\forall z \in \mathbb{R}^{nd},\ p_S(T-t,z) = \int_{\mathbb{R}^{nd}} p_M(T-t, z-\bar{z}) P_{N_{T-t}}(d\bar{z}), \tag{8.2}$$

where $P_{N_{T-t}}$ stands for the law of N_{T-t}. Now, the following properties hold for the Lévy-Itô decomposition.

Lemma 8.1 (Density estimate on the Martingale part and associated derivatives.)**.** *For all $m \geq 1$, there exists $C_m \geq 1$ s.t. for all $T-t > 0, z \in \mathbb{R}^{nd}$,*

$$p_M(T-t,z) \leq C_m (T-t)^{-nd/\alpha} \left(1 + \frac{|z|}{(T-t)^{1/\alpha}}\right)^{-m}.$$

Also, for all $m \geq 1$ and all multi-index β, $|\beta| \leq 2$,

$$|\partial_z^\beta p_M(T-t,z)| \leq C_m (T-t)^{-(nd+|\beta|)/\alpha} \left(1 + \frac{|z|}{(T-t)^{1/\alpha}}\right)^{-m}.$$

Proof. Similarly to the proof of Proposition 5.3 we write:

$$\begin{aligned} p_M(T-t,z) &= \frac{1}{(2\pi)^{nd}} \int_{\mathbb{R}^{nd}} dp e^{-i\langle p,z\rangle} \\ &\times \exp\left(-(T-t)\int_{\mathbb{R}^{nd}} \{1-\cos(\langle p,\xi\rangle)\} \mathbf{1}_{\{|\xi| \leq (T-t)^{1/\alpha}\}} \nu_S(d\xi)\right). \end{aligned}$$

Changing variables in $(T-t)^{1/\alpha} p = q$ yields:

$$\begin{aligned} p_M(T-t,z) &= \frac{1}{(2\pi)^{nd}} (T-t)^{-nd/\alpha} \int_{\mathbb{R}^{nd}} dq e^{-i\langle q, \frac{z}{(T-t)^{1/\alpha}}\rangle} \\ &\times \exp\left(-(T-t)\int_{\mathbb{R}^{nd}} \{1-\cos\langle q, \frac{\xi}{(T-t)^{1/\alpha}}\rangle\} \mathbf{1}_{\{|\xi| \leq (T-t)^{1/\alpha}\}} \nu_S(d\xi)\right). \end{aligned} \tag{8.3}$$

Let us now denote

$$\hat{f}_{T-t}(q) := \exp\left((T-t)\int_{\mathbb{R}^{nd}} \{\cos\langle q, \frac{\xi}{(T-t)^{1/\alpha}}\rangle - 1\} \mathbf{1}_{|\xi| \leq (T-t)^{1/\alpha}} \nu_S(d\xi)\right).$$

Since the Lévy measure in the above expression has finite support, we get from Theorem 3.7.13 in Jacob [Jac05] that $\hat{f}_{T-t}$ is infinitely differentiable as a function of q.

Moreover,

$$
\begin{aligned}
|\partial_q \hat f_{T-t}(q)| &\le (T-t)\int_{\mathbb{R}^{nd}} \frac{|\xi|}{(T-t)^{\frac{1}{\alpha}}}|\sin(\langle q, \frac{\xi}{(T-t)^{1/\alpha}}\rangle)|\mathbf{1}_{|\xi|\le (T-t)^{\frac{1}{\alpha}}}\nu_S(d\xi)\\
&\quad \times \exp\left((T-t)\int_{\mathbb{R}^{nd}}\{\cos(\langle q,\frac{\xi}{(T-t)^{1/\alpha}}\rangle)-1\}\mathbf{1}_{|\xi|\le (T-t)^{1/\alpha}}\nu_S(d\xi)\right).
\end{aligned}
$$

Write now:

$$
\begin{aligned}
&(T-t)\int_{\mathbb{R}^{nd}} \frac{|\xi|}{(T-t)^{\frac{1}{\alpha}}}|\sin(\langle q, \frac{\xi}{(T-t)^{1/\alpha}}\rangle)|\mathbf{1}_{|\xi|\le (T-t)^{\frac{1}{\alpha}}}\nu_S(d\xi)\\
\le\ & C(T-t)\int_{r\le (T-t)^{1/\alpha}} dr\frac{r^{nd-1}}{r^{d+1+\alpha}}\frac{r}{(T-t)^{\frac{1}{\alpha}}}(\mathbf{1}_{\alpha<1}+\mathbf{1}_{\alpha\ge 1}|q|\frac{r}{(T-t)^{1/\alpha}})\\
\le\ & C(T-t)\int_{r\le (T-t)^{1/\alpha}} dr\frac{r^{-\alpha}}{(T-t)^{\frac{1}{\alpha}}}(\mathbf{1}_{\alpha<1}+\mathbf{1}_{\alpha\ge 1}|q|\frac{r}{(T-t)^{1/\alpha}}) \le C(1+|q|).
\end{aligned}
$$

Thus there is $C\ge 1$:

$$
\begin{aligned}
|\partial_q \hat f_{T-t}(q)| &\le C(1+|q|)\exp\left((T-t)\int_{\mathbb{R}^{nd}}\{\cos(\langle q,\frac{\xi}{(T-t)^{1/\alpha}}\rangle)-1\}\nu_S(d\xi)\right)\\
&\quad \times\exp(2(T-t)\nu_S(B(0,(T-t)^{1/\alpha})^c)) \le C(1+|q|)\exp(-C^{-1}|q|^\alpha),
\end{aligned}
$$

since from (5.6), $\nu_S(B(0,(T-t)^{1/\alpha})^c)\le C/(T-t)$ and that the proof of Proposition 5.3 also yields that

$$
\exp\left((T-t)\int_{\mathbb{R}^{nd}}\{\cos(\langle q,\frac{\xi}{(T-t)^{1/\alpha}}\rangle)-1\}\nu_S(d\xi)\right)\le C\exp(-C^{-1}|q|^\alpha).
$$

Similarly, for all $l\in\mathbb{N}$:

$$
|\partial_q^l \hat f_{T-t}(q)| \le C_l(1+|q|^l)\exp(-C^{-1}|q|^\alpha),\ C_l\ge 1.
$$

Thus, $\hat f_{T-t}$ belongs the Schwartz space. Denoting by f_{T-t} its Fourier transform, we have:

$$
\forall m\ge 0,\ \forall z\in\mathbb{R}^{nd}, \exists C_m\ge 1 s.t. : |f_{T-t}(z)|\le C_m(1+|z|)^{-m}.
$$

Now since $p_M(T-t,z)=(T-t)^{-nd/\alpha}f_{T-t}(z/(T-t)^{1/\alpha})$, the announced bound follows. The control concerning the derivatives is derived similarly.

□

Besides, the following control holds for the Poisson measure.

Lemma 8.2 (Controls for the Poisson measure). *For all $T-t>0$, $P_{N_{T-t}}$ is a Poisson measure. Since* $\dim(\mathrm{supp}(\bar\mu)) = d$ *we have the following estimates. There exists a constant $C>0$ s.t. For all $z\in\mathbb{R}^{nd}, r>0$:*

$$P_{N_{T-t}}(B(z,r)) \le \frac{C}{\theta((T-t)^{1/\alpha})}(T-t)r^{d+1}\Big(1+\frac{r^\alpha}{T-t}\frac{\theta((T-t)^{1/\alpha})}{\theta(r)}\Big)|z|^{-(d+1+\alpha)}\theta(|z|). \tag{8.4}$$

Proof. In the stable case, i.e. $\theta=1$, this result is a consequence of Lemma 3.1 in [Wat07] and the intrinsic stable scaling. In the tempered case, it follows from Corollary 6 in Sztonyk [Szt10]. □

Let us observe that the above control also yields the upper-bound estimate for the density in Propositions 3.3, 3.4 in the off-diagonal regime. Precisely from (8.1), (8.2), Lemma 8.1 and (8.4) one gets:

$$\begin{aligned}
\tilde p_\alpha(t,T,x,y) &\le C_m \det(\mathbb{M}_{T-t})^{-1}(T-t)^{-nd/\alpha}\\
&\quad\times\int_{\mathbb{R}^{nd}}\left(1+\frac{|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)-\bar z|}{(T-t)^{1/\alpha}}\right)^{-m}P_{N_{T-t}}(d\bar z)\\
&\le C_m\det(\mathbb{T}^\alpha_{T-t})^{-1}\int_0^1 ds\\
&\quad\times P_{N_{T-t}}\left(\left\{\bar z\in\mathbb{R}^{nd}:\left(1+\frac{|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)-\bar z|}{(T-t)^{1/\alpha}}\right)^{-m}>s\right\}\right)\\
&\le C_m\det(\mathbb{T}^\alpha_{T-t})^{-1}\\
&\quad\times\int_0^1 P_{N_{T-t}}(B(\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y),s^{-1/m}(T-t)^{1/\alpha}))ds\\
&\le C_mC\det(\mathbb{T}^\alpha_{T-t})^{-1}\frac{(T-t)^{1+(d+1)/\alpha}}{\theta((T-t)^{1/\alpha})}\\
&\quad\times\int_0^1 s^{-(d+1)/m}\Big(1+s^{-\alpha/m}\frac{\theta((T-t)^{1/\alpha})}{\theta((T-t)^{1/\alpha}s^{-1/m})}\Big)ds\\
&\quad\times|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|^{-(d+1+\alpha)}\theta\left(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|\right)\\
&\le \frac{C_m}{\theta(1)}C\det(\mathbb{T}^\alpha_{T-t})^{-1}\frac{\theta\left(|\mathbb{M}_{T-t}^{-1}(R_{T,t}x-y)|\right)}{(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|)^{d+1+\alpha}}\\
&\quad\times\int_0^1[s^{-(d+1)/m}+s^{-(d+1+\alpha+\tilde\eta)/m}ds],
\end{aligned}$$

using for the last inequality that θ is non-increasing and exploiting that the doubling condition in **[T]** is equivalent to the fact that there exists $c>0,\ \tilde\eta\ge 0$ s.t.

$$\frac{\theta(r)}{\theta(R)}\le c\left(\frac{r}{R}\right)^{-\tilde\eta},\ 0<r\le R,$$

see e.g. [Bas95]. Choosing $m > d+1+\alpha+\tilde{\eta}$ then gives the result, i.e. there exists $C \geq 1$ s.t. for all $0 \leq t < T$, $(x,y) \in (\mathbb{R}^{nd})^2$, $\tilde{p}_\alpha(t,T,x,y) \leq C\bar{p}_\alpha(t,T,x,y)$.

Moreover, the previous procedure, associated with Lemma 8.1, allows to handle the small jumps in the estimation of

$$(L_t - \tilde{L}_t^{T,y})\tilde{p}_\alpha(t,T,x,y) = (L_t^M - \tilde{L}_t^{T,y,M})\tilde{p}_\alpha(t,T,x,y) + (L_t^N - \tilde{L}_t^N)\tilde{p}_\alpha(t,T,x,y).$$

Introducing

$$\nu(x,A) := \nu(\{z \in \mathbb{R}^d : \sigma(x)z \in A\}),$$

we write for a given parameter $a \in (0, 1 \wedge K)$ and $x \in \mathbb{R}^{nd}$:

$$\begin{aligned}
&(L_t^M - \tilde{L}_t^{T,y,M})\varphi(x)\\
= &\int_{\mathbb{R}^d} (\varphi(x+Bz) - \varphi(x) - \langle \nabla_{z^1}\varphi(x), z\rangle)\mathbf{1}_{|z|\leq a(T-t)^{1/\alpha}}(\nu(x,dz) - \nu(R_{t,T}y,dz)),\\
&(L_t^N - \tilde{L}_t^{T,y,N})\varphi(x)\\
= &\int_{\mathbb{R}^d} (\varphi(x+Bz) - \varphi(x))\mathbf{1}_{|z|>a(T-t)^{1/\alpha}}(\nu(x,dz) - \nu(R_{t,T}y,dz)).
\end{aligned}$$

The Lipschitz property of the density of the spectral measure μ in **[H-4]**, the non-degeneracy and Hölder continuity of σ and the properties concerning the tempering function in **[HT]** yield that $\nu(.,dz)$ is $\eta(\alpha \wedge 1)$ Hölder continuous w.r.t. its first parameter and there exists $C \geq 1$ s.t. uniformly in $z \in \mathbb{R}^d$,

$$|(\nu(x,dz) - \nu(R_{t,T}y,dz))| \leq C(\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)})\theta(|z|)|z|^{-(d+\alpha)}dz.$$

The condition that for all $r > 0$, $r\sup_{u\in[\kappa^{-1},\kappa]} g'(ur) \leq c\theta(r)$ appearing in **[HT]** is needed here to control the difference on the tempering functions. We now get:

$$\begin{aligned}
&|(L_t^M - \tilde{L}_t^{T,y,M})\tilde{p}_\alpha(t,T,x,y)|\\
= &\Bigg| \int_{\mathbb{R}^d} \big(\tilde{p}_\alpha(t,T,x+Bz,y) - \tilde{p}_\alpha(t,T,x,y) - \langle \nabla_{x^1} p_\alpha(t,T,x,y), z\rangle\big)\\
&(\nu(x,dz) - \nu(R_{t,T}y,dz))\Bigg|\\
\leq\ & C\det(\mathbb{M}_{T-t})^{-1}[\delta \wedge |x - R_{t,T}y|^{\eta(\alpha\wedge 1)}] \int_{\mathbb{R}^d} \mathbf{1}_{|z|\leq a(T-t)^{1/\alpha}}\theta(|z|)\frac{dz}{|z|^{d+\alpha}}\\
&\times\Bigg| \int_{\mathbb{R}^{nd}} \big\{p_M(T-t, \mathbb{M}_{T-t}^{-1}(R_{T,t}x + Bz - y) - \bar{z}) - p_M(T-t, \mathbb{M}_{T-t}^{-1}(R_{T,t}x - y) - \bar{z})\\
&-\langle \nabla_{x^1} p_M(T-t, \mathbb{M}_{T-t}^{-1}(R_{T,t}x - y) - \bar{z}), z\rangle \mathbf{1}_{\alpha\geq 1}\big\} P_{N_{T-t}}(d\bar{z})\Bigg|
\end{aligned}$$

The idea is now to perform a Taylor expansion on p_M to compensate the singularities in z. We assume for simplicity that $\alpha \in (0,1)$ which allows to perform the Taylor expansion at order 1 only. It suffices to expand at order 2 to handle the case $\alpha \in [1,2)$. We get from Lemma 8.1:

$$\begin{aligned}
& |(L_t^M - \tilde L_t^{T,y,M})\tilde p_\alpha(t,T,x,y)| \\
\le\ & C\det(\mathbb{M}_{T-t})^{-1}[\delta \wedge |x-R_{t,T}y|^{\eta(\alpha\wedge 1)}] \\
& \times \int_{\mathbb{R}^d} \mathbf{1}_{|z|\le a(T-t)^{1/\alpha}} \frac{dz}{|z|^{d+\alpha}} \int_{\mathbb{R}^{nd}} P_{N_{T-t}}(d\bar z) \\
& \times \sup_{|\tilde z|\in(0,a(T-t)^{1/\alpha}]} |\nabla_{x_1} p_M(T-t, \mathbb{M}_{T-t}^{-1}(R_{T,t}x + B\tilde z - y) - \bar z)||z| \\
\le\ & \frac{CC_m \det(\mathbb{T}^\alpha_{T-t})^{-1}}{(T-t)^{1/\alpha}}[\delta \wedge |x-R_{t,T}y|^{\eta(\alpha\wedge 1)}] \\
& \times \int_{\mathbb{R}^d} \mathbf{1}_{|z|\le a(T-t)^{1/\alpha}} |z|\frac{dz}{|z|^{d+\alpha}} \int_{\mathbb{R}^{nd}} P_{N_{T-t}}(d\bar z) \\
& \times \sup_{|\tilde z|\in(0,a(T-t)^{1/\alpha}]} \left(1 + \frac{|\mathbb{M}_{T-t}^{-1}(R_{T,t}x + B\tilde z - y) - \bar z)|}{(T-t)^{1/\alpha}}\right)^{-m} \\
\le\ & \frac{CC_m \det(\mathbb{T}^\alpha_{T-t})^{-1}}{(T-t)^{1/\alpha}}[\delta \wedge |x-R_{t,T}y|^{\eta(\alpha\wedge 1)}] \int_0^{at^{1/\alpha}} dr r^{-\alpha} \\
& \times \int_{\mathbb{R}^{nd}} \left((1-a) + \frac{|\mathbb{M}_{T-t}^{-1}(R_{T,t}x - y) - \bar z)|}{(T-t)^{1/\alpha}}\right)^{-m} P_{N_{T-t}}(d\bar z) \\
\le\ & \frac{C[\delta \wedge |x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]}{T-t}\bar p_\alpha(t,T,x,y). \qquad (8.5)
\end{aligned}$$

This therefore gives the expected control for the small jumps in the kernel, i.e. the operator $L_t^M - \tilde L_t^{T,y,M}$ acting on $\tilde p_\alpha(t,T,x,y)$ yields a bound homogeneous to the upper-bound $\bar p_\alpha(t,T,x,y)$ up to an additional multiplicative singularity of the form $\frac{C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge1)}]}{T-t}$.

The delicate part, yielding the *rediagonalization* phenomenon which might deteriorate the estimates in the degenerate framework, comes from the large jumps. We now specify how in the off-diagonal regime, when

$$\frac{|(x-R_{t,T}y)^1}{(T-t)^{1/\alpha}} \asymp |(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)| \overset{\text{Lemma 5.2}}{\asymp} |(\mathbb{T}^\alpha_{T-t})^{-1}(R_{T,t}x-y)|,$$

that is when the *slow* component dominates, a *bad* rediagonalization phenomenon can occur. Let us now discuss the various possible cases. Fix $\varepsilon > 0$.

- If $z \notin B\Big((R_{t,T}y-x)^1, \varepsilon|(x-R_{t,T}y)^1|\Big) := B_{\varepsilon,t,T,x,y}$, then

$$|(\mathbb{T}^\alpha_{T-t})^{-1}(x+Bz-R_{t,T}y)| \ge \frac{|z-(R_{t,T}y-x)^1|}{(T-t)^{1/\alpha}} \ge \varepsilon\frac{|(x-R_{t,T}y)^1|}{(T-t)^{1/\alpha}}.$$

Hence $\tilde p_\alpha(t,T,x+Bz,y)$ is off-diagonal and $\tilde p_\alpha(t,T,x+Bz,y) \le C\bar p_\alpha(t,T,x,y)$. Thus:

$$\begin{aligned}
&\int_{z\notin B_{\varepsilon,t,T,x,y}} |\{\tilde p_\alpha(t,T,x+Bz,y)-\tilde p_\alpha(t,T,x,y)\}|\mathbf{1}_{|z|>a(T-t)^{1/\alpha}}|\nu(x,dz)-\nu(R_{t,T}y,dz)|\\
\le\ & C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]\bar p_\alpha(t,T,x,y)\int_{\mathbb{R}^d}\mathbf{1}_{|z|>a(T-t)^{1/\alpha}}\frac{dz}{|z|^{d+\alpha}}\\
\le\ & \frac{C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]}{T-t}\bar p_\alpha(t,T,x,y). \qquad (8.6)
\end{aligned}$$

- If $z\in B_{\varepsilon,t,T,x,y}$ we can write:

$$\begin{aligned}
&\int_{z\in B_{\varepsilon,t,T,x,y}} |\tilde p_\alpha(t,T,x+Bz,y)-\tilde p_\alpha(t,T,x,y)|\mathbf{1}_{|z|>a(T-t)^{1/\alpha}}|\nu(x,dz)-\nu(R_{t,T}y,dz)|\\
\le\ & C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]\Bigg\{\frac{\theta(|(x-R_{t,T}y)^1|)}{|(x-R_{t,T}y)^1|^{d+\alpha}}\\
&\times\int_{z\in B_{\varepsilon,t,T,x,y}}\tilde p_\alpha(t,T,x+Bz,y)dz+\frac{\bar p_\alpha(t,T,x,y)}{T-t}\Bigg\}\\
\le\ & C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]\Bigg\{\frac{\bar p_\alpha(t,T,x,y)}{T-t}\\
&+\frac{\theta(|(x-R_{t,T}y)^1|)}{|(x-R_{t,T}y)^1|^{d+\alpha}}\int_{z\in B_{\varepsilon,t,T,x,y}}\frac{\det(\mathbb{T}^\alpha_{T-t})^{-1}}{(1+|(\mathbb{T}^\alpha_{T-t})^{-1}(x+Bz-R_{t,T}y)|)^{(d+1+\alpha)}}dz\Bigg\} \qquad (8.7)\\
\le\ & C[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}]\Bigg\{\frac{\bar p_\alpha(t,T,x,y)}{T-t}+\frac{\theta(|(x-R_{t,T}y)^1|)}{|(x-R_{t,T}y)^1|^{d+\alpha}}\\
&\times\frac{1}{(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}(1+|\{(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)\}^{2:n}|)^{1+\alpha}}\Bigg\}\\
\le\ & \frac{C}{T-t}[\delta\wedge|x-R_{t,T}y|^{\eta\alpha\wedge 1}](\bar p_\alpha+\breve p_\alpha)(t,T,x,y), \qquad (8.8)
\end{aligned}$$

using Propositions 3.3, 3.4 and Lemma 5.2 for the last but second inequality.

From (8.6) and (8.8) we derive:

$$|(L^N_t-\tilde L^{T,y,N}_t)\tilde p_\alpha(t,T,x,y)|\le\frac{C}{T-t}[\delta\wedge|x-R_{t,T}y|^{\eta(\alpha\wedge 1)}](\bar p_\alpha+\breve p_\alpha)(t,T,x,y),$$

which together with (8.5) gives the statement of Lemma 3.7 in the off-diagonal regime.

Remark 8.1 (About the rediagonalization)**.** Observe that a similar rediagonalization phenomenon occurs in the non-degenerate case as well. The fact is that, in that case we integrate a density in (8.7) and not a marginal. The decay of the jump measure gives in that case up to a multiplicative singularity in $(T-t)^{-1}$ the asymptotic behavior of the stable density. Namely when $n = 1$ we would have $d + \alpha$ instead of $d + 1 + \alpha$ in (8.7) and in the off-diagonal regime:

$$\begin{aligned}
\frac{1}{|x - R_{t,T}y|^{d+\alpha}} &= \frac{1}{T-t} \times \frac{T-t}{|x - R_{t,T}y|^{d+\alpha}} \\
&\leq \frac{C}{T-t} \frac{1}{(T-t)^{d/\alpha}\left(1 + \frac{|x-R_{t,T}y|}{(T-t)^{1/\alpha}}\right)^{d+\alpha}} \\
&:= \frac{C}{T-t}\bar{p}_\alpha(t,T,x,y),
\end{aligned}$$

where $\bar{p}_\alpha$ indeed corresponds to the upper bound for the large scale asymptotics of a stable process whose spectral measure is absolutely continuous, see again Proposition 3.3. In that framework, our proof provides an alternative to the Fourier arguments employed in [Kol00b].

Remark 8.2 (Loss of Concentration in the stable case)**.** From equations (8.7)-(8.8) we see that when $|\{(\mathbb{T}^\alpha_{T-t})^{-1}(x - R_{t,T}y)\}^{2:n}| \leq K$, i.e. the fast component in the backward dynamics are diagonal, we have a loss of concentration w.r.t. to the worst asymptotic bounds given in Proposition 3.3. Note also that in this case the lower bound in that proposition yields:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} \tilde{p}(t,T,x+Bz,y) - \tilde{p}(t,T,x,y)\mathbf{1}_{|z|>a(T-t)^{1/\alpha}} \frac{dz}{|z|^{d+\alpha}}\right| \geq -\frac{C}{(T-t)}\bar{p}_\alpha(t,T,x,y) \\
&+\frac{C^{-1}}{|(x-R_{t,T}y)^1|^{d+\alpha}} \frac{1}{(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}(1+|\{(\mathbb{T}^\alpha_{T-t})^{-1}(x-R_{t,T}y)\}^{2:n}|)^{nd(1+\alpha)-d}} \\
&\geq \frac{1}{(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}} \\
&\times\Big\{\frac{C^{-1}}{|(x-R_{t,T}y)^1|^{d+\alpha}(1+K)^{nd(1+\alpha)-d}} - \frac{C}{|(x-R_{t,T}y)^1|^{d+\alpha+1}}(T-t)^{1/\alpha}\Big\} \\
&\geq \frac{1}{2(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}} \frac{C^{-1}}{|(x-R_{t,T}y)^1|^{d+\alpha}(1+K)^{nd(1+\alpha)-d}},
\end{aligned}$$

if $|(x-R_{t,T}y)^1| \geq \bar{K}(T-t)^{1/\alpha}$ for $\bar{K}$ large enough. Hence, if $d = 1$ the previous bound is sharp provided $\sigma(t,x) - \sigma(t,R_{t,T}y) \geq \delta > 0$.

9 Steps for the proof of Theorem 2.1 in the nonlinear case.

We specify in this section how to modify the previous arguments to prove the well posedness of the martingale problem for the generator of (2.5). We focus on the case

$n = 2$. The constraint $d = 1$ will appear clearly during the proof. The first step consists in choosing a suitable parametrix. This is done similarly to the Gaussian case in [DM10]. Namely, we introduce for given $(T, y) \in \mathbb{R}^+ \times \mathbb{R}^d$ and $(t, x) \in [0, T) \times \mathbb{R}^d$ the *frozen process*:

$$\begin{aligned}
(\tilde{X}_s^{t,x,T,y})^1 &= x^1 + \int_t^s F_1(u, \phi_{u,T}(y))du + \int_t^s \sigma(u, \phi_{u,T}(y))dZ_u, \\
(\tilde{X}_s^{t,x,T,y})^2 &= x^2 + \int_t^s \Big(F_2(u, \phi_{u,T}(y)) \\
&\quad + \nabla_{x^1} F_2(u, \phi_{u,T}(y)) \Big((\tilde{X}_u^{t,x,T,y})^1 - \phi_{u,T}(y)^1 \Big) \Big) du,
\end{aligned} \tag{9.9}$$

denoting $\phi_{u,T}(y)$ the solution to $\phi_{T,T}(y) = y,\ \dot{\phi}_{u,T}(y) = F(u, \phi_{u,T}(y))$, i.e. backward flow associated with the deterministic differential system. This is a linear dynamics which once integrated through the resolvent yields:

$$\tilde{X}_s^{t,T,x,y} = \tilde{\phi}_{s,t}^{T,y}(x) + \int_t^s \tilde{R}_{s,u}^{T,y} B\sigma(u, \phi_{u,T}(y))dZ_u, \tag{9.10}$$

where recalling $F(t, x) = (F_1(t, x), F_2(t, x))^*$:

$$\begin{aligned}
\tilde{\phi}_{s,t}^{T,y}(x) &= x + \int_t^s F(u, \phi_{u,T}(y))du \\
&\quad + \int_t^s \begin{pmatrix} 0 & 0 \\ \nabla_{x^1} F_2(u, \phi_{u,T}(y)) & 0 \end{pmatrix} \Big(\tilde{\phi}_{u,t}^{T,y}(x) - \phi_{u,T}(y) \Big) du \\
&= \tilde{R}_{s,t}^{T,y}(x) + \int_t^s \tilde{R}_{s,u}^{T,y} \left\{ F(u, \phi_{u,T}(y)) - \begin{pmatrix} 0 & 0 \\ \nabla_{x^1} F_2(u, \phi_{u,T}(y)) & 0 \end{pmatrix} \phi_{u,T}(y) \right\} du, \\
\partial_s \tilde{R}_{s,t}^{T,y} &= \begin{pmatrix} 0 & 0 \\ \nabla_{x^1} F_2(s, \phi_{s,T}(y)) & 0 \end{pmatrix} \partial_s \tilde{R}_{s,t}^{T,y},\ \tilde{R}_{t,t}^{T,y} = I_{2\times 2}.
\end{aligned}$$

It can be shown, similarly to Section 2.3 in [DM10] that:

$$|(\mathbb{T}_{T-t}^{\alpha})^{-1}(\tilde{\phi}_{T,t}^{T,y}(x) - y)| \asymp |(\mathbb{T}_{T-t}^{\alpha})^{-1}(x - \phi_{t,T}(y))| \asymp |(\mathbb{T}_{T-t}^{\alpha})^{-1}(\phi_{T,t}(x) - y)|. \tag{9.11}$$

It can now be derived from equations (9.10), (9.11) similarly to the proof of Proposition 5.3 that

$$\begin{aligned}
\tilde{p}_{\alpha}(t, T, x, y) &\le \frac{C}{(T-t)^{(1+\frac{2}{\alpha})d}} \frac{1}{(1 + |(\mathbb{T}_{T-t}^{\alpha})^{-1}(\tilde{\phi}_{T,t}^{T,y}(x) - y)|)^{d+1+\alpha}} \\
&\le \frac{C}{(T-t)^{(1+\frac{2}{\alpha})d}} \frac{1}{(1 + |(\mathbb{T}_{T-t}^{\alpha})^{-1}(x - \phi_{t,T}(y))|)^{d+1+\alpha}} \\
&:= C\bar{p}_{\alpha,\phi}(t, T, x, y).
\end{aligned}$$

From the desintegration of the density in Appendix 8 (see equation (8.2), Lemma 8.1 and estimate (8.4)) we also derive the *global* gradient bounds:

$$|\partial_{x^1}\tilde p_\alpha(t,T,x,y)| \le \frac{C}{(T-t)^{\frac{1}{\alpha}}}\bar p_{\alpha,\phi}(t,T,x,y),$$
$$|\partial_{x^2}\tilde p_\alpha(t,T,x,y)| \le \frac{C}{(T-t)^{1+\frac{1}{\alpha}}}\bar p_{\alpha,\phi}(t,T,x,y).$$

On the other hand, the control of Lemma 3.7 concerning the kernel H now writes:

$$\begin{aligned}|H(t,T,x,y)| &\le C\frac{\delta\wedge|x-\phi_{t,T}(y)|^{\eta(\alpha\wedge 1)}}{T-t}\left\{\bar p_{\alpha,\phi}(t,T,x,y)+\breve p_{\alpha,\phi}(t,T,x,y)\right\}+\\ &C\left\{\frac{|x-\phi_{t,T}(y)|}{(T-t)^{\frac{1}{\alpha}}}+\left[\frac{|(x-\phi_{t,T}(y))^2|}{(T-t)^{1+\frac{1}{\alpha}}}+\frac{|(x-\phi_{t,T}(y))^1|(\delta\wedge|(x-\phi_{t,T}(y))^1|^\eta)}{(T-t)^{1+\frac{1}{\alpha}}}\right]\right\}\\ &\times\bar p_{\alpha,\phi}(t,T,x,y),\end{aligned}\tag{9.12}$$

where

$$\begin{aligned}\breve p_{\alpha,\phi}(t,T,x,y) &= \frac{\mathbf{1}_{|(x-\phi_{t,T}(y))^1|/(T-t)^{1/\alpha}|\asymp|(\mathbb{T}^\alpha_{T-t})^{-1}(x-\phi_{t,T}(y))|\ge K}}{(T-t)^{d/\alpha}(1+\frac{|(x-\phi_{t,T}(y))^1|}{(T-t)^{1/\alpha}})^{d+\alpha}}\\ &\times\frac{1}{(T-t)^{\frac{(n-1)d}{\alpha}+\frac{n(n-1)d}{2}}\left(1+\left|\left((\mathbb{T}^\alpha_{T-t})^{-1}(x-\phi_{t,T}(y))\right)^{2:n}\right|\right)^{1+\alpha}}.\end{aligned}$$

We emphasize that in the above controls on $\tilde p_\alpha, \nabla\tilde p_\alpha, H$, we have bounded the tempering function, appearing in the case **[HT]**, by a constant. Indeed, this term is not useful to investigate the martingale problem. Also, the additional contribution in H coming from the gradient term, which vanishes in the linear case, is derived writing:

$$\begin{aligned}&|\langle F(t,x)-(F(t,\phi_{t,T}(y))+\begin{pmatrix}0 & 0\\ \nabla_{x^1}F_2(t,\phi_{t,T}(y)) & 0\end{pmatrix}(x-\phi_{t,T}(y)),\nabla\tilde p_\alpha(t,T,x,y)\rangle|\\ &\le C\left\{\frac{|x-\phi_{t,T}(y)|}{(T-t)^{\frac{1}{\alpha}}}+\left[\frac{|(x-\phi_{t,T}(y))^2|}{(T-t)^{1+\frac{1}{\alpha}}}+\frac{|(x-\phi_{t,T}(y))^1|(\delta\wedge|(x-\phi_{t,T}(y))^1|^\eta)}{(T-t)^{1+\frac{1}{\alpha}}}\right]\right\}\\ &\qquad\times\bar p_{\alpha,\phi}(t,T,x,y).\end{aligned}$$

The contributions in (9.12) coming from the non-local part of H can be analyzed as previously. Let us now focus on the term

$$\frac{|(x-\phi_{t,T}(y))^1|(\delta\wedge|(x-\phi_{t,T}(y))^1|^\eta)}{(T-t)^{1+\frac{1}{\alpha}}}\bar p_{\alpha,\phi}(t,T,x,y) := G(t,T,x,y),$$

which is the trickiest among the new contributions. Indeed, it involves the first component, which has typical scale in $(T-t)^{1/\alpha}$, renormalized by the singularity deriving from the sensitivity w.r.t. the second one, i.e. $(T-t)^{-(1+1/\alpha)}$. Following the proof of Lemma 3.8 we write:

$$\begin{aligned} G(t,T,x,y) &\le C\int_{\mathbb{R}^{2d}} \frac{(T-t)^{\frac{1}{\alpha}}|Z^1|(\delta\wedge[(T-t)^{\frac{1}{\alpha}}|Z^1|]^\eta)}{(T-t)^{1+\frac{1}{\alpha}}}\frac{dZ}{(1+|Z|)^{d+1+\alpha}}\\ &\le \frac{C}{T-t}\int_{\mathbb{R}^d}|Z^1|(\delta\wedge[(T-t)^{\frac{1}{\alpha}}|Z^1|^\eta]))\frac{dZ^1}{(1+|Z^1|)^{1+\alpha}}\\ &\le C\{(T-t)^{\frac{\eta}{\alpha}-1}+\frac{1}{T-t}\int_{|Z^1|>K}(\delta^{1/\eta}\wedge[(T-t)^{\frac{1}{\alpha}}|Z^1|])^\eta)\frac{dZ^1}{|Z^1|^\alpha}\}\\ &\le C\{(T-t)^{\frac{\eta}{\alpha}-1}+\frac{1}{T-t}\int_{|Z^1|>K}(\delta^{1/\eta}\wedge[(T-t)^{\frac{1}{\alpha}}|Z^1|])^\varepsilon)\frac{dZ^1}{|Z^1|^\alpha}\}, \end{aligned}$$

for any $\varepsilon\in[0,\eta]$. Now, the above integral only converges if $d=1,\alpha>1$ and $\alpha-\varepsilon>1$ giving

$$G(t,T,x,y)\le C\{(T-t)^{\frac{\eta}{\alpha}-1}+(T-t)^{\frac{\varepsilon}{\alpha}-1}\},$$

which, once integrated in time yields the needed smoothing effect.

We conclude saying that it seems anyhow difficult to consider this case, i.e. a fully non linear unbounded drift, for the density estimate, since the additional contribution in (9.12) gives non-integrable singularities which we cannot here compensate as for the diffusion with a dependence on the fast variable.

However, it can be proved under **[HT]** for $d=1,n=2,\alpha>1$ and

$$F(t,x)=\begin{pmatrix}F_1(t,x)\\ \alpha_t x_1+\tilde F_2(t,x^2)\end{pmatrix},$$

where the coefficients are bounded measurable in time and s.t. $\alpha_t\in[c_0,c_0^{-1}]$, $c_0\in(0,1]$, and $F_1,\tilde F_2$ are Lipschitz continuous in space, that the density exists and that the estimates of Theorem 2.2 hold with the non-linear flow. In that case, the most singular term is linear and vanishes in H. Assumption **[HT]** is here crucial and would give instead of the previous control (9.12) that:

$$\begin{aligned}|H(t,T,x,y)|\le C\Bigg[&\frac{\delta\wedge|x-\phi_{t,T}(y)|^{\eta(\alpha\wedge1)}}{T-t}\{\bar p_{\alpha,\phi}(t,T,x,y)+\breve p_{\alpha,\phi}(t,T,x,y)\}+\\ &\left\{\frac{|x-\phi_{t,T}(y)|}{(T-t)^{\frac{1}{\alpha}}}+\frac{|(x-\phi_{t,T}(y))^2|}{(T-t)^{1+\frac{1}{\alpha}}}\right\}\bar p_{\alpha,\phi}(t,T,x,y)\Bigg]\theta(|\mathbb{M}_{T-t}^{-1}(x-\phi_{t,T}(y))|).\end{aligned}$$

The first contribution can be analyzed as previously, see Section 6. On the other hand the definition of $\Theta(r) := (1+r)\theta(r), r > 0$ gives:

$$\begin{aligned}\left\{\frac{|x-\phi_{t,T}(y)|}{(T-t)^{\frac{1}{\alpha}}} + \frac{|(x-\phi_{t,T}(y))^2|}{(T-t)^{1+\frac{1}{\alpha}}}\right\}\bar{p}_{\alpha,\phi}(t,T,x,y)\theta(|\mathbb{M}_{t,T}^{-1}(x-\phi_{t,T}(y))|)\\ \le \frac{C}{(T-t)^{\frac{1}{\alpha}}}\bar{p}_{\alpha,\phi,\Theta}(t,T,x,y),\end{aligned}$$

where $\bar{p}_{\alpha,\phi,\Theta}(t,T,x,y) := \bar{p}_{\alpha,\phi}(t,T,x,y)\Theta(|\mathbb{M}_{t,T}^{-1}(x-\phi_{t,T}(y))|)$.

Chapter 4

A Multi-step Richardson-Romberg extrapolation method for stochastic approximation

We obtain an expansion of the implicit weak discretization error for the target of stochastic approximation algorithms introduced and studied in [Fri13]. This allows us to extend and develop the Richardson-Romberg extrapolation method for Monte Carlo linear estimator (introduced in [TT90] and deeply studied in [Pag07]) to the framework of stochastic optimization by means of stochastic approximation algorithm. We notably apply the method to the estimation of the quantile of diffusion processes. Numerical results confirm the theoretical analysis and show a significant reduction in the initial computational cost.

1 Statement of the Problem

The aim of this paper is to combine a multistep Richardson-Romberg extrapolation method with stochastic approximation (SA) algorithms which are recursive simulation based procedures commonly used in the framework of stochastic optimization. Introduced by Robbins and Monro [RM51], SA algorithms aims at computing a zero of a continuous function $h : \mathbb{R}^d \to \mathbb{R}^d$ which is unknown to the experimenter but can only be estimated through experiments. In this general context, the function h writes $h(\theta) := \mathbb{E}[H(\theta, U)]$ where $H : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$ and U is a $\mathbb{R}^q$-valued random vector. To estimate a zero of h, one devises the following recursive algorithm

$$\theta_{p+1} = \theta_p - \gamma_{p+1} H(\theta_p, U^{p+1}),\ p \geq 0 \tag{1.1}$$

where $(U^p)_{p\geq 1}$ is an *i.i.d.* sequence of random variables with the same law as U defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, θ_0 is independent of the innovation of the algorithm with $\mathbb{E}[|\theta_0|^2] < +\infty$ and $\gamma = (\gamma_p)_{p\geq 1}$ is a deterministic and decreasing sequence of non-negative steps satisfying the usual assumption

$$\sum_{p\geq 1} \gamma_p = +\infty, \quad \text{and} \quad \sum_{p\geq 1} \gamma_p^2 < +\infty. \tag{1.2}$$

When the function h is the gradient of a convex potential, the recursive procedure (1.1) is a stochastic gradient algorithm. Indeed replacing $H(\theta_p, U^{p+1})$ by $h(\theta_p)$ in (1.1) leads to the usual deterministic descent gradient procedure.

In many applications, notably in computational finance, the sequence of random vectors $(U^p)_{p\geq 1}$ is not directly simulatable (at a reasonable cost) and can only be approximated by another sequence of easily simulatable random vectors $((U^n)^p)_{p\geq 1}$, $n > 0$, where U^n (weakly or strongly) approximates U as $n \to +\infty$ with a standard weak discretization error (or bias) $\mathbb{E}[f(U^n)] - \mathbb{E}[f(U)]$ that can be expanded in powers of $n^{-\alpha}$, $\alpha > 0$, for a specific class of functions $f \in \mathcal{C}$. One typical situation is when $U = X_T$, $X := (X_t)_{t\in[0,T]}$ being a q-dimensional diffusion process solution of a stochastic differential equation (SDE) and $U^n = X_T^n$ where $X^n := (X_t^n)_{t\in[0,T]}$ stands for its standard Euler-Maruyama discretization scheme with time step $\Delta = T/n$, $n \in \mathbb{N}^*$.

Since we are interested in the computation of the zero θ^* of h given by $h(\theta) := \mathbb{E}[H(\theta, U)]$ where $H : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$ and the function h is generally neither known nor computable since the random variable U cannot be easily simulated, estimating θ^* by devising directly the recursive scheme (1.1) is not possible. Therefore, two steps are needed to compute θ^*:

- the first step consists in approximating the zero θ^* of h by the zero $\theta^{*,n}$ of the function h^n defined by $h^n(\theta) := \mathbb{E}[H(\theta, U^n)]$, $\theta \in \mathbb{R}^d$. It induces *an implicit discretization error* which writes

$$\mathcal{E}_D(n) := \theta^* - \theta^{*,n}.$$

Under mild assumptions on h and h^n, it is proved in [Fri13] that $\theta^{*,n}$ converges to θ^* as n goes to infinity. Moreover, if the *standard weak discretization error* is of order $n^{-\alpha}$, $\alpha \in (0,1)$, that is $\forall \theta \in \mathbb{R}^d$, $h^n(\theta) - h(\theta) = \Lambda_1^0(\theta) n^{-\alpha} + o(n^{-\alpha})$, with $\Lambda_1^0 : \mathbb{R}^d \to \mathbb{R}^d$, then (under additional mild assumptions) this rate of convergence transfers to the *implicit discretization error* that is $\mathcal{E}_D(n) = \Theta_1 n^{-\alpha} + o(n^{-\alpha})$ for some $\Theta_1 \in \mathbb{R}^d$.

- the second step consists in approximating $\theta^{*,n}$ using $M \in \mathbb{N}^*$ steps of the following SA scheme

$$\theta_{p+1}^n = \theta_p^n - \gamma_{p+1} H(\theta_p^n, (U^n)^{p+1}), \; p \in [\![0, M-1]\!], \tag{1.3}$$

where $((U^n)^p)_{p\in[\![1,M]\!]}$ is an i.i.d. sequence of random variables with the same law as U^n, θ_0^n is independent of the innovation of the algorithm with $\sup_{n\geq 1} \mathbb{E}|\theta_0^n|^2 < +\infty$ and $\gamma = (\gamma_p)_{p\geq 1}$ is a sequence of non-negative deterministic and decreasing steps satisfying (1.2). This induces a *statistical error* which writes

$$\mathcal{E}_S(n, M) := \theta^{*,n} - \theta_M^n.$$

Regarding the *statistical error*, it is well-known that under mild assumptions the Robbins-Monro theorem guarantees that for each $n \in \mathbb{N}^*$, $\lim_{M \to +\infty} \mathcal{E}_S(n, M) = 0$. Moreover, under additional technical assumptions, a central limit theorem (CLT) holds at rate $\gamma^{-1/2}(M)$ that is $\gamma^{-1/2}(M)\mathcal{E}_S(n, M)$ converges in distribution to a normally distributed random variable. The reader may also refer to [FM12] and [FF13] for some recent developments on non-asymptotic deviation bounds for the statistical error.

The global error between θ^*, the quantity to estimate, and its implementable approximation θ^n_M can be decomposed as follows:

$$\begin{aligned}\mathcal{E}_{glob}(n, M) &= \theta^* - \theta^{*,n} + \theta^{*,n} - \theta^n_M \\ &:= \mathcal{E}_D(n) + \mathcal{E}_S(n, M).\end{aligned}$$

The first aim of this paper is to prove the existence of an expansion for the *implicit discretization error*, that is, under mild assumptions (see Section 2) on h and h^n, $\mathcal{E}_D(n)$ can be expanded as follows

$$\forall R \in \mathbb{N}^*, \quad \theta^{*,n} - \theta^* = \frac{C_1}{n^\alpha} + \cdots + \frac{C_R}{n^{\alpha R}} + o\left(\frac{1}{n^{\alpha R}}\right) \tag{1.4}$$

where $(C_1, \cdots, C_R) \in (\mathbb{R}^d)^R$. Then taking advantage of (1.4) we devise a multistep Richardson-Romberg extrapolation method for stochastic optimization by means of stochastic approximation algorithm. The principle of Richardson-Romberg extrapolation is to reduce the bias produced by the *implicit discretization error* by combining two estimators with different step size. To be more precise, one considers the two following weights $w_1 = (-1/(2^\alpha - 1))I_d$ and $w_2 = (2^\alpha/(2^\alpha - 1))I_d$, I_d is the identity matrix of dimension d and the Richardson-Romberg SA estimator

$$\Theta^{n,2n}_M = w_1\theta^n_M + w_2\theta^{2n}_M$$

where $(\theta^{2n}_M, \theta^n_M)$ is obtained using M steps of two SA schemes devised with the i.i.d. sequence $((U^{2n}, U^n)^p)_{p \in [\![1,M]\!]}$ of random variables with the same law as (U^{2n}, U^n). Under standard assumptions, this linear combination of SA estimators *a.s.* converges to the target $w_1\theta^{*,n} + w_2\theta^{*,2n}$ as the number of steps M goes to infinity. The key observation is that this new target satisfies the following implicit error expansion of order 2

$$w_1\theta^{*,n} + w_2\theta^{*,2n} - \theta^* = -\frac{C_2}{2^\alpha}\frac{1}{n^{2\alpha}} + o\left(\frac{1}{n^{2\alpha}}\right).$$

Moreover, in the spirit of [Pag07], we show how to control the asymptotic $L^1(\mathbb{P})$-norm of the distance between the new estimator $\Theta^{n,2n}_M$ and its target $w_1\theta^{*,n} + w_2\theta^{*,2n}$ as n goes to infinity. Then, it is natural to iterate this extrapolation to obtain a new SA estimator with an implicit discretization error of order $n^{-\alpha R}$ for any $R \in \mathbb{N}^*$. This extension called multi-step Richardson-Romberg extrapolation is deeply investigated in [Pag07] for Monte Carlo linear estimator in the framework of discretization of diffusion processes.

The aim of this paper is to investigate the Richardson-Romberg SA method. Our purpose is to show that the principle of multi-step Richardson-Romberg extrapolation for Monte Carlo linear estimator can be extended to the framework of stochastic optimization by means of SA algorithm. We notably prove that the new estimator outperforms the standard SA estimator in terms of computational cost.

The paper is organized as follows: in Section 2 we provide an expansion of the implicit discretization error in powers of $n^{-\alpha}$ under mild assumptions. Then we take advantage of this expansion to propose a multi-step Richardson-Romberg method by means of SA. In Section 3 is presented an illustration of the method to the estimation of the quantile of a stochastic differential equation (SDE) driven by a stable process. In Section 4 numerical results are carried out to confirm the theoretical analysis. Finally, Section 5 is devoted to theoretical results which are useful throughout the paper.

2 Main results

This section is divided in two parts. In the first one we obtain a general result concerning the expansion of the implicit discretization error. In the second one, we take advantage of this result to develop a Richardson-Romberg extrapolation method for stochastic optimization by means of SA algorithms.

2.1 Expansion of the implicit discretization error

We first provide a result concerning the convergence of the sequence $(\theta^{*,n})_{n\geq 1}$ towards θ^*. For a proof the reader may refer to [Fri13].

Proposition 2.1. *For all $n \in \mathbb{N}^*$, assume that h and h^n satisfy the mean reverting assumption:*

$$\forall \theta \neq \theta^*,\ \langle \theta - \theta^*, h(\theta)\rangle > 0 \quad \text{and} \quad \forall \theta \neq \theta^{*,n},\ \langle \theta - \theta^{*,n}, h^n(\theta)\rangle > 0.$$

Moreover, suppose that $(h^n)_{n\geq 1}$ converges locally uniformly towards h. Then, one has

$$\theta^{*,n} \to \theta^* \quad as \quad n \to +\infty.$$

Here we will investigate an expansion of the error term $\theta^{*,n} - \theta^*$ in powers of $n^{-\alpha}$. Through the document, we will refer to **[H-k]** the following set of assumptions:

1. For all $\theta \in \mathbb{R}^d$,
$$h(\theta) - h^n(\theta) = \frac{\Lambda_1^0(\theta)}{n^\alpha} + \cdots + \frac{\Lambda_k^0(\theta)}{n^{\alpha k}} + o\left(\frac{1}{n^{\alpha k}}\right). \tag{2.5}$$

2. $h, h^n \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^d)$ and for all $l \leq k-1$, for all $\theta \in \mathbb{R}^d$,
$$D^l h^n(\theta) - D^l h(\theta) = \frac{\Lambda_1^l(\theta)}{n^\alpha} + \cdots + \frac{\Lambda_{k-l}^l(\theta)}{n^{\alpha(k-l)}} + o\left(\frac{1}{n^{\alpha(k-l)}}\right) \tag{2.6}$$
where for all $\theta \in \mathbb{R}^d$, $\Lambda_1^l(\theta), \cdots, \Lambda_{k-l}^l(\theta)$ and $o(n^{-\alpha(k-l)})$ are multilinear maps from $(\mathbb{R}^d)^l$ to $\mathbb{R}^d$.

3. For all $l \in [\![1, k]\!]$, $(D^l h^n)_{n\geq 1}$ converges locally uniformly towards $D^l h$.

4. $Dh(\theta^*)$ is invertible.

Proposition 2.2. *Assume that $\theta^{*,n} \to \theta^*$ as $n \to +\infty$. Under* ***[H-1]****, one has*
$$n^\alpha\left(\theta^{*,n} - \theta^*\right) \underset{n\to\infty}{\longrightarrow} Dh(\theta^*)^{-1}\Lambda_1^0(\theta^*). \tag{2.7}$$

Proof. Observe that one has $h^n(\theta^{*,n}) - h^n(\theta^*) = -h^n(\theta^*) = h(\theta^*) - h^n(\theta^*)$. On the one hand, writing Taylor's formula with integral remainder yields:
$$h^n(\theta^{*,n}) - h^n(\theta^*) = \int_0^1 dt Dh^n(t\theta^{*,n} + (1-t)\theta^*)(\theta^{*,n} - \theta^*). \tag{2.8}$$
On the other hand, from the discretization error, we have $h(\theta^*) - h^n(\theta^*) = \Lambda_1^0(\theta^*)n^{-\alpha} + o\left(n^{-\alpha}\right)$.

Since $\theta^{*,n} \underset{n\to\infty}{\longrightarrow} \theta^*$, $Dh(\theta^*)$ is invertible, and $(Dh^n)_{n\geq 1}$ converges uniformly locally to Dh, for n large enough, the matrix $\int_0^1 Dh^n(t\theta^{*,n} + (1-t)\theta^*)dt$ is invertible. Multiplying both sides of (2.8) by n^α finally yields
$$n^\alpha\left(\theta^{*,n} - \theta^*\right) = \left(\int_0^1 Dh^n(t\theta^{*,n} + (1-t)\theta^*)dt\right)^{-1}\left(\Lambda_1^0(\theta^*) + o(1)\right) \underset{n\to\infty}{\longrightarrow} Dh(\theta^*)^{-1}\Lambda_1^0(\theta^*).$$
□

Let us note that Proposition 2.2 provides a first order expansion of $\theta^{*,n} - \theta^*$, that is $\theta^{*,n} - \theta^* = C_1 n^{-\alpha} + o(n^{-\alpha})$. We now give a generalization of this first result.

Theorem 2.3. *Assume that $\theta^{*,n} \to \theta^*$, $n \to +\infty$, and that* ***[H-p]*** *holds for some $p \in \mathbb{N}^*$. Then, $\theta^{*,n} - \theta^*$ has an expansion up to order p, that is, the following expansion holds:*
$$\theta^{*,n} - \theta^* = \frac{C_1}{n^\alpha} + \cdots + \frac{C_p}{n^{\alpha p}} + o\left(\frac{1}{n^{\alpha p}}\right).$$

Proof. If **[H-p]**, $p \in \mathbb{N}^*$, holds then Proposition 2.2 gives a first order expansion for $\theta^{*,n} - \theta^*$. We now prove the inductive step that is if $\theta^{*,n} - \theta^*$ has an expansion of order $k-1$ then an expansion holds at order k, for $k \leq p$. The basic idea does not change from the previous computation. From the development of the discretization error, we have:

$$h(\theta^*) - h^n(\theta^*) = \frac{\Lambda_1^0(\theta^*)}{n^\alpha} + \cdots + \frac{\Lambda_k^0(\theta^*)}{n^{\alpha k}} + o\left(\frac{1}{n^{\alpha k}}\right). \tag{2.9}$$

On the other hand, we write a Taylor's expansion of h^n up to the same order $k-1$:

$$\begin{aligned} h^n(\theta^{*,n}) - h^n(\theta^*) &= Dh^n(\theta^*)(\theta^{*,n} - \theta^*) + \cdots \\ &\quad + \frac{1}{(k-1)!} D^{k-1}h^n(\theta^*)(\theta^{*,n} - \theta^*)^{(k-1)} \\ &\quad + R_{k-1}^n(\theta^{*,n} - \theta^*), \end{aligned} \tag{2.10}$$

with the remainder in integral form satisfying:

$$\begin{aligned} R_{k-1}^n(\theta^{*,n} - \theta^*) &= \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k h^n(t\theta^{*,n} + (1-t)\theta^*)(\theta^{*,n} - \theta^*)^{(k)} dt \\ &= \frac{1}{k!} D^k h(\theta^*)(\theta^{*,n} - \theta^*)^{(k)} + o\left(\frac{1}{n^{\alpha k}}\right) \end{aligned}$$

where we used that $(D^k h^n)_{n \geq 1}$ converges locally uniformly to $D^k h$, $k \in [\![1, p]\!]$, and $\theta^{*,n} - \theta^* = O(n^{-\alpha})$ for the last equality. Let us note that for $l \in [\![1, k]\!]$, $D^l h(\theta^*)$ (as $\Lambda_j^l(\theta^*)$, $j = 1, \cdots, k-l$) is a multilinear maps from $(\mathbb{R}^d)^l$ to $\mathbb{R}^d$. The expansions (2.6) allow us to replace the derivatives of h^n by the derivatives of h in (2.10) at the cost of an error term, that is:

$$\begin{aligned} & h^n(\theta^{*,n}) - h^n(\theta^*) \\ = \;& Dh(\theta^*)(\theta^{*,n} - \theta^*) + \left(\frac{\Lambda_1^1(\theta^*)}{n^\alpha} + \cdots + \frac{\Lambda_{k-1}^1(\theta^*)}{n^{\alpha(k-1)}} + o\left(\frac{1}{n^{\alpha(k-1)}}\right)\right)(\theta^{*,n} - \theta^*) \\ & + \cdots + \frac{1}{(k-1)!}\left(D^{k-1}h(\theta^*) + \frac{\Lambda_1^{k-1}(\theta^*)}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right)\right)(\theta^{*,n} - \theta^*)^{(k-1)} \\ & + \frac{1}{k!} D^k h(\theta^*)(\theta^{*,n} - \theta^*)^{(k)} + o\left(\frac{1}{n^{\alpha k}}\right). \end{aligned}$$

Since $h^n(\theta^{*,n}) - h^n(\theta^*) = -h^n(\theta^*) = h(\theta^*) - h^n(\theta^*)$ and $Dh(\theta^*)$ is invertible, the previous equality implies

$$\frac{Dh(\theta^*)^{-1}\Lambda_1^0(\theta^*)}{n^\alpha} + \cdots + \frac{Dh(\theta^*)^{-1}\Lambda_k^0(\theta^*)}{n^{\alpha k}} + o\left(\frac{1}{n^{\alpha k}}\right) =$$
$$\theta^{*,n} - \theta^* + \left(\frac{Dh(\theta^*)^{-1}\Lambda_1^1(\theta^*)}{n^\alpha} + \cdots + \frac{Dh(\theta^*)^{-1}\Lambda_{k-1}^1(\theta^*)}{n^{\alpha(k-1)}} + o\left(\frac{1}{n^{\alpha(k-1)}}\right)\right)(\theta^{*,n} - \theta^*)$$
$$+ \cdots + \frac{1}{(k-1)!}\left(Dh(\theta^*)^{-1}D^{k-1}h(\theta^*) + \frac{Dh(\theta^*)^{-1}\Lambda_1^{k-1}(\theta^*)}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right)\right)(\theta^{*,n} - \theta^*)^{(k-1)}$$
$$+ \frac{1}{k!}Dh(\theta^*)^{-1}D^k h(\theta^*)(\theta^{*,n} - \theta^*)^{(k)} + o\left(\frac{1}{n^{\alpha k}}\right).$$

The last equation should be seen as a "bootstrap" for $\theta^{*,n} - \theta^*$, that is:

$$\begin{aligned}
\theta^{*,n} - \theta^* = {} & \frac{Dh(\theta^*)^{-1}\Lambda_1^0(\theta^*)}{n^\alpha} + \cdots + \frac{Dh(\theta^*)^{-1}\Lambda_k^0(\theta^*)}{n^{\alpha k}} + o\left(\frac{1}{n^{\alpha k}}\right) \\
& - \left(\frac{Dh(\theta^*)^{-1}\Lambda_1^1(\theta^*)}{n^\alpha} + \cdots + \frac{Dh(\theta^*)^{-1}\Lambda_{k-1}^1(\theta^*)}{n^{\alpha(k-1)}} + o\left(\frac{1}{n^{\alpha(k-1)}}\right)\right)(\theta^{*,n} - \theta^*) \\
& - \cdots \\
& - \frac{Dh(\theta^*)^{-1}}{(k-1)!}\left(D^{k-1}h(\theta^*) + \frac{\Lambda_1^{k-1}(\theta^*)}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right)\right)(\theta^{*,n} - \theta^*)^{(k-1)} \\
& - \frac{1}{k!}Dh(\theta^*)^{-1}D^k h(\theta^*)(\theta^{*,n} - \theta^*)^{(k)} + o\left(\frac{1}{n^{\alpha k}}\right),
\end{aligned} \tag{2.11}$$

The idea now is to plug the expansion of $\theta^{*,n} - \theta^*$ in the right hand side of (2.11) and check that the first remainder term comes at order $o(n^{-\alpha k})$. It is clear that on the first line the remainder term is of order $o(n^{-\alpha k})$. Moreover, for any $l \in [\![2, k]\!]$, the generic l-th term writes in the i-th component:

$$\begin{aligned}
& \frac{1}{l!}\left(\left(D^l h(\theta^*) + \frac{\Lambda_1^l(\theta^*)}{n^\alpha} + \cdots + \frac{\Lambda_{k-l}^l(\theta^*)}{n^{\alpha(k-l)}} + o\left(\frac{1}{n^{\alpha(k-l)}}\right)\right)(\theta^{*,n} - \theta^*)^{(l)}\right)_i \\
= & \sum_{i_1+\cdots+i_d=l} \frac{1}{i_1!\cdots i_d!}\Lambda_{i_1,\cdots,i_d}(\theta^{*,n} - \theta^*)_1^{i_1} \times \cdots \times (\theta^{*,n} - \theta^*)_d^{i_d} \\
& + o\left(\frac{1}{n^{\alpha(k-l)}}\right)(\theta^{*,n} - \theta^*)_1^{i_1} \times \cdots \times (\theta^{*,n} - \theta^*)_d^{i_d} \\
= & \sum_{i_1+\cdots+i_d=l} \frac{1}{i_1!\cdots i_d!}\Lambda_{i_1,\cdots,i_d}(\theta^{*,n} - \theta^*)_1^{i_1} \times \cdots \times (\theta^{*,n} - \theta^*)_d^{i_d} + o\left(\frac{1}{n^{\alpha k}}\right)
\end{aligned} \tag{2.12}$$

where $\Lambda_{i_1,\cdots,i_d} = \frac{\partial^l h_i}{\partial\theta_1^{i_1}\cdots\partial\theta_d^{i_d}}(\theta^*) + \frac{(\Lambda_1^l(\theta^*))_i}{n^\alpha} + \cdots + \frac{(\Lambda_{k-l}^l(\theta^*))_i}{n^{\alpha(k-l)}}$ with $(\Lambda_j^l(\theta^*))_i$ for $j \in [\![1, k-l]\!]$ satisfying $\frac{\partial^l h_i}{\partial\theta_1^{i_1}\cdots\partial\theta_d^{i_d}}(\theta^*) - \frac{\partial^l h_i^n}{\partial\theta_1^{i_1}\cdots\partial\theta_d^{i_d}}(\theta^*) = (\Lambda_1^l(\theta^*))_i/n^\alpha + \cdots + (\Lambda_{k-l}^l(\theta^*))_i/n^{\alpha(k-l)} + o(1/n^{\alpha(k-l)})$ and where we used that $(\theta^{*,n}-\theta^*)_1^{i_1} \times \cdots \times (\theta^{*,n}-\theta^*)_d^{i_d} = O(1/n^{\alpha l})$ for the last equality. Now, replacing $(\theta^{*,n}-\theta^*)_i$ by its expansion, we observe that the generic term in (2.12) satisfies

$$\begin{aligned}
&\Lambda_{i_1,\cdots,i_d}\left(\frac{C_1^1}{n^\alpha} + \cdots + \frac{C_{k-1}^1}{n^{\alpha(k-1)}} + o\left(\frac{1}{n^{\alpha(k-1)}}\right)\right)^{i_1} \times \\
&\cdots \times \left(\frac{C_1^d}{n^\alpha} + \cdots + \frac{C_{k-1}^d}{n^{\alpha(k-1)}} + o\left(\frac{1}{n^{\alpha(k-1)}}\right)\right)^{i_d} \\
&= \Lambda_{i_1,\cdots,i_d}\left(\frac{\tilde{C}}{n^{\alpha l}} + \cdots + o\left(\frac{1}{n^{\alpha(k+l-2)}}\right)\right) \\
&= \Lambda_{i_1,\cdots,i_d}\left(\frac{\tilde{C}}{n^{\alpha l}} + \cdots + o\left(\frac{1}{n^{\alpha k}}\right)\right)
\end{aligned}$$

where $\tilde{C} = (C_1^1)^{i_1} \times \cdots \times (C_1^d)^{i_d}$. We clearly see that the expression above yields an expansion in powers of $n^{-\alpha}$ with a remainder at order $o(n^{-\alpha k})$. Formally, as the power in the expansion (2.6) goes down, the power in the derivatives grows, compensating exactly and giving the right order in the remainder.

Finally, we expand the previous equation and group together the different terms with respect to the power of $n^{-\alpha}$. As we observed above, the remainder term is at order $o\left(n^{-\alpha k}\right)$, because of the compensation between the power in the expansion (2.6) and the order of the Taylor expansion. This completes the proof. □

2.2 Multi-step Richardson-Romberg extrapolation for stochastic approximation

Multi-step Richardson-Romberg extrapolation was successfully applied in the context of Monte Carlo linear estimator for the computation of $\mathbb{E}[f(X_T)]$, where $f : \mathbb{R}^d \to \mathbb{R}$ (with possible extension to the case of path-dependent options) and X is the (unique) strong solution to an SDE, see [Pag07]. In this section, we propose a multi-step Richardson-Romberg SA estimator with a control of the statistical error. We proceed as follows. Let $R \geq 2$ be an integer. To devise a SA estimator whose target has an implicit discretization error of order $n^{-\alpha R}$ as $n \to +\infty$, we introduce a sequence of R random vectors $\{U^{rn}, r \in [\![1, R]\!]\}$, $n \in \mathbb{N}^*$. Throughout this section we will assume that this sequence satisfies $U^{rn} \xrightarrow{\mathbb{P}} U^r$ as $n \to +\infty$ with $U^r \overset{d}{=} U$, $r \in [\![1, R]\!]$, all variables being defined on the same probability space. If assumption **[H-R]** holds then for all $r \in [\![1, R]\!]$ one gets

$$\theta^{*,rn} = \theta^* + \sum_{p=1}^{R-1} \frac{C_p}{r^{\alpha p}} \frac{1}{n^{\alpha p}} + \frac{C_R}{r^{\alpha R}} \frac{1}{n^{\alpha R}} \left(1 + \epsilon_r(n)\right)$$

with $\epsilon_r(n) \to 0$ as $n \to +\infty$. Then, one defines the Vandermonde $Rd \times (R-1)d$ matrix

$$V = \left[\frac{I_d}{r^{\alpha p}}\right]_{1\leq r\leq R, 1\leq p\leq R-1}$$

and the extended $Rd \times d$ unit matrix $\mathbf{I} = (I_d, \cdots, I_d)^T$ where I_d is the identity matrix of dimension d. Now we write

$$\begin{pmatrix} \vdots \\ \theta^{*,rn} \\ \vdots \end{pmatrix}_{1\leq r\leq R} = \mathbf{I}\theta^* + V \begin{pmatrix} \vdots \\ \frac{C_r}{n^{\alpha r}} \\ \vdots \end{pmatrix}_{1\leq r\leq R-1} + \begin{pmatrix} \vdots \\ \frac{C_R}{r^{\alpha R}} \frac{1}{n^{\alpha R}} \left(1 + \epsilon_r(n)\right) \\ \vdots \end{pmatrix}_{1\leq r\leq R}. \tag{2.13}$$

We consider the $Rd \times d$ weight matrix $\mathbf{w} = (\mathbf{w}_1, \cdots, \mathbf{w}_R)^T$, $\mathbf{w}_i$ being a $d \times d$ matrix for $i \in [\![1, R]\!]$ satisfying

$$\mathbf{w}^T\mathbf{I} = I_d \quad \text{and} \quad \mathbf{w}^T V = 0_{d\times d(R-1)} \tag{2.14}$$

which is equivalent to

$$\tilde{V}\mathbf{w} = E_1 \tag{2.15}$$

with $E_1 = (I_d, 0_{d\times d(R-1)})^T$ and $\tilde{V}$ is the Vandermonde matrix defined by

$$\tilde{V} = \begin{pmatrix} I_d & I_d & \cdots & I_d \\ I_d & \frac{I_d}{2^\alpha} & \cdots & \frac{I_d}{R^\alpha} \\ \vdots & \vdots & \cdots & \vdots \\ I_d & \frac{I_d}{2^{(R-1)\alpha}} & \cdots & \frac{I_d}{R^{(R-1)\alpha}} \end{pmatrix}.$$

Thanks to Cramer's rule, the solution $\mathbf{w}$ to (2.15) is explicitly given by

$$\forall r \in \{1, \cdots, R\}, \quad \mathbf{w}_r = (-1)^{R-r} \frac{r^{\alpha R}}{\prod_{j=0}^{r-1}(r^\alpha - j^\alpha) \prod_{j=r+1}^{R}(j^\alpha - r^\alpha)} I_d, \tag{2.16}$$

where we use the convention $\prod_{j=R+1}^{R}(j^\alpha - r^\alpha) = 1$. Let us note that when $\alpha = 1$ this last expression simplifies to $\mathbf{w}_r = (-1)^{R-r}(r^R/(r!(R-r)!))I_d$, $r = 1, \cdots, R$. The first condition in (2.14) reads $\sum_{r=1}^R \mathbf{w}_r = I_d$ which implies that $\lim_{n\to+\infty} \sum_{r=1}^R \mathbf{w}_r \theta^{*,rn} = \sum_{r=1}^R \mathbf{w}_r \theta^* = \theta^*$. Moreover, multiplying (2.13) on the left by $\mathbf{w}^T$ yields

$$\sum_{r=1}^{R} \mathbf{w}_r \theta^{*,rn} = \theta^* + C_R \frac{1}{n^{\alpha R}} \tilde{\mathbf{w}}_{R+1} \left(1 + \epsilon_{R+1}(n)\right) \tag{2.17}$$

where

$$\tilde{\mathbf{w}}_{R+1} = \sum_{r=1}^{R} (-1)^{R-r} \frac{r^{\alpha R}}{\prod_{j=0}^{r-1}(r^\alpha - j^\alpha)\prod_{j=r+1}^{R}(j^\alpha - r^\alpha)} \frac{1}{r^{\alpha R}} = \frac{(-1)^{R-1}}{R!^\alpha} \tag{2.18}$$

and

$$\epsilon_{R+1}(n) = \frac{1}{\tilde{\mathbf{w}}_{R+1}} \sum_{r=1}^{R} \frac{(-1)^{R-r}}{\prod_{j=0}^{r-1}(r^\alpha - j^\alpha)\prod_{j=r+1}^{R}(j^\alpha - r^\alpha)} \epsilon_r(n) \to 0, \quad \text{as} \quad n \to +\infty. \tag{2.19}$$

We now approximate the new target $\sum_{r=1}^{R} \mathbf{w}_r \theta^{*,rn}$, by means of $M \in \mathbb{N}^*$ steps of R SA schemes which write

$$\forall r \in [\![1, R]\!], \quad \theta^{rn}_{p+1} = \theta^{rn}_p - \gamma_{p+1} H(\theta^{rn}_p, (U^{rn})^{p+1}), \; p \in [\![0, M-1]\!] \tag{2.20}$$

where $((U^{rn})^p, r = 1, \cdots, R)_{p \in [\![1,M]\!]}$ is an i.i.d sequence with the same law as $(U^{rn}, r = 1, \cdots, R)$, θ^{rn}_0, $r = 1 \cdots, R$ are the initial conditions independent of the innovation sequence satisfying $\sup_{n\geq 1} \mathbb{E}|\theta^n_0|^2 < +\infty$ and the sequence $(\gamma_p)_{p\geq 1}$ satisfies (1.2). Now the new *statistical error* of the Richardson-Romberg extrapolation estimator writes

$$\mathcal{E}^{R-R}_S(n, M) := \sum_{r=1}^{R} \mathbf{w}_r (\theta^{*,rn} - \theta^{rn}_M).$$

We are looking for an efficient estimator among the family

$$\left\{ \sum_{r=1}^{R} \mathbf{w}_r \theta^{rn}_M, (n, M) \in (\mathbb{N}^*)^2 \right\}.$$

To be more precise, we will minimize the computational cost for a given $L^1(\mathbb{P})$-error $\varepsilon > 0$. We assume that the cost of a single simulation of U^n is proportional to n and is given by $K \times n$, where K is a generic positive constant independent of n. It notably corresponds to the case of discretization schemes of a stochastic process. In the case of the Richardson-Romberg method for SA, at each step $p \in [\![1, M]\!]$ of the procedure, for every $r \in [\![1, R]\!]$, one has to simulate the random vector $(U^n, U^{2n}, \cdots, U^{Rn})$ so that the global computational cost is given by

$$\text{Cost(R-R)} := KM \sum_{r=1}^{R} rn = KMn\frac{R(R+1)}{2}.$$

Hence the problem of interest writes

$$(n(\epsilon), M(\epsilon)) = argmin_{\mathbb{E}|\mathcal{E}^{R-R}_{glob}| \leq \varepsilon} \text{Cost(R-R)}.$$

From a practical point of view the constraint: $\mathbb{E}|\mathcal{E}_{glob}^{R-R}| \leq \varepsilon$ is not tractable since one does not have any explicit control on $\mathbb{E}|\mathcal{E}_{glob}^{R-R}|$. Hence one is led to consider some sharp upper bound of this $L^1(\mathbb{P})$-norm, namely

$$\begin{aligned}\mathbb{E}|\mathcal{E}_{glob}^{R-R}| &\leq \left|\sum_{r=1}^{R} \mathbf{w}_r \theta^{*,rn} - \theta^*\right| + \mathbb{E}\left[\left|\sum_{r=1}^{R} \mathbf{w}_r (\theta^{*,rn} - \theta_M^{rn})\right|\right] \\ &\leq \frac{|C_R|}{(R!n^R)^\alpha}\left(1 + |\epsilon_{R+1}(n)|\right) + \mathbb{E}\left[\left|\sum_{r=1}^{R} \mathbf{w}_r (\theta^{*,rn} - \theta_M^{rn})\right|\right]. \end{aligned} \tag{2.21}$$

Note that the bound (2.21) is not tractable since we do not have any closed form expression for the last term appearing in the right-hand side, namely the L^1-norm (or L^2-norm) of the statistical error of the Richardson-Romberg SA estimator. Again we will consider some sharp upper bound. In order to derive an explicit control we assume that the following conditions are in force:

(HUI) $\exists \delta > 0$, such that $\forall \theta \in \mathbb{R}^d$, $\sup_{n\in\mathbb{N}^*} \mathbb{E}[|H(\theta, U^n)|^{2+\delta}] < +\infty$.

(HC1) $\exists C > 0$ such that $\forall n \in \mathbb{N}^*, \forall \theta \in \mathbb{R}^d$, $\mathbb{E}[|H(\theta, U^n)|^2] \leq C(1 + |\theta - \theta^{*,n}|^2)$.

(HC2) $\forall \theta \in \mathbb{R}^d$, $\mathbb{P}(U \notin \mathcal{C}_\theta) = 0$ with $\mathcal{C}_\theta := \{x \in \mathbb{R}^q : x \mapsto H(\theta, x) \text{ is continuous at } x\}$.

(HRG) There exists $a \in (0, 1]$,

$$\sup_{n\in\mathbb{N}^*, (\theta,\theta')\in(\mathbb{R}^d)^2} \frac{\mathbb{E}|H(\theta, U^n) - H(\theta', U^n)|^2}{|\theta - \theta'|^{2a}} < +\infty.$$

(HUA) For each $n \in \mathbb{N}^*$, the map $h^n : \theta \in \mathbb{R}^d \mapsto \mathbb{E}[H(\theta, U^n)]$ is continuously differentiable with Dh^n Lipschitz-continuous uniformly in n and there exists $\underline{\lambda} > 0$ s.t. $\inf_{n\in\mathbb{N}^*, \theta\in\mathbb{R}^d} \lambda_{min}\left((Dh^n(\theta) + Dh^n(\theta)^T)/2\right) > \underline{\lambda}$ where $\lambda_{min}(A)$ denotes the lowest eigenvalue of the matrix A. (*Uniform Attractivity*).

(HS) The step sequence is given by $\gamma_p = \gamma(p)$, $p \geq 1$, where γ is a positive function defined on $[0, +\infty[$ decreasing to zero satisfying one of the following assumptions:

- γ varies regularly with exponent $(-\rho)$, $\rho \in (1/2, 1)$, that is, for any $x > 0$, $\lim_{t\to+\infty} \gamma(tx)/\gamma(t) = x^{-\rho}$.
- for $t \geq 1$, $\gamma(t) = \gamma_0/t$ and γ_0 satisfies $2\underline{\lambda}\gamma_0 > 1$.

Remark 2.1. Assumption **(HUA)** already appears in [Duf96] and [BMP90], see also [FM12] and [FF13] in another context. It allows to control the L^2-norm $\mathbb{E}|\theta_p^{rn} - \theta^{*,rn}|^2$, $r \in [\![1, R]\!]$ with respect to the step $\gamma(p)$ uniformly in n, see section 5, lemma 5.2 . As discussed in [KY03], (Chapter 10, Section 5, p.350, Theorem 5.2) if one considers the projected version of the algorithm (1.3) on a bounded convex set D, namely

$$\theta_{p+1}^n = \Pi_D\left[\theta_p^n - \gamma_{p+1} H(\theta_p^n, (U^n)^{p+1})\right],\ p \in [\![0, M-1]\!],$$

where Π_D denotes the orthogonal projection operator on D (for instance one may set $D = \Pi_{i=1}^d [a_i, b_i]$, $-\infty < a_i < b_i < +\infty$) and $\forall n \geq 1$, $\theta^{*,n} \in int(D)$, as very often happens from a practical point of view, then assumption **(HUA)** can be localized on D, that is $\inf_{n\in\mathbb{N}^*,\theta\in D} \lambda_{min}\left((Dh^n(\theta) + Dh^n(\theta)^T)/2\right) > \underline{\lambda}$.

We also want to point out that if assumption **(HUA)** is satisfied then passing to the limit as $n \to +\infty$ one easily shows that $\lambda_{min}\left((Dh(\theta^*) + Dh(\theta^*)^T)/2\right) \geq \underline{\lambda}$.

Proposition 2.4. *($L^1(\mathbb{P})$ control of the statistical error) Let $R \in \mathbb{N}^*$. Suppose that for $r \in [\![1, R]\!]$, $U^{rn} \xrightarrow{\mathbb{P}} U^r$ and $\theta_0^n \xrightarrow{\mathbb{P}} \theta_0$, as $n \to +\infty$. Under **(H-R)**, **(HUI)**, **(HC1)**, **(HC2)**, **(HRG)**, **(HS)** and **(HUA)**, one has for some positive constant $C := C(\gamma, \underline{\lambda})$*

$$\mathbb{E}[|\mathcal{E}_S^{R-R}|] \leq C\mathbb{E}\left[\left|\sum_{r=1}^R \mathbf{w}_r H(\theta^*, U^r)\right|^2\right]^{1/2} \gamma^{1/2}(M)\left(1 + \phi_1^R(n) + \phi_2^R(M)\right)$$

where ϕ_1^R, ϕ_2^R are two positive functions satisfying: $\phi_1^R(n) \to 0$ and $\phi_2^R(M) \to 0$ respectively as $M \to +\infty$, $n \to +\infty$ and ϕ_2^R is non-increasing.

Proof. We define for all $p \geq 1$,

$$\begin{aligned}
\Delta M_p^{rn} &:= h^{rn}(\theta_{p-1}^{rn}) - H(\theta_{p-1}^{rn}, (U^{rn})^p) \\
&= \mathbb{E}[H(\theta_{p-1}^{rn}, (U^{rn})^p) \big| \mathcal{F}_{p-1}] - H(\theta_{p-1}^{rn}, (U^{rn})^p).
\end{aligned}$$

Recalling that $((U^n, U^{2n}, \cdots, U^{rn}, \cdots, U^{Rn})^p)_{p\in[\![1,M]\!]}$ is a sequence of i.i.d. random variables we have that $(\Delta M_p^{rn})_{p\geq 1}$, $r \in [\![1, R]\!]$, are sequences of martingale increments w.r.t. the natural filtration of the stochastic approximation schemes $\mathcal{F} := (\mathcal{F}_p := \sigma(\theta_0^{rn}, (U^{rn})^1, \cdots, (U^{rn})^p, r = 1, \cdots, R); p \geq 1)$. Using Taylor's formula we get for $p \geq 0$ and $r \in [\![1, R]\!]$

$$\begin{aligned}
\theta_{p+1}^{rn} - \theta^{*,rn} &= \theta_p^{rn} - \theta^{*,rn} - \gamma_{p+1} h^{rn}(\theta_p^{rn}) + \gamma_{p+1}\Delta M_{p+1}^{rn} \\
&= \theta_p^{rn} - \theta^{*,rn} - \gamma_{p+1} Dh(\theta^*)(\theta_p^{rn} - \theta^{*,rn}) \\
&+ \gamma_{p+1}\left(Dh(\theta^*) - \int_0^1 d\lambda Dh^{rn}(\theta^{*,rn} + (1-\lambda)(\theta_p^{rn} - \theta^{*,rn}))\right)(\theta_p^{rn} - \theta^{*,rn}) \\
&+ \gamma_{p+1}\Delta M_{p+1}^{rn}.
\end{aligned}$$

Hence by a simple induction argument one has for $(r, M) \in [\![1, R]\!] \times \mathbb{N}^*$

$$\theta_M^{rn} - \theta^{*,rn} = \Pi_{1,M}(\theta_0^{rn} - \theta^{*,rn}) + \sum_{k=1}^M \gamma_k \Pi_{k+1,M}\Delta M_k^{rn} + \sum_{k=1}^M \gamma_k \Pi_{k+1,M} R_{k-1}^{rn} \tag{2.22}$$

where $R_k^{rn} = \left(Dh(\theta^*) - \int_0^1 d\lambda Dh^{rn}(\theta^{*,rn} + (1-\lambda)(\theta_k^{rn} - \theta^{*,rn}))\right)(\theta_k^{rn} - \theta^{*,rn})$ and $\Pi_{k,M} := \prod_{j=k}^{M}(I_d - \gamma_j Dh(\theta^*))$, with the convention that $\Pi_{M+1,M} = I_d$. Multiplying (2.22) on the left by $\mathbf{w}_r$ given by (2.16) and summing w.r.t r lead to

$$
\begin{aligned}
-\mathcal{E}_S^{R-R} = \Pi_{1,M}\left(\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn})\right) + \sum_{k=1}^{M}\gamma_k\Pi_{k+1,M}\left(\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn}\right) \\
+ \sum_{k=1}^{M}\gamma_k\Pi_{k+1,M}\left(\sum_{r=1}^{R}\mathbf{w}_r R_{k-1}^{rn}\right). \qquad (2.23)
\end{aligned}
$$

Ought to the Minkowski inequality it is sufficient to bound the $L^1(\mathbb{P})$-norm of each term in the above decomposition. First, since $-Dh(\theta^*)$ is a Hurwitz matrix, $\forall \lambda \in [0, \underline{\lambda})$, there exists $C > 0$ such that for any $k \leq n$, $\|\Pi_{k,n}\| \leq C\prod_{j=k}^{n}(1 - \lambda\gamma_j) \leq C\exp(-\lambda\sum_{j=k}^{n}\gamma_j)$. We refer to [Duf96] and [BMP90] for more details. Hence, one has for all $\eta \in (0, \underline{\lambda})$

$$
\begin{aligned}
\mathbb{E}[|\Pi_{1,M}(\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn}))|] &\leq ||\Pi_{1,M}||\mathbb{E}[|\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn})|] \\
&\leq Ce^{-(\underline{\lambda}-\eta)\sum_{k=1}^{M}\gamma_k}\mathbb{E}[|\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn})|].
\end{aligned}
$$

where $||.||$ stands for the matrix norm on $\mathbb{R}^d \otimes \mathbb{R}^d$. For the second term, recalling that $\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn}$ is a martingale increment, one has

$$
\mathbb{E}\left[\left|\sum_{k=1}^{M}\gamma_k\Pi_{k+1,M}(\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn})\right|^2\right]^{1/2} \leq \left(\sum_{k=1}^{M}\gamma_k^2||\Pi_{k+1,M}||^2\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn}\right|^2\right]\right)^{1/2}. \qquad (2.24)
$$

Similarly for the last term, one has

$$
\mathbb{E}\left[\left|\sum_{k=1}^{M}\gamma_k\Pi_{k+1,M}\left(\sum_{r=1}^{R}\mathbf{w}_r R_{k-1}^{rn}\right)\right|\right] \leq \sum_{k=1}^{M}\gamma_k||\Pi_{k+1,M}||\mathbb{E}\left|\sum_{r=1}^{R}\mathbf{w}_r R_{k-1}^{rn}\right|. \qquad (2.25)
$$

We now study the limit of each bound as n and M go to infinity. For the first term, observe that $\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn}) \xrightarrow{\mathbb{P}} \sum_{r=1}^{R}\mathbf{w}_r(\theta_0 - \theta^*) = \theta_0 - \theta^*$ as $n \to +\infty$. Moreover, since $\sup_{n\geq 1}\mathbb{E}|\theta_0^n|^2 < +\infty$, by uniform integrability one has $\mathbb{E}|\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn} - \theta^{*,rn})| \to \mathbb{E}|\theta_0 - \theta^*|$ as $n \to +\infty$. If $\gamma(p) = \gamma_0/p$ we select η such that

$2(\underline{\lambda}-\eta)\gamma_0 > 1$ otherwise we set $\eta < \underline{\lambda}$ which implies that $\exp(-(\underline{\lambda}-\eta)\sum_{j=k}^{n}\gamma_j) = \gamma^{1/2}(M)\phi_2^R(M)$ with $\phi_2^R(M) \to 0$ as $M \to +\infty$. Hence we get

$$\mathbb{E}[|\Pi_{1,M}(\sum_{r=1}^{R}\mathbf{w}_r(\theta_0^{rn}-\theta^{*,rn}))|] \leq C\gamma^{1/2}(M)\phi_2^R(M).$$

Let us now study the second term. Define for $k \geq 1$, $\Delta N_k^{rn} = h^{rn}(\theta^*) - H(\theta^*, (U^{rn})^k)$ then by the Cauchy-Schwarz inequality and **(HRG)** one has

$$\begin{aligned}
&\left|\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn}\right|^2\right] - \mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r\Delta N_k^{rn}\right|^2\right]\right| \\
&\leq C_R\left(\sum_{r=1}^{R}||\mathbf{w}_r||\mathbb{E}\left[|\Delta M_k^{rn}-\Delta N_k^{rn}|^2\right]\right)^{1/2} \\
&\times\left(\mathbb{E}[|H(\theta_{k-1}^{rn},(U^{rn})^k)|^2]^{1/2} + \mathbb{E}[|H(\theta^*,(U^{rn})^k)|^2]^{1/2}\right) \\
&\leq C_R\max_{1\leq r\leq R}\mathbb{E}[|\theta_{k-1}^{rn}-\theta^*|^{2a}]^{1/2} \\
&\leq C_R(\gamma_k^{a/2}+n^{-a\alpha})
\end{aligned}$$

where we used lemma 5.2 and $\max_{1\leq r\leq R}|\theta^{*,rn}-\theta^*| \leq Cn^{-\alpha}$ for the last inequality. Now observe that $\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r\Delta N_k^{rn}\right|^2\right] = \mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r(h^{rn}(\theta^*)-H(\theta^*,U^{rn}))\right|^2\right]$ so that using **(HC2)** and $U^{rn} \xrightarrow{\mathbb{P}} U^r$ as $n\to+\infty$, one has $\sum_{r=1}^{R}\mathbf{w}_r(h^{rn}(\theta^*)-H(\theta^*,U^{rn})) \xrightarrow{\mathbb{P}} -\sum_{r=1}^{R}\mathbf{w}_rH(\theta^*,U^r)$ as $n\to+\infty$. From **(HUI)** we deduce the L^2-uniform integrability of the family $\left\{\sum_{r=1}^{R}\mathbf{w}_r(h^{rn}(\theta^*)-H(\theta^*,U^{rn})), n\geq 1\right\}$ which yields

$$\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r\Delta N_k^{rn}\right|^2\right] \longrightarrow \mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_rH(\theta^*,U^r)\right|^2\right], \quad n\to+\infty.$$

Plugging the above estimates into (2.24), we derive the following bound

$$\begin{aligned}
&\mathbb{E}\left[\left|\sum_{k=1}^{M}\gamma_k\Pi_{k+1,M}(\sum_{r=1}^{R}\mathbf{w}_r\Delta M_k^{rn})\right|^2\right]^{1/2} \\
&\leq \mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_rH(\theta^*,U^r)\right|^2\right]^{1/2}\left(\sum_{k=1}^{M}\gamma_k^2||\Pi_{k+1,M}||^2\right)^{1/2}(1+\phi_1^R(n)) \qquad (2.26)\\
&+C_R\left(\sum_{k=1}^{M}\gamma_k^2\gamma_k^{a/2}||\Pi_{k+1,M}||^2\right)^{1/2},
\end{aligned}$$

with $\phi_1^R(n) \to 0$ as $n \to +\infty$. Using lemma 5.1, we successively derive that

$$\left(\sum_{k=1}^{M} \gamma_k^2 ||\Pi_{k+1,M}||^2\right)^{1/2} \leq C\gamma^{1/2}(M)$$

for some positive constant $C(\gamma, \underline{\lambda})$ and

$$\left(\sum_{k=1}^{M} \gamma_k^2 \gamma_k^{a/2} ||\Pi_{k+1,M}||^2\right)^{1/2} = o(\gamma^{1/2}(M)) = \gamma^{1/2}(M)\phi_2^R(M)$$

as $M \to +\infty$. We now focus on the last term. Let us first observe that using **(H-R)** and since Dh^{rn} is Lipschitz (uniformly in n) one has

$$\begin{aligned}|R_k^{rn}| = &\Bigg|\Bigg(Dh(\theta^*) - Dh^{rn}(\theta^*) \\ &+ \int_0^1 d\lambda\Big(Dh^{rn}(\theta^*) - Dh^{rn}(\theta^{*,rn} + (1-\lambda)(\theta_k^{rn} - \theta^{*,rn}))\Big)\Bigg)(\theta_k^{rn} - \theta^{*,rn})\Bigg| \\ &\leq C\left(\max_{1\leq r\leq R} ||Dh(\theta^*) - Dh^{rn}(\theta^*)|| + |\theta_k^{rn} - \theta^{*,rn}|\right)|\theta_k^{rn} - \theta^{*,rn}|\end{aligned}$$

so that plugging this estimate in (2.25) and using lemma 5.2 lead to

$$\begin{aligned}&\mathbb{E}\left[\left|\sum_{k=1}^{M} \gamma_k \Pi_{k+1,M}\left(\sum_{r=1}^{R} \mathbf{w}_r R_{k-1}^{rn}\right)\right|\right] \\ \leq\ & C\left(\sum_{k=1}^{M} (\gamma_k^{3/2} \max_{1\leq r\leq R} ||Dh(\theta^*) - Dh^{rn}(\theta^*)|| + \gamma_k^2)||\Pi_{k+1,M}||\right).\end{aligned}$$

Finally lemma 5.1 and since $\max_{1\leq r\leq R} ||Dh(\theta^*) - Dh^{rn}(\theta^*)|| \to 0$ as $n \to +\infty$ also imply

$$\max_{1\leq r\leq R} ||Dh(\theta^*) - Dh^{rn}(\theta^*)|| \left(\sum_{k=1}^{M} \gamma_k^{3/2} ||\Pi_{k+1,M}||\right) \leq C\gamma^{1/2}(M)\phi_1^R(n)$$

and applying again Lemma 5.1 with $a = 1/2$ and $v_k = \gamma_k^{1/2}$, one has:

$$\sum_{k=1}^{M} \gamma_k^2 ||\Pi_{k+1,M}|| = o(\gamma^{1/2}(M)) = \gamma^{1/2}(M)\phi_2^R(M).$$

□

From the previous computations we are naturally led to consider the following suboptimal computational cost optimization problem

$$(n(\epsilon), M(\epsilon)) = argmin_{\mu_R n^{-\alpha R}(1+|\epsilon_{R+1}(n)|)+\nu_R \gamma^{1/2}(M)(1+\phi_1^R(n)+\phi_2^R(M))\leq\varepsilon} \mathrm{Cost(R\text{-}R)} \quad (2.27)$$

where $\mu_R = \frac{|C_R|}{R!^\alpha}$ and $\nu_R = C\mathbb{E}\left[\left|\sum_{r=1}^R \mathbf{w}_r H(\theta^*, U^r)\right|^2\right]^{1/2}$.

Proposition 2.5. *(Computational cost optimization) Let $R \in \mathbb{N}^*$. Suppose that the assumptions of Proposition 2.4 are satisfied. Suppose that the step sequence γ is given by: $\gamma(p) = \gamma_0/p^\beta$, $\gamma_0 > 0$, $p > 0$, $\beta \in (1/2, 1]$. The multi-step Richardson-Romberg SA estimator of order R satisfies*

$$\inf_{\mu_R n^{-\alpha R}(1+|\epsilon_{R+1}(n)|)+\nu_R \gamma^{1/2}(M)(1+\phi_1^R(n)+\phi_2^R(M))\leq\varepsilon} \mathrm{Cost(R\text{-}R)}$$

$$\sim K\frac{R(R+1)}{2}\gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\mu_R^{\frac{1}{\alpha R}}\frac{1}{\varepsilon^{\frac{2}{\beta}+\frac{1}{\alpha R}}}\left(1+\frac{2\alpha R}{\beta}\right)^{\frac{1}{\alpha R}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}$$

as $\varepsilon \to 0$. Eventually this asymptotically optimal bound may be achieved with parameters satisfying:

$$n(\varepsilon) \sim \left(\frac{2\alpha R}{\beta}+1\right)^{\frac{1}{\alpha R}}\mu_R^{\frac{1}{\alpha R}}\varepsilon^{-\frac{1}{\alpha R}} \quad \textit{and} \quad M(\varepsilon) \sim \gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}} \quad \textit{as } \varepsilon \to 0. \quad (2.28)$$

Proof. Let us note that the cost minimization problem (2.27) is lower-bounded by the more tractable problem

$$\inf_{\mu_R n^{-\alpha R}+\nu_R\gamma^{1/2}(M)\leq\varepsilon} \mathrm{Cost(R\text{-}R)} = \inf_{\mu_R n^{-\alpha R}<\varepsilon} K\gamma^{-1}\left(\frac{(\epsilon-\mu_R n^{-\alpha R})^2}{\nu_R^2}\right) n\frac{R(R+1)}{2} \quad (2.29)$$

with $M = \gamma^{-1}\left(\frac{(\varepsilon-\mu_R n^{-\alpha R})^2}{\nu_R^2}\right) = \gamma_0^{1/\beta}\nu_R^{2/\beta}(\varepsilon-\mu_R n^{-\alpha R})^{-2/\beta}$. This optimization problem can be solved explicitly, more precisely the optimal parameters are given by

$$n(\varepsilon) = \left(\frac{2\alpha R}{\beta}+1\right)^{\frac{1}{\alpha R}}\mu_R^{\frac{1}{\alpha R}}\varepsilon^{-\frac{1}{\alpha R}}, \quad M(\varepsilon) = \gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}}.$$

The "liminf" side of the result clearly follows by plugging this solution into (2.29). Now set

$$\begin{aligned} n(\varepsilon) &= \left(\frac{2\alpha R}{\beta}+1\right)^{\frac{1}{\alpha R}}\mu_R^{\frac{1}{\alpha R}}\varepsilon^{-\frac{1}{\alpha R}}, \\ M(\varepsilon) &= \gamma^{-1}\left(\frac{(\varepsilon-\mu_R(1+|\varepsilon_{R+1}(n(\varepsilon))|)n^{-\alpha R}(\varepsilon))^2}{\nu_R^2\left(1+\frac{\beta}{2\alpha R}\right)^2\left(1+\phi_1^R(n(\varepsilon))+\phi_2^R(\gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}})\right)^2}\right). \end{aligned}$$

Since ϕ_2^R is non-increasing, the couple $(n(\varepsilon), M(\varepsilon))$ satisfies the constraint

$$\mu_R n^{-\alpha R}\left(1+|\epsilon_{R+1}(n)|\right)+\nu_R\gamma^{1/2}(M)(1+\phi_1^R(n)+\phi_2^R(M))\leq\varepsilon$$

so that the cost minimization problem (2.27) is upper-bounded by

$$\begin{aligned}K\frac{R(R+1)}{2}\mu_R^{\frac{1}{\alpha R}}\varepsilon^{-\frac{1}{\alpha R}}\left(1+\frac{2\alpha R}{\beta}\right)^{\frac{1}{\alpha R}}\gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}&\\ \times\varepsilon^{-\frac{2}{\beta}}\left(1-(1+|\varepsilon_{R+1}(n(\varepsilon))|)\frac{\beta}{2\alpha R+\beta}\right)^{\frac{2}{\beta}}&\\ \times\left(1+\phi_1^R(n(\varepsilon))+\phi_R^2(\gamma_0^{\frac{1}{\beta}}\nu_R^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha R}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}})\right)^{\frac{2}{\beta}}&\end{aligned}$$

and the result follows by letting ε goes to zero.

□

Remark 2.2. (Choice of the step sequence) According to Proposition 2.5, it is optimal to set $\beta=1$ to achieve a minimal asymptotic complexity. In this case a constraint appear on γ_0: $2\underline{\lambda}\gamma_0>1$. Let us note that for $\beta=1$ a simple computation shows that the constant C appearing in ν_R is equal to $\gamma_0/(2\underline{\lambda}\gamma_0-1)^{1/2}$ which reaches its minimum (as a function of γ_0) at $\gamma_0=1/\underline{\lambda}$. However the main drawback with this choice is that the constant $\underline{\lambda}$ is not known to the experimenter so that one is led to make a blind choice in practical implementation.

Remark 2.3. (Control of the variance) Let us note that when one decides to implement the Richardson-Romberg extrapolation SA scheme with an innovation satisfying $U^r=U$ a.s. $r=1,\cdots,R$ then one has $H(\theta^*,U^r)=H(\theta^*,U)$ *a.s.* for every $r\in[\![1,R]\!]$ so that using (2.14) yields

$$\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_rH(\theta^*,U^r)\right|^2\right]=\mathbb{E}\left[\left|(\sum_{r=1}^{R}\mathbf{w}_r)H(\theta^*,U)\right|^2\right]=\mathbb{E}\left[|H(\theta^*,U)|^2\right].$$

Hence we clearly see that this choice leads to a control in the L^1-norm of the statistical error of the multi-step Richardson-Romberg SA estimator. On the opposite considering mutually independent innovations U^r lead to an explosion of the previous control with respect to R. Indeed one has

$$\mathbb{E}\left[\left|\sum_{r=1}^{R}\mathbf{w}_r H(\theta^*,U^r)\right|^2\right] = \left(\sum_{r=1}^{R}\frac{r^{2\alpha R}}{\prod_{j=0}^{r-1}(r^\alpha - j^\alpha)^2\prod_{j=r+1}^{R}(j^\alpha - r^\alpha)^2}\right)\mathbb{E}\left[|H(\theta^*,U)|^2\right]$$
$$\geq \left(\frac{R^R}{R!}\right)^{2\alpha}\mathbb{E}\left[|H(\theta^*,U)|^2\right]$$
$$\sim \left(\frac{e^R}{\sqrt{2\pi}\sqrt{R}}\right)^{2\alpha}\mathbb{E}\left[|H(\theta^*,U)|^2\right] \quad \text{as } R\to+\infty,$$

where we used (2.16) for the first equality.

For instance when one is concerned with the discretization of a Brownian diffusion, the first aforementioned case consists in implementing the Richardson-Romberg method with R Euler schemes devised with the *same Brownian motion* W namely $W^r = W, \ \ r = 1,\cdots,R$ whereas the second case consists in implementing the method with mutually independent Brownian motions W^r. The optimality of this choice is discussed in [Pag07].

2.3 Comparison with the crude stochastic approximation estimator

Under the assumptions of Proposition 2.4 with $R = 1$, the global error for the crude SA estimator satisfies

$$\begin{aligned}\mathbb{E}\left[|\mathcal{E}_{glob}(M,\gamma,H)|\right] &= \mathbb{E}\left[|\theta^* - \theta^{*,n} + \theta^{*,n} - \theta^n_M|\right]\\ &\leq \frac{|C_1|}{n^\alpha}(1+|\varepsilon_1(n)|) + \mathbb{E}\left[|\theta^{*,n} - \theta^n_M|\right]\\ &\leq \frac{|C_1|}{n^\alpha}\left(1+|\epsilon_1(n)|\right)\\ &+ C\mathbb{E}\left[|H(\theta^*,U)|^2\right]^{\frac{1}{2}}\gamma^{\frac{1}{2}}(M)(1+\phi_1(n)+\phi_2(M)),\end{aligned}$$

with a computational cost given by $\mathrm{Cost(C\text{-}S)} := KMn$. Hence a similar result as in Proposition 2.5 holds.

Proposition 2.6. *Assume that the assumptions of Proposition 2.4 with $R = 1$ hold. Suppose that the step sequence γ is given by: $\gamma(p) = \gamma_0/p^\beta$, $\gamma_0 > 0$, $p > 0$, $\beta \in (1/2, 1]$. The crude SA estimator satisfies*

$$\inf_{|C_1|n^{-\alpha}(1+|\epsilon_1(n)|)+\nu_1\gamma^{1/2}(M)(1+\phi_1(n)+\phi_2(M))\leq\varepsilon}\mathrm{Cost(C\text{-}S)}$$
$$\sim K\gamma_0^{\frac{1}{\beta}}\nu_1^{\frac{2}{\beta}}|C_1|^{\frac{1}{\alpha}}\frac{1}{\varepsilon^{\frac{2}{\beta}+\frac{1}{\alpha}}}\left(1+\frac{2\alpha}{\beta}\right)^{\frac{1}{\alpha}}\left(1+\frac{\beta}{2\alpha}\right)^{\frac{2}{\beta}}$$

as $\varepsilon \to 0$ *with* $\nu_1 = C\mathbb{E}\left[|H(\theta^*,U)|^2\right]^{\frac{1}{2}}$. *Eventually this asymptotically optimal bound may be achieved with parameters satisfying:*

$$n(\varepsilon) \sim \left(\frac{2\alpha}{\beta}+1\right)^{\frac{1}{\alpha}} |C_1|^{\frac{1}{\alpha}}\varepsilon^{-\frac{1}{\alpha}} \quad \text{and} \quad M(\varepsilon) \sim \gamma_0^{\frac{1}{\beta}}\nu_1^{\frac{2}{\beta}}\left(1+\frac{\beta}{2\alpha}\right)^{\frac{2}{\beta}}\varepsilon^{-\frac{2}{\beta}} \quad \text{as } \varepsilon \to 0. \tag{2.30}$$

3 Application: Estimation of the quantile of a component of an SDE

In this section, we show how the previous results can be applied to the estimation of the quantile of a stochastic process solution to a stochastic differential equation. Also, when the exact value of a constant is not important we may repeat the same symbol for constants that may change from one line to next.

3.1 Notations and Hypotheses.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions and $(Z_t)_{t\geq 0}$ be a d-dimensional $(\mathcal{F}_t)_{t\geq 0}$ symmetric α-stable process, for $\alpha \in (0,2]$, that is a càdlàg process with independent and stationary increments with the scaling property $Z_{ct} \overset{(d)}{=} c^{1/\alpha}Z_t$. Note that the case $\alpha = 2$ corresponds to the standard Brownian motion. It is also the only case where Z is a continuous process. When $\alpha < 2$, the Stable process is discontinuous and its Lévy-Khintchine exponent writes for all $p \in \mathbb{R}^d$,

$$\mathbb{E}\left(e^{i\langle p, Z_t\rangle}\right) = \exp\left(-t\int_{S^{d-1}} |\langle p, \vartheta\rangle|^\alpha \mu(d\vartheta)\right).$$

We refer to the measure μ as the *spectral measure* of Z. It is related to the Lévy measure of the process Z as follows. Denote by ν the Lévy measure of Z, ν factorizes in $\nu(dz) = C_\alpha \frac{d|z|}{|z|^{1+\alpha}}\mu(\bar{z})$, where $z = (|z|, \bar{z}) \in \mathbb{R}_+ \times S^{d-1}$ stands for the polar coordinates. For the exact value of C_α, we refer to Sato [Sat05]. Let us consider a d-dimensional process $(X_t)_{t\geq 0} = (X_t^1, \dots, X_t^d)_{t\geq 0}$ with dynamics:

$$X_t = x + \int_0^t b(X_{s-})ds + \int_0^t \sigma(X_{s-})dZ_s, \tag{3.31}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$. We fix the time horizon $T = 1$. Let us denote by $\mathbb{P}_x$ (resp. $\mathbb{P}_{t,x}$, $t \in (0,1]$) the conditional probability given $\{X_0 = x\}$ (resp. $\{X_t = x\}$). For a given level $\ell \in (0,1)$, we are interested in the computation of the quantile at level ℓ of the random variable X_1^d defined as:

$$\theta^* = \inf\{\theta \in \mathbb{R} : \mathbb{P}_x(X_1^d \leq \theta) \geq \ell\}.$$

Since $\lim_{\theta\to+\infty}\mathbb{P}_x(X_1^d \le \theta) = 1$, we have $\{\theta \in \mathbb{R} : \mathbb{P}_x(X_1^d \le \theta) \ge \ell\} \ne \emptyset$. Moreover, we have $\lim_{\theta\to-\infty}\mathbb{P}_x(X_1^d \le \theta) = 0$, which implies that $\{\theta \in \mathbb{R} : \mathbb{P}_x(X_1^d \le \theta) \ge \ell\}$ is bounded from below so that θ^* always exists. Assuming that the distribution of X_1^d has no atoms, the quantile at level ℓ is the lowest solution of the equation:

$$\mathbb{P}_x(X_1^d \le \theta) = \ell.$$

If the distribution function is (strictly) increasing, which is notably the case if the process X solution of (3.31) admits a positive density $p(1, x, .)$, the solution to the above equation is unique, otherwise, there may be more than one solution. Now since the law of X_1^d is not known explicitly, the quantile θ^* cannot be computed and one has to approximate the dynamics by a discretization scheme that can be simulated. Let us note that the estimation of the quantile of a component of a Brownian diffusion process has already been investigated in [TZ04]. For a given time step $\Delta = \frac{1}{n}$, $n \in \mathbb{N}^*$, setting for all $i \in \mathbb{N}$, $t_i = i\Delta$, we consider the standard Euler scheme defined as follows:

$$X_t^n = x + \int_0^t b(X_{\phi(s)}^n)ds + \int_0^t \sigma(X_{\phi(s)}^n)dZ_s \quad \phi(s) = \sup\{t_i : \ t_i \le s\}. \tag{3.32}$$

Then one approximates θ^* by $\theta^{*,n}$ the quantile at level ℓ of $X_1^{n,d}$. We denote by **[A]** the following set of assumptions. Fix an integer $m \in \mathbb{N}$ which will hereafter refer to the regularity of the coefficients.

[A-1] $b \in \mathcal{C}^m(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in \mathcal{C}^m(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ with bounded derivatives. Also, when $\alpha \le 1$, we put $b = 0$.

[A-2] When $\alpha < 2$ for all $x, \xi \in \mathbb{R}^d$, there exists $C > 1$ such that:

$$C^{-1}|\xi|^2 \le \langle \xi, \sigma(x)\xi\rangle \le C|\xi|^2.$$

When $\alpha = 2$, setting $\Sigma(x) = \sigma(x)\sigma(x)^T$, for all $x, \xi \in \mathbb{R}^d$, there exists $C > 1$ such that:

$$C^{-1}|\xi|^2 \le \langle \xi, \Sigma(x)\xi\rangle \le C|\xi|^2.$$

[A-3] When $\alpha < 2$, the spectral measure μ has a $\mathcal{C}^m(S^{d-1})$ surface density and satisfies: for all $\xi \in \mathbb{R}^q$, there exists $C > 1$ such that:

$$C^{-1}|\xi|^\alpha \le \int_{S^{d-1}} |\langle \xi, \vartheta\rangle|^\alpha \mu(d\vartheta) \le C|\xi|^\alpha. \tag{3.33}$$

Proposition 3.1. *Assume that $\alpha \in (0, 2]$ and that **[A]** is in force. For every $t > 0$, the solutions X_t, X_t^n, of the SDE* (3.31) *and* (3.32) *have a strictly positive densities with respect to the Lebesgue measure. Consequently, the quantile is uniquely defined. Moreover, those densities are in $\mathcal{C}^m(\mathbb{R}^d, \mathbb{R}^d)$ if $\alpha > 1$, and in $\mathcal{C}^{m-1}(\mathbb{R}^d, \mathbb{R}^d)$ when $\alpha \le 1$.*

We refer to the work of Kolokoltsov [Kol00b] for the proof in the Stable case, who also derived Aronson's estimates with time singularity depending on the index α. In the Brownian case, i.e. $\alpha = 2$, if the drift b is a measurable bounded function and the diffusion coefficient σ is η-Hölder continuous, $\eta > 0$, and satisfies **[A-2]** then the aforementioned densities exists, are positive and satisfy Gaussian Aronson's estimates (see e.g. [Fri64] and [LM10] for the density of the Euler scheme).

Proposition 3.2. *For $\alpha \in (0,2)$ assume that **[A]** for $m \geq 2$. For $\alpha = 2$, assume the drift b and the diffusion coefficient σ are Lipschitz-continuous bounded functions and that σ satisfies **[A-2]**.Then one has*

$$\theta^{*,n} \to \theta^*, \ n \to +\infty.$$

Proof. Let $n \in \mathbb{N}^*$ and denote by F, F^n the distribution function of X_1^d and $X_1^{n,d}$ respectively. Since b and σ are Lipschitz we know that $(X_1^{n,d})_{n\geq 1}$ converges in distribution to X_1^d. Moreover, the function F is continuous so that $(F^n)_{n\geq 1}$ converges uniformly to F. Hence, we conclude that $F(\theta^{*,n}) \to \ell$, $n \to +\infty$. Now remark that from Proposition 3.1 since X_1^d has a strictly positive density the function F is one-to-one which in turn implies that F^{-1} exists and is continuous so that $\theta^{*,n} \to F^{-1}(\ell) = \theta^*$. $\square$

From Proposition 3.1 (existence of a positive density for $X_1^{n,d}$) the quantile $\theta^{*,n}$ at level ℓ of the random variable $X_1^{n,d}$ is the unique solution of the equation

$$\mathbb{P}_x\left(X_1^{n,d} \leq \theta\right) = \ell.$$

In this section, we are interested in giving an expansion for the error $\theta^* - \theta^{*,n}$ in powers of n^{-1}, using Theorem 2.3. Actually, we will prove that **[A]** implies **[H-k]**, for a desired $k > 0$. As we can see, Theorem 2.3 requires an expansion of $h^n - h$ and its derivatives up to order $k > 0$ in order to have an expansion of $\theta^* - \theta^{*,n}$ at the same order. Regularity of the function h may be obtained mainly by two means: either the function H is smooth w.r.t. the variable θ (with polynomial growth w.r.t θ and x) or the laws of X_T and X_T^n are smooth. Concerning the expansion of the difference $\partial_\theta^k h - \partial_\theta^k h^n$ it may also be obtained by two means: in the regular setting i.e. when the function $x \mapsto \partial_\theta^k H(\theta, x)$ and the coefficients b and σ are regular (say b, σ, $\partial_\theta^k H(\theta, .)$ are $\mathcal{C}_b^{R+5}$) one may use standard tools such as the one developed in Talay-Tubaro [TT90] (in the Brownian case); or in the (Hypo-)elliptic setting, the laws of X_T and X_T^n are smooth. Here, we are in the latter case. Indeed, the estimation of the quantile of a diffusion can be seen as an inverse problem, by setting $H(\theta, x) = 1 - \frac{1}{1-\ell}\mathbf{1}_{\{x^d \geq \theta\}}$. We thus see that regularity of H fails. However, for $\theta \in \mathbb{R}$, we have:

$$h(\theta) - h^n(\theta) = \frac{1}{1-\ell}\left(\mathbb{P}^x(X_1^d \leq \theta) - \mathbb{P}^x(X_1^{n,d} \leq \theta)\right).$$

Let $p(T,x,\theta)$ be the density of the diffusion, and $p_n(T,x,\theta)$ the density of the Euler scheme at time T. The derivative w.r.t. θ of the previous equality is $\forall k \geq 1, \forall(\theta,x) \in \mathbb{R}\times\mathbb{R}^d$:

$$\frac{d^k}{d\theta^k}h(\theta) - \frac{d^k}{d\theta^k}h^n(\theta) = \frac{1}{1-\ell}\left(\frac{\partial^{k-1}}{\partial\theta^{k-1}}p^{X_1^d}(1,x,\theta) - \frac{\partial^{k-1}}{\partial\theta^{k-1}}p_n^{X_1^{n,d}}(1,x,\theta)\right),$$

where we denote by $p^{X_1^d}(1,x,\theta)$ and $p_n^{X_1^{n,d}}(1,x,\theta)$ the marginal densities of X_1^d and $X_1^{n,d}$. Consequently, we observe that in order to apply Theorem 2.3, we have to give an expansion of the marginal densities and their derivatives, up to an order $k > 1$. Actually, we will show that the expansion holds for $p(1,x,\theta) - p_n(1,x,\theta)$ and its derivatives, the expansion for the marginals will follow from an integration over the $d-1$ first components.

3.2 Expansion for the densities.

Using a continuity technique known as the Parametrix expansion, Konakov and Mammen [KM02], in the Brownian case, and Konakov and Menozzi [KM10], in the stable case, successfully derive an expansion for the density of the solution of (3.31) to an arbitrary order, with explicit terms. The purpose of this section is to extend these results to the derivatives of the densities.

The Parametrix expansion consists in representing the density of the solution of (3.31) as a series involving the density of a *frozen* equation and the generators associated with (3.31) and the *frozen* density. We take a few lines here to describe this technique.

We define the following process as the frozen process. Recall $T = 1$ is a fixed deterministic time. For a given terminal point $y \in \mathbb{R}^d$, the frozen equation at point y is defined as:

$$\tilde{X}_t = x + b(y)t + \sigma(y)Z_t. \tag{3.34}$$

Thanks to the uniform ellipticity of σ, the process (3.34) has a density with respect to the Lebesgue measure. Recalling that $\Sigma(z) = \sigma(z)\sigma(z)^T$, the density is given by:

$$\tilde{p}_\alpha^y(t,x,y) = \begin{cases} \frac{\det(\Sigma(y))^{-1/2}}{(2\pi t)^{d/2}}\exp\left(-\frac{1}{2t}(y-x-b(y)t)^T\Sigma(y)^{-1}(y-x-b(y)t)\right), & \text{if } \alpha = 2 \\ \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dp e^{-i\langle p, y-x-b(y)t\rangle}\exp\left(-t\int_{S^{d-1}}|\langle p,\sigma(y)\vartheta\rangle|^\alpha\mu(d\vartheta)\right), & \text{if } \alpha \in (0,2). \end{cases}$$

We will often drop the superscript y with the convention $\tilde{p}_\alpha(t,x,y) = \tilde{p}_\alpha^y(t,x,y)$, when no ambiguity is possible. The distance between $p(t,x,y)$ and $\tilde{p}_\alpha(t,x,y)$ will then be quantified by the difference of the generators of (3.31) and (3.34). The generator of the SDE (3.31):

$$Lf(t,x,y)$$
$$= \begin{cases} \frac{1}{2}\mathrm{tr}\left(\Sigma(x)\partial_x^2 f(t,x,y)\right) + \langle b(x), \partial_x f(t,x,y)\rangle, & \text{if } \alpha = 2\ , \\ \langle b(x), \partial_x f(t,x,y)\rangle & \\ + \int_{\mathbb{R}^d} f(t, x+\sigma(x)z, y) - f(t,x,y) - \frac{\langle \nabla_x f(t,x,y), \sigma(x)z\rangle}{1+|z|^2}\nu(dz), & \text{if } \alpha \in (0,2). \end{cases}$$

Let us define the generator of the frozen process (3.34):

$$\tilde{L}^* f(t,x,y)$$
$$= \begin{cases} \frac{1}{2}\mathrm{tr}\left(\Sigma(y)\partial_x^2 f(t,x,y)\right) + \langle b(y), \partial_x f(t,x,y)\rangle, & \text{if } \alpha = 2, \\ \langle b(y), \partial_x f(t,x,y)\rangle & \\ + \int_{\mathbb{R}^d} f(t, x+\sigma(y)z, y) - f(t,x,y) - \frac{\langle \nabla_x f(t,x,y), \sigma(y)z\rangle}{1+|z|^2}\nu(dz), & \text{if } \alpha \in (0,2). \end{cases}$$

When $\alpha = 2$, these are differential operators of order 2. For $\alpha \in (0,2)$, these operators should be seen as fractional derivative of order α.

Theorem 3.3. *Under the assumptions **[A]**, the solution of* (3.31) *exists and has density with respect to the Lebesgue measure. Let $p(t,x,y)$ denote the density of* (3.31)*. It admits the following representation:*

$$p(t,x,y) = \sum_{k=0}^{\infty} \tilde{p}_\alpha \otimes H^{(k)}(t,x,y),$$

where we denoted $H(t,x,y) = (L - \tilde{L}^*)\tilde{p}_\alpha(t,x,y)$, and $\otimes$ is the space-time convolution:

$$f \otimes g\,(t,x,y) = \int_0^t \int_{\mathbb{R}^d} f\,(u,x,z)\, g\,(t-u,z,y)\, dzdu,$$

and $H^{(k)}(t,x,y) = H^{(k-1)} \otimes H(t,x,y)$, and $\tilde{p}_\alpha \otimes H^{(0)}(t,x,y) = \tilde{p}_\alpha(t,x,y)$.

This result has been investigated in the literature, let us mention Friedman [Fri64] for the Brownian case and Kolokoltsov [Kol00b] for the stable case. The proof relies on a precise study of the frozen density and its derivatives (fractional derivatives in the stable case), and show that in the time space convolution, the time singularities induced by the derivation can be compensated to get a convergent series.

Similarly, one gets an equivalent result for the density of the Euler scheme. We introduce the "frozen Markov chains" $(\tilde{X}^n_{t_k})_{k\in[\![0,n]\!]}$:

$$\tilde{X}^n_{t_k} = x,\ \tilde{X}^n_{t_{k+1}} = \tilde{X}^n_{t_k} + b(y)\Delta + \sigma(y)(Z_{t_{k+1}} - Z_{t_k}).$$

We denote the discrete generators:

$$\begin{aligned} & L_n f(t_k - t_j, x, y) \\ = \ & \Delta^{-1}\left(\int p_n(\Delta, x, z) f(t_k - t_{j+1}, z, y)dz - f(t_k - t_{j+1}, x, y)\right), \quad (3.35) \\ & \tilde{L}^*_n f(t_k - t_j, x, y) \\ = \ & \Delta^{-1}\left(\int \tilde{p}^y(\Delta, x, z) f(t_k - t_{j+1}, z, y)dz - f(t_k - t_{j+1}, x, y)\right). \quad (3.36) \end{aligned}$$

We then obtain a representation of the density of the Euler scheme using the frozen density and the discrete generators.

Theorem 3.4. *The density $p_n(t_k, x, y)$ of the Euler scheme admits the following representation:*

$$p_n(t_k - t_j, x, y) = \sum_{r=0}^{k-j} \tilde{p}_\alpha \otimes_n H_n^{(r,n)}(t_k - t_j, x, y),$$

where we denoted $H_n(t_k, x, y) = (L_n - \tilde{L}_n)\tilde{p}(t_k, x, y)$, and $\otimes_n$ is the discretized space-time convolution:

$$f \otimes_n g\,(t_k, x, y) = \frac{1}{n}\sum_{i=0}^{k-1}\int_{\mathbb{R}^d} f\,(t_i, x, z)\, g\,(t_k - t_i, z, y)\, dz,$$

and $H_n^{(r,n)}(t_k, x, y) = H_n^{(r-1,n)} \otimes_n H_n(t_k, x, y)$, where $\tilde{p}_\alpha \otimes H_n^{(0,n)}(t_k, x, y) = \tilde{p}_\alpha(t_k, x, y)$.

Remark 3.1. We use the notation $H_n^{(r,n)}(t_k, x, y)$ to emphasize the dependency in the discretization of the convolution. That is, the subscript n refers to the discrete generators, whereas the super script (r, n) refers respectively to the number of steps we iterate the convolution, and the number of discretization dates. Therefore, we have $H_n^{(1,n)}(t_k, x, y) = H_n(t, x, y)$. Also, using the convention $H_n^{(r,n)} = 0$ for $r > k - j$, we can write $p_n(t_k - t_j, x, y) = \sum_{r=0}^{+\infty} \tilde{p}_\alpha \otimes_n H_n^{(r,n)}(t_k - t_j, x, y)$.

Once again, these results have been investigated in the literature and we state them here without proof. The reader may consult [KM02, KM10] and the references therein.

Roughly speaking, we see that the differences between the two expansions of Theorems 3.3 and 3.4 come from the convolution and the kernel. Thus, in order to get an expansion for $p - p_n$, we introduce for all $k \in [\![0, n-1]\!]$:

$$\begin{aligned} p^d(t_k, x, y) &= \sum_{r=0}^{+\infty} \tilde{p}_\alpha \otimes_n H^{(r,n)}(t_k, x, y), \\ H^{(r,n)}(t_k, x, y) &= H^{(r-1,n)} \otimes_n H(t_k, x, y), \text{ where } \tilde{p}_\alpha \otimes H^{(0,n)}(t_k, x, y) = \tilde{p}_\alpha(t_k, x, y). \end{aligned}$$

Formally speaking, p^d is the series of Theorem 3.3, with discretized time integrals. We then look for an expansion for the two differences $p - p_n = p - p^d + p^d - p_n$. To that end, we define $\tilde{L}_* f(t, x, y) = \tilde{L}_x f(t, x, y)$, where :

$$\tilde{L}_\xi f(t, x, y) = \begin{cases} \frac{1}{2}\mathrm{tr}\left(\Sigma(\xi)\partial_x^2 f(t, x, y)\right) + \langle b(\xi), \partial_x f(t, x, y)\rangle, & \text{if } \alpha = 2, \\ \langle b(\xi), \partial_x f(t, x, y)\rangle - \int_{S^{d-1}} |\langle \partial_x, \sigma(\xi)\vartheta\rangle|^\alpha f(t, x, y)\mu(d\vartheta), & \text{if } \alpha \in (0, 2). \end{cases}$$

Note that both generators $\tilde{L}^*$ and $\tilde{L}_*$ depends on the freezing parameter y. This induces extra caution below, as we will be led to differentiate with respect to the freezing parameter.

Extending the results of Theorem 1.1 in Konakov and Mammen in [KM02], for the Brownian case, and Theorem 21 in Konakov and Menozzi in [KM10], for the Stable case, we have the following result.

Theorem 3.5. *Assume that **[A]** holds. Let $M \in \mathbb{N}^*$ be such that when $\alpha = 2$, $0 < M \le m/2$, and when $\alpha < 2$, we assume $m > d + 4$ and $0 < M \le m - (d + 4)$. Let $\gamma \in \mathbb{N}^d$, with $|\gamma| \le M$. Then, for all $x, y \in \mathbb{R}^d$, we have:*

$$\begin{aligned} &\partial_y^\gamma p(1, x, y) - \partial_y^\gamma p_n(1, x, y) \\ = &\sum_{k=1}^{M-1-|\gamma|} \frac{1}{(k+1)!n^k} \partial_y^\gamma \left(p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d\right)(1, x, y) \\ &- \frac{1}{(k+1)!n^k} \partial_y^\gamma \left(p^d \otimes_n \left(\tilde{L}_* - \tilde{L}^*\right)^{k+1} p_n\right)(1, x, y) + \frac{\partial_y^\gamma R(x, y)}{n^{M-|\gamma|}}. \end{aligned} \tag{3.37}$$

*Also, there is a constant $C > 0$ depending on the set of assumptions **[A]**, T, γ, and M such that the following bound holds for each term and the remainders:*

$$\begin{aligned} C\bar{p}_K^\alpha(t, x, y) \ \ge \ &\sum_{k=1}^{M-|\gamma|-1} \left|\partial_y^\gamma \left(p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d\right)(1, x, y)\right| \\ &+ \left|\partial_y^\gamma \left(p^d \otimes_n \left(\tilde{L}_* - \tilde{L}^*\right)^{k+1} p_n\right)(1, x, y)\right| \\ &+ |\partial_y^\gamma R(x, y)|, \end{aligned} \tag{3.38}$$

where for a given $K > 0$, we denoted $\bar{p}_K^\alpha(t,x,y)$ the following quantity:

$$\bar{p}_K^\alpha(t,x,y) = \begin{cases} t^{-d/2}\exp\left(-K\frac{|y-x|^2}{t}\right), & \text{if } \alpha = 2, \\ \frac{t^{-d/\alpha}}{\left[K \vee \frac{|y-x|}{t^{\frac{1}{\alpha}}}\right]^{d+\alpha}}, & \text{if } \alpha \in (0,2). \end{cases}$$

For $\gamma = 0$, expansion (3.37) is given in [KM02] in the Brownian case, and in [KM10] in the stable case. To get an expansion for $\partial_y^\gamma(p - p_n)(1,x,y)$, we take the derivative along y in each term in that expansion, and prove that each one is bounded by an α stable density.

Formally, $\bar{p}_K^\alpha(t,x,y)$ is a stable density (up to some normalizing constant depending on $K > 0$). Observe that $\bar{p}^\alpha$ satisfies a semi-group property in the following sense:

Proposition 3.6. *For all $\tau \in (0,t)$, for all $x, y \in \mathbb{R}^d$ for all $K_1, K_2 > 0$, there exists $K, C > 0$ depending on the set of assumptions **[A]** and the terminal time T, such that:*

$$\int_{\mathbb{R}^d} \bar{p}_{K_1}^\alpha(\tau,x,z)\bar{p}_{K_2}^\alpha(t-\tau,z,y)dz \le C\bar{p}_K^\alpha(t,x,y). \tag{3.39}$$

Proof. Indeed, for all $\alpha \in (0,2]$, we have that for $t > 0$, for all $x, y \in \mathbb{R}^d$, there exists $c, C, K > 0$ such that:

$$c\bar{p}_K^\alpha(t,x,y) \le \tilde{p}_\alpha^y(t,x,y) \le C\bar{p}_K^\alpha(t,x,y). \tag{3.40}$$

For the gaussian case, we refer to the seminal paper [Fri64] or Sheu [She91] for a stochastic control based approach. For the stable case $\alpha < 2$, the reader may consult and Kolokolstov [Kol00b]. Thus, one easily gets:

$$\begin{aligned} \int_{\mathbb{R}^d} \bar{p}_{K_1}^\alpha(\tau,x,z)\bar{p}_{K_2}^\alpha(t-\tau,z,y)dz &\le C\int_{\mathbb{R}^d} \tilde{p}_\alpha^y(\tau,x,z)\tilde{p}_\alpha^y(t-\tau,z,y)dz \\ &= C\tilde{p}_\alpha^y(t,x,y) \le C\bar{p}_K^\alpha(t,x,y). \end{aligned}$$

□

Using the previous density, we are able to bound the various terms appearing above.

Lemma 3.7. *For all multi index $\gamma, \eta \in \mathbb{N}^d$ such that $|\gamma| + |\eta| \le m$ if $\alpha > 1$, and $|\gamma| + |\eta| \le m - 1$ if $\alpha \le 1$, for all $x, y \in \mathbb{R}^d$, for all $t \in [0,T]$, for all $k \in [\![0, n-1]\!]$, there exists $C = C(\textbf{[A]}, T, \gamma, \eta) > 0$ such that the following bounds holds:*

$$|\partial_x^\gamma\partial_y^\eta p^d(t_k,x,y)| + |\partial_x^\gamma\partial_y^\eta p_n(t_k,x,y)| \le Ct_k^{-\frac{|\gamma|+|\eta|}{\alpha}}\bar{p}_K^\alpha(t_k,x,y), \tag{3.41}$$

$$|\partial_x^\gamma\partial_y^\eta p(t,x,y)| \le Ct^{-\frac{|\gamma|+|\eta|}{\alpha}}\bar{p}_K^\alpha(t,x,y), \tag{3.42}$$

Moreover, for all $\xi \in \mathbb{R}^d$,

$$|\partial_x^\gamma p^d(t_k, x, x+\xi)| + |\partial_x^\gamma p_n(t_k, x, x+\xi)| \le C\bar{p}_K^\alpha(t_k, x, x+\xi). \tag{3.43}$$

Eventually, when $\alpha < 2$, *denoting* $\Phi(t_k, x, y) = \sum_{r=1}^\infty H^{(r,n)}(t_k, x, y)$, *we have:*

$$\begin{aligned} \left|\partial_x^\gamma \partial_y^\eta \Phi(t_k, x, y)\right| &\le Ct_k^{-\frac{|\gamma|+|\eta|}{\alpha}} \bar{p}_K^\alpha(t_k, x, y)\left(1 + \frac{1 \wedge |x-y|}{t_k}\right), & (3.44)\\ |\partial_x^\gamma \Phi(t_k, x, x+\xi)| &\le C\bar{p}_K^\alpha(t_k, x, x+\xi)\left(1 + \frac{1 \wedge |\xi|}{t_k}\right). & (3.45)\end{aligned}$$

Remark 3.2. We point out that in equations (3.43) and (3.45), despite the presence of derivations, there are no singularities induced by them, as the derivation argument appears in both the forward and the backward arguments. This will be a key point in the proof of Theorem 3.5.

Proof. For the Brownian case, all the above estimates are proved in [KM02]. We thus focus on the stable case. In Konakov Menozzi [KM10], the bound (3.41) and (3.42) are given. To get the bound (3.43), we prove (3.44) and (3.45), using the following estimates proved in [KM10]:

$$\begin{aligned} \left|\partial_x^\gamma \partial_y^\eta H(t, x, y)\right| &\le C_1 t^{-\frac{|\gamma|+|\eta|}{\alpha}} \bar{p}_K^\alpha(t, x, y)\left(1 + \frac{1 \wedge |x-y|}{t}\right), & (3.46)\\ |\partial_x^\gamma H(t, x, x+\xi)| &\le C_1\bar{p}_K^\alpha(t, x, x+\xi)\left(1 + \frac{1 \wedge |\xi|}{t}\right). & (3.47)\end{aligned}$$

We then derive (3.43) for the derivative of the densities using the expansion:

$$p^d(1, x, y) = \tilde{p}_\alpha(t, x, y) + \frac{1}{n}\sum_{i=0}^{n-1}\int_{\mathbb{R}^d} \tilde{p}_\alpha(t_i, x, z)\, \Phi(1-t_i, z, y)\, dz. \tag{3.48}$$

To get the bound on p_n, one may proceed similarly. Denoting $\Phi_n(t_k, x, y) = \sum_{r=1}^\infty H_n^{(r,n)}(t_k, x, y)$, we investigate its derivatives and prove

$$\left|\partial_x^\gamma \partial_y^\eta \Phi_n(t_k, x, y)\right| \le Ct_k^{-\frac{|\gamma|+|\eta|}{\alpha}} \bar{p}_K^\alpha(t_k, x, y)\left(1 + \frac{1 \wedge |x-y|}{t_k}\right)$$

and

$$|\partial_x^\gamma \Phi_n(t_k, x, x+\xi)| \le C\bar{p}_K^\alpha(t_k, x, x+\xi)\left(1 + \frac{1 \wedge |\xi|}{t_k}\right),$$

by proving a similar estimate to (3.46) and (3.47) with H_n instead of H. The estimate on p_n will then be given by the counter part of representation (3.48) for p_n. We do not enter into the computational details.

We begin with (3.45). Observe that due to the presence of the derivation parameter in both the forward and the backward arguments, the derivatives does not yield any additional singularities. From (3.47), we prove by induction the following:

$$\left|\partial_x^\gamma H^{(r,n)}(t_k,x,x+\xi)\right| \le C_r t_k^{(r-1)\omega} \bar{p}_K^\alpha(t_k,x,x+\xi)\left(1+\frac{1\wedge|\xi|}{t_k}\right), \tag{3.49}$$

where $\omega = \frac{1}{\alpha}\wedge\alpha$, and the sequence of constants $(C_r)_{r\ge 0}$ is defined recursively by:

$$C_{r+1} = C_\gamma C_r C \max\left(\frac{1}{r\omega}, B\big((r-1)\omega+1,\omega\big)\right),\ C_1>0,$$

where C_1 is the constant appearing in bounds (3.47) and (3.49), and C is a positive constant independent of r,γ,x,ξ. For $r=1$, the bound is exactly (3.47). Suppose that it holds for $r\ge 1$. We have using the induction hypothesis, equation (3.47) and Leibnitz's formula:

$$\begin{aligned}
& \left|\partial_x^\gamma H^{(r+1,n)}(t_k,x,x+\xi)\right| \\
\le\ & \sum_{\eta=0}^{\gamma} C_\gamma^\eta \frac{1}{n}\sum_{i=0}^{k-1}\int_{\mathbb{R}^d}\left|\partial_x^\eta H^{(r,n)}(t_i,x,z+x)\right|\left|\partial_x^{\gamma-\eta}H(t_k-t_i,z+x,x+\xi)\right|dz \\
\le\ & C_\gamma C_r C\frac{1}{n}\sum_{i=0}^{k-1}\int_{\mathbb{R}^d} t_i^{(r-1)\omega}\bar{p}_K^\alpha(t_i,x,x+z)\left(1+\frac{1\wedge|z|}{t_i}\right) \\
& \quad \times\ \bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)\left(1+\frac{1\wedge|\xi-z|}{t_k-t_i}\right)dz.
\end{aligned} \tag{3.50}$$

We decompose, the integral:

$$\begin{aligned}
&\int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,x+z)\left(1+\frac{1\wedge|z|}{t_i}\right)\bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)\left(1+\frac{1\wedge|\xi-z|}{t_k-t_i}\right)dz \\
&= I_1+I_2+I_3+I_4
\end{aligned} \tag{3.51}$$

where:

$$\begin{aligned}
I_1 &= \int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,x+z)\bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)dz, \\
I_2 &= \int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,x+z)\bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)\frac{1\wedge|\xi-z|}{t_k-t_i}dz \\
I_3 &= \int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,x+z)\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)dz, \\
I_4 &= \int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,x+z)\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_k-t_i,x+z,x+\xi)\frac{1\wedge|\xi-z|}{t_k-t_i}dz.
\end{aligned}$$

The first one I_1 is bounded by $C\bar{p}_K^\alpha(t_k, x, x+\xi)$ thanks to the semi-group property (Proposition 3.6). By symmetry, I_2 and I_3 are treated the same way. We focus on I_2. In the rest, we denote by the symbol $\asymp$ the relation:

$$f(x) \asymp g(x) \Leftrightarrow \exists C > 1,\ \forall x \in \mathbb{R}^d:\ C^{-1}g(x) \leq f(x) \leq Cg(x).$$

We argue differently, according to the ratio $|\xi|/t_k^{1/\alpha}$.

- Suppose first that $|\xi| \leq Ct_k^{1/\alpha}$. Then, the diagonal estimate holds: $\bar{p}_K^\alpha(t_k, x, x+\xi) \asymp t_k^{-d/\alpha}$. On the one hand, if $i \geq k/2$, then $t_i \asymp t_k$. Since the diagonal estimate is a global bound, one has:

$$\bar{p}_K^\alpha(t_i, x, x+z) \leq Ct_i^{-d/\alpha} \asymp Ct_k^{-d/\alpha} \asymp C\bar{p}_K^\alpha(t_k, x, x+\xi).$$

  On the other hand, when $i \leq k/2$, then $t_k - t_i \asymp t_k$, and we have

$$\frac{1}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi) \leq C\frac{1}{t_k - t_i}(t_k - t_i)^{-d/\alpha} \asymp C\frac{1}{t_k}\bar{p}_K^\alpha(t_k, x, x+\xi).$$

- Suppose now that $|\xi| \geq Ct_k^{1/\alpha}$. Then, the off-diagonal estimate holds: $\bar{p}_K^\alpha(t_k, x, x+\xi) \asymp \frac{t_k}{|\xi|^{d+\alpha}}$. Now, since $|\xi| \leq |z| + |\xi - z|$, we have either $1/2|\xi| \leq |z|$, or $1/2|\xi| \leq |\xi - z|$. In the first case the off-diagonal estimate holds for the first density:

$$\bar{p}_K^\alpha(t_i, x, x+z) \asymp \frac{t_i}{|z|^{d+\alpha}} \leq C\frac{t_k}{|\xi|^{d+\alpha}} \asymp C\bar{p}_K^\alpha(t_k, x, x+\xi).$$

  In the second case, the second density is off-diagonal and we can write:

$$\begin{aligned}\frac{1}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi) &\leq C\frac{1}{t_k - t_i}\frac{t_k - t_i}{|\xi - z|^{d+\alpha}} \leq \frac{1}{t_k}\frac{t_k}{|\xi|^{d+\alpha}} \\ &\asymp C\frac{1}{t_k}\bar{p}_K^\alpha(t_k, x, x+\xi).\end{aligned}$$

Therefore, we always have the alternative:

$$\begin{aligned}\bar{p}_K^\alpha(t_i, x, x+z) &\leq C\bar{p}_K^\alpha(t_k, x, x+\xi), \text{ or} \\ \frac{1}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi) &\leq C\frac{1}{t_k}\bar{p}_K^\alpha(t_k, x, x+\xi).\end{aligned} \tag{3.52}$$

Combining this alternative with the smoothing effect of the Parametrix kernel H reflected in the bound (see Section 3 of Kolokoltsov [Kol00b]):

$$\forall \tau \in (0, T),\ \forall y \in \mathbb{R}^d,\ \int_{\mathbb{R}^d} \frac{1 \wedge |y - z|}{\tau}\bar{p}_K^\alpha(\tau, z, y)dz \leq \tau^{(\alpha\wedge 1)-1}, \tag{3.53}$$

gives that the second and third terms are bounded by:

$$I_2 + I_3 \leq C\bar{p}_K^\alpha(t_k, x, x+\xi)\left(t_i^{(\alpha\wedge 1)-1} + (t_k - t_i)^{(\alpha\wedge 1)-1} + \frac{1\wedge|\xi|}{t_k}\right).$$

We now turn to the last term in (3.51), that writes:

$$I_4 = \int_{\mathbb{R}^d} \frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_i, x, x+z)\frac{1\wedge|\xi - z|}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi)dz.$$

When $\bar{p}_K^\alpha(t_k, x, x+\xi)$ is in the diagonal regime, that is, when $|\xi| \leq Ct_k^{\frac{1}{\alpha}}$, we have:

$$\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_i, x, x+z) \leq Ct_i^{-d/\alpha}t_i^{\frac{1}{\alpha}-1}$$

and

$$\frac{1\wedge|\xi - z|}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi) \leq C(t_k - t_i)^{-d/\alpha}(t_k - t_i)^{\frac{1}{\alpha}-1}.$$

We prove the first inequality, the second one is obtained with the same arguments. Let us assume first that $|z| \leq Ct_i^{1/\alpha}$. In that case the diagonal estimate holds for $\bar{p}_K^\alpha(t_i, x, x+z)$, thus:

$$\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_i, x, x+z) \leq \frac{1\wedge|z|}{t_i}t_i^{-d/\alpha} \leq \frac{|z|}{t_i}t_i^{-d/\alpha} \leq Ct_i^{1/\alpha-1}\times t_i^{-d/\alpha}.$$

On the other hand, when the off-diagonal estimate holds for $\bar{p}_K^\alpha(t_i, x, x+z)$, that is when $|z| > Ct_i^{1/\alpha}$, we have:

$$\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_i, x, x+z) \leq \frac{1\wedge|z|}{t_i}\frac{t_i}{|z|^{d+\alpha}} \leq C\frac{1}{|z|^{d+\alpha-1}} \leq Ct_i^{-\frac{1}{\alpha}(d+\alpha-1)} = Ct_i^{-\frac{d}{\alpha}}t_i^{\frac{1}{\alpha}-1}.$$

Thus, in both cases, we obtained the announced bound.

Now, if $i \leq k/2$, $t_k \asymp t_k - t_i$, one has:

$$\frac{1\wedge|\xi - z|}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x+z, x+\xi) \leq C\bar{p}_K^\alpha(t_k, x, x+\xi)(t_k - t_i)^{\frac{1}{\alpha}-1}.$$

Then, using (3.53), $I_4 \leq \bar{p}_K^\alpha(t_k, x, x+\xi)(t_k - t_i)^{\frac{1}{\alpha}-1}t_i^{(\alpha\wedge 1)-1}$. Similarly, when $i > k/2$, we use that $t_k \asymp t_i$, to get

$$\frac{1\wedge|z|}{t_i}\bar{p}_K^\alpha(t_i, x, x+z) \leq C\bar{p}_K^\alpha(t_k, x, x+\xi)t_i^{\frac{1}{\alpha}-1}.$$

Consequently, when $|\xi| \leq Ct_k^{\frac{1}{\alpha}}$, I_4 is bounded in the following way:

$$I_4 \leq C\bar{p}_K^\alpha(t_k, x, x+\xi)\Big((t_k - t_i)^{\frac{1}{\alpha}-1}t_i^{(\alpha\wedge 1)-1} + t_i^{\frac{1}{\alpha}-1}(t_k - t_i)^{(\alpha\wedge 1)-1}\Big).$$

Assume now that $|\xi| > Ct_k^{\frac{1}{\alpha}}$. In that case using similar arguments one may prove that we have either:

$$\frac{1 \wedge |\xi - z|}{t_k - t_i}\bar{p}_K^\alpha(t_k - t_i, x + z, x + \xi) \le C\frac{1 \wedge |\xi|}{t_k}\bar{p}_K^\alpha(t_k, x, x + \xi),$$

or:

$$\frac{1 \wedge |z|}{t_i}\bar{p}_K^\alpha(t_i, x, x + z) \le C\frac{1 \wedge |\xi|}{t_k}\bar{p}_K^\alpha(t_k, x, x + \xi).$$

Thus, using (3.53), I_4 is now bounded as follows:

$$I_4 \le C\frac{1 \wedge |\xi|}{t_k}\bar{p}_K^\alpha(t_k, x, x + \xi)\Big(t_i^{(\alpha\wedge 1)-1} + (t_k - t_i)^{(\alpha\wedge 1)-1}\Big).$$

Plugging this estimate in (3.51) in turn implies:

$$\begin{aligned}\int_{\mathbb{R}^d} \bar{p}_K^\alpha(t_i, x, x + z)\left(1 + \frac{1 \wedge |z|}{t_i}\right)\bar{p}_K^\alpha(t_k - t_i, x + z, x + \xi)\left(1 + \frac{1 \wedge |y - z|}{t_k - t_i}\right)dz \\ \le C\bar{p}_K^\alpha(t_k, x, x + \xi)\Bigg(\Big((t_k - t_i)^{\frac{1}{\alpha}-1}t_i^{(\alpha\wedge 1)-1} + t_i^{\frac{1}{\alpha}-1}(t_k - t_i)^{(\alpha\wedge 1)-1}\Big) \\ + \frac{1 \wedge |\xi|}{t_k}\Big(t_i^{(\alpha\wedge 1)-1} + (t_k - t_i)^{(\alpha\wedge 1)-1}\Big)\Bigg). \end{aligned} \tag{3.54}$$

Plugging bound (3.54) in (3.50) yields:

$$\begin{aligned} &\left|\partial_x^\gamma H^{(r+1,n)}(t_k, x, x + \xi)\right| \\ \le\ & C_\gamma C_r C\bar{p}_K^\alpha(t_k, x, x + \xi)\frac{1}{n}\sum_{i=0}^{k-1} t_i^{(r-1)\omega}\Big((t_k - t_i)^{\frac{1}{\alpha}-1}t_i^{(\alpha\wedge 1)-1} + t_i^{\frac{1}{\alpha}-1}(t_k - t_i)^{(\alpha\wedge 1)-1}\Big) \\ & + C_\gamma C_r C\frac{1 \wedge |\xi|}{t_k}\bar{p}_K^\alpha(t_k, x, x + \xi)\frac{1}{n}\sum_{i=0}^{k-1} t_i^{(r-1)\omega}\Big(t_i^{(\alpha\wedge 1)-1} + (t_k - t_i)^{(\alpha\wedge 1)-1}\Big). \end{aligned} \tag{3.55}$$

Now, assume first that $\alpha \ge 1$. Recalling that $\omega = \frac{1}{\alpha} \wedge \alpha = \frac{1}{\alpha}$, the above bound becomes:

$$\begin{aligned}\left|\partial_x^\gamma H^{(r+1,n)}(t_k, x, x + \xi)\right| &\le C_\gamma C_r C\bar{p}_K^\alpha(t_k, x, x + \xi)\frac{1}{n}\sum_{i=0}^{k-1} t_i^{(r-1)\frac{1}{\alpha}}\Big((t_k - t_i)^{\frac{1}{\alpha}-1} + t_i^{\frac{1}{\alpha}-1}\Big) \\ &\quad + C_\gamma C_r C\frac{1 \wedge |\xi|}{t_k}\bar{p}_K^\alpha(t_k, x, x + \xi)\frac{1}{n}\sum_{i=0}^{k-1} t_i^{(r-1)\frac{1}{\alpha}} \\ &\le C_{r+1}t_k^{\frac{r}{\alpha}}\bar{p}_K^\alpha(t_k, x, x + \xi)\left(1 + \frac{1 \wedge |\xi|}{t_k}\right),\end{aligned}$$

where $C_{r+1} = C_\gamma C C_r \max\left(B\left((r-1)\frac{1}{\alpha}+1, \frac{1}{\alpha}\right), \frac{\alpha}{r}\right)$ and we used that $t_k^{(r-1)\frac{1}{\alpha}+1} \le t_k^{\frac{r}{\alpha}}$, since $\alpha \ge 1$ and $t_k \le 1$.

On the other hand, when $\alpha \le 1$, $\omega = \alpha$ one similarly proves that:

$$\left|\partial_x^\gamma H^{(r+1,n)}(t_k, x, x+\xi)\right| \le C_{r+1} t_k^{r\alpha} \bar{p}_K^\alpha(t_k, x, x+\xi)\left(1 + \frac{1 \wedge |\xi|}{t_k}\right),$$

with $C_{r+1} = C_\gamma C_r C \max\left(\frac{1}{r\alpha}, B\big((r-1)\alpha+1, \alpha\big)\right)$. This constant is coherent with the previous one, setting $C_{r+1} = C_\gamma C_r C \max\left(\frac{1}{r\omega}, B\big((r-1)\omega+1, \omega\big)\right)$. This concludes the proof of bound (3.49). Observe that by definition of Euler's Beta function, $(C_r)_{r\ge 0}$ produces a convergent series. To get the bound (3.45), we sum bounds (3.49). In order to get the bound (3.43), we now plug the bound (3.45) in equation (3.48), and from similar arguments, one derives (3.43).

To prove (3.44), we show by induction the following bound:

$$\left|\partial_x^\gamma \partial_y^\eta H^{(r,n)}(t_k, x, y)\right| \le C_r t_k^{(r-1)\omega - \frac{|\gamma|+|\eta|}{\alpha}} C_1 \bar{p}_K^\alpha(t_k, x, y)\left(1 + \frac{1 \wedge |x-y|}{t_k}\right). \qquad (3.56)$$

For $r = 1$, this bound is exactly (3.46). To get the estimate for $r+1$, we proceed as above.

$$\left|\partial_x^\gamma \partial_y^\eta H^{(r+1,n)}(t_k, x, y)\right| = \left|\partial_x^\gamma \partial_y^\eta \frac{1}{n}\sum_{i=0}^{k-1} \int_{\mathbb{R}^d} H^{(r,n)}(t_i, x, z) H(t_k - t_i, z, y) dz\right| \le I + II,$$

where

$$\begin{aligned} I &= \left|\partial_x^\gamma \partial_y^\eta \frac{1}{n}\sum_{i \le k/2} \int_{\mathbb{R}^d} H^{(r,n)}(t_i, x, z) H(t_k - t_i, z, y) dz\right|, \\ II &= \left|\partial_x^\gamma \partial_y^\eta \frac{1}{n}\sum_{i \ge k/2} \int_{\mathbb{R}^d} H^{(r,n)}(t_i, x, z) H(t_k - t_i, z, y) dz\right|. \end{aligned}$$

In I, the time parameter t_i is small, thus, the singularities induced by the derivation of $H^{(r,n)}(t_i, x, z)$ are the worst. In order to get rid of them, we make use of a change of variable to get:

$$\begin{aligned} I &= \left|\partial_x^\gamma \frac{1}{n}\sum_{i \le k/2} \int_{\mathbb{R}^d} H^{(r,n)}(t_i, x, z) \partial_y^\eta H(t_k - t_i, z, y) dz\right| \\ &= \left|\partial_x^\gamma \frac{1}{n}\sum_{i \le k/2} \int_{\mathbb{R}^d} H^{(r,n)}(t_i, x, z+x) \partial_y^\eta H(t_k - t_i, z+x, y) dz\right|. \end{aligned}$$

Now, from equations (3.49), (3.46) and Leibnitz's formula we derive:

$$\begin{aligned}
I &\leq \sum_{\beta=0}^{\gamma} C_\gamma^\beta \frac{1}{n} \sum_{i\leq k/2} \int_{\mathbb{R}^d} \left|\partial_x^\beta H^{(r,n)}(t_i, x, z+x)\right| \left|\partial_x^{\gamma-\beta}\partial_y^\eta H(t_k - t_i, z+x, y)\right| dz \\
&\leq C_\gamma C_n \frac{1}{n} \sum_{i\leq k/2} \int_{\mathbb{R}^d} t_i^{(r-1)\omega} \left(1 + \frac{1\wedge |z|}{t_i}\right) \bar{p}_K^\alpha(t_i, x, x+z) \\
&\quad \times (t_k - t_i)^{-\frac{|\gamma|+|\eta|}{\alpha}} \left(1 + \frac{1\wedge|y-x-z|}{t_k - t_i}\right) \bar{p}_K^\alpha(t_k - t_i, x+z, y) dz. \\
&\leq C_{n+1} t_k^{r\omega - \frac{|\gamma|+|\eta|}{\alpha}} \bar{p}_K^\alpha(t_k, x, y).
\end{aligned}$$

where we used that $t_k \asymp t_k - t_i$ for $i \leq k/2$ for the last inequality. Note that once again, the series $\sum_{r\geq 1} C_r$ converges. For II, we proceed with similar arguments. In this case, we use the change variables $w = z + y$ instead.

□

Proof of Theorem 3.5. The coefficients are the sum of two terms. We only focus on the first term, the second term can be treated similarly. From the definition of $\otimes_n$, we have:

$$\partial_y^\gamma p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d(1, x, y) = \partial_y^\gamma \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathbb{R}^d} p(t_i, x, z) \big(L - \tilde{L}^*\big)^{k+1} p^d(1 - t_i, z, y) dz.$$

To deal with the singularities coming from the derivatives we split the sum over i in two parts:

$$\begin{aligned}
\partial_y^\gamma p \otimes_n \left(L - \tilde{L}^*\right)^{k+1} p^d(1, x, y) &= \partial_y^\gamma \frac{1}{n} \sum_{i<n/2} \int_{\mathbb{R}^d} p(t_i, x, z) \big(L - \tilde{L}^*\big)^{k+1} p^d(1 - t_i, z, y) dz \\
&\quad + \partial_y^\gamma \frac{1}{n} \sum_{i\geq n/2} \int_{\mathbb{R}^d} p(t_i, x, z) \big(L - \tilde{L}^*\big)^{k+1} p^d(1 - t_i, z, y) dz \\
&= S_1 + S_2.
\end{aligned}$$

For S_1, the time parameter is small, thus the singularities brought by the derivation in y and the generators are negligible. Indeed, exchanging the derivation and the integral:

$$S_1 = \frac{1}{n} \sum_{i<n/2} \int_{\mathbb{R}^d} p(t_i, x, z) \partial_y^\gamma \left(\big(L - \tilde{L}^*\big)^{k+1} p^d(1 - t_i, z, y) \right) dz.$$

From bound (3.41) in Lemma 3.7, we derive:

$$\left|\partial_y^\gamma\left(\left(L-\tilde{L}^*\right)^{k+1}p^d(1-t_i,z,y)\right)\right| \;\leq\; C\,(1-t_i)^{-k-1-\frac{\gamma}{\alpha}}\,\bar{p}_K^\alpha\,(1-t_i,z,y)\,.$$

The right hand side of the previous equation is bounded uniformly in y, thus, from the Lebesgue theorem, we can derive under the integral. Now, since $p(t_i,x,z)\leq C\bar{p}_K^\alpha(t_i,x,z)$, this sum yields by a semi-group property:

$$\begin{aligned}
&\left|\partial_y^\gamma\frac{1}{n}\sum_{i<n/2}\int_{\mathbb{R}^d}p(t_i,x,z)\big(L-\tilde{L}^*\big)^{k+1}p^d(1-t_i,z,y)dz\right| \\
\leq\;& C\frac{1}{n}\sum_{i<n/2}\int_{\mathbb{R}^d}\bar{p}_K^\alpha(t_i,x,z)\,(1-t_i)^{-\frac{\gamma}{\alpha}-k-1}\,\bar{p}_K^\alpha\,(1-t_i,z,y)\,. \\
=\;& C\bar{p}_K^\alpha(1,x,y).
\end{aligned}$$

We now turn to the second sum. When $i\geq n/2$, by an integration by parts it follows

$$\begin{aligned}
S_2 &= \partial_y^\gamma\frac{1}{n}\sum_{i\geq n/2}\int_{\mathbb{R}^d}p(t_i,x,z)\big(L-\tilde{L}^*\big)^{k+1}p^d(1-t_i,z,y)dz \\
&= \partial_y^\gamma\frac{1}{n}\sum_{i\geq n/2}\int_{\mathbb{R}^d}\left(\big(L-\tilde{L}^*\big)^{k+1}\right)^T p(t_i,x,z)p^d(1-t_i,z,y)dz.
\end{aligned}$$

where $\left(\big(L-\tilde{L}^*\big)^{k+1}\right)^T$ stands for the adjoint of $\big(L-\tilde{L}^*\big)^{k+1}$, which is well defined thanks to the smoothness of the coefficients b and σ. The operator $L-\tilde{L}^*$ is an integro-differential operator (a derivative of order α), so that the operator $\left(\big(L-\tilde{L}^*\big)^{k+1}\right)^T$ is still an integro-differential operator which yields singularity of the same order as $(L-\tilde{L}^*)^{k+1}$, thus, applied to p yields singularities which are still negligible since $i\geq n/2$. However, the derivative ∂_y^γ will affect $p^d(1-t_i,z,y)$, thus, giving additional singularities so that beforehand we make use of the change of variable: $z=y-u$ to derive

$$S_2=\frac{1}{n}\sum_{i\geq n/2}\sum_{\eta=0}^{\gamma}C_\gamma^\eta\int_{\mathbb{R}^d}\partial_y^{\gamma-\eta}\left[\left(\big(L-\tilde{L}^*\big)^{k+1}\right)^T p\,(t_i,x,y-u)\right]\partial_y^\eta p^d\,(1-t_i,y-u,y)\,du.$$

Now, bound (3.43) of Lemma 3.7 gives: $\forall\eta\in\mathbb{N}^*$, $\exists C>0$ s.t.:

$$\left|\partial_y^\eta p^d\,(1-t_i,y-u,y)\right|\leq C\bar{p}_K^\alpha\,(1-t_i,y-u,y)\,.$$

On the other hand, since $i\geq n/2$, the singularities of the derivatives on the first density are negligible. We thus get from a semi group property:

$$\begin{aligned}|S_2| &\le \frac{C}{n}\sum_{i\ge n/2}\sum_{\eta=0}^{\gamma} C_\gamma^\eta 2^{\frac{|\gamma|+|\eta|}{\alpha}+k+1}\int_{\mathbb{R}^d}\bar p_K^\alpha\left(t_i,x,y-u\right)\bar p_K^\alpha\left(1-t_i,y-u,y\right)du\\ &\le C_\gamma \bar p_K^\alpha\left(1,x,y\right).\end{aligned}$$

In order to prove that the expansion (3.37) makes sense, it remains to prove the bound on the remainder. Since the expansion (3.37) is made of two contributions $p-p^d$ and p^d-p_n, the remainder $R(x,y)$ also splits in two terms

$$R(x,y)=R_1(1,x,y)+R_2(1,x,y),$$

each being a remainder respectively of the expansion of $p-p^d$ and p^d-p_n, with:

$$\begin{aligned}R_1(1,x,y) &= \sum_{r\ge 0}(Q_M\otimes_n H^{(r,n)})(1,x,y),\\ Q_M(t_i,x,y) &= \frac{1}{M!}\sum_{i=0}^{k-1}\int_{i/n}^{(i+1)/n}\left[n\left(u-i/n\right)\right]\\ &\quad\times\int_0^1(1-\delta)^{M-1}\int\frac{\partial^M}{\partial s^M}\left[p(s,x,z)H(t_i-s,z,y)\right]_{s=t_i+\delta(u-t_i)}dzd\delta du,\end{aligned}$$

and

$$\begin{aligned}R_2(t,x,y) &= \frac{1}{(M+1)!}\int_0^1(1-\tau)^M\left[p^d\otimes_n\left(\tilde L_*-\tilde L^*\right)^{M+1}\tilde p_\tau^\Delta\right](t,x,y)\\ \tilde p_\tau^\Delta(t,x,y) &= \sum_{r\ge 0}\tilde p_\tau\otimes_n H_n^{(r,n)}(t,x,y);\ H_n(t,x,y)=(L_n-\tilde L_n^*)\tilde p_\alpha(t,x,y),\end{aligned}$$

where $\tilde p_\tau(t,x,y)=\int_{\mathbb{R}^d}\tilde p_\alpha^x(\tau\Delta,x,z)\tilde p_\alpha^y(t-\tau\Delta,z,y)dz$ and L_n and $\tilde L_n^*$ stands for the discrete generators defined in equations (3.35) and (3.36). The reader may consult Konakov and Mammen [KM02] and Konakov and Menozzi [KM10] for more details.

We focus on the estimation of R_1. The arguments for R_2 are similar and hinted at the end of this proof. Observe that we can write:

$$R_1(1,x,y) = Q_M(1,x,y)+\frac{1}{n}\sum_{i=0}^{n-1}\int Q_M\left(t_i,x,z\right)\Phi\left(1-t_i,z,y\right)dz, \quad (3.57)$$

$$\Phi\left(1-t_i,z,y\right) = \sum_{r=1}^{+\infty}H^{(r,n)}\left(1-t_i,z,y\right). \quad (3.58)$$

Using Kolmogorov's forward and backward equations, one shows by induction on M that $Q_M(t,x,y)$ can be written as:

$$Q_M(t_k,x,y) = \frac{1}{M!}\sum_{i=0}^{k-1}\int_{t_i}^{(i+1)/n}[n\,(u-t_i)]\int_0^1(1-\delta)^{M-1}$$
$$\times\int p\left(t_i+\delta\,(u-t_i)\,,x,z\right)(L-\tilde{L}^*)^{M+1}\tilde{p}_\alpha\left(t_k-(t_i+\delta\,(u-t_i)),z,y\right)dzd\delta du.$$

See e.g. equation (4.5) in [KM02] for the brownian case and equation (4.4) in [KM10] for the stable case. Thus, apart from the additional integrations w.r.t. δ and u, we see that Q_M is of the same nature that the terms we already dealt with. Therefore, adapting the above arguments allows us to derive α-stable estimates for this term. In fact, a precise study of Q_M allows us to derive that $\forall i\in[\![1,n]\!]$, $\forall\gamma\geq 0$, $\exists C>0$ such that:

$$\left|\partial_y^\gamma Q_M\left(t_k,x,y\right)\right| \leq Ct_k^{-\frac{\gamma}{\alpha}-(M+1)}\bar{p}_K^\alpha\left(t_k,x,y\right). \tag{3.59}$$

To get this bound, we actually show that it holds for

$$\bar{Q}_M\left(t_k,x,y\right) = \int p\left(t,x,z\right)(L-\tilde{L}^*)^{M+1}\tilde{p}_\alpha\left(t_k-t,z,y\right)dz,$$

independently of $t\in[0,1]$. The arguments differs depending if t is closer to 0 or t_k. In the first case (say, $t\leq t_k/2$), the singularities induced by taking the derivative along y in $\tilde{p}\left(t_k-t,z,y\right)$ are bounded by to $t_k^{-\frac{\gamma}{\alpha}-(M+1)}$, which is the announced singularity. When t is closer to t_k (say $t>t_k/2$), we transfer the operator $(L-\tilde{L}^*)^{M+1}$ on p by taking the adjoint, and we change variables to $z=y-u$. The singularities in $\partial_y^\gamma\Big((L-\tilde{L}^*)^{M+1}\Big)^T p\left(t,x,y-u\right)$ then yields the announced $t_k^{-\frac{\gamma}{\alpha}-(M+1)}$, and we conclude using the semi-group property of Proposition 3.6. Thus, for $t_k=1$, we have $|\partial_y^\gamma Q_M(1,x,y)|\leq C\bar{p}_K^\alpha(1,x,y)$, which is the announced bound.

Now, for the second part of (3.57), we split the sum:

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}\int dzQ_M\left(t_i,x,z\right)\Phi\left(1-t_i,z,y\right) &= \frac{1}{n}\sum_{i\leq\frac{n}{2}}\int dzQ_M\left(t_i,x,z\right)\Phi\left(1-t_i,z,y\right)\\ &\quad+\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int dzQ_M\left(t_i,x,z\right)\Phi\left(1-t_i,z,y\right)\\ &= S_1(1,x,y)+S_2(1,x,y).\end{aligned}$$

In $S_1(1,x,y)$, the time parameter is such that when differentiating along y, the singularities are negligible. Thus, we derive under the integral, and use bounds (3.59) and (3.44), to get a convolution of $\bar{p}_K^\alpha$ functions. For $S_2(1,x,y)$, we make use of the change of variable $z=y-u$, to get:

$$S_2(1,x,y)=\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int duQ_M\left(t_i,x,y-u\right)\Phi\left(1-t_i,y-u,y\right).$$

In the stable case, when taking the derivative along y the bounds (3.59) and (3.45) yield:

$$\begin{aligned}\left|\partial_y^\gamma S_2\left(1,x,y\right)\right| &= \sum_{\eta=0}^{\gamma}C_\eta^\gamma\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int du\left|\partial_y^{\gamma-\eta}Q_M\left(t_i,x,y-u\right)\right|\left|\partial_y^\eta\Phi\left(1-t_i,y-u,y\right)\right|\\ &\leq C_\gamma\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int du\bar{p}_K^\alpha(t_i,x,y-u)\left(1+\frac{1\wedge|u|}{1-t_i}\right)\bar{p}_K^\alpha(1-t_i,y-u,y)\\ &\leq C\bar{p}_K^\alpha(1,x,y)\\ &\quad+\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int du\bar{p}_K^\alpha(t_i,x,y-u)\left(\frac{1\wedge|u|}{1-t_i}\right)\bar{p}_K^\alpha(1-t_i,y-u,y).\end{aligned}$$

Now, recall from definition of $\bar{p}_K^\alpha$, we have either $\bar{p}_K^\alpha(t_i,x,z)\leq C\bar{p}_K^\alpha(1,x,y)$ or $\frac{1}{1-t_i}\bar{p}_K^\alpha(1-t_i,z,y)\leq C\bar{p}_K^\alpha(1,x,y)$, so that:

$$\frac{1}{n}\sum_{i\geq\frac{n}{2}}\int du\bar{p}_K^\alpha(t_i,x,y-u)\left(\frac{1\wedge|u|}{1-t_i}\right)\bar{p}_K^\alpha(1-t_i,y-u,y)\leq\bar{p}_K^\alpha(1,x,y)\left(1+1\wedge|y-x|\right).$$

In the gaussian case, the proof is simpler, as the derivative of Φ is estimated by:

$$\left|\partial_y^\eta\Phi\left(1-t_i,y-u,y\right)\right|\leq\frac{1}{\sqrt{1-t_i}}\bar{p}_K^\alpha(1-t_i,y-u,y),$$

and we can directly conclude comparing the sum over η to a Beta function.

For R_2, we can take the derivative in y under the integral in the above expression to get:

$$\partial_y^\gamma R_2(t,x,y) = \frac{1}{(M+1)!}\int_0^1(1-\tau)^M\partial_y^\gamma\left[p^d\otimes_n\left(\tilde{L}_*-\tilde{L}^*\right)^{M+1}\tilde{p}_\tau^\Delta\right](t,x,y),$$

which is a term of the same nature as the second part of the expansion (3.37). In particular, one can show bounds on $\tilde{p}_\tau^\Delta(t,x,y)$, similar to those of Lemma 3.7. With these estimates at hand, we may use similar arguments to derive an α-stable estimate on R_2, for $\alpha\in(0,2]$. We leave the remaining details to the reader.

□

Remark 3.3. The terms in the expansion (3.37) depends on n. As already pointed out in [KM10] and [KM02], it is possible to make this expansion independent of n, using the bounds on the difference between the usual time space convolution $\otimes$ and its discretization $\otimes_n$. For $M = 2$, one derives the expansion:

$$\begin{aligned}
\partial_y^\gamma (p - p_n)(1,x,y) &= \frac{1}{2n}\partial_y^\gamma \Big(p \otimes_n (L - \tilde{L}^*)^2 p^d\Big)(1,x,y) \\
&\quad - \frac{1}{2n}\partial_y^\gamma \Big(p^d \otimes_n (\tilde{L}_* - \tilde{L}^*)^2 p_n\Big)(1,x,y) + \frac{1}{n^2}\partial_y^\gamma R(x,y) \\
&= \frac{1}{2n}\partial_y^\gamma \Big(p \otimes (L - \tilde{L}^*)^2 p\Big)(1,x,y) \\
&\quad - \frac{1}{2n}\partial_y^\gamma \Big(p \otimes (\tilde{L}_* - \tilde{L}^*)^2 p\Big)(1,x,y) + \frac{1}{n^2}\partial_y^\gamma \tilde{R}(x,y) \\
&= \frac{1}{2n}\partial_y^\gamma \Big(p \otimes (L^2 - (\tilde{L}^*)^2) p\Big)(1,x,y) + \frac{1}{n^2}\partial_y^\gamma \tilde{R}(x,y).
\end{aligned}$$

In the above expansion, $\partial_y^\gamma \tilde{R}(x,y)$ is a remainder term bounded by some stable density as $\partial_y^\gamma R(x,y)$.

Corollary 3.8. *Assume* ***[A]*** *holds. Recall m denotes the regularity of the coefficients of the SDE* (3.31)*. Let $M \in \mathbb{N}^*$, such that when $\alpha = 2$, $0 < M \le m/2$, and when $\alpha < 2$, we assume $m > d+4$ and $0 < M \le m-(d+4)$. Then, the following expansion holds:*

$$\theta^{*,n} - \theta^* = \frac{C_1}{n} + \cdots + \frac{C_p}{n^{M-1}} + o\left(\frac{1}{n^{M-1}}\right).$$

Proof. We prove that under the assumptions of Corollary 3.8, **[H-(M-1)]** holds. Let $M \le m/2$ for $\alpha = 2$ and $M \le m - (d+4)$ for $\alpha < 2$ and let $\gamma \in \mathbb{N}$ with $\gamma \le M - 1$. From Theorem 3.5, under **[A]**, expansion (3.37) holds up to order $M - 1$. Moreover, from Remark 3.3, this expansion can be made independent of n, namely

$$\partial_{y^d}^\gamma p(1,x,y) - \partial_{y^d}^\gamma p_n(1,x,y) = \sum_{k=1}^{M-\gamma-1} \frac{1}{n^k}\Gamma_k^\gamma(x,y) + \frac{\partial_y^\gamma \tilde{R}_n(x,y)}{n^{M-\gamma}}. \tag{3.60}$$

Thus, integrating equation (3.60) in the $d-1$ first variables yields for all $M \le m/2$ for $\alpha = 2$ and $M \le m - (d+4)$ for $\alpha < 2$ and $\gamma \le M - 1$:

$$\partial_{y^d}^\gamma p^{X_1^d}(1,x,y^d) - \partial_{y^d}^\gamma p_n^{X_1^{n,d}}(1,x,y^d) = \sum_{k=1}^{M-\gamma-1} \frac{1}{n^k}\bar{\Gamma}_k^\gamma(x,y^d) + \frac{\partial_{y^d}^\gamma \bar{R}_n^\gamma(x,y^d)}{n^{M-\gamma}}. \tag{3.61}$$

where we denoted $p^{X_1^d}$, $p_n^{X_1^{n,d}}$ the marginal densities of X_1^d and $X_1^{n,d}$, and

$$\bar{\Gamma}_k^\gamma(x,y^d) = \int_{\mathbb{R}^d\times\cdots\times\mathbb{R}^d} \Gamma_k^\gamma(x,y)dy_1\cdots dy_{d-1}.$$

The Gaussian bound on the remainder implies $\frac{\partial^l_{y_d}\bar{R}_n(x,y^d)}{n^{M-\gamma}} = O\left(n^{-(M-\gamma)}\right)$, so that we have for all $M \le m/2$ for $\alpha = 2$ and $M \le m-(d+4)$ for $\alpha < 2$ and $\gamma \le M-1$:

$$\partial^\gamma_{y^d} p^{X_1^d}(1,x,y^d) - \partial^\gamma_{y^d} p_n^{X_1^{n,d}}(1,x,y^d) = \sum_{k=1}^{M-\gamma-1} \frac{1}{n^k}\bar{\Gamma}^\gamma_k(x,y^d) + o(\frac{1}{n^{M-\gamma-1}}). \qquad (3.62)$$

Now, since $h(\theta) - h^n(\theta) = \mathbb{P}^x(X_1^d \le \theta) - \mathbb{P}^x(X_1^{n,d} \le \theta)$, taking $\gamma = 0$ in (3.62) yields:

$$h(\theta) - h^n(\theta) = \int_{-\infty}^{\theta} \left(p^{X_1^d}(1,x,y_d) - p_n^{X_1^{n,d}}(1,x,y_d)\right) dy_d = \sum_{k=1}^{M-1} \frac{\Lambda^0_k(\theta)}{n^k} + o(\frac{1}{n^{M-1}}),$$

where we denoted $\Lambda^0_k(\theta) = \int_{-\infty}^{\theta} \bar{\Gamma}^0_k(x,y_d)dy_d$. Thus, the first assumption in **[H-(M-1)]** holds.

We now turn to the expansion of the derivatives. From expansion (3.62) one easily gets:

$$\partial_\theta(h-h^n)(\theta) = (p^{X_1^d} - p_n^{X_1^{n,d}})(1,x,\theta) = \sum_{k=1}^{M-2} \frac{1}{n^k}\bar{\Gamma}^0_k(x,\theta) + o(\frac{1}{n^{M-2}})$$

and $\forall l \le M-2, \forall (x,\theta) \in \mathbb{R}^d \times \mathbb{R}$

$$\begin{aligned} \partial^l_\theta h(\theta) - \partial^l_\theta h^n(\theta) &= \partial^{l-1}_\theta p^{X_1^d}(1,x,\theta) - \partial^{l-1}_\theta p_n^{X_1^{n,d}}(1,x,\theta) \\ &= \sum_{k=1}^{M-l-1} \frac{1}{n^k}\Lambda^l_k(\theta) + o(\frac{1}{n^{M-1-l}}) \end{aligned}$$

where we denoted for consistency $\Lambda^l_k(\theta) = \bar{\Gamma}^{l-1}_k(x,\theta)$. Consequently, expansions (2.5) and (2.6) holds in **[H-(M-1)]**. It remains to check the local uniform convergence and the invertibility of $Dh(\theta^*)$. For the latter, recall that $Dh(\theta^*) = p^d(1,x,\theta^*)$. Also, we know that under **[A]**, stable bounds holds for $p(1,x,y)$ (see e.g. [Aro67] in the gaussian case, and [Kol00b] for the stable case), that is $p(1,x,y) \asymp \bar{p}^\alpha_K(1,x,y)$. Thus, the left hand side of the previous inequality gives that $p(1,x,y)$, and *a fortiori* $p^{X^d}(1,x,y)$, is never equal to zero. Hence $Dh(\theta^*)$ is invertible. Finally the local uniform convergence is a consequence of expansion (3.37). □

4 Numerical illustration

To illustrate the method we consider a geometric Brownian motion $(X_t)_{t\in[0,T]}$ with dynamics given by

$$X_t = x_0 \exp((r-\sigma^2/2)T + \sigma\sqrt{T}W_t),\ t \in [0,T]$$

for which the quantile is explicitly known at any level $\ell \in (0,1)$. Indeed a simple computation shows that

$$\theta^* = x_0 \exp((r - \sigma^2/2)T + \sigma\sqrt{T}\phi^{-1}(\ell))$$

where ϕ is the distribution function of the standard normal distribution $\mathcal{N}(0,1)$. Let us note that the assumptions of Section 3 are not satisfied in this example and that nobody would devise any kind of Monte Carlo simulation in practice since the law of X_T is explicitly known for any time T. However the Black-Scholes model and its Euler scheme appears as a natural and often used benchmark to test and evaluate the performance of Monte Carlo methods. We use the following values for the parameters: $x_0 = 100$, $r = 0.05$, $\sigma = 0.4$, $T = 1$, $\ell = 0.7$. The reference Black-Scholes quantile is $\theta^* = 119.69$. We set $\gamma(p) = \gamma_0/p$ with $\gamma_0 = 60$.

Let us note that in order to implement the Richardson-Romberg stochastic approximation estimator we need to simulate discretization schemes of the Brownian diffusion with different steps $\Delta_r = T/nr$, $r = 1, \cdots, R$. We thus need to simulate *consistent Brownian increments* on intervals of the form $[(k-1)T/(rn), kT/(rn)]$, $r = 1, \cdots, R$. The coefficients to compute by induction the Brownian increments from small intervals up to the root interval of length T/n have been computed up to $R = 5$ for $\alpha = 1$ and up to $R = 3$ for $\alpha = 1/2$ in [Pag07], Section 5.

In order to illustrate the result of Theroem 2.3, we plot in Figure 4.1 the behaviors of $\sum_{r=1}^{R} \mathbf{w}_r \theta^{*,rn} - \theta^*$ for $R = 2, 3, 4$ and $n = 2, \cdots, 15$. We estimate $\theta^{*,rn}$ by θ_M^{rn}, with $M = 10^6$ samples for $R = 2$ and $M = 10^8$ samples for $R = 3, 4$ using the same Brownian motion for each R (see Remark 2.3). We clearly see that the Richardson-Romberg estimator efficiency increases with R and the method gives satisfying results with $R = 3, 4$ for small values of n.

Let us observe that the asymptotic optimal parameters in Propositions 2.5 and 2.6 depend on the structural parameters: α, $|C_R|$, $C(\gamma, \underline{\lambda})$, $\mathbb{E}\left[|H(\theta^*, U)|^2\right]$, $|C_1|$. Let us note that in Proposition 2.2 and Theorem 2.3 one may show that the constants $|C_R|$ writes $C_R = Dh(\theta^*)^{-1}\tilde{C}_R$. Here one has $Dh(\theta^*) = p(1, x, \theta^*)/(1-\ell)$ so that $|C_R| = |\tilde{C}_R|(1-\ell)/p(1, x, \theta^*)$. We estimate $p(1, x, \theta^*)$ by $p_n(1, x, \theta_M^n) \approx (\mathbb{P}_x(X_1^{n,d} \leq \theta_M^n + \varepsilon) - \mathbb{P}_x(X_1^{n,d} \leq \theta_M^n - \varepsilon))/2\varepsilon$ which in turn is approximated by the crude Monte Carlo estimator $(2M)^{-1}\sum_{k=1}^{M} \mathbf{1}_{\left\{(X_1^{n,d})^k \leq \theta^* + \varepsilon\right\}} - \mathbf{1}_{\left\{(X_1^{n,d})^k \leq \theta^* - \varepsilon\right\}}$ with $M = 1000$, $n = 100$, $\varepsilon = 0.1$ leading to the value $Dh(\theta^*) = 2.56 \times 10^{-2}$. Finally estimating $|\tilde{C}_1|$ for the crude SA estimator and $|\tilde{C}_R|$ for the Richardson-Romberg extrapolation method is a challenging task. Consequently we implement these methods in a blind way setting $|\tilde{C}_R| = 1$ for every R. We also set $\underline{\lambda} = Dh(\theta^*)$ and $\gamma_0 = 1/\underline{\lambda}$. Note that we have $\mathbb{E}\left[|H(\theta^*, U)|^2\right] = \mathbb{E}\left[|1 - (1-\ell)^{-1}\mathbf{1}_{\left\{X_1^d \geq \theta^*\right\}}|^2\right] = \ell/(1-\ell)$. The optimal parameters for

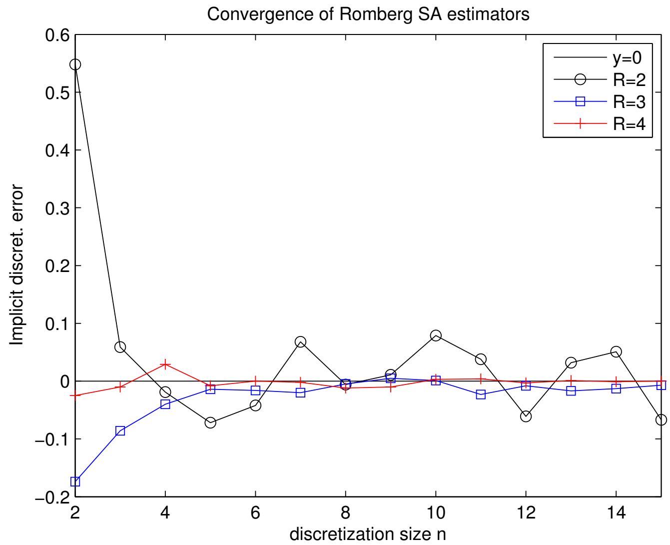

Figure 4.1: Richardson Romberg SA estimators: $\sum_{r=1}^{R} \mathbf{w}_r \theta^{*,rn} - \theta^*$ with respect to $n = 2, \cdots, 15$ for $R = 2, 3, 4$.

the Richardson-Romberg extrapolation method are set according to Proposition 2.5:

$$n(\varepsilon) = \left\lceil \left(\frac{2\alpha R}{\beta} + 1 \right)^{\frac{1}{\alpha R}} \mu_R^{\frac{1}{\alpha R}} \varepsilon^{-\frac{1}{\alpha R}} \right\rceil \quad \text{and} \quad M(\varepsilon) = \left\lceil \gamma_0^{\frac{1}{\beta}} \nu_R^{\frac{2}{\beta}} \left(1 + \frac{\beta}{2\alpha R} \right)^{\frac{2}{\beta}} \varepsilon^{-\frac{2}{\beta}} \right\rceil.$$

The target accuracy ε for the L^1-error has been set at $\varepsilon = 2^{-p}$, $p = 1, \cdots, 4$. The L^1-error is estimated using 400 runs of the algorithm. The results are summarized in Table 4.1 for the Richardson-Romberg extrapolation SA method and in Table 4.2 for the crude SA method.[1] Note that as expected the L^1-error is always lower than the specified ε for our estimators. Using the Richardson-Romberg SA scheme instead of the crude SA method leads to a gain in terms of CPU-time varying from 12 (for $\varepsilon = 5.00 \times 10^{-1}$) to 66 (for $\varepsilon = 6.25 \times 10^{-2}$).

5 Technical results

We provide here some useful technical results that are used repeatedly throughout the paper. For a proof the reader may refer to [Fri13].

[1]The computations were performed on a computer with 4 multithreaded(16) octo-core processors (Intel(R) Xeon(R) CPU E5-4620 @ 2.20GHz).

Target accuracy: ε	L^1-error	time (s)	R	n	M
5.00×10^{-1}	3.21×10^{-1}	0.9×10^{1}	2	14	8.69×10^{5}
2.50×10^{-1}	4.80×10^{-2}	5.15×10^{1}	2	20	3.48×10^{6}
1.25×10^{-1}	4.32×10^{-2}	1.70×10^{2}	3	8	1.21×10^{7}
6.25×10^{-2}	3.48×10^{-2}	7.92×10^{2}	3	10	4.85×10^{7}

Table 4.1: Richardson-Romberg SA estimators for the quantile at level ℓ of a geometric Brownian motion with a target accuracy $\varepsilon = 2^{-p}$, $p = 1, \cdots, 4$.

Target accuracy: ε	L^1-error	time (s)	n	M
5.00×10^{-1}	2.09×10^{-1}	$1,09 \times 10^{2}$	235	1.25×10^{6}
2.50×10^{-1}	3.84×10^{-2}	8.18×10^{2}	469	5.01×10^{6}
1.25×10^{-1}	3.48×10^{-2}	7.09×10^{3}	938	2.00×10^{7}
6.25×10^{-2}	2.91×10^{-2}	5.25×10^{4}	1876	8.01×10^{7}

Table 4.2: Crude SA estimators for the quantile at level ℓ of a geometric Brownian motion with a target accuracy $\varepsilon = 2^{-p}$, $p = 1, \cdots, 4$.

Lemma 5.1. *Let $a, b > 0$. Suppose that **(HUA)** is satisfied. Let $(\gamma_n)_{n\geq 1}$ be a sequence satisfying **(HS)**. If $\gamma(t) = \gamma_0/t$, $t \geq 1$, suppose $b\underline{\lambda}\gamma_0 > a$. Let $(v_n)_{n\geq 1}$ be a non-negative sequence. Then, for some positive constant $C := C(\underline{\lambda}, \gamma)$, one has*

$$\limsup_n \gamma_n^{-a} \sum_{k=1}^{n} \gamma_k^{1+a} ||\Pi_{k+1,n}||^b v_k \leq C \limsup_n v_n,$$

where $\Pi_{k,n} := \prod_{j=k}^{n}(I_d - \gamma_j Dh(\theta^*))$, with the convention that $\Pi_{n+1,n} = I_d$.

Lemma 5.2. *Let $(\theta_p^n)_{p\geq 0}$ be the scheme defined by (1.3), θ_0^n being independent of the innovation with $\sup_{n\geq 1}\mathbb{E}|\theta_0^n|^2 < +\infty$. Suppose that **(HUA)**, **(HC)** and **(HS)** hold. Then, for some constant $C > 0$, one has:*

$$\forall p \geq 1, \quad \sup_{n\geq 1} \mathbb{E}[|\theta_p^n - \theta^{*,n}|^2] \leq C\gamma(p)$$

Bibliography

[A$^+$81] R. Azencott et al. *Géodésiques et diffusions en temps petit*, volume 1984-1985. Astérisque, S.M.F., 1981.

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

[Aro59] D. G. Aronson. The fundamental solution of a linear parabolic equation containing a small parameter. *Ill. Journ. Math.*, 3:580–619, 1959.

[Aro67] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc. 73*, pages 890–896, 1967.

[BAL91] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.

[Bas88] R. F. Bass. Uniqueness in law for pure jump Markov processes. *Probab. Theory Related Fields*, 79(2):271–287, 1988.

[Bas95] R. F. Bass. *Probabilistic techniques in Analysis.* Springer, 1995.

[Bas97] R. F. Bass. *Diffusions and Elliptic Operators.* Springer, 1997.

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

[BC06] R.F. Bass and Z.Q. Chen. Systems of equations driven by stable processes. *Probab. Theory Related Fields*, 134(2):175–214, 2006.

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

[BKH14] V. Bally and A. Kohatsu-Higa. A probabilistic interpretation of the parametrix method. *Preprint*, 2014.

[BMP90] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. Translated from the French by Stephen S. Wilson.

[BP09] R.F. Bass and E.A. Perkins. A new technique for proving uniqueness for martingale problems. *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut*, pages 47–53, 2009.

[BPP09] S. Borovkova, F. J. Permana, and I. Pavlyukevich. Modelling electricity prices by potential lévy diffusions. *Journal of Energy Markets*, 2:83–110, 2009.

[BPV01] E. Barucci, S. Polidoro, and V. Vespri. Some results on partial differential equations and asian options. *Math. Models Methods Appl. Sci*, 3:475–497, 2001.

[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.

[BT09] Richard F. Bass and Huili Tang. The martingale problem for a class of stable-like processes. *Stochastic Process. Appl.*, 119(4):1144–1167, 2009.

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

[CKK08] Z.Q. Chen, P. Kim, and T. Kumagai. Weighted poincaré inequality and heat kernel estimates for finite range jump processes. *Mathematische Annalen*, 342(4):833–883, 2008.

[CMP15] C. Cinti, S. Menozzi, and S. Polidoro. Two-sided bounds for degenerate processes with densities supported in subsets of $\mathbb{R}^n$. *Potential Analysis*, 42-1:39–98, 2015.

[CW71] Ronald R. Coifman and Guido Weiss. *Analyse harmonique non-commutative sur certains espaces homogènes*. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971. Étude de certaines intégrales singulières.

[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.

[DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259(6):1577–1630, September 2010.

[Duf96] M. Duflo. *Algorithmes stochastiques*, volume 23 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996.

[EPRB99] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Comm. Math. Phys.*, 201–3:657–697, 1999.

[FF13] M. Fathi and N. Frikha. Transport-entropy inequalities and deviation estimates for stochastic approximation schemes. *Electron. J. Probab.*, 18:no. 67, 1–36, 2013.

[FM12] N. Frikha and S. Menozzi. Concentration bounds for stochastic approximations. *Electron. Commun. Probab.*, 17:no. 47, 15, 2012.

[Fra14] J. Franchi. Small time asymptotics for an example of strictly hypoelliptic heat kernel. *To appear in Séminaire de Probabilités*, 2014.

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

[Fri13] N. Frikha. Multi-level stochastic approximation algorithms. *hal-00870585v1*, page 47, 2013.

[Had32] Jacques Hadamard. *Le probleme de Cauchy et les équations aux dérivées partielles linéaires hyperboliques*, volume 220. Paris, 1932.

[Had64] Jacques Hadamard. *La théorie des équations aux dérivées partielles*. Éditions Scientifiques, Peking; Gauthier-Villars Éditeur, Paris, 1964.

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

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

[Jac96] N. Jacob. *Pseudo Differential Operators and Markov Process*. Akademie Verlag, 1996.

[Jac05] N. Jacob. *Pseudo differential operators and Markov processes*, volume 1. Imperial College Press, 2005.

[JKLS12] Niels Jacob, Victorya Knopova, Sandra Landwehr, and René L. Schilling. A geometric interpretation of the transition density of a symmetric Lévy process. *Sci. China Math.*, 55(6):1099–1126, 2012.

[JP98] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1):267–307, 1998.

[JYC09] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets*. Springer Finance. Springer-Verlag London Ltd., London, 2009.

[KK14] V. Knopova and A. Kulik. Parametrix method and the weak solution to an sde driven by an α-stable noise. *arXiv:1412.8732*, 2014.

[KM02] V. Konakov and E. Mammen. Edgeworth type expansions for euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.*, 8–3:271–285, 2002.

[KM10] V. Konakov and S. Menozzi. Weak error for stable driven sdes: expansion for the densities. *Journal of Theoretical Probability*, 24(2):454–478, 2010.

[KMM10] V. Konakov, S. Menozzi, and S. Molchanov. Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Annales de l'Institut Henri Poincaré, Série B*, 46–4:908–923, 2010.

[Kol00a] V. Kolokoltsov. *Semiclassical Analysis for Diffusions and Stochastic Processes*, volume LMN, 1724. Springer, 2000.

[Kol00b] V. Kolokoltsov. Symmetric stable laws and stable-like diffusion. *Proceedings of the London Mathematical Society*, 80(03):725–768, 2000.

[Kom84] Takashi Komatsu. Pseudodifferential operators and Markov processes. *J. Math. Soc. Japan*, 36(3):387–418, 1984.

[Kom08] Takashi Komatsu. On the martingale problem for pseudo-differential operators of variable order. *Theory Stoch. Process.*, 14(2):42–51, 2008.

[Kry87] N.V. Krylov. *Controlled diffusion processes*. Springer, 1987.

[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.

[KY03] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[Lar11] S. Laruelle. *Analysis of Stochastic Algorithms Applied to Finance*. Theses, Université Pierre et Marie Curie - Paris VI, December 2011.

[Léa85] R. Léandre. Régularité de processus de sauts dégénérés. *Annales de l'IHP Probabilités et statistiques*, 21:125–146, 1985.

[Léa88a] R. Léandre. Régularité de processus de sauts dégénérés (ii). In *Annales de l'IHP Probabilités et statistiques*, volume 24, pages 209–236. Elsevier, 1988.

[Léa88b] Rémi Léandre. Calcul des variations sur un brownien subordonné. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 414–433. Springer, Berlin, 1988.

[Lev07] E. E. Levi. Sulle equazioni lineari totalmente ellittiche alle derivate parziali. *Rendiconti del circolo Matematico di Palermo*, 24(1):275–317, 1907.

[LM10] V. Lemaire and S. Menozzi. On some non asyptotic bounds for the euler scheme. *Electronic Journal of Probability*, 15:1645–1681, 2010.

[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.

[MSH02] J. Mattingly, A. Stuart, and D. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101–2:185–232, 2002.

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

[Pag07] G. Pagès. Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods Appl.*, 13(1):37–70, 2007.

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

[PT69] W.E. Pruitt and S.J. Taylor. The potential kernel and hitting probabilities for the general stable process in rn. *Transactions of the American Mathematical Society*, pages 299–321, 1969.

[PZ09] E. Priola and J. Zabczyk. Densities for Ornstein–Uhlenbeck processes with jumps. *Bulletin of the London Mathematical Society*, 41(1):41–50, 2009.

[RM51] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Statistics*, 22:400–407, 1951.

[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.

[Sim11] Thomas Simon. On the absolute continuity of multidimensional Ornstein-Uhlenbeck processes. *Probab. Theory Related Fields*, 151(1-2):173–190, 2011.

[ST94] G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes: stochastic models with infinite variance*, volume 1. CRC Press, 1994.

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

[Tal02] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2:163–198, 2002.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and App.*, 8-4:94–120, 1990.

[TZ04] D. Talay and Z. Zheng. Approximation of quantiles of components of diffusion processes. *Stochastic Processes and their Applications*, 109-1:23–46, 2004.

[Wat01] Toshiro Watanabe. Temporal change in distributional properties of Lévy processes. In *Lévy processes*, pages 89–107. Birkhäuser Boston, Boston, MA, 2001.

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

[Yam11] M. Yamazato. Absolute continuity conditions for multivariate infinitely divisible distributions and their applications. *Teor. Veroyatn. Primen.*, 56(2):351–366, 2011.

[Zha14a] X. Zhang. Densities for SDEs driven by degenerate α-stable processes. *Ann. Probab.*, 42(5):1885–1910, 2014.

[Zha14b] X. Zhang. Nonlocal hormander's hypoellipticity theorem. *arXiv:1306.5016*, 2014.

[Zol86] V.M. Zolotarev. *One-dimensional stable distributions*, volume 65. American Mathematical Soc., 1986.