SCUOLA
NORMALE
SUPERIORE

Classe di Scienze

Corso di perfezionamento in Metodi Computazionali e Modelli Matematici per le Scienze e la Finanza

XXXV ciclo

On semilinear SDEs driven by α–stable noise, affine Volterra processes with jumps and their applications

Settore Scientifico Disciplinare **MAT/06**

Candidato
dr. Alessandro Bondi

Relatore
Prof. Franco Flandoli

Supervisione interna
Prof.ssa Giulia Livieri

Anno accademico 2022/2023

Abstract

This dissertation explores various topics related to stochastic dynamics with jumps, which find their initial motivation in numerical applications. It undertakes two distinct research trajectories, which ultimately converge in the concluding chapter of the thesis, where a potential interconnection between them is established from a theoretical perspective.

In the first line of study, the concepts of cylindrical Wiener process subordinated to a strictly $\alpha-$stable Lévy process, with $\alpha \in (0,1)$, and of the corresponding stochastic convolution are introduced in an infinite-dimensional, separable Hilbert space. The related Ornstein-Uhlenbeck (OU) process is then analyzed, with a focus on the regularizing properties of the associated Markov transition semigroup. In particular, an original formula –which is not of Bismut-Elworthy-Li's type– for the Gateaux derivatives of the functions generated by this semigroup is provided, together with an estimate for the norm of their gradients.

Taking $\alpha \in (\frac{1}{2}, 1)$, these results are applied to the study of semilinear, $N-$dimensional stochastic differential equations (SDEs) driven by the same additive, isotropic, stable Lévy noise. An important connection between the time-dependent Markov transition semigroup associated with their solutions and Kolmogorov backward equations in mild integral form is established via regularization-by-noise techniques. Such a link is the starting point for an iterative method which allows to approximate probabilities related to the SDEs with a single batch of Monte Carlo simulations as several parameters change, bringing a compelling computational advantage over the standard Monte Carlo approach. This method also pertains to the numerical computation of solutions of high-dimensional integro-differential Kolmogorov backward equations. The scheme, and in particular the first order approximation it provides, is then tested for two nonlinear vector fields in dimension $N = 100$ and shown to offer satisfactory results, especially when compared with the OU approximation.

Within this analysis, one of the concepts employed is the stochastic flow generated by SDEs with additive Lévy noise. In this dissertation, an extension to the case of multiplicative noise of some results known in the existing literature for the additive case is presented. More specifically, the existence of a sharp stochastic flow $X_t^{s,x}$ for an SDE of Itô's type with multiplicative noise which, $\mathbb{P}-$a.s., is simultaneously continuous in x (starting point) and càdlàg in s (starting time) and t (time) is proved. Remarkably, the study encompasses SDEs that include both compensated and non-compensated jump components, thereby addressing both small and large jumps, alongside a Brownian diffusion term. The theory is further expanded to cover controlled SDEs. Using the resulting sharp stochastic flow, a new dynamic programming principle is established with an argument that stands as an independently significant point of interest.

The second area of research regards the theory of affine processes, which has been recently extended to stochastic Volterra equations (SVEs) with continuous trajectories. These so-called affine Volterra processes overcome modeling shortcomings of classical affine processes because they may possess path-dependent features which introduce memory structures into the models. Furthermore, they can have

trajectories whose regularity is different from the paths of Brownian motion. In particular, singular kernels yield rough affine processes. In this thesis, a generalization of the above-mentioned theory by considering affine SVEs with jumps is studied. This extension is not straightforward because the jump structure together with possible singularities of the kernel may induce explosions of the trajectories. Nonetheless, the extended framework enables to obtain semi-explicit formulas for the conditional Fourier-Laplace transforms of the solutions via deterministic Riccati-Volterra equations. This study also provides exponential affine expressions for the conditional transforms of marked Hawkes processes.
Building upon this analysis, an extension of the rough Heston stochastic volatility model is introduced. In this extension, called the rough Hawkes Heston stochastic volatility model, the instantaneous spot variance is modeled as the solution of an affine SVE of convolution type with jumps. This setting takes into account both rough volatility and jump clustering phenomena, and employs the affine structure of the SVE to price options on the underlying (SPX) and on the related volatility index (VIX) via Fourier inversion techniques. A calibration of the model featuring a fractional kernel and an exponential distribution for the jumps is carried out, demonstrating its ability to accurately and simultaneously capture the volatility smiles of both SPX and VIX options. This is a remarkable result, especially considering the few parameters of the proposed model, namely five for the evolution of the dynamics and two for the term structure. Moreover, it proves the relevance, under an affine framework, of rough volatility and self-exciting jumps in order to describe the joint evolution of SPX and VIX.

Lastly, an exploration of the theoretical interconnection between the subjects of the two preceding areas of research is investigated. More precisely, a Volterra convolution equation in $\mathbb{R}^d$ perturbed with an additive fractional Brownian motion of Riemann–Liouville type characterized by a Hurst parameter $H \in (0,1)$ is considered. The solution of this equation is shown to satisfy a stochastic partial differential equation (SPDE) in the Hilbert space of square-integrable functions. This particular equation serves as the motivation for the study of an unconventional class of SPDEs, necessitating an original extension of the drift operator and its Fréchet differentials. It is demonstrated that these SPDEs generate a Markov stochastic flow which is twice Fréchet differentiable with respect to the initial data. This stochastic flow is subsequently employed to solve, in the classical sense of infinite-dimensional calculus, the path-dependent Kolmogorov equation corresponding to the SPDEs. In particular, a time-dependent infinitesimal generator is associated with the fractional Brownian motion. Certain challenges arise in the analysis of the mild formulation of the Kolmogorov equation for SPDEs driven by the same infinite-dimensional noise. This issue, which is relevant to the theory of regularization-by-noise, remains an open area for future research.

Contents

Introduction

The objective of this dissertation is to analyze a range of concepts related to stochastic dynamics with jumps, initially motivated by their relevance in numerical applications. The thesis pursues two separate research directions in Part I and Part II, ultimately exploring a connection between them from a theoretical perspective in Part III.
All the chapters of the thesis stem from a series of papers written during the PhD studies, see [34, 35, 36, 37, 38, 39].

Part I In this part, we study smoothing effect and derivative formulas for the transition semigroup of Ornstein–Uhlenbeck (OU) processes in Hilbert spaces driven by Brownian motions subordinated to $\alpha-$stable Lévy processes. In a finite–dimensional setting, such formulas are employed to develop an iterative scheme based on Kolmogorov equations in mild form. This scheme allows to numerically compute probabilities related to the solutions of semilinear stochastic differential equations (SDEs) driven by the same noise in an efficient way as several parameters of the dynamics change. Furthermore, we analyze the regularity of the stochastic flow generated by a wider class of SDEs with multiplicative noise and jumps. The *sharp* flow that we construct for controlled SDEs is then used to establish a new dynamic programming principle.

Chapter 1 presents the results of the paper [34]. Our aim here is to analyze OU processes Z^x, $x \in H$, driven by subordinated cylindrical Brownian noises, where $(H, \|\cdot\|_H)$ is an infinite–dimensional, separable Hilbert space. They are defined as the $H-$valued, mild solutions of the linear stochastic differential equations

$$\mathrm{d}Z_t^x = AZ_t^x\mathrm{d}t + \sqrt{Q}\,\mathrm{d}W_{L_t}, \quad Z_0^x = x \in H,$$

where $A\colon \mathcal{D}(A) \subset H \to H$ is a linear, selfadjoint, negative definite, unbounded operator, and $Q\colon H \to H$ is a linear, bounded, nonnegative definite operator. By construction, A and Q share a common complete orthonormal system (in short, CONS) of eigenvectors for H: it is denoted by $(e_n)_n$.
The main contribution of our research consists in obtaining regularization-by-noise results and gradient estimates for the Markov transition semigroup $R = (R_t)_{t\geq 0}$ associated with Z^x, $x \in H$, defined by

$$R_t\phi(x) = \mathbb{E}\left[\phi\left(Z_t^x\right)\right], \quad x \in H,\, \phi \in \mathcal{B}_b(H),\, t \geq 0,$$

by means of a derivative formula which is not of Bismut–Elworthy–Li's type. This is accomplished in correspondence of a particular structure for the noise W_L. Intuitively speaking, it can be thought of as

$$W_{L_t} = \sum_{n=1}^{\infty} \beta_{L_t}^n e_n, \quad t \geq 0,$$

where $(\beta^n)_n$ is a sequence of independent Brownian motions and $L = (L_t)_t$ is an independent, strictly α–stable subordinator (i.e., an increasing Lévy process) representing the random time–change, for $\alpha \in (0,1)$. Therefore the random perturbation W_L is a subordinated cylindrical Wiener process. The rigorous investigation of the convergence of the series defining W_L is detailed in Section 1.1, where we propose a natural procedure of independent interest essentially relying on Markov's inequality.

Similar models have been introduced, with different research purposes, by [44]. Here the authors define the noise by directly subordinating a cylindrical Wiener process, and the stochastic convolution by an integral with respect to a Lévy process taking values in a separable Banach space (the noise itself). An abstract approach to cylindrical Lévy processes in separable Banach spaces and to the theory of integration for deterministic, operator–valued integrands with respect to this type of noise is developed in [157, 158]. However, in our setting such theories are not necessary, as the choice of an $\alpha-$stable subordinator allows to devise an original technique which is more direct, as it is based on the one–dimensional integration theory.

In literature, the canonical case is the Gaussian one, which involves a cylindrical Wiener process $W_t = \sum_{n=1}^{\infty} \beta_t^n e_n$, $t \geq 0$. There is a well–established theory concerning this setting, and we may refer to the book [66] for an extensive collection of results on the subject.
Another important framework is the one proposed by [151], where the authors deal with a cylindrical, α–stable Lévy process $Z_t = \sum_{n=1}^{\infty} \zeta_t^n e_n$, $t \geq 0$. Here $(\zeta^n)_n$ are independent, real–valued, symmetric α–stable Lévy processes, for $\alpha \in (0,2)$. Despite the interesting generalization offered by this approach, the structure of the noise could be questionable in some applications, especially in physics. Indeed, fixing $t > 0$ and $N \in \mathbb{N}$, the Galerkin projection of Z_t has characteristic function

$$\mathbb{E}\left[e^{i\left\langle h, \sum_{n=1}^{N} \zeta_t^n e_n\right\rangle}\right] = e^{-t\gamma^\alpha \sum_{n=1}^{N} |\langle h, e_n\rangle|^\alpha}, \quad h \in H,$$

for some constant $\gamma > 0$. Thus, unlike the Brownian case, the rotational stability of the noise, often referred to as *isotropy*, is lost.

Motivated by this argument, it is worth studying the results contained in the aforementioned works [66, 151] also for the subordinated process W_L, since its Galerkin projections are 2α–stable, isotropic Lévy processes, as we shall discuss in Section 1.1. We focus on the linear case, which corresponds to OU processes. In this framework, a number of complications arises, the most evident being the lack of independence of the processes $(\beta_L^n)_n$. Generally speaking, this fact makes the techniques used in the other cases unfeasible. Nevertheless, the structure of the noise still enables to construct the objects of our interest and effectively advance our arguments. The intuition is that, upon conditioning with respect to the $\sigma-$algebra $\mathcal{F}^L$ generated by the subordinator L, we are dealing with time–shifted Brownian motions. Thus, our strategy consists of first studying separately, in a finite–dimensional framework, a deterministic time–change $\ell\colon \mathbb{R}_+ \to \mathbb{R}_+$, where ℓ is an increasing càdlàg function starting at 0 with $\ell > 0$ in $(0,\infty)$. This way to proceed is customary while working with subordinated Brownian motions, see, for instance, [127, 178], and also [69] for a related example.

The main achievement in Chapter 1 is Theorem 1.7, which establishes that, under suitable requirements on the operators A and Q, the map $R_t\phi$ belongs to $C_b^1(H)$ for every $\phi \in \mathcal{B}_b(H)$. Furthermore, it provides a gradient estimate for $R_t\phi$ (see (1.25)), and, when $\phi \in C_b(H)$, an expression for the Gateaux derivative $\langle \nabla R_t\phi\,(x)\,, h\rangle$, $x, h \in H$ (see (1.26)). Remarkably, this formula is not of Bismut–Elworthy–Li's type, which is consistent with the Gaussian framework, where it is preferable to employ the Bismut–Elworthy–Li's formula primarily in nonlinear scenarios. It is also worth highlighting a subtle difference between the finite– and infinite–dimensional cases: in the former, we obtain an expression for the Gateaux derivative of $R_t\phi$ for every $\phi \in \mathcal{B}_b(\mathbb{R}^N)$, see (1.18) in Theorem 1.6.

Assuming $\alpha \in (\frac{1}{2}, 1)$, Corollary 1.8 deduces from (1.25) the following gradient estimate, which holds for all $\phi \in \mathcal{B}_b(H)$ and $t > 0$:

$$\sup_{x\in H} \|\nabla R_t\phi(x)\|_H \leq \frac{C}{t^\gamma} \sup_{x\in H} |\phi(x)|, \quad \text{for some } C > 0 \text{ and } \gamma \in \Big[\frac{1}{2\alpha}, 1\Big).$$

This estimate is crucial, as it serves as the starting point for the analysis of the Kolmogorov equation in mild integral form employing fixed–point arguments. In the finite–dimensional setting, such an analysis is one of the topics discussed in Chapter 2.

Chapter 2 contains the findings of the paper [35]. Here, we consider the semilinear $N-$dimensional SDE

$$\begin{cases} \mathrm{d}X_t = (AX_t + B_0(t, X_t))\,\mathrm{d}t + \sqrt{Q}\,\mathrm{d}W_{L_t}, \quad t \in [s, T], \\ X_s = x \in \mathbb{R}^N, \end{cases} \tag{I.1}$$

with a specific interest in the case N high. As in Chapter 1, given $\alpha \in \left(\frac{1}{2}, 1\right)$, L is an $\alpha-$stable subordinator independent from $(\beta^n)_{n=1,\dots,N}$, which in turn are independent Brownian motions; we write $W = [\beta^1, \dots, \beta^N]^\top$. All these processes are defined in a common complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which we endow with the minimal augmented filtration generated by the subordinated Brownian motion W_L. Moreover, $T > 0$ is a finite time horizon and $s \in [0, T]$ is the initial time. As for $A, Q \in \mathbb{R}^{N\times N}$, they are diagonal matrices with A negative–definite and Q positive–definite. For our numerical experiments, see Section 2.6, we will consider $Q = \sigma^2\mathrm{Id}$, being $\mathrm{Id} \in \mathbb{R}^{N\times N}$ the identity matrix, so that $\sigma > 0$ is a parameter describing the strength of the noise. The nonlinear bounded vector field $B_0\colon [0, T] \times \mathbb{R}^N \to \mathbb{R}^N$ is subject to suitable regularity conditions which will be specified in the sequel and guarantee, among other things, the existence of a pathwise unique solution of (I.1). Such a solution will be denoted by $X^{s,x} = (X^{s,x}_t)_{t\in[s,T]}$.

Connected to the SDE (I.1), we have the following *Kolmogorov backward equation*:

$$\begin{cases} \partial_s u(s, x) = -\left\langle Ax + B_0(s, x), \nabla^\top u(s, x)\right\rangle \\ \qquad\qquad - \int_{\mathbb{R}^N} \left[u\left(s, x + \sqrt{Q}z\right) - u(s, x) - 1_D(z)\,\nabla u(s, x)\,\sqrt{Q}z\right]\nu(\mathrm{d}z), & s \in [0, t), \\ u(t, x) = \phi(x), & x \in \mathbb{R}^N, \end{cases} \tag{I.2}$$

where $\phi\colon \mathbb{R}^N \to \mathbb{R}$, $D = \{z \in \mathbb{R}^N, |z| \leq 1\}$ is the closed unit ball and we fix $t \in [0, T]$. Here $\nu(\mathrm{d}z)$ is the Lévy measure of W_L; up to a positive multiplicative constant, $\nu(\mathrm{d}z)$ is of the form (see, e.g., [164, Theorem 30.1])

$$\nu(\mathrm{d}z) = |z|^{-(N+2\alpha)}\,\mathrm{d}z.$$

The link between Equations (I.1) and (I.2) is provided by Theorem 2.7 (*ii*) (see also the book [126] for related results), where it is shown that the time–dependent Markov transition semigroup

$$P_{s,t}\phi(x) = \mathbb{E}\left[\phi\left(X^{s,x}_t\right)\right]$$

associated with (I.1) satisfies (I.2) in the closed interval $[0, t]$ for every $\phi \in C^3_b\left(\mathbb{R}^N\right)$. Moreover, we are able to extend the validity of this connection in $[0, t)$ to every function $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$ through an original procedure based on regularization-by-noise and a mild, integral formulation of (I.2) (see Remark 2.1 in Section 2.2).

In this chapter, our primary focus is precisely directed towards calculating expected values of the solution process $X^{s,x}$. We place particular attention to the case $\phi(x) = 1_{\{|x|>R\}}$, where $\mathbb{E}[\phi(X_t^{s,x})] = \mathbb{P}(|X_t^{s,x}| > R)$, for some threshold $R > 0$. Consequently, our objective is to describe a method which allows to compute probabilities related to the solution of the SDE (I.1).

Attempting to derive an estimate of these probabilities by numerically solving the integro–differential equation (I.2) is a typical example of *curse of dimensionality* (CoD). Given our intention to work in a high–dimensional context (we set $N = 100$ in our simulations), pursuing this strategy becomes unfeasible. The conventional approach to tackle our problem is the Monte Carlo method. Here, multiple paths of $X^{s,x}$ are simulated using the Euler–Maruyama scheme with a small time step. Subsequently, the average of the final points of these trajectories is computed to obtain an approximation of the desired expected values by virtue of the strong law of large numbers. This approach is known to be free from the CoD. However, if we were to follow this scheme, then we would have to start over the procedure every time we change the starting point x and the starting time s, the noise strength σ and even the nonlinearity B_0. This is a very common practice in a wide range of applications including weather forecasts and calibration of financial models, see [22] and references therein.

In order to overcome this setback, we aim at extending to our framework the ideas developed in the papers [84, 85] for the Gaussian case. In particular, we search for an iterative scheme $v_s^n(t,x)$, $n \in \mathbb{N} \cup \{0\}$, which relies on a single bulk of Monte Carlo simulations independent from the aforementioned parameters such that

$$P_{s,t}\phi(x) = \sum_{n=0}^{\infty} v_s^n(t,x). \tag{I.3}$$

More specifically, to approximate the value of the iterates $v_s^n(t,x)$, we only need to simulate, in a one–time effort, a large number of sample paths of the subordinator L and of the stochastic convolution $\widetilde{Z}_t^0 = \int_0^t e^{(t-r)A} \mathrm{d}W_{L_r}$ using the Euler–Maruyama scheme. We recall that $(\widetilde{Z}_t^0)_{t\in[0,T]}$ is the unique (up to indistinguishability) solution of the linear SDE

$$\mathrm{d}\widetilde{Z}_t^0 = A\widetilde{Z}_t^0 \mathrm{d}t + \mathrm{d}W_{L_t}, \qquad \widetilde{Z}_0^0 = 0.$$

The primary innovation of the approach that we propose lies in the structure of the noise W_L, which is a $2\alpha-$stable, rotation–invariant Lévy process (cfr. [164, Example 30.6]). The introduction of L makes the framework more complex compared to the Brownian one treated in [84, 85]. This fact leads us to develop an original procedure to obtain an expression for the iterates which is suitable for applications. Such a procedure is based on conditioning with respect to the $\sigma-$algebra $\mathcal{F}^L$ generated by the subordinator, and employs the derivative formulas established in Chapter 1.

The theoretical foundation of the iterative method (I.3), namely Theorem 2.3, has a remarkable interest on its own. Indeed, it establishes a connection between the time–dependent Markov transition semigroup associated with (I.1) and a mild, integral formulation of (I.2) (see Equation (I.4)) that, to the best of our knowledge, is new when it comes to isotropic, stable Lévy processes.
More precisely, given a continuous function $f\colon [0,T] \to \mathbb{R}^N$, $x \in \mathbb{R}^N$ and $0 \le s < T$, we denote by $Z^{s,x} = (Z_t^{s,x})_{t\in[s,T]}$ the OU process starting from x at time s, i.e., the unique solution of the linear SDE

$$\mathrm{d}Z_t^{s,x} = (AZ_t^{s,x} + f(t))\,\mathrm{d}t + \sqrt{Q}\,\mathrm{d}W_{L_t}, \quad Z_s^{s,x} = x.$$

The corresponding time–dependent, Markov transition semigroup $R = (R_{s,t})$, $0 \le s \le t \le T$, is defined by

$$R_{s,t}\phi = \mathbb{E}\left[\phi\left(Z_t^{s,\cdot}\right)\right]$$

and enables us to write the *Kolmogorov backward equation* in mild integral form associated with (I.1):

$$u_s^{\phi}(t,x) = R_{s,t}\phi(x) + \int_s^t R_{s,r}\left(\left\langle B_0(r,\cdot) - f(r), \nabla^{\top} u_r^{\phi}(t,\cdot)\right\rangle\right)(x)\,\mathrm{d}r, \quad s \in [0,t],\, x \in \mathbb{R}^N, \tag{I.4}$$

where $\phi \in \mathcal{B}_b(\mathbb{R}^N)$. Under appropriate regularity assumptions on B_0, Theorem 2.3 demonstrates that $P_{s,t}\phi$ is the unique solution of (I.4) in a specific function Banach space. Thus, the idea of the iterative scheme (I.3) arises from a direct study of (I.4) using a fixed–point argument, as outlined in Section 2.3. The proof of Theorem 2.3 is based on a regularization-by-noise technique, which utilizes the Bismut–Elworthy–Li's type formula in [178, Theorem 1.1] as initial step. Additionally, it takes into account the *sharp stochastic flow* generated by (I.1), a concept which has been studied in [149, 150]. It is worth noting that the analysis of the sharp stochastic flow generated by a more general class of SDEs with multiplicative noise forms the core of Chapter 3.

The final part of Chapter 2, Section 2.6, is dedicated to numerical experiments in dimension $N = 100$. Here, we test the iterative scheme for two choices of the nonlinear vector field B_0, with a primary focus on evaluating the first order approximation $v^0 + v^1$. The outcomes of these experiments show that, in the given examples, the first iteration of (I.3) provides a computationally efficient enhancement to the performance of the linear OU approximation.

Chapter 3 presents the results of the paper [38]. Here, we consider a Brownian motion W and a stationary Poisson point process p with values on a measurable space $(U,\mathcal{U})$ and characteristic measure $\nu(\mathrm{d}z)$. We denote by N_p [resp., $\widetilde{N}_p$] the Poisson [resp., compensated Poisson] random measure associated with p. Given a set $U_0 \in \mathcal{U}$ such that $\nu(U \setminus U_0) < \infty$, we study the following SDE of Itô's type:

$$\begin{cases} \mathrm{d}X_t = b(r,X_r)\,\mathrm{d}r + \alpha(r,X_r)\,\mathrm{d}W_r + \int_{U_0} g(X_{r-},r,z)\,\widetilde{N}_p(\mathrm{d}r,\mathrm{d}z) + \int_{U\setminus U_0} f(X_{r-},r,z)\,N_p(\mathrm{d}r,\mathrm{d}z), \\ X_s = x \in \mathbb{R}^d, \quad 0 \le s \le t \le T. \end{cases} \tag{I.5}$$

We require the coefficients b, α, g and f to be measurable. In addition, we assume Lipschitz–type and linear growth conditions on the $x-$variable of these coefficients, which ensure the existence of a pathwise unique strong solution $X = (X_t^{s,x})_{t\ge s}$ of (I.5). We remark that the Lipschitz–assumption on the coefficient f corresponding to the large–jumps part can be dispensed with (cfr. Section IV.9 in [106] and see Hypothesis 3.1). We refer to [15, 30, 106, 125, 167] for the theory of SDEs with jumps; see also Section 3.1 for more details.

In this chapter, we first deal with the problem of finding a version of the solution of (I.5) which depends in a regular way on all the variables (s,t,x). We prove, in particular, that there is a version of the solution X which is sharp in the following sense: there exists an almost sure event Ω' such that, for every $\omega \in \Omega'$, the map $(s,x,t) \mapsto X_t^{s,x}(\omega)$ is càdlàg in s (for t and x fixed), càdlàg in t (for s and x fixed) and continuous in x (for s and t fixed). Moreover, we prove the flow property

$$X_t^{s,x}(\omega) = X_t^{u,X_u^{s,x}(\omega)}(\omega), \quad s < u < t \le T, \tag{I.6}$$

and the stochastic continuity in s, locally uniformly in x and uniformly in t (see (3.4)). We call this version a *sharp stochastic flow* for Equation (I.5). We refer to Definition 3.1 and Theorems 3.1-3.2 for more general assertions.

We recall that, for SDEs driven by a Brownian motion, namely Equation (I.5) with $f \equiv 0$ and $g \equiv 0$, it is well known that there exists a stochastic flow $X = (X_t^{s,x})_{t\ge s}$ such that, for $\mathbb{P}-$a.e. ω, the mapping

$(s,x,t) \mapsto X_t^{s,x}(\omega)$ is continuous in s, t and x, and such that the flow property (I.6) holds. This is a consequence of the Komogorov–Chentsov test, which can be applied thanks to the estimate (see, for instance, [126, Theorem 3.4.3])

$$\mathbb{E}\Big[\big|X_t^{s,x} - X_{t'}^{s',x'}\big|^p\Big] \leq C\left(|x-x'|^p + |s-s'|^{p/2} + |t-t'|^{p/2}\right), \quad p \geq 2.$$

This continuous stochastic flow is deeply investigated in [124], where, in particular, it is employed to study first order stochastic PDEs when the coefficients are sufficiently smooth.

Concerning (I.5) in the case $f \equiv 0$, the problem of finding a sharp version of the solution also appears in the comment before Theorem 3.4.3 in [126]. Here, the author compares the stochastic flow available for SDEs driven by a Brownian motion with weaker results available in the jump case (see, in particular, [126, Theorem 3.4.1]). On this respect, note that the previous technique based on the Komogorov–Chentsov test can only give a continuous modification. As a consequence, it cannot be applied to SDEs with jumps in order to obtain a version of the solution which depends in a regular way on (s,t,x).

The main challenge in analyzing the regularity of the flow with respect to (s,t,x) in the jump case is to take care of the dependence on the initial time s. Such a problem has also been mentioned in [150, Remark 1.2] (see also [149, Introduction]). In this regard, after the proof of [125, Theorem 3.2] and in [126, Theorem 3.4.1], it is shown that, when s is fixed, there is a modification of the solution X with the following property: there exists an almost sure event Ω_s (possibly depending on s) such that, for every $\omega \in \Omega_s$, (I.6) holds and the map $(x,t) \mapsto X_t^{s,x}(\omega)$ is càdlàg in t (for s and x fixed) and continuous in x (for s and t fixed). However, since the regularity in s is not investigated, the resulting definition of stochastic flow given on pages 353–354 of [125] is not satisfactory[1]. Our work fills this gap.

We also mention [115], which considers the SDE

$$\mathrm{d}X_t = l(X_{t-})\,\mathrm{d}Z_t, \quad t \geq s, \qquad X_s = x \in \mathbb{R}^d, \tag{I.7}$$

from the point of view of random dynamical systems. Here $(Z_t)_{t\geq 0}$ is an $\mathbb{R}^k-$valued semimartingale and $l : \mathbb{R}^d \to \mathbb{R}^{d\times k}$ is Lipschitz continuous. Theorem 5 in [115] implies that there exists a version of the solution $X_t^{s,x}$ which is sharp in the variables (s,t,x) and satisfies the flow property (the stochastic continuity is not investigated in [115]). This result is applicable to the SDE (I.5) only under the assumption that the coefficients b and α are time–independent. Additionally, to apply this theorem, it is required that both g and f have a specific form (see Remark 3.1 in Section 3.1 for more details). Another related result is [50, Theorem 17.1.4], which investigates a general class of SDEs having the Markov property. In this theorem, the initial time s is fixed and one proves the existence of a version of the solution which depends in a measurable way on (t,x).

[1] We make some more comments on the stochastic flow for SDEs with "small jumps" studied in [125] and [126] (see Equations (3.4) and (3.37), respectively; these are similar to (I.5) with $f \equiv 0$).

(i) [125, Theorem 3.2] and [126, Theorem 3.4.1] show the existence of a modification of the solution which is, $\mathbb{P}-$a.s., continuous in the initial state $x \in \mathbb{R}^d$ and càdlàg in t: the regularity with respect to the initial time s is not considered. On the other hand, a version of the solution which is càdlàg in s is given by [126, Proposition 3.8.2], which requires to fix a time t. These results do not prove the simultaneous càdlàg property in s and t. Thus, the claims at the beginning of [125, Page 354] are not completely proved.

(ii) When the flow property (I.6) is stated in [125, Pages 353–354] and [126, Page 99], the dependence of the almost sure event on the variables s,t,x is not explicit.

One of the goals of our research consists in clarifying the points raised by (i) and (ii) with a thorough procedure. We do not, however, discuss the homeomorphism property of (I.5), which is the subject of [125, Section 3.4].

In Section 3.4 we argue that the sharp stochastic flow exists even when we consider more general coefficients depending, in a Lipschitz–continuous way, on an additional parameter $\mathbf{y} \in \mathbb{R}^k$. This fact is crucial for Section 3.5, where we focus on controlled SDEs and the corresponding dynamic programming principle (DPP), which is a fundamental concept in the theory of stochastic control (see, for instance, [27, 40, 108, 123, 145, 146, 148, 177]). In this part, we consider controlled SDEs like (I.5) where the coefficients b, α, g and f depend, in a Lipschitz–continuous way, also on an admissible simple (or step) control of the form

$$a(t,\omega) = \sum_{i=0}^{\bar{n}-1} Z_{t_i}(\omega) 1_{(t_i,t_{i+1}]}(t), \quad t \in [0,T],\, \omega \in \Omega, \tag{I.8}$$

for some $\bar{n} \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_{\bar{n}} = T$ and suitable square–integrable random variables Z_{t_i}, $i = 0, \ldots, \bar{n}-1$. We are able to construct a sharp stochastic flow in this framework, as well. In particular, such a flow enables us to consider identities like

$$X_r^{s,x,a} = X_r^{\theta, X_\theta^{s,x,a}, a},$$

which are meaningful even when θ is a stopping time taking values in (s,T). This kind of identities is useful for the proof of the DPP (see also Remark 3.5).
We prove a new DPP in the finite–horizon case, namely

$$v(s,x) = \sup_{a\in\mathcal{E}} \sup_{\theta\in\mathcal{T}_{s,T}} \mathbb{E}\left[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) \mathrm{d}r + v\left(\theta, X_\theta^{s,x,a}\right)\right], \quad s \in [0,T),\, x \in \mathbb{R}^d,$$

where

$$v(s,x) = \sup_{a\in\mathcal{E}} \mathbb{E}\left[\int_s^T h\left(r, X_r^{s,x,a}, a_r\right) \mathrm{d}r + j\left(X_T^{s,x,a}\right)\right].$$

We refer to (3.109) and Theorem 3.27 in Section 3.5 for the complete assertion. Here $\mathcal{E}$ [resp., $\mathcal{T}_{s,T}$] is the family of admissible simple controls [resp., stopping times with values in (s,T)]. To the best of our knowledge, the result is new even in the case when there are no large jumps (cfr. [40, 108, 148] focusing on special cases).

Moreover, our formulation of DPP is stronger than the usual one, which assumes the stopping time θ to be fixed, see [145, Remark 3.3.3], [146, Remark 3.2] for classes of SDEs without jumps and Remark 3.9. We also believe that the proof of the second part of the DPP is of independent interest, see, in particular, *Step II* of the proof of Theorem 3.27. Here, we consider controls as in (I.8) where t_i are dyadic numbers and Z_{t_i} are linear combinations of a finite number of elements in the Hilbert bases of suitable L^2-spaces. Such special controls allow to apply a classical measurable selection theorem in [42] (see also Remark 3.8).

Furthermore, we show that the value function v is lower semicontinuous (see Lemma 3.23). Additional properties of v might be investigated and will be the subject of a future research.

Outline of the proof of the sharp stochastic flow. To prove the existence of a sharp stochastic flow for (I.5) (see Theorem 3.1), we first consider the case $f \equiv 0$ (SDEs without large–jumps component).

In this case, the result can be deduced from a stronger one (see Theorem 3.2), which shows that the solution of (I.5) can be obtained employing a càdlàg stochastic process $Z = (Z_s)$ with values in $C(\mathbb{R}^d; \mathcal{D}_0)$ (see Section 3.1-3.2 for more details). In particular, for $\mathbb{P}-$a.e. ω,

$$[Z_s(\omega)](x) = X_\cdot^{s,x}(\omega) \in \mathcal{D}_0, \quad x \in \mathbb{R}^d.$$

Here $\mathcal{D}_0$ stands for the non–separable metric space of $\mathbb{R}^d-$valued, càdlàg functions endowed with the uniform norm in $[0,T]$. Indeed, we cannot use the Skorokhod topology J_1 on $\mathcal{D}_0$ to get our results (cfr. Remark 3.3).

In order to prove Theorem 3.2 we employ an extension of a càdlàg criterium in [28], which can be applied to the process Z taking values in the non–separable metric space $C(\mathbb{R}^d;\mathcal{D}_0)$. Such an extension is proved in the appendix (see Appendix 3.B), which also contains additional measure theoretic results that we have not found in the literature (see in particular Appendix 3.A).

The proof of Theorem 3.2 requires also Proposition 3.8, which is a variant of Theorem 3.7, a generalized Garsia–Rodemich–Rumsey type lemma due to [107]. Such proposition and its Corollary 3.9 allow us to estimate integrals like

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le T}\left|\int_s^t\int_{U_0} g\left(X^{s,x}_{r-},r,z\right)\widetilde{N}_p\left(\mathrm{d}r,\mathrm{d}z\right)\right|^{\gamma}\right] \tag{I.9}$$

which are crucial for the proof of Theorem 3.2 given in Subsections 3.2.2-3.2.3.

In Section 3.3 we consider the full SDE (I.5), i.e., the SDE (I.5) including also the large–jumps component determined by the coefficient f. This part is quite involved. The issue is that we cannot follow the standard interlacing procedure, see for instance [106, Section IV.9] and [43, Section 3.2], to preserve our sharp stochastic flow. Specifically, the main difficulty is to maintain the regularity of X with respect to s. To overcome this challenge, we carefully modify the interlacing method using the stochastic flow already obtained in Section 3.2. This also gives formulas for the solution of (I.5) which could be of independent interest (see, e.g., (3.86)).

Part II In this part, we investigate affine stochastic Volterra equations of convolution type with jumps. The focus is, in particular, on semi–explicit exponential affine formulas relying on deterministic Riccati–Volterra equations for the conditional Fourier–Laplace transforms of the solutions, called *affine Volterra processes with jumps*. These processes are then employed in a stochastic volatility model, named the *rough Hawkes Heston model*, which is applied with highly satisfactory results to the problem of the joint calibration of options on S&P 500 and VIX.

Chapter 4 contains the findings of the paper [37] and focuses on affine stochastic processes, which constitute unquestionably the most popular multi–factor framework to model rich and flexible stochastic dependence structures. Semi–explicit formulas for the Fourier–Laplace transform of affine processes make them numerically tractable, as Fourier transform–based methods can be used to perform fast calculations.

We recall that a conservative regular affine process $X=(X_t)_{t\ge0}$ with state space $E\subset\mathbb{R}^m$ is a stochastically continuous conservative Markov process having sufficiently regular Fourier–Laplace transforms given by exponential affine formulas in the initial state X_0. Such a process X can also be seen as a special semimartingale whose semimartingale characteristics (B,C,ν), with respect to the "truncation function" $h(\xi)=\xi$, are of the form

$$B_t=\int_0^t b(X_s)\,\mathrm{d}s,\quad C_t=\int_0^t a(X_s)\,\mathrm{d}s,\quad \nu(\mathrm{d}t,\mathrm{d}\xi)=\eta(X_t,\mathrm{d}\xi)\,\mathrm{d}t, \tag{I.10}$$

where, for every $x \in E$,

$$b(x) = b_0 + \sum_{k=1}^{m} x_k b_k, \quad a(x) = A_0 + \sum_{k=1}^{m} x_k A_k, \quad \eta(x, \mathrm{d}\xi) = \nu_0(\mathrm{d}\xi) + \sum_{k=1}^{m} x_k \nu_k(\mathrm{d}\xi). \tag{I.11}$$

In (I.11) we take $A_k \in \mathbb{R}^{m\times m}$, $b_k \in \mathbb{R}^m$, and $\nu_k(\mathrm{d}\xi)$ signed measures on $\mathbb{R}^m$ such that $\nu_k(\{0\}) = 0$ and $\int_{\mathbb{R}^m} |\xi|^2 |\nu_k|(\mathrm{d}\xi) < \infty$, for every $k = 0, \dots, m$. Additional conditions on the parameters A_k, b_k and ν_k have to imposed in order to guarantee existence and invariance results depending on the state space E; see, for instance, [74] for $E = \mathbb{R}^k_+ \times \mathbb{R}^l$ and [56] for E equal to the space of positive semidefinite matrices. The conditional Fourier–Laplace transform of the affine process X is given by

$$\mathbb{E}\left[\exp\left(\int_0^T f(T-s)^\top X_s \,\mathrm{d}s \right) \middle| \mathcal{F}_t \right] = \exp\left(\phi(T-t) + \int_0^t f(T-s)^\top X_s \,\mathrm{d}s + \psi(T-t)^\top X_t \right), \tag{I.12}$$

with ψ a $\mathbb{C}^m-$valued function that solves the deterministic Riccati equation

$$\psi(t) = \int_0^t \mathcal{R}\left(s, \psi(s)\right) \mathrm{d}s, \tag{I.13}$$

where

$$\mathcal{R}_k\left(s, z\right) = f_k\left(s\right) + \frac{1}{2} z^\top A_k z + z^\top b_k + \int_{\mathbb{R}^m} \left(\mathrm{e}^{z^\top \xi} - 1 - z^\top \xi \right) \nu_k(\mathrm{d}\xi), \quad k = 1, \dots, m, \tag{I.14}$$

and ϕ is the $\mathbb{C}-$valued function

$$\phi(t) = \int_0^t \left(\psi(s)^\top b_0 + \frac{1}{2} \psi(s)^\top A_0 \psi(s) + \int_{\mathbb{R}^m} \left(\mathrm{e}^{\psi(s)^\top \xi} - 1 - \psi(s)^\top \xi \right) \nu_0(\mathrm{d}\xi) \right) \mathrm{d}s. \tag{I.15}$$

The identity (I.12) is only valid under additional hypotheses on the $\mathbb{C}^m-$valued function f and $t, T \geq 0$ that imply appropriate conditions on the functions ϕ and ψ. [2]

The theory of affine processes was recently extended in [8, 92] to the framework of stochastic Volterra equations with continuous trajectories, where in general the semimartingale and Markov properties do not hold. These so–called affine Volterra processes overcome modeling shortcomings of affine processes because they may posses path–dependent features which introduce memory structures into the models. Furthermore, they can have trajectories whose Hölder's regularity is different from the Hölder's regularity of the paths of Brownian motion. More specifically, singular kernels yield rough processes in the spirit of [21, 79, 90]. The goal of this chapter is to extend the results in [8, 92] by considering general affine stochastic Volterra equations with jumps. This extension is not straightforward because the jump structure together with possible singularities of the kernel may induce explosions of the trajectories.

Our study can be motivated by financial models for stock volatility, in particular by the observation in [171] that a complete description of volatility should take into account both path roughness and jumps. We also refer to [176] for an interesting discussion on the topic. In this chapter, however, we concentrate on the mathematical properties of this family of processes and we address their possible applications in Chapter 5 (see also [39]).

[2] One of these conditions could be, for instance, the boundedness of the right term in (I.12). On a related subject, we also refer to [120], where the authors analyze the possible explosions of the associated Riccati equation (I.13).

We now summarize the framework and the main results of Chapter 4. Suppose that X is a predictable solution of a stochastic Volterra equation of the form

$$X_t = g_0(t) + \int_0^t K(t-s)\,\mathrm{d}Z_s, \quad \mathbb{P}\otimes \mathrm{d}t\text{–a.e.} \tag{I.16}$$

It is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ where the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions and has trajectories in $L^1_{\mathrm{loc}}(\mathbb{R}_+; E)$, for some state space $E \subset \mathbb{R}^m$. In (I.16) we take $g_0 \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^m)$, $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^{m\times d})$ a matrix–valued kernel, and Z a $d-$dimensional semimartingale whose characteristics depend on X. In order to have an affine structure, we suppose that Z has characteristics of the form (I.10)-(I.11), with $A_k \in \mathbb{R}^{d\times d}$, $b_k \in \mathbb{R}^d$, and ν_k signed measures on $\mathbb{R}^d$ such that $\nu_k(\{0\}) = 0$ and $\int_{\mathbb{R}^d} |\xi|^2 |\nu_k|(\mathrm{d}\xi) < \infty$. In this case, we call X an *affine Volterra process with jumps.* When $E = \mathbb{R}^m$, existence of weak solutions to (I.16) with trajectories in $L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^m)$ has been established in [3, Theorem 1.2]. For $E = \mathbb{R}_+$ and for a Volterra CIR–type of process with positive jumps, results in this direction can be found in [2, Theorem 2.13].

Fix $T > 0$ and $f \in C(\mathbb{R}_+; \mathbb{C}^m)$. By analogy with (I.13), assume that $\psi \in C(\mathbb{R}_+; \mathbb{C}^d)$ solves the deterministic Riccati–Volterra equation

$$\psi(t)^\top = \int_0^t \mathcal{R}(s, \psi(s))^\top K(t-s)\,\mathrm{d}s, \tag{I.17}$$

with $\mathcal{R}$ as in (I.14), and let ϕ be given by (I.15).
Our first main result is Theorem 4.5, which is a generalization of [8, Theorem 4.3] and provides a semi–explicit formula for the conditional Fourier–Laplace transform of X. This theorem shows that, in the above–mentioned framework, if we define

$$M_t = \exp\left(\phi(T-t) + \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_t^T \mathcal{R}(T-s, \psi(T-s))^\top g_t(s)\,\mathrm{d}s\right), \tag{I.18}$$

where $(g_t(\cdot))_{t\geq 0}$ denotes the adjusted forward process[3]

$$g_t(s) = g_0(s) + \int_0^t K(s-r)\,\mathrm{d}Z_r, \quad s > t, \tag{I.19}$$

then M is a local martingale. Moreover, if M is a martingale then one has the exponential affine formula

$$\mathbb{E}\left[\exp\left(\int_0^T f(T-s)^\top X_s\,\mathrm{d}s\right)\middle|\mathcal{F}_t\right] = M_t. \tag{I.20}$$

As a consequence, under these conditions, uniqueness in law holds for the stochastic Volterra equation (I.16).

In Subsection 4.2.1, we discuss how the arguments used to prove Theorem 4.5 can be adapted to obtain exponential affine formulas for the conditional Fourier–Laplace transforms of the marginal distributions of the solution process X and the semimartingale Z in (I.16). Notably, Theorem 4.6, which investigates the marginals of Z, can be applied to obtain original expressions for the conditional transforms of a multi–dimensional Hawkes process, see Corollary 4.7 and Remark 4.6.

[3] This adjusted forward process was also used in [4] and [121] to elucidate the affine structure of affine Volterra processes with continuous trajectories.

The second main result of Chapter 4 is Theorem 4.10, which establishes, under the assumption $m = d$ and additional conditions on the kernel K, an alternative formula for the local martingale M_t in (I.18) in terms of $(X_s)_{s\leq t}$ and Z_t only, namely

$$\log(M_t) = \phi(T-t) + \int_0^t f(T-s)^\top X_s \,\mathrm{d}s + \int_0^{T-t} \mathcal{R}(s, \psi(s))^\top g_0(T-s)\,\mathrm{d}s \\ + \psi(T-t)^\top Z_t + \Big(\pi_{T-t}^\top * (X - g_0)\Big)(t), \quad \text{(I.21)}$$

with ϕ as in (I.15) and $\pi_h \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^d)$, $h > 0$, a deterministic function that depends on K and ψ. This expression derives from a similar one for the adjusted forward process (I.19), shown in Lemma 4.9. The identity (I.21) can be used to show that (I.12) is a particular instance of (I.20) when g_0 is constant, and K is constant and equal to the identity matrix.

In Section 4.4, see Theorems 4.13 and 4.14, using our first two main results, we give a complete proof of the exponential affine formula (I.20) in the particular case $m = d = 1$, $E = \mathbb{R}_+$ and for a Volterra CIR–type process with positive jumps. The argument hinges on a novel comparison result between solutions of Riccati–Volterra equations, namely between a solution of (I.17) and a solution of an analogous equation in which the functions ψ and $\mathcal{R}$ are substituted with the corresponding real parts. This comparison result, together with the *affine with respect to the past formula* (I.21) of Theorem 4.10, yields the desired conclusion, because we can bound the complex–valued local martingale M (I.18) of Theorem 4.5 with a real–valued martingale.
It is also important at this point to mention [59], where the authors construct infinite–dimensional Markovian lifts of affine Volterra processes, possibly with jumps, and study affine transform formulas for these lifts. Such formulas are closely related to those of Chapter 4 within a one–dimensional setting, although in [59] the focus is on the marginal distributions of X (see Remark 4.7 in Subsection 4.2.1 for a precise comparison). The novelty of our work stems from the approach that we propose, which is inspired by the arguments in [8] and dispenses with the abstract infinite–dimensional theory in [59]. In addition, we carry out a complete analysis of the associated Riccati–Volterra equations, see Section 4.4. Furthermore, we obtain the *affine with respect to the past formula* (I.21), as well as new formulas for the semimartingale Z which can be applied to derive novel expressions for the conditional Fourier–Laplace transforms of multi–dimensional marked Hawkes processes, see Theorem 4.6 and Corollary 4.7. These last two points also distinguish our study from [2].

Chapter 5 presents the results of the article [39]. In particular, it introduces a stochastic volatility model, called the *rough Hawkes Heston model*, where the instantaneous variance σ^2 satisfies a stochastic affine Volterra equation of convolution type with jumps as in (I.16). A parsimonious specification of this model is calibrated using S&P 500 and VIX options data, demonstrating its remarkable precision in simultaneously replicating the behavior of implied volatility smiles for both S&P 500 and VIX options.

We now give a brief literature review to explain the choice of our framework. The Black–Scholes model, where volatility is constant, and more generally classical local volatility models, where volatility is a function of time and spot asset prices, fail to reproduce the dynamics of implied volatility smiles of options written on the underlying asset. To overcome this limitation, multiple stochastic, stochastic–local, and path–dependent volatility models have been developed and studied in recent years. The complexity of volatility modeling, however, has increased with the significant growth over time of markets on volatility indices, such as the VIX. The rise in popularity of these markets is explained, in part, by their relevance to protect portfolios [156] and has driven the standardization of contingent claims

written on the volatility indices themselves.
Volatility index markets have very unique features that, for VIX futures and exchange–traded products, are investigated in [14]. The intricacy of these markets is also exemplified by the difficulty to jointly model the behavior of the implied volatility smiles of vanilla options written on the underlying and its volatility index, see for instance [9, 141, 142, 155]. This longstanding puzzle is known as the S&P 500 (SPX)/VIX calibration puzzle, see [98, 99]. A growing body of literature explains the difficulty arguing that "the state-of-the-art stochastic volatility models in the literature cannot capture the S&P 500 and VIX option prices simultaneously", see [168]. As pointed out in [99], "all the attempts at solving the joint SPX/VIX smile calibration problem only produced imperfect, approximate fits." The problem is that usual stochastic models either fail to reproduce one or both shapes of the implied volatility for S&P 500 and VIX options or, when both the shapes are coherent, the implied volatility levels are incorrect. This modeling challenge has inspired the introduction of more sophisticated models, e.g. [6, 7, 57, 91, 99, 101, 159], that incorporate new features to the joint dynamics of the underlying and the volatility. In this chapter, we tackle the challenge by proposing a tractable affine model with rough volatility and volatility jumps, see (I.22) below. These jumps not only cluster but also exhibit an opposite direction compared to the jumps of the underlying prices, although they occur at the same time.

The dynamics of the VIX volatility index are highly complex. In particular, they exhibit large and systematically positive variations over very short periods, with a tendency to form clusters of spikes during difficult periods like the 2008 financial crisis and the beginning of the COVID 19 pandemic in 2020. This is accompanied by very long periods without any large fluctuation and a less important mean reversion speed. These observations are in line with an increasing number of studies that indicate the presence of jumps in the volatility [73, 171], on the underlying [19], and the fact these jumps are common to the volatility and underlying [166].

Access to high frequency data has improved our understanding of the microstructure of financial markets and the effects on volatility. In particular, recent studies indicate that non–Markovian models with rough volatility trajectories might be appropriate to better capture long time dependencies due to meta orders and the large contribution of automatic orders. This is examined in [53], which provides a general analysis of order–driven markets, in [51], which elucidates the memory–features of volatility, and in [76, 90], which give a micro–structural justification to the newly developed rough volatility models.

From a modeling point of view, affine models provide a convenient framework because they are flexible and, thanks to semi–explicit formulas for the Fourier–Laplace transform, fast computations can be performed using Fourier–based techniques [74, 75, 81]. The most popular affine stochastic volatility model is the Heston model [104], where the spot variance is a square–root mean–reverting CIR (Cox–Ingersoll–Ross [55]) process. This model is able to reproduce some stylized features like the mean–reverting property of the volatility and the leverage effect. It is, however, unable to reproduce other phenomena such as extreme paths of volatility during crisis periods (even for large values of the volatility of volatility parameter) and the at–the–money (ATM) skews of underlying options' implied volatility simultaneously for short and long maturities. These limitations, and the micro–structural behavior of markets described in the previous paragraph, motivated the introduction of the rough Heston model [78, 79]. The rough Heston model is tractable as it belongs to the class of affine Volterra models [8], and semi–explicit formulas for the Fourier–Laplace transform are still available. Unfortunately, this model cannot reproduce the features of options written on the volatility index and the underlying simultaneously (see Subsection 5.5.1).

In an attempt to improve the performance of the rough Heston model, we propose an extension of it which incorporates a self–exciting jump component. More precisely, we denote by $X = (X_t)_{t\geq 0}$ the log returns of the SPX index and by $\sigma^2 = (\sigma_t^2)_{t\geq 0}$ the spot variance. In order to model the joint behavior of SPX and VIX markets, under a risk neutral probability measure $\mathbb{Q}$ we consider the dynamics

$$\begin{cases} \mathrm{d}X_t = -\left(\frac{1}{2} + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 + \Lambda z\right) \nu\left(\mathrm{d}z\right)\right) \sigma_t^2 \mathrm{d}t \\ \qquad\qquad\qquad + \sigma_t \left(\sqrt{1-\rho^2}\, \mathrm{d}W_{1,t} + \rho\, \mathrm{d}W_{2,t}\right) - \Lambda \int_{\mathbb{R}_+} z\, \tilde{\mu}\left(\mathrm{d}t, \mathrm{d}z\right), \\ \sigma^2 \quad = g_0 + K * \mathrm{d}Z, \quad \mathbb{Q} \otimes \mathrm{d}t - \text{a.e.}, \end{cases} \tag{I.22}$$

where $K(t) = t^{\alpha-1}/\Gamma(\alpha)$, $\alpha \in (1/2, 1]$, and Z is the following semimartingale having jump measure $\mu(\mathrm{d}t, \mathrm{d}z)$ and compensator $\sigma_t^2 \mathrm{d}t \otimes \nu(\mathrm{d}z)$:

$$\mathrm{d}Z_t = b\, \sigma_t^2 \mathrm{d}t + \sqrt{c}\, \sigma_t\, \mathrm{d}W_{2,t} + \int_{\mathbb{R}_+} z\, \left(\mu\left(\mathrm{d}t, \mathrm{d}z\right) - \sigma_t^2 \mathrm{d}t \otimes \nu\left(\mathrm{d}z\right)\right), \quad Z_0 = 0.$$

Here $b \in \mathbb{R}$, $c > 0$, $\Lambda \geq 0$, $\rho \in [-1, 1]$, W_1 and W_2 are two independent Brownian motions, ν is a nonnegative measure on $\mathbb{R}_+$ such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z) < \infty$, and g_0 is a function representing the initial spot variance curve.

Consistently with empirical evidence discussed in the previous paragraphs, in (I.22) we add two specific features to the usual Heston model. First, we incorporate rough volatility by adding a power kernel proportional to $t^{\alpha-1}$, with $\alpha \in (1/2, 1]$, to the dynamics of the spot variance. Second, we postulate common jumps for the volatility and the underlying with a negative leverage, namely $-\Lambda$. The presence of jumps in both underlying and variance helps to reproduce a skewed implied volatility for vanilla options as in the Barndorff–Nielsen and Shephard model [16, 17]. Inspired by the Hawkes framework, taking into account jump–clustering and endogeneity of financial markets, we model the spot variance to be proportional to the intensity process of the jump component appearing in the dynamics of the spot variance itself and the log returns. For these reasons, we name our model (I.22) the *rough Hawkes Heston model*, see Section 5.1.

To keep mathematical and numerical tractability, we choose an affine setting. As a result, our model belongs to the class of affine Volterra processes [8], which has been recently extended to jump processes in [37, 58, 59] (see also Chapter 4). In particular, the Fourier–Laplace transform of the log returns and the square of the volatility index can be computed explicitly in terms of deterministic Riccati–Volterra equations, see Theorems 5.3 and 5.10. We approximate the solutions of these Riccati–Volterra equations via a multi–factor scheme as in [5], see Theorem 5.11. We leave for future study the implementation and analysis in our framework of other methods such as the Adams method [71, 72], asymptotic formulas inspired by forest expansions in the spirit of [10], and hybrid approximation techniques for Volterra equations similar to those in [45]. It might be also interesting to investigate an adaptation to Riccati–Volterra equations of the hybrid multi–factor approach proposed in [160] for the discretization of stochastic Volterra equations.

The affine property is an advantage of our modeling approach compared to other recent models proposed to solve the SPX/VIX calibration problem, such as [6, 57, 91, 102, 161], where pricing is done via Monte Carlo or machine learning techniques. In addition, our affine framework is convenient because variance swap rates and the square VIX index have explicit affine relations to the forward curve, see Corollary 5.7 and Remark 5.3. This is a generalization, to the affine Volterra setting, of the affine relation already pointed out in [118] within the classical exponential affine framework and empirically confirmed in [132].

Our approach shares similarities with the study in [26], which demonstrates that an exponential law for the jump size can capture upward VIX implied volatility within a Hawkes framework with an exponential kernel. Alternatively, other studies such as [114, 138] have explored the inclusion of more general jump measures. The advantage of the rough Hawkes Heston model over the aforementioned works is the possibility to achieve a joint calibration of the SPX/VIX smiles, while utilizing a simple and parsimonious specification for the jump distribution, namely an exponential law (see Section 5.5).

Previous literature on jump–diffusion models focusing on the evolution of S&P 500 and VIX proposes either high–dimensional models [54, 141, 166], or models based on hidden Markov chains [94, 143]. These models require a large number of parameters and suffer from the lack of interpretability of the random factors. Our approach to model the joint SPX/VIX dynamics is different. As in [26], we keep the number of parameters low by assuming that the jump intensity is proportional to the variance process itself, and jumps are common to the volatility and underlying with opposite signs. The main new ingredient of our model, compared to [26], is the addition of a Brownian component and a power kernel to the variance process. This generates by construction a jump clustering effect and takes into account related findings in the rough volatility literature [11, 12, 21, 24, 76, 79, 87, 90, 91, 131]. In particular, our model is consistent with the so–called *Zumbach effect*. Indeed, we suggest an extension of the rough Heston model, which reproduces the Zumbach effect according to [77].

The rough Hawkes Heston model is able to reconcile the shapes and levels of the S&P 500 and VIX volatility smiles. An important role is played by the parameter α characterizing the kernel, see Section 5.6. As is the case for other rough volatility models, this parameter controls the explosion rate of the term structure of ATM skews for SPX option smiles as maturity goes to zero. We show that when α is near to $1/2$, the rate of explosion is in the range $[0.5, 0.6]$. This is consistent with similar findings in the rough volatility literature [12, 21, 24, 76, 87, 90, 91]. In addition, in our framework, the parameter α plays a crucial role because it controls the level of the implied volatility of VIX options for short maturities. We observe that, as α approaches $1/2$, the levels of S&P 500 and VIX smiles are coherent.

To summarize, the model (I.22) that we propose in Chapter 5 shares many features with other existing models. These features are mainly: rough volatility [21, 76, 79, 87, 90, 91], jumps [16, 17, 19, 54, 97, 141, 166], the Hawkes/branching character of volatility [26, 41, 105], and the affine structure [8, 37, 74, 75, 81, 118, 114]. Consequently, we take advantage of the low regularity and memory features of rough volatility models, the large fluctuation of jumps, the clusters of Hawkes processes and the explicit Fourier–Laplace transform of the affine setup. The specification that we adopt for the joint SPX/VIX calibration is parsimonious with only five evolution–related parameters. Moreover, all the parameters have a financial interpretation. The parameter α in the kernel controls the decay of the volatility memory, SPX ATM skews and the level of VIX smiles. In addition, we have the classical parameters controlling the volatility mean reversion speed and the volatility of volatility, as well as two parameters related to the leverage effect. The latter specify the correlation between Brownian motions and between the jumps in the asset and its volatility. Despite its robustness, the rough Hawkes Heston stochastic volatility model captures remarkably well the implied volatility surfaces of S&P 500 and VIX at the same time, as demonstrated by the numerical experiments in Section 5.5.

Part III In the conclusive part of this dissertation, we investigate a theoretical connection between the subjects studied in Parts I and II. This is done in Chapter 6 by analyzing the stochastic flow and the path–dependent Kolmogorov equation associated with an unconventional class of stochastic partial differential equations (SPDEs). These SPDEs arise from finite–dimensional stochastic Volterra equations driven by additive fractional Brownian motions of Riemann–Liouville type. The results of Chapter 6 are contained in the paper [36].

More precisely, in Chapter 6 we consider the stochastic differential equation in $\mathbb{R}^d$ with additive noise

$$X_t = x_0 + \int_0^t k_1(t-s)\, b(s, X_s)\, \mathrm{d}s + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\, \mathrm{d}W_s, \tag{I.23}$$

where $x_0 \in \mathbb{R}^d$, $\alpha \in (\frac{1}{2}, 1)$, $W = (W_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^d$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a measurable vector field and k_1 is a locally square–integrable, $\mathbb{R}-$valued kernel that is continuous in $(0, \infty)$. This equation belongs to the class of stochastic Volterra equations (of convolution type), which is characterized by a wide and continuously expanding body of literature, see for instance [3, 8, 37, 103, 136, 152], and also Chapter 4. The additive noise driving the SDE (I.23) is a fractional Brownian motion (henceforth, fBm) of Riemann–Liouville type, with Hurst parameter $H = \alpha - \frac{1}{2} \in (0, \frac{1}{2})$. Our motivation for studying this random perturbation stems from its relevance in mathematical finance, particularly in the field of rough volatility models, see [39, 79, 91] and also Chapter 5. However, the theory that we develop in this chapter encompasses also the case $\alpha \in [1, \frac{3}{2})$, corresponding to a fBM with Hurst parameter $H \in [\frac{1}{2}, 1)$, exhibiting smoother trajectories and longer memory.

Inspired by [86], our aim is to insert the Volterra SDE (I.23) in a class of infinite–dimensional SDEs in a separable Hilbert space $(H, \langle \cdot, \cdot \rangle_H)$ of the form

$$w_t = \phi + \int_0^t B(s, w_s)\, \mathrm{d}s + \int_0^t \sigma(s)\, \mathrm{d}W_s, \quad \phi \in H, \tag{I.24}$$

where $\sigma \colon [0,T] \to \mathcal{L}(\mathbb{R}^d; H)$ and $B \colon [0,T] \times H \to H$. In order to achieve this objective, we need to consider a drift B with an unconventional structure. This motivates the study carried out in this chapter of a novel class of SPDEs and of the associated stochastic flow's regularity. Notably, these SPDEs require an extension of the drift operator and its Fréchet differentials, resulting in an abstract formulation of the problem that, to the best of our knowledge, is not covered by the existing literature. Given $\Phi \colon H \to \mathbb{R}$, we then study the following backward Kolmogorov equation associated with (I.24):

$$\begin{cases} \partial_t u(t,\phi) + \mathcal{A}_t u(t,\phi) = 0, \quad t \in [0,T], \phi \in H, \\ u(T,\phi) = \Phi(\phi), \end{cases}$$

which will be interpreted in integral form, see (6.67). Here $\mathcal{A}_t$, the time–dependent infinitesimal generator, is given by

$$\mathcal{A}_t u(t,\phi) = \frac{1}{2} \mathrm{Tr}\left(D^2 u(t,\phi)\, \sigma(t)\sigma(t)^*\right) + \langle B(t,\phi), \nabla u(t,\phi) \rangle_H .$$

As in [66, Chapter 9], the approach that we adopt for the existence of classical solutions of the Kolmogorov equation is based on a careful analysis of (I.24) and on the formula

$$u(t,\phi) = \mathbb{E}\left[\Phi\left(w_T^{t,\phi}\right)\right], \tag{I.25}$$

where $w_t^{t_0,\phi}$, $t \in [t_0, T]$, is the solution of an analogue of (I.24) starting at time t_0 instead of 0.

It is worth noting that we use classical tools of infinite–dimensional calculus, such as the Frechét derivative, when analyzing the Kolmogorov equation. This is a novelty compared to other studies addressing path–dependent PDEs related to Volterra SDEs, particularly [174] (see also [18] for a similar subject). In a sense, then, we unify the study of stochastic Volterra equations and fBm of Riemann–Liouville type to other infinite–dimensional systems. However, the assumptions imposed on B

are not entirely classical, resulting in an innovative abstract formulation of the problem. Consequently, the analysis developed here is only analogous to the classical one, not included into it.

A more direct approach to the Kolmogorov equation would be also of great interest for two reasons. Firstly, it would contribute to complete the comparison with the classical theory developed for other classes of problems, see [66]. Secondly, it could be used to study regularization-by-noise phenomena for SDEs driven by fractional Brownian motion, which are investigated in literature using different techniques, see, e.g., [88, 89, 103, 136, 139]. In fact, studying the Kolmogorov equation in mild form might allow to prove weak uniqueness of solutions of the underlying SDEs when the drift is not smooth, see [62, 63, 82, 169, 173]. In an attempt to develop such a direct approach, we have identified obstructions that we report in Section 6.4, so this problem remains open.

Part I

Chapter 1

Smoothing effect and derivative formulas for OU processes driven by subordinated cylindrical Brownian noises

In this chapter, we study the concept of cylindrical Wiener process subordinated to an $\alpha-$stable Lévy process, with $\alpha \in (0,1)$, in an infinite–dimensional, separable Hilbert space. We then investigate regularization-by-noise results, derivative formulas and gradient estimates for the Markov transition semigroup of OU processes driven by this random perturbation.

1.1 The framework

Let $(H, \langle\cdot,\cdot\rangle)$ be a separable Hilbert space and $(e_n)_n$ be a complete orthonormal system. We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and introduce a sequence of independent Brownian motions $(\beta^n)_n$ on it. Let $L = (L_t)_t$ be a strictly α–stable subordinator, i.e., an increasing Lévy process where the distribution of L_1 is characterized by

$$\mathbb{E}\left[e^{iuL_1}\right] = \exp\left\{-\bar{c}\,|u|^{\alpha}\left(1 - i\tan\frac{\pi\alpha}{2}\operatorname{sign}(u)\right)\right\}, \quad u \in \mathbb{R}. \tag{1.1}$$

Here $\bar{c} > 0$ and $\alpha \in (0,1)$. The Laplace transform of L_1 is given by

$$\mathbb{E}\left[e^{-uL_1}\right] = e^{-c'u^{\alpha}}, \quad u \geq 0, \tag{1.2}$$

with c' a constant depending on $\bar{c}$ (for an expression of c' we refer to [164, Example 24.12], but it is of no use in our work). Assuming L to be independent from $(\beta^n)_n$, [164, Theorem 30.1] implies that the subordinated Brownian motions $(\beta^n_L)_n$ are real–valued Lévy processes.

Denoting by $\mathcal{N}$ the family of $\mathcal{F}$–negligible sets, we introduce the augmented σ–algebra $\mathcal{F}^L := \sigma\left(\mathcal{F}^L_0 \cup \mathcal{N}\right)$, where $\mathcal{F}^L_0$ is the natural σ–algebra generated by the subordinator. Analogously, we consider the augmented σ–algebras $\mathcal{F}^{\beta^n}$ generated by the Brownian motions. Thanks to the hypotheses of independence, we have that $\mathcal{F}^L$, $\left(\mathcal{F}^{\beta^n}\right)_n$ are mutually independent.
In our context, it is natural to deal with different filtrations. Specifically, for every $n \in \mathbb{N}$ let $\mathbb{F}^n = \left(\mathcal{F}^{\beta^n}_t\right)_t$ be the minimal (i.e., smallest) augmented filtration generated by β^n, that is, $\mathcal{F}^{\beta^n}_t := \sigma\big(\big(\mathcal{F}^{\beta^n}_0\big)_t \cup \mathcal{N}\big)$ for every $t \geq 0$, where $\big(\big(\mathcal{F}^{\beta^n}_0\big)_t\big)_t$ is the natural filtration of the process β^n. According to [153, Theorem

I.31], $\mathbb{F}^n$ satisfies the usual hypotheses. Then we construct a complete filtration associated with the subordinated Brownian motions. It is denoted by $\mathbb{F}_L = (\mathcal{F}_t)_t$, where we define

$$\mathcal{F}_t \coloneqq \sigma\left(\bigcup_{n\in\mathbb{N}} \mathcal{F}_t^{\beta_L^n}\right), \quad t \geq 0,$$

with $\mathbb{F}_L^n = \left(\mathcal{F}_t^{\beta_L^n}\right)_t$ being the minimal augmented filtration associated with β_L^n.

Remark 1.1. *In the finite–dimensional case, we denote by $W_L^N = \left(W_{L_t}^N\right)_t$ the subordinated, $\mathbb{R}^N$–valued Brownian motion, meaning that*

$$W_{L_t}^N = \begin{bmatrix}\beta_{L_t}^1 & \cdots & \beta_{L_t}^N\end{bmatrix}^T, \quad t \geq 0.$$

By [164, Theorem 30.1], W_L^N is an $\mathbb{R}^N$–valued Lévy process, and it is easy to verify that its minimal augmented filtration $\left(\mathcal{F}_t^{W_L^N}\right)_t$ coincides with $\mathbb{F}_L$. This fact shows that the construction that we have carried out for $\mathbb{F}_L$ is natural.

Using the notation we have just introduced, in the general case the σ–algebras constituting $\mathbb{F}_L$ can be expressed as follows:

$$\mathcal{F}_t = \sigma\left(\bigcup_{N\in\mathbb{N}} \mathcal{F}_t^{W_L^N}\right), \quad t \geq 0.$$

1.1.1 The Subordinated Cylindrical Wiener Process

The aim of this subsection is to give a rigorous meaning to the formal notation $W_{L_t} = \sum_{n=1}^\infty \beta_{L_t}^n e_n$.

First, fix $h \in H$, $t > 0$ and notice that the series $\sum_{n=1}^\infty \beta_{L_t}^n \langle h, e_n\rangle$ converges in distribution. Indeed, even if the random variables $\left(\beta_{L_t}^n\right)_{n\in\mathbb{N}}$ are not independent due to the presence of the subordinator, we can still exploit the mutual independence of the σ–algebras $\left(\mathcal{F}^{\beta^n}\right)_n$ by conditioning with respect to $\mathcal{F}^L$, which in turn is independent from the previous ones. In order to do so, we use the law of total expectation together with (1.2) to get, for every $u \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left\{iu\sum_{n=1}^N \beta_{L_t}^n \langle h, e_n\rangle\right\}\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left\{iu\sum_{n=1}^N \beta_r^n \langle h, e_n\rangle\right\}\Bigg|_{r=L_t}\Bigg|\mathcal{F}^L\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\exp\left\{iu\sum_{n=1}^N \beta_r^n \langle h, e_n\rangle\right\}\right]\Bigg|_{r=L_t}\right] = \mathbb{E}\left[\prod_{n=1}^N \exp\left\{-\frac{1}{2}L_t\left|u\right|^2\left|\langle h, e_n\rangle\right|^2\right\}\right] \\
&= \exp\left\{-tc'\frac{1}{2^\alpha}\left|u\right|^{2\alpha}\left(\sum_{n=1}^N\left|\langle h, e_n\rangle\right|^2\right)^\alpha\right\} \underset{N\to\infty}{\longrightarrow} \exp\left\{-t\frac{c'}{2^\alpha}\left\|h\right\|_H^{2\alpha}\left|u\right|^{2\alpha}\right\}.
\end{aligned} \tag{1.3}$$

Hence applying Lévy's continuity theorem we see that the series $\sum_{n=1}^\infty \beta_{L_t}^n \langle h, e_n\rangle$ converges in distribution to a symmetric, 2α–stable random variable. Moreover, for every $n \in \mathbb{N}$, choosing $h = e_n$ and $N > n$ the computations in (1.3) provide the distribution of the Lévy process β_L^n, namely

$$\mathbb{E}\left[e^{iu\beta_{L_t}^n}\right] = \exp\left\{-t\frac{c'}{2^\alpha}\left|u\right|^{2\alpha}\right\}, \quad u \in \mathbb{R}, \text{ for any } t > 0. \tag{1.4}$$

The process $W_L = (W_{L_t})_t$ is a *subordinated cylindrical Wiener process*, but we might also call it *cylindrical, 2α–stable isotropic process.* In fact, for every $N \in \mathbb{N}$ and $t > 0$, if we denote by π_N the projection onto the first N Fourier components and by H_N its range, an argument analogous to the one in (1.3) yields:

$$\mathbb{E}\left[\exp\left\{i\left\langle z, \sum_{n=1}^{N}\beta^n_{L_t}e_n\right\rangle\right\}\right] = \exp\left\{-t\frac{c'}{2^\alpha}\left(\sum_{n=1}^{N}|\langle z, e_n\rangle|^2\right)^\alpha\right\}, \quad z \in H.$$

Hence canonically identifying H_N with $\mathbb{R}^N$, the Galerkin projection $\left(\sum_{n=1}^{N}\beta^n_{L_t}e_n\right)_t$ can be read as an $\mathbb{R}^N$–valued, 2α–stable, isotropic Lévy process.

Secondly, we consider a linear, bounded, nonnegative definite operator $Q : H \to H$ such that e_n is one of its eigenvectors corresponding to the eigenvalue $\sigma_n^2 \geq 0$, $n \in \mathbb{N}$. We study the convergence in probability –on an appropriate space– of the series:

$$\sqrt{Q}W_{L_t} = \sum_{n=1}^{\infty}\sigma_n\beta^n_{L_t}e_n, \quad t > 0.$$

Let us introduce a bounded sequence $(\rho_n)_n$ of strictly positive numbers such that $\sum_{n=1}^{\infty}\rho_n^{2r}\sigma_n^{2r} < \infty$ for some $r \in (0, \alpha)$, and consider the corresponding Hilbert space $(V, \langle\cdot,\cdot\rangle_V)$, where

$$V := \left\{h \in H : \sum_{n=1}^{\infty}\rho_n^{-2}|\langle h, e_n\rangle|^2 < \infty\right\} \quad \text{and} \quad \langle v, w\rangle_V := \sum_{n=1}^{\infty}\rho_n^{-2}\langle v, e_n\rangle\langle w, e_n\rangle, \quad v, w \in V. \tag{1.5}$$

Evidently $V \subset H$ with dense and continuous embedding, therefore using the concept of *Gelfand triple* we can think a generic $h \in H$ as an object in V', namely

$$\langle h, v\rangle_{V',V} = \sum_{n=1}^{\infty}\langle h, e_n\rangle\langle v, e_n\rangle, \quad v \in V.$$

Noticing that $\langle h, \cdot\rangle_{V',V} = \langle\tilde{v}, \cdot\rangle_V$, where $\tilde{v} := \sum_{n=1}^{\infty}\rho_n^2\langle h, e_n\rangle e_n \in V$, we can apply *Riesz representation theorem* to get $\|h\|^2_{V'} = \sum_{n=1}^{\infty}\rho_n^2|\langle h, e_n\rangle|^2$. Now we fix $t > 0$ and show that $\left(\sum_{n=1}^{N}\sigma_n\beta^n_{L_t}e_n\right)_N \subset V'$ is a Cauchy sequence in probability. Indeed, applying *Markov's inequality* and using the fact that the function $\phi(x) = x^r$, $x \geq 0$, is subadditive and strictly increasing as $0 < r < \alpha < 1$, for every $\epsilon > 0$ we get:

$$\begin{aligned}
\mathbb{P}\left(\left\|\sum_{n=p}^{q}\sigma_n\beta^n_{L_t}e_n\right\|_{V'} > \epsilon\right) &\leq \mathbb{P}\left(\phi\left(\left\|\sum_{n=p}^{q}\sigma_n\beta^n_{L_t}e_n\right\|^2_{V'}\right) > \phi\left(\epsilon^2\right)\right)\\
&\leq \frac{1}{\epsilon^{2r}}\mathbb{E}\left[\phi\left(\left\|\sum_{n=p}^{q}\sigma_n\beta^n_{L_t}e_n\right\|^2_{V'}\right)\right] = \epsilon^{-2r}\mathbb{E}\left[\phi\left(\sum_{n=p}^{q}\sigma_n^2\rho_n^2\left|\beta^n_{L_t}\right|^2\right)\right]\\
&\leq \epsilon^{-2r}\sum_{n=p}^{q}\mathbb{E}\left[\sigma_n^{2r}\rho_n^{2r}\left|\beta^n_{L_t}\right|^{2r}\right] = \epsilon^{-2r}\mathbb{E}\left[\left|\beta^1_{L_t}\right|^{2r}\right]\left(\sum_{n=p}^{q}\sigma_n^{2r}\rho_n^{2r}\right)\underset{p,q\to\infty}{\longrightarrow} 0,
\end{aligned}$$

where we use that by construction $\beta^n_{L_t} \sim \beta^1_{L_t}, n \in \mathbb{N}$, and that by (1.4) they all generate a 2α–stable distribution, which has finite moment of order $2r$ (see also Remark 1.2). By completeness, we can conclude the existence of an a.s. unique, V'–valued random variable $\sqrt{Q}W_{L_t}$ such that

$$\sqrt{Q}W_{L_t} = \mathbb{P} - \lim_{N\to\infty} \sum_{n=1}^{N} \sigma_n \beta^n_{L_t} e_n \quad \text{in } V'.$$

Actually such a convergence in probability is true also in the $\mathbb{P}$–a.s. sense, as the following, easy and general lemma proves.

Lemma 1.1. *Let $(X^n)_n$ be a sequence of real–valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and H be a separable Hilbert space admitting $(e_n)_n$ as CONS. If $\sum_{n=1}^{\infty} X^n e_n$ converges in probability, then it converges $\mathbb{P}$–a.s.*

Proof. Let $\tilde{S} := \mathbb{P} - \lim_{N\to\infty} \sum_{n=1}^{N} X^n e_n \colon \Omega \to H$. Obviously

$$\tilde{S}(\omega) = \sum_{n=1}^{\infty} \left\langle \tilde{S}(\omega), e_n \right\rangle e_n = H - \lim_{N\to\infty} \sum_{n=1}^{N} \left\langle \tilde{S}(\omega), e_n \right\rangle e_n, \quad \omega \in \Omega. \tag{1.6}$$

Convergence in measure implies a.s. convergence along a subsequence, hence we have

$$\tilde{S}(\omega) = H - \lim_{k\to\infty} \sum_{n=1}^{N_k} X^n(\omega) e_n \quad \text{for } \mathbb{P} - \text{a.e. } \omega \in \Omega.$$

Therefore, for $\mathbb{P}$–a.e. $\omega \in \Omega$, we see that the Fourier components of $\tilde{S}$ are

$$\left\langle \tilde{S}(\omega), e_{\bar{n}} \right\rangle = \lim_{k\to\infty} \left\langle \sum_{n=1}^{N_k} X^n(\omega) e_n, e_{\bar{n}} \right\rangle = X^{\bar{n}}(\omega) \quad \text{for every } \bar{n} \in \mathbb{N}.$$

Substituting in (1.6) we conclude

$$\tilde{S}(\omega) = H - \lim_{N\to\infty} \sum_{n=1}^{N} X^n(\omega) e_n \quad \text{for } \mathbb{P} - \text{a.e. } \omega \in \Omega,$$

as we stated. ■

Going back to $\sqrt{Q}W_{L_t}$, since $(\rho_n e_n)_n$ is a CONS for the Hilbert space V, Lemma 1.1 allows to write

$$\sqrt{Q}W_{L_t} = \lim_{N\to\infty} \left\langle \sum_{n=1}^{N} \sigma_n \beta^n_{L_t} e_n, \cdot \right\rangle_{V',V} = \lim_{N\to\infty} \sum_{n=1}^{N} \rho_n \sigma_n \beta^n_{L_t} \langle (\rho_n e_n), \cdot \rangle_V \quad \mathbb{P} - \text{a.s.}$$

It then follows that $\langle \sqrt{Q}W_{L_t}, v \rangle_{V',V} = \lim_{N\to\infty} \sum_{n=1}^{N} \sigma_n \beta^n_{L_t} \langle v, e_n \rangle$ for every $v \in V$, $\mathbb{P}$–a.s. Combining this with (1.3), we can see that $\langle \sqrt{Q}W_{L_t}, v \rangle_{V',V}$ has a symmetric, 2α–stable distribution. We collect the previous results in the next theorem.

Theorem 1.2. *1. Given $h \in H$ and $t > 0$, the series $\sum_{n=1}^{\infty} \beta^n_{L_t} \langle h, e_n \rangle$ converges in distribution to a real–valued, symmetric, 2α–stable random variable X_t whose characteristic function is*

$$\mathbb{E}\left[e^{iuX_t}\right] = \exp\left\{-t \frac{c'}{2^\alpha} \|h\|_H^{2\alpha} |u|^{2\alpha}\right\}, \quad u \in \mathbb{R}.$$

2. *Consider a linear, bounded, nonnegative definite operator $Q : H \to H$ such that $(e_n)_n$ is a basis of its eigenvectors corresponding to the eigenvalues $\left(\sigma_n^2\right)_n (\subset \mathbb{R}_+)$. Let $(\rho_n)_n$ be a bounded sequence of strictly positive weights such that $\sum_{n=1}^{\infty} \rho_n^{2r}\sigma_n^{2r} < \infty$ for some $0 < r < \alpha$. Then the corresponding Hilbert space $(V, \langle \cdot, \cdot \rangle_V)$ defined in (1.5) is continuously embedded with density in H and, for every $t > 0$, the random variable $\sqrt{Q}W_{L_t} : \Omega \to V'$ is defined as*

$$\sqrt{Q}W_{L_t} := \lim_{N\to\infty} \sum_{n=1}^{N} \sigma_n \beta_{L_t}^n e_n \quad \mathbb{P}-a.s.$$

In particular, for every $v \in V$,

$$\sum_{n=1}^{N} \beta_{L_t}^n \left\langle \sqrt{Q}v, e_n \right\rangle \underset{N\to\infty}{\longrightarrow} \left\langle \sqrt{Q}W_{L_t}, v \right\rangle_{V',V} \quad \mathbb{P}-a.s.$$

Remark 1.2. *We can state the finiteness of the absolute moment of order $2r$ of the random variable $\beta_{L_t}^1$ without explicitly knowing its distribution, i.e., without using (1.4). In fact, we can proceed as follows:*

$$\mathbb{E}\left[\left|\beta_{L_t}^1\right|^{2r}\right] = \mathbb{E}\left[\mathbb{E}\left[\left|\beta_{L_t}^1\right|^{2r} \middle| \mathcal{F}^L\right]\right] = \frac{2^r}{\sqrt{\pi}}\Gamma\left(\frac{2r+1}{2}\right)\mathbb{E}\left[L_t^r\right] < \infty,$$

recalling that $0 < r < \alpha$. This easy consideration allows to carry out the above procedure to define the noise $\sqrt{Q}W_L$ also for a subordinator L which is not $\alpha-$stable. More precisely, given $r \in (0,1]$ such that $\mathbb{E}[L_t^r] < \infty$ for some (hence, all) $t > 0$ and a bounded sequence $(\rho_n)_n$ of positive real numbers such that $\sum_{n=1}^{\infty} \rho_n^{2r}\sigma_n^{2r} < \infty$, then for every $t > 0$ it is possible to define the random variable $\sqrt{Q}W_{L_t}$ as in Theorem 1.2.

1.1.2 The Stochastic Convolution

Let $A : \mathcal{D}(A) \subset H \to H$ be a linear, selfadjoint, negative definite, unbounded operator that shares with Q a common basis of eigenvectors $(e_n)_n$. We denote by $(-\lambda_n)_n$, with $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le \cdots$ the corresponding eigenvalues, i.e., $Ae_n = -\lambda_n e_n$, $n \in \mathbb{N}$. Recalling that $\alpha \in (0,1)$ has been fixed at the beginning of Section 1.1, it is convenient to introduce the shorthand notation $X \sim \text{stable}(\alpha, \beta, \gamma, \delta)$ to denote a random variable X with characteristic function given by

$$\mathbb{E}\left[e^{iuX}\right] = \exp\left\{-\gamma^\alpha |u|^\alpha \left(1 - i\beta \tan\frac{\pi\alpha}{2}\text{sign}(u)\right) + i\delta u\right\}, \quad u \in \mathbb{R}.$$

where $|\beta| \le 1, \gamma > 0$ and $\delta \in \mathbb{R}$. Hence by (1.4), for every $n \in \mathbb{N}$ the Lévy process β_L^n has random variables distributed as

$$\beta_{L_t}^n \sim \text{stable}\left(2\alpha, 0, \left(t\frac{c'}{2^\alpha}\right)^{1/(2\alpha)}, 0\right), \quad t > 0.$$

We denote by $U^n = (U_t^n)_{t\ge0}$ the OU–process $U_t^n := \int_0^t e^{-\lambda_n(t-s)}\sigma_n \, d\beta_{L_s}^n$, $t \ge 0$: this is the unique (up to evanescence) solution of the one–dimensional stochastic differential equation

$$dU_t^n = -\lambda_n U_t^n dt + \sigma_n \, d\beta_{L_t}^n, \quad U_0^n = 0. \tag{1.7}$$

The processes $(U^n)_n$ are càdlàg and adapted to the filtration $\mathbb{F}_L$, and direct computations (see, e.g., [25, Proposition 3.2]) show that $U_t^n \sim \text{stable}(2\alpha, 0, \gamma_n(t), 0)$, where

$$\gamma_n(t) := \left(\frac{c'}{2^\alpha}\right)^{1/(2\alpha)} \left(\int_0^t e^{-2\alpha\lambda_n(t-s)} \sigma_n^{2\alpha}\, ds\right)^{1/(2\alpha)} = \sigma_n \left(\frac{c'}{2^{\alpha+1}\alpha}\right)^{1/(2\alpha)} \left(\frac{1-e^{-2\alpha\lambda_n t}}{\lambda_n}\right)^{1/(2\alpha)}, \quad t>0,\ n\in\mathbb{N}.$$

In the sequel, we require the next assumption.

Hypothesis 1.1. *For some $r \in (0,\alpha)$,*

$$\sum_{n=1}^{\infty} \frac{\sigma_n^{2r}}{\lambda_n^{r/\alpha}} < \infty.$$

We are now in position to construct the *stochastic convolution* and the corresponding OU–process.

Theorem 1.3. *Assume Hypothesis 1.1. Then, for all $t>0$, the series $\sum_{n=1}^{\infty} U_t^n e_n$ converges $\mathbb{P}-$a.s. to a random variable $\tilde{Z}_{A,Q}(t) = \int_0^t e^{(t-s)A}\sqrt{Q}\, dW_{L_s}$. The resulting process $\tilde{Z}_{A,Q} = \left(\tilde{Z}_{A,Q}(t)\right)_t$ is $\mathbb{F}_L$–adapted and is called* stochastic convolution.

The corresponding OU–process starting at $x\in H$, denoted by $Z^x = (Z_t^x)_t$ and defined by

$$Z_t^x := e^{tA}x + \int_0^t e^{(t-s)A}\sqrt{Q}\, dW_{L_s} = e^{tA}x + \tilde{Z}_{A,Q}(t), \quad t \geq 0,$$

is $\mathbb{F}_L$–adapted and Markovian with homogeneity in time.

Proof. Fix $t>0$. Thanks to the preceding discussion, we know that $U_t^n \sim \gamma_n(t) X$, $n\in\mathbb{N}$, where X is a random variable such that $X \sim \text{stable}(2\alpha,0,1,0)$. Then an application of Markov's inequality entails:

$$\begin{aligned}
\mathbb{P}\left(\left\|\sum_{n=p}^{q} U_t^n e_n\right\|_H > \epsilon\right) &\leq \epsilon^{-2r}\mathbb{E}\left[\phi\left(\left\|\sum_{n=p}^{q} U_t^n e_n\right\|_H^2\right)\right] = \epsilon^{-2r}\mathbb{E}\left[\phi\left(\sum_{n=p}^{q} |U_t^n|^2\right)\right] \\
&\leq \epsilon^{-2r}\sum_{n=p}^{q}\mathbb{E}\left[|U_t^n|^{2r}\right] = \epsilon^{-2r}\mathbb{E}\left[|X|^{2r}\right]\left(\frac{c'}{2^{\alpha+1}\alpha}\right)^{r/\alpha}\left(\sum_{n=p}^{q}\frac{\sigma_n^{2r}}{\lambda_n^{r/\alpha}}\left(1-e^{-2\alpha\lambda_n t}\right)^{r/\alpha}\right) \\
&\leq c(\epsilon)\left(\sum_{n=p}^{q}\frac{\sigma_n^{2r}}{\lambda_n^{r/\alpha}}\right) \underset{p,q\to\infty}{\longrightarrow} 0, \quad \epsilon>0,
\end{aligned}$$

with $c(\epsilon) := \epsilon^{-2r}\mathbb{E}\left[|X|^{2r}\right]\left(\frac{c'}{2^{\alpha+1}\alpha}\right)^{r/\alpha}$ and $\phi(x) = x^r$, as above. Therefore the series converges in probability:

$$\tilde{Z}_{A,Q}(t) = \int_0^t e^{(t-s)A}\sqrt{Q}\, dW_{L_s} := \mathbb{P} - \lim_{N\to\infty}\sum_{n=1}^{N} U_t^n e_n.$$

An application of Lemma 1.1 shows that such convergence is true in the $\mathbb{P}-$a.s. sense, as well. Obviously $\tilde{Z}_{A,Q}$ is an $\mathbb{F}_L$–adapted process, since the space $(\Omega,\mathcal{F},\mathbb{P})$ is complete by hypothesis, $\mathbb{F}_L$ is complete by construction and the one–dimensional OU–processes U^n are $\mathbb{F}_L$–adapted.

Concerning the OU–processes, for every $x \in H$ we can express the random variables of $Z^x = (Z_t^x)_t$ as follows:

$$Z_{t+h}^x \overset{a.s.}{=} e^{hA} Z_t^x + \int_t^{t+h} e^{(t+h-s)A} \sqrt{Q}\, dW_{L_s} = e^{hA} Z_t^x + \sum_{n=1}^{\infty} \left(\int_t^{t+h} e^{-\lambda_n (t+h-s)} \sigma_n \, d\beta_{L_s}^n \right) e_n,$$
$$t, h \geq 0.$$

At this point, the Markovianity of the process is a consequence of the independence of the random variable $\sum_{n=1}^{\infty} \left(\int_t^{t+h} e^{-\lambda_n (t+h-s)} \sigma_n d\beta_{L_s}^n \right) e_n$ from the $\sigma-$algebra $\mathcal{F}_t$. Indeed, such a random variable is measurable with respect to the $\sigma-$algebra

$$\mathcal{G}_t \coloneqq \sigma \left(\bigcup_{n=1}^{\infty} \sigma \left(\sigma \left(\beta_{L_u}^n - \beta_{L_t}^n,\, u \geq t \right) \cup \mathcal{N} \right) \right) = \sigma \left(\bigcup_{N=1}^{\infty} \sigma \left(\sigma \left(W_{L_u}^N - W_{L_t}^N,\, u \geq t \right) \cup \mathcal{N} \right) \right),$$

hence using the property of independence for the increments of a Lévy process one concludes $\mathcal{F}_t \perp\!\!\!\perp \mathcal{G}_t$, as desired. The time homogeneity is obtained by a standard argument relying on the stationarity of the increments of $(\beta_L^n)_n$ and the fact that the coefficients of the one–dimensional SDEs in (1.7) are time–autonomous. The proof is then complete. ■

We close this section with an example which analyzes a common framework in applications (see, e.g., [83]).

Example 1.1. Let $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ be the d–dimensional torus and denote by e_k the functions

$$e_k(x) \coloneqq \sqrt{2} \begin{cases} \cos(2\pi k \cdot x), & k \in \mathbb{Z}_+^d \\ \sin(2\pi k \cdot x), & k \in \mathbb{Z}_-^d \end{cases}, \quad x \in \mathbb{T}^d,$$

where $\mathbb{Z}_+^d \coloneqq \left\{ k \in \mathbb{Z}^d \setminus \{0\} : k_j > 0, \text{ where } j = 1, \ldots, d \text{ is the first index such that } k_j \neq 0 \right\}$ and $\mathbb{Z}_-^d \coloneqq -\mathbb{Z}_+^d$. Then $\left\{ e_k : k \in \mathbb{Z}_0^d \right\}$ constitute a complete orthonormal system for the Hilbert space

$$H = L_0^2 \left(\mathbb{T}^d; \mathbb{R} \right) \coloneqq \left\{ f \in L^2 \left(\mathbb{T}^d; \mathbb{R} \right) : \int_{\mathbb{T}^d} f(x)\, dx = 0 \right\},$$

where of course $\mathbb{Z}_0^d \coloneqq \mathbb{Z}^d \setminus \{0\}$. In particular, for every $f \in H$, we have

$$f = \sum_{k \in \mathbb{Z}_0^d} \hat{f}_k e_k, \quad \hat{f}_k \coloneqq \int_{\mathbb{T}^d} f(x)\, e_k(x)\, dx, \quad k \in \mathbb{Z}_0^d.$$

We first introduce the Sobolev spaces

$$W_0^{\beta,2} \left(\mathbb{T}^d \right) \coloneqq \left\{ f \in H : \sum_{k \in \mathbb{Z}_0^d} |k|^{2\beta} \hat{f}_k^2 < \infty \right\}, \quad \|f\|_{W_0^{\beta,2}}^2 \coloneqq \sum_{k \in \mathbb{Z}_0^d} |k|^{2\beta} \hat{f}_k^2,$$

and then define the linear operator A as follows:

$$A \colon W_0^{2,2} \left(\mathbb{T}^d \right) \to H \quad \text{such that} \quad Af = \Delta f = -(2\pi)^2 \sum_{k \in \mathbb{Z}_0^d} |k|^2 \hat{f}_k e_k.$$

In particular, the eigenvalues of A corresponding to e_k are $-\lambda_k = -(2\pi)^2 |k|^2$, hence A is unbounded and negative definite. Moreover it is selfadjoint, as well. Now we analyze Hypothesis 1.1 for two specifications of the linear, bounded, positive semidefinite operator $Q \colon H \to H$.

- Let $Q = \mathrm{Id}$. Then $\sigma_k = 1$, $k \in \mathbb{Z}_0^d$, and Hypothesis 1.1 reads

$$\frac{1}{(2\pi)^{2r/\alpha}} \sum_{k\in\mathbb{Z}_0^d} \frac{1}{|k|^{2r/\alpha}} < \infty \quad \text{for some } r \in (0,\alpha)\,,$$

which is satisfied if and only if $d = 1$. Hence the stochastic convolution is defined only in dimension $d = 1$.

- Set $Q = Q_\eta = (-\Delta)^{-\eta}$ for $\eta > 0$, the negative fractional power of the Laplacian, defined as an operator $Q_\eta \colon H \to H$ such that

$$Q_\eta f = \frac{1}{(2\pi)^{2\eta}} \sum_{k\in\mathbb{Z}_0^d} \frac{1}{|k|^{2\eta}} \hat{f}_k e_k, \quad f \in H.$$

In this case the convergence of the infinite sum in Hypothesis 1.1 amounts to requiring $\eta > \left(\frac{d}{2r} - \frac{1}{\alpha}\right) \vee 0$. Since r is chosen freely in the interval $(0,\alpha)$, Hypothesis 1.1 is satisfied if and only if

$$\eta > \left(\frac{d-2}{2\alpha}\right) \vee 0. \tag{1.8}$$

This fact can be interpreted as follows: the higher the dimension d, the weaker the effect of the noise on the high Fourier modes needs to be in order to have the well–posedness of the stochastic convolution.

1.2 Smoothing effect of the Markov Transition Semigroup

Let us introduce the Markov transition semigroup $R = (R_t)_t$ associated with the OU–processes $(Z^x)_{x\in H}$, which is given by

$$R_t \phi\,(x) \coloneqq \mathbb{E}\left[\phi\left(Z_t^x\right)\right], \quad x \in H,\, \phi \in \mathcal{B}_b\,(H)\,,\, t \geq 0,$$

where $\mathcal{B}_b\,(H)$ is the space of bounded, real–valued, Borel–measurable functions in H. Evidently, each R_t is linear and bounded from $C_b\,(H)$ into itself and R_0 is the identity. Our aim is to prove that, under suitable conditions, the operator R_t has a smoothing effect for every $t > 0$. Specifically, given a function $\phi \in \mathcal{B}_b\,(H)$, in the case $\alpha \in \left(\frac{1}{2}, 1\right)$ we are going to show that $R_t\phi \in C_b^1\,(H)$ and that the following gradient estimate holds:

$$\sup_{x\in H} \|\nabla R_t \phi\,(x)\|_H \leq \frac{C}{t^\gamma} \|\phi\|_\infty \quad \text{for every } t > 0, \text{ for some } 0 < \gamma < 1,\, C > 0. \tag{1.9}$$

1.2.1 Finite–dimensional case $H = \mathbb{R}^N$

Let $H = \mathbb{R}^N$ and $W^N = \begin{bmatrix}\beta^1 & \cdots & \beta^N\end{bmatrix}^T$. We start by presenting a theorem which allows to obtain an original derivation formula for the semigroup corresponding to the finite–dimensional OU processes $Z_t^\ell\,(x)$. They are defined as the unique, càdlàg solutions of the linear SDEs $dZ_t^\ell\,(x) = AZ_t^\ell\,(x)\,dt + \sqrt{Q}\,dW_{\ell_t}^N$, $Z_0^\ell\,(x) = x$, and can be expressed by the variation of constant formula as follows:

$$Z_t^\ell\,(x) = e^{At}x + \int_0^t e^{A(t-s)}\sqrt{Q}\,dW_{\ell_s}^N, \quad t \geq 0,\, \mathbb{P} - \text{a.s.}, \tag{1.10}$$

where $x \in \mathbb{R}^N$ and $\ell \colon \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing, càdlàg function such that $\ell_0 = 0$ and $\ell_t > 0$ for every positive t: the set of functions with these properties will be denoted by $\mathbb{S}$. Note that, for every $\ell \in \mathbb{S}$, $W^N_\ell = \left(W^N_{\ell_t}\right)_t$ is a càdlàg martingale with respect to the filtration $\left(\mathcal{F}^N_{\ell_t}\right)_t$, where $\left(\mathcal{F}^N_t\right)_t$ is the minimal augmented filtration generated by W^N. Analogously, for every $\ell \in \mathbb{S}$, we introduce the filtrations $\mathbb{F}^n_\ell = \left(\mathcal{F}^{\beta^n}_{\ell_t}\right)_t$, $n \in \mathbb{N}$, and observe that $\beta^n_\ell = \left(\beta^n_{\ell_t}\right)_t$ is a càdlàg, $\mathbb{F}^n_\ell$–martingale. The proof of such theorem is essentially based on the deterministic time–change procedure described by Zhang in [178, Section 2], but exploits the linear nature of our setting to avoid the application of the *Bismut–Elworthy–Li's formula* (see, e.g., [61, Proposition 8.21]). For the sake of completeness we report its main passages.

Theorem 1.4. *Let $t > 0$, $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$, $\ell \in \mathbb{S}$ and assume that $\sigma^2_n > 0$, $n = 1, \dots, N$. Then the function $\mathbb{E}\left[\phi\left(Z^\ell_t(\cdot)\right)\right]$ is differentiable at any point $x \in \mathbb{R}^N$ in every direction $h \in \mathbb{R}^N$, and*

$$\left\langle \nabla \mathbb{E}\left[\phi\left(Z^\ell_t(x)\right)\right], h\right\rangle = \mathbb{E}\left[\phi\left(Z^\ell_t(x)\right)\left(\sum_{n=1}^N \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}d\ell_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{\ell_s}\right)\right]. \tag{1.11}$$

Proof. For every $\epsilon > 0$ denote by $\ell^\epsilon(t) \coloneqq \frac{1}{\epsilon}\int_t^{t+\epsilon}\ell_s\,ds$, $t \geq 0$, the *Steklov's averages* of ℓ. They are strictly increasing, absolutely continuous functions such that, for every $t \geq 0$, $\ell^\epsilon_t \downarrow \ell_t$ as $\epsilon \downarrow 0$. Let $\gamma^\epsilon \coloneqq (\ell^\epsilon)^{-1} : [\ell^\epsilon_0, \infty) \to \mathbb{R}_+$ and define $Z^{\ell^\epsilon}(x)$ as in (1.10), i.e., for every $x \in \mathbb{R}^N$ the process $Z^{\ell^\epsilon}(x)$ is the unique solution of the linear SDE $dZ^{\ell^\epsilon}_t(x) = AZ^{\ell^\epsilon}_t(x)\,dt + \sqrt{Q}\,dW^N_{\ell^\epsilon_t}$, $Z^{\ell^\epsilon}_0 = x$. Now introduce the time–shifted processes $Y^{\ell^\epsilon}_t(x) \coloneqq Z^{\ell^\epsilon}_{\gamma^\epsilon_t}(x)$, $t \geq \ell^\epsilon_0$, and observe that

$$Y^{\ell^\epsilon}_t(x) = x + \int_{\ell^\epsilon_0}^t AY^{\ell^\epsilon}_s(x)\,\dot{\gamma}^\epsilon_s\,ds + \sqrt{Q}\left(W^N_t - W^N_{\ell^\epsilon_0}\right), \quad t \geq \ell^\epsilon_0,\ \mathbb{P}-\text{a.s.},$$

which shows that $dY^{\ell^\epsilon}_t(x) = AY^{\ell^\epsilon}_t(x)\,\dot{\gamma}^\epsilon_t dt + \sqrt{Q}\,dW^N_t$, $Y^{\ell^\epsilon}_{\ell^\epsilon_0}(x) = x$. Therefore,

$$Y^{\ell^\epsilon}_t(x) = e^{A\gamma^\epsilon_t}x + \int_{\ell^\epsilon_0}^t e^{A(\gamma^\epsilon_t - \gamma^\epsilon_s)}\sqrt{Q}\,dW^N_s, \quad t \geq \ell^\epsilon_0,\ \mathbb{P}-\text{a.s.}$$

In particular, since $\int_{\ell^\epsilon_0}^{\ell^\epsilon_t} e^{2A(t-\gamma^\epsilon_s)}Q\,ds = \int_0^t e^{2A(t-s)}Q\,d\ell^\epsilon_s$, where the integral is to be interpreted entrywise, we have

$$Z^{\ell^\epsilon}_t(x) = Y^{\ell^\epsilon}_{\ell^\epsilon_t}(x) \sim \mathcal{N}\left(e^{At}x, \int_0^t e^{2A(t-s)}Q\,d\ell^\epsilon_s\right).$$

At this point, we fix a generic $t > 0$, $x \in \mathbb{R}^N$ and use [178, Equation (2.6)] (it is just an application of *Gronwall lemma*) to get the convergence, in the L^2–sense, of $Z^{\ell^\epsilon}_t(x) \to Z^\ell_t(x)$ as $\epsilon \downarrow 0$. Moreover, recalling that $\ell^\epsilon_t \downarrow \ell_t$ as $\epsilon \downarrow 0$, we invoke *Helly's second theorem* (see [137, Theorem 7.3]) to get $\int_0^t e^{2A(t-s)}Q\,d\ell^\epsilon_s \to \int_0^t e^{2A(t-s)}Q\,d\ell_s$ as $\epsilon \downarrow 0$. Whence,

$$Z^\ell_t(x) \sim \mathcal{N}\left(e^{At}x, \int_0^t e^{2A(t-s)}Q\,d\ell_s\right). \tag{1.12}$$

Let us denote by $\Sigma_t \coloneqq \int_0^t e^{2A(t-s)}Q\,d\ell_s$. If we take $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$, then we are interested in differentiating the function

$$\mathbb{E}\left[\phi\left(Z^\ell_t(x)\right)\right] = \frac{1}{\sqrt{\det(2\pi\Sigma_t)}}\int_{\mathbb{R}^N}\phi(y)\exp\left\{-\frac{1}{2}\left\langle y - e^{At}x, \Sigma_t^{-1}\left(y - e^{At}x\right)\right\rangle\right\}dy.$$

An explicit computation simply based on the derivation of the normal density function implies, for every direction $h \in \mathbb{R}^N$,

$$\begin{aligned}&\left\langle \nabla \mathbb{E}\left[\phi\left(Z_t^\ell(x)\right)\right], h\right\rangle \\ &= \frac{1}{\sqrt{\det(2\pi\Sigma_t)}}\int_{\mathbb{R}^N}\phi(y)\exp\left\{-\frac{1}{2}\left\langle y-e^{At}x, \Sigma_t^{-1}\left(y-e^{At}x\right)\right\rangle\right\}\langle y-e^{At}x, \Sigma_t^{-1}e^{At}h\rangle dy \\ &= \mathbb{E}\left[\phi\left(Z_t^\ell(x)\right)\left\langle\left(\int_0^t e^{2A(t-s)}Q\,d\ell_s\right)^{-1}\left(\int_0^t e^{A(t-s)}\sqrt{Q}\,dW^N_{\ell_s}\right), e^{At}h\right\rangle\right],\end{aligned}$$

which coincides with (1.11) upon expanding the notation. ■

Now we investigate the subordinated Brownian motion case. The intuition behind the argument is to condition with respect to the σ–algebra $\mathcal{F}^L$, so that it is possible to apply the deterministic time–shift result we have just obtained in Theorem 1.4 upon changing the reference probability space. Let us denote by $\mathbb{W}$ the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^N$ vanishing at 0 and endow it with the Borel σ–algebra $\mathcal{B}(\mathbb{W})$ associated with the topology of locally uniform convergence. The pushforward probability measure generated by $W^N(\cdot)\colon(\Omega,\mathcal{F},\mathbb{P})\to(\mathbb{W},\mathcal{B}(\mathbb{W}))$ is denoted by $\mathbb{P}^{\mathbb{W}}$ and makes the canonical process $\mathfrak{x}=(x_t)_t$ a Brownian motion, where by definition

$$x_t(w) := w_t, \quad w\in\mathbb{W},\ t\ge 0.$$

We work with the usual completion $\left(\mathbb{W},\overline{\mathcal{B}(\mathbb{W})},\overline{\mathbb{P}^{\mathbb{W}}}\right)$ of this probability space: by [119, Theorem 7.9], $\mathfrak{x}$ is still a Brownian motion with respect to its minimal augmented filtration, which in turn satisfies the usual hypotheses and is denoted by $\mathbb{F}^{\mathbb{W}}$. In particular, note that the completeness of the space $(\Omega,\mathcal{F},\mathbb{P})$ implies the measurability of $W^N(\cdot)\colon(\Omega,\mathcal{F},\mathbb{P})\to\left(\mathbb{W},\overline{\mathcal{B}(\mathbb{W})}\right)$ and the fact that $\overline{\mathbb{P}^{\mathbb{W}}}$ is still the pushforward probability measure generated by $W^N(\cdot)$. Obviously, $W^N(\cdot)$ is independent from $\mathcal{F}^L$: as a consequence, a regular conditional probability of $W^N(\cdot)$ given $\mathcal{F}^L$ is the probability kernel

$$\begin{aligned}&\mathbb{P}\left(W^N(\cdot)\in\cdot\,\middle|\,\mathcal{F}^L\right):\Omega\times\overline{\mathcal{B}(\mathbb{W})}\to[0,1]\quad\text{such that}\quad\mathbb{P}\left(W^N(\cdot)\in A\right)(w):=\overline{\mathbb{P}^{\mathbb{W}}}(A),\\ &\omega\in\Omega,\ A\in\overline{\mathcal{B}(\mathbb{W})}.\end{aligned}\tag{1.13}$$

As regards the space $\mathbb{S}$, for every $t\ge 0$ we introduce the map $y_t\colon\mathbb{S}\to\mathbb{R}$ defined by $y_t(\ell):=\ell_t$, $\ell\in\mathbb{S}$, and consider the σ–algebra $\mathcal{F}^{\mathbb{S}}:=\sigma\left(y_t^{-1}(B),\,B\in\mathcal{B}(\mathbb{R}),\,t\ge 0\right)$. Since $L(\cdot)\colon\left(\Omega,\mathcal{F}^L,\mathbb{P}\right)\to\left(\mathbb{S},\mathcal{F}^{\mathbb{S}}\right)$ is measurable, we can construct the pushforward probability measure $\mathbb{P}^{\mathbb{S}}$ on $\left(\mathbb{S},\mathcal{F}^{\mathbb{S}}\right)$. At this point we take into account the product space $\left(\mathbb{W}\times\mathbb{S},\overline{\mathcal{B}(\mathbb{W})}\otimes\mathcal{F}^{\mathbb{S}},\overline{\mathbb{P}^{\mathbb{W}}}\otimes\mathbb{P}^{\mathbb{S}}\right)$ and note that, thanks to the mutual independence of $W^N(\cdot)$ and $L(\cdot)$, the product measure $\overline{\mathbb{P}^{\mathbb{W}}}\otimes\mathbb{P}^{\mathbb{S}}$ is indeed the pushforward probability measure generated by $\psi\colon\Omega\to\mathbb{W}\times\mathbb{S}$, $\psi(\omega):=\left(W^N_\cdot(\omega),L_\cdot(\omega)\right)$. Finally, we take the process $z=(z_t)_t$ defined by

$$z_t(w,\ell):=w_{\ell_t},\quad(w,\ell)\in\mathbb{W}\times\mathbb{S},\ t\ge 0,$$

and denote by $\mathbb{F}^z=(\mathcal{F}^z_t)_t$ its natural filtration. By construction, $W^N_{L_t}=z_t\circ\psi$ for every $t\ge 0$. Putting together all these properties, we can conclude that z is a Lévy process with respect to the right–continuous filtration $\mathbb{F}^z_+=\left(\mathcal{F}^z_{t+}\right)_t$, where

$$\mathcal{F}^z_{t+}:=\bigcap_{\epsilon>0}\mathcal{F}^z_{t+\epsilon},\quad t\ge 0.$$

Endowing the product space with this filtration, the stochastic integral of suitable processes with respect to z is well–defined. Let us consider then a deterministic, continuous, bounded, $\mathbb{R}^N$–valued process $\xi = (\xi_t)_t$: weaker assumptions can be done on it, but in our framework these are sufficient. Clearly the subordinated Brownian motion W_L^N is adapted with respect to the right–continuous filtration $\psi^{-1}\left(\mathbb{F}_+^z\right)$, therefore the usual rules of change of probability space (see, e.g., [109, §X-2]) entail

$$\int_0^t \xi_s \cdot dW_{L_s}^N = \left(\int_0^t \xi_s \cdot dz_s\right) \circ \psi, \quad t \geq 0, \, \mathbb{P} - \text{a.s.} \tag{1.14}$$

We conclude this preliminary discussion with an important substitution formula.

Lemma 1.5. *Let $\xi = (\xi_t)_t$ be a deterministic, continuous, bounded, $\mathbb{R}^N$–valued process. Then, for any $t > 0$,*

$$\left(\int_0^t \xi_s \cdot dz_s\right)(\cdot, \ell) = \int_0^t \xi_s \cdot dx_{\ell_s} \quad \overline{\mathbb{P}^{\mathbb{W}}} - a.s., \, for \ \mathbb{P}^{\mathbb{S}} - a.e. \ \ell \in \mathbb{S},$$

where the integral on the right–hand side of the equality is intended in the sense of stochastic integrals by càdlàg martingales on the filtered probability space $\left(\mathbb{W}, \overline{\mathcal{B}(\mathbb{W})}, \overline{\mathbb{P}^{\mathbb{W}}}; \mathbb{F}_\ell^{\mathbb{W}}\right)$.

Proof. Fix $t > 0$ and consider the elementary, predictable (with respect to both $\mathbb{F}_+^z$ and $\mathbb{F}_\ell^{\mathbb{W}}$, $\ell \in \mathbb{S}$) processes

$$\xi_s^m \coloneqq \xi_0 1_{\{0\}}(s) + \sum_{i=0}^{m-1} \xi_{t_i} 1_{]t_i, t_{i+1}]}(s), \quad s \geq 0, \, m \in \mathbb{N},$$

where $t_i = \frac{t}{m} i$, $i = 0, \dots, m$. The continuity of ξ implies that $\xi^m \to \xi$ pointwise in $[0, t]$ as $m \to \infty$; furthermore, since ξ is bounded, the sequence $(\xi^m)_m$ is uniformly bounded. Hence, by [110, Theorem 4.31, Chapter I],

$$\int_0^t \xi_s \cdot dz_s = \left(\overline{\mathbb{P}^{\mathbb{W}}} \otimes \mathbb{P}^{\mathbb{S}}\right) - \lim_{m \to \infty} \int_0^t \xi_s^m \cdot dz_s,$$

where the notation on the right–hand side denotes the limit in probability. Now convergence in probability implies almost–sure convergence along a subsequence, hence we can say that for $\mathbb{P}^{\mathbb{S}}$–a.e. $\ell \in \mathbb{S}$,

$$\left(\int_0^t \xi_s^{m_k} \cdot dz_s\right)(\cdot, \ell) \underset{k \to \infty}{\longrightarrow} \left(\int_0^t \xi_s \cdot dz_s\right)(\cdot, \ell) \quad \overline{\mathbb{P}^{\mathbb{W}}} - \text{a.s.} \tag{1.15}$$

With the same argument as above, we have

$$\int_0^t \xi_s \cdot dx_{\ell_s} = \overline{\mathbb{P}^{\mathbb{W}}} - \lim_{k \to \infty} \int_0^t \xi_s^{m_k} \cdot dx_{\ell_s} \quad \text{for every } \ell \in \mathbb{S}. \tag{1.16}$$

On the other hand, by the definition of stochastic integral of an elementary predictable process and noticing that $z_s(w, \ell) = w_{\ell_s} = x_{\ell_s}(w)$ for every $(w, \ell) \in \mathbb{W} \times \mathbb{S}$ and $s \geq 0$, one has

$$\begin{aligned}\left(\int_0^t \xi_s^{m_k} \cdot dz_s\right)(w, \ell) &= \sum_{i=0}^{m_k - 1} \xi_{t_i} \cdot \left(z_{t_{i+1}} - z_{t_i}\right)(w, \ell) \\ &= \sum_{i=0}^{m_k - 1} \xi_{t_i} \cdot \left(x_{\ell_{t_{i+1}}} - x_{\ell_{t_i}}\right)(w) = \left(\int_0^t \xi_s^{m_k} \cdot dx_{\ell_s}\right)(w),\end{aligned}$$

which holds true for all $(w,\ell)\in\mathbb{W}\times\mathbb{S}$. Combining the last equation with (1.15) and (1.16) we get

$$\left(\int_0^t \xi_s\cdot dz_s\right)(\cdot,\ell)=\int_0^t \xi_s\cdot dx_{\ell_s}\quad \overline{\mathbb{P}^{\mathbb{W}}}-\text{a.s., for } \mathbb{P}^{\mathbb{S}}-\text{a.e. } \ell\in\mathbb{S},$$

proving the thesis of the lemma. ∎

A useful result due to [33, Equation (14)] shows that there exists a constant $c>0$ such that, for every $t>0$, the density η_t of L_t satisfies

$$\eta_t(s)\le c\,t\,s^{-1-\alpha}e^{-ts^{-\alpha}},\quad s>0.$$

As a consequence, for every $p>0$,

$$\mathbb{E}\left[\frac{1}{L_t^p}\right]=\int_0^\infty \frac{1}{s^p}\eta_t(s)\,ds\le c\,t\int_0^\infty s^{-\alpha-1-p}e^{-ts^{-\alpha}}ds=\left(\frac{c}{\alpha}\int_0^1\left(\log\frac{1}{u}\right)^{p/\alpha}du\right)\frac{1}{t^{p/\alpha}},$$

performing the substitution $u=e^{-ts^{-\alpha}}$. Therefore ${L_t}^{-1}\in L^p$, and

$$\mathbb{E}\left[\frac{1}{L_t^p}\right]^{1/p}\le c_{\alpha,p}\frac{1}{t^{1/\alpha}}\quad\text{for some } c_{\alpha,p}>0. \tag{1.17}$$

We refer to [70, Theorem 2.1] for moment estimates concerning general subordinators (not necessarily with an $\alpha-$stable distribution). We are now in position to obtain the derivation formula for the Markov transition semigroup, together with an estimate on its gradient, in the subordinated Brownian motion case.

Theorem 1.6. *Let $t>0$, $\phi\in\mathcal{B}_b\left(\mathbb{R}^N\right)$ and assume that $\sigma_n^2>0$, $n=1,\dots,N$. Then the function $\mathbb{E}\left[\phi\left(Z_t\right)\right]$ is differentiable at any point $x\in\mathbb{R}^N$ in every direction $h\in\mathbb{R}^N$, and*

$$\langle\nabla\mathbb{E}\left[\phi\left(Z_t^x\right)\right],h\rangle=\mathbb{E}\left[\phi\left(Z_t^x\right)\left(\sum_{n=1}^N\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]. \tag{1.18}$$

In addition, there exists $c_\alpha>0$ such that the following gradient estimate holds:

$$\sup_{x\in\mathbb{R}^N}\left|\nabla\mathbb{E}\left[\phi\left(Z_t^x\right)\right]\right|\le c_\alpha\|\phi\|_\infty\sup_{n=1,\dots,N}\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}}e^{-\lambda_n t}\right)\quad\textit{for every } t>0. \tag{1.19}$$

Proof. Fix $t>0$ and $\phi\in\mathcal{B}_b\left(\mathbb{R}^N\right)$. In what follows, we denote by $\mathbb{E}^{\mathbb{W}}[\cdot]$ the expected value of a random variable defined on $\left(\mathbb{W},\overline{\mathcal{B}(\mathbb{W})},\overline{\mathbb{P}^{\mathbb{W}}}\right)$. Since $Z_t^x=e^{tA}x+\int_0^t e^{(t-s)A}\sqrt{Q}\,dW^N_{L_s}$, by (1.14) we have, for every $x\in\mathbb{R}^N$, $\mathbb{P}-$a.s.,

$$Z_t^x=\left(e^{tA}x+\int_0^t e^{(t-s)A}\sqrt{Q}\,dz_s\right)\circ\psi=\left(e^{tA}x+\int_0^t e^{(t-s)A}\sqrt{Q}\,dz_s\right)\left(W^N(\cdot),L(\cdot)\right)$$

Thus, recalling the expression (1.13) for the regular conditional probability $\mathbb{P}\left(W^N(\cdot)\in\cdot\middle|\mathcal{F}^L\right)$, we apply the *disintegration formula* for the conditional expectation to write, for every $x\in\mathbb{R}^N$,

$$\begin{aligned}\mathbb{E}\left[\phi\left(Z_t^x\right)\right]&=\mathbb{E}\left[\mathbb{E}\left[\phi\left(Z_t^x\right)\middle|\mathcal{F}^L\right]\right]=\mathbb{E}\left[\int_{\mathbb{W}}\phi\left(\left(e^{tA}x+\int_0^t e^{(t-s)A}\sqrt{Q}\,dz_s\right)(w,L(\cdot))\right)\overline{\mathbb{P}^{\mathbb{W}}}(dw)\right]\\&=\mathbb{E}\left[\mathbb{E}^{\mathbb{W}}\left[\phi\left(e^{tA}x+\int_0^t e^{(t-s)A}\sqrt{Q}\,dx_{\ell_s}\right)\right]\Big|_{\ell=L(\cdot)}\right]=\mathbb{E}\left[\mathbb{E}^{\mathbb{W}}\left[\phi\left(Z_t^\ell(x)\right)\right]\Big|_{\ell=L(\cdot)}\right],\end{aligned}$$

where in the second–to–last equality we use Lemma 1.5 and the fact that $\mathbb{P}^{\mathbb{S}}$ is the pushforward probability measure generated by $L(\cdot)$ on $\mathbb{S}$. Take $x \in \mathbb{R}^N$ and a direction $h \in \mathbb{R}^N$; if we can justify the derivation under the expected value, an application of (1.11) immediately leads to (1.18), as the following computations based on the previous argument show:

$$\langle \nabla \mathbb{E}[\phi(Z_t^x)], h\rangle = \mathbb{E}\left[\mathbb{E}^{\mathbb{W}}\left[\phi\left(Z_t^\ell(x)\right)\left(\sum_{n=1}^N \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}d\ell_s}\int_0^t e^{-\lambda_n(t-s)}dx^n_{\ell_s}\right)\right]\Bigg|_{\ell=L(\cdot)}\right] \tag{1.20}$$

$$= \mathbb{E}\left[\sum_{n=1}^N \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\left\{\int_{\mathbb{W}}\left(\phi\left(e^{tA}x + \int_0^t e^{(t-s)A}\sqrt{Q}\,dz_s\right)\right.\right.\right.$$
$$\left.\left.\left.\times\left(\int_0^t e^{-\lambda_n(t-s)}dz_s^n\right)\right)(w, L(\cdot))\,\overline{\mathbb{P}^{\mathbb{W}}}(dw)\right\}\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\phi(Z_t^x)\left(\sum_{n=1}^N \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\Bigg|\mathcal{F}^L\right]\right].$$

Indeed, such a derivation is licit, since *Jensen's inequality* and (1.12) entail

$$\left|\mathbb{E}^{\mathbb{W}}\left[\phi\left(Z_t^\ell(x)\right)\left(\sum_{n=1}^N \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}d\ell_s}\int_0^t e^{-\lambda_n(t-s)}dx^n_{\ell_s}\right)\right]\Bigg|_{\ell=L(\cdot)}\right|^2$$
$$\leq \|\phi\|_\infty^2 \sum_{n=1}^N \frac{1}{\sigma_n^2}\frac{e^{-2\lambda_n t}|\langle h, e_n\rangle|^2}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}, \tag{1.21}$$

with the right–hand side which does not depend on x and is integrable. In fact, for every $n = 1, \dots, N$, recalling that $L_1 \sim \text{stable}\left(\alpha, 1, \bar{c}^{1/\alpha}, 0\right)$ by (1.1), we have

$$\int_0^t e^{-2\lambda_n(t-s)}dL_s \sim \text{stable}\left(\alpha, 1, \bar{c}^{\frac{1}{\alpha}}\left(\frac{1-e^{-2\alpha\lambda_n t}}{2\alpha\lambda_n}\right)^{1/\alpha}, 0\right)$$
$$\Longrightarrow \int_0^t e^{-2\lambda_n(t-s)}dL_s \sim \left(\frac{1-e^{-2\alpha\lambda_n t}}{2\alpha\lambda_n}\right)^{\frac{1}{\alpha}} L_1,$$

hence by (1.17) there exists $c_\alpha > 0$ such that

$$\mathbb{E}\left[\sum_{n=1}^N \frac{1}{\sigma_n^2}\frac{e^{-2\lambda_n t}|\langle h, e_n\rangle|^2}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\right] \leq \mathbb{E}\left[\frac{1}{L_1}\right]\left(\sum_{n=1}^N \frac{e^{-2\lambda_n t}}{\sigma_n^2}\left(\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}\right)^{\frac{1}{\alpha}}|\langle h, e_n\rangle|^2\right)$$
$$\leq c_\alpha \sum_{n=1}^N \frac{e^{-2\lambda_n t}}{\sigma_n^2}\left(\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}\right)^{\frac{1}{\alpha}}|\langle h, e_n\rangle|^2. \tag{1.22}$$

Concerning the gradient estimate, it is sufficient to combine (1.20), (1.21) & (1.22) and to recall that the L^1–norm of a random variable is smaller than its L^2–norm to get

$$|\langle \nabla \mathbb{E}[\phi(Z_t^x)], h\rangle| \leq c_\alpha \|\phi\|_\infty \sup_{n=1,\dots,N}\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}}e^{-\lambda_n t}\right)|h|, \quad x, h \in \mathbb{R}^N,$$

where the constant c_α is allowed to be different from the one in (1.22). The desired inequality (1.19) is then recovered taking the sup for $|h| \leq 1$, and the proof is complete. ■

Remark 1.3. *The derivative formula in* (1.18) *is different from the Bismut–Elworthy–Li's type formula provided by [178]. More precisely, according to [178, Theorem* 1.1*], if* $\sigma_n^2 > 0$, $n = 1, \dots, N$ *and* $\phi \in C_b^1(\mathbb{R}^N)$*, then for every* $t > 0$, $x \in \mathbb{R}^N$ *and* $h \in \mathbb{R}^N$,

$$\langle \nabla \mathbb{E}\left[\phi\left(Z_t^x\right)\right], h\rangle = \mathbb{E}\left[\frac{1}{L_t}\phi\left(Z_t^x\right)\int_0^t \left(\sqrt{Q}\right)^{-1} e^{sA}h \cdot dW_{L_s}^N\right].$$

As a side note, we refer to [175] for a derivative formula of Bismut–Elworthy–Li's type for SDEs driven by a multiplicative Lévy noise (not necessarily $\alpha-$*stable) obtained by subordination of a Brownian motion.*

Remark 1.4. *Under the same hypotheses of Theorem 1.6, the derivative formula in* (1.18) *holds true also for a subordinator* L *which is not* $\alpha-$*stable and such that* $\mathbb{E}\left[L_t^{-1/2}\right] < \infty$ *for every* $t > 0$ *(sufficient conditions for this to happen can be found, e.g., in [70, Theorem* 2.1*]). The proof is the same as before, as long as we justify the differentiation under the expected value. To do that, we simply note that the expectation of the square root of the right–hand side in* (1.21) *is finite by the estimate*

$$\left(\int_0^t e^{-2\lambda_n(t-s)}dL_s\right)^{-1/2} \leq e^{\lambda_n t}L_t^{-1/2}, \quad n = 1, \dots, N.$$

1.2.2 Infinite–dimensional case

In this subsection we analyze the general case where H is infinite–dimensional. We work under the following assumption.

Hypothesis 1.2. *Suppose that* $\sigma_n^2 > 0$, $n \in \mathbb{N}$, *and that*

$$\sup_n\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}}e^{-\lambda_n t}\right) \leq C_t \qquad \textit{for every } t > 0, \textit{ for some function } C_t > 0.$$

In this setting, for every $h \in H$ and $t > 0$, we can define the real–valued random variable

$$\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta_{L_s}^n := L^2 - \lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta_{L_s}^n\right).$$

Indeed, with the same argument as the one in (1.22), Hypothesis 1.2 yields

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{n=m}^{M}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta_{L_s}^n\right|^2\right] &\leq c_\alpha\sum_{n=m}^{M}\frac{e^{-2\lambda_n t}}{\sigma_n^2}\left(\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}\right)^{\frac{1}{\alpha}}\left|\langle h, e_n\rangle\right|^2 \\ &\leq c_\alpha\, C_t^2\left(\sum_{n=m}^{M}\left|\langle h, e_n\rangle\right|^2\right)\underset{m,M\to\infty}{\longrightarrow} 0,\end{aligned}$$

where $c_\alpha > 0$. In particular,

$$\mathbb{E}\left[\left|\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h, e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta_{L_s}^n\right|^2\right]^{\frac{1}{2}} \leq \sqrt{c_\alpha}\, C_t\left\|h\right\|_H. \tag{1.23}$$

Hence the following, useful property holds:

$$\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h_m,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\xrightarrow{L^2}\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s},$$
$$\text{as } h_m\to h. \tag{1.24}$$

At this point we can present the main theorem of the chapter.

Theorem 1.7. *Assume Hypotheses 1.1 & 1.2. Then for every $\phi\in\mathcal{B}_b(H)$ and $t>0$ the function $R_t\phi\in C_b^1(H)$ and there exists $c_\alpha>0$ such that*

$$\sup_{x\in H}\|\nabla R_t\phi(x)\|_H\le c_\alpha\,C_t\,\|\phi\|_\infty \quad \textit{for every } t>0. \tag{1.25}$$

Moreover, given $\phi\in C_b(H)$ and $t>0$, for every $x,h\in H$ the Gateaux derivative of $R_t\phi$ at x along the direction h is given by

$$\langle\nabla R_t\phi(x),h\rangle=\mathbb{E}\left[\phi(Z_t^x)\left(\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]. \tag{1.26}$$

Proof. Fix $t>0$ and a function $\phi\in C_b(H)$.

We first consider the case $\dim H=N$, identifying H with $\mathbb{R}^N$, as usual. Evidently (1.26) coincides with (1.18) and the map $x\mapsto\nabla R_t(x)$ is a continuous function from $\mathbb{R}^N$ into itself: this follows from dominated convergence, together with $\phi\in C_b(\mathbb{R}^N)$ and $Z_t^{x_n}\to Z_t^x$ a.s. as $x_n\to x$. Moreover, Hypothesis 1.2 applied to (1.19) directly entails (1.25), therefore $R_t\phi\in C_b^1(\mathbb{R}^N)$. In order to pass to infinite dimension it is convenient to write

$$\begin{aligned}R_t\phi(x+h)-R_t\phi(x)&=\int_0^1\Big\langle\nabla R_t\phi\Big((1-\rho)x+\rho(x+h)\Big),h\Big\rangle d\rho\\&=\int_0^1\mathbb{E}\left[\phi\left(Z_t^{x+\rho h}\right)\left(\sum_{n=1}^{N}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]d\rho.\end{aligned}\tag{1.27}$$

We now consider the general case $\dim H=\infty$. Let π_N be the projection onto the first N Fourier components and H_N be its range. Due to the diagonal structure of our model, the projections $\pi_N Z_t^x$ of the OU–process are, $\mathbb{P}-$a.s.,

$$\pi_N Z_t^x=\sum_{n=1}^{N}e^{-\lambda_n t}\langle x,e_n\rangle e_n+\sum_{n=1}^{N}\left(\int_0^t e^{-\lambda_n(t-s)}\sigma_n d\beta^n_{L_s}\right)e_n,\quad N\in\mathbb{N}.$$

Therefore introducing the operators $A_N:=A\big|_{H_N}$ and $Q_N:=Q\big|_{H_N}$, which map H_N into itself, we can write $\pi_N Z_t^x=e^{tA_N}(\pi_N x)+\tilde{Z}_{A_N,Q_N}(t)$: this shows that such projections are OU–processes in H_N. Thus, the dominated convergence theorem together with the expression in (1.27) and the continuity of ϕ give

$$\begin{aligned}R_t\phi(x+h)-R_t\phi(x)&=\lim_{N\to\infty}\mathbb{E}\left[\phi\left(\pi_N Z_t^{x+h}\right)-\phi(\pi_N Z_t^x)\right]\\&=\lim_{N\to\infty}\int_0^1\mathbb{E}\left[\phi\left(\pi_N Z_t^{x+\rho h}\right)\left(\sum_{n=1}^{N}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]d\rho\\&=\int_0^1\mathbb{E}\left[\phi\left(Z_t^{x+\rho h}\right)\left(\sum_{n=1}^{\infty}\frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]d\rho.\end{aligned}$$

Now we can define $d_x(R_t\phi)(h) := \mathbb{E}\left[\phi(Z_t^x)\left(\sum_{n=1}^\infty \frac{1}{\sigma_n}\frac{e^{-\lambda_n t}\langle h,e_n\rangle}{\int_0^t e^{-2\lambda_n(t-s)}dL_s}\int_0^t e^{-\lambda_n(t-s)}d\beta^n_{L_s}\right)\right]$: it is the Fréchet differential of $R_t\phi$ at x. To see this, it is sufficient to note that the linear operator $d_x(R_t\phi)(\cdot)$ is continuous by the property in (1.24) and to apply Hölder's inequality, the dominated convergence theorem and (1.23) to get, for a positive constant c_α,

$$|R_t\phi(x+h) - R_t\phi(x) - d_x(R_t\phi)(h)| \le c_\alpha C_t \|h\|_H \int_0^1 \mathbb{E}\left[\left|\phi\left(Z_t^{x+\rho h}\right) - \phi(Z_t^x)\right|^2\right]^{1/2} d\rho = o(\|h\|_H).$$

As a consequence of the definition of $d_x(R_t\phi)(\cdot)$, the formula in (1.26) is verified. The upper bound (1.25) for the norm of the gradient is then obtained by (1.23) from the next, straightforward computation:

$$\|\nabla R_t\phi(x)\|_H = \sup_{\|h\|_H\le 1}|\langle\nabla R_t\phi(x), h\rangle| = \sup_{\|h\|_H\le 1}|d_x(R_t\phi)(h)| \le c_\alpha C_t \|\phi\|_\infty, \quad x\in H.$$

We also note that

$$\sup_{\|h\|_H\le 1}|(d_{x_n}(R_t\phi) - d_x(R_t\phi))(h)| \le c_\alpha C_t \mathbb{E}\left[|\phi(Z_t^{x_n}) - \phi(Z_t^x)|^2\right]^{1/2} \to 0 \quad \text{as } x_n \to x:$$

this proves the continuity of the map $x \mapsto d_\cdot(R_t\phi)$, hence $R_t\phi \in C_b^1(H)$.

Finally, we need to study the case where ϕ is just Borel measurable and bounded, without the hypothesis of continuity. In order to do this, it is sufficient to observe that by the *mean value theorem* and (1.25) we have, for every $\phi \in C_b^2(H)$,

$$|R_t\phi(x) - R_t\phi(y)| \le c_\alpha C_t \|\phi\|_\infty \|x-y\|_H, \quad x,y\in H. \tag{1.28}$$

Being R_t Markovian, [64, Lemma 7.1.5] implies that the same holds true for every $\phi \in \mathcal{B}_b(H)$. In particular, R_t maps bounded, Borel measurable functions in bounded, Lip–continuous functions. The semigroup law let us write $R_t\phi = R_s(R_{t-s}\phi)$ for some $0 < s < t$, which proves $R_t\phi \in C_b^1(H)$. The bound (1.25) follows from (1.28), hence the proof is complete. ■

We now focus on the gradient estimate (1.9). We need to substitute Hypothesis 1.2 with the following, stronger one.

Hypothesis 1.3. *Suppose that $\sigma_n^2 > 0$, $n\in\mathbb{N}$, and that*

$$\sup_n\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}}e^{-\lambda_n t}\right) \le C_0\frac{1}{t^\gamma}, \quad \text{for every } t>0, \text{ for some } C_0>0,\ 0<\gamma<1.$$

In other terms, in Hypothesis 1.2 we take $C_t := C_0 t^{-\gamma}$, $t>0$, for some $C_0>0$, $\gamma\in(0,1)$.

Remark 1.5. *Observe that, for every $n\in\mathbb{N}$, the term*

$$\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n t}}}e^{-\lambda_n t} \sim \frac{1}{\sigma_n}\frac{1}{t^{1/(2\alpha)}} \quad \text{as } t\downarrow 0.$$

Therefore, Hypothesis 1.3 should be verified only in the case $\alpha\in\left(\frac{1}{2},1\right)$ and for some $\gamma\in\left[\frac{1}{2\alpha},1\right)$.

It is also worth noticing that Hypothesis 1.3 is equivalent to the next condition.

Hypothesis 1.4. *For some $C_1 > 0$ and $\gamma \in \left[\frac{1}{2\alpha}, 1\right)$,*

$$\sigma_n \geq C_1 \, \lambda_n^{\frac{1}{2\alpha}-\gamma}, \quad n \in \mathbb{N}.$$

A short argument proving the equivalence of the two statements is shown in [151, Hypothesis (N)*].*

At this point the next result is immediate.

Corollary 1.8. *Consider $\alpha \in \left(\frac{1}{2}, 1\right)$ and assume Hypotheses 1.1 & 1.3. Then for every $\phi \in \mathcal{B}_b(H)$ the function $R_t\phi \in C_b^1(H)$, $t > 0$, and the gradient estimate* (1.9) *holds, namely there exists a constant $C > 0$ such that*

$$\sup_{x \in H} \|\nabla R_t \phi(x)\|_H \leq \frac{C}{t^\gamma} \|\phi\|_\infty \quad \text{for every } t > 0,$$

where $\gamma \in \left[\frac{1}{2\alpha}, 1\right)$ is the one appearing in Hypothesis 1.3.

Example 1.2. We investigate Hypothesis 1.3 –in its equivalent formulation Hypothesis 1.4– in the same framework as in Example 1.1. So we take $A = \Delta$ (hence $-\lambda_k = -(2\pi)^2 |k|^2$, $k \in \mathbb{Z}_0^d$) and study two possible choices for Q.

- If $Q = \mathrm{Id}$, then
$$1 \geq \frac{1}{(2\pi |k|)^{2\left(\gamma - \frac{1}{2\alpha}\right)}}, \quad k \in \mathbb{Z}_0^d$$
for every $\gamma \in \left[\frac{1}{2\alpha}, 1\right)$. Therefore, in dimension $d = 1$ both Hypotheses 1.1 & 1.3 are satisfied. In particular, motivated by the fact that R_t is a regularization operator with $R_0 = \mathrm{Id}$, we are interested in the behavior of $\nabla R_t \phi$ around 0, where $\phi \in \mathcal{B}_b(H)$. Therefore we choose $\gamma = \frac{1}{2\alpha}$ and Corollary 1.8 provides the next estimate:
$$\sup_{x \in H} \|\nabla R_t \phi(x)\|_H \leq C \|\phi\|_\infty \frac{1}{t^{2\alpha}} \quad \text{for every } t > 0,$$
for a positive constant C.

- If $Q = Q_\eta = (-\Delta)^{-\eta}$ for $\eta > 0$, then $\sigma_k^{(\eta)} = \lambda_k^{-\eta/2}$, $k \in \mathbb{Z}_0^d$, and Hypothesis 1.4 holds true if and only if $\eta \leq 2\gamma - \frac{1}{\alpha}$. Since we can take any $\gamma \in \left[\frac{1}{2\alpha}, 1\right)$, the aforementioned condition holds as soon as $\eta < 2 - \frac{1}{\alpha}$. Combining this result with (1.8) obtained in Example 1.1, we conclude that Hypotheses 1.1 & 1.3 simultaneously hold if and only if
$$\eta \in \left(\max\left\{ \frac{d-2}{2\alpha}, 0 \right\}, \, 2 - \frac{1}{\alpha} \right).$$
It then follows that there exist negative fractional powers of the Laplacian $Q_\eta = (-\Delta)^{-\eta}$ meeting the requirements of Corollary 1.8 up to dimension $d = 3$. Specifically, for $d = 1, 2$ there is a Q_η with the searched properties for every $\alpha \in \left(\frac{1}{2}, 1\right)$, whereas in dimension $d = 3$ we can find such a Q_η only for $\alpha \in \left(\frac{3}{4}, 1\right)$.

Chapter 2

Probability computation via mild Kolmogorov equations

In this chapter, given $\alpha \in (\frac{1}{2}, 1)$, we study an iterative scheme that allows to approximate the Markov transition semigroup associated with semilinear SDEs in $\mathbb{R}^N$ driven by $2\alpha-$stable, rotation–invariant Lévy processes obtained by subordination of Brownian motions, see (I.1)-(I.3). This iterative scheme arises from a mild integral formulation of the Kolmogorov equation corresponding to the SDEs, see (I.4), and relies on a single bulk of Monte Carlo simulations as several parameters of the dynamics change. In Section 2.6 we perform numerical experiments in dimension $N = 100$ for two choices of the nonlinear drift B_0. The outcomes show that, in these examples, the first iteration provides a compelling improvement over the linear OU approximation.

Notation Let $d, m, n \in \mathbb{N}$. In this chapter, elements of $\mathbb{R}^d$ are columns vectors. For any $u, v \in \mathbb{R}^d$, we denote by $|u|$ the Euclidean norm and by $\langle u, v\rangle = u^\top v$ the standard scalar product. For a matrix $A \in \mathbb{R}^{d\times m}$, $|A| = \sup_{x\in\mathbb{R}^m : |x|=1} |Ax|$ is the operator norm. Given a vector field $B \colon \mathbb{R}^d \to \mathbb{R}^{m\times n}$, the uniform norm is $\|B\|_\infty = \sup_{x\in\mathbb{R}^d} |B(x)|$. In particular, if $n = 1$ then the Jacobian matrix is denoted by $DB \in \mathbb{R}^{m\times d}$, and $D_h B = DBh$, $h \in \mathbb{R}^d$; if also $m = 1$ (so that B is a scalar function) then the gradient ∇B is a row vector and $D^2 B \in \mathbb{R}^{d\times d}$ represents the Hessian matrix. For an integer $k \in \mathbb{N} \cup \{0\}$, the space $C_b^k\left(\mathbb{R}^d; \mathbb{R}^{m\times n}\right)$ is constituted by the continuous vector fields B which are bounded, continuously differentiable up to order k with bounded derivatives. Taken $h = 1, \dots, k$ and $B \in C_b^k\left(\mathbb{R}^d; \mathbb{R}^{m\times n}\right)$, we write $\left\|\partial^h B\right\|_\infty = \sup_{i,j,\mathbf{h}} \|\partial_{\mathbf{h}} B_{i,j}\|_\infty$, where $B = (B_{i,j})$, $i = 1, \dots, m$, $j = 1, \dots, n$ and $\mathbf{h} \in (\mathbb{N} \cup \{0\})^d$ is a multi–index with length $\|\mathbf{h}\|_1 = h$.

2.1 Preliminaries and Kolmogorov backward equation in mild form

Fix $N \in \mathbb{N}$ and a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider N independent Brownian motions $(\beta^n)_{n=1,\dots,N}$: we write $W = \left[\beta^1, \dots, \beta^N\right]^\top$. Moreover, for $\alpha \in (0, 1)$ we take a strictly $\alpha-$stable subordinator $L = (L_t)_{t\ge 0}$ independent from $(\beta^n)_n$, and denote by $\mathcal{F}^L$ the augmented $\sigma-$algebra it generates, i.e., $\mathcal{F}^L = \sigma\left(\mathcal{F}_0^L \cup \mathcal{N}\right)$, where $\mathcal{F}_0^L$ is the natural $\sigma-$algebra generated by L and $\mathcal{N}$ is the family of $\mathcal{F}-$negligible sets. In other words, L is an increasing Lévy process with (cfr. [164, Example 24.12] and (1.1))

$$\mathbb{E}\left[e^{iuL_1}\right] = \exp\left\{-\bar{\gamma}^\alpha |u|^\alpha \left(1 - i \tan\frac{\pi\alpha}{2} \operatorname{sign} u\right)\right\}, \quad u \in \mathbb{R}, \text{ for some } \bar{\gamma} > 0. \tag{2.1}$$

Let us introduce the diagonal matrices $A = -\mathrm{diag}\left[\lambda_1, \ldots, \lambda_N\right]$ and $Q = \mathrm{diag}\left[\sigma_1^2, \ldots, \sigma_N^2\right]$, with $0 < \lambda_1 \leq \cdots \leq \lambda_N$ and $\sigma_n^2 > 0$, $n = 1, \ldots, N$. We endow Ω with the minimal augmented filtration $\mathbb{F} = \left(\mathcal{F}_t\right)_{t\geq 0}$ generated by W_L, which means $\mathcal{F}_t = \sigma\left(\mathcal{F}_{0,t}^{W_L} \cup \mathcal{N}\right)$ for $t \geq 0$, with $\left(\mathcal{F}_{0,t}^{W_L}\right)_{t\geq 0}$ being the natural filtration of W_L, see also Remark 1.1.

Given $T > 0$ and a continuous function $f\colon [0,T] \to \mathbb{R}^N$, if $x \in \mathbb{R}^N$ and $0 \leq s < T$ then $Z^{s,x} = (Z_t^{s,x})_{t\in[s,T]}$ is the OU process starting from x at time s, i.e., it is the unique solution of the next linear SDE

$$dZ_t^{s,x} = \left(AZ_t^{s,x} + f(t)\right) dt + \sqrt{Q}\, dW_{L_t}, \quad Z_s^{s,x} = x. \tag{2.2}$$

We denote by $R = (R_{s,t})$, $0 \leq s \leq t \leq T$, the time–dependent, Markov transition semigroup associated with this family of processes:

$$R_{s,t}\phi = \mathbb{E}\left[\phi\left(Z_t^{s,\cdot}\right)\right], \quad 0 \leq s \leq t < T,\ \phi \in \mathcal{B}_b\left(\mathbb{R}^N\right), \tag{2.3}$$

where $\mathcal{B}_b\left(\mathbb{R}^N\right)$ denotes the space of real–valued, Borel measurable and bounded functions defined on $\mathbb{R}^N$. The Chapman–Kolmogorov equations ensure that

$$R_{s,t}\left(R_{t,u}\phi\right) = R_{s,u}\phi, \quad 0 \leq s < t < u \leq T,\ \phi \in \mathcal{B}_b\left(\mathbb{R}^N\right). \tag{2.4}$$

For any $0 \leq s < t \leq T$, we define $I_{s,t}^L = \int_s^t e^{2(t-r)A} Q\, dL_r \colon \Omega \to \mathbb{R}^{N\times N}$ and $F_{s,t} = \int_s^t e^{(t-r)A} f(r)\, dr \in \mathbb{R}^N$. An adaptation of [34, Theorem 6] (see also Theorem 1.6 in Chapter 1) guarantees that, for any $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$, the function $R_{s,t}\phi$ is differentiable at any point $x \in \mathbb{R}^N$ in every direction $h \in \mathbb{R}^N$, with

$$\left\langle \nabla^\top R_{s,t}\phi(x), h\right\rangle = \mathbb{E}\left[\phi\left(Z_t^{s,x}\right)\left\langle \left(I_{s,t}^L\right)^{-1} e^{(t-s)A}h, Z_t^{s,x} - e^{(t-s)A}x - F_{s,t}\right\rangle\right]. \tag{2.5}$$

Moreover, $R_{s,t}\phi \in C_b^1\left(\mathbb{R}^N\right)$ and the following gradient estimate holds true for some constant $c_\alpha > 0$:

$$\left\|\nabla^\top R_{s,t}\phi\right\|_\infty \leq c_\alpha \left\|\phi\right\|_\infty \sup_{n=1,\ldots,N}\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n(t-s)}}}\,e^{-\lambda_n(t-s)}\right),$$
$$0 \leq s < t \leq T. \tag{2.6}$$

In the sequel, for every $x \in \mathbb{R}^N$ and $t \in (0,T]$ we are going to need the continuity of $R_{\cdot,t}\phi(x)$ in the interval $[0,t)$ [resp., in the closed interval $[0,t]$] when $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$ [resp., $\phi \in C_b\left(\mathbb{R}^N\right)$]. In order to prove this property, we first note that a variation of constants formula lets us consider, for $0 \leq s \leq t \leq T$ and $x \in \mathbb{R}^N$ (from (2.2))

$$Z_t^{s,x} = e^{(t-s)A}x + \int_s^t e^{(t-r)A} f(r)\, dr + \int_s^t e^{(t-r)A}\sqrt{Q}\, dW_{L_r}. \tag{2.7}$$

This expression shows that the process $(Z_t^{s,x})_{s\in[0,t]}$ is stochastically continuous (in the variable s). As a consequence, if $\phi \in C_b\left(\mathbb{R}^N\right)$, then we can easily deduce the continuity of $R_{\cdot,t}\phi(x)$ in $[0,t]$ applying the continuous mapping and Vitali's convergence theorems to (2.3). In the general case $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$, one can use the same argument combined with the regularizing property of R and (2.4) to obtain the continuity of $R_{\cdot,t}\phi(x)$ in $[0,t)$, as desired.

Finally, observe that there exists a constant $C = C(\alpha, A, Q) > 0$ such that

$$c_\alpha \sup_{n=1,\ldots,N}\left(\frac{1}{\sigma_n}\sqrt[2\alpha]{\frac{2\alpha\lambda_n}{1-e^{-2\alpha\lambda_n(t-s)}}}\,e^{-\lambda_n(t-s)}\right) \leq C\frac{1}{(t-s)^{1/(2\alpha)}}, \quad 0 \leq s < t \leq T.$$

We refer to [34, Remark 5] (see Remark 1.5) for a similar computation. Let us assume $\alpha \in \left(\frac{1}{2}, 1\right)$: in this way, denoting by $\gamma = 1/(2\alpha)$, we have $\gamma \in (0,1)$ and the bound in (2.6) entails

$$\left\| \nabla^\top R_{s,t}\phi \right\|_\infty \leq C \left\|\phi\right\|_\infty \frac{1}{(t-s)^\gamma}, \quad 0 \leq s < t \leq T,\, \phi \in \mathcal{B}_b\left(\mathbb{R}^N\right). \tag{2.8}$$

For a given measurable and bounded vector field $B\colon [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$, we are concerned with the analysis of the following *Kolmogorov backward equation* in mild, integral form:

$$\begin{aligned} u_s^\phi(t,x) &= R_{s,t}\phi(x) + \int_s^t R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top u_r^\phi(t,\cdot)\right\rangle\right)(x)\, dr, \\ &s \in [0,t],\, x \in \mathbb{R}^N, \end{aligned} \tag{2.9}$$

where $t \in (0,T]$ and $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$. We denote by $\left\|B\right\|_{0,T} = \sup_{0\leq t\leq T}\left\|B(t,\cdot)\right\|_\infty$. In order to study (2.9), for every $0 < t_1 < t_2 \leq T$, we consider the Banach space $\left(\Lambda_1^\gamma[t_1,t_2], \left\|\cdot\right\|_{\Lambda_1^\gamma[t_1,t_2]}\right)$ defined by

$$\begin{aligned} \Lambda_1^\gamma[t_1,t_2] = \Big\{ &V\colon [t_1,t_2]\times\mathbb{R}^N \to \mathbb{R} \text{ measurable} : V(\cdot,x) \in C([t_1,t_2]),\, x \in \mathbb{R}^N; \\ &V(s,\cdot) \in C_b^1\left(\mathbb{R}^N\right),\, s \in [t_1,t_2];\, \sup_{s\in[t_1,t_2]} s^\gamma \left\|V(s,\cdot)\right\|_1 < \infty \Big\}, \\ \left\|V\right\|_{\Lambda_1^\gamma[t_1,t_2]} = &\sup_{s\in[t_1,t_2]} s^\gamma \left\|V(s,\cdot)\right\|_1, \text{ where } \left\|V(s,\cdot)\right\|_1 = \left\|V(s,\cdot)\right\|_\infty + \left\|\partial^1 V(s,\cdot)\right\|_\infty. \end{aligned}$$

When $t_1 = 0$, we are careful to remove the left–end point of the interval $[t_1,t_2]$ in the previous definitions, so that we will be working with the space $\left(\Lambda_1^\gamma(0,t_2], \left\|\cdot\right\|_{\Lambda_1^\gamma(0,t_2]}\right)$. The following lemma proves the well–posedness of (2.9). We refer to [66, Theorem 9.38] for an analogous result concerning the *Kolmogorov forward equation* in mild form associated with OU processes in infinite dimension corresponding to Brownian motions.

Theorem 2.1. *Let $\alpha \in \left(\frac{1}{2},1\right)$ and $B\colon [0,T]\times\mathbb{R}^N \to \mathbb{R}^N$ be a measurable and bounded vector field. Then for every $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$ and $0 < t \leq T$, there exists a unique solution $u_s^\phi(t,x)$, $s \in [0,t]$, $x \in \mathbb{R}^N$, of (2.9) such that $u_{t-\diamond}^\phi(t,\cdot) \in \Lambda_1^\gamma(0,t]$, where $\gamma = 1/(2\alpha)$.*

Proof. Fix $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$, $t \in (0,T]$, $\bar{s} \in (0,t]$ and introduce the map $\Gamma_1\colon \Lambda_1^\gamma(0,\bar{s}] \to \Lambda_1^\gamma(0,\bar{s}]$ given by

$$\begin{aligned} \Gamma_1 V(s,x) &= R_{t-s,t}\phi(x) + \int_{t-s}^t R_{t-s,r}\left(\left\langle B(r,\cdot), \nabla^\top V(t-r,\cdot)\right\rangle\right)(x)\, dr, \\ &0 < s \leq \bar{s},\, x \in \mathbb{R}^N, \end{aligned} \tag{2.10}$$

for every $V \in \Lambda_1^\gamma(0,\bar{s}]$. Notice that such an application is well defined and with values in $\Lambda_1^\gamma(0,\bar{s}]$, thanks to the properties of R discussed above, the dominated convergence theorem and the next computations based on (2.8):

$$\begin{aligned} &\sup_{x\in\mathbb{R}^N}\left| \int_{t-s}^t \partial_{x_j} R_{t-s,r}\left(\left\langle B(r,\cdot), \nabla^\top V(t-r,\cdot)\right\rangle\right)(x)\, dr \right| \\ &\quad\leq NC\left\|B\right\|_{0,T}\left\|V\right\|_{\Lambda_1^\gamma(0,\bar{s}]} \int_{t-s}^t \frac{dr}{(r-(t-s))^\gamma (t-r)^\gamma} \\ &\quad\leq \frac{4^\gamma}{1-\gamma} NC\left\|B\right\|_{0,T}\left\|V\right\|_{\Lambda_1^\gamma(0,\bar{s}]} s^{1-2\gamma}, \quad 0 < s \leq \bar{s},\, j = 1,\ldots,N. \end{aligned} \tag{2.11}$$

Here $C = C(\alpha, A, Q) > 0$ is the same constant as in (2.8), and the last inequality is obtained using the bound

$$\int_{t-s}^{t} \frac{dr}{(r-(t-s))^{\gamma}(t-r)^{\gamma}} = \left\{ \int_{t-s}^{t-\frac{s}{2}} + \int_{t-\frac{s}{2}}^{t} \right\} \frac{dr}{(r-(t-s))^{\gamma}(t-r)^{\gamma}}$$
$$= 2\int_{t-s}^{t-\frac{s}{2}} \frac{dr}{(r-(t-s))^{\gamma}(t-r)^{\gamma}} \leq \frac{2}{1-\gamma}\left(\frac{2}{s}\right)^{\gamma}\left(\frac{s}{2}\right)^{1-\gamma} = \frac{4^{\gamma}}{1-\gamma}s^{1-2\gamma}, \quad (2.12)$$

where for the second equality we perform the substitution $u = 2t - s - r$. Estimates similar to those in (2.11) allow to write, for every $V_1, V_2 \in \Lambda_1^{\gamma}(0, \bar{s}]$,

$$\sup_{x\in\mathbb{R}^N} |(\Gamma_1 V_1 - \Gamma_1 V_2)(s,x)| + \sup_{x\in\mathbb{R}^N} \left|\partial_{x_j}(\Gamma_1 V_1 - \Gamma_1 V_2)(s,x)\right|$$
$$\leq \frac{4^{\gamma}}{1-\gamma} N \|B\|_{0,T}\left(s^{1-\gamma} + Cs^{1-2\gamma}\right)\|V_1 - V_2\|_{\Lambda_1^{\gamma}(0,\bar{s}]}, \quad 0 < s \leq \bar{s},\ j = 1,\dots,N.$$

Hence we obtain

$$\|\Gamma_1 V_1 - \Gamma_1 V_2\|_{\Lambda_1^{\gamma}(0,\bar{s}]} \leq \left[\frac{4^{\gamma}}{1-\gamma} N \|B\|_{0,T}\left(\bar{s} + C\bar{s}^{1-\gamma}\right)\right]\|V_1 - V_2\|_{\Lambda_1^{\gamma}(0,\bar{s}]}. \quad (2.13)$$

This shows that, for $\bar{s}$ sufficiently small, the map Γ_1 is a contraction in $\Lambda_1^{\gamma}(0, \bar{s}]$: we denote by $\overline{V}_1$ its unique fixed point. Now define

$$u_s^{\phi}(t,x) = R_{s,t}\phi(x) + \int_s^t R_{s,r}\left(\left\langle B(r,\cdot), \nabla^{\top}\overline{V}_1(t-r,\cdot)\right\rangle\right)(x)\,dr,$$
$$t - \bar{s} \leq s \leq t,\ x \in \mathbb{R}^N, \quad (2.14)$$

and notice that $u_{t-s}^{\phi}(t,x) = \overline{V}_1(s,x)$, $0 < s \leq \bar{s}$, $x \in \mathbb{R}^N$. Therefore $u_{\diamond}^{\phi}(t,\cdot)$ is the unique, local solution of (2.9) (in the strip $[t-\bar{s}, t] \times \mathbb{R}^N$) such that $u_{t-\diamond}^{\phi}(t,\cdot) \in \Lambda_1^{\gamma}(0,\bar{s}]$.

At this point, we can repeat the same procedure to construct the solution of (2.9) in the interval $[t - 2\bar{s}, t - \bar{s}]$, because the relation among constants in (2.13) –which is necessary to get a contraction– does not depend on the initial condition. Specifically, we take $\phi_1 = u_{t-\bar{s}}^{\phi}(t,\cdot) \in C_b^1(\mathbb{R}^N)$ and define the map

$$\Gamma_2 V(s,x) = R_{t-s,t-\bar{s}}\,\phi_1(x) + \int_{t-s}^{t-\bar{s}} R_{t-s,r}\left(\left\langle B(r,\cdot), \nabla^{\top}V(t-r,\cdot)\right\rangle\right)(x)\,dr,$$
$$\bar{s} \leq s \leq 2\bar{s},\ x \in \mathbb{R}^N,$$

for every $V \in \Lambda_1^{\gamma}[\bar{s}, 2\bar{s}]$. Computations analogous to the ones in the previous step show that the mapping $\Gamma_2\colon \Lambda_1^{\gamma}[\bar{s}, 2\bar{s}] \to \Lambda_1^{\gamma}[\bar{s}, 2\bar{s}]$ is a contraction: its unique fixed point is denoted by $\overline{V}_2$. Then we call

$$u_s^{\phi_1}(t-\bar{s},x) = R_{s,t-\bar{s}}\phi_1(x) + \int_s^{t-\bar{s}} R_{s,r}\left(\left\langle B(r,\cdot), \nabla^{\top}\overline{V}_2(t-r,\cdot)\right\rangle\right)(x)\,dr,$$
$$t - 2\bar{s} \leq s \leq t - \bar{s},\ x \in \mathbb{R}^N;$$

notice that $u_{t-s}^{\phi_1}(t-\bar{s},x)=\overline{V}_2(s,x)$, $\bar{s}\le s\le 2\bar{s}$, $x\in\mathbb{R}^N$, and that by the definition of ϕ_1, one has $u_{t-\bar{s}}^{\phi_1}(t-\bar{s},\cdot)=u_{t-\bar{s}}^{\phi}(t,\cdot)$. Now we extend the function $u_s^{\phi}(t,x)$ in (2.14) assigning

$$\bar{u}_s^{\phi}(t,x)=\begin{cases}u_s^{\phi}(t,x), & t-\bar{s}\le s\le t\\ u_s^{\phi_1}(t-\bar{s},x), & t-2\bar{s}\le s\le t-\bar{s}\end{cases},\quad x\in\mathbb{R}^N.$$

By the Chapman–Kolmogorov equations and Fubini's theorem we realize that $\bar{u}_\diamond^{\phi}(t,\cdot)$ is the unique local solution of (2.9) (in the strip $[t-2\bar{s},t]\times\mathbb{R}^N$) such that $\bar{u}_{t-\diamond}^{\phi}(t,\cdot)\in\Lambda_1^{\gamma}(0,2\bar{s}]$. In the sequel, we can simply denote it by $u_\diamond^{\phi}(t,\cdot)$.

This argument by steps of length $\bar{s}$ can be repeated iteratively to cover the whole interval $[0,t]$ and obtain the unique, global solution $u_\diamond^{\phi}(t,\cdot)$ of (2.9) such that $u_{t-\diamond}^{\phi}(t,\cdot)\in\Lambda_1^{\gamma}(0,t]$. Thus, the proof is complete. ■

If $\phi\in C_b^1(\mathbb{R}^N)$, then by (2.7) one can directly write $\nabla R_{s,t}\phi(x)=\mathbb{E}[\nabla\phi(Z_t^{s,x})]e^{(t-s)A}$. Next, considering that $\left|e^{(t-s)A}\right|\le 1$, $0\le s\le t\le T$, an application of (2.5)-(2.8) shows that $R_{s,t}\phi\in C_b^2(\mathbb{R}^N)$, with

$$\left\|\partial^2 R_{s,t}\phi\right\|_\infty\le C\left\|\partial^1\phi\right\|_\infty\frac{1}{(t-s)^{\gamma}},\quad 0\le s<t\le T,$$

where $C=C(\alpha,A,Q)>0$ is the same constant as in (2.8). This argument can be iterated to claim that, given an integer $n\ge 2$ and $\phi\in C_b^{n-1}(\mathbb{R}^N)$, $R_{s,t}\phi\in C_b^n(\mathbb{R}^N)$ and

$$\left\|\partial^n R_{s,t}\phi\right\|_\infty\le C\left\|\partial^{n-1}\phi\right\|_\infty\frac{1}{(t-s)^{\gamma}},\quad 0\le s<t\le T. \tag{2.15}$$

The previous consideration allows to extend Theorem 2.1. To this purpose, for an integer $n\ge 2$ and $0<t_1<t_2\le T$ we introduce the Banach space $\left(\Lambda_n^{\gamma}[t_1,t_2],\|V\|_{\Lambda_n^{\gamma}[t_1,t_2]}\right)$ defined by

$$\Lambda_n^{\gamma}[t_1,t_2]=\Big\{V\colon [t_1,t_2]\times\mathbb{R}^N\to\mathbb{R}\text{ measurable}: V(\cdot,x)\in C([t_1,t_2]),\ x\in\mathbb{R}^N;$$
$$V(s,\cdot)\in C_b^n(\mathbb{R}^N),\ s\in[t_1,t_2];\ \sup_{s\in[t_1,t_2]}s^{\gamma}\|V(s,\cdot)\|_n<\infty\Big\},$$

$$\|V\|_{\Lambda_n^{\gamma}[t_1,t_2]}=\sup_{s\in[t_1,t_2]}s^{\gamma}\|V(s,\cdot)\|_n,\text{ where }\|V(s,\cdot)\|_n=\|V(s,\cdot)\|_\infty+\sum_{j=1}^{n}\left\|\partial^j V(s,\cdot)\right\|_\infty.$$

As we have done before, when $t_1=0$ we remove the left–end point of $[t_1,t_2]$.

Corollary 2.2. *Let $\alpha\in\left(\frac{1}{2},1\right)$, $n\ge 2$ be an integer and $B\in C_b^{0,n-1}([0,T]\times\mathbb{R}^N;\mathbb{R}^N)$. Then for every $\phi\in C_b^{n-1}(\mathbb{R}^N)$ and $0<t\le T$, there exists a unique solution $u_s^{\phi}(t,x)$, $s\in[0,t]$, $x\in\mathbb{R}^N$, of (2.9) such that $u_{t-\diamond}^{\phi}(t,\cdot)\in\Lambda_n^{\gamma}(0,t]$, where $\gamma=1/(2\alpha)$.*

Proof. Take an integer $n\ge 2$; the argument parallels the one in the proof of Theorem 2.1, so here we only show that, for a given $\phi\in C_b^{n-1}(\mathbb{R}^N)$ and $\bar{s}\in(0,t]$ sufficiently small, the map $\Gamma_1\colon\Lambda_n^{\gamma}(0,\bar{s}]\to\Lambda_n^{\gamma}(0,\bar{s}]$ in (2.10) is well defined and a contraction. First, we note that for every $V\in\Lambda_n^{\gamma}(0,\bar{s}]$ and multi–index $\mathbf{j}$ such that $1\le\|\mathbf{j}\|_1\le n$,

$$\partial_{\mathbf{j}}\Gamma_1 V(s,x)=\partial_{\mathbf{j}}R_{t-s,t}\phi(x)+\int_{t-s}^{t}\partial_{\mathbf{j}}R_{t-s,r}\left(\left\langle B(r,\cdot),\nabla^{\top}V(t-r,\cdot)\right\rangle\right)(x)\,dr,$$
$$0<s\le\bar{s},\ x\in\mathbb{R}^N,$$

and that $\sup_{s\in(0,\bar{s}]} s^{\gamma} \left\|\partial^{\|\mathbf{j}\|_1} R_{t-s,t}\phi\right\|_{\infty} < \infty$ by (2.15). Secondly, invoking the estimates in (2.12) and (2.15), for every $0 < s \le \bar{s}$,

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^N} \left| \int_{t-s}^{t} \partial_{\mathbf{j}} R_{t-s,r} \left(\left\langle B(r,\cdot), \nabla^{\top} V(t-r,\cdot) \right\rangle \right)(x)\, dr \right| \\
&\quad \le N C_n C \|B\|_{n-1,T} \|V\|_{\Lambda_n^{\gamma}(0,\bar{s}]} \int_{t-s}^{t} \frac{dr}{(r-(t-s))^{\gamma} (t-r)^{\gamma}} \\
&\quad \le \frac{4^{\gamma}}{1-\gamma} N C_n C \|B\|_{n-1,T} \|V\|_{\Lambda_n^{\gamma}(0,\bar{s}]} s^{1-2\gamma}, \quad C_n = \binom{n-1}{[2^{-1}(n-1)]},
\end{aligned}$$

where $\|B\|_{n-1,T} = \sup_{0\le t\le T} \left(\|B(t,\cdot)\|_{\infty} + \sum_{j=1}^{n-1} \left\|\partial^{j} B(t,\cdot)\right\|_{\infty} \right)$ and $C = C(\alpha, A, Q) > 0$ is the same constant as in (2.8). It then follows that $\Gamma_1 V \in \Lambda_n^{\gamma}(0,\bar{s}]$, with ($V_1, V_2 \in \Lambda_n^{\gamma}(0,\bar{s}]$)

$$\|\Gamma_1 V_1 - \Gamma_1 V_2\|_{\Lambda_n^{\gamma}(0,\bar{s}]} \le \left[\frac{4^{\gamma}}{1-\gamma} N \|B\|_{n-1,T} \left(\bar{s} + n C_n C \bar{s}^{1-\gamma} \right) \right] \|V_1 - V_2\|_{\Lambda_n^{\gamma}(0,\bar{s}]},$$

which reduces to (2.13) when $n = 1$ and proves the contraction property of Γ_1 for $\bar{s}$ small enough. ■

2.2 The time–dependent Markov transition semigroup

Let $\alpha \in (0,1)$ and $B_0\colon [0,T]\times\mathbb{R}^N \to \mathbb{R}^N$ be a vector field such that $B_0 \in C_b^{0,1}\left([0,T]\times\mathbb{R}^N;\mathbb{R}^N\right)$. For every $x \in \mathbb{R}^N$ and $0 \le s \le T$, we define the process $X^{s,x} = (X_t^{s,x})_{t\in[s,T]}$ to be the unique (up to indistinguishability) solution of the semilinear stochastic differential equation

$$dX_t^{s,x} = \left(AX_t^{s,x} + B_0\left(t, X_t^{s,x}\right)\right) dt + \sqrt{Q}\, dW_{L_t}, \quad X_s^{s,x} = x \in \mathbb{R}^N. \tag{2.16}$$

We denote by $P = (P_{s,t})$, $0 \le s \le t \le T$, the corresponding time–dependent Markov transition semigroup given by

$$P_{s,t}\phi = \mathbb{E}\left[\phi\left(X_t^{s,\cdot}\right)\right], \quad \phi \in \mathcal{B}_b\left(\mathbb{R}^N\right).$$

The connection between the SDE in (2.16) and the Kolmogorov backward equation in mild integral form (2.9) is provided by the next, fundamental result.

Theorem 2.3. *Let $\alpha \in \left(\frac{1}{2}, 1\right)$, $B_0 \in C_b^{0,3}\left([0,T]\times\mathbb{R}^N;\mathbb{R}^N\right)$, $f \in C\left([0,T];\mathbb{R}^N\right)$ and define $B = B_0 - f$. Then, for every $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$ and $0 < t \le T$, the function $P_{s,t}\phi(x)$, $0 \le s \le t$, $x \in \mathbb{R}^N$, is the unique solution of (2.9) such that $P_{t-\diamond,t}\phi(\cdot) \in \Lambda_1^{\gamma}(0,t]$, where $\gamma = 1/(2\alpha)$.*

The purpose of this section is to develop a self–contained procedure which is specific to our framework and allows to prove Theorem 2.3 via important, preliminary results. In the case of time–independent nonlinearities and $f \equiv 0$ (hence for Kolmogorov forward equations in mild form), Theorem 2.3 is known for noises different from our W_L. As regards independent $\alpha-$stable Lévy processes in finite dimension, it has been established in [151, Lemma 5.12] (its proof relies on the theory of one–parameter semigroups, so it cannot be adapted to our framework). As for Brownian motions in infinite dimension, we refer to [66, Theorem 9.43].

Let $\alpha \in (0,1)$, $B_0 \in C_b^{0,1}\left([0,T]\times\mathbb{R}^N;\mathbb{R}^N\right)$ and recall that the subordinated Brownian motion W_L is an isotropic (i.e., rotation–invariant), $2\alpha-$stable, $\mathbb{R}^N-$valued Lévy process with compensator $\nu(dz) \asymp |z|^{-(N+2\alpha)}\, dz$ and no continuous martingale part (see [164, Theorem 30.1]). Here $\asymp$ denotes the equality up to a positive multiplicative constant. By [150, Theorem 3.1] (see also [149]) there is a sharp stochastic flow $X_t^{s,x}$ generated by the SDE (2.16) which is jointly measurable in (s,t,x,ω) and, $\mathbb{P}-$a.s., simultaneously continuous in x and càdlàg in s and t. More specifically, there exists an almost–sure event Ω' such that the following facts hold true for every $\omega \in \Omega'$:

- for every $x \in \mathbb{R}^N$ and $t \in [0,T]$, the mapping $s \mapsto X_t^{s,x}(\omega)$ is càdlàg in $[0,t]$;
- for every $x \in \mathbb{R}^N$ and $s \in [0,T]$, the mapping $t \mapsto X_t^{s,x}(\omega)$ is càdlàg in $[s,T]$;
- for every $0 \le s \le t \le T$, the mapping $x \mapsto X_t^{s,x}(\omega)$ is continuous in $\mathbb{R}^N$;
- the flow property is satisfied, namely $X_t^{s,x}(\omega) = X_t^{r,X_r^{s,x}(\omega)}(\omega)$ for every $x \in \mathbb{R}^N$, $0 \le s < r < t \le T$;
- for every $x \in \mathbb{R}^N$ and $0 \le s \le t \le T$, $X_t^{s,x}(\omega) = x + \int_s^t \left(AX_r^{s,x}(\omega) + B_0\left(r, X_r^{s,x}(\omega)\right)\right) dr$ $+ \sqrt{Q}\left(W_{L_t} - W_{L_s}\right)(\omega)$.

For every $\omega \in \Omega \setminus \Omega'$, we set $X_t^{s,x}(\omega) = x$, $(s,\, t) \in [0,T]^2$, $x \in \mathbb{R}^N$: from now on, we work with such a stochastic flow $X_t^{s,x}$. The next result shows that, under additional regularity requirements on B_0, it is differentiable with respect to x. Analogous claims concerning differentiability of stochastic flows can be found in literature in, e.g., [61, Theorem 8.18] for the Brownian case and in [126, Theorem 3.4.2] for the jumps one, although the latter requires regularity assumptions on the coefficients which are not fulfilled by our framework. The proof, which carries out a path–by–path argument thanks to the already mentioned properties guaranteed by [150], is postponed to Appendix 2.A.

Lemma 2.4. *Let $\alpha \in (0,1)$, $n \ge 2$ be an integer and $B_0 \in C_b^{0,n}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$. Then for every $\omega \in \Omega$ and $0 \le s \le t \le T$, the function $x \mapsto X_t^{s,x}(\omega)$ belongs to $C^n\left(\mathbb{R}^N\right)$, and there exists a constant $C > 0$ depending only on A, B_0, T, n and N such that*

$$\sum_{i=1}^n \left\|\partial^i X_t^{s,\cdot}(\omega)\right\|_\infty \le C, \quad 0 \le s \le t \le T,\, \omega \in \Omega. \tag{2.17}$$

The previous claim implies the following result regarding persistence of regularity.

Corollary 2.5. *Let $\alpha \in (0,1)$, $n \ge 2$ be an integer and $\phi \in C_b^n\left(\mathbb{R}^N\right)$. If $B_0 \in C_b^{0,n}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$, then for every $0 \le s \le t \le T$ the function $P_{s,t}\phi \in C_b^n\left(\mathbb{R}^N\right)$. In addition,*

$$\sup_{0 \le s \le t \le T} \left(\left\|P_{s,t}\phi\right\|_\infty + \sum_{i=1}^n \left\|\partial^i P_{s,t}\phi\right\|_\infty \right) < \infty. \tag{2.18}$$

Let $D = \left\{z \in \mathbb{R}^N,\, |z| \le 1\right\}$; we introduce the family of integro–differential operators $\left(A(s)\right)_{0 \le s \le T}$, defined on every $\psi \in C_b^2\left(\mathbb{R}^N\right)$ by

$$\begin{aligned} A(s)\,\psi(x) = {} & \left\langle Ax + B_0(s,x), \nabla^\top \psi(x) \right\rangle \\ & + \int_{\mathbb{R}^N} \left[\psi\left(x + \sqrt{Q}z\right) - \psi(x) - 1_D(z)\, \nabla\psi(x)\, \sqrt{Q}z\right] \nu(dz), \end{aligned} \tag{2.19}$$

where $x \in \mathbb{R}^N$. We need the next preparatory result.

Lemma 2.6. *(i) Let $\alpha \in \left(\frac{1}{2}, 1\right)$, $0 \le s \le T$ and $x \in \mathbb{R}^N$. If $B_0 \in C_b^{0,1}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$, then the mapping $r \mapsto P_{s,r} A(r) \phi(x)$ is continuous in $[s,T]$ for every $\phi \in C_b^2\left(\mathbb{R}^N\right)$;*

(ii) Let $\alpha \in (0,1)$ and $0 \le t \le T$. If $B_0 \in C_b^{0,3}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$, then for every $r \in [0,t]$ and $\phi \in C_b^3\left(\mathbb{R}^N\right)$ the mapping $x \mapsto A(r) P_{r,t}\phi(x)$ belongs to $C^1\left(\mathbb{R}^N\right)$.
Moreover, $\sup_{r\in[0,t]} \left\|1_B \nabla^\top A(r) P_{r,t}\phi\right\|_\infty < \infty$ for every bounded set $B \subset \mathbb{R}^N$.

Proof. We start off by proving Point (i). Fix $0 \le s \le T$ and $x \in \mathbb{R}^N$; from (2.16), Gronwall's lemma, [134, Theorem 3.2] and the continuity in probability of the Lévy process W_L we deduce that $\mathbb{E}\left[\sup_{t\in[s,T]} |X_t^{s,x}|^p\right] < \infty$ for every $p \in (1, 2\alpha)$, and that the process $X^{s,x}$ is stochastically continuous in $[s,T]$, as well. Consider $r \in [s,T]$ and a sequence $(r_n)_n \subset [s,T]$ such that $r_n \to r$ as $n \to \infty$. Given $\phi \in C_b^2\left(\mathbb{R}^N\right)$,

$$
\begin{aligned}
&P_{s,r_n} A(r_n) \phi(x) - P_{s,r} A(r) \phi(x) \\
&\qquad = P_{s,r_n}\left(A(r_n)\phi - A(r)\phi\right)(x) + \left(P_{s,r_n} A(r) \phi(x) - P_{s,r} A(r) \phi(x)\right) =: \mathbf{I}_n + \mathbf{II}_n.
\end{aligned}
$$

Since (2.19) entails $\left(A(r_n)\phi - A(r)\phi\right)(\cdot) = \left\langle B_0(r_n, \cdot) - B_0(r, \cdot), \nabla^\top \phi(\cdot)\right\rangle$ we have, by Vitali's and dominated convergence theorems,

$$
|\mathbf{I}_n| \le \left\|\nabla^\top \phi\right\|_\infty \left(2\left\|DB_0\right\|_{T,\infty} \mathbb{E}\left[\left|X_{r_n}^{s,x} - X_r^{s,x}\right|\right] + \mathbb{E}\left[\left|B_0(r_n, X_r^{s,x}) - B_0(r, X_r^{s,x})\right|\right]\right) \underset{n\to\infty}{\longrightarrow} 0,
$$

where we denote by $\|DB_0\|_{T,\infty} = \sup_{0\le t\le T} \|DB_0(t,\cdot)\|_\infty$. As for $\mathbf{II}_n$, note that $A(r)\phi$ is continuous in $\mathbb{R}^N$, and that for every $y \in \mathbb{R}^N$ (see (2.19)),

$$
\begin{aligned}
|A(r)\phi(y)| \le{}& \left\|\nabla^\top \phi\right\|_\infty \left(|A|\,|y| + \|B_0\|_{0,T}\right) \\
&+ \frac{1}{2}\left\|D^2\phi\right\|_\infty \int_{\mathbb{R}^N} 1_D(z) \left|\sqrt{Q}z\right|^2 \nu(dz) + 2\|\phi\|_\infty \int_{\mathbb{R}^N} 1_{D^c}(z)\,\nu(dz). \qquad (2.20)
\end{aligned}
$$

Therefore by the continuous mapping and Vitali's convergence theorem we obtain $\mathbf{II}_n \to 0$ as $n \to \infty$, proving Point (i).

We now move on to Point (ii), where it is sufficient to require $\alpha \in (0,1)$. Fix $0 \le r \le t \le T$; observe that for every $\psi \in C_b^3\left(\mathbb{R}^N\right)$ one has $A(r)\psi \in C^1\left(\mathbb{R}^N\right)$, with

$$
\begin{aligned}
\nabla A(r)\psi(x) ={}& \nabla\psi(x)\left(A + DB_0(r,x)\right) + \left(Ax + B_0(r,x)\right)^\top D^2\psi(x) \\
&+ \int_{\mathbb{R}^N} \left[\nabla\psi\left(x + \sqrt{Q}z\right) - \nabla\psi(x) - 1_D(z)\left(\sqrt{Q}z\right)^\top D^2\psi(x)\right]\nu(dz), \quad x \in \mathbb{R}^N.
\end{aligned}
$$

More specifically, in the previous computation we are allowed to differentiate under the integral sign because ($x \in \mathbb{R}^N$, $z \in D$)

$$
\left|\nabla^\top\psi\left(x + \sqrt{Q}z\right) - \nabla^\top\psi(x) - D^2\psi(x)\sqrt{Q}z\right| \le \frac{1}{2} N^{\frac{3}{2}} \left\|\partial^3\psi\right\|_\infty \left|\sqrt{Q}z\right|^2.
$$

The hypotheses prescribe $B_0 \in C_b^{0,3}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$ and $\phi \in C_b^3\left(\mathbb{R}^N\right)$, hence it is sufficient to invoke Corollary 2.5 to complete proof. ■

We are now in position to prove the following, crucial result concerning Kolmogorov equations (cfr. [126, Theorem 4.5.1] for an analogous claim in a different setting).

Theorem 2.7. *Take* $\alpha \in \left(\frac{1}{2}, 1\right)$.

(*i*) *Let* $0 \le s \le T$ *and* $x \in \mathbb{R}^N$. *If* $B_0 \in C_b^{0,1}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$ *and* $\phi \in C_b^2\left(\mathbb{R}^N\right)$, *then the function* $t \mapsto P_{s,t}\phi(x)$ *is continuously differentiable in* $[s,T]$ *and satisfies the Kolmogorov forward equation*

$$\partial_t P_{s,t}\phi(x) = P_{s,t}A(t)\phi(x); \tag{2.21}$$

(*ii*) *Let* $0 \le t \le T$ *and* $x \in \mathbb{R}^N$. *If* $B_0 \in C_b^{0,3}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$ *and* $\phi \in C_b^3\left(\mathbb{R}^N\right)$, *then the function* $s \mapsto P_{s,t}\phi(x)$ *is continuously differentiable in* $[0,t]$ *and satisfies the Kolmogorov backward equation*

$$\partial_s P_{s,t}\phi(x) = -A(s)P_{s,t}\phi(x). \tag{2.22}$$

Proof. Recall that by [164, Theorem 14.7 (iii)] the process W_L is centered in 0 when $\alpha \in \left(\frac{1}{2}, 1\right)$. As a consequence, denoting by N the Poisson random measure associated with its jumps and by $\widetilde{N}$ the compensated measure, $W_L = \int_0^\cdot \int_{\mathbb{R}^N} 1_D(z)\, z\widetilde{N}(ds, dz) + \int_0^\cdot \int_{\mathbb{R}^N} 1_{D^c}(z)\, zN(ds, dz)$ up to indistinguishability by [110, Theorem 2.34, Chapter II].

As for Point (*i*), take $0 \le s \le T$, $x \in \mathbb{R}^N$ and $\phi \in C_b^2\left(\mathbb{R}^N\right)$; by (2.16) an application of Itô formula ensures that

$$\begin{aligned}\phi\left(X_t^{s,x}\right) = \phi(x) &+ \int_s^t \left\langle AX_r^{s,x} + B_0\left(r, X_r^{s,x}\right), \nabla^\top \phi\left(X_r^{s,x}\right)\right\rangle dr \\ &+ \int_s^t \int_{\mathbb{R}^N} 1_D(z)\nabla\phi\left(X_{r-}^{s,x}\right)\sqrt{Q}z\,\widetilde{N}(dr, dz) \\ &+ \int_s^t \int_{\mathbb{R}^N} \left(\phi\left(X_{r-}^{s,x} + \sqrt{Q}z\right) - \phi\left(X_{r-}^{s,x}\right) - 1_D(z)\nabla\phi\left(X_{r-}^{s,x}\right)\sqrt{Q}z\right) N(dr, dz),\end{aligned}$$

which holds true $\mathbb{P}-$a.s. for every $t \in [s,T]$. Taking expectations in the previous equation and using Fubini's theorem we obtain

$$P_{s,t}\phi(x) = \phi(x) + \int_s^t \mathbb{E}\left[A(r)\phi\left(X_r^{s,x}\right)\right] dr = \phi(x) + \int_s^t P_{s,r}A(r)\phi(x)\,dr, \quad t \in [s,T],$$

which in turn implies (2.21) by Lemma 2.6 (*i*).

We now focus on Point (*ii*). Take $0 \le t \le T$ and $x \in \mathbb{R}^N$; arguing as in [126, Proposition 3.8.2] we see that $X_t^{\cdot,x}$ follows the backward dynamics ($\mathbb{P}-$a.s.)

$$\begin{aligned}X_t^{s,x} = x + \int_s^t DX_t^{r,x}\left(Ax + B_0(r,x)\right) dr + \int_s^t \int_{\mathbb{R}^N} \left(X_t^{r,x+\sqrt{Q}z} - X_t^{r,x}\right)\widetilde{N}(dr, dz) \\ + \int_s^t \int_{\mathbb{R}^N} \left[X_t^{r,x+\sqrt{Q}z} - X_t^{r,x} - 1_D(z)DX_t^{r,x}\sqrt{Q}z\right]\nu(dz)\,dr, \quad s \in [0,t].\end{aligned}$$

Hence invoking the backward Itô formula (see, e.g., [126, Theorem 2.7.1]) we deduce that, for every $\phi \in C_b^2\left(\mathbb{R}^N\right)$ and $s \in [0,t]$,

$$\begin{aligned}\phi\left(X_t^{s,x}\right) = \phi(x) &+ \int_s^t \int_{\mathbb{R}^N} \left(\phi\left(X_t^{r,x+\sqrt{Q}z}\right) - \phi\left(X_t^{r,x}\right)\right)\widetilde{N}(dr, dz) \\ &+ \int_s^t \nabla\phi\left(X_t^{r,x}\right) DX_t^{r,x}\left(Ax + B_0(r,x)\right) dr \\ &+ \int_s^t \int_{\mathbb{R}^N} \left[\phi\left(X_t^{r,x+\sqrt{Q}z}\right) - \phi\left(X_t^{r,x}\right) - 1_D(z)\nabla\phi\left(X_t^{r,x}\right) DX_t^{r,x}\sqrt{Q}z\right]\nu(dz)\,dr,\end{aligned}$$

which holds true $\mathbb{P}$−a.s. Taking expectations in the previous equation and using Fubini's theorem (remember Lemma 2.4) we obtain

$$P_{s,t}\phi(x) = \phi(x) + \int_s^t A(r) P_{r,t}\phi(x)\,dr, \quad s \in [0,t]. \tag{2.23}$$

Since by hypotheses we are working with $\phi \in C_b^3(\mathbb{R}^N)$ and $B_0 \in C_b^{0,3}([0,T]\times\mathbb{R}^N;\mathbb{R}^N)$, by Lemma 2.6 (*ii*) we can differentiate in x the expression in (2.23), showing the continuity of the mapping $r \mapsto \nabla P_{r,t}\phi(x)$ in $[0,t]$. This, together with (2.19), the fact that (2.23) also provides the continuity of the mapping $r \mapsto P_{r,t}\phi(x)$ in $[0,t]$ and a dominated convergence argument based on Corollary 2.5, ensures the continuity of the function $r \mapsto A(r) P_{r,t}\phi(x)$ in the same interval. Therefore differentiating (2.23) with respect to s we infer (2.22). The proof is now complete. ■

Another step that we need to prove Theorem 2.3 consists in a regularization result for the time–dependent Markov transition semigroup $P_{s,t}$ (see Lemma 2.10) which –at the best of our knowledge– is not established in literature with this type of noise. We start by recalling the *Bismut–Elworthy–Li's type formula* presented in [178, Theorem 1.1] (see also [175] for a related work treating multiplicative Lévy noise); such a formula is adapted to our framework, where we have to account for an initial time s not necessarily equal to 0.

Theorem 2.8 ([178])**.** *Let $\alpha \in \left(\frac{1}{2},1\right)$ and $B_0 \in C_b^{0,1}([0,T]\times\mathbb{R}^N;\mathbb{R}^N)$. Then for every $0 \le s < t \le T$ and $\phi \in C_b^1(\mathbb{R}^N)$, the function $P_{s,t}\phi$ is differentiable at x in every direction $h \in \mathbb{R}^N$ and*

$$\left\langle \nabla^\top P_{s,t}\phi(x), h \right\rangle = \mathbb{E}\left[\frac{1}{L_t - L_s}\phi\left(X_t^{s,x}\right)\int_s^t \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x}, dW_{L_r} \right\rangle\right]. \tag{2.24}$$

Furthermore, there exists a constant $C_\alpha > 0$ such that the next gradient estimate holds true:

$$\left\|\nabla^\top P_{s,t}\phi\right\|_\infty \le C_\alpha \|\phi\|_\infty \left|\left(\sqrt{Q}\right)^{-1}\right| e^{\left(|A| + \|DB_0\|_{T,\infty}\right)T} \frac{1}{(t-s)^{1/(2\alpha)}}, \quad 0 \le s < t \le T. \tag{2.25}$$

We are able to extend the previous claim to functions $\phi \in C_b(\mathbb{R}^N)$ with an approximation procedure, effectively making Theorem 2.8 a regularization–by–noise result. We need the next estimate, which derives from [33, Equation (14)] (see also (1.17)):

$$\mathbb{E}\left[\frac{1}{L_t^p}\right]^{\frac{1}{p}} \le c\,t^{-\frac{1}{\alpha}}, \quad t > 0, \text{ for some } c = c(\alpha,p) > 0, \text{ for every } p > 0. \tag{2.26}$$

Corollary 2.9. *Let $\alpha \in \left(\frac{1}{2},1\right)$ and $B_0 \in C_b^{0,1}([0,T]\times\mathbb{R}^N;\mathbb{R}^N)$. Then, for every $\phi \in C_b(\mathbb{R}^N)$ and $0 \le s < t \le T$, the function $P_{s,t}\phi$ is differentiable at $x \in \mathbb{R}^N$ in every direction $h \in \mathbb{R}^N$, and the expression in* (2.24) *holds true.*

Proof. Fix $x,h \in \mathbb{R}^N, 0 \le s < t \le T$, and $\phi \in C_b(\mathbb{R}^N)$. Since $C_b^\infty(\mathbb{R}^N)$ is dense in $C_b(\mathbb{R}^N)$, we can take a sequence $(\phi_n)_n \subset C_b^\infty(\mathbb{R}^N)$ such that $\|\phi_n - \phi\|_\infty \to 0$ as $n \to \infty$. Denote by $g_n(u) = P_{s,t}\phi_n(x+uh)$, $u \in \mathbb{R}$; by dominated convergence, for every $u \in \mathbb{R}$,

$$g_n(u) \to P_{s,t}\phi(x+uh) =: g(u), \quad \text{as } n \to \infty.$$

Now we invoke (2.24) to write

$$g_n'(u) = \lim_{v\to 0} \frac{\mathbb{E}\left[\phi_n\left(X_t^{s,x+uh+vh}\right)\right] - \mathbb{E}\left[\phi_n\left(X_t^{s,x+uh}\right)\right]}{v} = \left\langle \nabla^\top P_{s,t}\phi_n(x+uh), h\right\rangle$$
$$= \mathbb{E}\left[\frac{1}{L_t - L_s}\phi_n\left(X_t^{s,x+uh}\right)\int_s^t \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x+uh}, dW_{L_r}\right\rangle\right], \quad u\in\mathbb{R}.$$

Since $\alpha \in \left(\frac{1}{2}, 1\right)$, an application of [178, Theorem 3.2], (2.26), Hölder's inequality with $p\in(1,2\alpha)$ and Lemma 2.4 (see (2.17)) let us compute, for every $u\in\mathbb{R}$,

$$\mathbb{E}\left[\left|\frac{1}{L_t-L_s}\left(\phi_n\left(X_t^{s,x+uh}\right) - \phi\left(X_t^{s,x+uh}\right)\right)\int_s^t \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x+uh}, dW_{L_r}\right\rangle\right|\right]$$
$$\leq \left|\left(\sqrt{Q}\right)^{-1}\right| |h| \frac{c_1}{(t-s)^{1/\alpha}} \|\phi_n - \phi\|_\infty \to 0, \quad \text{as } n\to\infty, \tag{2.27}$$

where $c_1 = c_1(\alpha, p, A, B_0, T, N) > 0$. It follows that

$$g_n' \to \mathbb{E}\left[\frac{1}{L_t-L_s}\phi\left(X_t^{s,x+(\cdot)h}\right)\int_s^t \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x+(\cdot)h}, dW_{L_r}\right\rangle\right], \quad \text{uniformly in } \mathbb{R}.$$

This suffices to obtain the desired result, hence the proof is complete. ■

Note that for every $\phi\in C_b\left(\mathbb{R}^N\right)$ the expression on the right–hand side of (2.24) is continuous in x for every $h\in\mathbb{R}^N$. Indeed, let us fix $x\in\mathbb{R}^N$ and consider $(x_n)_n \subset \mathbb{R}^N$ such that $x_n\to x$ as $n\to\infty$. Then, using the same techniques as in the previous proof (cfr. (2.27)), together with Lemma 2.4 and a dominated convergence argument, we get (for some $p, q > 1$ determined by a generalized Holder's inequality, and $c = c(\alpha, p, q, A, B_0, Q, T, N) > 0$)

$$\left|\mathbb{E}\left[\frac{1}{L_t-L_s}\left(\phi\left(X_t^{s,x_n}\right)\int_s^t\left\langle\left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x_n}, dW_{L_r}\right\rangle - \phi\left(X_t^{s,x}\right)\int_s^t\left\langle\left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x}, dW_{L_r}\right\rangle\right)\right]\right|$$
$$\leq \|\phi\|_\infty \mathbb{E}\left[\frac{1}{L_t-L_s}\left|\int_s^t\left\langle\left(\sqrt{Q}\right)^{-1}\left(D_h X_r^{s,x_n} - D_h X_r^{s,x}\right), dW_{L_r}\right\rangle\right|\right]$$
$$+ \mathbb{E}\left[\frac{1}{L_t-L_s}\left|\int_s^t\left\langle\left(\sqrt{Q}\right)^{-1} D_h X_r^{s,x}, dW_{L_r}\right\rangle\right|\left|\phi\left(X_t^{s,x_n}\right) - \phi\left(X_t^{s,x}\right)\right|\right]$$
$$\leq \frac{c}{(t-s)^{1/\alpha}} \times \left[\|\phi\|_\infty\left(\int_s^t \mathbb{E}\left[\left|D_h X_r^{s,x_n} - D_h X_r^{s,x}\right|^{2\alpha}\right]dr\right)^{\frac{1}{2\alpha}} + |h|\,\mathbb{E}\left[\left|\phi\left(X_t^{s,x_n}\right) - \phi\left(X_t^{s,x}\right)\right|^q\right]^{\frac{1}{q}}\right]$$
$$\underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, $P_{s,t}\phi\in C_b^1\left(\mathbb{R}^N\right)$ for every $\phi\in C_b\left(\mathbb{R}^N\right)$. At this point, the next result is a straightforward consequence of the Chapman–Kolmogorov equations, the mean value theorem and [64, Lemma 7.1.5].

Lemma 2.10. *Let $\alpha\in\left(\frac{1}{2},1\right)$ and $B_0 \in C_b^{0,1}\left([0,T]\times\mathbb{R}^N;\mathbb{R}^N\right)$. Then, for every $\phi\in\mathcal{B}_b\left(\mathbb{R}^N\right)$ and $0\leq s<t\leq T$, one has $P_{s,t}\phi\in C_b^1\left(\mathbb{R}^N\right)$, and the gradient estimate in* (2.25) *holds true.*

Finally we are in position to prove Theorem 2.3.

Proof of Theorem 2.3. Fix $\alpha \in \left(\frac{1}{2}, 1\right)$, $0 < t \leq T$, $f \in C\left([0,T]; \mathbb{R}^N\right)$ and $B_0 \in C_b^{0,3}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$. Moreover, define $B = B_0 - f$. We first consider $\phi \in C_b^3\left(\mathbb{R}^N\right)$. Recalling (2.2), we introduce the family of integro–differential operators $\left(\widetilde{A}(s)\right)_{0\leq s\leq T}$, defined for every $\psi \in C_b^2\left(\mathbb{R}^N\right)$ by

$$\widetilde{A}(s)\psi(x) = \left\langle Ax + f(s), \nabla^\top \psi(x)\right\rangle + \int_{\mathbb{R}^N} \left[\psi\left(x + \sqrt{Q}z\right) - \psi(x) - 1_D(z)\nabla\psi(x)\sqrt{Q}z\right]\nu(dz),$$

where $x \in \mathbb{R}^N$. Let us take $0 \leq s < t$, $x \in \mathbb{R}^N$, and observe that by the definition in (2.19) and Corollary 2.5 there exists a constant $C > 0$ such that, for every $r_1, r_2 \in [s,t]$,

$$\sup_{u\in[s,t]} \left|A(u)P_{u,t}\phi\left(Z_{r_2}^{s,x}\right) - A(u)P_{u,t}\phi\left(Z_{r_1}^{s,x}\right)\right| \leq C\left|Z_{r_2}^{s,x} - Z_{r_1}^{s,x}\right| \\ \times \left[\left(|A|\left(1 + \left|Z_{r_1}^{s,x}\right|\right) + \|B_0\|_{1,T}\right) + \int_{\mathbb{R}^N}\left(1_D(z)\left|\sqrt{Q}z\right|^2 + 1_{D^c}(z)\right)\nu(dz)\right]. \quad (2.28)$$

We study the mapping $[s,t] \ni r \mapsto R_{s,r}(P_{r,t}\phi)(x)$: using (2.23) and (2.28), it is easy to argue that it is continuous in its domain by Theorem 2.7 (*ii*) coupled with Vitali's and dominated convergence theorems. It is also differentiable, with

$$\partial_r R_{s,r}(P_{r,t}\phi)(x) = R_{s,r}\left(\widetilde{A}(r)P_{r,t}\phi\right)(x) - R_{s,r}(A(r)P_{r,t}\phi)(x) \\ = -R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top P_{r,t}\phi\right\rangle\right)(x), \quad r \in [s,t]. \quad (2.29)$$

Indeed, take $r \in [s,t]$ and a generic sequence $(r_n)_n \subset [s,t] \setminus \{r\}$ such that $r_n \to r$ as $n \to \infty$; then

$$\frac{R_{s,r_n}(P_{r_n,t}\phi)(x) - R_{s,r}(P_{r,t}\phi)(x)}{r_n - r} \\ = R_{s,r_n}\left(\frac{P_{r_n,t}\phi - P_{r,t}\phi}{r_n - r}\right)(x) + \mathbb{E}\left[\frac{P_{r,t}\phi\left(Z_{r_n}^{s,x}\right) - P_{r,t}\phi\left(Z_r^{s,x}\right)}{r_n - r}\right] =: \mathbf{I}_n + \mathbf{II}_n.$$

We immediately notice that $\mathbf{II}_n \to R_{s,r}\left(\widetilde{A}(r)P_{r,t}\phi\right)(x)$ as $n \to \infty$ by Theorem 2.7 (*i*) and Corollary 2.5. As for $\mathbf{I}_n$, we split it again as follows:

$$\mathbf{I}_n = R_{s,r}\left(\frac{P_{r_n,t}\phi - P_{r,t}\phi}{r_n - r}\right)(x) + \mathbb{E}\left[\frac{P_{r_n,t}\phi - P_{r,t}\phi}{r_n - r}\left(Z_{r_n}^{s,x}\right) - \frac{P_{r_n,t}\phi - P_{r,t}\phi}{r_n - r}\left(Z_r^{s,x}\right)\right] =: \mathbf{III}_n + \mathbf{IV}_n.$$

By a dominated convergence argument based on (2.20), (2.23), Corollary 2.5 and Theorem 2.7 (*ii*) we have $\mathbf{III}_n \to -R_{s,r}(A(r)P_{r,t}\phi)(x)$ as $n \to \infty$. Finally we focus on $\mathbf{IV}_n$, estimating by (2.23)

$$|\mathbf{IV}_n| \leq \mathbb{E}\left[\sup_{u\in[s,t]}\left|A(u)P_{u,t}\phi\left(Z_{r_n}^{s,x}\right) - A(u)P_{u,t}\phi\left(Z_r^{s,x}\right)\right|\right].$$

Notice that the random variables inside the expected value in the previous inequality converge to 0 in probability as $n \to \infty$ by (2.28). Such a convergence is true also in the L^1-sense, thanks to the estimates in (2.20) and Vitali's convergence theorem. Thus, $\mathbf{IV}_n \to 0$ as $n \to \infty$, fact which completely shows (2.29). Observe that $\partial_r R_{s,r}(P_{r,t}\phi)(x)$ is continuous in $[s,t]$ by Vitali's and dominated convergence theorems, the mean value theorem, Corollary 2.5 and the continuity of the mapping $r \mapsto \nabla P_{r,t}\phi(x)$ in

$[s, t]$ (see (2.23) and the subsequent sentence). Therefore we can integrate it with respect to r on the interval $[s, t]$ and infer that

$$P_{s,t}\phi(x) = R_{s,t}\phi(x) + \int_s^t R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top P_{r,t}\phi\right\rangle\right)(x)\, dr, \tag{2.30}$$

which coincides with (2.9).

Next, we take $\phi \in C_b\left(\mathbb{R}^N\right)$ and consider a sequence $(\phi_n)_n \subset C_b^3\left(\mathbb{R}^N\right)$ such that $\|\phi_n - \phi\|_\infty \to 0$ as $n \to \infty$. Since by (2.25) and Lemma 2.10 (for some constant $C_\alpha > 0$)

$$\begin{aligned}
&\left|\int_s^t R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top P_{r,t}(\phi_n - \phi)\right\rangle\right)(x)\, dr\right| \\
&\qquad\leq C_\alpha \|B\|_{0,T} \|\phi_n - \phi\|_\infty \left|\left(\sqrt{Q}\right)^{-1}\right| e^{(|A| + \|DB_0\|_{T,\infty})T} \left(\int_s^t \frac{dr}{(t-r)^{1/(2\alpha)}}\right) \underset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

by dominated convergence it is immediate to get the validity of (2.30) for ϕ, as well.

Finally, we tackle the case $\phi \in \mathcal{B}_b\left(\mathbb{R}^N\right)$. We consider ϕ to be the indicator function of an open set to begin with. Then, by Urysohn's lemma there exists a sequence $(\phi_n)_n \subset C_b\left(\mathbb{R}^N\right)$ such that $0 \leq \phi_n \leq \phi$ and $\phi_n \to \phi$ pointwise as $n \to \infty$. By construction and dominated convergence we have

$$\lim_{n\to\infty}\left(P_{s,t}\phi_n(x) - R_{s,t}\phi_n(x)\right) = P_{s,t}\phi(x) - R_{s,t}\phi(x). \tag{2.31}$$

Now we focus on the integral term in (2.30). Let us fix $y, h \in \mathbb{R}^N$, $r \in (s,t)$ and $u \in (r,t)$. Then, exploiting the Chapman–Kolmogorov equations and (2.24), we write ($n \in \mathbb{N}$)

$$\begin{aligned}
\left\langle \nabla^\top P_{r,t}\phi_n(y), h\right\rangle &= \left\langle \nabla^\top \left(P_{r,u}\left(P_{u,t}\phi_n\right)\right)(y), h\right\rangle \\
&= \mathbb{E}\left[\frac{1}{L_u - L_r} P_{u,t}\phi_n\left(X_u^{r,y}\right) \int_r^u \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_v^{r,y}, dW_{L_v}\right\rangle\right].
\end{aligned} \tag{2.32}$$

Since, with the same argument as in (2.31), $P_{u,t}\phi_n \to P_{u,t}\phi$ pointwise in $\mathbb{R}^N$ as $n \to \infty$, and (see, e.g., (2.27))

$$\begin{aligned}
&\sup_{n\in\mathbb{N}}\left|\frac{P_{u,t}\phi_n\left(X_u^{r,y}\right)}{L_u - L_r} \int_r^u \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_v^{r,y}, dW_{L_v}\right\rangle\right| \\
&\qquad\leq \frac{1}{L_u - L_r}\left|\int_r^u \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_v^{r,y}, dW_{L_v}\right\rangle\right| \in L^1(\mathbb{P}),
\end{aligned}$$

we can pass to the limit in (2.32) to obtain, by dominated convergence,

$$\begin{aligned}
\lim_{n\to\infty}\left\langle \nabla^\top P_{r,t}\phi_n(y), h\right\rangle &= \mathbb{E}\left[\frac{1}{L_u - L_r} P_{u,t}\phi\left(X_u^{r,y}\right) \int_r^u \left\langle \left(\sqrt{Q}\right)^{-1} D_h X_v^{r,y}, dW_{L_v}\right\rangle\right] \\
&= \left\langle \nabla^\top \left(P_{r,u}\left(P_{u,t}\phi\right)\right)(y), h\right\rangle = \left\langle \nabla^\top P_{r,t}\phi(y), h\right\rangle.
\end{aligned}$$

Observe that the second–to–last equality in the previous equation is due to (2.24) and Lemma 2.10. As a consequence, for every $r \in (s,t)$ we infer that

$$\lim_{n\to\infty} R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top P_{r,t}\phi_n\right\rangle\right)(x) = R_{s,r}\left(\left\langle B(r,\cdot), \nabla^\top P_{r,t}\phi\right\rangle\right)(x),$$

where we use once again the dominated convergence theorem, thanks to the next bound that we get using (2.25) and Lemma 2.10:

$$\left\|\left\langle B\left(r,\cdot\right),\nabla^{\top}P_{r,t}\phi_{n}\right\rangle\right\|_{\infty}\leq C_{\alpha}\left\|B\right\|_{0,T}\left|\left(\sqrt{Q}\right)^{-1}\right|e^{\left(|A|+\|DB_{0}\|_{T,\infty}\right)T}\frac{1}{\left(t-r\right)^{1/(2\alpha)}}.$$

Moreover, this inequality also allows to pass the limit under the integral sign, so that we end up with

$$\lim_{n\to\infty}\int_{s}^{t}R_{s,r}\left(\left\langle B\left(r,\cdot\right),\nabla^{\top}P_{r,t}\phi_{n}\right\rangle\right)(x)\,dr=\int_{s}^{t}R_{s,r}\left(\left\langle B\left(r,\cdot\right),\nabla^{\top}P_{r,t}\phi\right\rangle\right)(x)\,dr. \tag{2.33}$$

Combining (2.31)-(2.33) we conclude that (2.30) holds true for ϕ, i.e., for every indicator function of an open set.

Note that the passages of the previous step do not require the continuity of the approximating functions $\left(\phi_{n}\right)_{n}$, as long as they are equibounded, satisfy (2.30) and converge pointwise to ϕ. Therefore, we can state that (2.30) holds true for every $\phi\in\mathcal{B}_{b}\left(\mathbb{R}^{N}\right)$ by the functional monotone class theorem (see, e.g., [32, Theorem 2.12.9]).

We notice that, from (2.30), the continuity of $P_{\cdot,t}\phi\left(x\right),\ x\in\mathbb{R}^{N}$, in the interval $[0,t)$ can be argued by dominated convergence (see (2.39) below for an analogous computation). Furthermore, the measurability of $P_{s,t}\phi\left(x\right)$ with respect to (s,x) is a consequence of the measurability of the stochastic flow $X_{t}^{s,x}\left(\omega\right)$ and Tonelli's theorem. These facts, together with Lemma 2.10 and the gradient estimate in (2.25), entail that $P_{t-\diamond,t}\phi\left(\cdot\right)\in\Lambda_{1}^{\gamma}\left(0,t\right],\ \gamma=1/\left(2\alpha\right)$. Recalling Theorem 2.1 the proof is complete. ■

Remark 2.1. *Suppose that the requirements of Theorem 2.3 are satisfied. Given* $0\leq s<t\leq T$ *and* $\phi\in\mathcal{B}_{b}\left(\mathbb{R}^{N}\right)$, *we consider* $r\in\left(s,t\right)$ *and call* $\widetilde{\phi}=P_{r,t}\phi$. *By Theorem 2.3 and the Chapman–Kolmogorov equations,*

$$P_{s,t}\phi\left(x\right)=P_{s,r}\widetilde{\phi}\left(x\right)=u_{s}^{\widetilde{\phi}}\left(r,x\right),\quad x\in\mathbb{R}^{N},$$

where $u_{s}^{\widetilde{\phi}}\left(r,x\right)$ *is the unique solution of* (2.9) *such that* $u_{r-\diamond}^{\widetilde{\phi}}\left(r,\cdot\right)\in\Lambda_{1}^{\gamma}\left(0,r\right],\ \gamma=1/\left(2\alpha\right)$. *Observing that* $\widetilde{\phi}\in C_{b}^{1}\left(\mathbb{R}^{N}\right)$ *by Lemma 2.10, we invoke Corollary 2.2 to say that* $P_{s,t}\phi\in C_{b}^{2}\left(\mathbb{R}^{N}\right)$. *An iteration of this argument shows that* $P_{s,t}\phi\in C_{b}^{4}\left(\mathbb{R}^{N}\right)$. *In particular, the Kolmogorov backward equation* (2.22) *holds true in the interval* $[0,t)$ *for every* $\phi\in\mathcal{B}_{b}\left(\mathbb{R}^{N}\right)$.

2.3 The iteration scheme

Let $\alpha\in\left(\frac{1}{2},1\right),\ t\in\left(0,T\right],\ u_{0}\in\mathcal{B}_{b}\left(\mathbb{R}^{N}\right),\ B_{0}\in C_{b}^{0,3}\left(\left[0,T\right]\times\mathbb{R}^{N};\mathbb{R}^{N}\right)$ and $f\in C\left(\left[0,T\right];\mathbb{R}^{N}\right)$, so that Theorem 2.3 holds true. The proof of Theorem 2.1 (see, in particular, (2.10)-(2.14)) suggests to approximate the unique solution $u_{s}^{u_{0}}\left(t,x\right)\left(=P_{s,t}u_{0}\left(x\right)\right)$ of (2.9) such that $u_{t-\diamond}^{u_{0}}\left(t,\cdot\right)\in\Lambda_{1}^{\gamma}\left(0,t\right],\ \gamma=1/\left(2\alpha\right)$, with the iterates

$$\begin{cases}u_{s}^{n+1}\left(t,x\right)=R_{s,t}u_{0}\left(x\right)+\int_{s}^{t}R_{s,r}\left(\left\langle B\left(r,\cdot\right),\nabla^{\top}u_{r}^{n}\left(t,\cdot\right)\right\rangle\right)(x)\,dr\\ u_{s}^{0}\left(t,x\right)=R_{s,t}u_{0}\left(x\right)\end{cases},$$

for $x\in\mathbb{R}^{N},\ s\in\left[0,t\right],\ n\in\mathbb{N}\cup\{0\}$. Here we recall that $B=B_{0}-f$. If we define $v_{s}^{0}\left(t,x\right)=u_{s}^{0}\left(t,x\right)$ and $v_{s}^{n+1}\left(t,x\right)=u_{s}^{n+1}\left(t,x\right)-u_{s}^{n}\left(t,x\right),\ n\in\mathbb{N}\cup\{0\}$, then these new functions satisfy the iteration scheme

$$\begin{cases}v_{s}^{n+1}\left(t,x\right)=\int_{s}^{t}R_{s,u}k_{u,t}^{n}\left(x\right)du\\ k_{u,t}^{n}\left(x\right)=\left\langle B\left(u,x\right),\nabla^{\top}v_{u}^{n}\left(t,x\right)\right\rangle,\quad x\in\mathbb{R}^{N},\ s\in\left[0,t\right],\ u\in\left[0,t\right),\ n\in\mathbb{N}\cup\{0\}.\\ v_{s}^{0}\left(t,x\right)=R_{s,t}u_{0}\left(x\right)\end{cases} \tag{2.34}$$

In the Brownian case, (2.34) has been investigated in [85]. In order to study the convergence of $\sum_{n=0}^{\infty} v_s^n(t,x)$ to $u_s^{u_0}(t,x)$ (in a sense that will be clarified later on), we need the next, preliminary result.

Lemma 2.11. *Let $\alpha \in \left(\frac{1}{2},1\right)$, $t \in (0,T]$, $n \in \mathbb{N} \cup \{0\}$ and denote by $\gamma = 1/(2\alpha)$. Then $k_{u,t}^n \in C_b\left(\mathbb{R}^N\right)$ and $v_s^n(t,\cdot) \in C_b^1\left(\mathbb{R}^N\right)$ for every u, $s \in [0,t)$.*

Moreover, there exists a constant $C = C(\alpha, A, Q) > 0$ such that, for every $n \in \mathbb{N}$ and $s \in [0,t)$,

$$\|v_s^n(t,\cdot)\|_\infty \le C^n \|B\|_{0,T}^n \|u_0\|_\infty \int_0^{t-s} ds_n \int_0^{s_n} ds_{n-1} \cdots \int_0^{s_2} ds_1 \prod_{i=1}^n \frac{1}{(s_{i+1}-s_i)^\gamma}, \tag{2.35}$$

and

$$\left\|\nabla^\top v_s^n(t,\cdot)\right\|_\infty \le C^{n+1} \|B\|_{0,T}^n \|u_0\|_\infty \int_0^{t-s} ds_n \int_0^{s_n} ds_{n-1} \ldots \int_0^{s_2} ds_1 \prod_{i=0}^n \frac{1}{(s_{i+1}-s_i)^\gamma}, \tag{2.36}$$

where $s_0 = 0$ and $s_{n+1} = t - s$.

We notice that the constant C in (2.35)-(2.36) is the same as the one appearing in the gradient estimate (2.8).

Proof. We proceed by induction to prove that, for every $u, s \in [0,t)$ and $n \in \mathbb{N} \cup \{0\}$, one has $v_s^n(t,\cdot) \in C_b^1\left(\mathbb{R}^N\right)$, $k_{u,t}^n \in C_b\left(\mathbb{R}^N\right)$ and

$$\left\|k_{u,t}^n\right\|_\infty \le C^{n+1} \|B\|_{0,T}^{n+1} \|u_0\|_\infty \int_u^t ds_1 \int_{s_1}^t ds_2 \cdots \int_{s_{n-1}}^t ds_n \prod_{i=0}^n \frac{1}{(s_{i+1}-s_i)^\gamma}, \tag{2.37}$$

where $C = C(\alpha, A, Q) > 0$ is the same constant as in (2.8). In (2.37), $s_0 = u$ and $s_{n+1} = t$. The estimates in (2.35)-(2.36) are an immediate consequence of (2.37) upon shifting the domain of integration and applying Tonelli's theorem.

For $n = 0$, the smoothing effect of the time–dependent Markov semigroup R guarantees that $v_s^0(t,\cdot) \in C_b^1\left(\mathbb{R}^N\right)$, which combined with the continuity of B yields $k_{u,t}^0 \in C_b\left(\mathbb{R}^N\right)$, with

$$\left\|k_{u,t}^0\right\|_\infty \le C \|B\|_{0,T} \|u_0\|_\infty \frac{1}{(t-u)^\gamma}. \tag{2.38}$$

To fix the ideas, consider the case $n = 1$. Since $k_{u,t}^0 \in C_b\left(\mathbb{R}^N\right)$ for every $0 \le u < t$, the dominated convergence theorem, (2.34) and (2.38) imply that $v_s^1(t,\cdot) \in C_b^1\left(\mathbb{R}^N\right)$, with $\nabla v_s^1(t,x) = \int_s^t \nabla R_{s,u} k_{u,t}^0(x)\, du$, for every $x \in \mathbb{R}^N$. Hence $k_{u,t}^1 \in C_b\left(\mathbb{R}^N\right)$, and by (2.8)-(2.38) we get

$$\left\|k_{u,t}^1\right\|_\infty \le C^2 \|B\|_{0,T}^2 \|u_0\|_\infty \int_u^t ds_1 \frac{1}{(s_1-u)^\gamma (t-s_1)^\gamma}.$$

Suppose now that our statement holds true at step $n \in \mathbb{N}$. Then by the same argument as before and (2.37) $v_s^{n+1}(t,\cdot) \in C_b^1\left(\mathbb{R}^N\right)$, with $\nabla v_s^{n+1}(t,x) = \int_s^t \nabla R_{s,u} k_{u,t}^n(x)\, du$. Therefore $k_{u,t}^{n+1} \in C_b\left(\mathbb{R}^N\right)$, with

$$\begin{aligned}\left\|k_{u,t}^{n+1}\right\|_\infty &\le C \|B\|_{0,T} \int_u^t ds_1 \frac{1}{(s_1-u)^\gamma} \left\|k_{s_1,t}^n\right\|_\infty \\ &\le C^{n+2} \|B\|_{0,T}^{n+2} \|u_0\|_\infty \int_u^t ds_1 \int_{s_1}^t ds_2 \cdots \int_{s_n}^t ds_{n+1} \prod_{i=0}^{n+1} \frac{1}{(s_{i+1}-s_i)^\gamma},\end{aligned}$$

where in the last inequality we apply the inductive hypothesis and consider $s_0 = u$, $s_{n+2} = t$. Thus, the claim is completely proved. ■

Another important property of the functions $v_{\cdot}^{n}(t,x)$, $x\in\mathbb{R}^N$, is the continuity in the interval $[0,t)$. In the case $n=0$, this follows from the property of R discussed in Section 2.1; for a generic $n\in\mathbb{N}$, it can be argued by (2.37) and dominated convergence writing

$$v_s^n(t,x)=\int_0^t 1_{\{u>s\}}R_{s,u}k_{u,t}^{n-1}(x)\,du. \tag{2.39}$$

Thanks to the estimates in (2.35)-(2.36), the convergence of the iteration scheme (2.34) is proved in the same way as in the Brownian case with no time–shift, see [84, Section 2.4]. Overall, the next result is true.

Theorem 2.12. *For every $\alpha\in\left(\frac{1}{2},1\right)$ and $0<t\le T$, the series $\sum_{n=0}^{\infty}v_s^n(t,x)$ converges uniformly in $[0,t]\times\mathbb{R}^N$, and the series $\sum_{n=0}^{\infty}\nabla^\top v_s^n(t,x)$ converges uniformly in $[0,t_0]\times\mathbb{R}^N$, for every $t_0\in(0,t)$. In particular,*

$$\sum_{n=0}^{\infty}v_s^n(t,x)=u_s^{u_0}(t,x),\quad s\in[0,t],\, x\in\mathbb{R}^N,$$

where $u_s^{u_0}(t,x)$ is the unique solution of (2.9) such that $u_{t-\diamond}^{u_0}(t,\cdot)\in\Lambda_1^\gamma(0,t]$, $\gamma=1/(2\alpha)$.

2.4 The first term of the iteration scheme

Let $\alpha\in\left(\frac{1}{2},1\right)$. The goal of this section is to study $v_s^1(t,x)=\int_s^t R_{s,u}k_{u,t}^0(x)\,du$ –the first term of (2.34)– for every $0\le s<t\le T$ and $x\in\mathbb{R}^N$. In particular, starting from

$$k_{u,t}^0(y)=\left\langle B(u,y),\nabla^\top R_{u,t}u_0(y)\right\rangle,\quad y\in\mathbb{R}^N,\, u\in(s,t), \tag{2.40}$$

we want to find an alternative, explicit expression (see Lemma 2.14) for

$$R_{s,u}k_{u,t}^0(x)=\mathbb{E}\left[k_{u,t}^0\left(Z_u^{s,x}\right)\right]. \tag{2.41}$$

In order to do this, we propose an approach which at first analyzes a deterministic time–shift, and then allows to recover the subordinated Brownian motion case by conditioning with respect to $\mathcal{F}^L$. The results of this part represent the base case for the induction argument that we will develop to compute the general term $v_s^{n+1}(t,x)$, $n\ge1$ (see Section 2.5).

2.4.1 Deterministic time–shift

Denote by $\mathbb{S}$ the set of real–valued, strictly increasing càdlàg functions defined on $\mathbb{R}_+$ and starting at 0. Take $\ell\in\mathbb{S}$ and note that $W_\ell=(W_{\ell_t})_{t\ge0}$ is a càdlàg martingale with respect to the filtration $\left(\mathcal{F}_{\ell_t}^W\right)_{t\ge0}$, where $\left(\mathcal{F}_t^W\right)_{t\ge0}$ is the minimal augmented filtration generated by W. For every $x\in\mathbb{R}^N$ and $0\le s<T$, the OU process $\left(Z_t^\ell(s,x)\right)_{t\in[s,T]}$ is the unique, càdlàg solution of the linear SDE

$$dZ_t^\ell(s,x)=\left(AZ_t^\ell(s,x)+f(t)\right)dt+\sqrt{Q}\,dW_{\ell_t},\quad Z_s^\ell(s,x)=x.$$

It can be expressed with a variation of constants formula as follows:

$$Z_t^\ell(s,x)=e^{(t-s)A}x+\int_s^t e^{(t-r)A}f(r)\,dr+\int_s^t e^{(t-r)A}\sqrt{Q}\,dW_{\ell_r},\quad t\in[s,T].$$

For every $0 \le s < t \le T$, define $I^{\ell}_{s,t} = \int_s^t e^{2(t-r)A} Q \, d\ell_r \in \mathbb{R}^{N\times N}$. It is possible to argue as in [34, Equation (12)] (see (1.12)) to deduce that

$$Z^{\ell}_t(s,x) \sim \mathcal{N}\left(e^{(t-s)A}x + F_{s,t},\, I^{\ell}_{s,t}\right).$$

Note that, for every $0 \le s < u < t \le T$,

$$Z^{\ell}_t(s,x) = e^{(t-u)A} Z^{\ell}_u(s,x) + F_{u,t} + \int_u^t e^{(t-r)A}\sqrt{Q}\, dW_{\ell_r}, \quad \mathbb{P}-\text{a.s.},$$

therefore $\left(Z^{\ell}(s,x)\right)_{x\in\mathbb{R}^N}$ is a family of $(\mathcal{F}_{\ell_t})_{t\in[s,T]}$ $-$Markov processes as s varies in $[0,T)$. In particular, its transition probability kernels $\mu^{\ell}_{u,t} \colon \mathbb{R}^N \times \mathcal{B}\left(\mathbb{R}^N\right) \to [0,1]$ are

$$\mu^{\ell}_{u,t}(y,\cdot) = \mathcal{N}\left(e^{(t-u)A}y + F_{u,t}, I^{\ell}_{u,t}\right), \quad y \in \mathbb{R}^N. \tag{2.42}$$

In the sequel, we denote by $\phi^{\ell}_{u,t}(y,\cdot)$ the density of $\mu^{\ell}_{u,t}(y,\cdot)$. Moreover, we define

$$\widetilde{F}_{u,t}(y) = e^{(t-u)A}y + F_{u,t}, \quad y \in \mathbb{R}^N,$$

in order to keep the notation shorter. Straightforward changes to [34, Theorem 4] (see Theorem 1.4 in Chapter 1) ensure that, for any $0 \le s < t \le T$, the function $\mathbb{E}\left[u_0\left(Z^{\ell}_t(s,\cdot)\right)\right] \in C^1_b\left(\mathbb{R}^N\right)$, with derivative at any point $x \in \mathbb{R}^N$ in every direction $h \in \mathbb{R}^N$ given by

$$\left\langle \nabla^{\top}\mathbb{E}\left[u_0\left(Z^{\ell}_t(s,x)\right)\right], h\right\rangle = \mathbb{E}\left[u_0\left(Z^{\ell}_t(s,x)\right)\left\langle \left(I^{\ell}_{s,t}\right)^{-1} e^{(t-s)A}h, Z^{\ell}_t(s,x) - \widetilde{F}_{s,t}(x)\right\rangle\right]. \tag{2.43}$$

With all these preliminaries in mind, we fix $\ell^0 \in \mathbb{S}$, $0 \le u < t \le T$ and define –by analogy with (2.40)– the function

$$k^{\ell^0}_{u,t}(y) = \left\langle B(u,y), \nabla^{\top}\mathbb{E}\left[u_0\left(Z^{\ell^0}_t(u,y)\right)\right]\right\rangle, \quad y \in \mathbb{R}^N. \tag{2.44}$$

Note that $k^{\ell^0}_{u,t} \in C_b\left(\mathbb{R}^N\right)$ because $B(u,\cdot)$ is continuous and bounded, as well. The next claim provides us an analogue of (2.41) in this framework.

Lemma 2.13. *Consider $0 \le s < t \le T$. Then for every $x \in \mathbb{R}^N$, $u \in (s,t)$ and $\ell^0, \ell^1 \in \mathbb{S}$, one has, $\mathbb{P}-$a.s., writing Z^{ℓ^1} for $Z^{\ell^1}(s,x)$,*

$$\begin{aligned} k^{\ell^0}_{u,t}\left(Z^{\ell^1}_u\right) = \mathbb{E}\Bigg[u_0\bigg(\left(I^{\ell^0}_{u,t}\right)^{\frac{1}{2}}\left(I^{\ell^1}_{u,t}\right)^{-\frac{1}{2}}\left(Z^{\ell^1}_t - \widetilde{F}_{u,t}\left(Z^{\ell^1}_u\right)\right) + \widetilde{F}_{u,t}\left(Z^{\ell^1}_u\right)\bigg) \\ \times\left\langle \left(I^{\ell^0}_{u,t}\right)^{-\frac{1}{2}} e^{(t-u)A}B\left(u, Z^{\ell^1}_u\right), \left(I^{\ell^1}_{u,t}\right)^{-\frac{1}{2}}\left(Z^{\ell^1}_t - \widetilde{F}_{u,t}\left(Z^{\ell^1}_u\right)\right)\right\rangle \Big|\sigma\left(Z^{\ell^1}_u\right)\Bigg]. \end{aligned} \tag{2.45}$$

Proof. Fix $x \in \mathbb{R}^N$, $0 \le s < u < t \le T$ and ℓ^0, $\ell^1 \in \mathbb{S}$; by (2.43) we have

$$\begin{aligned} k^{\ell^0}_{u,t}\left(Z^{\ell^1}_u(s,x)\right) = k^{\ell^0}_{u,t}(y)\Big|_{y=Z^{\ell^1}_u(s,x)} = \mathbb{E}\Bigg[u_0\left(Z^{\ell^0}_t(u,y)\right) \\ \times\left\langle \left(I^{\ell^0}_{u,t}\right)^{-1} e^{(t-u)A}B(u,y), Z^{\ell^0}_t(u,y) - \widetilde{F}_{u,t}(y)\right\rangle\Bigg]\Bigg|_{y=Z^{\ell^1}_u(s,x)}. \end{aligned} \tag{2.46}$$

Note that $Z_t^{\ell^0}(u,y) \sim \mu_{u,t}^{\ell^0}(y,\cdot)\,,\ y \in \mathbb{R}^N$; furthermore, direct computations show that, for every $y,\ \xi \in \mathbb{R}^N$,

$$\phi_{u,t}^{\ell^1}(y,\xi) = \det\left(\left(I_{u,t}^{\ell^0}\right)^{\frac{1}{2}}\left(I_{u,t}^{\ell^1}\right)^{-\frac{1}{2}}\right)\phi_{u,t}^{\ell^0}\left(y, \left(I_{u,t}^{\ell^0}\right)^{\frac{1}{2}}\left(I_{u,t}^{\ell^1}\right)^{-\frac{1}{2}}\left(\xi - \widetilde{F}_{u,t}(y)\right) + \widetilde{F}_{u,t}(y)\right).$$

Going back to (2.46) we write, substituting $\xi = \left(I_{u,t}^{\ell^0}\right)^{\frac{1}{2}}\left(I_{u,t}^{\ell^1}\right)^{-\frac{1}{2}}\left(\xi' - \widetilde{F}_{u,t}(y)\right) + \widetilde{F}_{u,t}(y)$ as suggested by the previous calculations and considering $y = Z_u^{\ell^1}(s,x)$,

$$\begin{aligned} k_{u,t}^{\ell^0}(y) &= \int_{\mathbb{R}^N} u_0(\xi)\left\langle \left(I_{u,t}^{\ell^0}\right)^{-1} e^{(t-u)A} B(u,y)\,, \xi - \widetilde{F}_{u,t}(y)\right\rangle \phi_{u,t}^{\ell^0}(y,\xi)\, d\xi \\ &= \int_{\mathbb{R}^N} u_0\left(\left(I_{u,t}^{\ell^0}\right)^{\frac{1}{2}}\left(I_{u,t}^{\ell^1}\right)^{-\frac{1}{2}}\left(\xi' - \widetilde{F}_{u,t}(y)\right) + \widetilde{F}_{u,t}(y)\right) \\ &\quad \times \left\langle \left(I_{u,t}^{\ell^0}\right)^{-\frac{1}{2}} e^{(t-u)A} B(u,y)\,, \left(I_{u,t}^{\ell^1}\right)^{-\frac{1}{2}}\left(\xi' - \widetilde{F}_{u,t}(y)\right)\right\rangle \phi_{u,t}^{\ell^1}(y,\xi')\, d\xi'\,. \end{aligned}$$

At this point we invoke the disintegration formula of the conditional expectation (see, e.g., [116, Theorem 5.4]) and (2.42) to deduce (2.45), completing the proof. ■

Remark 2.2. *The function* $k_{u,t}^{\ell^0}$, $\ell^0 \in \mathbb{S}$, $0 \le u < t \le T$, *does not depend on the probability space where the underlying OU processes* $Z_t^{\ell^0}(u,x)\,,\ x \in \mathbb{R}^N$, *are defined.*

2.4.2 Random time–shift

Here we investigate the subordinated Brownian motion case (see Lemma 2.14) after some further preparation. In what follows, we denote by Ω_k, $k \in \mathbb{N} \cup \{0\}$, copies of the probability space Ω. As in Subsection 1.2.1, we denote by $\mathbb{W}$ the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^N$ vanishing at 0 and endow it with the Borel σ–algebra $\mathcal{B}(\mathbb{W})$ associated with the topology of locally uniform convergence. The pushforward probability measure generated by $W(\cdot): (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{W}, \mathcal{B}(\mathbb{W}))$ is denoted by $\mathbb{P}^{\mathbb{W}}$ and makes the canonical process $\mathfrak{x} = (x_t)_{t\ge 0}$ a Brownian motion. We work with the usual completion $\left(\mathbb{W}, \overline{\mathcal{B}(\mathbb{W})}, \overline{\mathbb{P}^{\mathbb{W}}}\right)$ of this probability space: $\mathfrak{x}$ is still a Brownian motion with respect to its minimal augmented filtration (cfr. [119, Theorem 7.9]). The completeness of the space $(\Omega, \mathcal{F}, \mathbb{P})$ implies the measurability of $W(\cdot): (\Omega, \mathcal{F}, \mathbb{P}) \to \left(\mathbb{W}, \overline{\mathcal{B}(\mathbb{W})}\right)$ and the fact that $\overline{\mathbb{P}^{\mathbb{W}}}$ is still the pushforward probability measure generated by $W(\cdot)$. Since $W(\cdot)$ is independent from $\mathcal{F}^L$, a regular conditional distribution of $W(\cdot)$ given $\mathcal{F}^L$ is $\overline{\mathbb{P}^{\mathbb{W}}}(A)\,,\ A \in \overline{\mathcal{B}(\mathbb{W})}$. Moreover, we denote by (coherently with Subsection 2.4.1)

$$\overline{Z}_t^{\ell}(u,y) = \widetilde{F}_{u,t}(y) + \int_u^t e^{(t-r)A}\sqrt{Q}\, dx_{\ell_r} : \mathbb{W} \to \mathbb{R}^N, \quad 0 \le u \le t \le T,\ y \in \mathbb{R}^N,\ \ell \in \mathbb{S},$$

and by $\mathbb{E}_k[\cdot]$ [resp., $\mathbb{E}_{\mathbb{W}}[\cdot]$] the expectation of a random variable defined on Ω_k [resp., $\mathbb{W}$]. We are now in position to prove the next claim, which is the analogue of [85, Corollary 2.2].

Lemma 2.14. *For every $x \in \mathbb{R}^N$ and $0 \le s < t \le T$ one has*

$$v_s^1(t,x) = \int_s^t du\, (\mathbb{E}_0 \otimes \mathbb{E}_1) \left[u_0 \left(\left(I_{u,t}^L(\omega_0)\right)^{\frac{1}{2}} \left(I_{u,t}^L\right)^{-\frac{1}{2}} \left(Z_t^{s,x} - \widetilde{F}_{u,t}\left(Z_u^{s,x}\right)\right) + \widetilde{F}_{u,t}\left(Z_u^{s,x}\right) \right) \right.$$
$$\left. \times \left\langle \left(I_{u,t}^L(\omega_0)\right)^{-\frac{1}{2}} e^{(t-u)A} B\left(u, Z_u^{s,x}\right), \left(I_{u,t}^L\right)^{-\frac{1}{2}} \left(Z_t^{s,x} - \widetilde{F}_{u,t}\left(Z_u^{s,x}\right)\right) \right\rangle (\omega_1) \right]. \tag{2.47}$$

Proof. Fix $0 \le s < t \le T$; combining the definition in (2.40) and the expression in (2.5) we get, by the law of total expectation, for every $u \in (s,t)$ and $y \in \mathbb{R}^N$,

$$k_{u,t}^0(y) = \mathbb{E}_0\left[\mathbb{E}_0 \left[u_0\left(Z_t^{u,y}\right) \left\langle \left(I_{u,t}^L\right)^{-1} e^{(t-u)A} B(u,y), Z_t^{u,y} - \widetilde{F}_{u,t}(y) \right\rangle \Big| \mathcal{F}^L \right]\right]. \tag{2.48}$$

The discussion preceding this lemma together with the usual rules of change of probability space (see, e.g., [109, §X-2]) and the substitution formula in [34, Lemma 5] (see Lemma 1.5 in Chapter 1) lets us apply the disintegration formula for the conditional expectation to get, from (2.46)-(2.48) and Remark 2.2, for every $y \in \mathbb{R}^N$,

$$\begin{aligned} k_{u,t}^0(y) &= \mathbb{E}_0\left[\mathbb{E}_{\mathbb{W}} \left[u_0 \left(\overline{Z}_t^{\ell^0}(u,y)\right) \left\langle \left(I_{u,t}^{\ell^0}\right)^{-1} e^{(t-u)A} B(u,y), \overline{Z}_t^{\ell^0}(u,y) - \widetilde{F}_{u,t}(y) \right\rangle \right] \Big|_{\ell^0 = L(\omega_0)} \right] \\ &= \mathbb{E}_0 \left[k_{u,t}^{\ell^0}(y) \Big|_{\ell^0 = L(\omega_0)} \right]. \end{aligned} \tag{2.49}$$

Since we aim to compute (2.41), for a generic $x \in \mathbb{R}^N$ we focus on

$$R_{s,u} k_{u,t}^0(x) = \mathbb{E}_1 \left[\mathbb{E}_1 \left[k_{u,t}^0\left(Z_u^{s,x}\right) \Big| \mathcal{F}^L \right]\right] = \mathbb{E}_1 \left[\mathbb{E}_{\mathbb{W}} \left[k_{u,t}^0 \left(\overline{Z}_u^{\ell^1}(s,x)\right)\right] \Big|_{\ell^1 = L(\omega_1)} \right], \tag{2.50}$$

with the last equality which is obtained by the same argument as in (2.49). At this point we combine (2.49) and (2.50) to write, using Fubini's theorem,

$$R_{s,u} k_{u,t}^0(x) = \mathbb{E}_0 \left[\mathbb{E}_1 \left[\mathbb{E}_{\mathbb{W}} \left[k_{u,t}^{\ell^0} \left(\overline{Z}_u^{\ell^1}(s,x)\right)\right] \Big|_{\ell^1 = L(\omega_1)} \right] \Big|_{\ell^0 = L(\omega_0)} \right].$$

Recalling that (2.45) in Lemma 2.13 provides us with an expression for $k_{u,t}^{\ell^0}\left(\overline{Z}_u^{\ell^1}(s,x)\right)$, we can use the law of total expectation and reason backwards with the conditioning in $\mathcal{F}^L$ to conclude that

$$R_{s,u} k_{u,t}^0(x) = \mathbb{E}_0 \left[\mathbb{E}_1 \left[u_0 \left(\left(I_{u,t}^{\ell^0}\right)^{\frac{1}{2}} \left(I_{u,t}^L\right)^{-\frac{1}{2}} \left(Z_t^{s,x} - \widetilde{F}_{u,t}\left(Z_u^{s,x}\right)\right) + \widetilde{F}_{u,t}\left(Z_u^{s,x}\right) \right) \right.\right.$$
$$\left.\left. \times \left\langle \left(I_{u,t}^{\ell^0}\right)^{-\frac{1}{2}} e^{(t-u)A} B\left(u, Z_u^{s,x}\right), \left(I_{u,t}^L\right)^{-\frac{1}{2}} \left(Z_t^{s,x} - \widetilde{F}_{u,t}\left(Z_u^{s,x}\right)\right) \right\rangle (\omega_1) \right] \Big|_{\ell^0 = L(\omega_0)} \right].$$

Integrating the previous expression in the interval (s,t) with respect to u we obtain (2.47) completing the proof. ■

2.5 The general term of the iteration scheme

Let $\alpha \in \left(\frac{1}{2}, 1\right)$. We want to analyze the general term $v_s^{n+1}(t,x) = \int_s^t R_{s,u} k_{u,t}^n(x)\, du, 0 \le s < t \le T$, of the iteration (2.34) for an integer $n \ge 1$. Therefore we search for an explicit expression of

$$R_{s,u} k_{u,t}^n(x) = \mathbb{E}\left[k_{u,t}^n\left(Z_u^{s,x}\right)\right], \quad x \in \mathbb{R}^N,\ u \in (s,t). \tag{2.51}$$

2.5.1 Deterministic time–shift

We continue the construction carried out in Subsection 2.4.1. Specifically, fix an integer $n \ge 1$ and $t \in (0,T]$; for every $i = 1, \ldots, n$, $(i+1)$ −tuple $\mathfrak{s}_i = (s_{n-i+1}, s_{n-i+2}, \ldots, s_{n+1})$ such that $0 \le s_{n-i+1} < s_{n-i+2} < \cdots < s_{n+1} < t$ and $\mathfrak{l}_i = (\ell^0, \ldots, \ell^i) \in \mathbb{S}^{i+1}$, we define, for $y \in \mathbb{R}^N$ (see (2.44)),

$$k_{\mathfrak{s}_i,t}^{\mathfrak{l}_i}(y) = \left\langle B\left(s_{n-i+1}, y\right), \nabla^\top \mathbb{E}\left[k_{\mathfrak{s}_{i-1},t}^{\mathfrak{l}_{i-1}}\left(Z_{s_{n-i+2}}^{\ell^i}\left(s_{n-i+1}, y\right)\right)\right]\right\rangle, \tag{2.52}$$

where $\mathfrak{s}_{i-1} = (s_{n-i+2}, s_{n-i+3}, \ldots, s_{n+1})$ and $\mathfrak{l}_{i-1} = (\ell^0, \ldots, \ell^{i-1})$. To shorten the notation, we denote by $n_i = n - i$. Note that, by the continuity and boundedness of B, an induction argument shows that all these functions are well defined and in $C_b\left(\mathbb{R}^N\right)$. Moreover, as in Remark 2.2 we observe that their value does not depend on the probability space where the underlying OU processes are constructed. By (2.43) we have, for every $y \in \mathbb{R}^N$, writing Z^{ℓ^i} for $Z^{\ell^i}(s_{n_i+1}, y)$,

$$k_{\mathfrak{s}_i,t}^{\mathfrak{l}_i}(y) = \mathbb{E}\Bigg[k_{\mathfrak{s}_{i-1},t}^{\mathfrak{l}_{i-1}}\left(Z_{s_{n_i+2}}^{\ell^i}\right) \times \left\langle \left(I_{s_{n_i+1},s_{n_i+2}}^{\ell^i}\right)^{-1} e^{\left(s_{n_i+2}-s_{n_i+1}\right)A} B\left(s_{n_i+1}, y\right), Z_{s_{n_i+2}}^{\ell^i} - \widetilde{F}_{s_{n_i+1},s_{n_i+2}}(y)\right\rangle\Bigg]. \tag{2.53}$$

Motivated by (2.51), we seek an explicit formula for the term $k_{\mathfrak{s}_n,t}^{\mathfrak{l}_n}\left(Z_{s_1}^{\ell^{n+1}}(s,x)\right)$, where $\ell^{n+1} \in \mathbb{S}$, $0 \le s < s_1 < \cdots < s_{n+1} < t$ and $x \in \mathbb{R}^N$. A candidate for such an expression is given by (2.45) in Lemma 2.13, from which we deduce the next claim.

Lemma 2.15. *Consider $0 \le s < t \le T$ and an integer $n \ge 1$. Then, for every $x \in \mathbb{R}^N$, $i = 0, \ldots, n$, $(i+2)$ −tuple $(s_{n_i}, s_{n_i+1}, \ldots, s_{n+1})$ such that $s \le s_{n_i} < s_{n_i+1} < \cdots < s_{n+1} < t$ and $\ell^0, \ldots, \ell^{i+1} \in \mathbb{S}$, one has, writing $Z^{\ell^{i+1}}$ for $Z^{\ell^{i+1}}(s_{n_i}, x)$,*

$$\begin{aligned} k_{\mathfrak{s}_i,t}^{\mathfrak{l}_i}\left(Z_{s_{n_i+1}}^{\ell^{i+1}}\right) = \mathbb{E}\Bigg[& u_0\Bigg(\widetilde{F}_{s_{n_i+1},t}\left(Z_{s_{n_i+1}}^{\ell^{i+1}}\right) + \sum_{j=1}^{i+1} \\ & e^{\left(t-s_{n_j+3}\right)A}\left(I_{s_{n_j+2},s_{n_j+3}}^{\ell^{j-1}}\right)^{\frac{1}{2}}\left(I_{s_{n_j+2},s_{n_j+3}}^{\ell^{i+1}}\right)^{-\frac{1}{2}}\left(Z_{s_{n_j+3}}^{\ell^{i+1}} - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(Z_{s_{n_j+2}}^{\ell^{i+1}}\right)\right)\Bigg) \\ & \times \prod_{j=1}^{i+1}\Bigg\langle \left(I_{s_{n_j+2},s_{n_j+3}}^{\ell^{j-1}}\right)^{-\frac{1}{2}} e^{\left(s_{n_j+3}-s_{n_j+2}\right)A} B\Bigg(s_{n_j+2}, \widetilde{F}_{s_{n_i+1},s_{n_j+2}}\left(Z_{s_{n_i+1}}^{\ell^{i+1}}\right) + \sum_{k=j+1}^{i+1} \\ & e^{\left(s_{n_j+2}-s_{n_k+3}\right)A}\left(I_{s_{n_k+2},s_{n_k+3}}^{\ell^{k-1}}\right)^{\frac{1}{2}}\left(I_{s_{n_k+2},s_{n_k+3}}^{\ell^{i+1}}\right)^{-\frac{1}{2}}\left(Z_{s_{n_k+3}}^{\ell^{i+1}} - \widetilde{F}_{s_{n_k+2},s_{n_k+3}}\left(Z_{s_{n_k+2}}^{\ell^{i+1}}\right)\right)\Bigg), \\ & \left(I_{s_{n_j+2},s_{n_j+3}}^{\ell^{i+1}}\right)^{-\frac{1}{2}}\left(Z_{s_{n_j+3}}^{\ell^{i+1}} - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(Z_{s_{n_j+2}}^{\ell^{i+1}}\right)\right)\Bigg\rangle \Bigg| \sigma\left(Z_{s_{n_i+1}}^{\ell^{i+1}}\right)\Bigg], \quad \mathbb{P}-a.s., \end{aligned} \tag{2.54}$$

where $\mathfrak{s}_i = (s_{n_i+1}, \ldots, s_{n+1})$, $\mathfrak{l}_i = (\ell^0, \ldots, \ell^i)$ and $s_{n+2} = t$.

In the previous expression, we interpret the empty sum to be 0: we adopt this convention hereafter.

Proof. Fix $0 \leq s < t \leq T$ and an integer $n \geq 1$. We proceed by induction on i, observing that the base case $i = 0$ has been proven in (2.45), where $s_n = s$ and $s_{n+1} = u$.

For the induction step, suppose that the statement is valid for $i = m-1$, for some $m = 1, \dots, n$: our goal is to show that it holds true for $i = m$, as well. Take an $(m+2)$ −tuple $(s_{n_m}, s_{n_m+1}, \dots, s_{n+1})$ such that $s \leq s_{n_m} < s_{n_m+1} < \dots < s_{n+1} < t$ and $\ell^0, \dots, \ell^{m+1} \in \mathbb{S}$; write $\mathfrak{s}_m = (s_{n_m+1}, \dots, s_{n+1})$ and $\mathfrak{l}_m = (\ell^0, \dots, \ell^m)$. Recalling (2.53) and denoting by $s_{n+2} = t$, we apply the inductive hypothesis and the law of total expectation to deduce, for every $y \in \mathbb{R}^N$, writing Z^{ℓ^m} for $Z^{\ell^m}(s_{n_m+1}, y)$,

$$\begin{aligned} k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t}(y) = \mathbb{E}\Bigg[u_0\Bigg(&\widetilde{F}_{s_{n_m+2},t}\left(Z^{\ell^m}_{s_{n_m+2}}\right) + \sum_{j=1}^m \\ & e^{\left(t-s_{n_j+3}\right)A}\left(I^{\ell^{j-1}}_{s_{n_j+2},s_{n_j+3}}\right)^{\frac12}\left(I^{\ell^m}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac12}\left(Z^{\ell^m}_{s_{n_j+3}} - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(Z^{\ell^m}_{s_{n_j+2}}\right)\right)\Bigg) \\ & \times \prod_{j=1}^m \Big\langle \left(I^{\ell^{j-1}}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac12} e^{\left(s_{n_j+3}-s_{n_j+2}\right)A} B\Bigg(s_{n_j+2}, \widetilde{F}_{s_{n_m+2},s_{n_j+2}}\left(Z^{\ell^m}_{s_{n_m+2}}\right) + \sum_{k=j+1}^m \\ & e^{\left(s_{n_j+2}-s_{n_k+3}\right)A}\left(I^{\ell^{k-1}}_{s_{n_k+2},s_{n_k+3}}\right)^{\frac12}\left(I^{\ell^m}_{s_{n_k+2},s_{n_k+3}}\right)^{-\frac12}\left(Z^{\ell^m}_{s_{n_k+3}} - \widetilde{F}_{s_{n_k+2},s_{n_k+3}}\left(Z^{\ell^m}_{s_{n_k+2}}\right)\right)\Bigg), \\ & \left(I^{\ell^m}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac12}\left(Z^{\ell^m}_{s_{n_j+3}} - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(Z^{\ell^m}_{s_{n_j+2}}\right)\right)\Big\rangle \\ & \times \left\langle \left(I^{\ell^m}_{s_{n_m+1},s_{n_m+2}}\right)^{-1} e^{(s_{n_m+2}-s_{n_m+1})A} B\left(s_{n_m+1}, y\right), Z^{\ell^m}_{s_{n_m+2}} - \widetilde{F}_{s_{n_m+1},s_{n_m+2}}(y) \right\rangle \Bigg], \end{aligned} \tag{2.55}$$

where we also consider the $\sigma\big(Z^{\ell^m}_{s_{n_m+2}}\big)$−measurability of the random variable

$$\left\langle \left(I^{\ell^m}_{s_{n_m+1},s_{n_m+2}}\right)^{-1} e^{(s_{n_m+2}-s_{n_m+1})A} B\left(s_{n_m+1}, y\right), Z^{\ell^m}_{s_{n_m+2}} - \widetilde{F}_{s_{n_m+1},s_{n_m+2}}(y) \right\rangle.$$

To shorten the notation we write $k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t} = \mathbb{E}\big[f\big(Z^{\ell^m}_{s_{n_m+2}}, Z^{\ell^m}_{s_{n_m+3}}, \dots, Z^{\ell^m}_t\big)\big]$. Since $Z^{\ell^m}_r$, $r \in [s_{n_m+1}, t]$, is a Markov process, we know that (cfr. [116, Proposition 7.2])

$$\left(Z^{\ell^m}_{s_{n_m+2}}, Z^{\ell^m}_{s_{n_m+3}}, \dots, Z^{\ell^m}_t\right) \sim \mu^{\ell^m}_{s_{n_m+1},s_{n_m+2}}(y) \otimes \mu^{\ell^m}_{s_{n_m+2},s_{n_m+3}} \otimes \dots \otimes \mu^{\ell^m}_{s_{n+1},t}.$$

Hence, using the same notation as in the previous section,

$$\begin{aligned} k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t}(y) = \int_{\mathbb{R}^N} \phi^{\ell^m}_{s_{n_{m+1}},s_{n_m+2}}(y,\xi_1)\Bigg(\int_{\mathbb{R}^N} \phi^{\ell^m}_{s_{n_m+2},s_{n_m+3}}(\xi_1,\xi_2)\Bigg(\dots\Bigg(\\ \int_{\mathbb{R}^N} \phi^{\ell^m}_{s_{n+1},t}(\xi_m,\xi_{m+1}) f(\xi_1,\dots,\xi_{m+1})\, d\xi_{m+1}\Bigg)\dots\Bigg)d\xi_2\Bigg)d\xi_1. \end{aligned} \tag{2.56}$$

We wish to rewrite (2.56) as an integral in $\mu^{\ell^{m+1}}_{s_{n_m+1},s_{n_m+2}}(y) \otimes \mu^{\ell^{m+1}}_{s_{n_m+2},s_{n_m+3}} \otimes \dots \otimes \mu^{\ell^{m+1}}_{s_{n+1},t}$. In order to do so, we sequentially perform the following substitutions:

$$\begin{cases} \xi_1 = \left(I^{\ell^m}_{s_{n_m+1},s_{n_m+2}}\right)^{\frac{1}{2}}\left(I^{\ell^{m+1}}_{s_{n_m+1},s_{n_m+2}}\right)^{-\frac{1}{2}}\left(\xi_1' - \widetilde{F}_{s_{n_m+1},s_{n_m+2}}(y)\right) + \widetilde{F}_{s_{n_m+1},s_{n_m+2}}(y) =: g_1(\xi_1')\,; \\ \xi_h = \left(I^{\ell^m}_{s_{n_m+h},s_{n_m+h+1}}\right)^{\frac{1}{2}}\left(I^{\ell^{m+1}}_{s_{n_m+h},s_{n_m+h+1}}\right)^{-\frac{1}{2}}\left(\xi_h' - \widetilde{F}_{s_{n_m+h},s_{n_m+h+1}}\left(\xi_{h-1}'\right)\right) \\ \qquad + \widetilde{F}_{s_{n_m+h},s_{n_m+h+1}}\left(g_{h-1}\left(\xi_1',\dots,\xi_{h-1}'\right)\right) =: g_h\left(\xi_1',\dots,\xi_h'\right), \quad h = 2,\dots,m+1. \end{cases}$$

In this way, (2.56) becomes

$$k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t}(y) = \int_{\mathbb{R}^N} \phi^{\ell^{m+1}}_{s_{n_m+1},s_{n_m+2}}\left(y,\xi_1'\right)\left(\int_{\mathbb{R}^N} \phi^{\ell^{m+1}}_{s_{n_m+2},s_{n_m+3}}\left(\xi_1',\xi_2'\right)\left(\dots\left(\right.\right.\right.$$

$$\left.\left.\left.\int_{\mathbb{R}^N} \phi^{\ell^{m+1}}_{s_{n+1},t}\left(\xi_m',\xi_{m+1}'\right) f\left(g_1\left(\xi_1'\right),\dots,g_{m+1}\left(\xi_1',\dots,\xi_{m+1}'\right)\right) d\xi_{m+1}'\right)\dots\right)d\xi_2'\right)d\xi_1'.$$

Expanding the notation for f contained in (2.55), we exploit several cancellations to get

$$\begin{aligned} k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t}(y) = &\int_{\mathbb{R}^N} \phi^{\ell^{m+1}}_{s_{n_m+1},s_{n_m+2}}(y,\xi_1')\Bigg(\dots\Bigg(\int_{\mathbb{R}^N} \phi^{\ell^{m+1}}_{s_{n+1},t}(\xi_m',\xi_{m+1}')\,u_0\Bigg(\widetilde{F}_{s_{n_m+1},t}(y) + \sum_{j=1}^{m+1} \\ &e^{\left(t-s_{n_j+3}\right)A}\left(I^{\ell^{j-1}}_{s_{n_j+2},s_{n_j+3}}\right)^{\frac{1}{2}}\left(I^{\ell^{m+1}}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac{1}{2}}\left(\xi_{m_j+2}' - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(\xi_{m_j+1}'\right)\right)\Bigg)\Bigg) \\ &\times \prod_{j=1}^{m+1}\Bigg\langle \left(I^{\ell^{j-1}}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac{1}{2}} e^{\left(s_{n_j+3}-s_{n_j+2}\right)A} B\Bigg(s_{n_j+2}, \widetilde{F}_{s_{n_m+1},s_{n_j+2}}(y) + \sum_{k=j+1}^{m+1} \\ &e^{\left(s_{n_j+2}-s_{n_k+3}\right)A}\left(I^{\ell^{k-1}}_{s_{n_k+2},s_{n_k+3}}\right)^{\frac{1}{2}}\left(I^{\ell^{m+1}}_{s_{n_k+2},s_{n_k+3}}\right)^{-\frac{1}{2}}\left(\xi_{m_k+2}' - \widetilde{F}_{s_{n_k+2},s_{n_k+3}}\left(\xi_{m_k+1}'\right)\right)\Bigg)\Bigg), \\ &\left(I^{\ell^{m+1}}_{s_{n_j+2},s_{n_j+3}}\right)^{-\frac{1}{2}}\left(\xi_{m_j+2}' - \widetilde{F}_{s_{n_j+2},s_{n_j+3}}\left(\xi_{m_j+1}'\right)\right)\Bigg\rangle d\xi_{m+1}'\Bigg)\dots\Bigg)d\xi_1', \end{aligned} \tag{2.57}$$

where we denote by $\xi_0' = y$. Noticing that $\delta_{Z^{\ell^{m+1}}_{s_{n_m+1}}(s_{n_m},x)} \otimes \mu^{\ell^{m+1}}_{s_{n_m+1},s_{n_m+2}} \otimes \dots \otimes \mu^{\ell^{m+1}}_{s_{n+1},t}$, $x \in \mathbb{R}^N$, is a regular conditional distribution for

$$\mathbb{P}\Big(\left(Z^{\ell^{m+1}}_{s_{n_m+1}}(s_{n_m},x), Z^{\ell^{m+1}}_{s_{n_m+2}}(s_{n_m},x),\dots,Z^{\ell^{m+1}}_{t}(s_{n_m},x)\right) \in \cdot \,\Big|\, \sigma\left(Z^{\ell^{m+1}}_{s_{n_m+1}}(s_{n_m},x)\right)\Big)$$

thanks to [116, Propositions 5.6-7.2], (2.57) yields (2.54) by the disintegration formula of the conditional expectation. The proof is now complete. ■

2.5.2 Random time–shift

We argue by conditioning with respect to $\mathcal{F}^L$ as in Subsection 2.4.2. First, we present a result which generalizes (2.49) in the proof of Lemma 2.14.

Lemma 2.16. *Consider $0 \le s < t \le T$ and an integer $n \ge 1$. Then for all $i = 0,\dots,n$, $s_1 \in (s,t)$ and $y \in \mathbb{R}^N$,*

$$k^{i}_{s_1,t}(y)=\int_{s_1}^{t}ds_2\int_{s_2}^{t}ds_3\cdots\int_{s_i}^{t}ds_{i+1}\mathbb{E}_i\left[\dots\left[\mathbb{E}_0\left[k^{\mathfrak{l}_i}_{\mathfrak{s}_i,t}(y)\Big|_{\ell^0=L(\omega_0)}\right]\dots\right]\Big|_{\ell^i=L(\omega_i)}\right], \tag{2.58}$$

where $\mathfrak{s}_i=(s_1,\dots,s_{i+1})$ *and* $\mathfrak{l}_i=(\ell^0,\dots,\ell^i)$.

In this expression, we ignore the time–integrals when $i=0$.

Proof. Take an integer $n\geq 1$ and proceed by induction on i. For $i=0$, there are no integrals in time in (2.58), which then reduces to (2.49) with $s_1=u$.

Suppose that the statement holds for $i=m-1$, for some $m=1,\dots,n$: we want to prove its validity for $i=m$. To do so, let us fix $y\in\mathbb{R}^N$ and $s_1\in(s,t)$; recalling the definition of $k^m_{s_1,t}$ in (2.34), by Lemma 2.11 we can apply (2.5) to get

$$k^{m}_{s_1,t}(y)=\int_{s_1}^{t}ds_2\,\mathbb{E}_m\Bigg[\mathbb{E}_{\mathbb{W}}\Bigg[k^{m-1}_{s_2,t}\left(\overline{Z}^{\ell^m}_{s_2}\right)\\ \times\left\langle\left(I^{\ell^m}_{s_1,s_2}\right)^{-1}e^{(s_2-s_1)A}B(s_1,y),\overline{Z}^{\ell^m}_{s_2}-\widetilde{F}_{s_1,s_2}(y)\right\rangle\Bigg]\Bigg|_{\ell^m=L(\omega_m)}\Bigg].$$

Here we write $\overline{Z}^{\ell^m}$ for $\overline{Z}^{\ell^m}(s_1,y)$. By the inductive hypothesis, we substitute the expression for $k^{m-1}_{s_2,t}$, $s_2\in(s_1,t)$, in the previous equality to obtain (ignoring the inner time–integral when $m=1$)

$$k^{m}_{s_1,t}(y)=\int_{s_1}^{t}ds_2\,\mathbb{E}_m\Bigg[\mathbb{E}_{\mathbb{W}}\Bigg[\int_{s_2}^{t}ds_3\cdots\int_{s_m}^{t}ds_{m+1}\mathbb{E}_{m-1}\Bigg[\dots\Bigg[\mathbb{E}_0\Bigg[k^{\mathfrak{l}_{m-1}}_{\mathfrak{s}_{m-1},t}\left(\overline{Z}^{\ell^m}_{s_2}\right)\Bigg|_{\ell^0=L(\omega_0)}\\ \Bigg]\dots\Bigg]\Bigg|_{\ell^{m-1}=L(\omega_{m-1})}\Bigg]\left\langle\left(I^{\ell^m}_{s_1,s_2}\right)^{-1}e^{(s_2-s_1)A}B(s_1,y),\overline{Z}^{\ell^m}_{s_2}-\widetilde{F}_{s_1,s_2}(y)\right\rangle\Bigg]\Bigg|_{\ell^m=L(\omega_m)}\Bigg],$$

where $\mathfrak{s}_{m-1}=(s_2,\dots,s_{m+1})$ and $\mathfrak{l}_{m-1}=(\ell^0,\dots,\ell^{m-1})$. This equation can be rewritten by Fubini's theorem –whose application is guaranteed by [84, Lemma 2.12], upon carrying out computations similar to those in the proof of [34, Theorem 6] (see Theorem 1.6 in Chapter 1, and also [25, Proposition 3.2])– as follows:

$$k^{m}_{s_1,t}(y)=\int_{s_1}^{t}ds_2\int_{s_2}^{t}ds_3\cdots\int_{s_m}^{t}ds_{m+1}\mathbb{E}_m\Bigg[\mathbb{E}_{m-1}\Bigg[\dots\Bigg[\mathbb{E}_0\Bigg[\mathbb{E}_{\mathbb{W}}\Bigg[k^{\mathfrak{l}_{m-1}}_{\mathfrak{s}_{m-1},t}\left(\overline{Z}^{\ell^m}_{s_2}\right)\\ \times\left\langle\left(I^{\ell^m}_{s_1,s_2}\right)^{-1}e^{(s_2-s_1)A}B(s_1,y),\overline{Z}^{\ell^m}_{s_2}-\widetilde{F}_{s_1,s_2}(y)\right\rangle\Bigg]\Bigg|_{\ell^0=L(\omega_0)}\Bigg]\dots\Bigg]\Bigg|_{\ell^{m-1}=L(\omega_{m-1})}\Bigg]\Bigg|_{\ell^m=L(\omega_m)}\Bigg].$$

This gives (2.58), once we plug in the expression of $k^{\mathfrak{l}_m}_{\mathfrak{s}_m,t}(y)$ in (2.53), where $\mathfrak{s}_m=(s_1,\dots,s_{m+1})$ and $\mathfrak{l}_m=(\ell^0,\dots,\ell^m)$. Thus, the proof is complete. ■

According to (2.51), given $0\leq s<s_1<t\leq T$ we are interested in

$$R_{s,s_1}k^{n}_{s_1,t}(x)=\mathbb{E}_{n+1}\left[\mathbb{E}_{\mathbb{W}}\left[k^{n}_{s_1,t}\left(\overline{Z}^{\ell^{n+1}}_{s_1}\right)\right]\Big|_{\ell^{n+1}=L(\omega_{n+1})}\right]=\int_{s_1}^{t}ds_2\dots\int_{s_n}^{t}ds_{n+1}\\ \mathbb{E}_0\left[\dots\left[\mathbb{E}_{n+1}\left[\mathbb{E}_{\mathbb{W}}\left[k^{\mathfrak{l}_n}_{\mathfrak{s}_n,t}\left(\overline{Z}^{\ell^{n+1}}_{s_1}\right)\right]\Big|_{\ell^{n+1}=L(\omega_{n+1})}\right]\dots\right]\Big|_{\ell^0=L(\omega_0)}\right], \tag{2.59}$$

where we use Lemma 2.16 and Fubini's theorem for the second equality. Here $\overline{Z}_{s_1}^{\ell^{n+1}}$ represents $\overline{Z}_{s_1}^{\ell^{n+1}}(s,x)$, $\mathfrak{s}_n = (s_1, \dots, s_{n+1})$ and $\mathfrak{l}_n = (\ell^0, \dots, \ell^n)$. Since Lemma 2.15 in the previous subsection provides us with a formula for $k_{\mathfrak{s}_n,t}^{\mathfrak{l}_n}(\overline{Z}_{s_1}^{\ell^{n+1}})$ (see (2.54) with $s_0 = s$ and $i = n$), we just plug it into (2.59), apply the law of total expectation and reason backwards with the conditioning in $\mathcal{F}^L$ to deduce the next result (cfr. [85, Theorem 2.3]).

Theorem 2.17. *For every integer $n \geq 1$, $x \in \mathbb{R}^N$ and $0 \leq s < t \leq T$ one has*

$$
\begin{aligned}
v_s^{n+1}(t,x) = &\int_s^t ds_1 \int_{s_1}^t ds_2 \dots \int_{s_n}^t ds_{n+1} (\mathbb{E}_0 \otimes \mathbb{E}_1 \otimes \cdots \otimes \mathbb{E}_{n+1}) \Bigg[u_0 \Bigg(\widetilde{F}_{s_1,t}(Z_{s_1}^{s,x}) + \sum_{j=1}^{n+1} \\
& e^{\left(t - s_{n_j+3}\right)A} \Big(I^L_{s_{n_j+2}, s_{n_j+3}}(\omega_{j-1}) \Big)^{\frac{1}{2}} \Big(I^L_{s_{n_j+2}, s_{n_j+3}} \Big)^{-\frac{1}{2}} \Big(Z^{s,x}_{s_{n_j+3}} - \widetilde{F}_{s_{n_j+2}, s_{n_j+3}} \Big(Z^{s,x}_{s_{n_j+2}} \Big) \Big) \Bigg) \\
& \times \prod_{j=1}^{n+1} \Big\langle \Big(I^L_{s_{n_j+2}, s_{n_j+3}}(\omega_{j-1}) \Big)^{-\frac{1}{2}} e^{\left(s_{n_j+3} - s_{n_j+2}\right)A} B \Bigg(s_{n_j+2}, \widetilde{F}_{s_1, s_{n_j+2}}(Z_{s_1}^{s,x}) + \sum_{k=j+1}^{n+1} \\
& e^{\left(s_{n_j+2} - s_{n_k+3}\right)A} \Big(I^L_{s_{n_k+2}, s_{n_k+3}}(\omega_{k-1}) \Big)^{\frac{1}{2}} \Big(I^L_{s_{n_k+2}, s_{n_k+3}} \Big)^{-\frac{1}{2}} \Big(Z^{s,x}_{s_{n_k+3}} - \widetilde{F}_{s_{n_k+2}, s_{n_k+3}} \Big(Z^{s,x}_{s_{n_k+2}} \Big) \Big) \Bigg), \\
& \Big(I^L_{s_{n_j+2}, s_{n_j+3}} \Big)^{-\frac{1}{2}} \Big(Z^{s,x}_{s_{n_j+3}} - \widetilde{F}_{s_{n_j+2}, s_{n_j+3}} \Big(Z^{s,x}_{s_{n_j+2}} \Big) \Big) \Big\rangle (\omega_{n+1}) \Bigg],
\end{aligned}
\tag{2.60}
$$

where $s_{n+2} = t$.

2.6 Numerical simulations

In this section we report on the results obtained by implementing the iterative scheme described above for two choices of the nonlinear vector field B_0. We interpret the SDE in (2.16) as a finite–dimensional approximation of the reaction–diffusion SPDE

$$
\begin{cases}
dX(t,\xi) = (\Delta X(t,\xi) + B_0(t, X(t,\xi)))\, dt + \sigma\, dW_{L_t}, & t \geq s, \\
X(s,\xi) = x(\xi), & \xi \in \mathbb{T}^1,
\end{cases}
$$

where $\mathbb{T}^1 = \mathbb{R}^1/\mathbb{Z}^1$ is the one–dimensional torus. We refer to [34, Example 1], see Example 1.1 in Chapter 1, for an accurate description of this framework. Hence we consider $\lambda_k = |k|^2$, $k = 1, \dots, N$, and we take $Q = \sigma^2 \mathrm{Id}$. Here $\sigma > 0$ is a parameter describing the strength of the noise.

The reason why we choose such an SPDE is that we aim at applying our iterative scheme to random perturbations of fluid dynamic models, appearing for example in climate studies. We refer to the book [82, Chapters 3–5] for an extensive analysis of the topic in the case of Gaussian noise. The interest in considering W_L as the driver of randomness is that, departing from the Brownian setting, it allows to better capture extreme events thanks to the fat tails of its increments, while preserving the invariance by rotation, i.e., the isotropy. We refer to [48, Introduction] and the references therein for a wide range of applications characterized by strong non–Gaussianity. However, compared to [48], our approach to these high dimensional, non–Gaussian problems is completely different. Finally, we mention that, in order to tackle more realistic models involving, e.g., quadratic nonlinearities, the theoretical framework presented in this chapter has to be expanded, and in particular the hypothesis of boundedness of B_0 has to be overcome: this will be the focus of a future research.

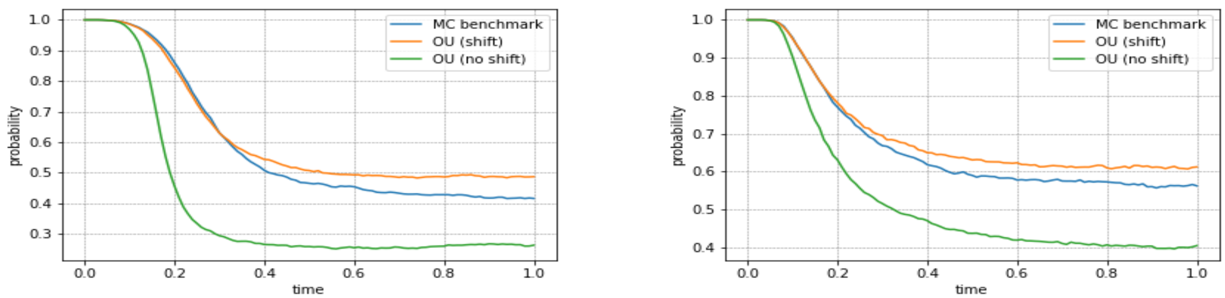

Figure 2.1: *Behavior in time of the OU approximations in the bounded cubic case with and without time–shift. The panel on the left refers to $\alpha = 0.55$, the one on the right to $\alpha = 0.85$. $\sigma = 0.5$ everywhere.*

Before moving to the application of the model, we have to determine the *time–shift* function $f \in C\left([0,T];\mathbb{R}^N\right)$ appearing in the OU process $Z^{s,x}$, $x \in \mathbb{R}^N$ (see (2.2)). Since we are dealing with a rotation–invariant noise and $\alpha \in \left(\frac{1}{2},1\right)$, $\mathbb{E}\left[W_{L_t}\right] = 0$, $t \geq 0$. As a consequence, the choice of f can be motivated as in [85, Introduction] for the Brownian case. In brief, we consider

$$f(t) = B_0\left(t, x(t)\right), \quad t \in [0,T],$$

where $x(\cdot): [0,T] \to \mathbb{R}^N$ is the unique solution of the integral equation

$$x(t) = x + \int_s^t \left(Ax(r) + B_0\left(r, x(r)\right)\right) dr, \quad t \in [s,T], \tag{2.61}$$

and $x(t) = x$, $t \in [0,s]$. Of course, $x(\cdot)$ is computed numerically. Note that (2.61) is the deterministic counterpart of the semilinear SDE (2.16), and that the expected value function of the OU process coincides with $x(\cdot)$ in the interval $[s,T]$ by the choice of f. The intuition is that, at least when the noise is weak, the trajectories of the semilinear solutions are "close" to $x(\cdot)$, allowing the $0-$th iterate to perform better than it would do with $f \equiv 0$. Figure 2.1 clearly displays this idea in the case of (bounded) cubic nonlinearity treated below (see (2.62)). Furthermore, in the sequel we monitor the effect of the time–shift on the first order approximation provided by our scheme. All the simulations are carried out using the High Performance Computing Center of the Scuola Normale Superiore (`https://hpccenter.sns.it`). We work in dimension $N = 100$, with $u_0(x) = 1_{\{|x|>R\}}$, $x \in \mathbb{R}^N$, for some $R > 0$, and we denote by $\mathbf{e} \in \mathbb{R}^N$ the vector with all components equal to 1. In particular, given $0 \leq s < t \leq 1$, we are interested in applying our iterates to approximate $P_{s,t}u_0(\mathbf{e}) = \mathbb{P}\left(\left|X_t^{s,\mathbf{e}}\right| > R\right)$, whose reference value is computed by averaging 10^5 samples of $X_t^{s,\mathbf{e}}$ obtained by the Euler–Maruyama scheme with time step 10^{-4}. The same strategy is used to obtain the $0-$th iterate $v_s^0(t,\mathbf{e}) = \mathbb{P}\left(\left|Z_t^{s,\mathbf{e}}\right| > R\right)$. In order to calculate the numerical integrals appearing in the formulas for $v_s^n(t,\mathbf{e})$, $n \in \mathbb{N}$ (see (2.47)-(2.60)), we use left Riemann sums in a uniform grid with mesh 10^{-2}. We will keep track of the relative error ϵ_r^n, defined by

$$\epsilon_r^n = \frac{P_{s,t}u_0(\mathbf{e}) - \sum_{i=0}^n v_s^i(t,\mathbf{e})}{P_{s,t}u_0(\mathbf{e})}, \quad n \in \mathbb{N} \cup \{0\}.$$

Finally, we will mainly focus on the first iteration, with the aim of understanding the possible improvements that it provides over the linear approximation of the OU process. In fact, although it is possible to implement our scheme up to any order thanks to (2.60), one needs an $n-$dimensional integral (in time) to get the iterate $v_s^n(t,\mathbf{e})$, $n \in \mathbb{N}$, fact which complicates the application of our method and

Table 2.1: First order approximation in the sine case with time–shift; noise strength $\sigma = 1$.

α	$\mathbb{P}(\|X_1^{0,\mathbf{e}}\| > 1)$	$v_0^0(1,\mathbf{e})$	ϵ_r^0	$v_0^1(1,\mathbf{e})$	ϵ_r^1
0.55	0.687	0.639	6.99e-2	0.012	5.24e-2
0.65	0.713	0.676	5.19e-2	1.34e-2	3.31e-2
0.75	0.794	0.737	7.18e-2	3.34e-2	2.97e-2
0.85	0.899	0.863	0.040	1.87e-2	1.92e-2

Table 2.2: Same setting as in Table 2.1, without time–shift ($f \equiv 0$).

α	$\mathbb{P}(\|X_1^{0,\mathbf{e}}\| > 1)$	$v_0^0(1,\mathbf{e})$	ϵ_r^0	$v_0^1(1,\mathbf{e})$	ϵ_r^1
0.55	0.691	0.502	0.274	0.101	0.127
0.65	0.720	0.558	0.225	0.110	7.22e-2
0.75	0.785	0.666	0.151	8.84e-2	0.039
0.85	0.896	0.840	6.25e-2	3.86e-2	1.94e-2

may result in losing its computational advantage over the classical Euler–Maruyama approach. In what follows, we fix the initial time $s = 0$ and the threshold $R = 1$. For the subordinator L, we set $\bar{\gamma} = 1$ in (2.1).

We first take $B_0(x)_k = \sin(x_k)$, $k = 1, \dots, N$. Table 2.1 shows the performance of the first order approximation of the iterative scheme with time–shift as α varies in $\left(\frac{1}{2}, 1\right)$, $\sigma = 1$ and $t = 1$. Table 2.2 is analogous, but it refers to $f \equiv 0$ (no time–shift). The first thing we notice is that in both cases the first iteration improves on the outcomes of the linear approximation. The role of the time–shift f is evident in the column ϵ_r^0: it allows $v_0^0(1,\mathbf{e})$ to be closer to the benchmark probability, and the first iterate builds on this to guarantee a better overall performance, particularly when α is close to $\frac{1}{2}$. Next, Figure 2.2 displays the behavior in time –up to $t = 1$– of the first order approximation in the case of time–shift for two strengths of noise ($\sigma = 0.1$ and $\sigma = 1.3$). Here $\alpha = 0.6$ is fixed. The panels of this figure highlight the benefits of considering $v_0^1(\cdot,\mathbf{e})$ over the starting OU estimates, especially when the noise is weak.

Secondly, we analyze the polynomial vector field

$$B_0(x)_k = b_0 \|\bar{y}\|_\infty \frac{(\bar{y}_k - x_k)\,|\bar{y}_k - x_k|^2}{b_0\|\bar{y}\|_\infty + \mathcal{S}\left(\mathcal{S}^+(\bar{y} - x)\right)^3}, \qquad k = 1, \dots, N, \tag{2.62}$$

where $\bar{y} \in \mathbb{R}^N$, $b_0 > 0$, $\mathcal{S}\colon \mathbb{R}^N \to \mathbb{R}$ and $\mathcal{S}^+\colon \mathbb{R}^N \to \mathbb{R}^N$, with ($x \in \mathbb{R}^N$)

$$\mathcal{S}(x) = \frac{\sum_{i=1}^N x_i e^{a x_i}}{\sum_{i=1}^N e^{a x_i}}; \qquad \mathcal{S}^+(x)_k = \frac{x_k e^{a x_k} - x_k e^{-a x_k}}{e^{a x_k} + e^{-a x_k}},\ k = 1, \dots, N.$$

The maps $\mathcal{S}, \mathcal{S}^+$ are smooth approximations of the maximum function and replace the infinity norm in (2.62), allowing $B_0 \in C_b^3\left(\mathbb{R}^N; \mathbb{R}^N\right)$, coherently with our theoretical framework. Therefore B_0 is to be interpreted as a cubic nonlinearity with a cutoff for large values of $\|x\|_\infty$. For our experiments, we consider $b_0 = 2$, $\bar{y} = 2\mathbf{e}$ and $a = 10^4$. In Tables 2.3-2.4 we report the outcomes of simulations with and without f, respectively, when $\sigma = 0.7$, $t = 1$ and α varies in $\left(\frac{1}{2}, 1\right)$. In particular, Table 2.3 shows that, in the case of time–shift, the first iterate always remarkably outperforms the linear approximation.

Table 2.3: First order approximation in the bounded cubic case with time–shift; noise strength $\sigma = 0.7$.

α	$\mathbb{P}(\lvert X_1^{0,\mathbf{e}}\rvert > 1)$	$v_0^0(1,\mathbf{e})$	ϵ_r^0	$v_0^1(1,\mathbf{e})$	ϵ_r^1
0.55	0.501	0.562	−0.122	−5.19e-2	−1.82e-2
0.65	0.531	0.594	−0.119	−6.65e-2	6.59e-3
0.75	0.587	0.648	−0.104	−6.40e-2	5.11e-3
0.85	0.679	0.743	−9.43e-2	−7.95e-2	2.28e-2

Table 2.4: Same setting as in Table 2.3, without time–shift ($f \equiv 0$).

α	$\mathbb{P}(\lvert X_1^{0,\mathbf{e}}\rvert > 1)$	$v_0^0(1,\mathbf{e})$	ϵ_r^0	$v_0^1(1,\mathbf{e})$	ϵ_r^1	$v_0^2(1,\mathbf{e})$	ϵ_r^2
0.55	0.495	0.374	0.244	−9.64e-4	0.246	0.109	2.62e-2
0.65	0.536	0.396	0.261	−2.74e-2	0.312	0.142	4.74e-2
0.75	0.586	0.462	0.212	−8.16e-2	0.351	0.191	2.49e-2
0.85	0.680	0.608	0.106	−8.20e-2	0.226	0.138	2.35e-2

On the contrary, when $f \equiv 0$ (Table 2.4), $v_0^1(1,\mathbf{e})$ deteriorates the OU estimate, and we are forced to implement the second iterate to get an accuracy similar to the one provided by the time–shift (compare the columns ϵ_r^1, Table 2.3, and ϵ_r^2, Table 2.4). Of course, the trade–off in the introduction of $v_0^2(1,\mathbf{e})$ consists in substantially increasing the computational time.
Finally, in Figure 2.3 we investigate the trajectories of $P_{0,\cdot}u_0(\mathbf{e})$ and of the first order approximation in the time interval $[0,1]$, as well as the corresponding absolute relative errors. Here we fix $\alpha = 0.6$ and consider two strengths of noise: $\sigma = 0.1$ and $\sigma = 1.3$. As already observed in the sine case, the advantages in introducing the first iterate are rather evident. Overall, we conclude that $v_0^1(\cdot,\mathbf{e})$ proves to be a versatile and computationally cheap method to improve on the performances of the linear approximation.

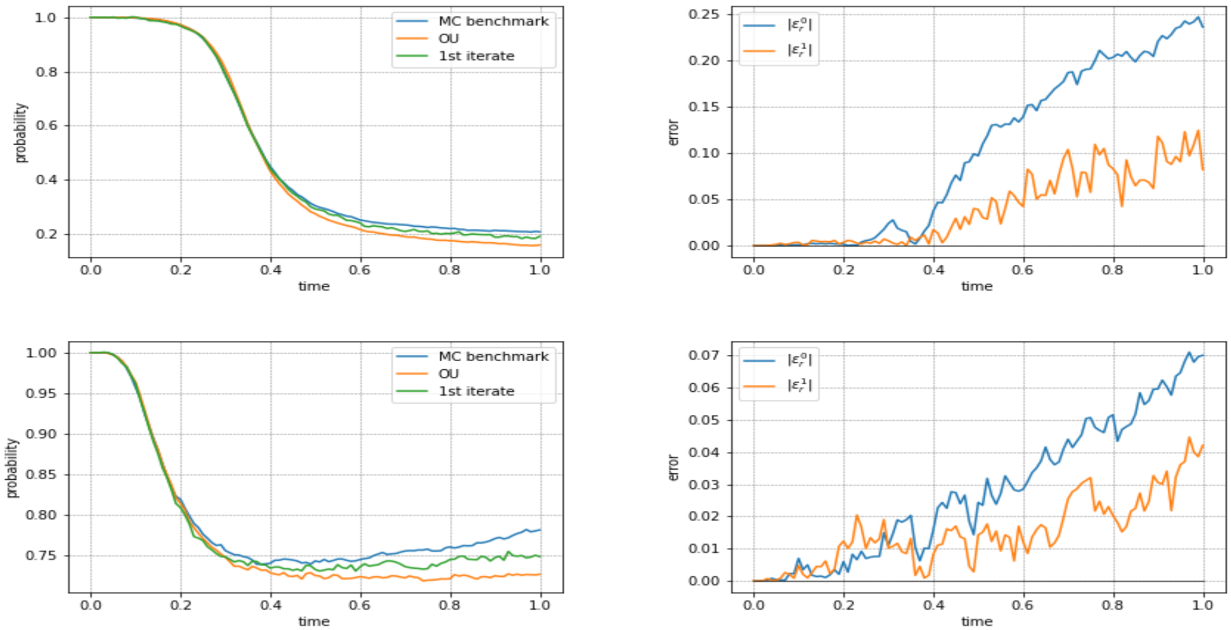

Figure 2.2: *Behavior in time of the first order approximation in the sine case with time–shift. In each line, the panel on the left shows the evolution of the probabilities, and the one on the right the corresponding errors. The top line refers to* $\sigma = 0.1$*, the bottom line to* $\sigma = 1.3$*.* $\alpha = 0.6$ *everywhere.*

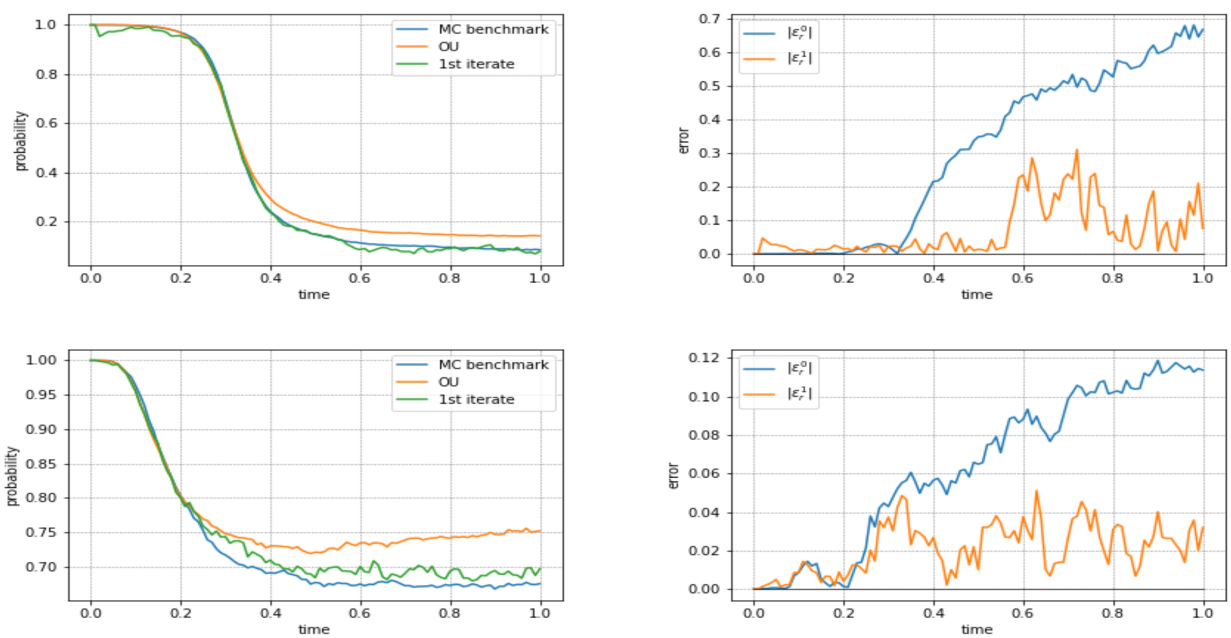

Figure 2.3: *Behavior in time of the first order approximation in the bounded cubic case with time–shift. In each line, the panel on the left shows the evolution of the probabilities, and the one on the right the corresponding errors. The top line refers to* $\sigma = 0.1$*, the bottom line to* $\sigma = 1.3$*.* $\alpha = 0.6$ *everywhere.*

Appendix 2.A Proof of Lemma 2.4

In this appendix we provide the proof of Lemma 2.4, a useful result for the arguments of Section 2.2.

Proof of Lemma 2.4. Let us fix $0 \le s \le T$, $x \in \mathbb{R}^N$ and a direction $h \in \mathbb{R}^N$; note that all the assertions of the statement are true for $\omega \in \Omega \setminus \Omega'$ by construction of the stochastic flow, hence we only focus on $\omega \in \Omega'$. For every $\epsilon \in (0,1]$ and $t \in [s,T]$ define the incremental ratio function

$$\begin{aligned} Y^1_{x,h}(\epsilon,t) &= \epsilon^{-1}\left(X^{s,x+\epsilon h}_t(\omega) - X^{s,x}_t(\omega)\right) \\ &= h + \int_s^t \left(AY^1_{x,h}(\epsilon,r) + \frac{B_0\left(r, X^{s,x}_r(\omega) + \epsilon Y^1_{x,h}(\epsilon,r)\right) - B_0\left(r, X^{s,x}_r(\omega)\right)}{\epsilon}\right) dr \\ &= h + \int_s^t \left(A + \int_0^1 DB_0\left(r, X^{s,x}_r(\omega) + \rho\epsilon Y^1_{x,h}(\epsilon,r)\right) d\rho\right) Y^1_{x,h}(\epsilon,r)\, dr. \end{aligned} \tag{2.63}$$

Notice that, for every $\epsilon \in (0,1]$ (omitting ω to keep notation short)

$$\left|X^{s,x+\epsilon h}_t - X^{s,x}_t\right| \le \epsilon\,|h| + \left(|A| + \|DB_0\|_{T,\infty}\right)\int_s^t \left|X^{s,x+\epsilon h}_r - X^{s,x}_r\right| dr, \quad t \in [s,T],$$

where we recall that $\|DB_0\|_{T,\infty} = \sup_{0\le t\le T}\|DB_0(t,\cdot)\|_\infty$. Thus, an application of Gronwall's lemma shows that $\left|Y^1_{x,h}(\epsilon,t)\right| \le |h|\, e^{\left(|A|+\|DB_0\|_{T,\infty}\right)T} =: C_1$ for all $t \in [s,T]$ and $\epsilon \in (0,1]$. Next, taking ϵ_1, $\epsilon_2 \in (0,1]$ and $t \in [s,T]$ we compute from (2.63)

$$\begin{aligned} \left|Y^1_{x,h}(\epsilon_2,t) - Y^1_{x,h}(\epsilon_1,t)\right| &\le \int_s^t |A|\left|Y^1_{x,h}(\epsilon_2,r) - Y^1_{x,h}(\epsilon_1,r)\right| dr \\ &\quad + \left|\int_s^t \left(\int_0^1 DB_0\left(r, X^{s,x}_r + \rho\epsilon_2 Y^1_{x,h}(\epsilon_2,r)\right) d\rho\, Y^1_{x,h}(\epsilon_2,r)\right.\right. \\ &\quad \left.\left. - \int_0^1 DB_0\left(r, X^{s,x}_r + \rho\epsilon_1 Y^1_{x,h}(\epsilon_1,r)\right) d\rho\, Y^1_{x,h}(\epsilon_1,r)\right) dr\right| \\ &\le \left(|A| + \|DB_0\|_{T,\infty} + \frac{N^2}{2}C_1\left\|\partial^2 B_0\right\|_{T,\infty}\right)\int_s^t \left|Y^1_{x,h}(\epsilon_2,r) - Y^1_{x,h}(\epsilon_1,r)\right| dr \\ &\quad + \frac{N^2}{2}C_1^2 T\left\|\partial^2 B_0\right\|_{T,\infty}|\epsilon_2 - \epsilon_1|, \end{aligned} \tag{2.64}$$

where $\left\|\partial^2 B_0\right\|_{T,\infty} = \sup_{0\le t\le T}\left\|\partial^2 B_0(t,\cdot)\right\|_\infty$. Therefore another application of Gronwall's lemma shows that the mapping $\epsilon \mapsto Y^1_{x,h}(\epsilon,t)$ is Lip–continuous in $(0,1]$ uniformly in $t \in [s,T]$, and by the theorem of extension of uniformly continuous functions we obtain the existence of $D_h X^{s,x}_t(\omega)$. Now by dominated convergence we are allowed to pass to the limit in (2.63), which yields

$$D_h X^{s,x}_t(\omega) = h + \int_s^t \left(A + DB_0\left(r, X^{s,x}_r(\omega)\right)\right) D_h X^{s,x}_r(\omega)\, dr, \quad t \in [s,T]. \tag{2.65}$$

Given the arbitrarity of h, $x \in \mathbb{R}^N$, this equation shows that the mapping $x \mapsto X^{s,x}_t(\omega)$ belongs to $C^1\left(\mathbb{R}^N\right)$, with $\left\|DX^{s,\cdot}_t(\omega)\right\|_\infty \le N\exp\left\{\left(|A| + \|DB_0\|_{T,\infty}\right)T\right\}$.

In order to analyze higher–order derivatives, we work by induction; fix $m = 1, \ldots, n-1$ and suppose as inductive hypothesis that $X_t^{s,\cdot}(\omega) \in C^m(\mathbb{R}^N)$, $t \in [s,T]$, with the estimate in (2.17) holding true for a sum from $i = 1$ to $i = m$. Moreover, assume that for every multi–index $\mathbf{h} \in (\mathbb{N} \cup \{0\})^N$ with length $1 \le \|\mathbf{h}\|_1 \le m$ one has, for any $t \in [s,T]$ (omitting ω)

$$D_{\mathbf{h}} X_t^{s,x} = \delta_{\mathbf{h}} + \int_s^t \left(\left(A + DB_0\left(r, X_r^{s,x}\right)\right) D_{\mathbf{h}} X_r^{s,x} + \mathcal{L}_{\mathbf{h}}(r,x) \right) dr,$$
$$\delta_{\mathbf{h}} = \begin{cases} e_j, & \text{if } \|\mathbf{h}\|_1 = 1 \text{ and } h_j = 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{2.66}$$

Here $(e_j)_{j=1,\ldots,N}$ is the canonical basis of $\mathbb{R}^N$ and $\mathcal{L}_{\mathbf{h}}(t,x) = (\mathcal{L}_{\mathbf{h},j}(t,x))_{j=1,\ldots,N}$, with $\mathcal{L}_{\mathbf{h},j}(t,x) \in \mathbb{R}$ denoting a sum of products where one factor is a (partial) derivative at $X_t^{s,x}$ of $B_{0,j}(t,\cdot)$ up to order $\|\mathbf{h}\|_1$ and the others are (partial) derivatives at x of $X_t^{s,\cdot}$ up to order $\|\mathbf{h}\|_1 - 1$. In particular, $\mathcal{L}_{\mathbf{h}}(t,x) = 0$ when $\|\mathbf{h}\|_1 = 1$ (cfr. (2.65)). At this point, consider x, $h \in \mathbb{R}^N$ and fix a multi–index $\mathbf{h}$ with length $\|\mathbf{h}\|_1 = m$; by analogy with (2.63), for any $\epsilon \in (0,1]$ and $t \in [s,T]$ define the incremental ratio function

$$\begin{aligned} Y_{x,h}^{m+1}(\epsilon,t) &= \epsilon^{-1}\left(D_{\mathbf{h}} X_t^{s,x+\epsilon h} - D_{\mathbf{h}} X_t^{s,x}\right) \\ &= \int_s^t \Bigg(\left(A + DB_0\left(r, X_r^{s,x}\right)\right) Y_{x,h}^{m+1}(\epsilon,r) + \epsilon^{-1}\left(\mathcal{L}_{\mathbf{h}}(r, x+\epsilon h) - \mathcal{L}_{\mathbf{h}}(r,x)\right) \\ &\quad + \frac{DB_0\left(r, X_r^{s,x+\epsilon h}\right) - DB_0\left(r, X_r^{s,x}\right)}{\epsilon} D_{\mathbf{h}} X_r^{s,x+\epsilon h} \Bigg) dr. \end{aligned}$$

Note that for any $j = 1, \ldots, N$ we can write ($t \in [s,T]$, $\epsilon \in (0,1]$)

$$\epsilon^{-1}\left(DB_0\left(t, X_t^{s,x+\epsilon h}\right) - DB_0\left(t, X_t^{s,x}\right)\right)_{j,\cdot} = \left(\left(\int_0^1 D^2 B_{0,j}\left(t, X_t^{s,x} + \rho\epsilon Y_{x,h}^1(\epsilon,t)\right) d\rho \right) Y_{x,h}^1(\epsilon,t) \right)^{\top},$$

and that, further, the inductive hypothesis of boundedness for the derivatives of $X_t^{s,\cdot}$ (see (2.17)), together with the structure of $\mathcal{L}_{\mathbf{h}}$ and $B_0 \in C_b^{m+1}\left([0,T] \times \mathbb{R}^N; \mathbb{R}^N\right)$ ensures that

$$\epsilon^{-1}\left|\mathcal{L}_{\mathbf{h}}(t, x+\epsilon h) - \mathcal{L}_{\mathbf{h}}(t,x)\right| \le C_2 |h|, \quad t \in [s,T], \epsilon \in (0,1],$$

for some constant $C_2 = C_2(A, B_0, T, m, N) > 0$. These facts, the Lip–continuity of the map $\epsilon \mapsto Y_{x,h}^1(\epsilon,t)$ in $(0,1]$ uniformly in $t \in [s,T]$ and computations analogous to those in (2.64) entail that there exists $D_h D_{\mathbf{h}} X_t^{s,x}(\omega)$. The arbitrarity of x, h and $\mathbf{h}$ coupled with Gronwall's lemma provides us with the desired bound (2.17) for the derivatives of order $m+1$, and finally by dominated convergence the validity of (2.66) for a multi–index of length $m+1$ is a consequence of the chain rule. In particular, $X_t^{s,\cdot}(\omega) \in C^{m+1}(\mathbb{R}^N)$. The proof is then complete, considering that the base case is provided by (2.65). ∎

Chapter 3

A sharp càdlàg property for jump diffusions and dynamic programming principle

In this chapter, we prove the existence of a version of the stochastic flow $X = (X_t^{s,x})_{t \geq s}$ generated by (I.5) which depends in a regular way on the variables (s, t, x). In particular, the flow X is *sharp* in the following sense: there exists an almost sure event Ω' such that, for every $\omega \in \Omega'$, the map $(s, x, t) \mapsto X_t^{s,x}(\omega)$ is càdlàg in s (for t and x fixed), càdlàg in t (for s and x fixed) and continuous in x (for s and t fixed). In the case of SDEs with only small jumps, i.e., with $f \equiv 0$, this result solves an open problem which also appears in [126]. In our proof, we deal with non–separable spaces of càdlàg functions involving supremum norms and, when $f \equiv 0$, we employ an extension of the càdlàg criterion in [28], which is proved in the appendix, see Appendix 3.B. We then extend our approach to encompass controlled SDEs having also a large–jumps component. Using our sharp stochastic flow we obtain a new dynamic programming principle, whose proof is of independent interest.

3.1 Preliminaries and main results on the sharp càdlàg property

In this chapter, $|\cdot|$ denotes the Euclidean norm in any $\mathbb{R}^m$, $m \geq 1$. Fix $T > 0$ and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ satisfying the usual hypotheses. On this probability space, we take an $m-$dimensional $\mathbb{F}-$Brownian motion $W = (W_t)_{0 \leq t \leq T}$. Moreover, given a measurable space $(U, \mathcal{U})$, we consider a stationary Poisson point process p on U with intensity measure $dt \otimes \nu(dz)$, where $\nu(dz)$ is a $\sigma-$finite measure on U (see [106, Section 9, Chapter I]). In particular, for every $\omega \in \Omega$, $p(\omega) \colon D_p(\omega) \to U$, where $D_p(\omega)$ is a countable subset of $(0, \infty)$. Let N_p be the counting measure associated with p, namely

$$N_p\left((0,t] \times V\right)(\omega) = \#\left\{s \in D_p(\omega) \cap (0,t] : [p(\omega)](s) \in V\right\}, \quad t > 0,\, V \in \mathcal{U},\, \omega \in \Omega;$$

this is a Poisson random measure on $(0, \infty) \times U$. In the sequel, we write $p_s(\omega) = [p(\omega)](s)$ to have a compact notation. We denote by $\widetilde{N}_p(dt, dz) = N_p(dt, dz) - dt \otimes \nu(dz)$ the compensated Poisson random measure. We suppose that p is $\mathbb{F}-$adapted, in the sense that $N_p((0,t] \times V)$ is $\mathcal{F}_t-$measurable for every $V \in \mathcal{U}$ and $t > 0$.

Fix a measurable set $U_0 \in \mathcal{U}$ such that $\nu(U \setminus U_0) \in (0, \infty)$. We are going to consider SDEs with coefficients satisfying the next requirements.

Hypothesis 3.1. *Consider the drift coefficient* $b(t,x) = (b_j(t,x))_{j=1,\dots,d}$, *the diffusion matrix* $\alpha(t,x) = (\alpha_{i,j}(t,x))_{i=1,\dots,d;\, j=1,\dots,m}$ *and the small–jumps coefficient* $g(x,t,z) = (g_j(x,t,z))_{j=1,\dots,d}$. *We require* $b\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$, $\alpha\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ *and* $g\colon \mathbb{R}^d\times[0,T]\times U \to \mathbb{R}^d$ *to be jointly measurable in their domains.*

We assume that b, α *and* g *satisfy linear growth and Lipschitz–type conditions, see [106]. More precisely, for every* $p \geq 2$, *there exists a constant* K_p *such that*

$$|b(t,x)|^p + |\alpha(t,x)|^p + \int_{U_0} |g(x,t,z)|^p\,\nu(dz) \leq K_p(1+|x|^p), \quad x\in\mathbb{R}^d,\ t\in[0,T], \tag{3.1}$$

and

$$|b(t,x)-b(t,y)|^p + |\alpha(t,x)-\alpha(t,y)|^p + \int_{U_0} |g(x,t,z)-g(y,t,z)|^p\,\nu(dz) \leq K_p|x-y|^p, \quad x,y\in\mathbb{R}^d,\ t\in[0,T]. \tag{3.2}$$

Here, $|\alpha|^2 = \sum_{i,j}|\alpha_{i,j}|^2$. *We also consider a large–jumps coefficient* $f\colon \mathbb{R}^d\times[0,T]\times U\to\mathbb{R}^d$, *supposing that* f *is a jointly measurable function which is continuous in the first argument.*

In this chapter, we study the SDE (cfr. (I.5) in Introduction)

$$X_t = x + \int_s^t b(r,X_r)\,dr + \int_s^t \alpha(r,X_r)\,dW_r + \int_s^t\int_{U_0} g\left(X_{r-},r,z\right)\widetilde{N}_p(dr,dz) + \int_s^t\int_{U\setminus U_0} f\left(X_{r-},r,z\right)N_p(dr,dz), \quad t\in[s,T], \tag{3.3}$$

where $s\in[0,T)$ and $x\in\mathbb{R}^d$. In particular, the small–jumps case $f\equiv 0$ is investigated in Section 3.2, while the large–jumps case $f\not\equiv 0$ is analyzed in Section 3.3.
A solution to (3.3) is a càdlàg, $\mathbb{R}^d-$valued, $\mathbb{F}-$adapted process $X=(X_t)_{s\leq t\leq T}$ satisfying (3.3) up to indistinguishability. We extend the trajectories of X in the whole interval $[0,T]$ by setting $X_t = X_s$, $t\in[0,s]$. Under Hypothesis 3.1, it is known that (3.3) admits a pathwise unique solution for every initial condition (s,x) (see Sections 3.2-3.3 for the details). Our goal is to prove the existence of a *sharp* version of the solution X which is simultaneously càdlàg in the time variables s,t, and continuous in the space variable x; furthermore, X is stochastically continuous in s. More precisely, we search for a *sharp stochastic flow* generated by (3.3) according to the following definition, where we denote by $\mathcal{D}_0$ the complete metric space of $\mathbb{R}^d-$valued, càdlàg functions on $[0,T]$ endowed with the uniform norm.

Definition 3.1. Let $X\colon\Omega\times[0,T]\times\mathbb{R}^d\times[0,T]\to\mathbb{R}^d$ be an $\mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}^d\times[0,T])-$measurable function and denote by $X_t^{s,x}(\omega) = X(\omega,s,x,t)$. We say that X is the *sharp stochastic flow* generated by (3.3) if there exists an a.s. event Ω' –independent from s,t,x– such that the four following requirements are fulfilled for every $\omega\in\Omega'$, $s\in[0,T)$ and $x\in\mathbb{R}^d$.

1. The process $(X_t^{s,x})_{t\in[s,T]}$ satisfies (3.3) in Ω';
2. (i) The map $X_\cdot^{s,x}(\omega):[0,T]\to\mathbb{R}^d$ is càdlàg;
   (ii) The map $X^{s,\cdot}(\omega):\mathbb{R}^d\to\mathcal{D}_0$ is continuous;
   (iii) The map $X^{\cdot,x}(\omega):[0,T]\to\mathcal{D}_0$ is càdlàg, locally uniformly in x;

3. The flow property holds: $X_t^{s,x}(\omega) = X_t^{u, X_u^{s,x}(\omega)}(\omega), \quad s < u < t \le T.$

4. The function X is stochastically continuous in the following sense: for every $\epsilon > 0$ and $M > 0$,

$$\lim_{r \to s} \mathbb{P}\Big(\sup_{|x| \le M} \sup_{0 \le t \le T} |X_t^{r,x} - X_t^{s,x}| > \epsilon \Big) = 0, \quad s \in [0,T]. \tag{3.4}$$

Notice that, by the pathwise uniqueness of (3.3) and Point 2. in Definition 3.1, a sharp stochastic flow generated by (3.3) is unique up to an a.s. event. The next theorem shows the existence of the sharp stochastic flow associated with (3.3).

Theorem 3.1. *Under Hypothesis 3.1, the sharp stochastic flow generated by* (I.5) *exists.*

When $f \equiv 0$, i.e., in the small–jumps case, we deduce the previous result from a stronger one, which is presented in Theorem 3.2 after introducing some notations. Let $\mathcal{C}_0 = (C(\mathbb{R}^d; \mathcal{D}_0), d_0^{lu})$ be the metric space of continuous functions defined on $\mathbb{R}^d$ with values in $\mathcal{D}_0$ with the usual distance d_0^{lu} (defined below in (3.20)). We endow $\mathcal{C}_0$ with the $\sigma-$algebra $\mathcal{C}$ generated by the projections $\pi_x \colon C(\mathbb{R}^d; \mathcal{D}_0) \to (\mathcal{D}_0, \mathcal{D})$, $x \in \mathbb{R}^d$, defined by $\pi_x(f) = f(x)$, $f \in \mathcal{C}_0$. Here $\mathcal{D}$ is the $\sigma-$algebra on $\mathcal{D}_0$ generated by the Skorokhod topology (see the discussion around (3.11)).

Theorem 3.2. *When $f \equiv 0$, under Hypothesis 3.1, the sharp stochastic flow generated by* (I.5) *exists and is a stochastically continuous, càdlàg $(\mathcal{C}_0, \mathcal{C})-$valued process.*

Remark 3.1. *Despite some differences with the assertions in Theorem 3.1, [115, Theorem 5] gives a version of the solution of the SDE* (I.7) *which is sharp in the variables (s,t,x) and satisfies the flow property. [115, Theorem 5] can be applied to the SDE* (3.3) *when the coefficients are time–independent and g and f have a special form; it cannot be used to study controlled SDEs. More precisely, α and b must be time–independent and $g(x,r,z) = g_1(x) g_2(r,z)$, $f(x,r,z) = f_1(x) f_2(r,z)$, where f_2, g_2 are measurable in their domains with values in $\mathbb{R}^k$ and g_2 verifies $\int_0^T dr \int_{U_0} |g_2(r,z)|^2 \nu(dz) < \infty$. Moreover, one has to require that $g_1, f_1 : \mathbb{R}^d \to \mathbb{R}^{d \times k}$ are Lipschitz continuous. In this case,*

$$Z_t = \Big(t, W_t, \int_0^t \int_{U_0} g_2(r,z) \widetilde{N}_p(dr,dz), \int_0^t \int_{U \setminus U_0} f_2(r,z) N_p(dr,dz) \Big) \in \mathbb{R}^{1+m+2k}.$$

The proof of [115, Theorem 5] is very different from the one of Theorem 3.1, which relies on Theorem 3.2. On the other hand, in [115] there are no results related to Theorem 3.2 (the space $\mathcal{C}_0 = (C(\mathbb{R}^d; \mathcal{D}_0), d_0^{lu})$ is introduced in this chapter).

3.2 SDEs with small jumps

In this section, we are interested in the study of (3.3) with $f \equiv 0$, i.e., the SDE with *small jumps.* In particular, we consider the following SDE:

$$Y_t = \eta + \int_s^t b(r, Y_r)\, dr + \int_s^t \alpha(r, Y_r)\, dW_r + \int_s^t \int_{U_0} g(Y_{r-}, r, z) \widetilde{N}_p(dr,dz), \quad t \in [s,T], \tag{3.5}$$

where $s \in [0,T)$ and

$$\eta \in L^0(\mathcal{F}_s)$$

i.e., η is an $\mathcal{F}_s-$measurable random variable with values in $\mathbb{R}^d$. A solution of this equation is a càdlàg, $\mathbb{R}^d-$valued, $\mathbb{F}-$adapted process $Y = (Y_t)_{s\leq t\leq T}$ satisfying (3.5) up to indistinguishability.

Conditions (3.1)-(3.2) guarantee the existence of a solution Y to (3.5) for every $\eta \in L^p(\Omega)\cap L^0(\mathcal{F}_s)$, with $p \geq 2$, see [125, Theorem 3.1] or [167, Theorem 117]. Such a solution is pathwise unique and satisfies

$$\mathbb{E}\big[\sup_{s\leq t\leq T} |Y_t|^p\big] < \infty.$$

We denote by $Y^{s,\eta}$ the solution of (3.5) starting from η at time s. We also set $Y^{s,\eta}_r = Y^{s,\eta}_s$ if $0 \leq r < s$.

3.2.1 Flow property and continuity in x

The pathwise uniqueness of (3.5) immediately implies the *cocycle property*: for every $x \in \mathbb{R}^d$ and $0 \leq s < u \leq T$, there exists an a.s. event $\Omega_{s,u,x}$ such that

$$Y_t^{u,Y_u^{s,x}}(\omega) = Y_t^{s,x}(\omega), \quad t \in [u,T], \omega \in \Omega_{s,u,x}. \tag{3.6}$$

The notation $\Omega_{s,u,x}$ indicates an (a.s.) event which may depend on s, u and x (it is independent of t). This notation will be adopted for the rest of the chapter.

The next result is an extension of [125, Equation (3.7)] to random initial conditions, see also [126, Lemma 3.3.3]. The proof contains useful estimates (in particular, see (3.9)-(3.10)) which will be used several times hereinafter.

Lemma 3.3. *Fix $p \geq 2$. Then, for every $\mathcal{F}_s-$measurable random variables $\xi, \eta \in L^p(\Omega)$, one has*

$$\mathbb{E}\left[\sup_{s\leq t\leq T}\left|Y_t^{s,\eta} - Y_t^{s,\xi}\right|^p\right] \leq 4^{p-1}e^{c(T-s)}\mathbb{E}\left[|\eta-\xi|^p\right], \quad s \in [0,T), \tag{3.7}$$

where $c > 0$ is a constant depending only on p, d, m, T, K_2, K_p.

Proof. Fix $0 \leq s < T$ and two $\mathcal{F}_s-$measurable random variables $\xi, \eta \in L^p(\Omega)$. By (3.5), in an a.s. event $\Omega_{s,\xi,\eta}$ we have, using Hölder's inequality and the Lipschitz condition (3.2) on b,

$$\begin{aligned}&\left|Y_t^{s,\eta} - Y_t^{s,\xi}\right|^p\\ &\leq 4^{p-1}\Bigg(|\eta-\xi|^p + K_pT^{p-1}\int_s^t \sup_{s\leq u\leq r}\left|Y_u^{s,\eta} - Y_u^{s,\xi}\right|^p dr + \sup_{s\leq u\leq t}\left|\int_s^u \left(\alpha\left(r,Y_r^{s,\eta}\right) - \alpha\left(r,Y_r^{s,\xi}\right)\right)dW_r\right|^p\\ &\quad + \sup_{s\leq u\leq t}\left|\int_s^u\int_{U_0}\left(g\left(Y_{r-}^{s,\eta},r,z\right) - g\left(Y_{r-}^{s,\xi},r,z\right)\right)\widetilde{N}_p(dr,dz)\right|^p\Bigg), \quad t\in[s,T].\end{aligned}$$

Taking the supremum and the expectation we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{s\leq u\leq t}\left|Y_u^{s,\eta} - Y_u^{s,\xi}\right|^p\right] &\leq 4^{p-1}\Bigg(\mathbb{E}\left[|\eta-\xi|^p\right]\\ &\quad + K_pT^{p-1}\int_s^t \mathbb{E}\left[\sup_{s\leq u\leq r}\left|Y_u^{s,\eta} - Y_u^{s,\xi}\right|^p\right]dr + \mathbb{E}\left[\sup_{s\leq u\leq t}\left|\int_s^u\left(\alpha\left(r,Y_r^{s,\eta}\right) - \alpha\left(r,Y_r^{s,\xi}\right)\right)dW_r\right|^p\right]\\ &\quad + \mathbb{E}\left[\sup_{s\leq u\leq t}\left|\int_s^u\int_{U_0}\left(g\left(Y_{r-}^{s,\eta},r,z\right) - g\left(Y_{r-}^{s,\xi},r,z\right)\right)\widetilde{N}_p(dr,dz)\right|^p\right]\Bigg), \quad t\in[s,T].\end{aligned} \tag{3.8}$$

By the Burkholder–Davis–Gundy inequality and the Lipschitz condition (3.2) on α we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\le u\le t}\left|\int_s^u \left(\alpha\left(r,Y_r^{s,\eta}\right)-\alpha\left(r,Y_r^{s,\xi}\right)\right)dW_r\right|^p\right]\\
&\quad\le c_p\,(dm)^p\,T^{\frac{p}{2}-1}\mathbb{E}\left[\int_s^t\left|\alpha\left(r,Y_r^{s,\eta}\right)-\alpha\left(r,Y_r^{s,\xi}\right)\right|^p dr\right]\\
&\quad\le c_p\,(dm)^p\,T^{\frac{p}{2}-1}K_p\int_s^t\mathbb{E}\left[\sup_{s\le u\le r}\left|Y_u^{s,\eta}-Y_u^{s,\xi}\right|^p\right]dr,\quad t\in[s,T]\,.
\end{aligned}\tag{3.9}$$

where $c_p>0$ is a constant only depending on p. As for the integral with respect to $\widetilde{N}_p$, [125, Theorem 2.11] yields, for every $t\in[s,T]$,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\le u\le t}\left|\int_s^u\int_{U_0}\left(g\left(Y_{r-}^{s,\eta},r,z\right)-g\left(Y_{r-}^{s,\xi},r,z\right)\right)\widetilde{N}_p\left(dr,dz\right)\right|^p\right]\\
&\quad\le c_{1,p}\,d^p\left(\mathbb{E}\left[\left(\int_s^t dr\int_{U_0}\nu\left(dz\right)\left|g\left(Y_{r-}^{s,\eta},r,z\right)-g\left(Y_{r-}^{s,\xi},r,z\right)\right|^2\right)^{\frac{p}{2}}\right]\right.\\
&\qquad\left.+\ \mathbb{E}\left[\int_s^t dr\int_{U_0}\nu\left(dz\right)\left|g\left(Y_{r-}^{s,\eta},r,z\right)-g\left(Y_{r-}^{s,\xi},r,z\right)\right|^p\right]\right)\\
&\quad\le c_{1,p}\,d^p\left[T^{\frac{p}{2}-1}K_2^{\frac{p}{2}}+K_p\right]\int_s^t\mathbb{E}\left[\sup_{s\le u\le r}\left|Y_u^{s,\eta}-Y_u^{s,\xi}\right|^p\right]dr,
\end{aligned}\tag{3.10}$$

where $c_{1,p}>0$. Going back to (3.8), we combine (3.9) and (3.10) to get the existence of a constant $c=c\left(p,d,m,T,K_2,K_p\right)>0$ such that

$$\mathbb{E}\Big[\sup_{s\le u\le t}\left|Y_u^{s,\eta}-Y_u^{s,\xi}\right|^p\Big]\le 4^{p-1}\mathbb{E}\Big[|\eta-\xi|^p\Big]+c\int_s^t\mathbb{E}\Big[\sup_{s\le u\le r}\left|Y_u^{s,\eta}-Y_u^{s,\xi}\right|^p\Big]dr,\quad t\in[s,T]\,.$$

At this point Gronwall's lemma provides us with the assertion. ■

Denote by $\mathcal{D}_0=\left(\mathcal{D}\left([0,T]\,;\mathbb{R}^d\right),\|\cdot\|_0\right)$, the metric space of $\mathbb{R}^d-$valued, càdlàg functions with the uniform norm $\|\cdot\|_0$ in $[0,T]$: $\mathcal{D}_0$ is complete but not separable. Inspired by [147, Chapter V], we endow $\mathcal{D}_0$ with the $\sigma-$algebra $\mathcal{D}$ generated by the projections

$$\pi_t\colon\mathcal{D}_0\to\mathbb{R}^d,\ t\in[0,T],\quad\text{defined by }\pi_t(f)=f(t),\ f\in\mathcal{D}_0.\tag{3.11}$$

It is well known that $\mathcal{D}$ coincides with the Borel $\sigma-$algebra generated by the Skorokhod topology J_1, see [31, Theorem 12.5], [110, Theorem 1.14, Chapter VI]) and [111, Corollary 2.4]. On the contrary, $\mathcal{D}$ is strictly smaller than the Borel $\sigma-$algebra of the uniform distance (cfr. [31, Eq. (15.2)]). Notice that the difference of two càdlàg functions, considered as a mapping from $(\mathcal{D}_0\times\mathcal{D}_0,\mathcal{D}\otimes\mathcal{D})$ to $(\mathcal{D}_0,\mathcal{D})$, is measurable. Indeed, this is an immediate consequence of the measurability of the following map:

$$(\mathcal{D}_0\times\mathcal{D}_0,\mathcal{D}\otimes\mathcal{D})\to\mathbb{R}^{2d},\quad(x,y)\mapsto\left(x\left(t\right),y\left(t\right)\right),\quad t\in[0,T]\,.$$

Moreover, observe that also $\|\cdot\|_0:(\mathcal{D}_0,\mathcal{D})\to\mathbb{R}$ is measurable, because the càdlàg property allows to compute the supremum on a countable dense set of $[0,T]$.
Let us fix $s\in[0,T)$ and consider the random field $Y^{s,\cdot}=\left(Y^{s,x}\right)_{x\in\mathbb{R}^d}$. For every $x\in\mathbb{R}^d$, the map

$Y^{s,x}\colon \Omega \to (\mathcal{D}_0, \mathcal{D})$ is a random variable, i.e., it is $\mathcal{D}-$measurable. Hence, by the previous discussion the function $\omega \mapsto \sup_{0\le t\le T} |Y_t^{s,x}(\omega) - Y_t^{s,y}(\omega)|$ is measurable for every x, $y \in \mathbb{R}^d$. Thanks to (3.7), choosing $p > d$ we can apply the *Kolmogorov–Chentsov continuity criterion* as in [30, Lemma A.2.37] to find a continuous modification $\widetilde{Y}^{s,\cdot}$ of $Y^{s,\cdot}$. Hence there exist a.s. events Ω_s and $\Omega_{s,x}$ such that

$$\widetilde{Y}^{s,x} = Y^{s,x} \text{ in } \Omega_{s,x}, \quad \text{and} \quad \widetilde{Y}^{s,\cdot}(\omega) : \mathbb{R}^d \to \mathcal{D}_0 \text{ is continuous for any } \omega \in \Omega_s. \tag{3.12}$$

By setting $\widetilde{Y}_t^{s,x}(\omega) = x$ for every $x \in \mathbb{R}^d$, $t \in [0,T]$ and $\omega \in \Omega_s^c$, we get the continuity of $\widetilde{Y}^{s,\cdot}(\omega)$ for all $\omega \in \Omega$. From now on, we will always work with this continuous version, which we keep denoting by $Y^{s,\cdot}$.

The following result shows the dependence of $\mathbb{P}-$a.s. path of the solution on the initial condition.

Proposition 3.4. *For every $s \in [0,T)$ and $\eta \in L^p(\Omega) \cap L^0(\mathcal{F}_s)$, $p \ge 2$, there exists an a.s. event $\Omega_{s,\eta}$ such that*

$$Y_t^{s,\eta}(\omega) = Y_t^{s,\eta(\omega)}(\omega), \quad t \in [s,T], \omega \in \Omega_{s,\eta}. \tag{3.13}$$

Proof. Fix $p \ge 2$ and $s \in [0,T)$. First, we notice that (3.13) is an immediate consequence of the pathwise uniqueness of the solutions to (3.5) when $\eta \in L^p(\Omega) \cap L^0(\mathcal{F}_s)$ is simple, namely $\eta = \sum_{k=1}^n a_k 1_{A_k}$, where $n \in \mathbb{N}$, $a_k \in \mathbb{R}^d$ and $(A_k)_k$ is an $\mathcal{F}_s-$measurable partition of Ω, $k = 1, \ldots, n$.

Secondly, we consider $\eta \in L^p(\Omega) \cap L^0(\mathcal{F}_s)$ and take a sequence $(\eta_n)_n$ of simple, $\mathcal{F}_s-$measurable random variables converging to it both in the L^p-sense and almost surely. By the previous step, we can find an a.s. event $\Omega'_{s,\eta}$ (independent from n) such that

$$Y^{s,\eta_n}(\omega) = Y^{s,\eta_n(\omega)}(\omega), \quad n \in \mathbb{N}, \omega \in \Omega'_{s,\eta}.$$

Without loss of generality, suppose that $\eta_n \to \eta$ pointwise on $\Omega'_{s,\eta}$. The continuity of the random field $(Y^{s,x})_{x\in\mathbb{R}^d}$ in Ω yields

$$\mathcal{D}_0 - \lim_{n\to\infty} Y^{s,\eta_n(\omega)}(\omega) = Y^{s,\eta(\omega)}(\omega), \quad \omega \in \Omega'_{s,\eta}.$$

On the other hand, an application of Lemma 3.3 shows that, possibly passing to a subsequence,

$$\mathcal{D}_0 - \lim_{k\to\infty} Y^{s,\eta_{n_k}}(\omega) = Y^{s,\eta}(\omega), \quad \omega \in \Omega_{s,\eta},$$

with an a.s. event $\Omega_{s,\eta} \subset \Omega'_{s,\eta}$. By the three previous assertions we infer (3.13) and the proof is complete. ■

Combining the cocycle property in (3.6) with Proposition 3.4 we get the *flow property* expressed in the next corollary. This result improves (3.6), because in (3.6) the a.s. event $\Omega_{s,u,x}$ possibly depends on $x \in \mathbb{R}^d$.

Corollary 3.5. *For every $0 \le s < u \le T$, there exists an a.s. event $\Omega_{s,u}$ such that*

$$Y_t^{u,Y_u^{s,x}(\omega)}(\omega) = Y_t^{s,x}(\omega), \quad t \in [u,T], x \in \mathbb{R}^d, \omega \in \Omega_{s,u}. \tag{3.14}$$

Proof. Fix $0 \le s < u \le T$. Equations (3.6)-(3.13) imply, for every $x \in \mathbb{R}^d$, the existence of an a.s. event $\Omega_{s,u,x}$ where (3.14) holds. Therefore, it is sufficient to remove the dependence of this event from x to prove the assertion. Let $\Omega_{s,u} = \cap_{x\in\mathbb{Q}^d} \Omega_{s,u,x}$; then $\mathbb{P}(\Omega_{s,u}) = 1$ and

$$Y_t^{u,Y_u^{s,x}(\omega)}(\omega) = Y_t^{s,x}(\omega), \quad t \in [u,T], x \in \mathbb{Q}^d, \omega \in \Omega_{s,u}. \tag{3.15}$$

For a point $x \in \mathbb{R}^d \setminus \mathbb{Q}^d$, take a sequence $(x_n)_n \subset \mathbb{Q}^d$ such that $x_n \to x$. Given $\omega \in \Omega_{s,u}$, by the continuity of the random field $Y^{s,\cdot}$ (see (3.12) and the subsequent comment), one can pass to the limit in (3.15) to show that (3.15) holds in x, as well. This gives (3.14), completing the proof. ■

Observe that each process $Y^{s,x}$, $x \in \mathbb{R}^d$, satisfies (3.5) in an a.s. event $\Omega_{s,x}$: we now want to find a common a.s. event Ω_s –independent from x– where (3.5) holds (with $\eta = x$). To do this, we consider suitable modifications of the stochastic integrals. We start off by taking $\Omega'_s = \cap_{x\in\mathbb{Q}^d}\Omega_{s,x}$, so that, for every $\omega \in \Omega'_s$,

$$
\begin{aligned}
&Y_t^{s,x} - x - \int_s^t b\left(r, Y_r^{s,x}\right) dr \\
&\qquad = \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right), \quad t \in [s,T], \, x \in \mathbb{Q}^d. \qquad (3.16)
\end{aligned}
$$

We can now construct two continuous random fields corresponding to the addends on the right–hand side of the previous equation. Specifically, for the term $(Z_1^{s,x})_{x\in\mathbb{R}^d} = \left(\int_s^\cdot \alpha\left(r, Y_r^{s,x}\right) dW_r\right)_{x\in\mathbb{R}^d}$, we combine (3.7) with the estimate in (3.9) to write, for any $p > d$,

$$
\mathbb{E}\Big[\sup_{s\le t\le T}\left|\int_s^t \left(\alpha\left(r, Y_r^{s,y}\right) - \alpha\left(r, Y_r^{s,x}\right)\right) dW_r\right|^p\Big] \le C\left|x-y\right|^p, \quad x,y \in \mathbb{R}^d,
$$

where $C = C\left(T, p, d, m, K_2, K_p\right) > 0$. Hence the Kolmogorov–Chentsov criterion ensures the existence of a version of this random field which is continuous in an a.s. event Ω''_s. We set this modification equal to 0 outside Ω''_s, so that it is continuous in the whole space Ω, and keep denoting it by $Z_1^{s,x} = \int_s^\cdot \alpha\left(r, Y_r^{s,x})\right) dW_r$. Moreover, we can think of $(Z_1^{s,x})_x$ as a continuous, $(\mathcal{D}_0, \mathcal{D})-$valued random field defining $Z_{1,t}^{s,x}(\omega) = Z_{1,s}^{s,x}(\omega)$, for every $t \in [0,s)$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$.

As for $Z_2^{s,x} = \int_s^\cdot\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)$, $x \in \mathbb{R}^d$, the argument to obtain a $(\mathcal{D}_0, \mathcal{D})-$valued, continuous modification is the same once we consider the estimate in (3.10). This construction ensures that, in an a.s. event $\Omega_s \subset \Omega'_s$, (3.16) holds with the right–hand side being the sum of continuous random fields.

Finally, it is easy to see that $(Z_3^{s,x})_{x\in\mathbb{R}^d} = \left(\int_s^\cdot b\left(r, Y_r^{s,x}\right) dr\right)_{x\in\mathbb{R}^d}$ is a continuous random field in Ω by the continuity of $Y^{s,\cdot}$ and the condition (3.2) on b. If we define $Z_3^{s,x}(\omega)(t) = 0$, $t \in [0,s)$, $\omega \in \Omega$, $x \in \mathbb{R}^d$, then $(Z_{3,x}^s)_x$ is a continuous, $(\mathcal{D}_0, \mathcal{D})-$valued random field. Going back to (3.16), we deduce that

$$
\begin{aligned}
&Y_t^{s,x} = x + \int_s^t b\left(r, Y_r^{s,x}\right) dr + \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right), \\
&\qquad t \in [0,T], \, x \in \mathbb{R}^d, \, \omega \in \Omega_s. \qquad (3.17)
\end{aligned}
$$

We conclude this subsection with a corollary showing the consequences of the cocycle property (3.6) of $(Y^{s,x})_{x\in\mathbb{R}^d}$ (see also Corollary 3.5) on the continuous vector fields $(Z_i^{s,x})_{x\in\mathbb{R}^d}$, $i = 1,2,3$.

Corollary 3.6. *For every $s \in [0,T)$ and $\eta \in L^p(\Omega) \cap L^0(\mathcal{F}_s)$, $p \ge 2$, there exists an a.s. event $\Omega_{s,\eta}$ such that*

$$
\begin{aligned}
&\left(\int_s^t \alpha\left(r, Y_r^{s,\eta}\right) dW_r\right)(\omega) = Z_{1,t}^{s,\eta(\omega)}(\omega), \quad \left(\int_s^t \int_{U_0} g\left(Y_{r-}^{s,\eta}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right)(\omega) = Z_{2,t}^{s,\eta(\omega)}(\omega), \\
&\int_s^t b\left(r, Y_r^{s,\eta}(\omega)\right) dr = Z_{3,t}^{s,\eta(\omega)}(\omega), \qquad (3.18)
\end{aligned}
$$

for all $t \in [s,T]$ and $\omega \in \Omega_{s,\eta}$. Furthermore, for every $u \in (s,T]$, there exists an a.s. event $\Omega_{s,u}$ such that

$$
Z_{i,u}^{s,x} + Z_{i,t}^{u,Y_u^{s,x}(\omega)}(\omega) = Z_{i,t}^{s,x}(\omega), \quad t \in [u,T], \, x \in \mathbb{R}^d, \, i = 1,2,3, \, \omega \in \Omega_{s,u}. \qquad (3.19)
$$

Proof. The equalities in (3.18) can be inferred with the same argument as in the proof of Proposition 3.4, recalling the estimates (3.8), (3.9) and (3.10) in the proof of Lemma 3.3.
As for (3.19), we focus only on Z_1, being the other cases analogous. Fix $0 \le s < u \le T$ and compute, by the cocycle property in (3.6) and (3.18),

$$Z^{s,x}_{1,t}(\omega) = \left(\int_s^u \alpha\left(r, Y^{s,x}_r\right) dW_r + \int_u^t \alpha\left(r, Y^{u, Y^{s,x}_u}_r\right) dW_r \right)(\omega) = Z^{s,x}_{1,u}(\omega) + Z^{u, Y^{s,x}_u(\omega)}_{1,t}(\omega),$$

which holds for every $t \in [u,T]$, $x \in \mathbb{Q}^d$ and $\omega \in \Omega_{s,u}$, where $\Omega_{s,u}$ is an a.s. event independent from x. In fact, the previous equation is valid also for $x \in \mathbb{R}^d \setminus \mathbb{Q}^d$, by the continuity of the $(\mathcal{D}_0, \mathcal{D})-$valued random fields $Y^{s,\cdot}$, $Z^{s,\cdot}_1$ and $Z^{u,\cdot}_1$. Hence we recover (3.19), completing the proof. ■

3.2.2 The stochastic continuity in the initial time s

Let $\mathcal{C}_0 = (C(\mathbb{R}^d; \mathcal{D}_0), d^{lu}_0)$ be the metric space of continuous functions defined on $\mathbb{R}^d$ with values in $\mathcal{D}_0$, where the distance d^{lu}_0 is given by

$$d^{lu}_0(f,g) = \sum_{N=1}^{\infty} \frac{1}{2^N} \frac{\sup_{|x|\le N} \|f(x) - g(x)\|_0}{1 + \sup_{|x|\le N} \|f(x) - g(x)\|_0}, \quad f, g \in C\left(\mathbb{R}^d; \mathcal{D}_0\right). \tag{3.20}$$

The space $\mathcal{C}_0$ is complete but not separable. Hence, instead of endowing it with the Borel $\sigma-$algebra associated with d^{lu}_0, we consider the $\sigma-$algebra $\mathcal{C}$ generated by the projections $\pi_x \colon C(\mathbb{R}^d; \mathcal{D}_0) \to (\mathcal{D}_0, \mathcal{D})$, $x \in \mathbb{R}^d$, defined by $\pi_x(f) = f(x)$, $f \in \mathcal{C}_0$. In Appendix 3.A (see Lemma 3.28), we prove that $\mathcal{C}$ can be read as a Borel $\sigma-$algebra of $C(\mathbb{R}^d; \mathcal{D}_0)$ endowed with the metric d^{lu}_S. Although we are not going to use Lemma 3.28 in this chapter, it is worth presenting because it is an analogue of the fact that $\mathcal{D}$ coincides with the Borel $\sigma-$algebra generated by J_1 in $\mathcal{D}([0,T]; \mathbb{R}^d)$, see the discussion around (3.11) and references therein.
Arguments similar to those in Subsection 3.2.1 about the space $(\mathcal{D}_0, \mathcal{D})$ show that $d^{lu}_0 \colon (\mathcal{C}_0 \times \mathcal{C}_0, \mathcal{C} \otimes \mathcal{C}) \to \mathbb{R}$ is measurable. Indeed, for every $N \in \mathbb{N}$, the map $(f,g) \mapsto \sup_{|x|\le N} \|f(x) - g(x)\|_0$ from $(\mathcal{C}_0 \times \mathcal{C}_0, \mathcal{C} \otimes \mathcal{C})$ to $\mathbb{R}$ is measurable (by continuity, the sup can be computed on a countable dense subset of $\mathbb{R}^d$).

We consider the process $Y = (Y_s)_{s \in [0,T]}$, where

$$Y_s = Y^{s,\cdot},\ 0 \le s < T, \quad \text{and} \quad Y_T(\omega) : \mathbb{R}^d \to \mathcal{D}_0,\ \ [Y_T(\omega)(x)](t) = Y^{T,x}_t(\omega) = x,\ \omega \in \Omega, \tag{3.21}$$

$0 \le t \le T$, $x \in \mathbb{R}^d$ (see (3.17)). Since $(Y^{s,x})^{-1}(A) \in \mathcal{F}$ for every $A \in \mathcal{D}$ and $x \in \mathbb{R}^d$, the map $Y_s \colon \Omega \to (\mathcal{C}_0, \mathcal{C})$ is a random variable for all $s \in [0,T]$. This fact coupled with the above discussion shows that

$$\omega \mapsto d^{lu}_0\left(Y_s(\omega), Y_t(\omega)\right) \text{ is measurable for every } s, t \in [0,T]. \tag{3.22}$$

As in [149], we want to apply [28, Theorem 4.2] to show the càdlàg property of the $(\mathcal{C}_0, \mathcal{C})-$valued process Y. The aforementioned theorem requires the stochastic continuity of Y, which is then the aim of this subsection.

Before presenting our result (see Lemma 3.10), we need some preparation. An important tool that we are going to use is [107, Theorem 1.1] (see also [149, Theorem 3.1]), which in turn is based on a generalized *Garsia–Rodemich–Rumsey* type lemma (see [13]). For the reader's convenience we report its statement, where we denote by $\log^+$ the positive part of the logarithm, namely $\log^+ x = \log x \vee 0$, $x \in \mathbb{R}_+$.

Theorem 3.7 ([107]). *Consider a separable metric space (M,ρ) and an $\mathcal{F}\otimes\mathcal{B}\left(\mathbb{R}^d\right)/\mathcal{B}(M)-$measurable map $\phi\colon \Omega\times\mathbb{R}^d\to M$ such that $\phi(\omega,\cdot)$ is continuous for every $\omega\in\Omega$. Suppose that there are $p>2d$ and $c>0$ such that*

$$\mathbb{E}\left[\left(\rho\left(\phi\left(\cdot,x\right),\phi\left(\cdot,y\right)\right)\right)^p\right]\le c\left|x-y\right|^p,\quad x,y\in\mathbb{R}^d.$$

For any $\alpha>1$, define the map $f_\alpha(x)=([x^d\left(\log^+ x\right)^\alpha]\vee 1)^{-1}$, $x\in\mathbb{R}_+$. Then the function $Z\colon\Omega\to[0,\infty]$ given by

$$Z\left(\omega\right)=\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{\rho\left(\phi\left(\omega,x\right),\phi\left(\omega,y\right)\right)}{|x-y|}\right)^p f_\alpha\left(|x|\right)f_\alpha\left(|y|\right)dx\,dy\right)^{\frac{1}{p}},\quad \omega\in\Omega, \tag{3.23}$$

is a $p-$integrable random variable satisfying

$$\begin{aligned}\rho\left(\phi\left(\omega,x\right),\phi\left(\omega,y\right)\right)\le c_0 Z\left(\omega\right)|x-y|^{1-\frac{2d}{p}}\left(\left[\left(|x|\vee|y|\right)^{\frac{2d}{p}}\left(\log^+\left(|x|\vee|y|\right)\right)^{\frac{2\alpha}{p}}\right]\vee 1\right),\\ x,y\in\mathbb{R}^d,\omega\in\Omega,\end{aligned} \tag{3.24}$$

where c_0 is a positive constant depending on α,p,d.

We wish to apply the previous result to $(\mathcal{D}_0,\mathcal{D})-$valued, continuous random fields. Although $\mathcal{D}_0$ is not separable and $\mathcal{D}$ is not the Borel $\sigma-$algebra generated by $\|\cdot\|_0$, this can be done thanks to the following proposition.

Proposition 3.8. *Theorem 3.7 holds substituting $(\mathcal{D}_0,\|\cdot\|_0,\mathcal{D})$ for $(M,\rho,\mathcal{B}(M))$.*

Proof. We note that the map $\omega\mapsto\|\phi(\omega,x)-\phi(\omega,y)\|_0$ is measurable for every x, $y\in\mathbb{R}^d$ (see Subsection 3.2.1, where this fact is proved for $Y^{s,\cdot}$). As a consequence, since $\phi\colon\Omega\times\mathbb{R}^d\to\mathcal{D}_0$ is $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^d)/\mathcal{D}-$measurable and $\phi(\omega,\cdot)$ is continuous for each $\omega\in\Omega$ by hypothesis, the function $K\colon\Omega\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ defined by

$$K\left(\omega,x,y\right)=\left\|\phi\left(\omega,x\right)-\phi\left(\omega,y\right)\right\|_0,\quad x,y\in\mathbb{R}^d,\,\omega\in\Omega$$

is jointly measurable. Looking now at the proof of [107, Theorem 1.1] and the results cited therein, it turns out that the separability of the arrival space (M,ρ) is only used to ensure the measurability of the function Z in (3.23). When $M=(\mathcal{D}_0,\mathcal{D})$, this property can be inferred directly. Indeed, for any $p>0$ and $\alpha>1$,

$$Z(\omega)=\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{K\left(\omega,x,y\right)}{|x-y|}\right)^p f_\alpha\left(|x|\right)f_\alpha\left(|y|\right)dx\,dy\right)^{1/p},\quad\omega\in\Omega;$$

since K is non–negative, as well as f_α and $|x-y|^{-1}$, $x\neq y$, the desired measurability for Z is given by Tonelli's theorem. ■

Fix $\alpha=2$ and denote by $f=f_\alpha$. Notice that, for any $\gamma>0$, $(\log(x))^\gamma=\mathrm{O}(x)$ as $x\to\infty$. Hence, combining Proposition 3.8 with (3.7) in Lemma 3.3 we obtain the next corollary.

Corollary 3.9. *For every $p>2d$ and $s\in[0,T)$, the $p-$integrable random variable $U_{s,p}$ defined by*

$$U_{s,p}(\omega)=\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{\|Y^{s,x}(\omega)-Y^{s,y}(\omega)\|_0}{|x-y|}\right)^p f\left(|x|\right)f\left(|y|\right)dx\,dy\right)^{\frac{1}{p}},\quad\omega\in\Omega,$$

is such that

$$\sup_{0\le t\le T}\left|Y_t^{s,x}(\omega)-Y_t^{s,y}(\omega)\right|\le c_1\,U_{s,p}(\omega)\left|x-y\right|^{1-2d/p}\left[\left(|x|\vee|y|\right)^{\frac{2d+1}{p}}\vee 1\right],\quad x,y\in\mathbb{R}^d,\ \omega\in\Omega,\quad (3.25)$$

where $c_1=c_1(d,p)>0$. *Furthermore,*

$$\sup_{s\in[0,T)}\mathbb{E}\left[U_{s,p}^p\right]\le 4^{p-1}e^{cT}c_2,\quad (3.26)$$

where $c>0$ *is the same constant as in* (3.7) *and* $c_2=\left(\int_{\mathbb{R}^d}f(|x|)\,dx\right)^2<\infty$.

We remark that results similar to Corollary 3.9 hold with the continuous random fields $Z_i^{s,\cdot}$, $i=1,2,3$, instead of $Y^{s,\cdot}$. We are now ready to prove the main result of this subsection.

Lemma 3.10. *The* $(\mathcal{C}_0,\mathcal{C})$*−valued process* $Y=(Y_s)_{s\in[0,T]}$ *considered in* (3.21) *is continuous in probability.*

Proof. Fix $s\in[0,T]$ and take a sequence $(s_n)_n\subset[0,T]$ such that $s_n\to s$ as $n\to\infty$. We want to show that

$$\mathbb{E}\Big[\sup_{|x|\le N}\left\|Y^{s,x}-Y^{s_n,x}\right\|_0\Big]=\mathbb{E}\Big[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\Big]\underset{n\to\infty}{\longrightarrow}0,\quad N\ge 1.\quad (3.27)$$

Indeed, this is a sufficient condition to obtain the stochastic continuity of Y in s, as the next argument explains. By definition of continuity in probability, we aim to prove that $\lim_{n\to\infty}\mathbb{P}\left(d_0^{lu}(Y_{s_n},Y_s)>\epsilon\right)=0$ for any $\epsilon>0$, which is equivalent to

$$\lim_{n\to\infty}\mathbb{E}\left[d_0^{lu}(Y_{s_n},Y_s)\left(1+d_0^{lu}(Y_{s_n},Y_s)\right)^{-1}\right]=0.$$

Therefore, it is enough to show that

$$\lim_{n\to\infty}\mathbb{E}\left[d_0^{lu}(Y_{s_n},Y_s)\right]=\lim_{n\to\infty}\sum_{N=1}^{\infty}\frac{1}{2^N}\mathbb{E}\left[\frac{\sup_{|x|\le N}\left\|Y^{s,x}-Y^{s_n,x}\right\|_0}{1+\sup_{|x|\le N}\left\|Y^{s,x}-Y^{s_n,x}\right\|_0}\right]=0.\quad (3.28)$$

The dominated convergence theorem –endowing $\mathbb{N}$ with the canonical counting measure– gives (3.28) knowing (3.27).

Fix $N\ge 1$. We start off by proving the right stochastic continuity in a point $s\in[0,T)$. Take a sequence $(s_n)_n\subset(s,T)$ such that $s_n\downarrow s$ and split the expectation in (3.27) as follows:

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right]\le\mathbb{E}\left[\sup_{|x|\le N}\sup_{0\le t\le s}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right]+\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right]$$
$$+\mathbb{E}\left[\sup_{|x|\le N}\sup_{s_n\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right],\quad n\in\mathbb{N}.\quad (3.29)$$

We analyze the second and third addends in the right–hand side of (3.29), the first being 0. As for the second, by (3.17), for every $n\in\mathbb{N}$, we can find an a.s. event Ω_{s,s_n} independent of x where

$$\sup_{s\le t\le s_n}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\le\sup_{s\le t\le s_n}\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|+\sup_{s\le t\le s_n}\left|\int_s^t\alpha\left(r,Y_r^{s,x}\right)dW_r\right|$$
$$+\sup_{s\le t\le s_n}\left|\int_s^t\int_{U_0}g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p(dr,dz)\right|,\quad x\in\mathbb{R}^d.\quad (3.30)$$

Thus, we prove that, as $n \to \infty$,

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|\right] \to 0, \qquad \mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t \alpha\left(r, Y_r^{s,x}\right)dW_r\right|\right] \to 0,$$
$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p\left(dr, dz\right)\right|\right] \to 0. \tag{3.31}$$

The limits in (3.31) are all dealt with using the same technique, so we just focus on the one appearing in the second line. In particular, we want to apply Proposition 3.8 to the continuous, $(\mathcal{D}_0, \mathcal{D})-$valued random field $(Z_2^{s,x})_{x\in\mathbb{R}^d}$. To do this, we take $\gamma > 2d$ and write

$$\mathbb{E}\Big[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Z_{2,t}^{s,x}\right|^\gamma\Big] \le 2^{\gamma-1}\Big(\mathbb{E}\Big[\sup_{s\le t\le s_n}\left|Z_{2,t}^{s,0}\right|^\gamma\Big] + \mathbb{E}\Big[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Z_{2,t}^{s,x} - Z_{2,t}^{s,0}\right|^\gamma\Big]\Big)$$
$$=: 2^{\gamma-1}\left(\mathbf{I}_n + \mathbf{II}_n\left(N\right)\right). \tag{3.32}$$

As for $\mathbf{I}_n$, by the linear growth condition (3.1) for g and [125, Theorem 2.11] we infer that, for some $C > 0$,

$$\begin{aligned}\mathbf{I}_n &= \mathbb{E}\left[\sup_{s\le t\le s_n}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,0}, r, z\right)\widetilde{N}_p\left(dr, dz\right)\right|^\gamma\right]\\ &\le c_{d,\gamma}\left(\mathbb{E}\left[\left(K_2\int_s^{s_n}\left(1+\left|Y_{r-}^{s,0}\right|\right)^2 dr\right)^{\frac{\gamma}{2}}\right] + \mathbb{E}\left[K_\gamma\int_s^{s_n}\left(1+\left|Y_{r-}^{s,0}\right|\right)^\gamma dr\right]\right)\\ &\le c_{d,\gamma}\left(\left(s_n-s\right)^{\frac{\gamma}{2}}K_2^{\frac{\gamma}{2}} + \left(s_n-s\right)K_\gamma\right)\mathbb{E}\left[\sup_{s\le t\le T}\left(1+\left|Y_t^{s,0}\right|\right)^\gamma\right] \le C\cdot \mathrm{o}\left(1\right), \quad \text{as } n\to\infty,\end{aligned} \tag{3.33}$$

where in the last inequality we use [125, Equation (3.6)]. As regards $\mathbf{II}_n$, by the estimates in (3.7) and (3.10) we have, for some $C_0 = C_0\left(\gamma, d, m, T, K_2, K_\gamma\right) > 0$,

$$\mathbb{E}\left[\sup_{s\le t\le s_n}\left|\int_s^t\int_{U_0}\left(g\left(Y_{r-}^{s,x}, r, z\right) - g\left(Y_{r-}^{s,y}, r, z\right)\right)\widetilde{N}_p\left(dr, dz\right)\right|^\gamma\right]$$
$$\le C_0\ \left(s_n - s\right)\left|x-y\right|^\gamma, \quad x, y\in\mathbb{R}^d. \tag{3.34}$$

Hence Proposition 3.8 yields, arguing as in Corollary 3.9 with $Y^{s,\cdot}$ replaced by $Z_2^{s,\cdot}$,

$$\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Z_{2,t}^{s,x} - Z_{2,t}^{s,0}\right| \le c\left(d,p\right)\widetilde{U}^{(2)}_{s,s_n,\gamma}N^{1+\frac{1}{\gamma}}, \quad \text{in } \Omega,$$

where

$$\widetilde{U}^{(2)}_{s,s_n,\gamma}(\omega) = \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{\sup_{s\le t\le s_n}\left|Z_{2,t}^{s,x}(\omega) - Z_{2,t}^{s,y}(\omega)\right|}{|x-y|}\right)^\gamma f\left(|x|\right)f\left(|y|\right)dx\,dy\right)^{\frac{1}{\gamma}}, \quad \omega\in\Omega.$$

In particular, $\widetilde{U}^{(2)}_{s,s_n,\gamma}$ is a $\gamma-$integrable random variable such that, by (3.34), $\mathbb{E}\big[\big(\widetilde{U}^{(2)}_{s,s_n,\gamma}\big)^\gamma\big] \le C_1\left(s_n - s\right)$, where $C_1 = C_1(\gamma, d, m, T, K_2, K_\gamma)$. Consequently,

$$\mathbf{II}_n\left(N\right) \le c^\gamma C_1 N^{\gamma+1}\left(s_n - s\right) \to 0 \quad \text{as } n\to\infty. \tag{3.35}$$

Combining (3.33)-(3.35) in (3.32) we deduce that

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{\gamma}\right]\to 0 \quad \text{as } n\to\infty. \tag{3.36}$$

Note that (3.36) holds for every $\gamma>0$ by Jensen's inequality, hence we recover the second line of (3.31) as a particular case. In an analogous way one can prove that, for any $\gamma>0$,

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|^{\gamma}\right]\underset{n\to\infty}{\longrightarrow}0,\quad \mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|\int_s^t \alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{\gamma}\right]\underset{n\to\infty}{\longrightarrow}0, \tag{3.37}$$

whence

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right]\to 0 \quad \text{as } n\to\infty. \tag{3.38}$$

Next, we study the third addend in (3.29) using the flow property in Corollary 3.5. In particular, by (3.14) there exists an almost sure event Ω_{s,s_n} independent of x where

$$\sup_{s_n\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|=\sup_{s_n\le t\le T}\left|Y_t^{s_n,Y_{s_n}^{s,x}}-Y_t^{s_n,x}\right|\le\sup_{0\le t\le T}\left|Y_t^{s_n,Y_{s_n}^{s,x}}-Y_t^{s_n,x}\right|,\quad x\in\mathbb{R}^d.$$

Now we choose $p\ge 2d+1$ and apply (3.25) in Corollary 3.9 to deduce that

$$\sup_{0\le t\le T}\left|Y_t^{s_n,Y_{s_n}^{s,x}}-Y_t^{s_n,x}\right|\le c_1\,U_{s_n,p}\left[\left(|x|\vee\left|Y_{s_n}^{s,x}\right|\right)^{\frac{2d+1}{p}}\vee 1\right]\left|x-Y_{s_n}^{s,x}\right|^{1-\frac{2d}{p}},\quad x\in\mathbb{R}^d,$$

which holds in the whole space Ω. Moreover, by (3.17), there exists a full probability set Ω_s where

$$\left|Y_{s_n}^{s,x}\right|\le N+\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|,\quad |x|\le N.$$

Combining the three previous expressions we get the existence of an almost sure event Ω'_{s,s_n} where

$$\begin{aligned}\sup_{|x|\le N}\sup_{s_n\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\le c_1\,U_{s_n,p}\sup_{|x|\le N}\Bigg\{\Bigg[N^{\frac{2d+1}{p}}\\
+\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}}+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{\frac{2d+1}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{\frac{2d+1}{p}}\Bigg]\\
\times\Bigg[\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{1-\frac{2d}{p}}+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{1-\frac{2d}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{1-\frac{2d}{p}}\Bigg]\Bigg\}.\end{aligned} \tag{3.39}$$

In the estimates in (3.39), we use the subadditivity property of the function x^r, for $r=(2d+1)/p$ or $r=1-2d/p$. Let us denote by $C_2 \coloneqq \sup_{0<t<T}\mathbb{E}\left[U_{t,p}^2\right]$, which is finite by (3.26) in Corollary 3.9. Taking the expected value in (3.39), we apply the Cauchy–Schwarz inequality to write

$$\begin{aligned}\mathbb{E}\left[\sup_{|x|\le N}\sup_{s_n\le t\le T}\left|Y_t^{s,x}-Y_t^{s_n,x}\right|\right]^2\le 12\,c_1\,C_2\,\mathbb{E}\Bigg[\sup_{|x|\le N}\Bigg\{\\
N^{\frac{4d+2}{p}}+\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{\frac{4d+2}{p}}+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{\frac{4d+2}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{\frac{4d+2}{p}}\Bigg\}\\
\times\sup_{|x|\le N}\Bigg\{\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{2-\frac{4d}{p}}+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{2-\frac{4d}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{2-\frac{4d}{p}}\Bigg\}\Bigg].\end{aligned}$$

Invoking the Cauchy–Schwarz inequality one more time,

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s_n\le t\le T}|Y_t^{s,x}-Y_t^{s_n,x}|\right]^2 \le 12\sqrt{12}\,c_1\,C_2\Bigg(N^{\frac{4d+2}{p}}+\Bigg(\mathbb{E}\Bigg[\sup_{|x|\le N}\Bigg\{\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{\frac{8d+4}{p}}$$

$$+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{\frac{8d+4}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{\frac{8d+4}{p}}\Bigg\}\Bigg]\Bigg)^{\frac12}\Bigg)\Bigg(\mathbb{E}\Bigg[$$

$$\sup_{|x|\le N}\Bigg\{\left|\int_s^{s_n} b\left(r,Y_r^{s,x}\right)dr\right|^{4-\frac{8d}{p}}+\left|\int_s^{s_n}\alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{4-\frac{8d}{p}}+\left|\int_s^{s_n}\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{4-\frac{8d}{p}}\Bigg\}\Bigg]\Bigg)^{\frac12}.$$

Notice that the right–hand side of the previous equation goes to 0 as $n\to\infty$ by (3.36)-(3.37). Summing up,

$$\mathbb{E}\left[\sup_{|x|\le N}\sup_{s_n\le t\le T}|Y_t^{s,x}-Y_t^{s_n,x}|\right]\to 0 \qquad \text{as } n\to\infty. \tag{3.40}$$

Combining (3.40) with (3.38) we obtain (3.27), which yields the desired right–stochastic continuity for the process Y in s. The left–continuity of Y in $(0,T]$ can be argued in a similar way. The proof is complete. ■

Recall the continuous, $(\mathcal{D}_0,\mathcal{D})-$valued random fields $Z_i^{s,\cdot}$, $i=1,2,3$, introduced at the end of Subsection 3.2.1 and appearing on the right–hand side of (3.17). For every $i=1,2,3$, we define the $(\mathcal{C}_0,\mathcal{C})-$valued stochastic process $Z_i=(Z_{i,s})_{s\in[0,T]}$, whose random variables are $Z_{i,s}=Z_i^{s,\cdot}$, $0\le s<T$, and we set, for every $\omega\in\Omega$,

$$Z_{i,T}\left(\omega\right):\mathbb{R}^d\to\mathcal{D}_0,\qquad \left[Z_{i,T}\left(\omega\right)\left(x\right)\right]\left(t\right)=Z_{i,t}^{T,x}\left(\omega\right)=0,\quad x\in\mathbb{R}^d,\ 0\le t\le T.$$

Using a strategy similar to the one followed in Lemma 3.10 for the process Y, in the next result we show the continuity in probability of the processes Z_i.

Corollary 3.11. *For every $i=1,2,3$, the $(\mathcal{C}_0,\mathcal{C})-$valued process $Z_i=\left(Z_{i,s}\right)_{s\in[0,T]}$ is continuous in probability.*

Proof. Fix $i=1,2,3$ and $s\in[0,T)$. As in Lemma 3.10, we only prove the right stochastic continuity of Z_i in s. Hence we consider a sequence $(s_n)_n\subset(s,T)$ such that $s_n\downarrow s$ and we show that (cfr. (3.27))

$$\mathbb{E}\left[\sup_{|x|\le N}\|Z_i^{s,x}-Z_i^{s_n,x}\|_0\right]=\mathbb{E}\left[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|\right]\underset{n\to\infty}{\longrightarrow}0,\quad N\ge 1. \tag{3.41}$$

Taking $N\ge1$, we split the expectation in (3.41) as follows:

$$\begin{aligned}\mathbb{E}\left[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|\right]&\le\mathbb{E}\left[\sup_{|x|\le N}\sup_{0\le t\le s}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|\right]+\mathbb{E}\left[\sup_{|x|\le N}\sup_{s\le t\le s_n}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|\right]\\&\quad+\mathbb{E}\left[\sup_{|x|\le N}\sup_{s_n\le t\le T}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|\right]\\&=:\left(\mathbf{I}_n^{(i)}+\mathbf{II}_n^{(i)}+\mathbf{III}_n^{(i)}\right)(N),\quad n\in\mathbb{N}.\end{aligned}$$

Since $\mathbf{I}_n^{(i)}(N)=0$, $n\in\mathbb{N}$, by construction, and $\lim_{n\to\infty}\mathbf{II}_n^{(i)}(N)=0$ by (3.31), we only study $\mathbf{III}_n^{(i)}(N)$. In particular, we prove that $\lim_{n\to\infty}\mathbf{III}_n^{(i)}(N)=0$. To do this, we invoke Corollary 3.6, precisely (3.19), which guarantees the existence of an almost sure event Ω_{s,s_n} –independent from x– where, for every $x\in\mathbb{R}^d$,

$$\sup_{s_n\le t\le T}\left|Z_{i,t}^{s,x}-Z_{i,t}^{s_n,x}\right|=\sup_{s_n\le t\le T}\left|Z_{i,s_n}^{s,x}+Z_{i,t}^{s_n,Y_{s_n}^{s,x}}-Z_{i,t}^{s_n,x}\right|$$
$$\le\sup_{s\le t\le s_n}\left|Z_{i,t}^{s,x}\right|+\sup_{0\le t\le T}\left|Z_{i,t}^{s_n,Y_{s_n}^{s,x}}-Z_{i,t}^{s_n,x}\right|. \tag{3.42}$$

Choose $p\ge 2d+1$. By Lemma 3.3 and the estimates in its proof, we can apply Proposition 3.8 to deduce the existence of a $p-$integrable random variable $\widetilde{U}_{s_n,p}^{(i)}$ such that, for some $c=c(p,d)>0$,

$$\sup_{0\le t\le T}\left|Z_{i,t}^{s_n,Y_{s_n}^{s,x}}-Z_{i,t}^{s_n,x}\right|\le c\,\widetilde{U}_{s_n,p}^{(i)}\left[\left(|x|\vee\left|Y_{s_n}^{s,x}\right|\right)^{\frac{2d+1}{p}}\vee 1\right]\left|x-Y_{s_n}^{s,x}\right|^{1-\frac{2d}{p}},\quad x\in\mathbb{R}^d,$$

which holds in the whole space Ω. Since (again by Lemma 3.3) $\sup_{n\in\mathbb{N}}\mathbb{E}\big[\big(\widetilde{U}_{s_n,p}^{(i)}\big)^2\big]<\infty$, we can proceed as in (3.39) and the subsequent estimates to obtain

$$\lim_{n\to\infty}\mathbb{E}\Big[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Z_{i,t}^{s_n,Y_{s_n}^{s,x}}-Z_{i,t}^{s_n,x}\right|\Big]=0.$$

Taking now the supremum over $|x|\le N$ and expectations in (3.42), we have

$$\mathbf{III}_n^{(i)}(N)\le\mathbf{II}_n^{(i)}(N)+\mathbb{E}\Big[\sup_{|x|\le N}\sup_{0\le t\le T}\left|Z_{i,t}^{s_n,Y_{s_n}^{s,x}}-Z_{i,t}^{s_n,x}\right|\Big]\to 0\quad\text{as } n\to\infty,$$

completing the proof. ■

3.2.3 The càdlàg property in the initial time s

In this subsection we study the existence of a càdlàg modification of the process $Y=(Y_s)_s$ in (3.21). Such a property is obtained by Theorem 3.12, which is a reformulation of [28, Theorem 4.2] (see also [29, Theorem 1.2]). We compare Theorem 3.12 with the original result in [28]-[29] in Remark 3.2.

Theorem 3.12. *Let $X=(X_t)_{t\in[0,T]}$ be a family of functions defined on a complete probability space $(\Xi,\mathcal{G},\mathbb{Q})$ with values in a complete metric space (E,Δ). Let $0\le s\le t\le u\le T$ and denote by $\Delta(s,t,u)=\Delta(X_s,X_t)\wedge\Delta(X_t,X_u)$. Suppose that the map $\omega\mapsto\Delta(X_s(\omega),X_t(\omega))$ is measurable for any $s,t\in[0,T]$, and that X is continuous in probability. If there exist continuous, increasing functions $\delta\colon[0,T]\to\mathbb{R}_+$ and $\theta\colon[0,T\vee 1]\to\mathbb{R}_+$, with $\delta(0)=\theta(0)=0$ and θ concave, such that*

$$\mathbb{E}_{\mathbb{Q}}\left[\Delta(s,t,u)1_A\right]\le\delta\left(u-s\right)\cdot\theta\left(\mathbb{Q}\left(A\right)\right),\quad 0\le s\le t\le u\le T,\ A\in\mathcal{G}, \tag{3.43}$$

and that

$$\int_0^T u^{-2}\delta\left(u\right)\theta\left(u\right)du<\infty, \tag{3.44}$$

then X has a càdlàg modification X' (modification means that the map $\omega\mapsto\Delta(X_t(\omega),X'_t(\omega))$ is equal to 0, $\mathbb{Q}-$a.s., for any $t\in[0,T]$).

Remark 3.2. *Compared to the original assertion in [28]-[29], in Theorem 3.12 we do not require the functions $X_t \colon \Xi \to (E, \Delta)$ to be measurable with respect to the Borel $\sigma-$algebra of E. This is crucial for our arguments, because we are going to apply Theorem 3.12 to the $(\mathcal{C}_0, \mathcal{C})-$valued process Y and $\mathcal{C}$ is not the Borel $\sigma-$algebra of $\mathcal{C}_0$. The hypothesis on the measurability of the map $\omega \mapsto \Delta(X_s(\omega), X_t(\omega))$ is inspired by an analogous assumption in [30, Lemma A.2.37], where the Kolmogorov–Chentsov continuity criterion is proved without supposing the separability of the arrival space. In any case, such an hypothesis does not alter the strategy of the proof of Theorem 3.12, which is presented in Appendix 3.B for the sake of completeness.*

The next corollary, which gives a sufficient condition for the existence of a càdlàg version, can be easily deduced from Theorem 3.12.

Corollary 3.13. *Under the same hypotheses as in Theorem 3.12, if there exist $q > 1/2$, $r > 0$ and $C > 0$ such that*

$$\mathbb{E}_{\mathbb{Q}} \left[\Delta \left(X_s, X_t \right)^q \cdot \Delta \left(X_t, X_u \right)^q \right] \leq C \left| u - s \right|^{1+r}, \quad 0 \leq s \leq t \leq u \leq T, \tag{3.45}$$

then X has a càdlàg modification.

Proof. Define $\theta(h) = h^{1-\frac{1}{2q}}$ and $\delta(h) = C^{\frac{1}{2q}} h^{\frac{1+r}{2q}}$, $h \geq 0$. Then (3.43) can be deduced from (3.45) as in [149, Corollary 4.2], while (3.44) is satisfied because $r > 0$. ∎

We are now ready to present the main result of Section 3.2. In the proof, we employ the concept of strong solution to Equation (3.5) (see (3.55)), which allows to follow an argument relying on conditional expectations with respect to the augmented $\sigma-$algebra generated by W and N_p.

Theorem 3.14. *There exists a càdlàg version Z of the $(\mathcal{C}_0, \mathcal{C})-$valued process $Y = (Y_s)_s$.*

Proof. We will apply Corollary 3.13. Note that, by the completeness of the probability space, the càdlàg version Z will be automatically a $(\mathcal{C}_0, \mathcal{C})-$valued process.

Recall that in (3.22) we have shown the measurability of the map $\omega \mapsto d_0^{lu}(Y_s(\omega), Y_t(\omega))$, $s, t \in [0, T]$. Thus, according to Corollary 3.13, in order to find a càdlàg modification of the stochastically continuous process Y it is sufficient to determine $q > 1/2, r > 0$ and a constant $C > 0$ such that

$$\mathbb{E} \left[d_0^{lu} \left(Y_s, Y_u \right)^q \cdot d_0^{lu} \left(Y_u, Y_v \right)^q \right] \leq C \left| v - s \right|^{1+r}, \quad 0 \leq s < u < v \leq T. \tag{3.46}$$

Note that using the completeness of the probability space, it is not difficult to prove that

Let us take a triplet of times (s, u, v), with $0 \leq s < u < v \leq T$, and denote by $\rho = v - s$. We can assume $\rho < 1$, otherwise (3.46) is trivially satisfied for any choice of $q, r \in \mathbb{R}_+$ and $C \geq 1$. By the computations in Subsection 3.2.2, precisely (3.30)-(3.39), there exists an a.s. event $\Omega_{s,u}$ where, for every

$p \geq 2d+1$ and $N \geq 1$,

$$
\begin{aligned}
&\sup_{|x|\leq N} \sup_{0\leq t\leq T} |Y_t^{s,x} - Y_t^{u,x}| \\
&\leq \sup_{|x|\leq N} \left\{ \sup_{s\leq t\leq u} \left| \int_s^t b\left(r, Y_r^{s,x}\right) dr \right| + \sup_{s\leq t\leq u} \left| \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r \right| + \sup_{s\leq t\leq u} \left| \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right| \right\} \\
&\quad + c_1\, U_{u,p} \sup_{|x|\leq N} \Bigg\{ \\
&\quad \left[N^{\frac{2d+1}{p}} + \left| \int_s^u b\left(r, Y_r^{s,x}\right) dr \right|^{\frac{2d+1}{p}} + \left| \int_s^u \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^{\frac{2d+1}{p}} + \left| \int_s^u \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^{\frac{2d+1}{p}} \right] \\
&\quad \times \left[\left| \int_s^u b\left(r, Y_r^{s,x}\right) dr \right|^{1-\frac{2d}{p}} + \left| \int_s^u \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^{1-\frac{2d}{p}} + \left| \int_s^u \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^{1-\frac{2d}{p}} \right] \Bigg\} \\
&=: \mathbf{I}_1 + c_1 U_{u,p} \mathbf{II}_1. \qquad (3.47)
\end{aligned}
$$

If we compute the product appearing in $\mathbf{II}_1$, then introducing the set $A = \left\{ \frac{2d+1}{p}, 1 - \frac{2d}{p} \right\}$ we can estimate

$$
\begin{aligned}
\mathbf{II}_1 \leq \sup_{|x|\leq N} \Bigg\{ & N^{\frac{2d+1}{p}} \sup_{s\leq t\leq u} \left| \int_s^t b\left(r, Y_r^{s,x}\right) dr \right|^{1-\frac{2d}{p}} + N^{\frac{2d+1}{p}} \sup_{s\leq t\leq u} \left| \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^{1-\frac{2d}{p}} \\
& + N^{\frac{2d+1}{p}} \sup_{s\leq t\leq u} \left| \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^{1-\frac{2d}{p}} + \sup_{s\leq t\leq u} \left| \int_s^t b\left(r, Y_r^{s,x}\right) dr \right|^{1+\frac{1}{p}} \\
& + \sup_{s\leq t\leq u} \left| \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^{1+\frac{1}{p}} + \sup_{s\leq t\leq u} \left| \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^{1+\frac{1}{p}} \Bigg\} + S\left(N\right) =: \mathbf{III}_1,
\end{aligned}
$$

where we denote by $\mathcal{S}\left(N\right)$, $N \geq 1$, the quantity

$$
\begin{aligned}
\mathcal{S}\left(N\right) = \sum_{i,j\in A,\, i\neq j} \sup_{|x|\leq N} \sup_{s\leq t\leq u} \Bigg\{ & \\
& \left| \int_s^t b\left(r, Y_r^{s,x}\right) dr \right|^i \left| \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^j + \left| \int_s^t b\left(r, Y_r^{s,x}\right) dr \right|^i \left| \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^j \\
& + \left| \int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r \right|^i \left| \int_s^t \int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right|^j \Bigg\}. \qquad (3.48)
\end{aligned}
$$

Furthermore, notice that $\mathbf{I}_1 \leq \mathbf{III}_1$. Therefore, going back to (3.47) we have

$$\sup_{|x|\leq N}\sup_{0\leq t\leq T}\left|Y_t^{s,x}-Y_t^{u,x}\right| \leq \left(1+c_1\,U_{u,p}\right)\mathbf{III}_1 \leq \left(1+c_1\,U_{u,p}\right)$$
$$\times\Bigg[N^{\frac{2d+1}{p}}\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|^{1-\frac{2d}{p}} + N^{\frac{2d+1}{p}}\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t \alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{1-\frac{2d}{p}}$$
$$+ N^{\frac{2d+1}{p}}\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{1-\frac{2d}{p}}$$
$$+\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|^{1+\frac{1}{p}} + \sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t \alpha\left(r,Y_r^{s,x}\right)dW_r\right|^{1+\frac{1}{p}}$$
$$+\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|^{1+\frac{1}{p}} + \mathcal{S}\left(N\right)\Bigg]. \tag{3.49}$$

Since we can pick $p \geq 2d+1$ arbitrarily large, we consider $p^{-1}(2d+1) < 1-2dp^{-1}$. Moreover, note that given two numbers $a, b \geq 0$,

$$(ab)\wedge 1 \leq (1+a)\,(b\wedge 1) \quad \text{and} \quad (a+b)\wedge 1 \leq a\wedge 1 + b\wedge 1. \tag{3.50}$$

With these considerations, using the inequality $x/(1+x) \leq 1\wedge x$, $x \geq 0$, we simplify the expression in (3.49) to deduce that

$$\frac{\sup_{|x|\leq N}\left\|Y^{s,x}-Y^{u,x}\right\|_0}{1+\sup_{|x|\leq N}\left\|Y^{s,x}-Y^{u,x}\right\|_0} \leq \sup_{|x|\leq N}\sup_{0\leq t\leq T}\left|Y_t^{s,x}-Y_t^{u,x}\right|\wedge 1$$
$$\leq \left(2+c_1\,U_{u,p}\right)\left(N^{\frac{2d+1}{p}}+2\right)\Bigg[\mathcal{S}\left(N\right)\wedge 1 + \sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|\wedge 1\right)^{\frac{2d+1}{p}}$$
$$+\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|\wedge 1\right)^{\frac{2d+1}{p}} + \sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t \alpha\left(r,Y_r^{s,x}\right)dW_r\right|\wedge 1\right)^{\frac{2d+1}{p}}\Bigg].$$

Multiplying by 2^{-N} and then summing over N, by the definition in (3.20) we infer that

$$d_0^{lu}\left(Y_s,Y_u\right)$$
$$\leq 2\left(2+c_1\,U_{u,p}\right)\sum_{N=1}^{\infty}\frac{N+1}{2^N}\Bigg[\mathcal{S}\left(N\right)\wedge 1 + \sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right)\right|\wedge 1\right)^{\frac{2d+1}{p}}$$
$$+\sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t b\left(r,Y_r^{s,x}\right)dr\right|\wedge 1\right)^{\frac{2d+1}{p}} + \sup_{|x|\leq N}\sup_{s\leq t\leq u}\left(\left|\int_s^t \alpha\left(r,Y_r^{s,x}\right)dW_r\right|\wedge 1\right)^{\frac{2d+1}{p}}\Bigg], \tag{3.51}$$

which holds in $\Omega_{s,u}$. Now we want to split the series in (3.51). To do this, we first notice that the function $2^{-x}\,(x+1)$ is strictly decreasing in $[1,\infty)$, hence we can estimate

$$\sum_{N=\overline{N}+1}^{\infty}\frac{N+1}{2^N} \leq \int_{\overline{N}}^{\infty}\frac{x+1}{2^x}\,dx = \frac{1}{(\log 2)^2}\,\frac{\left(\overline{N}+1\right)\log 2+1}{2^{\overline{N}}}, \quad \overline{N} \geq 1.$$

Therefore, for any $\gamma > 0$ –a new leverage parameter *which has to be fixed*– there exists a constant $c_\gamma > 0$ such that $\sum_{N=\overline{N}+1}^{\infty}(N+1)2^{-N} \le c_\gamma(\overline{N}+1)^{-\gamma}$, $\overline{N} \ge 1$.
Secondly, we introduce another parameter $\sigma > 0$ –once again, *to be determined*– and set $\overline{N} = \overline{N}(\rho) = [\rho^{-\sigma}]$ (recall that $\rho = v - s < 1$). Note that $\sum_{N=[\rho^{-\sigma}]+1}^{\infty}(N+1)2^{-N} \le c_\gamma([\rho^{-\sigma}]+1)^{-\gamma} \le c_\gamma\rho^{\sigma\gamma}$, and that $\sum_{N=1}^{\infty}(N+1)2^{-N} = 3$. Hence from (3.51) we write

$$d_0^{lu}(Y_s, Y_u) \le 2\,(2 + c_1\,U_{u,p})\Bigg[3\Bigg(\mathcal{S}\left(\rho^{-\sigma}\right) \wedge 1 + \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{\frac{2d+1}{p}}$$
$$+ \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r, Y_r^{s,x}\right)dW_r\right|^{\frac{2d+1}{p}}\Bigg) + 4\,c_\gamma\,\rho^{\sigma\gamma}\Bigg]. \quad (3.52)$$

At this point, we revert to the definition of $\mathcal{S}(\cdot)$ in (3.48). By (3.50),

$$\left(\left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{1-\frac{2d}{p}}\right) \wedge 1$$
$$\le \left(1 + \left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}}\right)\left(\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right| \wedge 1\right)^{1-\frac{2d}{p}}$$
$$\le \left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{\frac{2d+1}{p}} + \left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}}, \quad x \in \mathbb{R}^d,\ t \in [s,u].$$

Repeating the previous argument we find an upper bound for $\mathcal{S}(\rho^{-\sigma}) \wedge 1$, namely

$$\mathcal{S}\left(\rho^{-\sigma}\right) \wedge 1 \le 4\Bigg(\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{\frac{2d+1}{p}}$$
$$+ \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r, Y_r^{s,x}\right)dW_r\right|^{\frac{2d+1}{p}}\Bigg).$$

Therefore (3.52) becomes

$$d_0^{lu}(Y_s, Y_u) \le 2\,(2 + c_1\,U_{u,p})\Bigg[15\Bigg(\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{\frac{2d+1}{p}}$$
$$+ \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r, Y_r^{s,x}\right)dr\right|^{\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r, Y_r^{s,x}\right)dW_r\right|^{\frac{2d+1}{p}}\Bigg) + 4\,c_\gamma\,\rho^{\sigma\gamma}\Bigg].$$

Analogously, in an a.s. event $\Omega_{u,v}$ one has, using also that $d_0^{lu}(f,g) \le 1$ for every $f, g \in \mathcal{C}_0$,

$$d_0^{lu}(Y_u, Y_v) \le 30\,(2 + c_1\,U_{v,p})\Bigg[\Bigg(\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t\int_{U_0} g\left(Y_{r-}^{u,x}, r, z\right)\widetilde{N}_p(dr,dz)\right|^{\frac{2d+1}{p}}$$
$$+ \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t b\left(r, Y_r^{u,x}\right)dr\right|^{\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t \alpha\left(r, Y_r^{u,x}\right)dW_r\right|^{\frac{2d+1}{p}} + c_\gamma\,\rho^{\sigma\gamma}\Bigg) \wedge 1\Bigg].$$

If we multiply the two previous expressions, raise both sides to a power $q \in (1/2, p)$ and take expectations, then

$$\mathbb{E}\left[d_0^{lu}(Y_s,Y_u)^q \cdot d_0^{lu}(Y_u,Y_v)^q\right] \le 30^{2q}\left(4^{q-1}\vee 1\right)^2 \mathbb{E}\Bigg[(2+c_1\,U_{u,p})^q\,(2+c_1\,U_{v,p})^q$$
$$\times\Bigg\{\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\int_{U_0} g\left(Y^{s,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{q\frac{2d+1}{p}}$$
$$+\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r,Y^{s,x}_r\right)dr\right|^{q\frac{2d+1}{p}}+\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r,Y^{s,x}_r\right)dW_r\right|^{q\frac{2d+1}{p}}+c_\gamma^q\,\rho^{\sigma\gamma q}\Bigg\}$$
$$\times\Bigg(\Bigg\{\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t\int_{U_0} g\left(Y^{u,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{q\frac{2d+1}{p}}$$
$$+\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t b\left(r,Y^{u,x}_r\right)dr\right|^{q\frac{2d+1}{p}}+\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t \alpha\left(r,Y^{u,x}_r\right)dW_r\right|^{q\frac{2d+1}{p}}+c_\gamma^q\,\rho^{\sigma\gamma q}\Bigg\}\wedge 1\Bigg)\Bigg].$$

Notice that the definition in Corollary 3.9 and Lemma 3.3 yield the $\eta-$integrability of the random variables $U_{u,p}$ and $U_{v,p}$ for every $\eta > 1$. We fix $\eta = \frac{5}{4}$ and apply Hölder's inequality with exponents η and $\eta' = \eta(\eta-1)^{-1} = 5$ to deduce that

$$\mathbb{E}\left[d_0^{lu}(Y_s,Y_u)^q \cdot d_0^{lu}(Y_u,Y_v)^q\right] \le 30^{2q}4^{2\frac{\eta-1}{\eta}}\left(4^{q-1}\vee 1\right)^2\left(\mathbb{E}\left[(2+c_1\,U_{u,p})^{5q}\,(2+c_1\,U_{v,p})^{5q}\right]\right)^{1/5}$$
$$\times\Bigg(\mathbb{E}\Bigg[\Bigg\{c_\gamma^{\eta q}\,\rho^{\eta\sigma\gamma q}+\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\int_{U_0} g\left(Y^{s,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{\eta q\frac{2d+1}{p}}$$
$$+\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r,Y^{s,x}_r\right)dr\right|^{\eta q\frac{2d+1}{p}}+\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r,Y^{s,x}_r\right)dW_r\right|^{\eta q\frac{2d+1}{p}}\Bigg\}$$
$$\times\Bigg(\Bigg\{c_\gamma^{\eta q}\rho^{\eta\sigma\gamma q}+\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t\int_{U_0} g\left(Y^{u,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{\eta q\frac{2d+1}{p}}$$
$$+\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t b\left(r,Y^{u,x}_r\right)dr\right|^{\eta q\frac{2d+1}{p}}+\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t \alpha\left(r,Y^{u,x}_r\right)dW_r\right|^{\eta q\frac{2d+1}{p}}\Bigg\}\wedge 1\Bigg)\Bigg]\Bigg)^{\frac{1}{\eta}}. \quad (3.53)$$

For every couple of time indeces $0 \le t_1 \le t_2 \le T$, let us denote by

$$\mathcal{F}^{W,N_p}_{t_1,t_2} = \sigma\left(\{W_t - W_{t_1},\, N_p\left((t_1,t]\times E\right),\, t\in[t_1,t_2]\,,\, E\in\mathcal{U}\}\cup\mathcal{N}\right), \quad (3.54)$$

where $\mathcal{N}$ is the family of negligible events in $(\Omega,\mathcal{F},\mathbb{P})$: we set $\mathbb{F}^{W,N_p}_{t_1} = (\mathcal{F}^{W,N_p}_{t_1,t})_{t\in[t_1,T]}$. In particular, $\mathbb{F}^{W,N_p}_0 = (\mathcal{F}^{W,N_p}_{0,t})_{t\in[0,T]}$ is the augmented filtration generated by W and N_p. We are considering the SDE (3.5) with deterministic coefficients, which are obviously predictable with respect to $\mathbb{F}^{W,N_p}_{t_1}$, for every $t_1 \in [0,T)$, once restricted to the interval $[t_1,T]$. Hence we can invoke [167, Theorem 117] (see also [15, Theorem 2]) to claim that, for every $t_1 \in [0,T)$ and $x \in \mathbb{R}$,

$$Y^{t_1,x} \quad \text{is a strong solution of (3.5).} \quad (3.55)$$

This means that $Y^{t_1,x}$ is $\mathbb{F}^{W,N_p}_{t_1}$−adapted (in fact, it is also an $\mathbb{F}^{W,N_p}_{t_1}$−Markov process). As a consequence, for all $x \in \mathbb{R}^d$, $i = 1,2,3$, and $t \in [u,v]$, the random variable $Z^{u,x}_{i,t}$ (defined at the end of Subsection 3.2.1) is $\mathcal{F}^{W,N_p}_{u,v}$−measurable. Recalling the continuity of the $(\mathcal{D}_0,\mathcal{D})$−valued random fields $Z^{u,\cdot}_i$ and using the fact that countable sup of measurable functions is measurable, we deduce that the random variables $\sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}|Z^{u,x}_{i,t}|$, $i=1,2,3$, appearing in (3.53) are $\mathcal{F}^{W,N_p}_{u,v}$−measurable. It then follows that they are independent from $\mathcal{F}^{W,N_p}_{s,u}$. Indeed, since W and N_p are mutually independent (see [106, Theorem 6.3, Chapter II], or [125, Proposition 2.6]), a standard argument based on *Dynkin's theorem* ensures that $\mathcal{F}^{W,N_p}_{t_1,t_2}$ and $\mathcal{F}^{W,N_p}_{t_2,t_3}$ are independent, for any $0 \le t_1 < t_2 < t_3 \le T$. Therefore

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg(\Bigg\{c_\gamma^{\eta q}\rho^{\eta\sigma\gamma q} + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t\int_{U_0} g\left(Y^{u,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{\eta q\frac{2d+1}{p}}\\
&\quad + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t b\left(r,Y^{u,x}_r\right)dr\right|^{\eta q\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t \alpha\left(r,Y^{u,x}_r\right)dW_r\right|^{\eta q\frac{2d+1}{p}}\Bigg\}\wedge 1\Bigg)\Bigg|\mathcal{F}^{W,N_p}_{s,u}\Bigg]\\
&= \mathbb{E}\Bigg[\Bigg(\Bigg\{c_\gamma^{\eta q}\rho^{\eta\sigma\gamma q} + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t\int_{U_0} g\left(Y^{u,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right|^{\eta q\frac{2d+1}{p}}\\
&\quad + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t b\left(r,Y^{u,x}_r\right)dr\right|^{\eta q\frac{2d+1}{p}} + \sup_{|x|\le\rho^{-\sigma}}\sup_{u\le t\le v}\left|\int_u^t \alpha\left(r,Y^{u,x}_r\right)dW_r\right|^{\eta q\frac{2d+1}{p}}\Bigg\}\wedge 1\Bigg)\Bigg]. \qquad (3.56)
\end{aligned}$$

From now on, we denote by $\tilde{p} = \eta q(2d+1)p^{-1}$ and enumerate positive constants implicitly assuming their dependence on the parameters $\gamma,\eta,p,q,d,m,T,\{K_r,\ r\ge 2\}$. By virtue of Corollary 3.9 and [125, Equation (3.6)], for every $\tilde{q} > 2d$,

$$\begin{aligned}
\mathbb{E}\Big[\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|Y^{s,x}_t\right|^{\tilde{q}}\Big] &\le 2^{\tilde{q}-1}\Big(\mathbb{E}\Big[\sup_{s\le t\le u}\left|Y^{s,0}_t\right|^{\tilde{q}}\Big] + \mathbb{E}\Big[\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|Y^{s,x}_t - Y^{s,0}_t\right|^{\tilde{q}}\Big]\Big)\\
&\le c_2\left(1+\rho^{-\sigma(\tilde{q}+1)}\right) \le 2c_2\rho^{-\sigma(\tilde{q}+1)}. \qquad (3.57)
\end{aligned}$$

Taking $\tilde{p} > d$, (3.57) with $\tilde{q} = 2\tilde{p}$, (3.1) and Hölder's inequality entail

$$\mathbb{E}\left[\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t b\left(r,Y^{s,x}_r\right)dr\right|^{2\tilde{p}}\right] \le c_3\,\rho^{2\tilde{p}-\sigma(2\tilde{p}+1)}. \qquad (3.58)$$

In order to estimate the expected value in the right–hand side of (3.53), we also want to apply Proposition 3.8 to the continuous, $(\mathcal{D}_0,\mathcal{D})$−valued random fields

$$Z^{\xi,x}_1 = \int_\xi^\cdot \alpha\left(r,Y^{\xi,x}_r\right)dW_r, \qquad Z^{\xi,x}_2 = \int_\xi^\cdot\int_{U_0} g\left(Y^{\xi,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz), \quad x\in\mathbb{R}^d,$$

where $\xi = s,u$. We only show the case $\xi = s$, being $\xi = u$ analogous. We start from $Z^{s,\cdot}_1$, writing

$$\begin{aligned}
\mathbb{E}\left[\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r,Y^{s,x}_r\right)dW_r\right|^{2\tilde{p}}\right] &\le c_4\Bigg(\mathbb{E}\left[\sup_{s\le t\le u}\left|\int_s^t \alpha\left(r,Y^{s,0}_r\right)dW_r\right|^{2\tilde{p}}\right]\\
&\quad + \mathbb{E}\left[\sup_{|x|\le\rho^{-\sigma}}\sup_{s\le t\le u}\left|\int_s^t\left(\alpha\left(r,Y^{s,x}_r\right)-\alpha\left(r,Y^{s,0}_r\right)\right)dW_r\right|^{2\tilde{p}}\right]\Bigg) =: c_4\left(\mathbf{I}_2+\mathbf{II}_2\right).
\end{aligned}$$

Regarding $\mathbf{I}_2$, by (3.1), the Burkholder–Davis–Gundy inequality and [125, Equation (3.6)] we obtain $\mathbf{I}_2 \leq c_5\rho^{\tilde{p}}$. As for $\mathbf{II}_2$, since $\tilde{p} > d$ we can apply Proposition 3.8, which implies that, in the whole space Ω,

$$\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u} \left| \int_s^t \left(\alpha\left(r, Y_r^{s,x}\right) - \alpha\left(r, Y_r^{s,0}\right)\right) dW_r \right| \leq c_6 \widetilde{U}^{(1)}_{s,u,\tilde{p}} \rho^{-\sigma\left(1+\frac{1}{2\tilde{p}}\right)}.$$

Here $\widetilde{U}^{(1)}_{s,u,\tilde{p}}$ is the $2\tilde{p}-$integrable random variable

$$\widetilde{U}^{(1)}_{s,u,\gamma}(\omega) = \Big(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Big(\frac{\sup_{s\leq t\leq u}\left|Z_{1,t}^{s,x}(\omega) - Z_{1,t}^{s,y}(\omega)\right|}{|x-y|}\Big)^{2\tilde{p}} f\left(|x|\right) f\left(|y|\right) dx\, dy\Big)^{\frac{1}{2\tilde{p}}}, \quad \omega\in\Omega,$$

which satisfies, by (3.2), Lemma 3.3 and the Burkholder–Davis–Gundy inequality,

$$\mathbb{E}\left[\left(\widetilde{U}^{(1)}_{s,u,\tilde{p}}\right)^{2\tilde{p}}\right] \leq c_7\, \rho^{\tilde{p}}.$$

Thus,

$$\mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u} \left|\int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r\right|^{2\tilde{p}}\right] \leq c_8\left[\rho^{\tilde{p}} + \rho^{\tilde{p}-\sigma(2\tilde{p}+1)}\right] \leq 2\, c_8\, \rho^{\tilde{p}-\sigma(2\tilde{p}+1)}. \tag{3.59}$$

Moving on to $Z_2^{s,\cdot}$, we can argue as in (3.32) to conclude that

$$\mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u} \left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right|^{2\tilde{p}}\right] \leq c_9\left(\rho + \rho^{1-\sigma(2\tilde{p}+1)}\right) \leq 2\, c_9\, \rho^{1-\sigma(2\tilde{p}+1)}. \tag{3.60}$$

Analogously, if we require $\tilde{p} > 2d$, then

$$\mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u} \left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right|^{\tilde{p}}\right] \leq c_{10}\, \rho^{1-\sigma(\tilde{p}+1)}. \tag{3.61}$$

Going back to (3.53), by (3.56) and the law of iterated expectations with respect to $\mathcal{F}_{s,u}^{W,N_p}$ we compute

$$\begin{aligned}
&\mathbb{E}\left[d_0^{lu}\left(Y_s, Y_u\right)^q \cdot d_0^{lu}\left(Y_u, Y_v\right)^q\right] \leq c_{11}\Bigg\{\rho^{\sigma\gamma q} \\
&\quad + \left(\mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u}\left|\int_s^t b\left(r, Y_r^{s,x}\right) dr\right|^{\tilde{p}} + \sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u}\left|\int_s^t \alpha\left(r, Y_r^{s,x}\right) dW_r\right|^{\tilde{p}}\right]\right)^{\frac{1}{\eta}} \\
&\quad + \left(\mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{s\leq t\leq u}\left|\int_s^t\int_{U_0} g\left(Y_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right|^{\tilde{p}}\right] \left(\rho^{\eta\sigma\gamma q} + \mathbb{E}\left[\sup_{|x|\leq\rho^{-\sigma}} \sup_{u\leq t\leq v}\left|\int_u^t b\left(r, Y_r^{u,x}\right) dr\right|^{\tilde{p}}\right.\right.\right. \\
&\quad \left.\left.\left. + \sup_{|x|\leq\rho^{-\sigma}} \sup_{u\leq t\leq v}\left|\int_u^t\int_{U_0} g\left(Y_{r-}^{u,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right|^{\tilde{p}} + \sup_{|x|\leq\rho^{-\sigma}} \sup_{u\leq t\leq v}\left|\int_u^t \alpha\left(r, Y_r^{u,x}\right) dW_r\right|^{\tilde{p}}\right)\right]\right)^{\frac{1}{\eta}}\Bigg\} \\
&\quad =: c_{11}\left\{\rho^{\sigma\gamma q} + \mathbf{I}_3 + \mathbf{II}_3\right\}.
\end{aligned} \tag{3.62}$$

To perform this passage, we observe that the argument of the second expected value in (3.53) is a product between a sum $\overline{S}$, which includes an integral in $\widetilde{N}_p$, and $\min\{S,1\}$, where S is another sum. In (3.62), $\mathbf{II}_3$ is obtained multiplying the integral with respect to $\widetilde{N}_p$ in $\overline{S}$ with S, while $\rho^{\sigma\gamma q}$ and $\mathbf{I}_3$ multiplying the remaining terms in $\overline{S}$ by 1.

By (3.58)-(3.59) and the Cauchy–Schwarz inequality,

$$\mathbf{I}_3 \leq c_{12}\rho^{\frac{\tilde{p}}{2\eta}-\frac{\sigma}{2\eta}(2\tilde{p}+1)}.$$

The same results together with (3.60)-(3.61) also yield

$$\mathbf{II}_3 \leq c_{13}\rho^{\frac{1}{\eta}-\frac{\sigma}{\eta}(\tilde{p}+1)}\left(\rho^{\sigma\gamma q}+\rho^{\frac{1}{2\eta}-\frac{\sigma}{2\eta}(2\tilde{p}+1)}\right).$$

Combining the two previous estimates in (3.62) we deduce that

$$\mathbb{E}\left[d_0^{lu}\left(Y_s,Y_u\right)^q\cdot d_0^{lu}\left(Y_u,Y_v\right)^q\right] \leq c_{14}\left(\rho^{\sigma\gamma q}+\rho^{\frac{\tilde{p}}{2\eta}-\frac{\sigma}{2\eta}(2\tilde{p}+1)}+\rho^{\sigma\gamma q+\frac{1}{\eta}(1-\sigma(\tilde{p}+1))}+\rho^{\frac{3}{2\eta}-\frac{\sigma}{\eta}\left(2\tilde{p}+\frac{3}{2}\right)}\right). \quad (3.63)$$

At this point, it only remains to select appropriate parameters to recover (3.46) from (3.63). Recall that $\eta=\frac{5}{4}$, hence $\frac{3}{2\eta}=\frac{6}{5}>1$. Collecting the conditions written throughout the lines above, we pick (p,q,σ,γ) according to the following steps.

1. $p\in(4d+1,\infty)$, so that $p^{-1}(2d+1)<1-2dp^{-1}$;
2. $q\in(p(2d+\frac{1}{2})(2d+1)^{-1},\,p)$, where the lower bound ensures that $q\,(2d+1)\,p^{-1}>2d+\frac{1}{2}$. In turn, this yields
$$q>1/2, \qquad \tilde{p}=\eta q\frac{2d+1}{p}>2d, \qquad \frac{\tilde{p}}{2\eta}>\frac{3}{2\eta};$$
3. $\sigma\in(0,(8\tilde{p}+6)^{-1})$, i.e., σ is so small that
$$\frac{3}{2\eta}-\frac{\sigma}{\eta}\left(2\tilde{p}+\frac{3}{2}\right)>1. \quad (3.64)$$
This bound also guarantees that $\sigma\,(\tilde{p}+1)<1$;
4. $\gamma\in\left(3(2\eta\sigma q)^{-1},\infty\right)$, so that $\sigma\gamma q>\frac{3}{2\eta}$.

With the previous prescriptions and noticing that $\frac{2\tilde{p}+1}{2}<2\tilde{p}+\frac{3}{2}$, from (3.63) we conclude that

$$\mathbb{E}\left[d_0^{lu}\left(Y_s,Y_u\right)^q\cdot d_0^{lu}\left(Y_u,Y_v\right)^q\right] \leq c_{15}\rho^{\frac{3}{2\eta}-\frac{\sigma}{\eta}\left(2\tilde{p}+\frac{3}{2}\right)} =: c_{15}\rho^{1+r}, \quad \text{where}$$
$$r=\frac{3}{2\eta}-\frac{\sigma}{\eta}\left(2\tilde{p}+\frac{3}{2}\right)-1. \quad (3.65)$$

Since $r>0$ by (3.64), recalling that $\rho=v-s<1$, we see that Equation (3.65) reduces to (3.46). The proof is then complete. ■

Using the càdlàg version Z of the process Y given by Theorem 3.14, we consider

$$Z_{1,t}^{s,x}=\int_s^t \alpha\left(r,Z_r^{s,x}\right)dW_r, \quad Z_{2,t}^{s,x}=\int_s^t\int_{U_0} g\left(Z_{r-}^{s,x},r,z\right)\widetilde{N}_p\left(dr,dz\right), \quad Z_{3,t}^{s,x}=\int_s^t b\left(r,Z_r^{s,x}\right)dr. \quad (3.66)$$

Thanks to (3.42), for every $i = 1, 2, 3$, the estimate (3.51) for $d_0^{lu}(Y_s, Y_u)$ constitutes an upper bound for $d_0^{lu}(Z_{i,s}, Z_{i,u})$, upon substituting $\left(3 + c_1 \widetilde{U}_{u,p}^{(i)}\right)$ for $(2 + c_1 U_{u,p})$ in the right–hand side. Here the random variables $\widetilde{U}_{u,p}^{(i)}$ are defined according to Proposition 3.8 applied to the $(\mathcal{D}_0, \mathcal{D})$−valued random fields $Z_i^{u,\cdot}$. Therefore, recalling that $Z_i = (Z_{i,s})_{s\in[0,T]}$ are stochastically continuous processes with values in $(\mathcal{C}_0, \mathcal{C})$ by Corollary 3.11, the same computations as in the proof of Theorem 3.14 allow to invoke Corollary 3.13, which yields the following result.

Corollary 3.15. *For any $i = 1, 2, 3$, there exists a càdlàg version of the $(\mathcal{C}_0, \mathcal{C})$−valued process $Z_i = (Z_{i,s})_s$.*

In order not to complicate the notation, we keep denoting by Z_i, $i = 1, 2, 3$, the càdlàg processes given by Corollary 3.15. Without loss of generality, we assume that Z and Z_i are càdlàg in the whole space Ω.
At the end of Subsection 3.2.1 (see (3.17)), we have determined the existence of an a.s. event Ω_s independent from x where the SDE (3.5) is satisfied. Now, combining Theorem 3.14 with Corollary 3.15, we can get rid of the dependence of such Ω_s from the initial time s. This is done in the next lemma, where we are also able to establish the flow property (3.14) in an a.s. event independent from the space and time variables.

Lemma 3.16. *There exists an a.s. event Ω' (independent from x, s and t) such that*

$$Z_t^{s,x} = x + \int_s^t b\left(r, Z_r^{s,x}\right) dr + \int_s^t \alpha\left(r, Z_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} g\left(Z_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right),$$
$$s \in [0,T)\,,\ t \in [0,T]\,,\ x \in \mathbb{R}^d,\ \omega \in \Omega'. \tag{3.67}$$

Furthermore,

$$Z_t^{s,x}(\omega) = Z_t^{u, Z_u^{s,x}(\omega)}(\omega)\,, \quad 0 \le s < u < t \le T,\ x \in \mathbb{R}^d,\ \omega \in \Omega'. \tag{3.68}$$

Proof. By (3.17), there exists an a.s. event Ω_1 –independent from $s \in [0,T)$ and $x \in \mathbb{R}^d$– such that

$$Z_t^{s,x}(\omega) = x + \sum_{i=1}^{3} Z_{i,t}^{s,x}(\omega)\,, \quad s \in [0,T) \cap \mathbb{Q},\ t \in [0,T]\,,\ x \in \mathbb{R}^d,\ \omega \in \Omega_1. \tag{3.69}$$

Thanks to the càdlàg property of the $(\mathcal{C}_0, \mathcal{C})$−valued processes Z and Z_i, $i = 1, 2, 3$, a standard approximation argument in s ensures that (3.69) holds for every $s \in [0,T)$. Hence (3.67) is satisfied in Ω_1.

As for the flow property in (3.68), note that by (3.14) in Corollary 3.5 there is an a.s. event Ω_2 –independent from x, s, u– such that

$$Z_t^{s,x}(\omega) = Z_t^{u, Z_u^{s,x}(\omega)}(\omega)\,, \quad 0 \le s < u < t \le T,\ s, u \in \mathbb{Q},\ x \in \mathbb{R}^d,\ \omega \in \Omega_2. \tag{3.70}$$

Fix $x \in \mathbb{R}^d$, $\omega \in \Omega_2$ and $s, u \in \mathbb{R} \setminus \mathbb{Q}$ such that $0 \le s < u < T$. Consider $t \in (u, T]$ and a sequence $(s_n)_n \subset [0,u) \cap \mathbb{Q}$ such that $s_n \downarrow s$ as $n \to \infty$: we know that $\lim_{n\to\infty} Z_t^{s_n,x}(\omega) = Z_t^{s,x}(\omega)$. Moreover, we take a sequence $(u_n)_n \subset (u,t) \cap \mathbb{Q}$, with $u_n \downarrow u$ as $n \to \infty$, so that $\lim_{n\to\infty} Z_{u_n}^{s_n,x}(\omega) = Z_u^{s,x}(\omega)$. At this point, from the càdlàg property of the $(\mathcal{C}_0, \mathcal{C})$−valued process Z, we deduce that $Z^{u,y}(\omega) = \mathcal{D}_0 - \lim_{n\to\infty} Z^{u_n,y}(\omega)$ locally uniformly in $y \in \mathbb{R}^d$. As a consequence,

$$Z^{u, Z_u^{s,x}(\omega)}(\omega) = \mathcal{D}_0 - \lim_{n\to\infty} Z^{u_n, Z_{u_n}^{s_n,x}(\omega)}(\omega)\,.$$

By (3.70), $Z_t^{s_n,x}(\omega) = Z_t^{u_n, Z_{u_n}^{s_n,x}(\omega)}(\omega)$, $n \in \mathbb{N}$, hence we can pass to the limit as $n \to \infty$ to obtain (3.68) in Ω_2.

The a.s. event Ω' is obtained by setting $\Omega' = \Omega_1 \cap \Omega_2$, completing the proof. ■

Theorem 3.14 and Lemma 3.16 coupled with Lemma 3.10 show that *Z is the sharp stochastic flow generated by the SDE* (3.5) (without large jumps) according to Theorem 3.2.

More precisely, Z satisfies Points 1.-3. in Definition 3.1 by Lemma 3.16, while the fact that Z is a stochastically continuous, càdlàg process with values in $(\mathcal{C}_0, \mathcal{C})$ –which entail Points 2.-4. in Definition 3.1– is guaranteed by Theorem 3.14 and Lemma 3.10.

Combining stochastic continuity and càdlàg property, we also infer that the process $Z = (Z_s)_{s\in[0,T]}$ *has no fixed–time discontinuities,* meaning that, for every $s \in [0, T]$,

$$Z_{s-}(\omega) = Z_s(\omega), \quad \omega \in \Omega_s. \tag{3.71}$$

We conclude this part by stating a couple of lemmas discussing further properties of the sharp flow $Z_t^{s,x}$: they will be used in Section 3.3 while studying the SDE (3.3) with large jumps. Their proofs are postponed to Appendix 3.C. The first result regards the joint–measurability.

Lemma 3.17. *For any $\bar{s}$, $\bar{t} \in [0, T]$, the mapping $Z\colon \Omega \times [0, \bar{s}] \times \mathbb{R}^d \times [0, \bar{t}] \to \mathbb{R}^d$ defined by $Z(\omega, s, x, t) = Z_t^{s,x}(\omega)$ is $\mathcal{F}_{\bar{t}} \otimes \mathcal{B}([0, \bar{s}] \times \mathbb{R}^d \times [0, \bar{t}])-$measurable.*

The second lemma shows that the flow $Z_t^{s,x}$ can be used to construct a solution to (3.5) when the initial condition η is only a measurable random variable.

Lemma 3.18. *Fix $s \in [0, T]$ and let $\eta \in L^0(\mathcal{F}_s)$. Then, denoting by $Z^{s,\eta}$ the process defined by $Z_t^{s,\eta}(\omega) = Z_t^{s,\eta(\omega)}(\omega)$, $\omega \in \Omega$, $t \in [0, T]$, there exists an a.s. event $\Omega_{s,\eta}$ where the following equation is satisfied:*

$$Z_t^{s,\eta} = \eta + \int_0^t 1_{\{r>s\}} b\left(r, Z_r^{s,\eta}\right) dr + \int_0^t 1_{\{r>s\}} \alpha\left(r, Z_r^{s,\eta}\right) dW_r$$
$$+ \int_0^t \int_{U_0} 1_{\{r>s\}} g\left(Z_{r-}^{s,\eta}, r, z\right) \widetilde{N}_p(dr, dz), \quad t \in [0, T]. \tag{3.72}$$

In particular, $Z^{s,\eta}$ is the pathwise unique solution of (3.72)*. Further, for every $\omega \in \Omega_{s,\eta}$, $t \in [0, T]$ (cfr.* (3.66)*),*

$$\begin{aligned} &\left(\int_0^t 1_{\{r>s\}} \alpha\left(r, Z_r^{s,\eta}\right) dW_r\right)(\omega) = Z_{1,t}^{s,\eta(\omega)}(\omega), \qquad \int_0^t 1_{\{r>s\}} b\left(r, Z_r^{s,\eta}(\omega)\right) dr = Z_{3,t}^{s,\eta(\omega)}(\omega), \\ &\left(\int_0^t \int_{U_0} 1_{\{r>s\}} g\left(Z_{r-}^{s,\eta}, r, z\right) \widetilde{N}_p(dr, dz)\right)(\omega) = Z_{2,t}^{s,\eta(\omega)}(\omega). \end{aligned} \tag{3.73}$$

We observe that, using (3.73) and arguing as in the proof of Lemma 3.16, it is possible to show that $Z^{s,\eta}$ in Lemma 3.18 satisfies (3.72) in an a.s. event Ω_η depending only on η.

Remark 3.3. *In Subsections 3.2.2-3.2.3 we have considered the process Y with values in the complete metric space $\mathcal{C}_0 = (C(\mathbb{R}^d; \mathcal{D}_0), d_0^{lu})$. Since $\mathcal{C}_0$ is not separable, we have endowed it with the $\sigma-$algebra $\mathcal{C}$ generated by the projections π_x –strictly smaller than the Borel $\sigma-$algebra– in order to overcome measurability issues.*

An alternative approach which, at a first glance, might appear to be more natural is the following one. Denote by $\mathcal{D}_S$ *the space of* $\mathbb{R}^d-$*valued, càdlàg functions on* $[0,T]$ *endowed with the Skorokhod topology* J_1*, i.e.,* $\mathcal{D}_S = \left(\mathcal{D}\left([0,T];\mathbb{R}^d\right), J_1\right)$*. According to [31, Section 12, Chapter 3],* $\mathcal{D}_S$ *is a Polish space with the following metric defining the topology:*

$$d_S(x,y) = \inf_{\lambda\in\Lambda}\left\{\|\lambda\|^0 \vee \|x - y\circ\lambda\|_0\right\}, \quad x,y\in\mathcal{D}\left([0,T];\mathbb{R}^d\right). \tag{3.74}$$

Here Λ *is the set of continuous and strictly increasing functions* λ *such that* $\lambda(0)=0$ *and* $\lambda(T)=T$*, and* $\|\lambda\|^0 = \sup_{s<t}\left|\log\left(\frac{\lambda(t)-\lambda(s)}{t-s}\right)\right|$*. Note that* J_1 *is weaker than the topology generated by the uniform convergence. Indeed, taking* $\lambda = I$*,*

$$d_S(x,y) \le \|x-y\|_0, \quad x,y\in\mathcal{D}\left([0,T];\mathbb{R}^d\right). \tag{3.75}$$

Hence, for every $s\in[0,T]$*,* $Y_s\in C(\mathbb{R}^d;\mathcal{D}_S)$*. By [122], the complete metric space* $\left(C\left(\mathbb{R}^d;\mathcal{D}_S\right), d_S^{lu}\right)$*, where*

$$d_S^{lu}(f,g) = \sum_{N=1}^{\infty}\frac{1}{2^N}\frac{\sup_{|x|\le N} d_S(f(x),g(x))}{1+\sup_{|x|\le N} d_S(f(x),g(x))}, \quad f,g\in C\left(\mathbb{R}^d;\mathcal{D}_S\right),$$

is also separable. Therefore we can argue as at the beginning of page 702 in [149] to infer the measurability of Y_s *with respect to the Borel* $\sigma-$*algebra associated with* d_S^{lu}*. Observe that, by* (3.75) *and the fact that* $x(1+x)^{-1}$ *is increasing in* $\mathbb{R}_+$*,* $d_S^{lu}(f,g)\le d_0^{lu}(f,g)$*,* $f,g\in C\left(\mathbb{R}^d;\mathcal{D}_S\right)$*. Thus, we can exploit the same computations as those presented in the chapter to obtain the existence of a càdlàg modification* $\widetilde{Y}$ *of the* $C\left(\mathbb{R}^d;\mathcal{D}_S\right)-$*valued process* Y*. Moreover, using [110, Proposition 2.1, Chapter VI] we can prove Lemma 3.16, as well. However,* $\widetilde{Y}$ *is not the* sharp *stochastic flow associated with* (3.5) *according to Definition 3.1, because (ii) and (iii)] in Point 2. hold in a weaker sense, namely replacing* $\mathcal{D}_0$ *with* $\mathcal{D}_S$*. As a consequence, for every* $\bar{s}\in[0,T]$*,* $t\in[0,T]$*,* $x\in\mathbb{R}^d$ *and* $\omega\in\Omega$*, we can not deduce that* $\lim_{x\to\bar{x}}\widetilde{Y}_t^{\bar{s},x}(\omega) = \widetilde{Y}_t^{\bar{s},\bar{x}}(\omega)$ *or* $\lim_{s\downarrow\bar{s}}\widetilde{Y}_t^{s,\bar{x}}(\omega) = \widetilde{Y}_t^{\bar{s},\bar{x}}(\omega)$*.*

3.3 SDEs with large jumps

In this section, we investigate the SDE (3.3) with $f\neq 0$. Given $s\in[0,T)$ and $\eta\in L^0(\mathcal{F}_s)$, we study

$$\begin{aligned} X_t^{s,\eta} &= \eta + \int_s^t b\left(r, X_r^{s,\eta}\right)dr + \int_s^t \alpha\left(r, X_r^{s,\eta}\right)dW_r \\ &\quad + \int_s^t\int_{U_0} g\left(X_{r-}^{s,\eta}, r, z\right)\widetilde{N}_p(dr,dz) + \int_s^t\int_{U\setminus U_0} f\left(X_{r-}^{s,\eta}, r, z\right)N_p(dr,dz), \quad t\in[s,T]. \end{aligned} \tag{3.76}$$

Compared to the SDE (3.5) that we have been discussing in Section 3.2, (3.76) presents an additional integral with respect to the (non–compensated) Poisson random measure N_p. For this reason, (3.76) is often referred to as an SDE with *large jumps*. In particular, given $\omega\in\Omega$, one can read

$$\left(\int_s^t\int_{U\setminus U_0} f\left(X_{r-}^{s,\eta}, r, z\right)N_p(dr,dz)\right)(\omega) = \sum_{r\in D_p(\omega)\cap(s,t]} 1_{U\setminus U_0}\left(p_r(\omega)\right) f(X_{r-}^{s,\eta}(\omega), r, p_r(\omega)),$$

with the sum on the right–hand side which is finite $\mathbb{P}-$a.s., because $\nu(U\setminus U_0)<\infty$ implies that $D_p(\omega)$ is discrete, $\mathbb{P}-$a.s. We study (3.76) by adapting an interlacing method described, for example, in

[43, 106, 128]. Such an adaptation is not trivial for our scope of finding a sharp stochastic flow generated by (3.76), as detailed in Remark 3.4.

Recall that a solution to (3.76) is a càdlàg, $\mathbb{R}^d-$valued, $\mathbb{F}-$adapted process $X^{s,\eta} = (X_t^{s,\eta})_{s\le t\le T}$ satisfying (3.76) up to indistinguishability. As usual, we extend the trajectories of $X^{s,\eta}$ in the whole interval $[0,T]$ by setting $X_t^{s,\eta} = X_s^{s,\eta}$, $t \in [0,s]$. Under our assumptions on the coefficients (see Section 3.1), there exists a pathwise unique solution of (3.76).

Remark 3.4. *The existence of a pathwise unique solution of* (3.76) *can be proven by adapting the interlacing procedure described in, e.g, [43, Subsection 3.2] and [106, Section 9, Chapter IV]) to the case* $s \neq 0$. *To do this, starting from* W, p *and* $\mathbb{F}$, *we construct a Brownian motion* $W^{(s)}$ *and a stationary Poisson point process* $p^{(s)}$ *with respect to a filtration* $\mathbb{F}^{(s)}$, *for every* $s \in (0,T)$. *On the other hand, it is not clear how to use this approach to prove the existence of a* sharp *stochastic flow generated by* (3.76) *according to Definition 3.1. In particular, it is not clear how to analyze the regularity of the flow with respect to the initial time* s. *To overcome this issue, we follow an argument relying on the sharp stochastic flow* $Z_t^{s,x}$ *generated by the SDE* (3.5) *with small jumps (see Section 3.2). Remarkably, we are also able to obtain an explicit expression –based on* $Z_t^{s,x}$*– for the solution of* (3.76), *from which we deduce the regularity properties that we are looking for.*

We now prove Theorem 3.1, which asserts the existence of a sharp stochastic flow generated by (3.76) according to Definition 3.1. In order to make the proof easier to follow, in Theorem 3.19 we reformulate the statement of Theorem 3.1 in an expanded version.

Theorem 3.19. *There exist an* $\mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}^d\times[0,T])-$*measurable function* $X\colon \Omega\times[0,T]\times\mathbb{R}^d\times[0,T]\to\mathbb{R}^d$, *denoted by* $X_t^{s,x}(\omega) = X(\omega,s,x,t)$, *and an almost sure event* Ω'' *(independent of* s,t *and* x*) such that*

$$X_t^{s,x} = x + \int_s^t b\left(r, X_r^{s,x}\right) dr + \int_s^t \alpha\left(r, X_r^{s,x}\right) dW_r + \int_s^t\int_{U_0} g\left(X_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr,dz\right)$$
$$+ \int_s^t\int_{U\setminus U_0} f\left(X_{r-}^{s,x}, r, z\right) N_p\left(dr,dz\right), \quad s\in[0,T),\, t\in[0,T],\, x\in\mathbb{R}^d,\, \omega\in\Omega'', \tag{3.77}$$

and such that the flow property holds:

$$X_t^{s,x}(\omega) = X_t^{u,X_u^{s,x}(\omega)}(\omega), \quad 0\le s<u<t\le T,\, x\in\mathbb{R}^d,\, \omega\in\Omega''. \tag{3.78}$$

Furthermore, the process $(X^s)_{s\in[0,T]}$ *is stochastically continuous in the sense of Point 4. in Definition 3.1, and, for every* $\omega\in\Omega$, *the mapping* $(s,t,x)\mapsto X_t^{s,x}(\omega)$ *satisfies Point 2. in Definition 3.1.*

Proof. In order not to complicate the notation, we are going to construct the flow $X_t^{s,x}$ generated by (3.77) with $t\in[0,T)$, excluding the upper bound T. Since the proof is rather long, we divide it into several steps.

Step I: Construction of the flow $X_t^{s,x}$. Denote by $P_t = N_p((0,t]\times(U\setminus U_0))$, $t>0$, and set $P_0 = 0$: since $\nu(U\setminus U_0)\in(0,\infty)$, $P = (P_t)_{t\ge0}$ is a Poisson process with intensity $\nu(U\setminus U_0)$. Let τ_n, $n\in\mathbb{N}$, be the arrival times for the jumps of P. It is well known that $\tau_n(\omega)\uparrow\infty$ as $n\to\infty$, for every $\omega\in\Omega_1$, where Ω_1 is an a.s. event (see, e.g., [164, Theorem 21.3] and the subsequent comment). Notice that P is càdlàg and continuous in probability, hence it does not jump at time T, $\mathbb{P}-$a.s. Thus, we suppose that $\tau_n\neq T$ in Ω_1, for every $n\in\mathbb{N}$.

To construct the solution of (3.77) we use $Z = (Z_s)_{s\in[0,T]}$: the càdlàg, $(\mathcal{C}_0,\mathcal{C})-$valued process studied in Subsection 3.2.3 satisfying (3.67), see Lemma 3.16. Note that (3.67) is the analogous of (3.77) without the integral in N_p, i.e., without the "large jumps". We argue that $Z^{s,x}_\cdot(\omega)$ does not jump at $\tau_n(\omega)\in(s,T)$, for every $s\in[0,T)$, $x\in\mathbb{R}^d$ and $\omega\in\Omega_2$, where Ω_2 is an a.s. event. To see this, we take a sequence $(U_n)_n\subset\mathcal{U}$ such that $\nu(U_n)<\infty$ and $\cup_n U_n = U$, which exists because $\nu(dz)$ is $\sigma-$finite. Moreover, we denote by

$$g_n(r,z,\omega) = 1_{(-n,n)}\left(g\left(Z^{s,x}_{r-}(\omega),r,z\right)\right)g\left(Z^{s,x}_{r-}(\omega),r,z\right),\quad n\in\mathbb{N}.$$

By construction of the stochastic integral with respect to $\widetilde{N}_p$ (see [106, Section 3, Chapter II]),

$$\lim_{n\to\infty}\int_s^t\int_{U_0\cap U_n} g_n(r,z,\cdot)\,\widetilde{N}_p(dr,dz) = \int_s^t\int_{U_0} g\left(Z^{s,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz),$$

where the limit is uniform on compacts in probability. It follows that, $\mathbb{P}-$a.s.,

$$\sup_{t\in[s,T]}\left|\int_s^t\int_{U_0\cap U_{n_k}} g_{n_k}(r,z,\cdot)\,\widetilde{N}_p(dr,dz) - \int_s^t\int_{U_0} g\left(Z^{s,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right| \underset{k\to\infty}{\longrightarrow} 0. \tag{3.79}$$

Since $p_{\tau_n}\in U\setminus U_0$ and, for every $k\in\mathbb{N}$, for $\mathbb{P}-$a.s. $\omega\in\Omega$,

$$\left(\int_s^t\int_{U_0\cap U_{n_k}} g_{n_k}(r,z,\cdot)\,\widetilde{N}_p(dr,dz)\right)(\omega) = \sum_{r\in D_p(\omega)\cap(s,t]} g_{n_k}(r,p_r(\omega),\omega)\,1_{U_0\cap U_{n_k}}(p_r(\omega)) - \int_s^t dr\int_{U_0\cap U_{n_k}} g_{n_k}(r,z,\omega)\,\nu(dz),\quad t\in[s,T],$$

these approximating processes do not jump at time $\tau_n\in(s,T)$ for all $n\in\mathbb{N}$, $\mathbb{P}-$a.s. Therefore, by (3.79), the process $\int_s^\cdot\int_{U_0} g\left(Z^{s,x}_{r-},r,z\right)\widetilde{N}_p(dr,dz)$ does not jump either at time $\tau_n\in(s,T)$ in an a.s. event depending on s and x. Whence the same conclusion holds for $Z^{s,x}_\cdot$ in an a.s. event $\Omega_{s,x}$, by (3.67). Define the a.s. event

$$\Omega_2 = \bigcap_{s\in[0,T)\cap\mathbb{Q}}\ \bigcap_{x\in\mathbb{Q}^d}\Omega_{s,x}:$$

we are going to show that $Z^{s,x}_\cdot(\omega)$ does not jump at $\tau_n(\omega)\in(s,T)$ for every $x\in\mathbb{R}^d$, $s\in[0,T)$ and $\omega\in\Omega_2$. Fix $\omega\in\Omega_2$, $s\in[0,T)\cap\mathbb{Q}$, $x\in\mathbb{R}^d$ and take a sequence $(x_m)_m\in\mathbb{Q}^d$ such that $x_m\to x$ as $m\to\infty$. Since the map $Z^{s,\cdot}(\omega):\mathbb{R}^d\to\mathcal{D}_0$ is continuous and $Z^{s,x_m}_\cdot(\omega)$ is continuous at $\tau_n(\omega)$, for all $m\in\mathbb{N}$, we conclude that $Z^{s,x}_\cdot(\omega)$ is continuous at $\tau_n(\omega)$, as well. Indeed, uniform convergence in t preserves continuity. An analogous argument relying on the càdlàg property of the map $Z^{\cdot,x}(\omega):[0,T]\to\mathcal{D}_0$ allows to deduce the continuity of $Z^{s,x}_\cdot(\omega)$ at $\tau_n(\omega)$ for every $s\in[0,T)$, as desired.

Recalling the almost certain event Ω' given by Lemma 3.16, we define $\Omega_3=\Omega_1\cap\Omega_2\cap\Omega'$. Without loss of generality, we suppose that $Z^{s,x}_t(\omega)=x$ for every $0\le t\le s\le T$, $x\in\mathbb{R}^d$ and $\omega\in\Omega_3$. For the sake of shortness, from now on

$$\text{we denote by }\ \tau_n\ \text{ the random variable }\tau_n\wedge T,\ n\in\mathbb{N}. \tag{3.80}$$

We now construct the solution X of (3.76) using an ad hoc, path–by–path, interlacing procedure (see Remark 3.4). For $s=T$, we just assign $X^{T,x}_t=x$. Take $s\in[0,T)$ and $x\in\mathbb{R}^d$. First, we set $X^{s,x}_t(\omega)=x$

for $\omega \in \Omega \setminus \Omega_3$, $t \in [0,T]$. Secondly, fix $\omega \in \Omega_3$ and denote by $\tau_{n(s,\omega)}(\omega) = \min_n\{\tau_n(\omega) > s\}$. In words, if $\tau_{n(s,\omega)}(\omega) < T$, then it represents the first jump time of P occurring (strictly) after time s and before time T. In the sequel, we omit ω to keep notation simple. We define

$$X_t^{s,x} = Z_t^{s,x}, \quad t \in \left[0, \tau_{n(s)}\right). \tag{3.81}$$

If $\tau_{n(s)} = T$ the construction is over. Otherwise, for $t = \tau_{n(s)}$ we set

$$\begin{aligned} X_{\tau_{n(s)}}^{s,x} &= X_{\tau_{n(s)}-}^{s,x} + f\left(X_{\tau_{n(s)}-}^{s,x}, \tau_{n(s)}, p_{\tau_{n(s)}}\right) = Z_{\tau_{n(s)}-}^{s,x} + f\left(Z_{\tau_{n(s)}-}^{s,x}, \tau_{n(s)}, p_{\tau_{n(s)}}\right) \\ &= Z_{\tau_{n(s)}}^{s,x} + f\left(Z_{\tau_{n(s)}}^{s,x}, \tau_{n(s)}, p_{\tau_{n(s)}}\right), \end{aligned} \tag{3.82}$$

where the last equality is due to the fact that $\tau_{n(s)}$ is not a jump time for $Z_\cdot^{s,x}$, because $\omega \in \Omega_3 \subset \Omega_2$. Next, we define

$$X_t^{s,x} = Z_t^{\tau_{n(s)}, X_{\tau_{n(s)}}^{s,x}}, \quad t \in \left[\tau_{n(s)}, \tau_{n(s)+1}\right). \tag{3.83}$$

This argument by steps can be repeated to cover the whole interval $[s,T)$. More precisely, for every $m \in \mathbb{N}$ such that $\tau_{n(s)+m} < T$, we define recursively

$$X_t^{s,x} = \begin{cases} X_{\tau_{n(s)+m}-}^{s,x} + f\left(X_{\tau_{n(s)+m}-}^{s,x}, \tau_{n(s)+m}, p_{\tau_{n(s)+m}}\right), & t = \tau_{n(s)+m}, \\ Z_t^{\tau_{n(s)+m}, X_{\tau_{n(s)+m}}^{s,x}}, & t \in [\tau_{n(s)+m}, \tau_{n(s)+m+1}), \end{cases} \tag{3.84}$$

In particular, since $\omega \in \Omega_3 \subset \Omega_2$ we observe that

$$X_{\tau_{n(s)+m}-}^{s,x} = Z_{\tau_{n(s)+m}}^{\tau_{n(s)+m-1}, X_{\tau_{n(s)+m-1}}^{s,x}}. \tag{3.85}$$

We finally extend the map $X_\cdot^{s,x}$ to $[0,T]$ by setting $X_T^{s,x} = X_{T-}^{s,x}$.

<u>Step II</u>: *The process* $(X_t^{s,x})_{t\in[0,T]}$ *is* $\mathbb{F}-$*adapted.* The claim is trivial if $s = T$ because $X_t^{T,x} = x$, so we consider $s \in [0,T)$. For every $t \in [0,T)$, setting $\tau_0 = 0$ we have (recall (3.80))

$$X_t^{s,x} = x1_{\Omega\setminus\Omega_3} + x1_{\{t\le s\}}1_{\Omega_3} + \sum_{n=1}^{\infty} Z_t^{\tau_{n-1}\vee s, X_{\tau_{n-1}\vee s}^{s,x}} 1_{\{\tau_{n-1}\vee s \le t < \tau_n \vee s\}}1_{\Omega_3}. \tag{3.86}$$

Notice that the series in (3.86) is actually a finite sum, as $\tau_n(\omega) = T$ definitively in Ω_3, hence $[\tau_{n-1}(\omega)\vee s, \tau_n \vee s(\omega)) = \emptyset$ definitively in Ω_3. In what follows, we write $\tau_n^s = \tau_n \vee s$, $n \in \mathbb{N}_0$. Recalling that the filtration $\mathbb{F}$ is complete, $x1_{\Omega\setminus\Omega_3}$ and $x1_{\{t\le s\}}1_{\Omega_3}$ are $\mathcal{F}_t-$measurable. Since $(\tau_n^s)_n$ is a sequence of $\mathbb{F}-$stopping times, the sets $\{\omega : \tau_{n-1}^s(\omega) \le t < \tau_n^s(\omega)\}$, $n \in \mathbb{N}$, are $\mathcal{F}_t-$measurable. As a consequence, $Z_t^{s,x}1_{\{s\le t<\tau_1^s\}}1_{\Omega_3}$ –the first term of the series in (3.86)– is $\mathcal{F}_t-$measurable. Moreover, by Lemma 3.17, $Z_{\tau_1^s}^{s,x}1_{\{\tau_1^s\le t\}}$ is $\mathcal{F}_t-$measurable, so (by (3.82))

$$X_{\tau_1^s}^{s,x}1_{\{\tau_1^s\le t\}}1_{\Omega_3} = \left(Z_{\tau_1^s}^{s,x} + 1_{\{\tau_1^s>s\}}f\left(Z_{\tau_1^s}^{s,x}, \tau_1^s, p_{\tau_1^s}\right)\right)1_{\{\tau_1^s\le t\}}1_{\Omega_3}$$

is $\mathcal{F}_t-$measurable, too. Hence another application of Lemma 3.17 yields the $\mathcal{F}_t-$measurability of the second term of the series in (3.86), i.e.,

$$Z_t^{\tau_1^s, X_{\tau_1^s}^{s,x}} 1_{\{\tau_1^s\le t<\tau_2^s\}}1_{\Omega_3}.$$

At this point, an induction argument based on the recursive definition law in (3.84) allows us to conclude that all the addends in the series (3.86) are $\mathcal{F}_t$–measurable. Therefore, considering also that $X^{s,x}$ is left–continuous in T, we deduce that the process $X^{s,x}$ is $\mathbb{F}$–adapted, as desired.

Step III: The regularity of the flow $X_t^{s,x}$. The aim of this part is to prove (i)-(ii)-(iii) in Definition 3.1. We only analyze the case $\omega \in \Omega_3$, being the other one trivial ($X_t^{s,x}(\omega) = x$, $\omega \in \Omega \setminus \Omega_3$). Conditions (i)-(ii) are immediate also for $s = T$ because $X_t^{T,x} = x$, so we consider $s \in [0,T)$. Recalling that the series in (3.86) is actually a finite sum, for every $x \in \mathbb{R}^d$ the càdlàg property with respect to $t \in [0,T]$ is evident, because the path $X_\cdot^{s,x}(\omega)$ is constructed by combining a finite number of càdlàg trajectories of the flow $Z_t^{s,x}$. Hence (i) is verified.
To study the continuity in x in the sense of (ii), we take $x \in \mathbb{R}^d$ and a sequence $(x_j)_j \subset \mathbb{R}^d$ such that $x_j \to x$ as $j \to \infty$. Since $X_t^{s,x_j} = Z_t^{s,x_j}$ and $X_t^{s,x} = Z_t^{s,x}$ for $t \in [0, \tau_{n(s)})$, and $Z_s \in \mathcal{C}_0$,

$$\lim_{j\to\infty} \sup_{0\le t<\tau_{n(s)}} \left| X_t^{s,x_j} - X_t^{s,x} \right| = 0.$$

If $\tau_{n(s)} < T$, then by (3.82) and the continuity of f in the first argument we have $X_{\tau_{n(s)}}^{s,x_j} \to X_{\tau_{n(s)}}^{s,x}$ as $j \to \infty$, from which we deduce, by (3.83),

$$\lim_{j\to\infty} \sup_{t\in(\tau_{n(s)},\tau_{n(s)+1})} \left| X_t^{s,x_j} - X_t^{s,x} \right| = 0.$$

In general, using (3.84)-(3.85) we can work by induction to obtain (ii).
Finally we study the càdlàg property in the variable $s \in [0,T]$ according to (iii). Firstly, we analyze the right–continuity in $s \in [0,T)$. Fix $M > 0$ and take a sequence $(s_j)_j \subset (s,T)$ such that $s_j \to s$ as $j \to \infty$. We assume, without loss of generality, that $\tau_{n(s_j)} = \tau_{n(s)}$ for all j. Since Z is a $(\mathcal{C}_0, \mathcal{C})$–valued càdlàg process, by construction

$$\lim_{j\to\infty} \sup_{|x|\le M} \sup_{0\le t<\tau_{n(s)}} \left| X_t^{s_j,x} - X_t^{s,x} \right| = 0.$$

If $\tau_{n(s)} < T$, notice that the set $\left\{ Z_{\tau_{n(s)}}^{s_j,x}, Z_{\tau_{n(s)}}^{s,x}, \text{ with } |x| \le M,\, j \in \mathbb{N} \right\}$ is bounded. Considering that f is uniformly continuous in the first variable on compact sets, by (3.82) we deduce that

$$\begin{aligned}\sup_{|x|\le M} \left| X_{\tau_{n(s)}}^{s_j,x} - X_{\tau_{n(s)}}^{s,x} \right| &\le \sup_{|x|\le M} \left| Z_{\tau_{n(s)}}^{s_j,x} - Z_{\tau_{n(s)}}^{s,x} \right| \\ &\quad + \sup_{|x|\le M} \left| f\left(Z_{\tau_{n(s)}}^{s_j,x}, \tau_{n(s)}, p_{\tau_{n(s)}} \right) - f\left(Z_{\tau_{n(s)}}^{s,x}, \tau_{n(s)}, p_{\tau_{n(s)}} \right) \right| \underset{j\to\infty}{\longrightarrow} 0.\end{aligned}$$

As $x \mapsto Z_\cdot^{\tau_{n(s)},x}$ is a continuous function from $\mathbb{R}^d$ to $\mathcal{D}_0$, it is uniformly continuous on compact sets of $\mathbb{R}^d$. Moreover, the previous equation coupled with

$$\sup_{|x|\le M} \left| X_{\tau_{n(s)}}^{s_j,x} \right| < \infty,\ j \in \mathbb{N}, \qquad \sup_{|x|\le M} \left| X_{\tau_{n(s)}}^{s,x} \right| < \infty \quad \text{(by the continuity in } x\text{)}$$

ensures that the set $\left\{ X_{\tau_{n(s)}}^{s_j,x}, X_{\tau_{n(s)}}^{s,x}, \text{ with } |x| \le M,\, j \in \mathbb{N} \right\}$ is bounded. Combining these two facts, by (3.83)

$$\lim_{j\to\infty} \sup_{|x|\le M} \sup_{\tau_{n(s)}<t<\tau_{n(s)+1}} \left| X_t^{s_j,x} - X_t^{s,x} \right| = 0.$$

Using (3.84)-(3.85) we argue by induction to infer the right–continuity in s in the sense of (iii). Secondly, we can prove the left–continuity of $X_t^{\cdot,x}$ in $s \in (0,T]$ in a similar way, exploiting the left–continuity of the process $(Z_s)_s$. We only note that, in this case, it is possible that $s = \tau_n$ for some $n \in \mathbb{N}$. Hence given a sequence $(s_j)_j \subset (0,s)$ such that $s_j \to s$ as $j \to \infty$, we might have $\tau_{n(s_j)} = s < \tau_{n(s)}$ for j large enough. This, however, does not affect the existence of the left–limits because $\tau_{n(s_j)}$ are definitively all equal.

Step IV: The stochastic continuity of the flow $X_t^{s,x}$. To obtain the stochastic continuity in the sense of Point 4. in Definition 3.1, it is sufficient to prove that, for every $s \in (0,T]$, there exists an a.s. event Ω_s where

$$X_t^{s-,x} = X_t^{s,x}, \quad x \in \mathbb{R}^d,\, t \in [0,T]. \tag{3.87}$$

Indeed, combining this equality with the càdlàg property we have just proved, we deduce that

$$\begin{aligned} 0 = \lim_{r \uparrow s} \mathbb{P}\left(\sup_{|x|\le M} \sup_{0\le t \le T} \left| X_t^{r,x} - X_t^{s-,x} \right| > \epsilon \right) &= \lim_{r\uparrow s} \mathbb{P}\left(\sup_{|x|\le M} \sup_{0\le t\le T} |X_t^{r,x} - X_t^{s,x}| > \epsilon \right) \\ &= \lim_{r\downarrow s} \mathbb{P}\left(\sup_{|x|\le M} \sup_{0\le t\le T} |X_t^{r,x} - X_t^{s,x}| > \epsilon \right) = 0, \quad \epsilon,\, M > 0. \end{aligned}$$

The case $s = T$ is the easier one: by construction and (3.71) we have, in an a.s. event contained in Ω_3 and depending on T,

$$X_t^{T-,x} = \lim_{r\uparrow T} X_t^{r,x} = \lim_{r \uparrow T} Z_t^{r,x} = Z_t^{T-,x} = Z_t^{T,x} = x = X_t^{T,x}, \quad t \in [0,T),\, x \in \mathbb{R}^d;$$

the final time $t = T$ can be recovered by passing to the limit as $t \to T$, because $X_\cdot^{T-,x}$ and $X_\cdot^{T,x}$ are left–continuous in T. As for $s \in [0,T)$, we can argue as at the beginning of this proof to construct an a.s. event $\Omega_s \subset \Omega_3$ such that $\tau_n(s) \neq s$ for all $n \in \mathbb{N}$, and that (recall (3.71)) $Z_{s-} = Z_s$. Then, by (3.81),

$$X_t^{s-,x} = Z_t^{s-,x} = Z_t^{s,x} = X_t^{s,x}, \quad t \in [0, \tau_{n(s)}),\, x \in \mathbb{R}^d, \text{ in } \Omega_s :$$

if $\tau_{n(s)} < T$, this equality holds also for $t = \tau_{n(s)}$ by (3.82). Employing (3.83)-(3.84)-(3.85) and the left–continuity of $X_\cdot^{s-,x}$, $X_\cdot^{s,x}$ in T, we reason by induction to obtain (3.87).

Step V: The stochastic flow $X_t^{s,x}$ satisfies (3.77). Recall that $\tau_m^s = \tau_m \vee s$, $m \in \mathbb{N}_0$, where τ_m is given in (3.80). We now argue –using Lemma 3.18– that the process $X_t^{s,x}$ satisfies (3.76) with $\eta = x$ in $t \in [0,T)$. This is equivalent to showing that, for every $m \in \mathbb{N}_0$, there is an a.s. event $\Omega_{1,m}(s,x)$ such that, for all $t \in [\tau_m^s, \tau_{m+1}^s)$ (note that $[\tau_m^s, \tau_{m+1}^s)$ can also be empty),

$$\begin{aligned} X_t^{s,x} = x 1_{\{\tau_m^s \le s\}} + \left(X_{\tau_m^s -}^{s,x} + f\left(X_{\tau_m^s -}^{s,x}, \tau_m^s, p_{\tau_m^s} \right)\right) 1_{\{\tau_m^s > s\}} + \int_s^t 1_{\{r > \tau_m^s\}} b\left(r, X_r^{s,x}\right) dr \\ + \int_s^t 1_{\{r>\tau_m^s\}} \alpha\left(r, X_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} 1_{\{r > \tau_m^s\}} g\left(X_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right). \end{aligned} \tag{3.88}$$

Indeed, the stochastic integrals appearing in the previous expression can be read as differences involving truncated processes, see, for instance, [106, Section 3, Chapter II] and [110, Property 4.37, Chapter I]. More precisely, $\mathbb{P}-$a.s., for every $t \in [s,T)$,

$$\int_s^t 1_{\{r>\tau_m^s\}} \alpha\left(r, X_r^{s,x}\right) dW_r = \int_s^t \alpha\left(r, X_r^{s,x}\right) dW_r - \left(\int_s^\cdot \alpha\left(r, X_r^{s,x}\right) dW_r \right)_{\tau_m^s \wedge t}$$

and similarly

$$\int_s^t \int_{U_0} 1_{\{r>\tau_m^s\}} g\left(X_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) = \int_s^t \int_{U_0} g\left(X_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right) - \left(\int_s^\cdot \int_{U_0} g\left(X_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right)\right)_{\tau_m^s \wedge t}.$$

In view of the interlacing construction carried out above (cfr. (3.86)), in order to verify (3.88) we search for an a.s. event $\Omega_{1,m}(s,x) \subset \Omega_3$ such that, for all $t \in [\tau_m^s, \tau_{m+1}^s)$,

$$Z_t^{s,x} = x + \int_s^t 1_{\{r>\tau_m^s\}} b\left(r, Z_r^{s,x}\right) dr + \int_s^t 1_{\{r>\tau_m^s\}} \alpha\left(r, Z_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} 1_{\{r>\tau_m^s\}} g\left(Z_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right), \quad \text{in } \Omega_{1,m}(s,x) \cap \{\tau_m^s \le s\},$$

and

$$Z_t^{\tau_m^s, X_{\tau_m^s}^{s,x}} = X_{\tau_m^s}^{s,x} + \int_s^t 1_{\{r>\tau_m^s\}} b\left(r, Z_r^{\tau_m^s, X_{\tau_m^s}^{s,x}}\right) dr + \int_s^t 1_{\{r>\tau_m^s\}} \alpha\left(r, Z_r^{\tau_m^s, X_{\tau_m^s}^{s,x}}\right) dW_r + \int_s^t \int_{U_0} 1_{\{r>\tau_m^s\}} g\left(Z_{r-}^{\tau_m^s, X_{\tau_m^s}^{s,x}}, r, z\right) \widetilde{N}_p\left(dr, dz\right), \quad \text{in } \Omega_{1,m}(s,x) \cap \{\tau_m^s > s\}. \tag{3.89}$$

In $\{\tau_m^s \le s\}$, the former equation can be rewritten without the indicator functions, namely

$$Z_t^{s,x} = x + \int_s^t b\left(r, Z_r^{s,x}\right) dr + \int_s^t \alpha\left(r, Z_r^{s,x}\right) dW_r + \int_s^t \int_{U_0} g\left(Z_{r-}^{s,x}, r, z\right) \widetilde{N}_p\left(dr, dz\right),$$

which, by (3.67), holds in the whole $\Omega_3 \subset \Omega'$. Thus, we only focus on (3.89). Note that, in (3.89), we can insert 0 instead of s as lower bound for the integrals because we are working in $\{\tau_m^s > s\}$.

Consider a non–increasing sequence of simple random variables $(\tau_{m,n})_n$, with $\tau_{m,n} \le T$, such that $\tau_{m,n} \downarrow \tau_m^s$ as $n \to \infty$ in an a.s. event $\Omega_{2,m}(s) \subset \Omega_3$: (3.89) holds if we replace τ_m^s with $\tau_{m,n}$. More precisely, we write $\tau_{m,n} = \sum_{k=1}^{N_n} a_k^n 1_{A_k^n}$ for some $N_n \in \mathbb{N}$, $(a_k^n)_k \subset (-\infty, T]$ and some measurable partition $(A_k^n)_k \subset \mathcal{F}$ of Ω, $k = 1, \dots, N_n$. Then, by Lemma 3.18, there exists an a.s. event where, for every $t \in [0,T]$ and $n \in \mathbb{N}$,

$$\begin{aligned} Z_t^{\tau_{m,n}, X_{\tau_{m,n}}^{s,x}} &= \sum_{k=1}^{N_n} Z_t^{a_k^n, X_{a_k^n}^{s,x}} 1_{A_k^n} \\ &= \sum_{k=1}^{N_n} \left[X_{a_k^n}^{s,x} + \int_0^t 1_{\{r>a_k^n\}} b\left(r, Z_r^{a_k^n, X_{a_k^n}^{s,x}}\right) dr + \int_0^t 1_{\{r>a_k^n\}} \alpha\left(r, Z_r^{a_k^n, X_{a_k^n}^{s,x}}\right) dW_r \right. \\ &\qquad \left. + \int_0^t \int_{U_0} 1_{\{r>a_k^n\}} g\left(Z_{r-}^{a_k^n, X_{a_k^n}^{s,x}}, r, z\right) \widetilde{N}_p\left(dr, dz\right) \right] 1_{A_k^n}, \end{aligned}$$

Note that here we do not insert $1_{A_n^k}$ inside the stochastic integrals in order not to lose the adaptedness of the integrands. Invoking (3.73), the previous equation can be rewritten as follows:

$$Z_t^{\tau_{m,n}, X_{\tau_{m,n}}^{s,x}} = X_{\tau_{m,n}}^{s,x} + Z_{1,t}^{\tau_{m,n}, X_{\tau_{m,n}}^{s,x}} + Z_{2,t}^{\tau_{m,n}, X_{\tau_{m,n}}^{s,x}} + Z_{3,t}^{\tau_{m,n}, X_{\tau_{m,n}}^{s,x}}, \quad t \in [0,T],\ n \in \mathbb{N}, \tag{3.90}$$

which holds in an a.s. event $\Omega_{3,m}(s,x) \subset \Omega_3$. Since $X_t^{s,x}$ is a càdlàg function of t and Z is a $\mathcal{C}_0-$valued càdlàg process,

$$\lim_{n\to\infty} X^{s,x}_{\tau_{m,n}} = X^{s,x}_{\tau^s_m}, \qquad \lim_{n\to\infty} Z_t^{\tau_{m,n}, X^{s,x}_{\tau_{m,n}}} = Z_t^{\tau^s_m, X^{s,x}_{\tau^s_m}}, \quad \text{in } \Omega_{2,m}(s).$$

Thus, recalling that also the processes Z_1, Z_2 and Z_3 are càdlàg with values in the space $\mathcal{C}_0$, we can pass to the limit in (3.90) as $n \to \infty$ to deduce that, in the a.s. event $\Omega_{1,m}(s,x) = \Omega_{2,m}(s) \cap \Omega_{3,m}(s,x)$,

$$Z_t^{\tau^s_m, X^{s,x}_{\tau^s_m}} = X^{s,x}_{\tau^s_m} + Z_{1,t}^{\tau^s_m, X^{s,x}_{\tau^s_m}} + Z_{2,t}^{\tau^s_m, X^{s,x}_{\tau^s_m}} + Z_{3,t}^{\tau^s_m, X^{s,x}_{\tau^s_m}}, \quad t \in [0,T]. \tag{3.91}$$

Therefore (3.89) is satisfied in $\Omega_{1,m}(s,x)$ on the entire $[0,T]$, proving that $X^{s,x}$ is a solution to (3.88). Hence $X^{s,x}$ solves (3.76) with $\eta = x$.

It remains to find an a.s. event Ω'' –not depending on s, x and t– where (3.77) is satisfied. If we define

$$\Omega_{1,m} = \bigcap_{s\in[0,T)\cap\mathbb{Q}} \bigcap_{x\in\mathbb{Q}^d} \Omega_{1,m}(s,x), \quad m \in \mathbb{N}_0,$$

then (3.91) is simultaneously satisfied in $\Omega_{1,m}$ for every $s \in [0,T)\cap\mathbb{Q}$ and $x \in \mathbb{Q}^d$. In fact, the continuity of the flow $X_t^{s,x}$ in x implies that (3.91) holds for every $x \in \mathbb{R}^d$ in $\Omega_{1,m}$, with s being a rational number in $[0,T)$.

If $s \in [0,T)\cap(\mathbb{R}\setminus\mathbb{Q})$, then we just consider $(s_n)_n \subset (s,T)\cap\mathbb{Q}$ such that $s_n \downarrow s$ as $n \to \infty$, and another limiting argument based on the regularity of Z, Z_1, Z_2, Z_3 and $X_t^{s,x}$ shows that (3.91) holds for this choice of s, too. Summarizing, (3.91) holds in $\Omega_{1,m}$ for all $s \in [0,T)$ and $x \in \mathbb{R}^d$. Therefore the flow $X_t^{s,x}$ satisfies (3.77) in

$$\Omega'' = \cap_{m=0}^{\infty} \Omega_{1,m}.$$

Step VI: The flow property (3.78). Note that, for every $\bar{t} \in [0,T]$, the function $X\colon \Omega \times [0,T] \times \mathbb{R}^d \times [0,\bar{t}] \to \mathbb{R}^d$ defined by $X(\omega,s,x,t) = X_t^{s,x}(\omega)$ is $\mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}^d\times[0,\bar{t}])-$measurable by (3.86) and Lemma 3.17. As a consequence, for every $s \in [0,T)$, $\eta \in L^0(\mathcal{F}_s)$ and $t \in [s,T]$, the random variable $X_t^{s,\eta(\cdot)}(\cdot)$ is $\mathcal{F}_t-$measurable. Denote by $X^{s,\eta}$ the process defined by $X_t^{s,\eta}(\omega) = X_t^{s,\eta(\omega)}(\omega)$, $\omega \in \Omega$, $t \in [s,T]$; by the same arguments as those used to prove (3.88), with η instead of x, we deduce that $X^{s,\eta}$ solves (3.76) with initial condition (s,η). In particular, for every $0 \le s < u < t \le T$ and $x \in \mathbb{R}^d$, thanks to the pathwise uniqueness of (3.76) we infer the existence of an a.s. event $\Omega_{s,u,x}$ such that

$$X_t^{s,x}(\omega) = X_t^{u,X_u^{s,x}(\omega)}(\omega), \quad t \in (u,T], \omega \in \Omega_{s,u,x}.$$

Since $X_t^{s,x}$ satisfies Point 2. in Definition 3.1, we can proceed as in Corollary 3.5 and Lemma 3.16 to obtain (3.78), i.e., to establish the previous equation in an a.s. event not depending on s, u and x.

The proof is now complete. ■

3.4 Towards controlled SDEs

The results of Theorem 3.1 continue to hold if we consider more general coefficients depending on additional variables for the SDE (3.3), under requirements similar to Hypothesis 3.1. More precisely, we suppose that b, α, g and f introduced in Section 3.1 also depend on $\mathbf{y} \in \mathbb{R}^k$, for some $k \in \mathbb{N}$, and that they are jointly measurable in their domains. We require that $f\colon \mathbb{R}^d \times [0,T] \times U \times \mathbb{R}^k \to \mathbb{R}^d$ is continuous in the first and last arguments, and that b, α, g satisfy the linear growth and Lipschitz–type conditions in (3.1)-(3.2), uniformly in $\mathbf{y} \in \mathbb{R}^k$. Moreover, we assume that, for every $p \geq 2$, there exists a constant K_p such that

$$\begin{aligned}&|b(t,x,\mathbf{y}_1) - b(t,x,\mathbf{y}_2)|^p + |\alpha(t,x,\mathbf{y}_1) - \alpha(t,x,\mathbf{y}_2)|^p \\ &+ \int_{U_0} |g(x,t,z,\mathbf{y}_1) - g(x,t,z,\mathbf{y}_2)|^p \,\nu(dz) \leq K_p |\mathbf{y}_1 - \mathbf{y}_2|^p, \quad x \in \mathbb{R}^d,\ \mathbf{y}_1,\ \mathbf{y}_2 \in \mathbb{R}^k,\ t \in [0,T]. \end{aligned} \tag{3.92}$$

Given $s \in [0,T)$, $\eta \in L^0(\mathcal{F}_s)$ and $\mathbf{y} \in \mathbb{R}^k$, the corresponding SDE is the following:

$$\begin{aligned} X_t^{s,\eta,\mathbf{y}} = \eta &+ \int_s^t b\left(r, X_r^{s,\eta,\mathbf{y}}, \mathbf{y}\right) dr + \int_s^t \alpha\left(r, X_r^{s,\eta,\mathbf{y}}, \mathbf{y}\right) dW_r + \int_s^t \int_{U_0} g\left(X_{r-}^{s,\eta,\mathbf{y}}, r, z, \mathbf{y}\right) \widetilde{N}_p(dr,dz) \\ &+ \int_s^t \int_{U \setminus U_0} f\left(X_{r-}^{s,\eta,\mathbf{y}}, r, z, \mathbf{y}\right) N_p(dr,dz), \quad t \in [s,T]. \end{aligned} \tag{3.93}$$

Thanks to (3.92), the SDE (3.93) admits a pathwise unique solution.

The arguments used in Sections 3.2 and 3.3 can be easily adapted to encompass the additional variable $\mathbf{y}$. Here we provide some insights on how these changes can be implemented. We mainly focus on the small–jumps case $f \equiv 0$, discussing some important steps of Section 3.2 in the new framework.

By (3.92), an analogue of Lemma 3.3 holds: for every $s \in [0,T)$ and $p \geq 2$, for some $C > 0$,

$$\mathbb{E}\left[\sup_{s \leq t \leq T} \left| X_t^{s,\eta,\mathbf{y}_1} - X_t^{s,\xi,\mathbf{y}_2} \right|^p \right] \leq C\left(\mathbb{E}\left[|\eta - \xi|^p\right] + |\mathbf{y}_1 - \mathbf{y}_2|^p\right), \quad \eta, \xi \in L^p(\Omega) \cap L^0(\mathcal{F}_s),\ \mathbf{y}_1,\ \mathbf{y}_2 \in \mathbb{R}^k.$$

Hence, invoking the Kolmogorov–Chentsov criterion, there is a version of the solution $X_t^{s,x,\mathbf{y}}$ such that, for every $\omega \in \Omega$, the map $X^{s,\cdot,\cdot}(\omega)\colon \mathbb{R}^d \times \mathbb{R}^k \to (\mathcal{D}_0, \mathcal{D})$ is continuous. This allows to prove the flow property

$$X_t^{s,x,\mathbf{y}}(\omega) = X_t^{u, X_u^{s,x,\mathbf{y}}(\omega), \mathbf{y}}(\omega), \quad t \in [u,T],\ x \in \mathbb{R}^d,\ \mathbf{y} \in \mathbb{R}^k,\ \omega \in \Omega_{s,u},$$

for any $0 \leq s < u \leq T$.

As in Corollary 3.9 (see also [107]), given $p > 2(d+k)$, there exist random variables $U_{s,p}$, $s \in [0,T)$, with $\sup_{s \in [0,T)} \mathbb{E}\left[|U_{s,p}|^p\right] < \infty$, such that

$$\begin{aligned} &\sup_{0 \leq t \leq T} \left| X_t^{s,x,\mathbf{y}_1}(\omega) - X_t^{s,y,\mathbf{y}_2}(\omega) \right| \\ &\qquad \leq c\, U_{s,p}(\omega) \left(|x-y| + |\mathbf{y}_1 - \mathbf{y}_2|\right)^{1 - 2\frac{d+k}{p}} \left[\max\left\{|x| + |\mathbf{y}_1|, |y| + |\mathbf{y}_2|\right\}^{\frac{2(d+k)+1}{p}} \vee 1\right], \end{aligned} \tag{3.94}$$

where $x, y \in \mathbb{R}^d$, $\mathbf{y}_1$, $\mathbf{y}_2 \in \mathbb{R}^k$, $\omega \in \Omega$. Similar results hold for the continuous, $(\mathcal{D}_0, \mathcal{D})-$valued random fields $Z_1^{s,x,\mathbf{y}} = \int_s^{\cdot} \alpha\left(r, X_r^{s,x,\mathbf{y}}, \mathbf{y}\right)) dW_r$, $Z_2^{s,x,\mathbf{y}} = \int_s^{\cdot} \int_{U_0} g\left(X_{r-}^{s,x,\mathbf{y}}, r, z, \mathbf{y}\right) \widetilde{N}_p(dr,dz)$ and $Z_3^{s,x,\mathbf{y}} = \int_s^{\cdot} b\left(r, X_r^{s,x,\mathbf{y}}, \mathbf{y}\right) dr$.

Equation (3.94) enables to estimate fundamental quantities like (cfr. (I.9) in Introduction)

$$\mathbb{E}\left[\sup_{|x| \leq N} \sup_{|\mathbf{y}| \leq N} \sup_{s \leq t \leq T} \left| \int_s^t \int_{U_0} g\left(X_{r-}^{s,x,\mathbf{y}}, r, z, \mathbf{y}\right) \widetilde{N}_p(dr,dz) \right|^\gamma \right], \quad \gamma > 2(d+k),\ N > 0,$$

by writing

$$\mathbb{E}\left[\sup_{|x| \leq N} \sup_{|\mathbf{y}| \leq N} \sup_{s \leq t \leq T} \left| Z_{2,t}^{s,x,\mathbf{y}} \right|^\gamma \right] \leq 2^{\gamma-1} \left(\mathbb{E}\left[\sup_{s \leq t \leq T} \left| Z_{2,t}^{s,0,\mathbf{0}} \right|^\gamma \right] + \mathbb{E}\left[\sup_{|x| \leq N} \sup_{|\mathbf{y}| \leq N} \sup_{s \leq t \leq T} \left| Z_{2,t}^{s,x,\mathbf{y}} - Z_{2,t}^{s,0,\mathbf{0}} \right|^\gamma \right] \right),$$

and then proceeding as in, e.g., (3.33)-(3.35).

Consider now the space $\widetilde{\mathcal{C}}_0 = C(\mathbb{R}^d \times \mathbb{R}^k; \mathcal{D}_0)$ endowed with the $\sigma-$algebra $\widetilde{\mathcal{C}}$ generated by the projections $\pi_{x,\mathbf{y}} \colon C(\mathbb{R}^d \times \mathbb{R}^k; \mathcal{D}_0) \to (\mathcal{D}_0, \mathcal{D})$, $x \in \mathbb{R}^d$, $\mathbf{y} \in \mathbb{R}^k$, defined by $\pi_{x,\mathbf{y}}(f) = f(x, \mathbf{y})$, $f \in \widetilde{\mathcal{C}}_0$. The previous considerations allow to deduce the stochastic continuity [resp., the càdlàg property] of the $(\widetilde{\mathcal{C}}_0, \widetilde{\mathcal{C}})-$valued process $(X^s)_{s\in[0,T]}$ arguing as in Subsection 3.2.2 [resp., Subsection 3.2.3]. Therefore Theorem 3.2 holds when the coefficients b, α, g depend also on $\mathbf{y}$.

In the large–jumps case, the same observations, coupled with the arguments in Section 3.3, yield the validity of Theorem 3.1 for (3.93), as well.

As a consequence, the SDE (3.93) generates a *sharp stochastic flow* $X = X_t^{s,x,\mathbf{y}}$ in the following sense.

Definition 3.2. Let $X \colon \Omega \times [0,T] \times \mathbb{R}^d \times [0,T] \times \mathbb{R}^k \to \mathbb{R}^d$ be an $\mathcal{F} \otimes \mathcal{B}([0,T] \times \mathbb{R}^d \times [0,T] \times \mathbb{R}^k)-$measurable function and denote by $X_t^{s,x,\mathbf{y}}(\omega) = X(\omega, s, x, t, \mathbf{y})$. We say that X is the *sharp stochastic flow* generated by (3.93) if there exists an a.s. event Ω' –independent from $s, t, x, \mathbf{y}$– such that the four following requirements are fulfilled for every $\omega \in \Omega'$, $s \in [0,T)$, $x \in \mathbb{R}^d$ and $\mathbf{y} \in \mathbb{R}^k$.

1. The process $(X_t^{s,x,\mathbf{y}})_{t\in[s,T]}$ satisfies (3.93) in Ω';
2. (i) The map $X_\cdot^{s,x,\mathbf{y}}(\omega) : [0,T] \to \mathbb{R}^d$ is càdlàg;
   (ii) The map $X^{s,\cdot,\cdot}(\omega) : \mathbb{R}^d \times \mathbb{R}^k \to \mathcal{D}_0$ is continuous;
   (iii) The map $X^{\cdot,x,\mathbf{y}}(\omega) : [0,T] \to \mathcal{D}_0$ is càdlàg, locally uniformly in $x, \mathbf{y}$;
3. The flow property holds: $\quad X_t^{s,x,\mathbf{y}}(\omega) = X_t^{u, X_u^{s,x,\mathbf{y}}(\omega), \mathbf{y}}(\omega), \quad s < u < t \le T;$
4. The function X is stochastically continuous in the following sense: for every $\epsilon > 0$ and $M > 0$,
$$\lim_{r \to s} \mathbb{P}\Big(\sup_{|x|\le M} \sup_{|\mathbf{y}|\le M} \sup_{0\le t\le T} \big| X_t^{r,x,\mathbf{y}} - X_t^{s,x,\mathbf{y}} \big| > \epsilon \Big) = 0, \quad s \in [0,T].$$

3.5 The dynamic programming principle

The aim of this section is to establish a dynamic programming principle for controlled SDEs, employing the properties of the sharp stochastic flow constructed in the previous sections (see Theorem 3.1 and Section 3.4). According to [117, Chapter 4], in this section we suppose that the measurable space $(U, \mathcal{U})$ is in fact a Polish space, with $\mathcal{U}$ being the Borel $\sigma-$algebra.

Moreover, we will use extensively

$$\mathcal{F}_{t_1,t_2}^{W,N_p} = \sigma\left(\{W_t - W_{t_1},\, N_p\left((t_1,t] \times E\right),\, t \in [t_1,t_2],\, E \in \mathcal{U}\} \cup \mathcal{N}\right), \quad 0 \le t_1 \le t_2 \le T, \tag{3.95}$$

where $\mathcal{N}$ is the family of negligible events in $(\Omega, \mathcal{F}, \mathbb{P})$: we set $\mathbb{F}_{t_1}^{W,N_p} = (\mathcal{F}_{t_1,t}^{W,N_p})_{t\in[t_1,T]}$ (cfr. (3.54)). In particular, $\mathbb{F}^{W,N_p} = \mathbb{F}_0^{W,N_p} = (\mathcal{F}_{0,t}^{W,N_p})_{t\in[0,T]}$ is the augmented filtration generated by W and N_p.

3.5.1 Controlled SDEs

Fix $l \in \mathbb{N}$. Here we assume that the coefficients b, α, g and f introduced in Section 3.1 also depend on $\mathbf{y} \in \mathbb{R}^l$, and that they are jointly measurable in their domains. We require that $f \colon \mathbb{R}^d \times [0,T] \times U \times \mathbb{R}^l \to$

$\mathbb{R}^d$ is continuous in the first and last arguments, and that b, α, g satisfy the linear growth and Lipschitz–type conditions in (3.1)-(3.2), uniformly in $\mathbf{y} \in \mathbb{R}^l$. Moreover, we assume that, for every $p \geq 2$, condition (3.92) holds.

We are considering controlled SDEs of the form

$$\begin{aligned} X_t^{s,x,a} &= x + \int_s^t b\left(r, X_r^{s,x,a}, a_r\right) dr + \int_s^t \alpha\left(r, X_r^{s,x,a}, a_r\right) dW_r \\ &\quad + \int_s^t \int_{U_0} g\left(X_{r-}^{s,x,a}, r, z, a_r\right) \widetilde{N}_p\left(dr, dz\right) + \int_s^t \int_{U \setminus U_0} f\left(X_{r-}^{s,x,a}, r, z, a_r\right) N_p\left(dr, dz\right), \end{aligned} \tag{3.96}$$

where $a \in \mathcal{E}$, i.e., a is a *simple (or step) control.* More precisely, $\mathcal{E} = \cup_{\bar{n} \in \mathbb{N}} \mathcal{E}_{\bar{n}}$, where we denote by $\mathcal{E}_{\bar{n}}$ the space of $\mathbb{F}^{W,N_p}$−adapted square–integrable simple processes a of the form

$$a(t, \omega) = \sum_{i=0}^{\bar{n}-1} Z_i(\omega) 1_{(t_i, t_{i+1}]}(t), \quad t \in [0, T],\, \omega \in \Omega, \tag{3.97}$$

where $Z = (Z_i)_i \subset L^2(\Omega; \mathbb{R}^l)$, Z_i being $\mathcal{F}_{0,t_i}^{W,N_p}$−measurable. Here $\Pi_{\bar{n}} = (t_0, t_1, \ldots, t_{\bar{n}})$ is an $(\bar{n}+1)$−tuple, $0 = t_0 < t_1 < \cdots < t_{\bar{n}} = T$. Without loss of generality, we may assume that the processes in $\mathcal{E}$ starts from 0.

We denote by $X^{s,x,a}$ the pathwise unique strong solution of the controlled SDE (3.96); see also (3.55). *We now show that there exists a sharp stochastic flow generated by* (3.96)*.*

Consider $\bar{n} \in \mathbb{N}$, $\Pi_{\bar{n}}$ as before and a vector $\mathbf{y}(\bar{n}) = (y_i)_{i=0,\ldots,\bar{n}-1} \in (\mathbb{R}^l)^{\bar{n}}$. Given $s \in [0, T)$ and $\eta \in L^0(\mathcal{F}_s)$, we analyze the following SDE:

$$\begin{aligned} X_t^{s,\eta,\mathbf{y}(\bar{n})} &= \eta + \sum_{i=0}^{\bar{n}-1} \Bigg(\int_s^t 1_{\{t_i < r \leq t_{i+1}\}} b\left(r, X_r^{s,\eta,\mathbf{y}(\bar{n})}, y_i\right) dr + \int_s^t 1_{\{t_i < r \leq t_{i+1}\}} \alpha\left(r, X_r^{s,\eta,\mathbf{y}(\bar{n})}, y_i\right) dW_r \\ &\quad + \int_s^t \int_{U_0} 1_{\{t_i < r \leq t_{i+1}\}} g\left(X_{r-}^{s,\eta,\mathbf{y}(\bar{n})}, r, z, y_i\right) \widetilde{N}_p\left(dr, dz\right) \\ &\quad + \int_s^t \int_{U \setminus U_0} 1_{\{t_i < r \leq t_{i+1}\}} f\left(X_{r-}^{s,\eta,\mathbf{y}(\bar{n})}, r, z, y_i\right) N_p\left(dr, dz\right) \Bigg), \quad t \in [s, T]. \end{aligned} \tag{3.98}$$

If we define $\tilde{b}$: $[0, T] \times \mathbb{R}^d \times \left(\mathbb{R}^l\right)^{\bar{n}} \to \mathbb{R}^d$ by

$$\tilde{b}\left(t, x, \mathbf{z}\right) = \sum_{i=0}^{\bar{n}-1} 1_{\{t_i < t \leq t_{i+1}\}} b\left(t, x, z_i\right), \quad t \in [0, T],\, x \in \mathbb{R}^d,\, \mathbf{z} = (z_0, \ldots, z_{\bar{n}-1}) \in \left(\mathbb{R}^l\right)^{\bar{n}},$$

with analogous definitions for the coefficients $\tilde{\alpha}$, $\tilde{g}$ and $\tilde{f}$, then (3.98) can be rewritten as

$$\begin{aligned} X_t^{s,\eta,\mathbf{y}(\bar{n})} &= \eta + \int_s^t \tilde{b}\left(r, X_r^{s,\eta,\mathbf{y}(\bar{n})}, \mathbf{y}(\bar{n})\right) dr + \int_s^t \tilde{\alpha}\left(r, X_r^{s,\eta,\mathbf{y}(\bar{n})}, \mathbf{y}(\bar{n})\right) dW_r \\ &\quad + \int_s^t \int_{U_0} \tilde{g}\left(X_{r-}^{s,\eta,\mathbf{y}(\bar{n})}, r, z, \mathbf{y}(\bar{n})\right) \widetilde{N}_p\left(dr, dz\right) + \int_s^t \int_{U \setminus U_0} \tilde{f}\left(X_{r-}^{s,\eta,\mathbf{y}(\bar{n})}, r, z, \mathbf{y}(\bar{n})\right) N_p\left(dr, dz\right), \end{aligned}$$

for $t \in [s, T]$. We note that $\tilde{b}$, $\tilde{\alpha}$, $\tilde{g}$ and $\tilde{f}$ satisfy the requirements of Section 3.4 (see (3.92)), hence the SDE (3.98) generates a *sharp stochastic flow* $X^{\Pi_{\bar{n}}}$ in the sense of Definition 3.2.

Recalling (3.97), if we set $[\mathbf{y}(\bar{n})](\omega) = Z(\omega)$, then arguments similar to those in the proof of Lemma 3.18 yield

$$X_t^{s,x,a}(\omega) = X_t^{s,x,[\mathbf{y}(\bar{n})](\omega)}(\omega), \quad \omega \in \Omega. \tag{3.99}$$

Therefore there exists a version of $X_t^{s,x,a}$ which is a sharp stochastic flow in the sense of Definition 3.1. In particular, the stochastic continuity in Point 4. can be inferred using equalities like (3.71)-(3.87) (see also (3.112) below) and Point 2. (iii) in Definition 3.1.

Remark 3.5. *The existence of a sharp stochastic flow has been established for different classes of controlled SDEs without jumps. In fact, this is a useful tool to prove the DPP (see [146] and the references therein). In the special case of controlled jump–diffusions driven by a compound Poisson measure, the stochastic flow has been also used in the proof of [40, Proposition 5.4], where, however, the details about its validity are missing.*

In order to prove the dynamic programming principle (see, in particular, Step II in the proof of Theorem 3.27), we want to use a classical measurable selection theorem, namely [42, Theorem 2]. To this purpose, it is useful to introduce simple processes with jump times in a countable, dense subset of $[0, T]$.

Thus, for every $n \in \mathbb{N}$, we consider the set $S_n = \{t_i^{(n)},\ i = 0, \dots, 2^n\}$ of dyadic points of $[0, T]$ with mesh $T2^{-n}$, and call $S = \cup_{n\in\mathbb{N}} S_n$. For every $t \in [0, T]$, we define $t_n^- = \max\{s \in S_n : s \le t\}$ and $t_n^+ = \min\{s \in S_n : s > t\}$, with $T_n^+ = \infty$. We denote by $\mathcal{A}_n$ the subspace of simple processes $a \in \mathcal{E}$ such that

$$a(t, \omega) = \sum_{i=0}^{2^n-1} Z_i(\omega) 1_{\left(t_i^{(n)}, t_{i+1}^{(n)}\right]}(t), \quad t \in [0, T],\ \omega \in \Omega,$$

for some $Z = (Z_i)_i \subset L^2(\Omega; \mathbb{R}^l)$, where Z_i is $\mathcal{F}^{W,N_p}_{0,t_i^{(n)}}$−measurable. Therefore every simple process $a \in \mathcal{A}_n$ is identified by Z: we are going to write $a \sim Z$. We denote by

$$\mathcal{Z}_n = \left\{(Z_i)_{i=0,\dots,2^n-1} \subset L^2(\Omega; \mathbb{R}^l) \text{ such that } Z_i \text{ is } \mathcal{F}^{W,N_p}_{0,t_i^{(n)}}\text{−measurable}\right\}.$$

In the sequel, we set

$$\mathcal{A} = \cup_{n\in\mathbb{N}} \mathcal{A}_n \quad \text{and} \quad \mathcal{Z} = \cup_{n\in\mathbb{N}} \mathcal{Z}_n. \tag{3.100}$$

The following lemma shows that it is possible to approximate, in probability, the solution $X_t^{s,x,a}$, $a \in \mathcal{E}$, of (3.96) with solutions of the same equation corresponding to controls in $\mathcal{A}$.

Lemma 3.20. *For every $s \in [0, T)$, $x \in \mathbb{R}^d$ and $a \in \mathcal{E}$, there exists a sequence $(a_n)_n \subset \mathcal{A}$ such that $a_n \to a$ in $L^2([0, T] \times \Omega)$ and*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{0\le t\le T} |X_t^{s,x,a} - X_t^{s,x,a_n}| > \epsilon\right) = 0, \quad \epsilon > 0.$$

Proof. Fix $s \in [0, T)$, $x \in \mathbb{R}^d$, $\bar{n} \in \mathbb{N}$ and $a \in \mathcal{E}_{\bar{n}}$. By definition, there exists an $(\bar{n}+1)$−tuple $\bar{t} = (t_0, \dots, t_{\bar{n}})$, with $0 = t_0 < t_1 < \dots < t_{\bar{n}} = T$, and $\mathcal{F}^{W,N_p}_{0,t_i}$−measurable, $\mathbb{R}^l$−valued random variables Z_i such that

$$a_{t_i} = a(t, \cdot) = Z_i, \quad t \in (t_i, t_{i+1}],\ i = 0, \dots, \bar{n} - 1.$$

We consider controls $a' \in \mathcal{A}$ depending on the random variables a_{t_i}, $i = 0, \ldots, \bar{n} - 1$, of the form

$$a'(t, \omega) = \sum_{i=0}^{\bar{n}-1} a_{t_i}(\omega) 1_{(q_i, q_{i+1}]}(t), \quad t \in [0, T],\, \omega \in \Omega,$$

where $\bar{q} = (q_0, q_1, \ldots, q_{\bar{n}})$ is an $(\bar{n}+1)-$tuple of dyadic points $0 = q_0 < q_1 < \cdots < q_{\bar{n}} = T$ such that $t_i < q_i < t_{i+1}$, for any $i = 1, \ldots, \bar{n} - 1$. Clearly, a' depends also on $\bar{q}$, and

$$\mathbb{E} \int_0^T \left| a_r - a'_r \right|^2 dr = \sum_{i=1}^{\bar{n}-1} \int_{t_i}^{q_i} \mathbb{E} \left| a_{t_i} - a_{t_{i-1}} \right|^2 dr = \sum_{i=1}^{\bar{n}-1} (q_i - t_i) \mathbb{E} \left| a_{t_i} - a_{t_{i-1}} \right|^2. \tag{3.101}$$

Consequently, $\lim_{\bar{q} \to \bar{t}} \mathbb{E} \int_0^T |a_r - a'_r|^2 dr = 0$. We prove that

$$\lim_{\bar{q} \to \bar{t}} \mathbb{P} \left(\sup_{0 \le t \le T} \left| X_t^{s,x,a} - X_t^{s,x,a'} \right| > \epsilon \right) = 0, \quad \epsilon > 0, \tag{3.102}$$

i.e., we prove the convergence in probability of $X_t^{s,x,a'}$ to $X_t^{s,x,a}$ as $\bar{q} \to \bar{t}$, uniformly on $t \in [0, T]$.

In order to do this, we use an induction argument. By the pathwise uniqueness of (3.96), the assertion holds when $t \in [0, t_1]$. We then suppose that it is true on $[0, t_k]$, $k = 1, \ldots, \bar{n} - 1$, and we aim to prove that it holds on $[0, t_{k+1}]$, as well. To this purpose, it is enough to check that

$$\lim_{\bar{q} \to \bar{t}} \mathbb{P} \Big(\sup_{t_k \le t \le t_{k+1}} \left| X_t^{s,x,a} - X_t^{s,x,a'} \right| > \epsilon \Big) = 0, \quad \epsilon > 0. \tag{3.103}$$

First we work on the interval $[t_k, q_k]$, namely we show that

$$\lim_{\bar{q} \to \bar{t}} \mathbb{P} \Big(\sup_{t_k \le t \le q_k} \left| X_t^{s,x,a} - X_t^{s,x,a'} \right| > \epsilon \Big) = 0, \quad \epsilon > 0. \tag{3.104}$$

We introduce the control $\tilde{a} \in \mathcal{E}$ given by

$$\tilde{a}(t, \omega) = \sum_{i=0,\, i \neq k}^{\bar{n}-1} a_{t_i}(\omega) 1_{(t_i, t_{i+1}]}(t) + a_{t_{k-1}}(\omega) 1_{(t_k, t_{k+1}]}(t), \quad t \in [0, T],\, \omega \in \Omega,$$

and note that, for every $t \in [t_k, q_k]$, by the flow property in Point 3. of Definition 3.1 we have

$$X_t^{s,x,a} = X_t^{t_k, X_{t_k}^{s,x,a}, a}, \qquad X_t^{s,x,a'} = X_t^{t_k, X_{t_k}^{s,x,a'}, a'}, \quad \mathbb{P} - \text{a.s.}$$

Using this equality we can write, $\mathbb{P}-$a.s., for every $t \in [t_k, q_k]$,

$$X_t^{s,x,a} - X_t^{s,x,a'} = \left[X_t^{t_k, X_{t_k}^{s,x,a}, a} - X_t^{t_k, X_{t_k}^{s,x,a}, \tilde{a}} \right] + \left[X_t^{t_k, X_{t_k}^{s,x,a}, \tilde{a}} - X_t^{t_k, X_{t_k}^{s,x,a'}, a'} \right] =: \mathbf{I}(t) + \mathbf{II}(t).$$

Noticing that $a'_t = \tilde{a}_t$ when $t \in [t_k, q_k]$, by (3.96)-(3.99) we have

$$X_t^{t_k, X_{t_k}^{s,x,a'}, a'} = X_t^{t_k, X_{t_k}^{s,x,a'}, \tilde{a}}, \quad t \in [t_k, q_k], \quad \mathbb{P} - \text{a.s.}$$

Since, by the inductive hypothesis, $X_{t_k}^{s,x,a'} \to X_{t_k}^{s,x,a}$ in probability as $\bar q \to \bar t$, using the properties of our stochastic flow in Definition 3.1 we infer that

$$\lim_{\bar q \to \bar t} \mathbb{P}\left(\sup_{t_k \le t \le q_k} |\mathbf{II}(t)| > \epsilon \right) = 0, \quad \epsilon > 0.$$

Hence, denoting by $\eta = X_{t_k}^{s,x,a}$, to obtain (3.104) we just need to show that

$$\lim_{\bar q \to \bar t} \mathbb{P}\left(\sup_{t_k \le t \le q_k} |\mathbf{I}(t)| > \epsilon \right) = \lim_{\bar q \to \bar t} \mathbb{P}\Big(\sup_{t_k \le t \le q_k} \Big| X_t^{t_k,\eta,a} - X_t^{t_k,\eta,\tilde a} \Big| > \epsilon \Big) = 0, \quad \epsilon > 0. \tag{3.105}$$

In particular, we prove convergence results similar to (3.105) for each of the following terms:

$$\int_{t_k}^t b\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) dr, \qquad \int_{t_k}^t \int_{U_0} g(X_{r-}^{t_k,\eta,a}, r, z, a_{t_k}) \widetilde N_p\,(dr,dz)\,,$$
$$\int_{t_k}^t \alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) dW_r, \qquad \int_{t_k}^t \int_{U\setminus U_0} f(X_{r-}^{t_k,\eta,a}, r, z, a_{t_k}) N_p\,(dr,dz)\,.$$

We only consider the stochastic integrals with respect to dW_r and N_p: the others can be treated in a similar way. As for the integral in dW_r, we define, for every $\delta > 0$,

$$\sigma_{\bar q}(\delta) = \inf\left\{ t \in [0,T] : \int_0^t 1_{(t_k,q_k]}(r) \left| \alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right|^2 dr \ge \delta \right\},$$

with $\inf \emptyset = \infty$. This is an $\mathcal{F}_{0,t}^{W,N_p}$–stopping time. By the dominated convergence theorem, using the Lipschitz continuity of $\alpha(r,\cdot,\cdot)$, we obtain

$$\lim_{\bar q \to \bar t} \int_0^T 1_{(t_k,q_k]}(r) \left| \alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right|^2 dr = 0, \quad \mathbb{P}-\text{a.s.}$$

It follows that $\sigma_{\bar q}(\delta) \to \infty$ as $\bar q \to \bar t$. In particular, $\lim_{\bar q \to \bar t} \mathbb{P}\left(\sigma_{\bar q}(\delta) \le T\right) = 0$. Hence, by Markov's inequality, for some $c > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t_k \le t \le q_k} \left| \int_{t_k}^t \left[\alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right] dW_r \right| > \epsilon \right) \\
&= \mathbb{P}\left(\sup_{0 \le t \le T} \left| \int_0^t 1_{(t_k,q_k]}(r) \left[\alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right] dW_r \right| > \epsilon \right) \\
&\le \mathbb{P}\left(\sup_{0 \le t \le T} \left| \int_0^t 1_{(t_k,q_k]}(r) \left[\alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right] dW_r \right| > \epsilon,\ \sigma_{\bar q}(\delta) \ge T \right) + \mathbb{P}\left(\sigma_{\bar q}(\delta) \le T\right) \\
&\le \frac{c}{\epsilon^2} \mathbb{E}\left[\int_0^{\sigma_{\bar q}(\delta)} 1_{(t_k,q_k]}(r) \left| \alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right|^2 dr \right] + \mathbb{P}\left(\sigma_{\bar q}(\delta) \le T\right) \\
&\le \frac{c\delta}{\epsilon^2} + \mathbb{P}\left(\sigma_{\bar q}(\delta) \le T\right),
\end{aligned}$$

for every $\epsilon > 0$, which shows that

$$\lim_{\bar q \to \bar t} \mathbb{P}\left(\sup_{t_k \le t \le q_k} \left| \int_{t_k}^t \left[\alpha\left(r, X_r^{t_k,\eta,a}, a_{t_k}\right) - \alpha\left(r, X_r^{t_k,\eta,\tilde a}, a_{t_{k-1}}\right)\right] dW_r \right| > \epsilon \right) = 0, \quad \epsilon > 0.$$

As for the stochastic integral with respect to N_p, note that, for every $\omega \in \Omega$,

$$\sup_{t_k \le t \le q_k} \left| \int_{t_k}^{t} \int_{U \setminus U_0} \Big(f(X_{r-}^{t_k,\eta,a}, r, z, a_{t_k}) - f(X_{r-}^{t_k,\eta,\tilde{a}}, r, z, a_{t_{k-1}}) \Big) N_p(dr, dz) \right| (\omega)$$
$$\le \sum_{r \in D_p(\omega) \cap (t_k, q_k]} 1_{U \setminus U_0} (p_r(\omega)) \left| f\Big(X_{r-}^{t_k,\eta,a}(\omega), r, p_r(\omega), a_{t_k}(\omega)\Big) - f\Big(X_{r-}^{t_k,\eta,\tilde{a}}(\omega), r, p_r(\omega), a_{t_{k-1}}(\omega)\Big) \right|.$$

This is a random sum which is finite, $\mathbb{P}-$a.s. Thus, it is straightforward to prove that it converges, $\mathbb{P}-$a.s., to 0 as $\bar{q} \to \bar{t}$. The proof of (3.105) is complete, so (3.104) holds.

It remains to consider $t \in [q_k, t_{k+1}]$. In this case, by the flow property in Point 3. of Definition 3.1 we have

$$X_t^{s,x,a} = X_t^{q_k, X_{q_k}^{s,x,a}, a}, \qquad X_t^{s,x,a'} = X_t^{q_k, X_{q_k}^{s,x,a'}, a'}, \quad t \in [q_k, t_{k+1}], \, \mathbb{P} - \text{a.s.}$$

Since $a'_t = a_t$ when $t \in (q_k, t_{k+1}]$, by (3.96)-(3.99) we have

$$X_t^{q_k, X_{q_k}^{s,x,a'}, a'} = X_t^{q_k, X_{q_k}^{s,x,a'}, a}, \quad t \in [q_k, t_{k+1}], \quad \mathbb{P} - \text{a.s.}$$

Note that $q_k \downarrow t_k$ as $\bar{q} \to \bar{t}$. By (3.104) and the càdlàg property of the process $X_{\cdot}^{s,x,a}$, we know that $X_{q_k}^{s,x,a} \to X_{t_k}^{s,x,a}$ and $X_{q_k}^{s,x,a'} \to X_{t_k}^{s,x,a}$ in probability as $\bar{q} \to \bar{t}$. Therefore, by the regularity of the sharp stochastic flow $X_t^{s,x,a}$ in Point 2. of Definition 3.1, we conclude that

$$\mathbb{P}\left(\sup_{q_k \le t \le t_{k+1}} \left| X_t^{s,x,a} - X_t^{s,x,a'} \right| > \epsilon \right) \le \mathbb{P}\left(\sup_{0 \le t \le T} \left| X_t^{q_k, X_{q_k}^{s,x,a}, a} - X_t^{q_k, X_{q_k}^{s,x,a'}, a} \right| > \epsilon \right) \underset{\bar{q} \to \bar{t}}{\longrightarrow} 0, \quad \epsilon > 0.$$

This equation and (3.104) yield (3.103), hence (3.102) holds and the proof is complete. ■

3.5.2 On $\mathcal{F}_{0,t}^{W,N_p}-$measurable random variables

Let Ω_W be the space of continuous functions from $[0,T]$ to $\mathbb{R}^m$ and, for every $t \in [0,T]$, let $\mathcal{B}_t^W$ be the smallest $\sigma-$algebra on Ω_W which makes the projections $\pi_s \colon \Omega_W \to \mathbb{R}^m$, $s \in [0,t]$, measurable. Notice that the Brownian motion $W \colon (\Omega, \mathcal{F}_{0,t}^{W,N_p}) \to (\Omega_W, \mathcal{B}_t^W)$ is a measurable map. We denote by Ω_N the set of integer–valued measures defined on $E = (0,T] \times U$ with values in $\mathbb{N} \cup \{\infty\}$. As in [106], we endow Ω_N with the minimal $\sigma-$algebra $\mathcal{G}$ which makes measurable all the mappings: $\mu \mapsto \mu(A)$, with $A \in \mathcal{B}((0,T]) \otimes \mathcal{U}$. The Poisson random measure N_p discussed in the previous sections is a measurable map from Ω into Ω_N. For every $t \in [0,T]$, we consider the minimal $\sigma-$algebra $\mathcal{B}_t^N$ on Ω_N which makes measurable all the mappings

$$\mu \mapsto \mu((0,s] \times A), \quad s \le t, \;\; A \in \mathcal{U};$$

notice that $N_p \colon (\Omega, \mathcal{F}_{0,t}^{W,N_p}) \to (\Omega_N, \mathcal{B}_t^N)$ is measurable.

Inspired by the proof of [170, Lemma 3.11], we now clarify how to rewrite a random variable $Y : \big(\Omega, \mathcal{F}_{0,t}^{W,N_p}\big) \to \mathbb{R}^l$ in terms of W and N_p, for all $t \in [0,T]$. We introduce the measurable function $\widetilde{T}_t \colon \big(\Omega, \mathcal{F}_{0,t}^{W,N_p}\big) \to \big(\Omega_W \times \Omega_N, \mathcal{B}_t^W \otimes \mathcal{B}_t^N\big)$ defined by

$$\widetilde{T}_t(\omega) = (W(\omega), N_p(\omega)), \quad \omega \in \Omega.$$

In the sequel, we write $\left(\Omega_\times, \mathcal{B}_t^\times\right)$ instead of $\left(\Omega_W \times \Omega_N, \mathcal{B}_t^W \otimes \mathcal{B}_t^N\right)$ to keep the notation short. Denoting by $\mathbb{F}^{W,N_p,0} = (\mathcal{F}_t^{W,N_p,0})_{t\in[0,T]}$ the natural filtration generated by W and N_p, namely

$$\mathcal{F}_t^{W,N_p,0} = \sigma\left(\{W_r,\, N_p\left((0,r]\times E\right),\, r\in[0,t]\,,\, E\in\mathcal{U}\}\right), \quad t\in[0,T],$$

we observe that

$$\sigma\big(\widetilde{T}_t\big) = \mathcal{F}_t^{W,N_p,0}. \tag{3.106}$$

As a consequence, [177, Theorem 1.7] yields the existence of a measurable mapping

$$F:\left(\Omega_\times, \mathcal{B}_t^\times\right) \to \mathbb{R}^l \quad \text{such that } \; Y(\omega) = F\left(\widetilde{T}_t(\omega)\right), \text{ for } \mathbb{P}\text{−a.s. } \omega\in\Omega.$$

Consider an $\mathbb{F}^{W,N_p}$−stopping time θ with values in $[0,T]$ and denote by $\theta_t = \theta\wedge t$, for every $t\in[0,T]$. Note that, given $(w,\mu)\in\Omega_\times$, one can write, for every $\omega\in\Omega$,

$$w = w(\theta_t(\omega)\wedge\cdot) + \left[w(\theta_t(\omega)\vee\cdot) - w(\theta_t(\omega))\right], \quad \mu = \mu\left(\cdot\cap\left((0,\theta_t(\omega)]\times U\right)\right) + \mu\left(\cdot\cap\left((\theta_t(\omega),T]\times U\right)\right).$$

It follows that, for every $\omega\in\Omega$,

$$\begin{aligned} F\left(\widetilde{T}_t(\omega)\right) = F\left(W(\omega), N_p(\omega)\right) = F\big(&W_{\theta_t(\omega)\wedge\cdot}(\omega) + \left[W_{\theta_t(\omega)\vee\cdot} - W_{\theta_t(\omega)}\right](\omega), \\ &N_p\left(\cdot\cap\left((0,\theta_t(\omega)]\times U\right)\right)(\omega) + N_p\left(\cdot\cap\left((\theta_t(\omega),T]\times U\right)\right)(\omega)\big). \end{aligned}$$

We observe that the random variable from $(\Omega, \mathcal{F}_t^{W,N_p,0})$ to $\left(\Omega_\times, \mathcal{B}_t^\times\right)$ defined by

$$\omega \mapsto \left(W_{\theta_t(\omega)\wedge\cdot}(\omega), N_p\left(\cdot\cap\left((0,\theta_t(\omega)]\times U\right)\right)(\omega)\right)$$

is measurable with respect to $\mathcal{F}_{\theta_t}^{W,N_p}$, the σ−algebra generated by the stopping time θ_t relative to the filtration $\mathbb{F}^{W,N_p}$. On the other hand, the $\left(\Omega_\times, \mathcal{B}_t^\times\right)$ −valued random variable

$$\omega \mapsto \left(\left[W_{\theta_t(\omega)\vee\cdot} - W_{\theta_t(\omega)}\right](\omega), N_p\left(\cdot\cap\left((\theta_t(\omega),T]\times U\right)\right)(\omega)\right)$$

is independent from $\mathcal{F}_{\theta_t}^{W,N_p}$. Therefore, if Y is integrable, we can compute its conditional expectation with respect to $\mathcal{F}_{\theta_t}^{W,N_p}$ to deduce that, for $\mathbb{P}$−a.s. $\omega\in\Omega$,

$$\begin{aligned} \mathbb{E}\Big[Y\big|\mathcal{F}_{\theta_t}^{W,N_p}\Big](\omega) &= \mathbb{E}\left[F\left(\widetilde{T}_t\right)|\mathcal{F}_{\theta_t}^{W,N_p}\right](\omega) \\ &= \mathbb{E}\left[F\left(W_{\theta_t(\omega)\wedge\cdot}(\omega) + [W_{\theta_t\vee\cdot} - W_{\theta_t}], N_p\left(\cdot\cap\left((0,\theta_t(\omega)]\times U\right)\right)(\omega) + N_p\left(\cdot\cap\left((\theta_t,T]\times U\right)\right)\right)\right]. \end{aligned}$$

In particular, given $\mathcal{A}_n \ni a \sim Z \in \mathcal{Z}_n$, for every $i = 0,\ldots,2^n-1$, $n\in\mathbb{N}$, there exists a measurable function

$$F_i^{(n)}:\left(\Omega_\times, \mathcal{B}_{t_i^{(n)}}^\times\right) \to \mathbb{R}^l \quad \text{such that } \; Z_i(\omega) = F_i^{(n)}\left(\widetilde{T}_{t_i^{(n)}}(\omega)\right), \text{ for } \mathbb{P}\text{−a.s. } \omega\in\Omega.$$

Moreover, for every $\mathbb{F}^{W,N_p}$−stopping time θ we have, $\mathbb{P}$−a.s.,

$$\begin{aligned} \mathbb{E}\left[Z_i|\mathcal{F}_{\theta_s}^{W,N_p}\right](\omega) = \mathbb{E}\Big[F_i^{(n)}\big(&W_{\theta_s(\omega)\wedge\cdot}(\omega) + \left[W_{\theta_s\vee\cdot} - W_{\theta_s}\right], \\ &N_p\left(\cdot\cap\left((0,\theta_s(\omega)]\times U\right)\right)(\omega) + N_p\left(\cdot\cap\left((\theta_s,T]\times U\right)\right)\big)\Big], \quad s\in\left[0, t_i^{(n)}\right]. \end{aligned}$$

Given $\bar{\omega} \in \Omega$ and $s \in [0,T]$, we denote by $a^{\theta,\bar{\omega}}_{s,n}$ the following $\mathbb{F}^{W,N_p}-$adapted simple process obtained from a:

$$a^{\theta,\bar{\omega}}_{s,n}(t,\omega) = 1_{\{t \le s_n^+\}} a(t,\bar{\omega}) + 1_{\{t > s_n^+\}} \sum_{i=0}^{2^n-1} F_i^{(n)} \left(W_{\theta_s(\bar{\omega})\wedge \cdot}(\bar{\omega}) + \left[W_{\theta_s(\omega)\vee \cdot} - W_{\theta_s(\omega)} \right](\omega), \right.$$
$$\left. N_p\left(\cdot \cap ((0,\theta_s(\bar{\omega})] \times U)\right)(\bar{\omega}) + N_p\left(\cdot \cap ((\theta_s(\omega),T] \times U)\right)(\omega) \right) 1_{\left(t_i^{(n)},t_{i+1}^{(n)}\right]}(t), \quad (3.107)$$

where $t \in [0,T]$, $\omega \in \Omega$. We conclude this subsection with a result which will be useful for the proof of the DPP (see Subsection 3.5.3).

Lemma 3.21. *For any $0 \le s \le t \le T$, $l \in \mathbb{N}$, the Hilbert space*

$$H = L^2(\Omega, \mathcal{F}^{W,N_p}_{s,t}; \mathbb{R}^l) \ \text{ is separable.} \quad (3.108)$$

Proof. The claim is trivial for $s = t$, so we consider $0 \le s < t \le T$. Let $\Omega^{s,t}_W$ be the space of continuous functions from $[s,t]$ to $\mathbb{R}^m$ and, for every $r \in [s,t]$, let $\mathcal{B}^W_{s,t}$ be the smallest $\sigma-$algebra on $\Omega^{s,t}_W$ which makes the projections $\pi_r - \pi_s \colon \Omega^{s,t}_W \to \mathbb{R}^m$, $r \in [s,t]$, measurable.
Similarly, we denote by $\Omega^{s,t}_N$ the set of integer–valued measures defined on $(s,t] \times U$ with values in $\mathbb{N} \cup \{\infty\}$. We endow $\Omega^{s,t}_N$ with the minimal $\sigma-$algebra $\mathcal{B}^N_{s,t}$ which makes measurable all the mappings

$$\mu \mapsto \mu((s,r] \times A), \quad s < r \le t, \ \ A \in \mathcal{U}.$$

We write $\left(\Omega^{s,t}_\times, \mathcal{B}^\times_{s,t}\right)$ instead of $\left(\Omega^{s,t}_W \times \Omega^{s,t}_N, \mathcal{B}^W_{s,t} \otimes \mathcal{B}^N_{s,t}\right)$ to keep the notation short.

Note that H is isomorphic to $K = L^2\left(\Omega^{s,t}_\times, \mathcal{B}^\times_{s,t}, \mathbb{Q}^{s,t}; \mathbb{R}^l\right)$, where $\mathbb{Q}^{s,t}$ is the image law of $\mathbb{P}$ under the random variable

$$\omega \mapsto \left(W_\cdot(\omega) - W_s(\omega), N_p\left((s,\cdot] \times \cdot\right)(\omega)\right).$$

We know from [117, Theorem 4.2] that $(\Omega^{s,t}_\times, \mathcal{B}^\times_{s,t})$ is metrizable and that it can be considered as a Polish space (this fact has been also remarked in [15, Section 2]). It is not difficult to prove that $L^2(E, \mathcal{B}, \mu; \mathbb{R}^l)$ is separable when E is a Polish space and μ is a probability measure defined on the $\sigma-$algebra of Borel sets $\mathcal{B}$. This shows (3.108). ∎

3.5.3 Statement and proof of DPP

Given a *jointly measurable map* $h \colon [0,T] \times \mathbb{R}^d \times \mathbb{R}^l \to \mathbb{R}$, *with h continuous in the second and third arguments, and a continuous map* $j \colon \mathbb{R}^d \to \mathbb{R}$, we define the gain function

$$J(s,x,a) = \mathbb{E}\left[\int_s^T h\left(r, X^{s,x,a}_r, a_r\right) dr + j\left(X^{s,x,a}_T\right) \right], \quad s \in [0,T),\, x \in \mathbb{R}^d,\, a \in \mathcal{E}. \quad (3.109)$$

We also require *h and j to be bounded in their domains,* so that J is well defined.

The value function v associated with J is

$$v(s,x) = \sup_{a \in \mathcal{E}} J(s,x,a), \quad s \in [0,T),\, x \in \mathbb{R}^d.$$

The next result shows that it is not restrictive to consider only controls $a \in \mathcal{A}$, i.e., simple processes with jump times in S, when defining v.

Lemma 3.22. *The following equality holds for every $s \in [0,T)$ and $x \in \mathbb{R}^d$:*

$$v(s,x) = \sup_{a\in\mathcal{A}} J(s,x,a). \tag{3.110}$$

Proof. Fix $s \in [0,T)$ and $x \in \mathbb{R}^d$. Since $\mathcal{A} \subset \mathcal{E}$, we only focus on $v(s,x) \le \sup_{a\in\mathcal{A}} J(s,x,a)$. According to Lemma 3.20, for every $a \in \mathcal{E}$ there exists a sequence $(a_n)_n \subset \mathcal{A}$ such that $a_n \to a$ in $L^2\big([0,T]\times\Omega\big)$ and that

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{0\le t\le T} |X_t^{s,x,a_n} - X_t^{s,x,a}| > \epsilon\right) = 0, \quad \epsilon > 0. \tag{3.111}$$

We now prove that there exists a subsequence $(a_{n_k})_k$ such that

$$\lim_{k\to\infty} J(s,x,a_{n_k}) = J(s,x,a).$$

This easily implies the assertion in (3.110). Recall that

$$J(s,x,a_n) = \mathbb{E}\left[\int_s^T h\left(r, X_r^{s,x,a_n}, a_n(r)\right) dr + j\left(X_T^{s,x,a_n}\right)\right], \quad n \in \mathbb{N}.$$

By (3.111) and Vitali's convergence theorem, $\mathbb{E}\left[j\left(X_T^{s,x,a_n}\right)\right] \to \mathbb{E}\left[j\left(X_T^{s,x,a}\right)\right]$ as $n\to\infty$. Regarding the other addend, Fubini's theorem yields

$$\mathbb{E}\left[\int_s^T \left(h\left(r, X_r^{s,x,a_n}, a_n(r)\right) - h\left(r, X_r^{s,x,a}, a(r)\right)\right) dr\right] \\ = \int_s^T \mathbb{E}\left[h\left(r, X_r^{s,x,a_n}, a_n(r)\right) - h\left(r, X_r^{s,x,a}, a(r)\right)\right] dr.$$

Using (3.111) and the fact that there exists a subsequence a_{n_k} such that $a_{n_k}(r) \to a(r)$ in probability as $k\to\infty$, for a.e. $r \in [s,T]$, by Vitali's convergence theorem we infer that

$$\lim_{k\to\infty} \mathbb{E}\left[h\left(r, X_r^{s,x,a_{n_k}}, a_{n_k}(r)\right) - h\left(r, X_r^{s,x,a}, a(r)\right)\right] = 0, \quad \text{for a.e. } r \in [s,T],$$

which in turn implies, by dominated convergence,

$$\lim_{k\to\infty} \int_s^T \mathbb{E}\left[h\left(r, X_r^{s,x,a_{n_k}}, a_{n_k}(r)\right) - h\left(r, X_r^{s,x,a}, a(r)\right)\right] dr = 0.$$

This computation finishes the proof. ∎

Remark 3.6. *By Lemma 3.22, we can consider the value function v as in* (3.110)*, i.e., v is computed taking the supremum over controls $a \in \mathcal{A}$. Moreover, using the identification $\mathcal{A}_n \ni a \sim Z \in \mathcal{Z}_n$, $n \in \mathbb{N}$, in the sequel we might also write $J(s,x,Z) = J(s,x,a)$.*

Thanks to (3.99), we can exploit the regularity of the flow $X^{\Pi_{\bar{n}}}$ generated by (3.98) (see Definition 3.2) to deduce the following property of v.

Lemma 3.23. *The function $v\colon [0,T) \times \mathbb{R}^d \to \mathbb{R}$ is lower semicontinuous.*

Proof. Fix $n \in \mathbb{N}$, $s \in [0,T)$, $x \in \mathbb{R}^d$ and $a \in \mathcal{A}_n$. By Points 2. (iii) and 4. in Definition 3.2 (see also (3.71)-(3.87)) we infer that there exists an a.s. event Ω_s such that

$$X_t^{s-,x,\mathbf{y}}(\omega) = X_t^{s,x,\mathbf{y}}(\omega), \quad x \in \mathbb{R}^d,\ \mathbf{y} \in \left(\mathbb{R}^l\right)^{2^n},\ t \in [0,T],\ \omega \in \Omega_s. \tag{3.112}$$

Hence by (3.99) we infer that

$$X_t^{s-,x,a}(\omega) = \lim_{r \uparrow s} X_t^{r,x,a}(\omega) = X_t^{s,x,a}(\omega), \quad x \in \mathbb{R}^d,\ t \in [0,T],\ \omega \in \Omega_s.$$

Thus, if we take a sequence $(s_k, x_k)_k \subset [0,T) \times \mathbb{R}^d$ such that $(s_k, x_k) \to (s,x)$ as $k \to \infty$, then the dominated convergence theorem yields $J(s,x,a) = \lim_{k\to\infty} J(s_k, x_k, a)$. Since a is chosen arbitrarily, we deduce that

$$v(s,x) = \sup_{a \in \mathcal{A}} \lim_{k\to\infty} J(s_k, x_k, a) \le \liminf_{k\to\infty} \sup_{a\in\mathcal{A}} J(s_k,x_k,a) = \liminf_{k\to\infty} v\left(s_k, x_k\right),$$

which completes the proof. ■

The following, technical result is about a continuity property of J with respect to the control $a \in \mathcal{A}$.

Lemma 3.24. *For every $n \in \mathbb{N}$, consider $Z \in \mathcal{Z}_n$ and a sequence $(Z_k)_k \subset \mathcal{Z}_n$ such that $\lim_{k\to\infty}(Z_k)_i = Z_i$, $i = 0, \ldots, 2^n - 1$, $\mathbb{P}-$a.s. (see (3.100)). Then, identifying $\mathcal{A}_n \ni a_k \sim Z_k$ and $\mathcal{A}_n \ni a \sim Z$, the following holds for every $s \in [0,T)$ and $x \in \mathbb{R}^d$:*

$$J(s,x,a) = \lim_{k\to\infty} J(s,x,a_k).$$

Proof. Fix $n \in \mathbb{N}$, $s \in [0,T)$ and $x \in \mathbb{R}^d$; from (3.99) and Point 2. (ii) in Definition 3.2 we have

$$\lim_{k\to\infty} X_t^{s,x,a_k} = X_t^{s,x,a}, \quad t \in [0,T],\ \mathbb{P}-\text{a.s.}$$

Then, writing $J(s,x,a_k) = \sum_{i=0}^{2^n-1} \mathbb{E}\left[\int_s^T 1_{\{t_i^{(n)} < r \le t_{i+1}^{(n)}\}} h\left(r, X_r^{s,x,a_k}, (Z_k)_i\right) dr\right] + \mathbb{E}\left[j\left(X_T^{s,x,a_k}\right)\right]$, $k \in \mathbb{N}$, the continuity hypotheses on h and j allow us to invoke the dominated convergence theorem and complete the proof. ■

Taking into account Remark 3.8 we now restrict the set of controls. To do this, fix $n \in \mathbb{N}$; recall the set $S_n = \{t_i^{(n)},\ i = 0, \ldots, 2^n\}$ of dyadic points of $[0,T]$ with mesh $T2^{-n}$.

Note that, as U is a Polish space, according to (3.108) the space $L_i^{(n)} = L^2(\Omega, \mathcal{F}^{W,N_p}_{0,t_i^{(n)}}; \mathbb{R}^l)$ is separable, $i = 0, \ldots, 2^n - 1$. It is then possible to consider an Hilbert basis $\left(e_{i,m}^{(n)}\right)_m$ for each $L_i^{(n)}$. We denote by

$$L_{i,M}^{(n)} \text{ the subspace of } L_i^{(n)} \text{ generated by } e_{i,1}^{(n)}, \ldots, e_{i,M}^{(n)}, \text{ for every } M \in \mathbb{N}. \tag{3.113}$$

We introduce the subspace $\mathcal{A}_{n,M} \subset \mathcal{A}_n$ given by the simple processes $a \in \mathcal{A}_n$ of the form

$$a(t,\omega) = \sum_{i=0}^{2^n-1} Z_i(\omega) 1_{\left(t_i^{(n)}, t_{i+1}^{(n)}\right]}(t), \quad t \in [0,T],\ \omega \in \Omega,$$

for some family $Z = (Z_i)_i \subset L^{(n)}_{i,M}$. As before, we define

$$\mathcal{Z}_{n,M} = \left\{(Z_i)_{i=0,\dots,2^n-1} \text{ such that } Z_i \in L^{(n)}_{i,M}\right\},$$

and we identify $\mathcal{A}_{n,M} \ni a \sim Z \in \mathcal{Z}_{n,M}$; moreover we set $\widetilde{\mathcal{A}} = \cup_{n,M\in\mathbb{N}}\mathcal{A}_{n,M}$ and $\widetilde{\mathcal{Z}} = \cup_{n,M\in\mathbb{N}}\mathcal{Z}_{n,M}$. Since $L^{(n)}_i = \overline{\cup_M L^{(n)}_{i,M}}$, where the closure is relative to the L^2-norm, Lemma 3.24 entails that

$$v(s,x) = \sup_{a\in\widetilde{\mathcal{A}}} J(s,x,a). \tag{3.114}$$

Thus, at the level of the value function v, working with $\widetilde{\mathcal{A}}$ instead of $\mathcal{A}$ is not restrictive.

Fix $s \in [0,T]$ and n, $M \in \mathbb{N}$. For every $i = 0,\dots,2^n-1$ such that $t^{(n)}_i \geq s$, we define

$$H^{(n)}_{s,i} = L^2\left(\Omega, \mathcal{F}^{W,N_p}_{s,t^{(n)}_i}; \mathbb{R}^l\right): \tag{3.115}$$

by (3.108), $H^{(n)}_{s,i}$ is separable, hence it has an Hilbert basis $\left(e^{(n)}_{s,i,m}\right)_m$. We denote by $H^{(n)}_{s,i,M}$ the subspace of $H^{(n)}_{s,i}$ generated by $e^{(n)}_{s,i,1},\dots,e^{(n)}_{s,i,M}$, and by

$$\mathcal{Z}^s_{n,M} = \left\{\left(Z_i\right)_{i=0,\dots,2^n-1} \in \mathcal{Z}_{n,M} \text{ such that } Z_i \in H^{(n)}_{s,i,M}, \text{ for all } i : t^{(n)}_i \geq s\right\}:$$

the corresponding simple, finitely generated controls are denoted by $\mathcal{A}^s_{n,M}$. Moreover, we set

$$\widetilde{\mathcal{A}}^s = \bigcup_{n,M\in\mathbb{N}} \mathcal{A}^s_{n,M}, \qquad \widetilde{\mathcal{Z}}^s = \bigcup_{n,M\in\mathbb{N}} \mathcal{Z}^s_{n,M}, \tag{3.116}$$

and we define the function

$$\tilde{v}(s,x) = \sup_{a\in\widetilde{\mathcal{A}}^s} J(s,x,a), \quad s\in[0,T),\ x\in\mathbb{R}^d.$$

By analogy with (3.114), if we call

$$\mathcal{Z}^s_n = \left\{(Z_i)_{i=0,\dots,2^n-1} \in \mathcal{Z}_n \text{ such that } Z_i \text{ is } \mathcal{F}^{W,N_p}_{s,t^{(n)}_i}\text{−measurable, for all } i : t^{(n)}_i \geq s\right\}$$

and denote by $\mathcal{A}^s_n$ the corresponding simple controls, then by Lemma 3.24 we obtain

$$\tilde{v}(s,x) = \sup_{a\in\cup_{n\in\mathbb{N}}\mathcal{A}^s_n} J(s,x,a), \quad s\in[0,T),\ x\in\mathbb{R}^d. \tag{3.117}$$

The next lemma clarifies that we can restrict ourselves to controls independent of $\mathcal{F}^{W,N_p}_{0,s}$ (cfr. [40, Remark 5.2] and [146, Remark 3.1]).

Lemma 3.25. *The following equality holds for every $s \in [0,T)$ and $x \in \mathbb{R}^d$:*

$$v(s,x) = \tilde{v}(s,x). \tag{3.118}$$

Proof. We only focus on the inequality $v(s,x) \le \tilde{v}(s,x)$, being the other one trivial. Fix $s \in [0,T)$, $x \in \mathbb{R}^d$ and a simple control $a \in \mathcal{A}_{n,M}$, with $n, M \in \mathbb{N}$. Using

$$a^{s,\bar{\omega}}_{s,n}(t,\omega) = 1_{\{t \le s_n^+\}} a(t,\bar{\omega}) + 1_{\{t > s_n^+\}} \sum_{i=0}^{2^n-1} F_i^{(n)} \left(W_{s\wedge\cdot}(\bar{\omega}) + [W_{s\vee\cdot} - W_s](\omega),\right.$$
$$\left. N_p\left(\cdot \cap ((0,s] \times U)\right)(\bar{\omega}) + N_p\left(\cdot \cap ((s,T] \times U)\right)(\omega)\right) 1_{\left(t_i^{(n)}, t_{i+1}^{(n)}\right]}(t),$$

where $t \in [0,T]$, ω, $\bar{\omega} \in \Omega$ (see (3.107)), we compute

$$J(s,x,a) = \mathbb{E}\left[\mathbb{E}\left[\int_s^T h\left(r, X_r^{s,x,a}, a_r\right) dr + j\left(X_T^{s,x,a}\right) \middle| \mathcal{F}_{0,s}^{W,N_p}\right]\right] = \mathbb{E}\left[J(s,x,a^{s,\cdot}_{s,n})\right] \le \tilde{v}(s,x).$$

Here the last inequality is due to (3.117) and the fact that $a^{s,\omega}_{s,n} \in \mathcal{A}^s_n$, for every $\omega \in \Omega$. Taking the supremum over $a \in \widetilde{\mathcal{A}}$, by (3.114) we deduce (3.118) and complete the proof. ∎

The next remarks will be important in the proof of the DPP (see, in particular, Step II).

Remark 3.7. *For every $\bar{\theta} \in [0,T]$, the set $\widetilde{\mathcal{Z}}^{\bar{\theta}}$ (see (3.116)) can be identified with a Borel subset of the separable space*

$$\prod_{\substack{i,n\,:\,t_i^{(n)} < \bar{\theta} \\ M\in\mathbb{N}}} \left[L_{i,M}^{(n)}\right] \times \prod_{\substack{i,n\,:\,\bar{\theta} \le t_i^{(n)} < T \\ M\in\mathbb{N}}} \left[H_{\bar{\theta},i,M}^{(n)}\right].$$

Therefore, by [27, Proposition 7.12], $\widetilde{\mathcal{Z}}^{\bar{\theta}}$ endowed with the trace topology is a Borel space. Moreover, for every $n, M \in \mathbb{N}$ and $i = 0, \dots, 2^n - 1$ such that $t_i^{(n)} \ge \bar{\theta}$ [resp., $t_i^{(n)} < \bar{\theta}$], the projection map $\pi^{(n)}_{\bar{\theta},i,M} \colon \widetilde{\mathcal{Z}}^{\bar{\theta}} \to H^{(n)}_{\bar{\theta},i,M}$ [resp., $\pi^{(n)}_{\bar{\theta},i,M} \colon \widetilde{\mathcal{Z}}^{\bar{\theta}} \to L^{(n)}_{i,M}$] is continuous, hence Borel measurable.

Remark 3.8. *Fix a separable Hilbert space $H = L^2(\Omega, \mathcal{G}; \mathbb{R}^k)$ with basis $\mathbf{e} = (e_m)_m$. We can choose a representative for every element e_m, $m \in \mathbb{N}$, of the basis $\mathbf{e}$. In this way, it makes sense to consider $e_m(\omega)$, $\omega \in \Omega$. Given $M \in \mathbb{N}$, we denote by $F_M = span\{e_1, \dots, e_M\}$; for every $y \in F_M$, there exists a unique representative $\bar{y}$ of y such that $\bar{y}(\omega) = \sum_{m=1}^M \langle y, e_m \rangle_H e_m(\omega)$, $\omega \in \Omega$. Note that $F_M \subset H$ is isomorphic to $\mathbb{R}^{kM}$. We can define a map $T \colon F_M \times (\Omega, \mathcal{G}) \to \mathbb{R}$ by*

$$T(y,\omega) = \bar{y}(\omega), \quad y \in F_M,\, \omega \in \Omega.$$

Note that T is well defined, and depends on the choice of the representatives for $(e_m)_m$.
Furthermore, we remark that T is measurable with respect to the product $\sigma-$algebra. *Indeed, this is a consequence of the fact that $T(y,\cdot)$ is $\mathcal{G}-$measurable for every $y \in F_M$, and that $T(\cdot,\omega) \colon F_M \to \mathbb{R}$ is continuous for every $\omega \in \Omega$. This observation will be important in* (3.123) *(see also* (3.126)*).*

For the proof of DPP, see Theorem 3.27, we need to use a classical *measurable selection theorem*. In particular, we employ a simplified version of [42, Theorem 2] (see Theorem 3.26), an important result which is also considered, in a more general form, in [27, Proposition 7.50(b)]. Note that, according to [42, Remark 2, Page 909], we can work with functions defined in separable, absolute Borel sets (or Borel spaces) instead of Polish spaces. Recall that a topological space X is said to be a Borel space if there exists a Polish space Z such that X is homeomorphic to a member of the Borel $\sigma-$algebra of Z. We also recall that, given two Borel spaces X and Y, a function $g : X \to Y$ is *universally measurable* if it is $(\mathcal{G}, \mathcal{B}(Y))-$measurable, where $\mathcal{G}$ is the intersection of the completions of the Borel $\sigma-$algebra $\mathcal{B}(X)$ with respect to all the Borel probability measures on X.

Theorem 3.26 ([42]). *Let X and Y be Borel spaces and let $f : X \times Y \to \mathbb{R}$ be a Borel measurable bounded function. Then, for any $\epsilon > 0$, there exists a universally measurable function $\varphi_\epsilon : X \mapsto Y$ such that*

$$f(x, \varphi_\epsilon(x)) \leq \inf_{y \in Y} f(x, y) + \epsilon, \quad x \in X.$$

We are now ready to state the dynamic programming principle (or Bellman's principle).

Theorem 3.27. *Consider* (3.96) *and* (3.109). *Fix $s \in [0, T)$ and denote by $\mathcal{T}_{s,T}$ the set of stopping times with respect to the filtration $\mathbb{F}^{W,N_p}$ taking values in (s, T). Then the following holds:*

$$\begin{aligned} v(s,x) &= \sup_{a \in \mathcal{E}} \inf_{\theta \in \mathcal{T}_{s,T}} \mathbb{E}\Big[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr + v\left(\theta, X_\theta^{s,x,a}\right) \Big] \\ &= \sup_{a \in \mathcal{E}} \sup_{\theta \in \mathcal{T}_{s,T}} \mathbb{E}\Big[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr + v\left(\theta, X_\theta^{s,x,a}\right) \Big], \quad s \in [0,T),\, x \in \mathbb{R}^d. \end{aligned} \tag{3.119}$$

Remark 3.9. *Equation* (3.119) *in Theorem 3.27 gives a stronger version of the DPP, which is typically formulated as follows: for any stopping time $\theta \in \mathcal{T}_{s,T}$,*

$$v(s,x) = \sup_{a \in \mathcal{E}} \mathbb{E}\Big[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr + v\left(\theta, X_\theta^{s,x,a}\right) \Big], \quad s \in [0,T),\, x \in \mathbb{R}^d.$$

Proof. We divide the proof into two steps.

Step I: We show that, for every $s \in [0, T)$ and $x \in \mathbb{R}^d$,

$$v(s,x) \leq \sup_{a \in \mathcal{A}} \inf_{\theta \in \mathcal{T}_{s,T}} \mathbb{E}\left[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr + v\left(\theta, X_\theta^{s,x,a}\right) \right]. \tag{3.120}$$

Since $\mathcal{A} \subset \mathcal{E}$, the estimate in (3.120) will hold replacing $\mathcal{A}$ with $\mathcal{E}$, too. Fix $s \in [0, T)$, $x \in \mathbb{R}^d$, $\theta \in \mathcal{T}_{s,T}$ and $a \in \mathcal{A}_q$, for some $q \in \mathbb{N}$. By the flow property in Point 3. of Definition 3.1,

$$\begin{aligned} J(s,x,a) &= \mathbb{E}\left[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr \right] + \mathbb{E}\left[\int_\theta^T h\left(r, X_r^{\theta, X_\theta^{s,x,a}, a}, a_r\right) dr + j\left(X_T^{\theta, X_\theta^{s,x,a}, a} \right) \right] \\ &= \mathbb{E}\left[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr \right] \\ &\quad + \sum_{k=0}^{2^q - 1} \mathbb{E}\left[1_{\left(t_k^{(q)}, t_{k+1}^{(q)}\right]}(\theta) \left(\int_{\theta_{k+1}}^T h\left(r, X_r^{\theta_{k+1}, X_{\theta_{k+1}}^{s,x,a}, a}, a_r\right) dr + j\left(X_T^{\theta_{k+1}, X_{\theta_{k+1}}^{s,x,a}, a} \right) \right) \right], \end{aligned} \tag{3.121}$$

where we denote by $\theta_{k+1} = \theta \wedge t_{k+1}^{(q)}$, $k = 0, \ldots, 2^q - 1$.
We focus on the second addend in (3.121). Observe that $X_{\theta_{k+1}}^{s,x,a}$ is measurable with respect to the $\sigma-$algebra $\mathcal{F}_{\theta_{k+1}}^{W,N_p}$ generated by the stopping time θ_{k+1} relative to the filtration $\mathbb{F}^{W,N_p}$. Therefore, thanks to the arguments in Subsection 3.5.2 (see, in particular, (3.107)), we condition with respect to

$\mathcal{F}^{W,N_p}_{\theta_{k+1}}$ to deduce that, for every $k = 0, \dots, 2^q - 1$, for $\mathbb{P}-$a.s. $\omega \in \Omega$,

$$\mathbb{E}\left[1_{\left(t_k^{(q)}, t_{k+1}^{(q)}\right]}(\theta)\left(\int_{\theta_{k+1}}^T h\left(r, X_r^{\theta_{k+1}, X^{s,x,a}_{\theta_{k+1}}, a}, a_r\right) dr + j\left(X_T^{\theta_{k+1}, X^{s,x,a}_{\theta_{k+1}}, a}\right)\right) \Big| \mathcal{F}^{W,N_p}_{\theta_{k+1}}\right](\omega)$$
$$= 1_{\left(t_k^{(q)}, t_{k+1}^{(q)}\right]}(\theta_{k+1}(\omega)) J\left(\theta_{k+1}(\omega), X^{s,x,a}_{\theta_{k+1}}(\omega), a^{\theta_{k+1},\omega}_{\theta_{k+1}(\omega),q}\right)$$
$$= 1_{\left(t_k^{(q)}, t_{k+1}^{(q)}\right]}(\theta(\omega)) J\left(\theta(\omega), X^{s,x,a}_{\theta}(\omega), a^{\theta,\omega}_{\theta(\omega),q}\right).$$

Since $a^{\theta,\omega}_{\theta(\omega),q} \in \mathcal{A}_q$, by the definition of the value function v we deduce that

$$J\left(\theta(\omega), X^{s,x,a}_{\theta}(\omega), a^{\theta,\omega}_{\theta(\omega),q}\right) \le v(\theta(\omega), X^{s,x,a}_{\theta}(\omega)), \quad \omega \in \Omega.$$

Going back to (3.121) we conclude that, by the law of total expectation,

$$J(s,x,a) \le \mathbb{E}\Big[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr\Big] + \mathbb{E}\left[v(\theta, X_\theta^{s,x,a})\right].$$

Since $a \in \mathcal{A}$ and $\theta \in \mathcal{T}_{s,T}$ are arbitrary, we obtain (3.120).

<u>*Step II:*</u> We show that, for every $s \in [0,T)$ and $x \in \mathbb{R}^d$,

$$v(s,x) \ge \sup_{a\in\mathcal{E}} \sup_{\theta\in\mathcal{T}_{s,T}} \mathbb{E}\Big[\int_s^\theta h\left(r, X_r^{s,x,a}, a_r\right) dr + v\left(\theta, X_\theta^{s,x,a}\right)\Big]. \tag{3.122}$$

Fix $s \in [0,T)$, $x \in \mathbb{R}^d$ and $a \in \mathcal{E}$; first, we assume that $\theta = \bar\theta \in (s,T) \cap S$. Our idea is to apply a measurable selection argument, specifically Theorem 3.27. However, before doing this, some necessary preparation is required.

Consider $n, M \in \mathbb{N}$ and $i = 0, \dots, 2^n - 1$ such that $t_i^{(n)} \ge \bar\theta$; recall the space

$$H^{(n)}_{\bar\theta,i,M} \subset H^{(n)}_{\bar\theta,i} \text{ generated by } e^{(n)}_{\bar\theta,i,m},\ m = 1, \dots, M,$$

see (3.115) and the subsequent line. We write e_m for $e^{(n)}_{\bar\theta,i,m}$ to keep notation simple and fix a representative for every function e_m, $m = 1, \dots, M$. Denoting by $\bar y$ the unique representative of $y \in H^{(n)}_{\bar\theta,i,M}$ such that $\bar y(\omega) = \sum_{m=1}^M \langle y, e_m\rangle_{L^2} e_m(\omega)$, $\omega \in \Omega$, by Remark 3.8 we can define the measurable map

$$g_{t_i^{(n)}} \colon H^{(n)}_{\bar\theta,i,M} \times \left(\Omega, \mathcal{F}^{W,N_p}_{\bar\theta, t_i^{(n)}}\right) \to \mathbb{R}^l, \qquad g_{t_i^{(n)}}(y,\omega) = \bar y(\omega). \tag{3.123}$$

When $t_i^{(n)} < \bar\theta$, we can follow the same argument to define, recalling also (3.113), the measurable map $g_{t_i^{(n)}} \colon L^{(n)}_{i,M} \times \left(\Omega, \mathcal{F}^{W,N_p}_{0,t_i^{(n)}}\right) \to \mathbb{R}^l$ as in (3.123).

We now focus on the application of Theorem 3.27. Denote by $\mathcal{X} = \mathcal{D}([s,\bar\theta]; \mathbb{R}^d)$ the usual space of càdlàg functions endowed with the Skorokhod topology: since $\mathcal{X}$ is a Polish space, it is a Borel space too. Recalling (3.116), we introduce the function $f \colon \mathcal{X} \times \widetilde{\mathcal{Z}}^{\bar\theta} \to \mathbb{R}$ defined by, for every $\xi \in \mathcal{X}$ and $y \in \widetilde{\mathcal{Z}}^{\bar\theta}$,

$$f(\xi, y) = -J(\bar\theta, \pi_{\bar\theta}(\xi), y), \quad \text{ where } \pi_{\bar\theta}(\xi) = \xi(\bar\theta).$$

The map f is Borel measurable and bounded. Thus, by Remark 3.7 we can apply Theorem 3.27, which yields, for any $\epsilon > 0$, the existence of a universally measurable function $c_\epsilon\colon \mathcal{X} \to \widetilde{\mathcal{Z}}^{\bar\theta}$ such that

$$J(\bar\theta, \pi_{\bar\theta}(\xi), c_\epsilon(\xi)) \geq \sup_{y \in \widetilde{\mathcal{Z}}^{\bar\theta}} J(\bar\theta, \pi_{\bar\theta}(\xi), y) - \epsilon = v(\bar\theta, \pi_{\bar\theta}(\xi)) - \epsilon, \quad \xi \in \mathcal{X}. \tag{3.124}$$

Note that the last equality in (3.124) is due to Lemma 3.25. Consequently,

$$\mathbb{E}[J(\bar\theta, X^{s,x,a}_{\bar\theta}, c_\epsilon(X^{s,x,a}_\cdot))] \geq \mathbb{E}[v(\bar\theta, X^{s,x,a}_{\bar\theta})] - \epsilon; \tag{3.125}$$

for every $\omega \in \Omega$, we identify

$$\widetilde{\mathcal{A}}^{\bar\theta} \ni \tilde c_\epsilon(\omega) \sim c_\epsilon(X^{s,x,a}_\cdot(\omega)) \in \widetilde{\mathcal{Z}}^{\bar\theta}.$$

At this point, we modify the control $a \in \mathcal{E}$ after time $\bar\theta$ using the processes $\tilde c_\epsilon(\cdot)$, with the aim of invoking (3.125) and the flow property in Point 3. of Definition 3.1. However, since $\tilde c_\epsilon(\omega) \in \mathcal{A}^{\bar\theta}_{n,M}$ for some integers n, M *depending on* ω, we need to consider suitable approximated controls to make the procedure rigorous.

For every $\omega \in \Omega$ and $n, M \in \mathbb{N}$, define $c^{(n)}_{\epsilon,M}(\omega) \in \mathcal{Z}^{\bar\theta}_{n,M}$ by (recall Remark 3.7)

$$\left(c^{(n)}_{\epsilon,M}(\omega)\right)_i = \pi^{(n)}_{\bar\theta,i,M}\left(c_\epsilon(X^{s,x,a}_\cdot(\omega))\right), \quad i = 0, \dots, 2^n - 1;$$

we identify $\quad \mathcal{A}^{\bar\theta}_{n,M} \ni c_{\epsilon,M,n}(\omega) \sim c^{(n)}_{\epsilon,M}(\omega) \in \mathcal{Z}^{\bar\theta}_{n,M}$.

For every $K \in \mathbb{N}$, we denote by $\Phi_K\colon \mathbb{R}^l \to \mathbb{R}^l$ the truncation function defined by $\Phi_K(x) = x1_{\{|x| \leq K\}}$, $x \in \mathbb{R}^l$. Then we consider the quantity $c^{(n)}_{\epsilon,M,K}(\omega) \in \mathcal{Z}^{\bar\theta}_{n,M}$, given by

$$\left(c^{(n)}_{\epsilon,M,K}(\omega)\right)_i(\cdot) = \Phi_K\left(\left(c^{(n)}_{\epsilon,M}(\omega)\right)_i(\cdot)\right), \quad i = 0, \dots, 2^n - 1,\ K \in \mathbb{N};$$

we identify $\quad \mathcal{A}^{\bar\theta}_{n,M} \ni c_{\epsilon,M,n,K}(\omega) \sim c^{(n)}_{\epsilon,M,K}(\omega) \in \mathcal{Z}^{\bar\theta}_{n,M}$.

Recalling (3.123) and the subsequent comment, we introduce the $\mathcal{F}^{W,N_p}_{0,t^{(n)}_i}$–measurable random variable

$$\tilde c^{(n)}_{\epsilon,M,i}(\omega) \coloneqq \left(c^{(n)}_{\epsilon,M}(\omega)\right)_i(\omega) = g_{t^{(n)}_i}\left(\pi^{(n)}_{\bar\theta,i,M}\left(c_\epsilon(X^{s,x,a}_\cdot(\omega))\right), \omega\right), \quad \omega \in \Omega,\ i = 0, \dots, 2^n - 1.$$

Once again, for every $K \in \mathbb{N}$ and $i = 0, \dots, 2^n - 1$, we define

$$\tilde c^{(n)}_{\epsilon,M,i,K}(\omega) \coloneqq \Phi_K\left(\tilde c^{(n)}_{\epsilon,M,i}(\omega)\right) = \Phi_K\left(g_{t^{(n)}_i}\left(\pi^{(n)}_{\bar\theta,i,M}\left(c_\epsilon(X^{s,x,a}_\cdot(\omega))\right), \omega\right)\right). \tag{3.126}$$

Note that the map Φ_K ensures that the random variables $\tilde c^{(n)}_{\epsilon,M,i,K}$ in (3.126) are square–integrable, which allows us to identify $\quad \mathcal{Z}_n \ni \left(\tilde c^{(n)}_{\epsilon,M,i,K}\right)_i \sim \tilde c_{\epsilon,M,n,K} \in \mathcal{A}_n$.

Consider $n, M, K \in \mathbb{N}$ and suppose that $\bar\theta \in \Pi_n$. We introduce the simple process $a_{\epsilon,M,n,K}$ defined by

$$a_{\epsilon,M,n,K}(t,\omega) = 1_{\{t \leq \bar\theta\}} a(t,\omega) + 1_{\{t > \bar\theta\}} \tilde c_{\epsilon,M,n,K}(t,\omega), \quad t \in [0,T],\ \omega \in \Omega.$$

The control $a_{\epsilon,M,n,K} \in \mathcal{E}$ because $\tilde c_{\epsilon,M,n,K} \in \mathcal{A}_n$: this explains the need for the truncation map Φ_K. Thus, by the flow property of $X^{s,x,(a_{\epsilon,M,n,K})}$ in Point 3. of Definition 3.1, we have

$$\begin{aligned} v(s,x) \geq{}& \mathbb{E}\left[\int_s^{\bar\theta} h\left(r, X^{s,x,a}_r, a_r\right) dr\right] \\ &+ \mathbb{E}\left[\int_{\bar\theta}^T h\left(r, X_r^{\bar\theta, X^{s,x,a}_{\bar\theta}, (\tilde c_{\epsilon,M,n,K})}, \left(\tilde c_{\epsilon,M,n,K}\right)_r\right) dr + j\left(X_T^{\bar\theta, X^{s,x,a}_{\bar\theta}, (\tilde c_{\epsilon,M,n,K})}\right)\right]. \end{aligned} \tag{3.127}$$

Observe that, for every $y \in \mathbb{R}^d$, $\mathbf{y}(n) \in (\mathbb{R}^l)^{2^n}$ and $r \in (\bar\theta, T]$, by [167, Theorem 117] and the construction carried out in the proof of Theorem 3.19, the random variable $X_r^{\bar\theta,y,\mathbf{y}(n)}$ is independent from $\mathcal{F}^{W,N_p}_{0,\bar\theta}$. Moreover, by (3.126) we infer that, for every $i = 0, \dots, 2^n - 1$ such that $t_i^{(n)} \geq \bar\theta$, for $\mathbb{P}-$a.s. $\omega \in \Omega$,

$$\begin{aligned}\mathbb{E}\left[\tilde c^{(n)}_{\epsilon,M,i,K} \middle| \mathcal{F}^{W,N_p}_{0,\bar\theta}\right](\omega) &= \mathbb{E}\left[\Phi_K\left(g_{t_i^{(n)}}\left(\pi^{(n)}_{\bar\theta,i,M}\left(c_\epsilon(\xi)\right),\cdot\right)\right)\right]_{\xi = X^{s,x,a}_\cdot(\omega)}\\ &= \mathbb{E}\left[\Phi_K\left(g_{t_i^{(n)}}\left(\pi^{(n)}_{\bar\theta,i,M}\left(c_\epsilon(X^{s,x,a}_\cdot(\omega))\right),\cdot\right)\right)\right] = \mathbb{E}\left[\left(c^{(n)}_{\epsilon,M,K}(\omega)\right)_i(\cdot)\right].\end{aligned}$$

Therefore, $\mathbb{P}-$a.s.,

$$\begin{aligned}\mathbb{E}\left[\int_{\bar\theta}^T h\left(r, X_r^{\bar\theta, X^{s,x,a}_{\bar\theta},(\tilde c_{\epsilon,M,n,K})}, (\tilde c_{\epsilon,M,n,K})_r\right) dr + j\left(X_T^{\bar\theta, X^{s,x,a}_{\bar\theta},(\tilde c_{\epsilon,M,n,K})}\right) \middle| \mathcal{F}^{W,N_p}_{0,\bar\theta}\right]\\ = J\left(\bar\theta, X^{s,x,a}_{\bar\theta}, c_{\epsilon,M,n,K}\right).\end{aligned}$$

Going back to (3.127), by the law of total expectation we can write

$$v(s,x) \geq \mathbb{E}\Big[\int_s^{\bar\theta} h\left(r, X_r^{s,x,a}, a_r\right) dr\Big] + \mathbb{E}\left[J\left(\bar\theta, X^{s,x,a}_{\bar\theta}(\cdot), c_{\epsilon,M,n,K}(\cdot)\right)\right]. \tag{3.128}$$

Observing that, for every $\omega \in \Omega$,

$$\lim_{K\to\infty}\left(c^{(n)}_{\epsilon,M,K}(\omega)\right)_i(\omega') = \left(c^{(n)}_{\epsilon,M}(\omega)\right)_i(\omega'), \quad i = 0,\dots,2^n-1,\ \omega' \in \Omega,$$

by Lemma 3.24 we can pass to the limit as $K \to \infty$ in (3.128) to obtain, by dominated convergence,

$$v(s,x) \geq \mathbb{E}\left[\int_s^{\bar\theta} h\left(r, X_r^{s,x,a}, a_r\right) dr\right] + \mathbb{E}\left[J\left(\bar\theta, X^{s,x,a}_{\bar\theta}(\cdot), c_{\epsilon,M,n}(\cdot)\right)\right]. \tag{3.129}$$

Notice that, for every $\omega \in \Omega$, when n and M are sufficiently large then $\widetilde{\mathcal{A}}^{\bar\theta} \ni \tilde c_\epsilon(\omega) = c_{\epsilon,M,n}(\omega) \in \mathcal{A}^{\bar\theta}_{n,M}$. Hence, choosing $M = n$ we have

$$J\left(\bar\theta, X^{t,x,a}_{\bar\theta}(\omega), \tilde c_\epsilon(\omega)\right) = \lim_{n\to\infty} J\left(\bar\theta, X^{s,x,a}_{\bar\theta}(\omega), c_{\epsilon,n,n}(\omega)\right), \quad \omega \in \Omega,$$

which by dominated convergence implies

$$\mathbb{E}\left[J\left(\bar\theta, X^{t,x,a}_{\bar\theta}(\cdot), \tilde c_\epsilon(\cdot)\right)\right] = \lim_{n\to\infty}\mathbb{E}\left[J\left(\bar\theta, X^{s,x,a}_{\bar\theta}(\cdot), c_{\epsilon,n,n}(\cdot)\right)\right]. \tag{3.130}$$

Combining (3.129)-(3.130) with (3.125) we conclude that

$$v(s,x) \geq \mathbb{E}\left[\int_s^{\bar\theta} h\left(r, X_r^{s,x,a}, a_r\right) dr\right] + \mathbb{E}\left[J\left(\bar\theta, X^{t,x,a}_{\bar\theta}, \tilde c_\epsilon\right)\right] \geq \mathbb{E}\left[\int_s^{\bar\theta} h\left(r, X_r^{s,x,a}, a_r\right) dr + v(\bar\theta, X^{t,x,a}_{\bar\theta})\right] - \epsilon.$$

Suppose now that θ is a simple, $\mathbb{F}^{W,N_p}-$stopping time with values in $(s,T)\cap S$. Then we can write $\theta = \sum_{k=1}^N \theta_k 1_{A_k}$, where $\theta_k \in (s,T)\cap S$ and $(A_k)_k$ is a partition of Ω such that $A_k \in \mathcal{F}^{W,N_p}_{0,\theta_k}$, $k = 1,\dots,N$.

We can invoke the measurable selection theorem (see Theorem 3.26) $N-$times to deduce, for any $\epsilon > 0$, the existence of a universally measurable function $c_{\epsilon,k}\colon \mathcal{D}\left([s,\theta_k];\mathbb{R}^d\right) \to \widetilde{\mathcal{Z}}^{\theta_k}$ such that (cfr. (3.124))

$$J(\theta_k, X^{s,x,a}_{\theta_k}(\omega), c_{\epsilon,k}(X^{s,x,a}_{\cdot}(\omega))) \geq v\left(\theta_k, X^{s,x,a}_{\theta_k}(\omega)\right) - \epsilon, \quad k = 1,\dots,N,\ \omega \in \Omega.$$

If we identify, for every $k = 1,\dots,N$, $\widetilde{\mathcal{A}}^{\theta_k} \ni \tilde{c}_{\epsilon,k}(\omega) \sim c_{\epsilon,k}(X^{s,x,a}_{\cdot}(\omega)) \in \widetilde{\mathcal{Z}}^{\theta_k}$, $\omega \in \Omega$, then conditioning with respect to $\mathcal{F}^{W,N_p}_{0,\theta_k}$ we can follow the previous arguments to obtain

$$\begin{aligned} v(s,x) &\geq \mathbb{E}\left[\int_s^\theta h\left(r, X^{s,x,a}_r, a_r\right) dr\right] + \sum_{k=1}^N \mathbb{E}\left[1_{A_k} J\left(\theta_k, X^{s,x,a}_{\theta_k}(\cdot), \tilde{c}_{\epsilon,k}(\cdot)\right)\right] \\ &\geq \mathbb{E}\left[\int_s^\theta h\left(r, X^{s,x,a}_r, a_r\right) dr + v(\theta, X^{s,x,a}_\theta)\right] - \epsilon. \end{aligned} \tag{3.131}$$

Finally, we show that (3.122) holds for all $\theta \in \mathcal{T}_{s,T}$. We consider a sequence $(\tilde{\theta}_n)_{n\in\mathbb{N}} \subset \mathcal{T}_{s,T}$ of simple $\mathbb{F}^{W,N_p}-$stopping times with values in $(s,T)\cap S$ such that $\tilde{\theta}_n \downarrow \theta$ as $n \to \infty$, $\mathbb{P}-$a.s. An application of Lemma 3.23 yields, for $\mathbb{P}-$a.s. $\omega \in \Omega$,

$$v\left(\theta(\omega), X^{t,x,a}_\theta(\omega)\right) \leq \liminf_{n\to\infty} v\left(\tilde{\theta}_n(\omega), X^{t,x,a}_{\tilde{\theta}_n}(\omega)\right).$$

Hence by (3.131) and Fatou's lemma, which can be applied because v is bounded, we deduce that

$$\begin{aligned} v(s,x) &\geq \liminf_{n\to\infty}\left(\mathbb{E}\left[\int_s^{\tilde{\theta}_n} h\left(r, X^{s,x,a}_r, a_r\right) dr + v\left(\tilde{\theta}_n, X^{s,x,a}_{\tilde{\theta}_n}\right)\right]\right) - \epsilon \\ &\geq \mathbb{E}\left[\int_s^\theta h\left(r, X^{s,x,a}_r, a_r\right) dr + v(\theta, X^{s,x,a}_\theta)\right] - \epsilon. \end{aligned}$$

Since $\epsilon > 0$, $a \in \mathcal{E}$ and $\theta \in \mathcal{T}_{s,T}$ are chosen arbitrarily, the previous equation entails (3.122).

Combining (3.122) with (3.120) we obtain (3.119). The proof is now complete. ■

Appendix 3.A On the $\sigma-$algebra $\mathcal{C}$

Recall that $\mathcal{D}([0,T];\mathbb{R}^d)$ is the set of $\mathbb{R}^d-$valued, càdlàg functions defined on $[0,T]$, and that we denote by $\mathcal{D}_0 = (\mathcal{D}([0,T];\mathbb{R}^d), \|\cdot\|_0)$ and by $\mathcal{D}_S = (\mathcal{D}([0,T];\mathbb{R}^d), J_1)$. In particular, the Skorokhod topology J_1 is generated by the metric d_S, see Remark 3.3.

Notice that $C(\mathbb{R}^d;\mathcal{D}_0) \subset C(\mathbb{R}^d;\mathcal{D}_S)$, because $d_S(x,y) \leq \|x-y\|_0$ for every $x,y \in \mathcal{D}([0,T];\mathbb{R}^d)$. Since $(C(\mathbb{R}^d;\mathcal{D}_S), d^{lu}_S)$ is separable (see [122]), where

$$d^{lu}_S(f,g) = \sum_{N=1}^\infty 2^{-N} \frac{\sup_{|x|\leq N} d_S(f(x), g(x))}{1 + \sup_{|x|\leq N} d_S(f(x), g(x))}, \quad f,g \in C\left(\mathbb{R}^d;\mathcal{D}_S\right),$$

it follows that $(C(\mathbb{R}^d;\mathcal{D}_0), d^{lu}_S)$ is separable, too. We denote $\mathcal{C}_S$ the corresponding Borel $\sigma-$algebra.

Lemma 3.28. *The following equality between $\sigma-$algebras holds:*

$$\mathcal{C} = \mathcal{C}_S. \tag{3.132}$$

Proof. First we prove the inclusion $\mathcal{C} \subset \mathcal{C}_S$. Fix $x \in \mathbb{R}^d$ and consider the projection $\pi_x : C(\mathbb{R}^d; \mathcal{D}_0) \to \mathcal{D}_0$ defined by $\pi_x(f) = f(x)$, $f \in C(\mathbb{R}^d; \mathcal{D}_0)$. Notice that $\pi_x : (C(\mathbb{R}^d; \mathcal{D}_0), d_S^{lu}) \to \mathcal{D}_S$ is continuous, because if $f_n, f \in C(\mathbb{R}^d; \mathcal{D}_0)$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} d_S^{lu}(f_n, f) = 0$, then

$$\lim_{n\to\infty} d_S(f_n(y), f(y)) = 0, \quad \text{for every } y \in \mathbb{R}^d.$$

As $\mathcal{D}$ is the Borel $\sigma-$algebra generated by J_1, we infer that π_x is $\mathcal{C}_S/\mathcal{D}-$measurable, whence $\mathcal{C} \subset \mathcal{C}_S$.

Secondly, we focus on the inclusion $\mathcal{C}_S \subset \mathcal{C}$. Since $(C(\mathbb{R}^d; \mathcal{D}_0), d_S^{lu})$ is separable, it suffices to show that

$$B(f_0, R) = \left\{ f \in C(\mathbb{R}^d; \mathcal{D}_0) : d_S^{lu}(f_0, f) \le R \right\} \in \mathcal{C}, \quad f_0 \in C(\mathbb{R}^d; \mathcal{D}_0),\ R > 0.$$

Hence we fix $f_0 \in C(\mathbb{R}^d; \mathcal{D}_0)$ and consider the map $C(\mathbb{R}^d; \mathcal{D}_0) \ni f \mapsto d_S^{lu}(f_0, f) \in \mathbb{R}$: we argue that it is $\mathcal{C}-$measurable. Indeed, this is a consequence of the fact that, for every $m \in \mathbb{N}$, the map

$$C(\mathbb{R}^d; \mathcal{D}_0) \ni f \mapsto \frac{\sup_{|x|\le m} d_S(f_0(x), f(x))}{1 + \sup_{|x|\le m} d_S(f_0(x), f(x))} \in \mathbb{R} \quad \text{is } \mathcal{C}-\text{measurable}.$$

Note that the supremum can be computed over $x \in \mathbb{Q}^d$ because $f_0,\ f : \mathbb{R}^d \to \mathcal{D}_S$ are continuous. Since $x/(1 + x)$, $x \in \mathbb{R}$, is measurable, we only need to verify that

$$h_m : C(\mathbb{R}^d; \mathcal{D}_0) \to \mathbb{R} \quad \text{defined by} \quad h_m(f) = \sup_{|x|\le m,\, x\in\mathbb{Q}^d} d_S(f_0(x), f(x)) \quad \text{is } \mathcal{C}-\text{measurable}. \tag{3.133}$$

For every $c \in \mathbb{R}$ and $r > 0$, denoting by $B(c, r) \subset \mathbb{R}$ the closed ball of radius R and center c in $\mathbb{R}$,

$$\begin{aligned} h_m^{-1}(B(c,r)) &= \left\{ f \in C(\mathbb{R}^d; \mathcal{D}_0) : \sup_{|x|\le m,\, x\in\mathbb{Q}^d} d_S(f_0(x), f(x)) \in B(c,r) \right\} \\ &= \bigcap_{|x|\le m,\, x\in\mathbb{Q}^d} \pi_x^{-1}\left(d_S^{-1}(f_0(x), \cdot)(B(c,r)) \right) \in \mathcal{C}, \end{aligned}$$

where in the last step we use that $d_S(f_0(x), \cdot) : \mathcal{D}_S \to \mathbb{R}$ is continuous (hence measurable) and that $\mathcal{D}$ is the Borel $\sigma-$algebra generated by J_1. Thus, (3.133) is satisfied, whence $\mathcal{C}_S \subset \mathcal{C}$.

The double inclusion proves (3.132), completing the proof. ■

Appendix 3.B Proof of Theorem 3.12

In this section we use the same notation and work under the same hypotheses as in Theorem 3.12. First of all, by Jensen's inequality, which can be invoked since θ is concave, we have

$$\sup\left\{ \sum_{j=1}^{n} \theta(p_j),\ p_j \ge 0 \text{ such that } \sum_{j=1}^{n} p_j = 1 \right\} = n\theta\left(\frac{1}{n}\right), \quad n \in \mathbb{N}. \tag{3.134}$$

Before presenting the proof of Theorem 3.12, we introduce an approximation scheme. For every $n \in \mathbb{N}$, we denote by S_n the set of dyadic points

$$S_n = \left\{ Tj2^{-n},\ j = 0, 1, \ldots, 2^n \right\}.$$

For any $t \in [0,T]$, let $t_n^- = \max\{s \in S_n : s \leq t\}$ and $t_n^+ = \min\{s \in S_n : s > t\}$, where we set $T_n^+ = \infty$. Next, for every $n \in \mathbb{N}$ we define the function $g_n \colon \Xi \times S_{n+1} \to S_n$ by

$$g_n(\omega,t) = \begin{cases} t_n^-, & \Delta\left(X_t(\omega), X_{t_n^-}(\omega)\right) = \Delta\left(t_n^-, t, t_n^+\right)(\omega) \\ t_n^+, & \text{otherwise} \end{cases}, \quad \omega \in \Xi,\, t \in S_{n+1}.$$

In our framework, the map $\omega \mapsto \Delta(X_{t_n^\pm}(\omega), X_t(\omega))$ is $\mathcal{G}-$measurable. It follows that also the function $\omega \mapsto \Delta(X_t(\omega), X_{g_n(w,t)}(\omega)) = \Delta(t_n^-, t, t_n^+)(\omega)$ is $\mathcal{G}-$measurable, because it is the minimum of two random variables. Finally, for all $m,n \in \mathbb{N}$ such that $n \geq m$, define $f_{m,n} = g_m \circ \cdots \circ g_n$. Note that $f_{m,n}$ maps $\Xi \times S_{n+1}$ into S_m. The following lemma is a fundamental tool in the proof of Theorem 3.12.

Lemma 3.29. *Let $m,n \in \mathbb{N}$ be such that $n \geq m$ and $\omega \in \Xi$. Then the map $g_n(\omega,\cdot)$ is non–decreasing in S_{n+1} and, restricted to S_n, is the identity. As a consequence, the map $f_{m,n}(\omega,\cdot)$ is non–decreasing in S_{n+1}.*

Furthermore, there exists a family of increasing, càdlàg step functions $f_n \colon \Xi \times [0,T] \to S_n$, $n \in \mathbb{N}$, such that

$$f_{m,n}(\omega,\cdot) \circ f_{n+1}(\omega,\cdot) = f_m(\omega,\cdot), \quad n \geq m,\, \omega \in \Xi, \tag{3.135}$$

and that

$$|t - f_n(\omega,t)| \leq T2^{-n}, \quad \omega \in \Xi,\, t \in [0,T]. \tag{3.136}$$

Proof. Fix $m,n \in \mathbb{N}$ such that $n \geq m$ and $\omega \in \Xi$. In the sequel, we do not explicitly write the dependence of g_n and $f_{m,n}$ on ω to keep notation simple. By definition, $S_n \subset S_{n+1}$ and $t = t_n^-$ for every $t \in S_n$. This implies that $g_n(t) = t$. Consider now $s,t \in S_{n+1}$ such that $s < t$: since $s_n^+ \leq t_n^-$, the function g_n is non–decreasing in S_{n+1}, as desired. Therefore the same property holds for $f_{m,n} = g_m \circ \cdots \circ g_n$, too.

For the second part of the statement, consider $n \in \mathbb{N}$, $\omega \in \Xi$, $s \in S_n$, and for every integer $k \geq n$ define $T(k,n,s)(\omega) = \min\{t \in S_{k+1} : f_{n,k}(\omega,t) = s\}$. Since g_{k+1} restricted to S_{k+1} is the identity, the sequence $(T(k,n,s)(\omega))_{k\geq n}$ is non–increasing, hence there exists $T(n,s)(\omega) = \lim_{k\to\infty} T(k,n,s)(\omega)$. The monotonicity of the map $f_{n,k}$ proved in the previous point yields

$$T(n,s)(\omega) \in \left[s - T2^{-n}, s\right], \quad s \in S_n \setminus \{0\}, \tag{3.137}$$

and

$$T(n,s)(\omega) \leq T\left(n, s_n^+\right)(\omega), \quad \text{setting } T\left(n, T_n^+\right) = T. \tag{3.138}$$

Thanks to (3.138), we can construct the function $f_n \colon \Xi \times [0,T] \to S_n$ defining, for all $\omega \in \Xi$,

$$f_n(\omega,t) = \begin{cases} s, & t \in \left[T(n,s)(\omega), T\left(n,s_n^+\right)(\omega)\right),\ S_n \ni s < T, \\ T, & t \in \left[T(n,T)(\omega), T\right]. \end{cases}$$

It is immediate to notice that f_n is a càdlàg, increasing step function, while (3.136) is guaranteed by (3.137).

It only remains to prove the composition property in (3.135). Consider $\omega \in \Xi$, two integers $n \geq m$ and $s \in S_{n+1}$. Note that, by definition (omitting ω as before),

$$T(k,n+1,s) \geq T(k,n,g_n(s)) \geq \cdots \geq T(k,m,f_{m,n}(s)), \quad k > n,$$

hence passing to the limit as $k \to \infty$,

$$T(n+1,s) \geq T(n,g_n(s)) \geq \cdots \geq T(m,f_{m,n}(s)). \tag{3.139}$$

Take $t \in [0,T]$ and denote by $\tilde{s} \in S_{n+1}$ the unique element in S_{n+1} such that $t \in [T(n+1,\tilde{s}), T(n+1,\tilde{s}^+_{n+1})[$. Notice that this interval is closed when $\tilde{s} = T$. Analogously, let $\bar{s} \in S_m$ be such that $t \in [T(m,\bar{s}), T(m,\bar{s}^+_m)[$, again closing the interval when $\bar{s} = T$. Since, by (3.139), $t \geq T(n+1,\tilde{s}) \geq T(m, f_{m,n}(\tilde{s}))$, from (3.138) we infer that $f_{m,n}(\tilde{s}) \leq \bar{s}$, whence

$$f_{m,n}\left(f_{n+1}\left(t\right)\right) = f_{m,n}(\tilde{s}) \leq \bar{s} = f_m\left(t\right). \tag{3.140}$$

On the other hand, we argue by cases. Firstly, note that $T(m,\bar{s}) \leq T(n+1,\bar{s})$ by (3.139). Secondly, we observe that either $t \geq T(n+1,\bar{s})$ or $t \in [T(m,\bar{s}), T(n+1,\bar{s}))$. In the former case, (3.138) implies that $\bar{s} \leq f_{m,n}(\tilde{s})$. In the latter (where in particular $\bar{s} \neq 0$), there exists $u \in [T(m,\bar{s}), T(n+1,\bar{s})) \cap S_{k+1}$, for some integer $k > n$, such that

$$\bar{s} = f_{m,k}\left(u\right) = f_{m,n}\left(f_{n+1,k}\left(u\right)\right).$$

Since $u < T(n+1,\bar{s})$, $f_{n+1,k}\left(u\right) \in (\bar{s} - T2^{-m}, \bar{s}) \cap S_{n+1}$. Hence we can define $\bar{u} = \min\{v \in (\bar{s} - T2^{-m}, \bar{s}) \cap S_{n+1} : f_{m,n}(v) = \bar{s}\}$, and it is easy to argue by contradiction that $T(n+1,\bar{u}) \leq T(m,\bar{s})$. Consequently,

$$f_{m,n}\left(f_{n+1}\left(t\right)\right) \geq f_{m,n}\left(\bar{u}\right) = \bar{s} = f_m\left(t\right). \tag{3.141}$$

Combining (3.140) and (3.141) we obtain (3.135), completing the proof. ■

We are now ready to prove Theorem 3.12.

Proof of Theorem 3.12 For every $n \in \mathbb{N}$, consider the finite partition $\{A^{(n)}_t,\, t \in S_{n+1} \setminus S_n\}$ of Ξ, defined by

$$\begin{aligned} A^{(n)}_t &= \left\{\omega \in \Xi : \Delta\left(X_t\left(\omega\right), X_{g_n(\omega,t)}\left(\omega\right)\right) = \max_{u \in S_{n+1}} \Delta\left(X_u\left(\omega\right), X_{g_n(\omega,u)}\left(\omega\right)\right)\right\} \\ &= \left(\Delta\left(X_t, X_{g_n(\cdot,t)}\right) - \max_{u \in S_{n+1}} \Delta\left(X_u, X_{g_n(\cdot,u)}\right)\right)^{-1}\{0\}, \quad t \in S_{n+1} \setminus S_n. \end{aligned}$$

Noticing that $\Xi \ni \omega \mapsto \max_{u \in S_{n+1}} \Delta\left(X_u\left(\omega\right), X_{g_n(\omega,u)}\left(\omega\right)\right)$ is $\mathcal{G}-$measurable because it is the maximum of $2^{n+1}+1$ random variables, we deduce that $A^{(n)}_t \in \mathcal{G}$ for every $t \in S_{n+1} \setminus S_n$. Then, by the hypothesis in (3.43),

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left[\max_{u \in S_{n+1}} \Delta\left(X_u, X_{g_n(\cdot,u)}\right)\right] &= \sum_{t \in S_{n+1} \setminus S_n} \mathbb{E}_{\mathbb{Q}}\left[\Delta\left(t^-_n, t, t^+_n\right) 1_{A^{(n)}_t}\right] \\ &\leq \delta\left(T2^{-n}\right) \sum_{t \in S_{n+1} \setminus S_n} \theta\left(\mathbb{Q}\left(A^{(n)}_t\right)\right) \leq \delta\left(T2^{-n}\right) 2^n \theta\left(2^{-n}\right), \end{aligned} \tag{3.142}$$

where in the last step we use (3.134) and the fact that the cardinality of $S_{n+1} \setminus S_n$ is 2^n. Now, since for every $\omega \in \Xi$ and $m, n \in \mathbb{N}$ such that $n \geq m$, (omitting ω to save space)

$$\begin{aligned} \Delta\left(X_t, X_{f_{m,n}(t)}\right) &\leq \Delta\left(X_t, X_{g_n(t)}\right) + \Delta\left(X_{g_n(t)}, X_{g_{n-1}(g_n(t))}\right) + \cdots + \Delta\left(X_{f_{m+1,n}(t)}, X_{f_{m,n}(t)}\right) \\ &\leq \sum_{j=m}^{n} \max_{u \in S_{j+1}} \Delta\left(X_u, X_{g_j(u)}\right), \quad t \in S_{n+1}, \end{aligned}$$

we obtain

$$\sup_{n \geq m} \max_{t \in S_{n+1}} \Delta\left(X_t\left(\omega\right), X_{f_{m,n}(\omega,t)}\left(\omega\right)\right) \leq \sum_{n=m}^{\infty} \max_{u \in S_{n+1}} \Delta\left(X_u\left(\omega\right), X_{g_n(\omega,u)}\left(\omega\right)\right) =: \mathbf{I}_m\left(\omega\right). \tag{3.143}$$

Consider m so big that $2^{1-m} \leq T$ and denote by $\overline{T} = T \vee 1$. From the integral condition in (3.44), (3.142) and recalling that the functions δ, θ are non–decreasing, we invoke the dominated convergence theorem to conclude that

$$\mathbb{E}_{\mathbb{Q}}\left[\mathbf{I}_m\right] \leq \sum_{n=m}^{\infty} \delta\left(T2^{-n}\right) 2^n \theta\left(2^{-n}\right) \leq \overline{T} \sum_{n=m}^{\infty} \delta\left(\overline{T}2^{-n}\right) \frac{2^n}{\overline{T}} \theta\left(\overline{T}2^{-n}\right) \leq 2\overline{T} \int_{m-1}^{\infty} \delta\left(\overline{T}2^{-x}\right) \frac{2^x}{\overline{T}} \theta\left(\overline{T}2^{-x}\right) dx$$
$$= \frac{2\overline{T}}{\log 2} \int_0^T 1_{\{y \leq \overline{T}2^{1-m}\}} y^{-2} \delta\left(y\right) \theta\left(y\right) dy \underset{m\to\infty}{\longrightarrow} 0.$$

Therefore, there exists a subsequence $(\mathbf{I}_{m_k})_k$ such that $\lim_{k\to\infty} \mathbf{I}_{m_k} = 0$ almost uniformly in Ω. This means that there exists a sequence of $\mathcal{G}-$measurable sets $(\Xi_N)_N$, with $\mathbb{Q}(\Xi_N) < N^{-1}$ and $\Xi_{N+1} \subset \Xi_N$, such that

$$\lim_{k\to\infty} \sup_{\omega\in\Xi\setminus\Xi_N} \mathbf{I}_{m_k}\left(\omega\right) = 0.$$

Using a diagonalization argument, we can find a subsequence $(\mathbf{I}_{m_{k_p}})_p$ with the following property:

$$\mathbf{I}_{m_{k_q}}\left(\omega\right) < 2^{-q}, \quad \omega \in \Xi\setminus\Xi_p,\, q \geq p, \text{ for every } p \in \mathbb{N}.$$

Let us define the almost sure event $\Xi_0 = \Xi \setminus \cap_{N=1}^{\infty}\Xi_N$: note that for any $\omega \in \Xi_0$, $\omega \in \Xi \setminus \Xi_N$ for all N sufficiently large. Going back to (3.143), the previous relation gives

$$\sup_{n\geq m_{k_p}} \max_{t\in S_{n+1}} \Delta\left(X_t\left(\omega\right), X_{f_{m_{k_p},n}(\omega,t)}\left(\omega\right)\right) \leq \mathbf{I}_{m_{k_p}}\left(\omega\right) < 2^{-p}, \quad \omega\in\Xi_0,\, p\geq p\left(\omega\right)\in\mathbb{N},$$

which is equivalent to writing, for every $\omega \in \Xi \setminus \Xi_{p(\omega)}$,

$$\Delta\left(X_t\left(\omega\right), X_{f_{m_{k_p},n}(\omega,t)}\left(\omega\right)\right) < 2^{-p}, \quad t\in S_{n+1},\, n\geq m_{k_p},\, p\geq p\left(\omega\right).$$

The composition property in Lemma 3.29 (see (3.135)) yields

$$\Delta\left(X_{f_{n+1}(\omega,t)}\left(\omega\right), X_{f_{m_{k_p}}(\omega,t)}\left(\omega\right)\right) < 2^{-p}, \quad t\in[0,T],\, n\geq m_{k_p},\, p\geq p\left(\omega\right). \tag{3.144}$$

Note that $X_{f_n(\omega,\cdot)}\left(\omega\right) : [0,T] \to E$ is càdlàg for every $n \in \mathbb{N}$ and $\omega \in \Xi$ because $f_n(\omega,\cdot)$ is a càdlàg, step function. Thus, (3.144) shows that the sequence $(X_{f_{m_{k_p}}(\omega,\cdot)}(\omega))_p$ is Cauchy in the metric space of the $E-$valued càdlàg functions defined in $[0,T]$ endowed with the uniform distance. This space is complete since E is complete, hence there exists a càdlàg function $\widetilde{X}(\omega)\colon [0,T] \to E$ such that

$$\lim_{p\to\infty} \sup_{0\leq t\leq T} \Delta\left(\widetilde{X}\left(\omega\right)\left(t\right), X_{f_{m_{k_p}}(\omega,t)}(\omega)\right) = 0, \quad \omega\in\Xi_0.$$

Finally, we define the function $Z\colon \Xi\times[0,T] \to E$ by

$$Z_t\left(\omega\right) = \begin{cases} \widetilde{X}\left(w\right)\left(t\right), & \omega\in\Xi_0, \\ 0, & \text{otherwise.} \end{cases}$$

By construction Z is càdlàg. Moreover, notice that $f_n(\omega,t) \in \{t_n^-, t_n^+\}$ by (3.136), for every $\omega \in \Xi$, $n \in \mathbb{N}$ and $t \in [0,T]$. Recall that X is continuous in probability, hence for every $t \in [0,T]$ there exists a full probability set Ξ_t and a subsequence $(m_{k_{p_q}})_q$ depending on t such that, denoting by $\tilde{q} = \tilde{q}(q) = m_{k_{p_q}}$,

$$\lim_{q\to\infty} \Delta\left(X_t\left(\omega\right), X_{t_{\tilde{q}}^-}\left(\omega\right)\right) = \lim_{q\to\infty} \Delta\left(X_t\left(\omega\right), X_{t_{\tilde{q}}^+}\left(\omega\right)\right) = 0, \quad \omega\in\Xi_t.$$

Then, for all $\omega \in \Xi_0 \cap \Xi_t$,

$$\begin{aligned}\Delta(X_t(\omega), Z_t(\omega)) &= \lim_{q\to\infty} \Delta\Big(X_t(\omega), X_{f_{m_{k_{pq}}}(\omega,t)}(\omega)\Big) \\ &\le \lim_{q\to\infty} \Delta\left(X_t(\omega), X_{t^-_{\bar q}}(\omega)\right) + \Delta\left(X_t(\omega), X_{t^+_{\bar q}}(\omega)\right) = 0.\end{aligned}$$

In conclusion, Z is the càdlàg version of X that we are looking for as $\mathbb{Q}(\Xi_0 \cap \Xi_t) = 1$. The proof is now complete. ■

Appendix 3.C Proofs of Lemma 3.17-3.18

Proof of Lemma 3.17. In order not to complicate the notation, we carry out the proof in the case $\bar s = \bar t = T$, hence $s, t \in [0,T]$. The assumption $\bar t = T$ is not restrictive because the $\mathbb{R}^d-$valued random variables $Z^{s,x}_t$ are $\mathcal{F}_{\bar t}-$measurable for every $t \in [0, \bar t]$, so that all the following passages can be easily adapted to treat a general $\bar t \in [0,T]$. As for $\bar s \in [0,T]$, this case can be treated by considering only dyadic points in $[0,T]$ up to $\bar s$ in the procedure that we are about to explain.

Consider the function $g_1 : (\Omega \times [0,T], \mathcal{F}_T \otimes \mathcal{B}([0,T])) \to (\mathcal{C}_0, \mathcal{C})$ defined by $g_1(\omega, s) = Z_s(\omega)$. Let $S_n = \{Tj2^{-n}, j = 0, \dots, 2^n\}$ be the set of dyadic points and denote by $s^+_n = \min\{t \in S_n : t \ge s\}$, $s \in [0,T]$. Then, we define $g_{1,n}(\omega, s) = g_1(\omega, s^+_n)$. For every $A \in \mathcal{B}(\mathbb{R}^d)$, $n \in \mathbb{N}$, $x \in \mathbb{R}^d$ and $t \in [0,T]$, we have

$$\begin{aligned}g^{-1}_{1,n}\left(\pi^{-1}_x\left(\pi^{-1}_t(A)\right)\right) &= \left(\left(Z^{0,x}_t\right)^{-1}(A) \times \{0\}\right) \\ &\bigcup\Bigg(\bigcup_{s\in S_n\setminus\{0\}} \left((Z^{s,x}_t)^{-1}(A) \times \left(s - T2^{-n}, s\right]\right)\Bigg) \in \mathcal{F}_T \otimes \mathcal{B}([0,T]).\end{aligned}$$

Recalling that $\mathcal{C}$ is the $\sigma-$algebra generated by $\pi_x : \mathcal{C}_0 \to (\mathcal{D}_0, \mathcal{D})$, $x \in \mathbb{R}^d$, and that $\mathcal{D}$ is the $\sigma-$algebra generated by $\pi_t : \mathcal{D}_0 \to \mathbb{R}^d$, $t \in [0,T]$, the previous computation shows that

$$g_{1,n} \text{is } \mathcal{F}_T \otimes \mathcal{B}([0,T])/\mathcal{C}-\text{measurable}.$$

The càdlàg property of $g_1(\omega, \cdot)$ ensured by Theorem 3.14 yields $\lim_{n\to\infty} g_{1,n} = g_1$ pointwise in $\Omega \times [0,T]$, hence g_1 is measurable, as well. Next, consider $g_2 : (\mathcal{C}_0 \times \mathbb{R}^d, \mathcal{C} \otimes \mathcal{B}(\mathbb{R}^d)) \to (\mathcal{D}_0, \mathcal{D})$ given by $g_2(f, x) = \pi_x(f) = f(x)$. For every $n \in \mathbb{N}$, let $\Pi_n = \{2^{-n}z, z \in \mathbb{Z}^d\}$ be the set of lattice points in $\mathbb{R}^d$ with mesh 2^{-n}. Denoting by $\bar x_n = 2^{-n}[2^n x] \in \Pi_n$, $x \in \mathbb{R}^d$, we define $g_{2,n}(f, x) = g_2(f, \bar x_n)$. Note that $g_{2,n}$ is $\mathcal{C} \otimes \mathcal{B}(\mathbb{R}^d)/\mathcal{D}-$measurable for all $n \in \mathbb{N}$. Indeed, for every $A \in \mathcal{B}(\mathbb{R}^d)$ and $t \in [0,T]$,

$$\begin{aligned}g^{-1}_{2,n}\left(\pi^{-1}_t(A)\right) &= \Big\{(f, x) \in \mathcal{C}_0 \times \mathbb{R}^d : [g_{2,n}(f, x)](t) \in A\Big\} \\ &= \bigcup_{x\in\Pi_n} \left(\pi^{-1}_x\left(\pi^{-1}_t(A)\right) \times \left(\left[x_1 + 2^{-n}\right) \times \cdots \times \left[x_d + 2^{-n}\right)\right)\right) \in \mathcal{C} \otimes \mathcal{B}\big(\mathbb{R}^d\big).\end{aligned}$$

Since, by continuity, $\lim_{n\to\infty} g_{2,n} = g_2$ pointwise in $\mathcal{C}_0 \times \mathbb{R}^d$, we conclude that g_2 is measurable. Finally, we introduce the map $g_3 : (\mathcal{D}_0 \times [0,T], \mathcal{D} \otimes \mathcal{B}([0,T])) \to \mathbb{R}^d$ defined by $g_3(f, t) = f(t)$: arguing as we have done for g_1, we infer that g_3 is measurable. At this point, we read the function Z in the statement of this lemma as the following composition, where $\mathrm{Id} : [0,T] \to [0,T]$ and $\mathrm{Id}_d : \mathbb{R}^d \to \mathbb{R}^d$ are the identity maps:

$$Z = g_3 \circ (g_2, \mathrm{Id}) \circ (g_1, \mathrm{Id}_d, \mathrm{Id}).$$

The previous argument allows then to deduce that Z is $\mathcal{F}_T \otimes \mathcal{B}([0,T]\times\mathbb{R}^d\times[0,T])-$measurable. ■

Proof of Lemma 3.18. Since (3.72) can be obtained from (3.76) by setting $f=0$, the existence of a pathwise unique solution of (3.72) can be argued as in Remark 3.4. Thus, we only focus on showing that the process $Z^{s,\eta}$ solves (3.72).

All the assertions of the lemma are trivially satisfied when $s=T$. Thus, we fix $s\in[0,T)$ and take $\eta\in L^0(\mathcal{F}_s)$. By construction, $Z^{s,\eta}_t=\eta$ for every $t\in[0,s]$, $\mathbb{P}-$a.s., so we only focus on $t\in[s,T]$. First of all, notice that $Z^{s,\eta}_t$ is $\mathcal{F}_t-$measurable by Lemma 3.17. Next, consider a sequence of simple, $\mathcal{F}_s-$measurable, $\mathbb{R}^d-$valued random variables $(\eta_n)_n$ such that $\eta_n\to\eta$ as $n\to\infty$, $\mathbb{P}-$a.s. Specifically, for every $n\in\mathbb{N}$, let $\eta_n=\sum_{k=1}^{N_n} a^n_k 1_{A^n_k}$, for some $N_n\in\mathbb{N}$, $(a^n_k)_k\subset\mathbb{R}^d$ and some partition $(A^n_k)_k\subset\mathcal{F}_s$, $k=1,\dots,N_n$. By (3.67), using [106, Section 3, Chapter II] and [110, Property 4.37, Chapter I] we have, $\mathbb{P}-$a.s., for any $n\in\mathbb{N}$,

$$\begin{aligned}
Z^{s,\eta_n}_t &= \sum_{k=1}^{N_n} Z^{s,a^n_k}_t 1_{A^n_k}\\
&= \sum_{k=1}^{N_n}\left[a^n_k+\int_s^t b\left(r,Z^{s,a^n_k}_r\right)dr+\int_s^t \alpha\left(r,Z^{s,a^n_k}_r\right)dW_r+\int_s^t\int_{U_0} g\left(Z^{s,a^n_k}_{r-},r,z\right)\widetilde{N}_p(dr,dz)\right]1_{A^n_k}\\
&= \eta_n+\int_s^t b\left(r,Z^{s,\eta_n}_r\right)dr+\int_s^t \alpha\left(r,Z^{s,\eta_n}_r\right)dW_r+\int_s^t\int_{U_0} g\left(Z^{s,\eta_n}_{r-},r,z\right)\widetilde{N}_p(dr,dz),\quad t\in[s,T]. \quad (3.145)
\end{aligned}$$

In order to recover (3.72), we want to take limits in (3.145) as $n\to\infty$. From the continuity of $Z^{s,x}_t$ in x (see (ii) in Definition 3.1), we infer that $\lim_{n\to\infty} Z^{s,\eta_n}_t = Z^{s,\eta}_t$ uniformly in $t\in[s,T]$, $\mathbb{P}-$a.s. Next, by dominated convergence and (3.2),

$$\lim_{n\to\infty}\int_s^t b\left(r,Z^{s,\eta_n}_r\right)dr=\int_s^t b\left(r,Z^{s,\eta}_r\right)dr,\quad t\in[s,T],\ \mathbb{P}-\text{a.s.} \quad (3.146)$$

The convergence of the stochastic integrals in (3.145) is studied via a localization procedure. As for the integral with respect to the Brownian motion, for every $\epsilon>0$ we define

$$\sigma_n(\epsilon)=\inf\left\{u\in[s,T]:\int_s^u\left|\alpha\left(r,Z^{s,\eta_n}_r\right)-\alpha\left(r,Z^{s,\eta}_r\right)\right|^2dr\ge\epsilon\right\},\quad \text{with } \inf\emptyset=\infty.$$

Since, by the dominated convergence theorem, $\lim_{n\to\infty}\int_s^T\left|\alpha\left(r,Z^{s,\eta_n}_r\right)-\alpha\left(r,Z^{s,\eta}_r\right)\right|^2dr=0$, $\mathbb{P}-$a.s., we have $\sigma_n(\epsilon)\to\infty$ as $n\to\infty$, $\mathbb{P}-$a.s. In particular, $\lim_{n\to\infty}\mathbb{P}\left(\sigma_n(\epsilon)\le T\right)=0$. Hence by Markov's inequality, for every $\delta>0$, for some $c>0$,

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[s,T]}\left|\int_s^t\left(\alpha\left(r,Z^{s,\eta_n}_r\right)-\alpha\left(r,Z^{s,\eta}_r\right)\right)dW_r\right|\ge\delta\Big)\\
&\quad\le\frac{c}{\delta^2}\mathbb{E}\Big[\int_s^T 1_{[s,\sigma_n(\epsilon)]}(r)\left|\alpha\left(r,Z^{s,\eta_n}_r\right)-\alpha\left(r,Z^{s,\eta}_r\right)\right|^2dr\Big]+\mathbb{P}\left(\sigma_n(\epsilon)\le T\right)\\
&\quad\le\frac{c\epsilon}{\delta^2}+\mathbb{P}\left(\sigma_n(\epsilon)\le T\right),\quad \epsilon>0,
\end{aligned}$$

which proves that $\lim_{n\to\infty}\int_s^t\alpha\left(r,Z^{s,\eta_n}_r\right)dW_r=\int_s^t\alpha\left(r,Z^{s,\eta}_r\right)dW_r$ uniformly on $[s,T]$ in probability. This in turn yields the existence of a subsequence such that

$$\lim_{k\to\infty}\int_s^t\alpha\left(r,Z^{s,\eta_{n_k}}_r\right)dW_r=\int_s^t\alpha\left(r,Z^{s,\eta}_r\right)dW_r,\quad \text{uniformly in } t\in[s,T],\ \mathbb{P}-\text{a.s.} \quad (3.147)$$

The convergence of the integral with respect to $\widetilde{N}_p$ in (3.145) can be treated analogously. More precisely, by (3.2), $\lim_{n\to\infty}\int_s^T\int_{U_0}\left|g\left(Z_{r-}^{s,\eta_n},r,z\right)-g\left(Z_{r-}^{s,\eta},r,z\right)\right|^2\nu(dz)\,dr=0$, $\mathbb{P}-$a.s. If we introduce the stopping times

$$\widetilde{\sigma}_n(\epsilon)=\inf\left\{u\in[s,T]:\int_s^u\int_{U_0}\left|g\left(Z_{r-}^{s,\eta_n},r,z\right)-g\left(Z_{r-}^{s,\eta},r,z\right)\right|^2\nu(dz)\,dr\geq\epsilon\right\},\quad \epsilon>0,$$

then we can proceed as before to deduce that, $\mathbb{P}-$a.s.,

$$\lim_{k\to\infty}\int_s^t\int_{U_0}g\left(Z_{r-}^{s,\eta_{n_k}},r,z\right)\widetilde{N}_p(dr,dz)$$
$$=\int_s^t\int_{U_0}g\left(Z_{r-}^{s,\eta},r,z\right)\widetilde{N}_p(dr,dz),\quad \text{uniformly in } t\in[s,T]\,. \tag{3.148}$$

Thus, combining (3.146)-(3.147)-(3.148), we can pass to the limit in (3.145) along a suitable subsequence to get (3.72).

The equalities in (3.73) are obtained using (3.146)-(3.147)-(3.148) and recalling the continuity of $Z_i^{s,x}$ in the space variable x. The lemma is now completely proved. ■

Part II

Chapter 4

Affine Volterra processes with jumps

In this chapter, we delve into the theory of affine processes, which has been recently extended to stochastic Volterra equations with continuous trajectories. These *affine Volterra processes* possibly incorporate path–dependent features which allow to go beyond the Markovian framework. Furthermore, they can have trajectories whose regularity is different from the paths of Brownian motion. More specifically, singular kernels yield rough affine processes. We extend the theory by considering affine stochastic Volterra equations with jumps. This extension poses nontrivial challenges because the jump structure, together with possible singularities of the kernel, may induce explosions of the trajectories. We also provide semi–explicit exponential affine formulas for the conditional Fourier–Laplace transforms of marked Hawkes processes.

Notation Throughout the chapter, elements of $\mathbb{R}^k$ and $\mathbb{C}^k$ are column vectors. Given a matrix $A \in \mathbb{C}^{k\times l}$, the element in row i and column j is A^{ij}, $A^\top \in \mathbb{C}^{l\times k}$ is its transpose matrix, and $|A|$ is the Frobenius norm. We also use the notation $\mathbb{R}^k_+ = \left\{x \in \mathbb{R}^k : x_i \geq 0,\, i = 1,\dots,k\right\}$ and $\mathbb{C}^k_- = \left\{x \in \mathbb{C}^k : \mathrm{Re}\,(x_i) \leq 0,\, i = 1,\dots,k\right\}$, where, for $z \in \mathbb{C}$, $\mathrm{Re}(z)$ denotes its real part. The imaginary part of a complex number z is $\mathrm{Im}\, z$. We adopt the convolution notation $(f * g)(t) = \int_0^t f(t-s)\, g(s)\,\mathrm{d}s$ for functions f, g.

4.1 Preliminaries on affine Volterra processes with jumps

Fix $d, m \in \mathbb{N}$. Let $g_0 \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R}^m)$, $K \in L^2_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}^{m\times d}\right)$ be a matrix–valued kernel and $E \subset \mathbb{R}^m$ be a subset which will be the state–space that we consider. We also introduce a *characteristic triplet* (b, a, η) consisting of the measurable maps $b\colon \mathbb{R}^m \to \mathbb{R}^d$, $a\colon \mathbb{R}^m \to \mathbb{R}^{d\times d}$ and the (positive) transition kernel $\eta(x, \mathrm{d}\xi)$ from $\mathbb{R}^m$ to $\mathbb{R}^d$. We require this triplet to be affine on E, meaning that, for every $x \in E$,

$$b(x) = b_0 + \sum_{k=1}^m x_k b_k, \qquad a(x) = A_0 + \sum_{k=1}^m x_k A_k, \qquad \eta(x, \mathrm{d}\xi) = \nu_0(\mathrm{d}\xi) + \sum_{k=1}^m x_k \nu_k(\mathrm{d}\xi). \tag{4.1}$$

Here $b_0, b_1, \dots, b_m \in \mathbb{R}^d$, $A_0, A_1, \dots, A_m \in \mathbb{R}^{d\times d}$, and $(\nu_k)_{k=0,\dots,m}$ are signed measures on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |\xi|^2 |\nu_k|(\mathrm{d}\xi) < \infty$, with $\nu_k(\{0\}) = 0$. Throughout the chapter, we denote by $X = (X_t)_{t\geq 0}$ a predictable process with trajectories in $L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R}^m)$ and such that $X \in E$, $\mathbb{P}\otimes \mathrm{d}t$–a.e. It is defined on a filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ where the filtration $\mathbb{F}$ satisfies the usual conditions. Moreover, we assume that X solves the following affine stochastic Volterra equation of convolution type

$$X_t = g_0(t) + \int_0^t K(t-s)\,\mathrm{d}Z_s, \quad \mathbb{P}-\text{a.s., for a.e. } t \in \mathbb{R}_+. \tag{4.2}$$

Here Z is a $d-$dimensional semimartingale starting at 0 whose differential characteristics with respect to the Lebesgue measure are $(b(X_t), a(X_t), \eta(X_t, \mathrm{d}\xi))$, $t \geq 0$. These characteristics are taken with respect to the "truncation function" $h(\xi) = \xi$, $\xi \in \mathbb{R}^d$, which can be chosen because Z is a special semimartingale due to [110, Proposition 2.29, Chapter II] and the local integrability of the trajectories of X. In the sequel, we denote by $\mu(\mathrm{d}t, \mathrm{d}\xi)$ the measure associated with the jumps of Z and by $\nu(\mathrm{d}t, \mathrm{d}\xi) = \eta(X_t, \mathrm{d}\xi)\,\mathrm{d}t$ its compensator. We remark that $\nu(\mathrm{d}t, \mathrm{d}\xi)$ is positive, even though ν_k, $k = 0, \dots, m$, is a signed measure and we do not impose any requirements on the sign of the components of $x \in E$. Indeed, from the definition of (positive) transition kernel, $\eta(x, \mathrm{d}\xi)$ is a positive measure on $\mathbb{R}^d$ for every $x \in \mathbb{R}^m$. This is coherent with the fact that $\nu(\mathrm{d}t, \mathrm{d}\xi)$ should be positive, as it is the compensator of the positive jump–measure $\mu(\mathrm{d}t, \mathrm{d}\xi)$.

It is worth discussing the good definition of the stochastic integral in (4.2). Recalling that $X \in E$, $\mathbb{P} \otimes \mathrm{d}t-$a.e., the canonical representation theorem for semimartingales (see [110, Proposition 2.34, Chapter II]) shows that Z admits the decomposition

$$Z_t = \int_0^t b(X_s)\,\mathrm{d}s + M_t^c + M_t^d = b_0 t + \sum_{k=1}^d b_k \int_0^t X_{k,s}\,\mathrm{d}s + M_t^c + M_t^d, \quad t \geq 0,\ \mathbb{P}\text{–a.s.},$$

where $\mathrm{d}M_t^d = \int_{\mathbb{R}^d} \xi\,(\mu - \nu)(\mathrm{d}t, \mathrm{d}\xi)$ is an $\mathbb{R}^d-$valued, purely discontinuous local martingale and M^c is a $d-$dimensional, continuous local martingale satisfying $\mathrm{d}\langle M^c, M^c\rangle_t = a(X_t)\,\mathrm{d}t$. Now if we introduce, for every $j = 1, \dots, d$, the increasing process $C_t^j = \int_0^t \int_{\mathbb{R}^d} |\xi_j|^2\,\nu(\mathrm{d}s, \mathrm{d}\xi)$, $t \geq 0$, then we have

$$C_t^j = \left(\int_{\mathbb{R}^d} |\xi_j|^2\,\nu_0(\mathrm{d}\xi)\right) t + \sum_{k=1}^m \int_{\mathbb{R}^d} |\xi_j|^2\,\nu_k(\mathrm{d}\xi) \left(\int_0^t X_{k,s}\,\mathrm{d}s\right), \quad t \geq 0,\ \mathbb{P}-\text{a.s.}$$

As a consequence of this expression, the local integrability of the paths of X implies that C^j is locally integrable. Hence [110, Theorem 1.33 (a), Chapter II] yields that M^d is a locally square–integrable martingale with

$$\mathrm{d}\left\langle M^{d,j}, M^{d,j}\right\rangle_t = \left[\int_{\mathbb{R}^d} |\xi_j|^2\,\nu_0(\mathrm{d}\xi) + \sum_{k=1}^m \left(\int_{\mathbb{R}^d} |\xi_j|^2\,\nu_k(\mathrm{d}\xi)\right) X_{k,t}\right]\mathrm{d}t, \tag{4.3}$$

where $M^{d,j}$ is the $j-$th component of M^d, $j = 1, \dots, d$. It is convenient to introduce the locally square–integrable martingale $\widetilde{Z} = M^c + M^d$, which satisfies

$$\widetilde{Z}_t = Z_t - \int_0^t b(X_s)\,\mathrm{d}s = Z_t - b_0 t - \sum_{k=1}^d b_k \int_0^t X_{k,s}\,\mathrm{d}s, \quad t \geq 0,\ \mathbb{P}-\text{a.s.} \tag{4.4}$$

Given an integer $l \in \mathbb{N}$ and $F \in L^2_{\text{loc}}\left(\mathbb{R}_+; \mathbb{R}^{l\times d}\right)$, we define the $l-$dimensional random variable

$$\left(F * \mathrm{d}\widetilde{Z}\right)_T = (F * \mathrm{d}M^c)_T + \left(F * \mathrm{d}M^d\right)_T = \int_0^T F(T-s)\,\mathrm{d}M_s^c + \int_0^T F(T-s)\,\mathrm{d}M_s^d.$$

This is well–defined for a.e. $T \in \mathbb{R}_+$. Indeed, consider the stopping times $\tau_n = \inf\left\{t \geq 0 : \int_0^t |X_s|\,\mathrm{d}s > n\right\}$ for all $n \in \mathbb{N}$. Since $X_\cdot(\omega) \in L^1_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^m)$, $\tau_n \to \infty$ as $n \to \infty$ in Ω. Then for every $\overline{T} > 0$, we can apply the Young's type inequality in [3, Lemma A.1] with $p=q=r=1$ and Tonelli's theorem to deduce that

$$\int_0^{\overline{T}}\left(\mathbb{E}\Big[\int_0^{T\wedge\tau_n}|F(T-s)|^2|X_{k,s}|\,\mathrm{d}s\Big]\right)\mathrm{d}T=\int_0^{\overline{T}}\left(\int_0^T|F(T-s)|^2\,\mathbb{E}\left[1_{\{s\le\tau_n\}}|X_{k,s}|\right]\mathrm{d}s\right)\mathrm{d}T$$
$$\le\|F\|^2_{L^2([0,\overline{T}];\mathbb{R}^{l\times d})}\,\mathbb{E}\Big[\int_0^{\overline{T}\wedge\tau_n}|X_{k,s}|\,\mathrm{d}s\Big]\le n\,\|F\|^2_{L^2([0,\overline{T}];\mathbb{R}^{l\times d})}<\infty,\quad k=1,\dots,m.$$

This ensures that $\mathbb{E}\big[\int_0^{T\wedge\tau_n}|F(T-s)|^2X_{k,s}\,\mathrm{d}s\big]<\infty$, $k=1,\dots,m$, $n\in\mathbb{N}$, for a.e. $T\in\mathbb{R}_+$, say for every $T\in\mathbb{R}_+\setminus N$, where $N\subset\mathbb{R}_+$ is a $\mathrm{d}t-$null set. As a consequence, it is straightforward to conclude that the processes

$$\left(\int_0^t F(T-s)\,\mathrm{d}M^c_s\right)_{t\in[0,T]},\qquad\left(\int_0^t F(T-s)\,\mathrm{d}M^d_s\right)_{t\in[0,T]},\tag{4.5}$$

are locally square–integrable martingales for every $T\in\mathbb{R}_+\setminus N$. Indeed, denoting by $M^{c,j}$ the $j-$th component of M^c, $j=1,\dots,d$, for every $n\in\mathbb{N}$ we can write

$$\sum_{j=1}^d\mathbb{E}\Big[\int_0^{T\wedge\tau_n}|F(T-s)|^2\,\mathrm{d}\left\langle M^{c,j},M^{c,j}\right\rangle_s\Big]=\sum_{j=1}^d\mathbb{E}\Big[\int_0^{T\wedge\tau_n}|F(T-s)|^2\left(A_0^{jj}+\sum_{k=1}^m X_{k,s}A_k^{jj}\right)\mathrm{d}s\Big]<\infty,$$

and (by (4.3))

$$\sum_{j=1}^d\mathbb{E}\Big[\int_0^{T\wedge\tau_n}|F(T-s)|^2\,\mathrm{d}\left\langle M^{d,j},M^{d,j}\right\rangle_s\Big]$$
$$=\sum_{j=1}^d\mathbb{E}\Big[\int_0^{T\wedge\tau_n}|F(T-s)|^2\left(\int_{\mathbb{R}^d}|\xi_j|^2\,\nu_0(\mathrm{d}\xi)+\sum_{k=1}^m X_{k,s}\int_{\mathbb{R}^d}|\xi_j|^2\,\nu_k(\mathrm{d}\xi)\right)\mathrm{d}s\Big]<\infty.$$

We always work with a jointly measurable version of the stochastic convolution $F*\mathrm{d}\widetilde{Z}$ defined on $\Omega\times\mathbb{R}_+$ (such a modification exists, see, e.g., [162, Theorem 3.5]).
As for the convolution of F with the drift part of Z, using [96, Theorem 2.2 (i), Chapter 2] we compute

$$\mathbb{E}\Big[\int_0^T\left(\int_0^T 1_{\{t\le\tau_n\}}1_{\{s\le t\}}|F(t-s)|\left(|b_0|+\sum_{k=1}^m|b_k|\,|X_{k,s}|\right)\mathrm{d}s\right)\mathrm{d}t\Big]$$
$$=\mathbb{E}\Big[\int_0^{T\wedge\tau_n}\left(\int_0^t|F(t-s)|\left(|b_0|+\sum_{k=1}^m|b_k|\,|X_{k,s}|\right)\mathrm{d}s\right)\mathrm{d}t\Big]$$
$$\le\|F\|_{L^1([0,T];\mathbb{R}^{l\times d})}\left[|b_0|\,T+n\left(\sum_{k=1}^m|b_k|\right)\right]<\infty,\quad T>0.$$

This shows that there exists a $\mathbb{P}\otimes\mathrm{d}t-$null set $N_1\subset\Omega\times\mathbb{R}_+$ such that the next expression is well–defined:

$$1_{\{t\le\tau_n(\omega)\}}\int_0^t F(t-s)\left(b_0+\sum_{k=1}^m b_kX_{k,s}(\omega)\right)\mathrm{d}s,\quad n\in\mathbb{N},\ (\omega,t)\in(\Omega\times\mathbb{R}_+)\setminus N_1.$$

By Fubini's theorem the resulting processes are jointly measurable in $(\Omega \times \mathbb{R}_+) \setminus N_1$, hence passing to the limit as $n \to \infty$, we obtain the jointly measurable process $\int_0^t F(t-s)\left(b_0 + \sum_{k=1}^m b_k X_{k,s}\right) \mathrm{d}s$ (defined on the same set). Finally we introduce

$$(F * b(X))(\omega, t) = \begin{cases} \int_0^t F(t-s)\left(b_0 + \sum_{k=1}^m b_k X_{k,s}(\omega)\right) \mathrm{d}s, & (\omega, t) \in (\Omega \times \mathbb{R}_+) \setminus N_1, \\ 0, & (\omega, t) \in N_1. \end{cases}$$

This is a jointly measurable process defined on the whole $\Omega \times \mathbb{R}_+$. This machinery for constructing jointly measurable modifications of given processes will be used several times in the sequel.
Overall, the previous argument proves that the integral on the right side of (4.2) is well–defined $\mathbb{P}-$a.s., for a.e. $t \in \mathbb{R}_+$. We denote by $(F * \mathrm{d}Z) = (F * b(X)) + \left(F * \mathrm{d}\widetilde{Z}\right)$; with this notation, Equation (4.2) can be written as follows

$$X = g_0 + (K * \mathrm{d}Z) = g_0 + (K * b(X)) + \left(K * \mathrm{d}\widetilde{Z}\right), \quad \mathbb{P} \otimes \mathrm{d}t - \text{a.e.} \tag{4.6}$$

The following lemma will be useful in the sequel.

Lemma 4.1. *For every $T > 0$,*

$$\mathbb{E}\left[\|X\|_{L^1([0,T];\mathbb{R}^m)}\right] < \infty. \tag{4.7}$$

Proof. The proof follows the same steps as those in [3, Theorem 1.4]. The difference is that the affine structure of our model guaranteed by (4.1) is substituted for [3, Condition (1.5)], and makes the $L^1_{\mathrm{loc}}-$integrability of the paths of X sufficient (instead of the $L^p_{\mathrm{loc}}-$integrability, $p \geq 2$ required in [3]). ■

Knowing the additional property in (4.7), the same argument as the one above (without stopping times) shows that the processes in (4.5) are indeed square–integrable martingales for a.e. $T \in \mathbb{R}_+$.

Remark 4.1. *We refer to [3] for a general solution theory concerning equations of the type in* (4.6) *when $g_0 \in L^p_{loc}(\mathbb{R}_+;\mathbb{R}^m)$, $p \geq 2$, and $E = \mathbb{R}^m$.*
In the case $m = d = 1$ and $E = \mathbb{R}_+$, if one defines $Y_t = \int_0^t X_s \,\mathrm{d}s$, $t \geq 0$, then $Y = (Y_t)_{t \geq 0}$ is a nondecreasing process and an application of [3, Lemma 3.2*] shows*

$$Y_t = \int_0^t g_0(s)\,\mathrm{d}s + \int_0^t K(t-s) Z_s \,\mathrm{d}s = \int_0^t g_0(s)\,\mathrm{d}s + (K * Z)_t, \quad t \geq 0, \mathbb{P} - a.s.$$

This type of stochastic Volterra equations is analyzed in [2] for locally integrable kernels $K \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R})$.

4.1.1 Stochastic convolution for processes with jumps

The goal of this subsection is to develop technical results concerning the stochastic convolution. In particular, we aim to make Lemma 2.1 and Lemma 2.6 in [8] feasible in our context, where we are dealing with discontinuities for Z and, more importantly, with a process X which a priori is not bounded. This requires to modify the statements and the proofs of the aforementioned results, which are crucial for the development of the theory. Such changes are important from a conceptual point of view and after every result we add a remark showing the parallel with the setting in [8].

We start with a preliminary claim.

Lemma 4.2. *Fix $p \in \mathbb{N}$. Let $F, G \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{p\times d})$ and $S \subset \mathbb{R}_+$ be such $\mathbb{R}_+ \setminus S$ is $\mathrm{d}t-$null set. Suppose that $F = G$ a.e. in $\mathbb{R}_+$. Then*

$$\int_0^T F(T-s)\,\mathrm{d}Z_s = \int_0^T 1_S(s)\,G(T-s)\,\mathrm{d}Z_s, \quad \mathbb{P}-a.s., \text{ for a.e. } T \in \mathbb{R}_+. \tag{4.8}$$

In particular,

$$F * \mathrm{d}Z = G * \mathrm{d}Z, \quad \mathbb{P}\otimes \mathrm{d}t - a.e. \tag{4.9}$$

Proof. It is sufficient to prove (4.8) replacing Z with $\widetilde{Z}$, because trivially $F * b(X) = G * (1_S b(X))$, $\mathbb{P}\otimes \mathrm{d}t-$a.e. on $\Omega \times \mathbb{R}_+$. Moreover, we only work with the stochastic integral in $\mathrm{d}M^c$, as by (4.3) we can repeat the next procedure (component–wise) for the convolution in $\mathrm{d}M^d$ recovering (4.8).

The argument above in the section implies the existence of a $\mathrm{d}t-$null set $N \subset \mathbb{R}_+$ such that, for every $T \in \mathbb{R}_+ \setminus N$, we have, $\mathbb{P}-$a.s.,

$$\int_0^T 1_S(s)\,G(T-s)\,\mathrm{d}M^c_s - \int_0^T F(T-s)\,\mathrm{d}M^c_s = \int_0^T \left(1_S(s)\,G(T-s) - F(T-s)\right)\mathrm{d}M^c_s. \tag{4.10}$$

Consider $Q = \left(\int_0^t \left(1_S(s)\,G(T-s) - F(T-s)\right)\mathrm{d}M^c_s\right)_{t\le T}$. Note that Q is a $p-$dimensional, square–integrable martingale whose predictable quadratic covariation is, due to the hypotheses,

$$\langle Q, Q\rangle_t = \int_0^t \left(1_S(s)\,G(T-s) - F(T-s)\right) a(X_s) \left(1_S(s)\,G(T-s) - F(T-s)\right)^\top \mathrm{d}s = 0,$$

which holds for every $t \in [0,T]$, $\mathbb{P}-$a.s. Since Q starts at 0, we can conclude that $Q = 0$ up to evanescence, hence (4.8) follows.

Regarding (4.9), it is an immediate consequence of (4.8) with $S = \mathbb{R}_+$ and the joint measurability of the stochastic convolutions, which allows to state an equality $\mathbb{P}\otimes \mathrm{d}t-$a.e. This completes the proof. ∎

Remark 4.2. *In [8], the authors consider the stochastic convolution of a function $F \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{p\times d})$ with respect to a continuous local martingale M with predictable quadratic covariation $d\langle M, M\rangle_t = a_t\,\mathrm{d}t$, where (a_t) is an adapted, locally bounded process. These assumptions allow to define $(F * \mathrm{d}M)_t$ for every $t \in \mathbb{R}_+$. In particular, two jointly measurable versions of the stochastic convolution are equal $\mathbb{P}-$a.s., for every $t \ge 0$. This concept is stronger than the $\mathbb{P}\otimes \mathrm{d}t-$uniqueness that we have in our framework. As for (4.9) in Lemma 4.2, in the continuous case it can be stated as follows: for every $F, G \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{p\times d})$, with $F = G$ a.e. in $\mathbb{R}_+$, one has*

$$(F * \mathrm{d}M)_t = (G * \mathrm{d}M)_t, \quad \mathbb{P}-a.s.,\ t \ge 0.$$

Now we state a result concerning the associativity of the stochastic convolution.

Lemma 4.3. *Fix $p, q \in \mathbb{N}$. Let $\rho \in L^1_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{q\times p})$ and $F \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{p\times d})$. Then*

$$((\rho * F) * \mathrm{d}Z)_t = (\rho * (F * \mathrm{d}Z))(t), \quad \mathbb{P}-a.s., \text{ for a.e. } t \in \mathbb{R}_+. \tag{4.11}$$

Proof. Also in this case we just need to show the statement with $\mathrm{d}\widetilde{Z}$ in place of $\mathrm{d}Z$, because an application of Fubini's theorem provides $(\rho * F) * b(X) = \rho * (F * b(X))$, $\mathbb{P}\otimes \mathrm{d}t-$a.e. on $\Omega \times \mathbb{R}_+$. In addition it is sufficient to focus only on the stochastic convolutions in $\mathrm{d}M^c$, as discussed in the preceding proof. By linearity we can assume $d = p = q = 1$ without loss of generality, and we consider $\rho \ge 0$ to keep the notation simple, otherwise we should split it into positive and negative part.

First note that the function $\rho * F \in L^2_{\text{loc}}(\mathbb{R}_+;\mathbb{R})$, hence for every $t \in \mathbb{R}_+ \setminus N_1$, being N_1 a d$t-$null set, we have

$$((\rho * F) * \mathrm{d}M^c)_t = \int_0^t \left(1_{\{(t-u)\in S\}} \int_0^{t-u} F(t-u-s)\,\rho(s)\,\mathrm{d}s\right) \mathrm{d}M_u^c$$
$$= \int_0^t \left(\int_0^t 1_{\{(t-u)\in S\}} 1_{\{s\le t-u\}} F(t-u-s)\,\rho(s)\,\mathrm{d}s\right) \mathrm{d}M_u^c, \quad \mathbb{P}-\text{a.s}, \qquad (4.12)$$

where $S \subset \mathbb{R}_+$ is such that $\int_0^t F(t-s)\,\rho(s)\,\mathrm{d}s$, $t \in S$, is well–defined. In particular, $\mathbb{R}_+ \setminus S$ is a d$t-$null set. Our goal is to apply the stochastic Fubini's theorem (see, e.g., [153, Theorem 65, Chapter IV]), but before we can do that we need a preliminary step. For every $T > 0$, a change of variables, sequential applications of Tonelli's theorem and Young's inequality yield (in the whole Ω)

$$\int_0^T \left[\int_0^t \left(\int_0^t 1_{\{(t-u)\in S\}} 1_{\{s\le t-u\}} |F(t-s-u)|^2 \rho(s)\,\mathrm{d}s\right) |X_u|\,\mathrm{d}u\right] \mathrm{d}t$$
$$\le \int_0^T \left[\int_0^t \left(\int_0^{t-s} |F(t-s-u)|^2 |X_u|\,\mathrm{d}u\right) \rho(s)\,\mathrm{d}s\right] \mathrm{d}t = \int_0^T \left[\int_s^T \left(|F|^2 * |X|\right)(t-s)\,\mathrm{d}t\right] \rho(s)\,\mathrm{d}s$$
$$= \int_0^T \left[\int_0^{T-s} \left(|F|^2 * |X|\right)(t)\,\mathrm{d}t\right] \rho(s)\,\mathrm{d}s \le \|\rho\|_{L^1([0,T])} \|F\|^2_{L^2([0,T])} \|X\|_{L^1([0,T])}.$$

Taking expectation and recalling (4.7) we have

$$\int_0^T \mathbb{E}\left[\int_0^t \left(\int_0^t 1_{\{(t-u)\in S\}} 1_{\{s\le t-u\}} |F(t-s-u)|^2 \rho(s)\,\mathrm{d}s\right) |X_u|\,\mathrm{d}u\right] \mathrm{d}t$$
$$\le \|\rho\|_{L^1([0,T])} \|F\|^2_{L^2([0,T])} \mathbb{E}\left[\|X\|_{L^1([0,T])}\right] < \infty.$$

This proves that there exists $N_2 \subset \mathbb{R}_+$ such that

$$\mathbb{E}\left[\int_0^t \left(\int_0^t 1_{\{(t-u)\in S\}} 1_{\{s\le t-u\}} |F(t-s-u)|^2 \rho(s)\,\mathrm{d}s\right) X_u\,\mathrm{d}u\right] < \infty, \quad t \in \mathbb{R}_+ \setminus N_2. \qquad (4.13)$$

Taking $t \in \mathbb{R}_+ \setminus (N_1 \cup N_2)$, thanks to (4.13) and Lemma 4.2 (see (4.8)) we can apply the stochastic Fubini's theorem in (4.12) to deduce that

$$((\rho * F) * \mathrm{d}M^c)_t$$
$$= \int_0^t \left(\int_0^t 1_{\{(t-u)\in S\}} 1_{\{s\le t-u\}} F(t-u-s)\,\rho(s)\mathrm{d}s\right) \mathrm{d}M_u^c = \int_0^t \left(\int_0^{t-s} 1_{\{(t-u)\in S\}} F(t-s-u)\,\mathrm{d}M_u^c\right) \rho(s)\,\mathrm{d}s$$
$$= \int_0^t (F * \mathrm{d}M^c)_{t-s}\,\rho(s)\,\mathrm{d}s = (\rho * (F * \mathrm{d}M^c))(t), \quad \mathbb{P}-\text{a.s.},$$

and the proof is complete. ■

Remark 4.3. *The previous result is the analogue of [8, Lemma 2.1], where the authors are able –in the framework described in Remark 4.2– to handle a generic signed measure of locally bounded variation L. Essentially they can do so because the convolution $F * \mathrm{d}M$ is defined as a stochastic integral for every $t \in \mathbb{R}_+$. As a consequence, it is unique up to a $\mathbb{P} \otimes |L|-$null set, being $|L|$ the total variation measure*

of L.
In contrast with this, notice that in our setting it is not possible to make sense of the right side of (4.11) *for a fixed time $t > 0$ when ρ is replaced by L. Indeed, $F * \mathrm{d}Z$ is only defined up to a $\mathbb{P} \otimes \mathrm{d}t-$null set, therefore the value of $(L * (F * \mathrm{d}Z))(t)$ would depend on the modification one chooses. However, Lemma 4.3 can be slightly extended by replacing ρ in* (4.11) *with an $\mathbb{R}^{q\times p}-$valued measure which is the sum of a locally integrable function and a point mass in* 0 *(this extension can be inferred directly from* (4.11)*). We are going to need this final comment in Section 4.3.*

We are now ready to state an analogue of [8, Lemma 2.6].

Proposition 4.4. *Assume that $m = d$, and that the kernel $K \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{d\times d})$ admits a resolvent of the first kind L*[1]*. Let $F \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{d\times d})$ be such that $F * L$ is locally absolutely continuous. Then*

$$(F * \mathrm{d}Z)_t = (F * L)(0)(X - g_0)(t) + \big((F * L)' * (X - g_0)\big)(t), \quad \text{for a.e. } t \in \mathbb{R}_+, \mathbb{P} - a.s. \tag{4.14}$$

Proof. By Lebesgue's fundamental theorem of calculus we can write (denoting by I the identity matrix in $\mathbb{R}^{d\times d}$)

$$(F * L)(t) = (F * L)(0) + \int_0^t (F * L)'(s)\,\mathrm{d}s = (F * L)(0) + \big((F * L)' * I\big)(t), \quad t \geq 0,$$

which implies, convolving with K, using [96, Theorem 6.1 (ix), Chapter 3] and a change of variables,

$$\int_0^t F(s)\,\mathrm{d}s = (F * L)(0)\int_0^t K(s)\,\mathrm{d}s + \int_0^t \big((F * L)' * K\big)(s)\,\mathrm{d}s, \quad t \geq 0.$$

We can differentiate both sides of the previous equation, as they are absolutely continuous functions, and we obtain

$$F(t) = (F * L)(0)K(t) + \big((F * L)' * K\big)(t), \quad \text{for a.e. } t \in \mathbb{R}_+.$$

Then convolving with $\mathrm{d}Z$ yields

$$\begin{aligned}(F * \mathrm{d}Z)_t &= (F * L)(0)(K * \mathrm{d}Z)_t + \big(\big((F * L)' * K\big) * \mathrm{d}Z\big)_t \\ &= (F * L)(0)(K * \mathrm{d}Z)_t + \big((F * L)' * (K * \mathrm{d}Z)\big)(t), \quad \mathbb{P} - \text{a.s, for a.e. } t \in \mathbb{R}_+, \end{aligned} \tag{4.15}$$

where in the first equality we use Lemma 4.2 (see (4.9)) and in the second Lemma 4.3 with $\rho = (F * L)'$. The crucial point here is to pass to the trajectories. In order to do so, observe that by (4.6) we have

$$X_t - g_0(t) = (K * \mathrm{d}Z)_t, \quad \text{for a.e. } t \in \mathbb{R}_+, \mathbb{P} - \text{a.s.},$$

hence $\big((F * L)' * (K * \mathrm{d}Z)\big)(t) = \big((F * L)' * (X - g_0)\big)(t)$, $\mathbb{P}-$a.s., for a.e. $t \in \mathbb{R}_+$. Moreover we can consider a jointly measurable modification of the process $\big((F * L)' * (X - g_0)\big)$ thanks to Fubini's theorem, which in turn can be applied as

$$\begin{aligned}&\mathbb{E}\Big[\int_0^T \Big(\int_0^T 1_{\{s\leq t\}} \big|(F * L)'(s)\big| \left|(X - g_0)(t - s)\right| \mathrm{d}s\Big)\mathrm{d}t\Big] \\ &\quad = \mathbb{E}\Big[\int_0^T \Big(\int_0^t \big|(F * L)'(s)\big| \left|(X - g_0)(t - s)\right| \mathrm{d}s\Big)\,\mathrm{d}t\Big] \\ &\quad \leq \big\|(F * L)'\big\|_{L^1([0,T];\mathbb{R}^{d\times d})}\Big(\mathbb{E}\Big[\|X\|_{L^1([0,T];\mathbb{R}^d)}\Big] + \|g_0\|_{L^1([0,T];\mathbb{R}^d)}\Big) < \infty, \quad T > 0,\end{aligned} \tag{4.16}$$

---

[1]Given a kernel $K \in L^1_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{d\times d})$, an $\mathbb{R}^{d\times d}-$valued measure L is called its *(measure) resolvent of the first kind* if $L * K = K * L = I$, where $I \in \mathbb{R}^{d\times d}$ is the identity matrix. L does not always exists, but if it does then it is unique (cfr. [96, Theorem 5.2, Chapter 5]).

by Tonelli's theorem, Equation (4.7) and [96, Theorem 2.2 (i), Chapter 2]. Substituting this term in (4.15) and recalling once again (4.6) we deduce that

$$(F * \mathrm{d}Z)_t = (F * L)(0)(X - g_0)(t) + \big((F * L)' * (X - g_0)\big)(t), \quad \mathbb{P} - \text{a.s., for a.e. } t \in \mathbb{R}_+.$$

This equality can be understood up to a $\mathbb{P} \otimes \mathrm{d}t-$null set because it involves only jointly measurable processes. Therefore (4.14) holds true and the proposition is completely proved. ■

Remark 4.4. *In [8, Lemma 2.6] the authors require $F * L$ to be right–continuous and of locally bounded variation. The loss of generality in Proposition 4.4, where we assume the local absolute continuity for the same function, is triggered by Lemma 4.3 and Remark 4.3.*

4.2 Towards the conditional Fourier–Laplace transform

In this section we are going to introduce processes $V^T = \left(V_t^T\right)_{t\in[0,T]}$ which will be used to find an *ansatz* for the conditional Fourier–Laplace transform of $(f^\top * X)(T)$, $T > 0$, where f is a suitable given function. The procedure that we employ can also be adapted to characterize the marginal distributions of the solution process X and of the semimartingale Z in (4.2), see Subsection 4.2.1.

We first introduce some notation. For a $\mathbb{C}-$valued function $g \in L^1(\nu_k)$, $k = 0, 1, \dots, m$, we denote

$$\langle \eta(x, \mathrm{d}\xi), g(\xi)\rangle = \int_{\mathbb{R}^d} g(\xi)\,\nu_0(\mathrm{d}\xi) + \sum_{k=1}^m \left(\int_{\mathbb{R}^d} g(\xi)\,\nu_k(\mathrm{d}\xi)\right) x_k, \quad x \in E;$$

$$\nu(g(\xi)) = \left[\int_{\mathbb{R}^d} g(\xi)\,\nu_1(\mathrm{d}\xi) \quad \int_{\mathbb{R}^d} g(\xi)\,\nu_2(\mathrm{d}\xi) \quad \dots \quad \int_{\mathbb{R}^d} g(\xi)\,\nu_m(\mathrm{d}\xi)\right]^\top \in \mathbb{C}^m.$$

Note that $\langle \eta(x, \mathrm{d}\xi), g(\xi)\rangle = \int_{\mathbb{R}^d} g(\xi)\,\nu_0(\mathrm{d}\xi) + \nu(g(\xi))^\top x$ for every $x \in E$. In addition, we consider

$$B = \begin{bmatrix} b_1 & b_2 & \dots & b_m \end{bmatrix} \in \mathbb{R}^{d\times m},$$

$$A(u) = \begin{bmatrix} u^\top A_1 u & u^\top A_2 u & \dots & u^\top A_m u \end{bmatrix}^\top \in \mathbb{C}^m, \quad u \in \mathbb{C}^d.$$

Notice that $b(x) = b_0 + Bx$, and $u^\top a(x) u = u^\top A_0 u + A(u)^\top x$, for every $x \in E$, $u \in \mathbb{C}^d$.

Take $f \in C(\mathbb{R}_+; \mathbb{C}^m)$ and denote by

$$D_\nu = \left\{u \in \mathbb{C}^d \;:\; \left(e^{u^\top \cdot} - 1 - u^\top \cdot\,\right) \in L^1(\nu_k),\ \text{for every } k = 0, \dots, m\right\}. \tag{4.17}$$

Recalling that the signed measures ν_k, $k = 0, \dots, m$, have finite second moment, one can show that $i\mathbb{R}^d \subset D_\nu$, and that $D_\nu \cap \mathbb{R}^d$ is a star–shaped domain at 0.

Consider a subset $D \subset D_\nu$ where the $\mathbb{C}-$valued functions $u \mapsto \int_{\mathbb{R}^d} \left(e^{u^\top \xi} - 1 - u^\top \xi\right)\nu_k(\mathrm{d}\xi)$ are locally bounded, for every $k = 0, \dots, m$. We define the map $\mathscr{R}\colon \mathbb{R}_+ \times D \to \mathbb{C}^m$ as follows:

$$\mathscr{R}(t, u) = f(t) + B^\top u + \frac{1}{2}A(u) + \nu\left(e^{u^\top \xi} - 1 - u^\top \xi\right), \quad (t, u) \in \mathbb{R}_+ \times D. \tag{4.18}$$

Notice that, by the definition of D, the mapping $\mathscr{R}$ is locally bounded. The following hypothesis introduces the deterministic Riccati–Volterra equation that allows us to determine a semi–explicit exponential affine formula for $\mathbb{E}[\exp\{(f^\top * X)(T)\}|\mathcal{F}_t]$, with $T > 0$ and $t \in [0, T]$.

Hypothesis 4.1. *There exists a continuous global solution $\psi\colon \mathbb{R}_+ \to D$ to the deterministic Riccati–Volterra equation*

$$\psi(t)^\top = \int_0^t \mathcal{R}(s, \psi(s))^\top K(t-s)\,\mathrm{d}s = \Big(\mathcal{R}(\cdot, \psi(\cdot))^\top * K\Big)(t), \quad t \geq 0. \tag{4.19}$$

Since ψ is continuous and takes values in D, the $\mathbb{C}^m$−valued map $s \mapsto \mathcal{R}(s, \psi(s))$ is locally bounded.

We introduce the $\mathbb{C}$−valued function $\phi\colon \mathbb{R}_+ \to \mathbb{C}$ given by

$$\phi(t) = \int_0^t \left(\psi(s)^\top b_0 + \frac{1}{2}\psi(s)^\top A_0 \psi(s) + \int_{\mathbb{R}^d} \left(\mathrm{e}^{\psi(s)^\top \xi} - 1 - \psi(s)^\top \xi\right) \nu_0(\mathrm{d}\xi)\right) \mathrm{d}s, \quad t \geq 0. \tag{4.20}$$

For every $T > 0$ we define the following càdlàg, adapted, $\mathbb{C}$−valued semimartingale on $\Omega \times [0, T]$:

$$\begin{aligned} V_t^T = V_0^T - \int_0^t \left[\frac{1}{2}\psi(T-s)^\top a(X_s)\,\psi(T-s) + \left\langle \eta(X_s, \mathrm{d}\xi), e^{\psi(T-s)^\top \xi} - 1 - \psi(T-s)^\top \xi \right\rangle\right] \mathrm{d}s \\ + \int_0^t \psi(T-s)^\top \mathrm{d}\widetilde{Z}_s, \end{aligned} \tag{4.21}$$

$$\begin{aligned} V_0^T = \int_0^T \Bigg(f(T-s) + B^\top \psi(T-s) + \frac{1}{2} A(\psi(T-s)) \\ + \nu\left(e^{\psi(T-s)^\top \xi} - 1 - \psi(T-s)^\top \xi\right)\Bigg)^\top g_0(s)\,\mathrm{d}s + \phi(T). \end{aligned} \tag{4.22}$$

Observe that V^T is left–continuous in T because $\psi(0) = 0$ by (4.19). This process is the natural extension of [8, Equations (4.4) – (4.5)] to the framework with jumps. Moreover, one can write

$$V_0^T = \phi(T) + \int_0^T \mathcal{R}(T-s, \psi(T-s))^\top g_0(s)\,\mathrm{d}s. \tag{4.23}$$

Our aim is to find, using the stochastic Fubini's theorem, an alternative expression for the random variables V_t^T by means of integrals in time of the trajectories of suitable processes.
In the case $b \equiv 0$, we are going to use the paths of the forward process. Precisely, for a fixed $t \in [0, T]$, by (4.70) in Appendix 4.A we have

$$\mathbb{E}\left[X_s \middle| \mathcal{F}_t\right] = g_0(s) + \int_0^t K(s-r)\,\mathrm{d}\widetilde{Z}_r, \quad \mathbb{P}-\text{a.s., for a.e. } s > t. \tag{4.24}$$

Hence requiring the kernel K to be continuous on $(0, \infty)$, the process on the right side of the previous equation has a jointly measurable version that we denote by $\widetilde{g}_t(s)$, $s > t$. Note that it makes sense to integrate in time the trajectories of such $\widetilde{g}_t(\cdot)$ since it is unique up to a $\mathbb{P} \otimes \mathrm{d}t$−null set.
In the case $b \neq 0$ we consider the paths of a process $g_t(\cdot)$ such that

$$g_t(s) = g_0(s) + \int_0^t K(s-r)\,\mathrm{d}Z_r, \quad \mathbb{P}-\text{a.s., } s > t. \tag{4.25}$$

Also in this case we assume K to be continuous on $(0, \infty)$, so that $g_t(\cdot)$ can be taken jointly measurable on $\Omega \times (t, \infty)$ and is uniquely defined up to a $\mathbb{P} \otimes \mathrm{d}t$−null set. An application of the stochastic Fubini's theorem (see, e.g., [153, Theorem 65, Chapter IV]) shows that the trajectories of $g_t(\cdot)$ are locally integrable in (t,∞). Note that when $t=0$ we have an abuse of notation, as g_0 represents both the initial input curve in (4.2) and the process just defined in (4.25). This, however, is not an issue as these two concepts coincide $\mathbb{P}\otimes \mathrm{d}t-$a.e. in $\Omega\times(0,\infty)$. In the following, we continue to consider g_0 as the initial input curve. Finally, notice that

$$g_t(s)=\mathbb{E}\left[X_s-\int_0^{s-t}K(s-t-r)\,b(X_{t+r})\,\mathrm{d}r\,\middle|\,\mathcal{F}_t\right],\quad \mathbb{P}-\text{a.s., for a.e. } s>t. \tag{4.26}$$

For this reason $g_t(\cdot)$ is called *adjusted forward process.*

Theorem 4.5. *Assume Hypothesis 4.1. Let $K\in L^2_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}^{m\times d}\right)$ be a continuous kernel on $(0,\infty)$ and define, for every $t\in[0,T]$,*

$$\widetilde{V}_t^T=\phi(T-t)+\int_0^t f(T-s)^\top X_s\,\mathrm{d}s+\int_t^T \mathcal{R}(T-s,\psi(T-s))^\top g_t(s)\,\mathrm{d}s. \tag{4.27}$$

Then

$$V_t^T=\widetilde{V}_t^T,\quad \mathbb{P}-a.s.,\ t\in[0,T]. \tag{4.28}$$

In addition, the process $\left(\exp\left\{V_t^T\right\}\right)_{t\in[0,T]}$ is a $\mathbb{C}-$valued local martingale, and if it is a true martingale then

$$\mathbb{E}\left[\exp\left\{\left(f^\top * X\right)(T)\right\}\middle|\mathcal{F}_t\right]=\exp\left\{\widetilde{V}_t^T\right\},\quad \mathbb{P}-a.s.,\ t\in[0,T]. \tag{4.29}$$

Proof. It is straightforward to check that (4.28) holds for $t=0$.

Focusing on the case $t\in(0,T]$, we rewrite the definition of $\widetilde{V}_t^T$ in (4.27) as follows

$$\begin{aligned}\widetilde{V}_t^T&=\phi(T-t)+\int_0^t f(T-s)^\top X_s\,\mathrm{d}s\\&\quad+\int_t^T\mathcal{R}(T-s,\psi(T-s))^\top g_0(s)\,\mathrm{d}s+\int_t^T\mathcal{R}(T-s,\psi(T-s))^\top(g_t-g_0)(s)\,\mathrm{d}s.\end{aligned} \tag{4.30}$$

It is convenient to introduce the process

$$\bar{g}_t(s)=\begin{cases}X_s, & s\le t\\ g_t(s), & s>t\end{cases}. \tag{4.31}$$

Recall that by (4.25) $g_t(s)=g_0(s)+\int_0^t K(s-r)\,\mathrm{d}Z_r$, $\mathbb{P}-$a.s. for a.e. $s>t$, and that by (4.6) $X_s=g_0(s)+\int_0^s K(s-r)\,\mathrm{d}Z_r$, $\mathbb{P}-$a.s., for a.e. $s\in[0,t]$. Therefore $\bar{g}_t(\cdot)$ is a jointly measurable modification of the process $g_0(\cdot)+\int_0^t 1_{\{r\le\cdot\}}K(\cdot-r)\,\mathrm{d}Z_r$. Invoking the stochastic Fubini's theorem in [153, Theorem 65, Chapter IV] and recalling the Riccati–Volterra equation in (4.19), after a suitable change of variables we obtain

$$\begin{aligned}&\int_0^T\mathcal{R}(T-s,\psi(T-s))^\top(\bar{g}_t-g_0)(s)\,\mathrm{d}s\\&=\int_0^T\mathcal{R}(T-s,\psi(T-s))^\top\left[\int_0^t 1_{\{r\le s\}}K(s-r)\,\mathrm{d}Z_r\right]\mathrm{d}s\\&=\int_0^t\left[\int_r^T\mathcal{R}(T-s,\psi(T-s))^\top K(s-r)\,\mathrm{d}s\right]\mathrm{d}Z_r=\int_0^t\left[\int_0^{T-r}\mathcal{R}(s,\psi(s))^\top K(T-r-s)\,\mathrm{d}s\right]\mathrm{d}Z_r\\&=\int_0^t\psi(T-r)^\top\,\mathrm{d}Z_r,\quad \mathbb{P}-\text{a.s.}\end{aligned} \tag{4.32}$$

Such an application is legitimate, because the boundedness of $\mathscr{R}(\cdot,\psi(\cdot))$ in $[0,T]$ by a positive constant C_T and a change of variables yield, for every $k=1,\dots,m$,

$$\begin{aligned}
&\int_0^t\left[\int_0^T 1_{\{r\le s\}}\left|\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)\right|^2\left|K\left(s-r\right)\right|^2\mathrm{d}s\right]\left|X_{k,r}\right|\mathrm{d}r\\
&\qquad=\int_0^t\left[\int_0^{T-r}\left|\mathscr{R}\left(s,\psi\left(s\right)\right)\right|^2\left|K\left(T-r-s\right)\right|^2\mathrm{d}s\right]\left|X_{k,r}\right|\,\mathrm{d}r\\
&\qquad\le C_T^2\left\|K\right\|^2_{L^2\left([0,T];\mathbb{R}^{m\times d}\right)}\left\|X\right\|_{L^1([0,t];\mathbb{R}^m)}\,,
\end{aligned}\tag{4.33}$$

so the expectation of the leftmost side is finite thanks to (4.7). As for the drift part,

$$\begin{aligned}
&\int_0^t\left(\int_0^T 1_{\{r\le s\}}|\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)|^2|K\left(s-r\right)|^2\,\mathrm{d}s\right)^{\frac{1}{2}}\left|X_{k,r}\right|\mathrm{d}r\\
&\qquad\qquad\le C_T\left\|K\right\|_{L^2\left([0,T];\mathbb{R}^{m\times d}\right)}\left\|X\right\|_{L^1([0,t];\mathbb{R}^m)}\,.
\end{aligned}\tag{4.34}$$

Going back to (4.30) and recalling the definitions of V^T in (4.21)–(4.23) we obtain, $\mathbb{P}$–a.s.,

$$\begin{aligned}
\widetilde{V}_t^T&=\phi\left(T-t\right)+\int_0^t f\left(T-s\right)^\top X_s\,\mathrm{d}s+\int_t^T\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)^\top g_0\left(s\right)\mathrm{d}s\\
&\quad+\int_0^t\psi\left(T-s\right)^\top\mathrm{d}Z_s-\int_0^t\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)^\top\left(X_s-g_0\left(s\right)\right)\mathrm{d}s\\
&=\phi\left(T-t\right)+\int_0^T\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)^\top g_0\left(s\right)\mathrm{d}s+\int_0^t\psi\left(T-s\right)^\top\mathrm{d}Z_s\\
&\quad-\int_0^t\left[B^\top\psi\left(T-s\right)+\frac{1}{2}A\left(\psi\left(T-s\right)\right)+\nu\left(e^{\psi(T-s)^\top\xi}-1-\psi\left(T-s\right)^\top\xi\right)\right]^\top X_s\,\mathrm{d}s\\
&=\phi\left(T\right)+\int_0^T\mathscr{R}\left(T-s,\psi\left(T-s\right)\right)^\top g_0\left(s\right)+\int_0^t\psi\left(T-s\right)^\top\mathrm{d}\widetilde{Z}_s\\
&\quad-\int_0^t\left[\frac{1}{2}\psi\left(T-s\right)^\top a\left(X_s\right)\psi\left(T-s\right)+\left\langle\eta\left(X_s,\mathrm{d}\xi\right),e^{\psi(T-s)^\top\xi}-1-\psi\left(T-s\right)^\top\xi\right\rangle\right]\mathrm{d}s\\
&=V_t^T,
\end{aligned}\tag{4.35}$$

where in the second–to–last equality we use (4.4). This proves (4.28).

Moving on to the next assertion, denote by $H^T=\left(H_t^T\right)_{t\in[0,T]}=\left(\exp\left\{V_t^T\right\}\right)_{t\in[0,T]}$. By Itô's formula and the dynamics in (4.21) we have

$$\begin{aligned}
\mathrm{d}H_t^T&=H_{t-}^T\\
&\times\left[-\left(\frac{1}{2}\psi\left(T-t\right)^\top a\left(X_t\right)\psi\left(T-t\right)+\left\langle\eta\left(X_t,\mathrm{d}\xi\right),e^{\psi(T-t)^\top\xi}-1-\psi\left(T-t\right)^\top\xi\right\rangle\right)\mathrm{d}t+\psi\left(T-t\right)^\top\mathrm{d}\widetilde{Z}_t\right]\\
&+\frac{1}{2}H_{t-}^T\psi\left(T-t\right)^\top a\left(X_t\right)\psi\left(T-t\right)\mathrm{d}t+H_{t-}^T\int_{\mathbb{R}^d}\left(e^{\psi(T-t)^\top\xi}-1-\psi\left(T-t\right)^\top\xi\right)\mu\left(\mathrm{d}t,\mathrm{d}\xi\right)\\
&=H_{t-}^T\left[\psi\left(T-t\right)^\top\mathrm{d}M_t^c+\int_{\mathbb{R}^d}\left(e^{\psi(T-t)^\top\xi}-1\right)\left(\mu-\nu\right)\left(\mathrm{d}t,\mathrm{d}\xi\right)\right],\quad H_0^T=\exp\left(V_0^T\right).
\end{aligned}$$

We define $N^T = (N_t^T)_{t\in[0,T]}$ by $\mathrm{d}N_t^T = \psi(T-t)^\top \mathrm{d}M_t^c + \int_{\mathbb{R}^d}\left(e^{\psi(T-t)^\top\xi}-1\right)(\mu-\nu)(\mathrm{d}t,\mathrm{d}\xi)$, $N_0^T=0$. Then N^T is a local martingale and the previous computations show that $H^T = \exp\{V_0^T\}\mathcal{E}(N^T)$ up to evanescence, where $\mathcal{E}$ denotes the Doléans–Dade exponential. Therefore H^T is a local martingale, as stated. Finally, in case it is a true martingale, (4.29) directly follows from (4.28), and the proof is complete. ■

Remark 4.5. *Assuming $m=d$, it is possible to find an expression for V^T in terms of the* true *forward process even in the case $b\neq 0$. Indeed, by (4.73) in Appendix 4.A*

$$\mathbb{E}\left[X_s\middle|\mathcal{F}_t\right] = \left(g_0 - (R_B * g_0) + (E_B * b_0)\right)(s) + \int_0^t E_B(s-r)\,\mathrm{d}\widetilde{Z}_r, \quad \mathbb{P}-a.s., \text{ for a.e. } s>t. \tag{4.36}$$

Here R_B is the resolvent of the second kind [2] *of $-KB$ and $E_B = K - R_B * K$. If K is continuous on $(0,\infty)$, then E_B is continuous on the same interval, as well. Thus, one can choose a jointly measurable version $f_t(s)$, $s>t$, of the process on the right side of (4.36), which is unique up to a $\mathbb{P}\otimes\mathrm{d}t-$null set. Arguing as in [8, Lemma 4.4], we obtain the variation of constants formula*

$$\psi(t)^\top = \int_0^t \left[f(s) + \frac{1}{2}A(\psi(s)) + \nu\left(e^{\psi(s)^\top\xi} - 1 - \psi(s)^\top\xi\right)\right]^\top E_B(t-s)\,\mathrm{d}s, \quad t\geq 0,$$

which combined with the strategy in the proof of Theorem 4.5 leads to

$$\begin{aligned} V_t^T = \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_t^T \bigg[& \left(\mathcal{R}(T-s,\psi(T-s)) - B^\top\psi(T-s)\right)^\top f_t(s) \\ & + \frac{1}{2}\psi(T-s)^\top A_0\psi(T-s) + \int_{\mathbb{R}^d}\left(e^{\psi(T-s)^\top\xi} - 1 - \psi(T-s)^\top\xi\right)\nu_0(\mathrm{d}\xi)\bigg]\mathrm{d}s, \quad \mathbb{P}-a.s. \end{aligned}$$

However in the framework of jumps it is preferable to work with the adjusted forward process, because –as will become clear in the next section– certain properties can be assumed for the kernel K, but they can be neither required (i.e. it would not be a reasonable hypothesis) nor inferred for E_B.

4.2.1 The marginal distributions of X and Z

The procedure that we have used above to obtain a formula for $\mathbb{E}[\exp\{(f^\top * X)(T)\}|\mathcal{F}_t]$ (see (4.29) in Theorem 4.5) can also be followed to deduce an exponential affine expression for the conditional Fourier–Laplace transform of the marginal distributions of the solution process X and the semimartingale Z in (4.2), i.e.,

$$\mathbb{E}[\exp\{u_1^\top Z_T\}|\mathcal{F}_t], \qquad \mathbb{E}[\exp\{u_2^\top X_T\}|\mathcal{F}_t],$$

for suitable $u_1\in\mathbb{C}^d$ and $u_2\in\mathbb{C}^m$.

We start off by showing the formula for the conditional transform of the $\mathbb{R}^d-$valued càdlàg semimartingale Z.

[2] Given $K\in L^1_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}^{d\times d}\right)$, its *resolvent of the second kind* is the unique solution $R\in L^1_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}^{d\times d}\right)$ of the two equations $K*R = R*K = K-R$ (cfr. [96, Theorem 3.1, Chapter 2] and the subsequent definition).

Theorem 4.6. *Assume that* $K \in L^2_{\text{loc}}(\mathbb{R}_+;\mathbb{R}^{m\times d})$ *is continuous in* $(0,\infty)$. *Given* $u_1 \in \mathbb{C}^d$ *and* $f \in C(\mathbb{R}_+;\mathbb{C}^m)$, *suppose that there exists a continuous global solution* $\psi_1 \colon \mathbb{R}_+ \to D$ *of the deterministic Riccati–Volterra equation*

$$\psi_1(t)^\top = u_1^\top + \int_0^t \mathcal{R}(s,\psi_1(s))^\top K(t-s)\,\mathrm{d}s, \quad t \geq 0,$$

where $\mathcal{R}\colon \mathbb{R}_+ \times D \to \mathbb{C}^m$ *is given in* (4.18). *Then, for every* $T > 0$, *defining* $\widetilde{V}^T = (\widetilde{V}^T_t)_{t\in[0,T]}$ *as in* (4.27) *with* ψ_1 *instead of* ψ *and* $\widetilde{V}^T_1 = \left(\widetilde{V}^T_{1,t}\right)_{t\in[0,T]}$ *by*

$$\widetilde{V}^T_{1,t} = \widetilde{V}^T_t + u_1^\top Z_t, \quad t \in [0,T],$$

the process $\exp\{\widetilde{V}^T_1\}$ *is a local martingale. In particular, if* $\exp\{\widetilde{V}^T_1\}$ *is a true martingale, then*

$$\mathbb{E}\left[\exp\left\{u_1^\top Z_T + (f^\top * X)(T)\right\}\middle|\mathcal{F}_t\right] = \exp\left\{\widetilde{V}^T_{1,t}\right\}, \quad \mathbb{P}-a.s.,\ t\in[0,T].$$

Proof. We use the same argument as in the proof of Theorem 4.5. It can be split into two parts.

In the first step, we define the process $V^T_1 = \left(V^T_{1,t}\right)_{t\in[0,T]}$ as in (4.21)-(4.23), replacing ψ with ψ_1. Then, using the stochastic Fubini's theorem, we prove that V^T_1 and $\widetilde{V}^T_1$ are versions of each other.

In the second step, we use Itô's formula to infer that $\exp\{V^T_1\}$ is a local martingale, since it can be seen as the Doléans–Dade exponential of a local martingale. ■

We now fix an i.i.d. sequence $(Y_{i,k})_k$ of $\mathcal{F}_0-$measurable, $\mathbb{R}-$valued random variables distributed according to a probability measure θ_i on $\mathbb{R}$ with finite first and second moments, for every $i = 1,\dots,d$. Denoting by $(e_i)_{i=1,\dots,d}$ the canonical basis of $\mathbb{R}^d$, we extend θ_i to a probability measure on $\mathbb{R}^d$ by setting

$$\bar{\theta}_i(A) = \theta_i\left(\pi_i\left(A\cap \operatorname{span}(e_i)\right)\right), \quad A \in \mathcal{B}\left(\mathbb{R}^d\right),$$

where $\pi_i\colon \mathbb{R}^d \to \mathbb{R}$ is defined by $\pi_i(x) = x^\top e_i$ and $\mathcal{B}\left(\mathbb{R}^d\right)$ are the Borel–measurable sets of $\mathbb{R}^d$. We also take an increasing sequence $(T_{i,k})_{k\in\mathbb{N}}$ of $\mathbb{F}-$stopping times independent from $(Y_{i,k})_k$ and such that $T_{i,k} \to \infty$ as $k\to\infty$, $\mathbb{P}-$a.s., for every $i = 1,\dots,d$.
We focus on a $d-$dimensional marked point process $N = (N_t)_{t\geq 0} = [N_{1,t},\dots,N_{d,t}]^\top_{t\geq 0}$ with jump times $(T_{i,k})_{k\in\mathbb{N}}$ and jump sizes $(Y_{i,k})_{k\in\mathbb{N}}$, namely,

$$N_{i,t} = \sum_{t\geq T_{i,k}} Y_{i,k} = \sum_{k\in\mathbb{N}} Y_{i,k} 1_{\{t\geq T_{i,k}\}}, \quad t\geq 0,\ i=1,\dots,d.$$

Given a vector $\lambda_0 \in \mathbb{R}^m$ and two matrices $\overline{\Lambda}_0, \overline{\Lambda}_1 \in \mathbb{R}^{d\times m}$, we suppose that the conditional intensity of N is $\overline{\Lambda}_1\lambda$, where $\lambda = (\lambda_t)_{t\geq 0} = [\lambda_{1,t},\dots,\lambda_{m,t}]^\top_{t\geq 0}$ is an $\mathbb{R}^m-$valued process satisfying

$$\lambda_t = \lambda_0 + \int_0^t K(t-s)\overline{\Lambda}_0\lambda_s \mathrm{d}s + \sum_{i=1}^d \sum_{t>T_{i,k}} Y_{i,k}K(t-T_{i,k})e_i, \quad t\geq 0. \tag{4.37}$$

This means that, denoting by $\mu^N(\mathrm{d}t,\mathrm{d}\xi)$ the measure on $\mathbb{R}_+\times\mathbb{R}^d$ associated with the jumps of N and by $\nu^N(\mathrm{d}t,\mathrm{d}\xi)$ its compensator, then

$$\nu^N(\mathrm{d}t,\mathrm{d}\xi) = \sum_{i=1}^d \left(\overline{\Lambda}_1\lambda_t\right)_i \mathrm{d}t\otimes\bar{\theta}_i(\mathrm{d}\xi). \tag{4.38}$$

We call N a *marked Hawkes process* in $\mathbb{R}^d$. In particular, note that

$$N_{i,t} - \int_0^t \int_{\mathbb{R}^d} \xi_i \, \nu^N(\mathrm{d}s, \mathrm{d}\xi) = N_{i,t} - \left(\int_{\mathbb{R}} \xi_i \theta_i(\mathrm{d}\xi_i) \right) \int_0^t \left(\overline{\Lambda}_1 \lambda_s \right)_i \mathrm{d}s, \quad t \geq 0, \, i = 1, \dots, d,$$

are local martingales. Consequently, the $i-$th component N_i of N is a 1−dimensional marked point process with intensity $(\overline{\Lambda}_1 \lambda)_i$ and marks $Y_{i,k} \sim \theta_i$, $k \in \mathbb{N}$, for every $i = 1, \dots, d$.
We define the process

$$\overline{N} = N + \int_0^{\cdot} \overline{\Lambda}_0 \lambda_s \, \mathrm{d}s$$

and the measures

$$\widetilde{\theta}_i \, (\mathrm{d}\xi) = \sum_{j=1}^{d} \overline{\Lambda}_1^{ji} \bar{\theta}_j(\mathrm{d}\xi), \quad i = 1, \dots, m; \tag{4.39}$$

in the sequel, for a function $g \in L^1(\widetilde{\theta}_i)$, $i = 1, \dots, m$, we write

$$\widetilde{\theta}(g(\xi)) = \left[\int_{\mathbb{R}^d} g(\xi) \, \widetilde{\theta_1}(\mathrm{d}\xi) \quad \dots \quad \int_{\mathbb{R}^d} g(\xi) \, \widetilde{\theta_m}(\mathrm{d}\xi) \right]^\top \in \mathbb{C}^m.$$

By analogy with (4.17), we introduce the set

$$D_{\widetilde{\theta}} = \left\{ u \in \mathbb{C}^d \; : \; \left(e^{u^\top \cdot} - 1 - u^\top \cdot \right) \in L^1(\widetilde{\theta}_k), \text{ for every } k = 1, \dots, m \right\},$$

and we consider a subset $\overline{D} \subset D_{\widetilde{\theta}}$ where the $\mathbb{C}-$valued functions $u \mapsto \int_{\mathbb{R}^d} \left(e^{u^\top \xi} - 1 - u^\top \xi \right) \widetilde{\theta}_k(\mathrm{d}\xi)$ are locally bounded, for every $k = 1, \dots, m$. In this framework, Theorem 4.6 is relevant because it provides a formula for the conditional Fourier–Laplace transform of N and $\overline{N}$. More precisely, the following corollary holds.

Corollary 4.7. *Assume that the kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^{m \times d})$ is continuous in $(0, \infty)$. Given $u_1 \in \mathbb{C}^d$ and $f \in C(\mathbb{R}_+; \mathbb{C}^m)$, suppose that there exists a continuous global solution $\bar{\psi}_1 \colon \mathbb{R}_+ \to \overline{D}$ of the deterministic Riccati–Volterra equation*

$$\bar{\psi}_1 \, (t)^\top = u_1^\top + \int_0^t \overline{\mathcal{R}}_1 \left(s, \bar{\psi}_1 \, (s) \right)^\top K \, (t - s) \, \mathrm{d}s, \quad t \geq 0, \tag{4.40}$$

where $\overline{\mathcal{R}}_1 \colon \mathbb{R}_+ \times \overline{D} \to \mathbb{C}^m$ is defined by

$$\overline{\mathcal{R}}_1(t, u) = f(t) + \overline{\Lambda}_0^\top u + \widetilde{\theta} \left(e^{u^\top \xi} - 1 \right).$$

For every $T > 0$, consider the process $\overline{V}_1^T = \left(\overline{V}_{1,t}^T \right)_{t \in [0,T]}$ given by

$$\overline{V}_{1,t}^T = u_1^\top \overline{N}_t + \int_0^t f(T - s)^\top \lambda_s \, \mathrm{d}s + \int_t^T \overline{\mathcal{R}}_1(T - s, \bar{\psi}_1(T - s))^\top \tilde{g}_t \, (s) \, \mathrm{d}s, \quad t \in [0, T],$$

where $\tilde{g}_t(\cdot)$ is defined as in (4.25) with $\overline{N}$ instead of Z. Then the process $\exp\{\overline{V}_1^T\}$ is a local martingale. In particular, if $\exp\{\overline{V}_1^T\}$ is a true martingale, then

$$\mathbb{E} \left[\exp \left\{ u_1^\top \overline{N}_T + (f^\top * \lambda)(T) \right\} \middle| \mathcal{F}_t \right] = \exp \left\{ \overline{V}_{1,t}^T \right\}, \quad \mathbb{P} - a.s., \, t \in [0, T] \,. \tag{4.41}$$

Proof. We first observe that the differential characteristics of the special semimartingale $\overline{N}$ with respect to the "truncation function" $h(\xi) = \xi$ are

$$\left(\left(\overline{\Lambda}_0 + \operatorname{diag}\left[\int_{\mathbb{R}} \xi_1 \theta_1(\mathrm{d}\xi_1), \dots, \int_{\mathbb{R}} \xi_d \theta_d(\mathrm{d}\xi_d)\right] \overline{\Lambda}_1\right) \lambda_t, \quad 0, \quad \sum_{i=1}^m \lambda_{i,t} \widetilde{\theta}_i(\mathrm{d}\xi)\right).$$

Indeed, by (4.39), the compensator $\nu^{\overline{N}}(\mathrm{d}t, \mathrm{d}\xi)$ of the jump–measure of $\overline{N}$, which coincides with $\nu^N(\mathrm{d}t, \mathrm{d}\xi)$ in (4.38), can be rewritten as

$$\nu^{\overline{N}}(\mathrm{d}t, \mathrm{d}\xi) = \sum_{i=1}^d \left(\overline{\Lambda}_1 \lambda_t\right)_i \mathrm{d}t \otimes \bar{\theta}_i(\mathrm{d}\xi) = \sum_{i=1}^m \lambda_{i,t} \, \mathrm{d}t \otimes \widetilde{\theta}_i(\mathrm{d}\xi).$$

By (4.37) we infer that λ is a càg and adapted, hence predictable process with locally (square–)integrable trajectories in the state space $E = \mathbb{R}^m$. In addition, since

$$\sum_{t > T_{i,k}} K(t - T_{i,k}) = \sum_{t \geq T_{i,k}} K(t - T_{i,k}), \quad t \in \mathbb{R}_+ \setminus \{T_{i,k}, \, k \in \mathbb{N}\}, \, \mathbb{P}\text{–a.s.}, \, i = 1, \dots, d,$$

then λ satisfies the following stochastic Volterra equation of convolution type with jumps (cfr. (2.50)):

$$\lambda = \lambda_0 + \left(K * \mathrm{d}\overline{N}\right), \quad \mathbb{P} \otimes \mathrm{d}t - \text{a.e.}$$

Thus, the statements of the corollary are an immediate consequence of Theorem 4.6, obtained by replacing Z with $\overline{N}$ and, respectively, B, A and $\nu(e^{\cdot^\top \xi} - 1 - \cdot^\top \xi)$ in the definition of $\mathscr{R}$ in (4.18) with

$$\overline{\Lambda}_0 + \operatorname{diag}\left[\int_{\mathbb{R}} \xi_1 \theta_1(\mathrm{d}\xi_1), \dots, \int_{\mathbb{R}} \xi_d \theta_d(\mathrm{d}\xi_d)\right] \overline{\Lambda}_1, \quad 0, \quad \widetilde{\theta}(e^{\cdot^\top \xi} - 1 - \cdot^\top \xi).$$

This completes the proof. ∎

Remark 4.6. *To the best of our knowledge, Corollary 4.7 contains an original expression for the conditional Fourier–Laplace transform of the Hawkes process N, see* (4.41) *with $f = -\overline{\Lambda}_0^\top u_1$. Comparing our findings to the existing literature, we note that in the one–dimensional case, results akin to Equation* (4.41) *can be found in [92]. To be precise, in [92, Theorem 3.1] the authors, albeit concentrating on other quantities, outline a procedure that can be adapted to derive an expression for $\mathbb{E}[\exp\{u_1 N_T\} | \mathcal{F}_t]$. Here, $N = (N_t)_{t \geq 0}$ is a Hawkes process with intensity $\lambda = \lambda_0 + K * \mathrm{d}\widetilde{N}$, where $\widetilde{N}$ represents the compensated process of N. Notably, this formula coincides with Equation* (4.41) *when we set $\overline{\Lambda}_0 = -\int_{\mathbb{R}} \xi \theta(\mathrm{d}\xi)$, $\overline{\Lambda}_1 = 1$ and $f = u_1 \int_{\mathbb{R}} \xi \theta(\mathrm{d}\xi)$, although the Riccati–Volterra equations in [92] have a different formulation from* (4.41)*, see for instance [92, Equation (3.4)]. In particular, one retrieves a solution to* (4.41) *by convolving a solution of the corresponding equation in [92] with the kernel K. Thus, even within this restricted framework, Corollary 4.7 stands as an original contribution because it covers Hawkes processes with more general intensities, e.g., $\lambda = \lambda_0 + K * \mathrm{d}N$.*

We now analyze the conditional distributions of the solution process X of (4.6). In particular, in Theorem 4.8 we obtain an expression for $\mathbb{E}[\exp\{u_2^\top X_T\} | \mathcal{F}_t]$, for some $u_2 \in \mathbb{C}^m$. However, since X is defined up to a $\mathbb{P} \otimes \mathrm{d}t$–null set, such a formula will be meaningful only for a.e. $T > 0$. This, in conjunction with the fact that we may be dealing with unbounded solutions of Riccati–Volterra equations in the new setting, makes the applicability of the arguments in Theorem 4.5 more delicate, as detailed in the proof of Theorem 4.8.

Theorem 4.8. *Assume that the kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R}^{m\times d})$ is continuous in $(0,\infty)$. Given $u_2 \in \mathbb{C}^m$ and $f \in C(\mathbb{R}_+;\mathbb{C}^m)$, suppose that there exists a locally square–integrable function $\psi_2 \colon \mathbb{R}_+ \to D_\nu$ such that the mappings $s \mapsto \int_{\mathbb{R}^d}\big(e^{\psi(s)^\top\xi} - 1 - \psi(s)^\top\xi\big)\nu_k(\mathrm{d}\xi)$, $k = 0,\dots,m$, are locally integrable on $\mathbb{R}_+$ and that the following deterministic Riccati–Volterra equation is satisfied:*

$$\psi_2\,(t)^\top = u_2^\top K + \int_0^t \mathcal{R}\,(s,\psi_2\,(s))^\top\,K\,(t-s)\,\mathrm{d}s, \quad \textit{for a.e. } t \geq 0,$$

where D_ν is given in (4.17) and $\mathcal{R}\colon \mathbb{R}_+ \times D_\nu \to \mathbb{C}^m$ in (4.18). Then, for a.e. $T > 0$, defining $\widetilde{V}^T = (\widetilde{V}^T_t)_{t\in[0,T]}$ as in (4.27) with ψ_2 instead of ψ and $\widetilde{V}^T_2 = \big(\widetilde{V}^T_{2,t}\big)_{t\in[0,T]}$ by

$$\widetilde{V}^T_{2,t} = \widetilde{V}^T_t + u_2^\top \int_0^t K(T-s)\,\mathrm{d}Z_s, \quad t \in [0,T], \tag{4.42}$$

the process $\exp\{\widetilde{V}^T_2\}$ is a local martingale. In particular, if $\exp\{\widetilde{V}^T_2\}$ is a true martingale, then

$$\mathbb{E}\left[\exp\left\{u_2^\top X_T + (f*X)(T)\right\}\Big|\mathcal{F}_t\right] = \exp\left\{\widetilde{V}^T_{2,t}\right\}, \quad \mathbb{P}-a.s.,\ t\in[0,T]\,. \tag{4.43}$$

Proof. In this proof, we highlight the main changes that need to be implemented into the arguments employed in Theorem 4.5.

Notice that, contrary to before, we work with a solution ψ_2 of (4.8) which is supposed to be only square–integrable, with potential explosions precluding the boundedness on compact sets of $\mathcal{R}(\cdot,\psi_2(\cdot))$ and $\int_{\mathbb{R}^d}\big(e^{\psi_2(\cdot)^\top\xi} - 1 - \psi_2(\cdot)^\top\xi\big)\nu_k(\mathrm{d}\xi)$, $k = 0,\dots,m$. As a consequence, recalling also the construction of the stochastic convolution in Section 4.1, we can define the process $V^T_2 = (V^T_{2,t})_{t\in[0,T]}$ as in (4.21)-(4.23) with ψ_2 instead of ψ only for a.e. $T > 0$.
We now discuss the good definition of the process $\widetilde{V}^T = (\widetilde{V}^T_t)_{t\in[0,T]}$ in (4.27) (with ψ replaced by ψ_2) for a.e. $T > 0$. By (4.26), the conditional Jensen's inequality, the law of total expectation and the change of variables $r' = r + t$, we observe that, for a.e. $s > t$,

$$\begin{aligned}
1_{\{s>t\}}\mathbb{E}\left[|g_t(s)|\right] &\leq c\,\mathbb{E}\left[|X_s| + 1_{\{s>t\}}\int_0^{s-t}|K(s-t-r)|\,|b(X_{t+r})|\,\mathrm{d}r\right]\\
&\leq c\left(\mathbb{E}\left[|X_s|\right] + \int_0^s \left|K(s-r')\right|\mathbb{E}\left[|b\left(X_{r'}\right)|\right]\mathrm{d}r'\right)\\
&\leq c\left(\mathbb{E}\left[|X_s|\right] + \|K\|_{L^1([0,s];\mathbb{R}^{m\times d})} + (|K| * \mathbb{E}[|X_\cdot|])(s)\right),
\end{aligned} \tag{4.44}$$

for some constant $c = c(d,m,b_0,\dots,b_m) > 0$ allowed to change from line to line. This estimate enables us to show that $\int_t^T \mathcal{R}\,(T-s,\psi_2\,(T-s))^\top\,g_t\,(s)\,\mathrm{d}s$, which appears in the definition of $\widetilde{V}^T_t$ in (4.27), exists $\mathbb{P}-$a.s. in $\mathbb{C}$, for every $t\in[0,T]$, for a.e. $T > 0$. Indeed, recalling Lemma 4.1, for every $\overline{T} > 0$ we have

$$\begin{aligned}
&\int_0^{\overline{T}}\left(\int_0^T \left|\mathcal{R}\,(T-s,\psi_2\,(T-s))\right|\left(\mathbb{E}\left[|X_s|\right] + (|K| * \mathbb{E}[|X_\cdot|])(s)\right)\mathrm{d}s\right)\mathrm{d}T\\
&\qquad\leq \|\mathcal{R}(\cdot,\psi_2(\cdot))\|_{L^1([0,\overline{T}];\mathbb{C}^m)}\,\mathbb{E}\left[\|X\|_{L^1([0,\overline{T}];\mathbb{R}^m)}\right]\left(1 + \|K\|_{L^1([0,\overline{T}];\mathbb{R}^{m\times d})}\right) < \infty,
\end{aligned}$$

hence

$$\int_0^T \left|\mathcal{R}\,(T-s,\psi_2\,(T-s))\right|\left(\mathbb{E}\left[|X_s|\right] + (|K| * \mathbb{E}[|X_\cdot|])(s)\right)\mathrm{d}s < \infty, \quad \text{for a.e. } T > 0.$$

Combining this equation with (4.44) we conclude that, for a.e. $T>0$, for every $t\in[0,T]$,

$$\int_0^T 1_{\{s>t\}}\left|\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right|\mathbb{E}\left[\left|g_t(s)\right|\right]\mathrm{d}s<\infty. \tag{4.45}$$

Thus, $\widetilde{V}^T$ is well defined for a.e. $T>0$. Moreover, as $K\in L^2_{\text{loc}}\left(\mathbb{R}_+;\mathbb{R}^{m\times d}\right)$, the stochastic convolution $(K*\mathrm{d}Z)_T$ is well defined for a.e. $T>0$, see Section 4.1. As a result, the process $\widetilde{V}_2^T$ in (4.42) is well defined, for a.e. $T>0$.

Since, by the same computations as in the proof of Theorem 4.5, $\exp\{V_2^T\}$ is a local martingale, the final step required to conclude the current proof is to establish the equality $V_2^T=\widetilde{V}_2^T$. To do this, since $\mathscr{R}(\cdot,\psi_2(\cdot))$ is neither locally bounded nor locally square–integrable anymore, we cannot apply the stochastic Fubini's theorem as in (4.32). Indeed, the conditions in (4.33)-(4.34) are no longer satisfied. To overcome this problem, we employ an approximation argument.
For every $n\in\mathbb{N}$, define the truncation function $h_n(z)=z1_{\{|z|\le n\}}$, $z\in\mathbb{C}^m$. Recalling the process $\bar{g}_t(\cdot)$ in (4.31), the stochastic Fubini's theorem, whose application is possible thanks to computations similar to (4.33)-(4.34), yields, $\mathbb{P}-$a.s.,

$$\begin{aligned}\int_0^T h_n\left(\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right)^\top\left(\bar{g}_t-g_0\right)(s)\,\mathrm{d}s&\\ =\int_0^t\left[\int_r^T h_n\left(\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right)^\top K\left(s-r\right)\mathrm{d}s\right]\mathrm{d}Z_r.&\end{aligned} \tag{4.46}$$

Since, for any $z\in\mathbb{C}^m$, $h_n(z)\to z$ as $n\to\infty$, and (see (4.45)) $\int_0^T\left|\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right|\left|\bar{g}_t-g_0\right|(s)\,\mathrm{d}s<\infty$, $\mathbb{P}-$a.s., for every $t\in[0,T]$, for a.e. $T>0$, by the dominated convergence theorem we infer that

$$\begin{aligned}\lim_{n\to\infty}\int_0^T h_n(\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right))^\top\left(\bar{g}_t-g_0\right)(s)\,\mathrm{d}s&\\ =\int_0^T\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)^\top\left(\bar{g}_t-g_0\right)(s)\,\mathrm{d}s,\quad\mathbb{P}-\text{a.s.}&\end{aligned} \tag{4.47}$$

As for the right side of (4.46), we observe that, again by dominated convergence, for a.e. $T>0$,

$$\begin{aligned}\lim_{n\to\infty}\int_r^T h_n\left(\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right)^\top K\left(s-r\right)\mathrm{d}s&\\ =\int_r^T\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)^\top K\left(s-r\right)\mathrm{d}s,\quad\text{for a.e. } r\in[0,T].&\end{aligned}$$

Moreover, noticing that the map $|\mathscr{R}(\cdot,\psi_2(\cdot))|*|K|\in L^2_{\text{loc}}(\mathbb{R}_+)$, for $i=1,2$ we have

$$\begin{aligned}&\mathbb{E}\left[\int_0^t\left(\int_r^T\left|\mathscr{R}\left(T-s,\psi_2\left(T-s\right)\right)\right|\left|K(s-r)\right|\mathrm{d}s\right)^i\left|X_r\right|\mathrm{d}r\right]\\ &\quad=\int_0^t\left(\int_0^{T-r}\left|\mathscr{R}\left(T-r-s,\psi_2\left(T-r-s\right)\right)\right|\left|K(s)\right|\mathrm{d}s\right)^i\mathbb{E}\left[\left|X_r\right|\right]\mathrm{d}r\\ &\quad\le\int_0^T\left(\left(\left|\mathscr{R}(\cdot,\psi_2(\cdot))\right|*|K|\right)(T-r)\right)^i\mathbb{E}\left[\left|X_r\right|\right]\mathrm{d}r<\infty,\quad\text{for every } t\in[0,T],\end{aligned}$$

which holds for a.e. $T > 0$. Thus, we can apply the dominated convergence theorem for stochastic integrals (see, for instance, [153, Theorem 32, Chapter IV]) to claim that, for a.e. $T > 0$ and for every $t \in [0,T]$,

$$\mathbb{P} - \lim_{n\to\infty} \int_0^t \left[\int_r^T h_n \left(\mathcal{R} \left(T-s, \psi_2 \left(T-s \right) \right) \right)^\top K \left(s-r \right) \mathrm{d}s \right] \mathrm{d}Z_r$$
$$= \int_0^t \left[\int_r^T \mathcal{R} \left(T-s, \psi_2 \left(T-s \right) \right)^\top K \left(s-r \right) \mathrm{d}s \right] \mathrm{d}Z_r, \quad (4.48)$$

where $\mathbb{P} - \lim$ denotes the limit in probability. Combining (4.47)-(4.48) in (4.46), for a.e. $T > 0$ and $t \in [0,T]$ we obtain

$$\int_0^T \mathcal{R} \left(T-s, \psi_2 \left(T-s \right) \right)^\top \left(\bar{g}_t - g_0 \right) (s) \, \mathrm{d}s = \int_0^t \left[\int_r^T \mathcal{R} \left(T-s, \psi_2 \left(T-s \right) \right)^\top K \left(s-r \right) \mathrm{d}s \right] \mathrm{d}Z_r$$
$$= \int_0^t \left(\psi_2 \left(T-r \right)^\top - u_2^\top K(T-r) \right) \mathrm{d}Z_r, \quad \mathbb{P}-\text{a.s.}$$

Therefore, an analogue of the equality in (4.32) holds; proceeding as in (4.35), it allows to show that V_2^T and $\widetilde{V}_2^T$ are versions of each other. This completes the proof. ■

Remark 4.7. *When $m = d = 1$, $u_2 \in \mathbb{R}_-$, $\phi \equiv 0$, $f \equiv 0$ and g_0 is the Laplace transform of a finite signed measure on $\mathbb{R}_+$, Equation* (4.43) *reduces to the expression in [59, Theorem 5.7 (v)] for the (conditional) Laplace transform of the marginal distributions of X. To deduce such a result, however, in [59] the authors employ an abstract approach based on infinite dimensional Markovian lifts of stochastic affine Volterra processes, which is completely different from our procedure inspired by [8].*

4.3 An expression for V^T affine in the past trajectory of X

In this section we consider $m = d$ and aim to find an alternative formula for V^T (see (4.21)-(4.22)) which is affine in the past trajectory of X. This new expression can be used to prove the martingale property of the complex–valued process $\exp\left\{V^T\right\}$ in particular cases (see Section 4.4). Similar formulas might also be obtained for the processes introduced in Subsection 4.2.1 to study the marginal distributions of X and Z, see Theorems 4.6-4.8.

Due to the lack of regularity of the trajectories of both X and the stochastic convolution in $\mathrm{d}Z$, we are going to require mild, additional conditions on the kernel K, in particular on the shifted kernels $\Delta_h K$ for $h > 0$. We start with a preliminary result providing an alternative expression for the adjusted forward process $g_t(\cdot)$.

Lemma 4.9. *Assume that $K \in L^2_{\text{loc}}\left(\mathbb{R}_+; \mathbb{R}^{d\times d}\right)$ is continuous on $(0,\infty)$ and that it admits a resolvent of the first kind L with no point masses in $(0,\infty)$. In addition, suppose that for every $h > 0$ the shifted kernel $\Delta_h K$ is differentiable, with derivative $(\Delta_h K)' \in C\left(\mathbb{R}_+; \mathbb{R}^{d\times d}\right)$. Then, for every $T > 0$, for every $t \in [0,T)$*

$$g_t(T) = g_0(T) + K(T-t)\, Z_t + \left(\left((\Delta_{T-t} K)' * L \right) * (X - g_0) \right)(t), \quad \mathbb{P}-a.s. \quad (4.49)$$

Proof. Fix $h > 0$. We first show that the stochastic convolution $\Delta_h K * \mathrm{d}Z$ has a càdlàg version. Indeed, for every $t \in \mathbb{R}_+$, $(\Delta_h K * \mathrm{d}Z)_t = \int_0^t f_{t,h}(s)\,\mathrm{d}Z_s$, $\mathbb{P}-$a.s., with $f_{t,h}(s) = K(t+h-s)$, $s \in [0,t]$. Integration by parts yields

$$\begin{aligned}\int_0^t f_{t,h}(s)\,\mathrm{d}Z_s &= f_{t,h}(t)\,Z_t - f_{t,h}(0)\,Z_0 - \int_0^t (f_{t,h})'(s)\,Z_{s-}\,\mathrm{d}s\\ &= K(h)\,Z_t + \int_0^t (\Delta_h K)'(t-s)\,Z_s\,\mathrm{d}s, \quad \mathbb{P}-\text{a.s.},\end{aligned}$$

where we also note that $Z_{t-} = Z_t$ for a.e. $t > 0$, $\mathbb{P}-$a.s. Since the rightmost side of the previous equality is a càdlàg process we obtain the desired claim. Hence in what follows we consider $\Delta_h K * \mathrm{d}Z$ to be right–continuous. In particular, the process $(\Delta_h K - K(h)) * \mathrm{d}Z$ is continuous.

Thanks to the assumptions on the kernel, we apply [96, Corollary 7.3, Chapter 3] to claim that the function $(\Delta_h K - K(h)) * L$ is locally absolutely continuous in $\mathbb{R}_+$, with

$$((\Delta_h K - K(h)) * L)' = (\Delta_h K)' * L, \quad \text{a.e. in } \mathbb{R}_+.$$

In particular, the function $(\Delta_h K)' * L \in C\left(\mathbb{R}_+; \mathbb{R}^{d\times d}\right)$ by [96, Corollary 6.2 (iii), Chapter 3], the absence of point masses of L in $(0,\infty)$ and the continuity of $(\Delta_h K)'$. Therefore we invoke Proposition 4.4 to obtain

$$\begin{aligned}((\Delta_h K - K(h)) * \mathrm{d}Z)_t &= ((\Delta_h K - K(h)) * L)(0)\,(X - g_0)(t) + \left(\left((\Delta_h K)' * L\right) * (X - g_0)\right)(t)\\ &= \left(\left((\Delta_h K)' * L\right) * (X - g_0)\right)(t), \qquad \text{for a.e. } t \in \mathbb{R}_+,\ \mathbb{P}-\text{a.s.}\end{aligned}$$

Note that the last equality involves continuous processes, so it is indeed true for every $t \geq 0$ up to a $\mathbb{P}-$null set. Thus,

$$(\Delta_h K * \mathrm{d}Z)_t = K(h)\,Z_t + \left(\left((\Delta_h K)' * L\right) * (X - g_0)\right)(t), \quad t \geq 0,\ \mathbb{P}-\text{a.s.} \tag{4.50}$$

At this point, take $t < T$ and recall that, by (4.25),

$$\begin{aligned}g_t(T) = g_0(T) + \int_0^t K(T-s)\,\mathrm{d}Z_s &= g_0(T) + \int_0^t (\Delta_{T-t} K)(t-s)\,\mathrm{d}Z_s\\ &= g_0(T) + (\Delta_{T-t} K * \mathrm{d}Z)_t, \quad \mathbb{P}-\text{a.s.}\end{aligned}$$

It suffices to take $h = T - t$ in (4.50) to deduce that

$$(\Delta_{T-t} K * \mathrm{d}Z)_t = K(T-t)\,Z_t + \left(\left((\Delta_{T-t} K)' * L\right) * (X - g_0)\right)(t), \quad \mathbb{P}-\text{a.s.}.$$

Hence, combining the two previous equations, we conclude

$$g_t(T) = g_0(T) + K(T-t)\,Z_t + \left(\left((\Delta_{T-t} K)' * L\right) * (X - g_0)\right)(t), \quad \mathbb{P}-\text{a.s.},$$

completing the proof. ■

Fix a generic $T > 0$. By Equation (4.49) we can write, for every $t \in [0,T)$,

$$g_t(s) = g_0(s) + K(s-t)\,Z_t + \left(\left((\Delta_{s-t} K)' * L\right) * (X - g_0)\right)(t), \quad \mathbb{P}-\text{a.s.},\ s \in (t,T). \tag{4.51}$$

Intuitively speaking, we want to plug this expression in (4.27), so that we obtain an alternative formulation for V_t^T which is an affine function on the past trajectory $\{X_s, s \leq t\}$. This is done in the next theorem, which extends [8, Theorem 4.5] under further conditions on the kernel K. These addtional assumptions hold for instance in the one–dimensional case if K is completely monotone (recall that a function f is called completely monotone on $(0,\infty)$ if it is infinitely differentiable there with $(-1)^k f^{(k)}(t) \geq 0$ for all $t > 0$ and $k = 0, 1, \dots$).

Theorem 4.10. *Assume that $K \in L^2_{\text{loc}}\left(\mathbb{R}_+; \mathbb{R}^{d\times d}\right)$ is continuous on $(0,\infty)$ and that it admits a resolvent of the first kind L with no point masses in $(0,\infty)$. In addition, suppose that for every $h > 0$ the shifted kernel $\Delta_h K$ is differentiable, with $(\Delta_h K)'$ continuous on $\mathbb{R}_+$. Under Hypothesis 4.1, if the total variation bound*

$$\sup_{h\in\left(0,\overline{T}\right]} \|\Delta_h K * L\|_{TV\left(\left[0,\overline{T}\right]\right)} < \infty, \quad \text{for all } \overline{T} > 0, \tag{4.52}$$

holds, then for every $h > 0$ the $\mathbb{C}^d$–valued function

$$\pi_h(r) = \left(\mathcal{R}(\cdot, \psi(\cdot))^\top * \left((\Delta_\cdot K)' * L\right)(r)\right)(h)^\top \tag{4.53}$$

is well–defined for a.e. $r \in \mathbb{R}_+$ and belongs to $L^1_{\text{loc}}\left(\mathbb{R}_+; \mathbb{C}^d\right)$. Moreover, $\mathbb{P}$–a.s., for a.e. $t \in (0,T)$,

$$\begin{aligned} V_t^T = \phi(T-t) + \int_0^t f(T-s)^\top X_s \, \mathrm{d}s + \int_0^{T-t} \mathcal{R}(s, \psi(s))^\top g_0(T-s) \, \mathrm{d}s \\ + \psi(T-t)^\top Z_t + \left(\pi_{T-t}{}^\top * (X - g_0)\right)(t), \end{aligned} \tag{4.54}$$

where ϕ is defined in (4.20).

Proof. Fix $h > 0$; expanding the notation in (4.53) for π_h we have

$$\pi_h(r)^\top = \int_0^h \mathcal{R}(s, \psi(s))^\top \left[\int_0^r (\Delta_{h-s} K)'(r-u) \, L(\mathrm{d}u)\right] \mathrm{d}s.$$

In order to see that it is well–defined a.e. on $\mathbb{R}_+$ and belongs to $L^1_{\text{loc}}\left(\mathbb{R}_+; \mathbb{C}^d\right)$, first note that for every positive s, the continuity of $(\Delta_s K)'$ and the absence of point masses for L in $(0,\infty)$ allow to apply [96, Corollary 6.2 (iii), Chapter 3], which ensures the continuity on $\mathbb{R}_+$ of $(\Delta_s K)' * L$. As a consequence, we can define the $\mathbb{C}^d$–valued measurable function

$$\left[\mathcal{R}(s, \psi(s))^\top \left((\Delta_{h-s} K)' * L\right)(r)\right]^\top, \quad (s, r) \in (0, h) \times \mathbb{R}_+.$$

Recalling the previous proof, we see that $(\Delta_h K)' * L$ is, almost everywhere, the derivative of the locally absolutely continuous function $(\Delta_h K - K(h)) * L$. The boundedness of $\mathcal{R}(\cdot, \psi(\cdot))$ on $[0, h]$ by a constant $C_h > 0$ (see Hypothesis 4.1 and the subsequent comment) coupled with Condition (4.52), Tonelli's theorem and [96, Theorem 6.1 (v), Chapter 3] yields, for a generic $\overline{T} > h$,

$$\begin{aligned}
\int_0^{\overline{T}} \left[\int_0^h |\mathcal{R}(s,\psi(s))| \left|\left((\Delta_{h-s}K)' * L\right)(r)\right| \mathrm{d}s \right] \mathrm{d}r &= \int_0^h |\mathcal{R}(s,\psi(s))| \left[\int_0^{\overline{T}} \left|\left((\Delta_{h-s}K)' * L\right)(r)\right| \mathrm{d}r \right] \mathrm{d}s \\
&\le d^2 \int_0^h |\mathcal{R}(s,\psi(s))| \left[\|(\Delta_{h-s}K - K(h-s)) * L\|_{\mathrm{TV}([0,\overline{T}])} \right] \mathrm{d}s \\
&\le d^2 \left[\int_0^h |\mathcal{R}(s,\psi(s))| \, \|\Delta_{h-s}K * L\|_{\mathrm{TV}([0,\overline{T}])} \,\mathrm{d}s + |L|\left([0,\overline{T}]\right) \int_0^h |\mathcal{R}(s,\psi(s))| \, |K(h-s)| \,\mathrm{d}s \right] \\
&\le d^2 \left[\sup_{s \in (0,\overline{T}]} \|\Delta_s K * L\|_{\mathrm{TV}([0,\overline{T}])} C_h h + |L|\left([0,\overline{T}]\right) \int_0^h |\mathcal{R}(s,\psi(s))| \, |K(h-s)| \,\mathrm{d}s \right] < \infty.
\end{aligned}$$

Hence the conclusion on π_h follows. Furthermore, by Lebesgue's fundamental theorem of calculus, the $\mathbb{C}^d$−valued function $\Pi_h(r) = \int_0^r \pi_h(u)\,\mathrm{d}u$, $r \in \mathbb{R}_+$, is locally absolutely continuous on $\mathbb{R}_+$, with $\Pi_h' = \pi_h$ a.e. Using Fubini's theorem we can obtain the following explicit expression for such Π_h

$$\Pi_h(r)^\top = \int_0^h \mathcal{R}(s,\psi(s))^\top \left((\Delta_{h-s}K - K(h-s)) * L\right)(r)\,\mathrm{d}s, \quad r \in \mathbb{R}_+. \tag{4.55}$$

At this point we observe that for every function $g \in L^1_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}^d\right)$ we have, reasoning as before and using the boundedness of $\mathcal{R}(\cdot,\psi(\cdot))$ on $[0,T]$ by a positive constant C_T,

$$\begin{aligned}
&\int_0^T \left[\int_0^t |g(t-u)| \left(\int_0^{T-t} |\mathcal{R}(s,\psi(s))| \left|\left((\Delta_{T-t-s}K)' * L\right)(u)\right| \mathrm{d}s \right) \mathrm{d}u \right] \mathrm{d}t \\
&\quad = \int_0^T \left[\int_0^t |g(t-u)| \left(\int_0^{T-t} |\mathcal{R}(T-t-s,\psi(T-t-s))| \left|\left((\Delta_s K)' * L\right)(u)\right| \mathrm{d}s \right) \mathrm{d}u \right] \mathrm{d}t \\
&\quad = \int_0^T \left[\int_0^{T-t} |\mathcal{R}(T-t-s,\psi(T-t-s))| \left(\int_0^t |g(t-u)| \left|\left((\Delta_s K)' * L\right)(u)\right| \mathrm{d}u \right) \mathrm{d}s \right] \mathrm{d}t \\
&\quad = \int_0^T \left[\int_0^{T-s} |\mathcal{R}(T-s-t,\psi(T-s-t))| \left(\int_0^t |g(t-u)| \left|\left((\Delta_s K)' * L\right)(u)\right| \mathrm{d}u \right) \mathrm{d}t \right] \mathrm{d}s \\
&\quad \le C_T d^2 \|g\|_{L^1\left([0,T];\mathbb{R}^d\right)} \int_0^T \|(\Delta_s K - K(s)) * L\|_{\mathrm{TV}([0,T])} \,\mathrm{d}s \\
&\quad \le C_T d^2 \|g\|_{L^1\left([0,T];\mathbb{R}^d\right)} \left[T \sup_{s \in (0,T]} \|\Delta_s K * L\|_{\mathrm{TV}([0,T])} + |L|([0,T]) \int_0^T |K(s)|\,\mathrm{d}s \right] < \infty,
\end{aligned} \tag{4.56}$$

where we apply Tonelli's theorem, together with [96, Theorem 2.2 (i), Chapter 2] and a change of variables. Consequently, for almost every $t \in (0,T)$ we can apply Fubini's theorem to obtain

$$\int_0^t \pi_{T-t}(u)^\top g(t-u)\,\mathrm{d}u = \int_0^{T-t} \mathcal{R}(s,\psi(s))^\top \left[\int_0^t \left((\Delta_{T-t-s}K)' * L\right)(u)\, g(t-u)\,\mathrm{d}u \right] \mathrm{d}s. \tag{4.57}$$

Computations analogous to those in (4.56) (with g [resp., $|\mathcal{R}(\cdot,\psi(\cdot))|$] substituted by $X - g_0$ [resp., 1]) let us conclude, by Fubini's theorem and Equation (4.7), that there is a jointly measurable modification of the process $\left(\left((\Delta_{\cdot-t}K)' * L\right) * (X - g_0)\right)(t)$ on $\Omega \times (t,T)$ for a.e. $t \in (0,T)$. Therefore we interpret (4.51) pathwise, namely the equality holds almost everywhere in (t,T) up to a $\mathbb{P}$−null set.

Now we focus on $\widetilde{V}_t^T$. The previous analysis together with (4.27) and a suitable change of variables yields

$$\begin{aligned}\widetilde{V}_t^T &= \phi(T-t) + \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_0^{T-t} \mathcal{R}(s,\psi(s))^\top g_t(T-s)\,\mathrm{d}s \\ &= \left\{\phi(T-t) + \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_0^{T-t} \mathcal{R}(s,\psi(s))^\top g_0(T-s)\,\mathrm{d}s\right\} \\ &\quad + \left\{\left(\int_0^{T-t} \mathcal{R}(s,\psi(s))^\top K(T-t-s)\,\mathrm{d}s\right) Z_t\right\} \\ &\quad + \left\{\int_0^{T-t} \mathcal{R}(s,\psi(s))^\top \left(((\Delta_{T-t-s}K)' * L) * (X-g_0)\right)(t)\,\mathrm{d}s\right\} \\ &= \mathbf{I}_t + \mathbf{II}_t + \mathbf{III}_t, \quad \mathbb{P}-\text{a.s.}, \ t \in (0,T).\end{aligned}$$

The idea is to analyze separately the addends that we have singled out in the previous computations. Note that $\mathbf{III}_t$ is finite because $\widetilde{V}_t^T, \mathbf{I}_t, \mathbf{II}_t$ are so, and that we can consider a jointly measurable modification of this process in $\Omega \times (0,T)$, again by Fubini's theorem and Equation (4.7) (see (4.56)). Taking into account (4.28) we have

$$V_t^T = \mathbf{I}_t + \mathbf{II}_t + \mathbf{III}_t, \quad \text{for a.e. } t \in (0,T), \, \mathbb{P}-\text{a.s.}, \tag{4.58}$$

where the equality can be understood pathwise as it involves jointly measurable processes.
Regarding $\mathbf{II}_t$, since ψ solves the Riccati–Volterra equation in (4.19) we have

$$\mathbf{II}_t = \psi(T-t)^\top Z_t.$$

As for $\mathbf{III}_t$, by (4.57) we have

$$\mathbf{III}_t = \int_0^t \pi_{T-t}(u)^\top (X-g_0)(t-u)\,\mathrm{d}u = \left(\pi_{T-t}{}^\top * (X-g_0)\right)(t), \quad \text{for a.e. } t \in (0,T), \, \mathbb{P}-\text{a.s.}$$

Substituting the two previous equations in (4.58) we conclude

$$\begin{aligned}V_t^T = \phi(T-t) + \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_0^{T-t} \mathcal{R}(s,\psi(s))^\top g_0(T-s)\,\mathrm{d}s \\ + \psi(T-t)^\top Z_t + \left(\pi_{T-t}{}^\top * (X-g_0)\right)(t),\end{aligned}$$

for almost every $t \in (0,T)$, $\mathbb{P}-$a.s. The proof is now complete. ■

If the resolvent of the first kind L is the sum of a locally integrable function and a point mass in 0, then recalling (4.6) we can apply Lemma 4.3 (see also the final comment in Remark 4.3) and argue as in (4.16) to see that $Z_t = (L * (X-g_0))(t)$, for a.e. $t > 0$, $\mathbb{P}-$a.s. In addition, for every $h > 0$ we define the $\mathbb{C}^d-$valued function

$$\begin{aligned}\widetilde{\Pi}_h(r)^\top &= \Pi_h(r)^\top + \psi(h)^\top L(\{0\}) + \psi(h)^\top L((0,r]) \\ &= \int_0^h \mathcal{R}(s,\psi(s))^\top (\Delta_{h-s}K * L)(r)\,\mathrm{d}s, \quad r \in \mathbb{R}_+,\end{aligned} \tag{4.59}$$

where the second equality is due to (4.55). Note that $\widetilde{\Pi}_h$ is locally absolutely continuous on $\mathbb{R}_+$, and that

$$\begin{aligned}\left(\pi_{T-t}{}^\top * (X-g_0)\right)(t) &= \left(\mathrm{d}\Pi_{T-t}{}^\top * (X-g_0)\right)(t)\\ &= \left(\mathrm{d}\widetilde{\Pi}_{T-t}{}^\top * (X-g_0)\right)(t) - \psi(T-t)^\top (L*(X-g_0))(t) + \psi(T-t)^\top L(\{0\})(X-g_0)(t),\end{aligned}$$

which holds for a.e. $t \in (0,T)$, $\mathbb{P}-$a.s. Substituting in (4.54) we immediately deduce the following result.

Corollary 4.11. *Under the same hypotheses of Theorem 4.10, if the resolvent of the first kind L is the sum of a locally integrable function and a point mass in 0, then $\mathbb{P}-$a.s., for a.e. $t \in (0,T)$*

$$\begin{aligned}V_t^T = \phi(T-t) &+ \int_0^t f(T-s)^\top X_s\,\mathrm{d}s + \int_0^{T-t} \mathcal{R}(s,\psi(s))^\top g_0(T-s)\,\mathrm{d}s\\ &+ \psi(T-t)^\top L(\{0\})(X-g_0)(t) - \left(\mathrm{d}\widetilde{\Pi}_{T-t}{}^\top * g_0\right)(t) + \left(\mathrm{d}\widetilde{\Pi}_{T-t}{}^\top * X\right)(t). \quad (4.60)\end{aligned}$$

4.4 The 1−dimensional Volterra square root diffusion with jumps

In this section we discuss a one–dimensional example ($m = d = 1$) where not only are we able to infer the assumptions made in the previous arguments, such as the existence of solutions to the stochastic Volterra equation (4.2) and the Riccati–Volterra equation (4.19) (i.e., Hypothesis 4.1), but also we can prove the martingale property of the process $\exp\{V^T\}$. In order to develop the theory we need to require more properties for the kernel K. In particular, we consider a hypothesis which is standard in the theory of stochastic Volterra equations, that is (see [2, Condition (2.10)], and also [8, Condition (3.4)] and [5, Assumption B.2])

Hypothesis 4.2. *The kernel K is nonnegative, nonincreasing, not identically zero and continuously differentiable on $(0,\infty)$. Furthermore, its resolvent of the first kind L is nonnegative and nonincreasing, i.e., $s \mapsto L([s,s+t])$ is nonincreasing for every $t \geq 0$.*

Notice that, under Hypothesis 4.2, the map $s \mapsto L(\{s\})$ is nonincreasing, where L is the resolvent of the first kind of K. Combining this fact with Lebesgue's decomposition theorem, which ensures that L has at most a countable number of point masses, we deduce that L has no point masses in $(0,\infty)$.
In the sequel, we suppose that K and the shifted kernels $\Delta_{1/n}K$, $n \in \mathbb{N}$, satisfy Hypothesis 4.2. This is the case, for example, when K is a completely monotone function not identically equal to 0.

We focus on the following stochastic Volterra equation of convolution type:

$$X = g_0 + (K * \mathrm{d}Z), \quad \mathbb{P}\otimes \mathrm{d}t - \text{a.e.}, \quad (4.61)$$

where Z is a real–valued semimartingale with differential characteristics (with respect to $h(\xi) = \xi$, $\xi \in \mathbb{R}$) given by $(b(X_t), a(X_t), \eta(X_t, \mathrm{d}\xi))$, $t \geq 0$, with

$$b(x) = bx, \qquad a(x) = c\,x, \qquad \eta(x,\mathrm{d}\xi) = x\nu(\mathrm{d}\xi), \quad x \geq 0.$$

Here $b \in \mathbb{R}$, $c \geq 0$ and ν is a nonnegative measure on $\mathbb{R}_+$ such that $\int_{\mathbb{R}_+} |\xi|^2 \nu(\mathrm{d}\xi) < \infty$. The function $g_0 \colon \mathbb{R}_+ \to \mathbb{R}$ is an admissible input curve in either one of the following two forms

i. g_0 is continuous and non–decreasing, with $g_0(0) \geq 0$;

ii. $g_0(t) = x_0 + \int_0^t K(t-s)\,\theta(s)\,\mathrm{d}s$, $t \geq 0$, where $x_0 \geq 0$ and $\theta\colon \mathbb{R}_+ \to \mathbb{R}_+$ is locally bounded.

Notice that (4.61) describes a 1−dimensional Volterra square root diffusion with jumps. In this framework, we can invoke [2, Theorem 2.13] to claim the existence of a weak, predictable solution $X = (X_t)_{t\geq 0}$ of (4.61) with trajectories in $L^1_{\text{loc}}(\mathbb{R}_+)$ such that $X \geq 0$, $\mathbb{P}\otimes \mathrm{d}t$−a.e. Actually, if $g_0 \in L^2_{\text{loc}}(\mathbb{R}_+)$, the paths of this solution X are in $L^2_{\text{loc}}(\mathbb{R}_+)$, $\mathbb{P}$−a.s., as the next result shows.

Lemma 4.12. *Suppose that $g_0 \in L^2_{\text{loc}}(\mathbb{R}_+)$ and let X be a solution of (4.61) with trajectories in $L^1_{\text{loc}}(\mathbb{R}_+)$ such that $X \geq 0$, $\mathbb{P}\otimes \mathrm{d}t$−a.e. Then, for every $T > 0$, $\mathbb{E}\big[\big(\int_0^T |X_t|^2\,\mathrm{d}t\big)^{1/2}\big] < \infty$.*

Proof. The convolution equation (4.61) enables us to write, $\mathbb{P}$−a.s.,

$$|X_t|^2 \leq 4\left(|g_0(t)|^2 + |b|^2\,|(K*X)(t)|^2 + |(K*\mathrm{d}M^c)_t|^2 + \left|\left(K*\mathrm{d}M^d\right)_t\right|^2\right), \quad \text{for a.e. } t \geq 0.$$

Integrating over the interval $(0,T)$, $T>0$, we have

$$\left(\int_0^T |X_t|^2\,\mathrm{d}t\right)^{\frac{1}{2}} \leq 2\bigg(2 + \|g_0\|_{L^2([0,T])} + |b|\,\|K*X\|_{L^2([0,T])} + \int_0^T |(K*\mathrm{d}M^c)_t|^2\,\mathrm{d}t + \int_0^T \left|\left(K*\mathrm{d}M^d\right)_t\right|^2\mathrm{d}t\bigg), \quad \mathbb{P}-\text{a.s.},$$

where we also use that $\sqrt{x} \leq 1 + x$, $x \in \mathbb{R}_+$. By [96, Theorem 2.2 (i), Chapter 2], $\|K*X\|_{L^2([0,T])} \leq \|K\|_{L^2([0,T])}\,\|X\|_{L^1([0,T])}$, hence taking expectation in the previous inequality we obtain, using Tonelli's theorem,

$$\mathbb{E}\left[\left(\int_0^T |X_t|^2\,\mathrm{d}t\right)^{\frac{1}{2}}\right] \leq 2\bigg(2 + \|g_0\|_{L^2([0,T])} + |b|\,\|K\|_{L^2([0,T])}\,\mathbb{E}\left[\|X\|_{L^1([0,T])}\right] + \int_0^T \mathbb{E}\left[|(K*\mathrm{d}M^c)_t|^2\right]\mathrm{d}t + \int_0^T \mathbb{E}\left[\left|\left(K*\mathrm{d}M^d\right)_t\right|^2\right]\mathrm{d}t\bigg). \tag{4.62}$$

Recall that $(K*\mathrm{d}M^c)_t = \int_0^t K(t-s)\,\mathrm{d}M^c_s$, $\mathbb{P}$−a.s. for a.e. $t \geq 0$; therefore we use the Burkholder–Davis–Gundy inequality and the Young's type inequality in [3, Lemma A.1] to write (always bearing in mind Tonelli's theorem)

$$\int_0^T \mathbb{E}\left[|(K*\mathrm{d}M^c)_t|^2\right]\mathrm{d}t \leq c\cdot c_1\mathbb{E}\left[\int_0^T\left(\int_0^t |K(t-s)|^2\,X_s\,\mathrm{d}s\right)\mathrm{d}t\right] \leq c\cdot c_1\,\|K\|^2_{L^2([0,T])}\,\mathbb{E}\left[\|X\|_{L^1([0,T])}\right], \quad \text{for some } c_1 > 0.$$

Analogously, we invoke [134, Theorem 3.2] to assert

$$\int_0^T \mathbb{E}\left[\left|\left(K*\mathrm{d}M^d\right)_t\right|^2\right]\mathrm{d}t \leq 2\left(\int_{\mathbb{R}_+}|\xi|^2\,\nu(\mathrm{d}\xi)\right)c_2\,\|K\|^2_{L^2([0,T])}\,\mathbb{E}\left[\|X\|_{L^1([0,T])}\right], \quad \text{for some } c_2 > 0.$$

Now substituting the previous two bounds in (4.62) we see that the right side is finite by (4.7). This concludes the proof. ∎

Fix $f \in C(\mathbb{R}_+;\mathbb{C}_-)$. Since $\nu(\mathrm{d}\xi)$ is a positive measure on $\mathbb{R}_+$ with finite second moment and, for every $z \in \mathbb{C}_-$,

$$
\begin{aligned}
|\mathrm{Re}\,(e^z - 1 - z)| &\le \left|e^{\mathrm{Re}\, z} - 1 - \mathrm{Re}\, z\right| + e^{\mathrm{Re}\, z}\left(1 - \cos\left(\mathrm{Im}\, z\right)\right) \le |z|^2, \\
|\mathrm{Im}\,(e^z - 1 - z)| &= \left|e^{\mathrm{Re}\, z}\sin\left(\mathrm{Im}\, z\right) - \mathrm{Im}\, z\right| \le |\sin\left(\mathrm{Im}\, z\right) - \mathrm{Im}\, z| + |\mathrm{Im}\, z|\left(1 - e^{\mathrm{Re}\, z}\right) \le 2\,|z|^2,
\end{aligned}
$$

we have $\mathbb{C}_- \subset D_\nu$, see (4.17). We then set $D = \mathbb{C}_-$ and define the map $\mathcal{R}\colon \mathbb{R}_+ \times D \to \mathbb{C}$ as in (4.18). Notice that, by the dominated convergence theorem, $\mathcal{R}$ is continuous in its domain $\mathbb{R}_+ \times \mathbb{C}_-$, hence it is locally bounded. In addition to Equation (4.19), we consider the deterministic Riccati–Volterra equation

$$
\overline{\psi}(t) = \int_0^t K(t-s)\,\overline{\mathcal{R}}\left(s, \overline{\psi}(s)\right)\mathrm{d}s, \quad t \ge 0, \tag{4.63}
$$

where $\overline{\mathcal{R}}\colon \mathbb{R}_+ \times \mathbb{C}_- \to \mathbb{C}$ is defined by

$$
\overline{\mathcal{R}}(t,u) = \mathrm{Re}\, f(t) + bu + \frac{c}{2}u^2 + \int_{\mathbb{R}_+}\left(e^{u\xi} - 1 - u\xi\right)\nu(\mathrm{d}\xi), \quad (t,u) \in \mathbb{R}_+ \times \mathbb{C}_- \tag{4.64}
$$

and $\mathrm{Re}\, f\colon \mathbb{R}_+ \to \mathbb{R}_-$ denotes the real part of f. We observe that also $\overline{\mathcal{R}}$ is continuous in $\mathbb{R}_+ \times \mathbb{C}_-$. The next theorem shows the existence of global solutions to (4.19) and (4.63) (in particular, Hypothesis 4.1 is verified), as well as a comparison result between them which is crucial for the subsequent argument on the martingale property.

Theorem 4.13. *Let $f \in C(\mathbb{R}_+;\mathbb{C}_-)$ and assume Hypothesis 4.2.*

(i) *There exist a continuous global solution $\psi \in C(\mathbb{R}_+;\mathbb{C}_-)$ of (4.19) and a real–valued, continuous global solution $\overline{\psi} \in C(\mathbb{R}_+;\mathbb{R}_-)$ of (4.63).*

(ii) *Given $\psi \in C(\mathbb{R}_+;\mathbb{C}_-)$ and $\overline{\psi} \in C(\mathbb{R}_+;\mathbb{R}_-)$ satisfying (4.19) and (4.63), respectively, the following inequality holds:*

$$
\mathrm{Re}\,\psi(t) \le \overline{\psi}(t), \quad t \ge 0. \tag{4.65}
$$

Proof. The proof of (i) is in Appendix 4.B.1, and the one of (ii) is in Appendix 4.B.2. ■

In what follows, we take two continuous functions $\psi, \overline{\psi}$ as in Theorem 4.13 (i) and fix $T > 0$. We aim to prove the martingale property of the process $\exp\left\{V^T\right\}$, where V^T is given by (4.21)–(4.22). For this purpose, we define the process $\overline{V}^T$ as in (4.21)–(4.22), substituting $\mathrm{Re}\, f$ [resp., $\overline{\psi}$] for f [resp., ψ]. Theorem 4.5 shows that $\overline{V}_t^T = \widetilde{\overline{V}}_t^T$, $\mathbb{P}$−a.s., for every $t \in [0,T]$, where of course we define $\widetilde{\overline{V}}^T$ as in (4.27) with the same substitution as before. It is known that $\exp\{\overline{V}^T\}$ is a true, real–valued martingale. This is due to [2, Lemma 6.1], which in turn is an interesting application of the Novikov–type condition in [129, Theorem IV.3]. The idea of the present section consists in using the expression (4.60) in order to prove the bound $|\exp\{V^T\}| \le C\exp\{\overline{V}^T\}$ up to indistinguishability for some $C > 0$, so that we can conclude that $\exp\left\{V^T\right\}$ is a martingale, too.

Direct computations based on the Riccati–Volterra equation (4.19) yield, for every $h > 0$,

$$
\Delta_h\psi(r) = \left(\Delta_h\left(\mathcal{R}\left(\cdot, \psi(\cdot)\right)\right) * K\right)(r) + \left(\mathcal{R}\left(\cdot, \psi(\cdot)\right) * \Delta_r K\right)(h), \quad r \ge 0.
$$

Focusing on the second addend on the right side, if we convolve it with L then

$$((\mathcal{R}(\diamond,\psi(\diamond)) * \Delta_{\cdot}K)(h) * L)(r) = \int_0^r \left[\int_0^h \mathcal{R}(s,\psi(s))(\Delta_{r-u}K)(h-s)\,\mathrm{d}s\right] L(\mathrm{d}u)$$
$$= \int_0^h \mathcal{R}(s,\psi(s))((\Delta_{h-s}K) * L)(r)\,\mathrm{d}s, \quad r \geq 0,$$

where the application of Fubini's theorem is justified because K is nonnegative and nonincreasing and L is a nonnegative measure. Whence, since $\Delta_h(\psi * L)(r) = (\mathcal{R}(\cdot,\psi(\cdot)) * 1)(r+h)$, recalling (4.59) we can write

$$\widetilde{\Pi}_h(r) = (\Delta_h\psi * L)(r) - \Delta_h(\psi * L)(r) + \int_0^h \mathcal{R}(s,\psi(s))\,\mathrm{d}s$$
$$= -\int_{(0,h]} \psi(h-s)\,L(r+\mathrm{d}s) + \int_0^h \mathcal{R}(s,\psi(s))\,\mathrm{d}s, \quad r \geq 0,$$

and in particular

$$\mathrm{Re}\left(\widetilde{\Pi}_h(r)\right) = -\int_{(0,h]} \mathrm{Re}(\psi(h-s))\,L(r+\mathrm{d}s) + \int_0^h \mathrm{Re}(\mathcal{R}(s,\psi(s)))\,\mathrm{d}s, \quad r \geq 0.$$

Repeating the same argument for $\overline{\psi}$ we also obtain

$$\widetilde{\overline{\Pi}}_h(r) = -\int_{(0,h]} \overline{\psi}(h-s)\,L(r+\mathrm{d}s) + \int_0^h \overline{\mathcal{R}}\left(s,\overline{\psi}(s)\right)\mathrm{d}s, \quad r \geq 0.$$

Taking the difference between the two previous equations we infer, for every $r \geq 0$,

$$\widetilde{\overline{\Pi}}_h(r) - \mathrm{Re}\left(\widetilde{\Pi}_h(r)\right) = -\int_{(0,h]} \left[\overline{\psi} - \mathrm{Re}\,\psi\right](h-s)\,L(r+\mathrm{d}s) + \int_0^h \left[\overline{\mathcal{R}}\left(\cdot,\overline{\psi}(\cdot)\right) - \mathrm{Re}(\mathcal{R}(\cdot,\psi(\cdot)))\right](s)\,\mathrm{d}s$$
$$= \left(\overline{\psi} - \mathrm{Re}\,\psi\right)(h)L(\{r\}) - \int_{[0,h]} \left[\overline{\psi} - \mathrm{Re}\,\psi\right](h-s)\,L(r+\mathrm{d}s)$$
$$+ \int_0^h \left[\overline{\mathcal{R}}\left(\cdot,\overline{\psi}(\cdot)\right) - \mathrm{Re}(\mathcal{R}(\cdot,\psi(\cdot)))\right](s)\,\mathrm{d}s. \tag{4.66}$$

Hence, we see that this function is increasing on the interval $(0,\infty)$ by (4.65) in Theorem 4.13 (ii) and Hypothesis 4.2 (see also the subsequent comment). We are now in position to prove the next, important result.

Theorem 4.14. *Assume that the kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R})$ satisfies the requirements of Corollary 4.11 together with Hypothesis 4.2. Then there exists a constant $C > 0$ such that*

$$\left|\exp\left\{V_t^T\right\}\right| \leq C\exp\left\{\overline{V}_t^T\right\}, \quad t \in [0,T],\ \mathbb{P} - a.s. \tag{4.67}$$

In particular, $\left(\exp\left\{V_t^T\right\}\right)_{t\in[0,T]}$ is a complex–valued martingale.

Proof. First of all note that $\left|\exp\left\{V^T\right\}\right| = \exp\left\{\mathrm{Re}\left(V^T\right)\right\}$. For the reader's convenience, we write the expression for $\mathrm{Re}\left(V^T\right)$ provided by (4.60)

$$\begin{aligned}\mathrm{Re}\left(V_t^T\right) = &\int_0^t \mathrm{Re}\, f\left(T-s\right) X_s\,\mathrm{d}s\\ &+\int_0^{T-t} \mathrm{Re}\left(\mathcal{R}\left(s,\psi\left(s\right)\right)\right) g_0\left(T-s\right)\mathrm{d}s + \mathrm{Re}\,\psi\left(T-t\right) L\left(\{0\}\right)\left(X-g_0\right)\left(t\right)\\ &-\left(\mathrm{d}\left(\mathrm{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*g_0\right)\left(t\right)+\left(\mathrm{d}\left(\mathrm{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*X\right)\left(t\right), \quad \text{for a.e. } t\in\left(0,T\right), \mathbb{P}-\text{a.s.}\end{aligned}$$

The idea of the proof is simply to compare, term by term, the addends of this sum with the corresponding ones in the expansion of $\overline{V}^T$ according to (4.60). We are going to consider a common set $\Omega_0 \subset \Omega$, with $\mathbb{P}\left(\Omega_0\right) = 1$, such that both the expressions for $\mathrm{Re}\left(V_t^T\right)$ and $\overline{V}_t^T$ are valid on $(0,T)\setminus N^\omega$, being $N^\omega \subset (0,T)$ a $\mathrm{d}t-$null set for every $\omega\in\Omega_0$.

Regarding the random terms, recall that $X \geq 0$, $\mathbb{P}\otimes\mathrm{d}t-$a.e. Therefore, without loss of generality, we can assume that for every $\omega\in\Omega_0$ and $t\in(0,T)\setminus N^\omega$ we have $X_t\left(\omega\right)\geq 0$. As a consequence (by (4.66))

$$\begin{aligned}&\left(\mathrm{d}\left(\widetilde{\overline{\Pi}}_{T-t}-\mathrm{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*X_\cdot\left(\omega\right)\right)\left(t\right)\geq 0\\ &\qquad\Longrightarrow\left(\mathrm{d}\left(\mathrm{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*X_\cdot\left(\omega\right)\right)\left(t\right)\leq\left(\mathrm{d}\widetilde{\overline{\Pi}}_{T-t}*X_\cdot\left(\omega\right)\right)\left(t\right).\end{aligned}$$

It is important to stress the fact that such an inequality can be stated because the measure L is absolutely continuous with respect to the Lebesgue measure on the interval $(0,\infty)$. Summing up,

$$\left(\mathrm{d}\left(\mathrm{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*X\right)\left(t\right)\leq\left(\mathrm{d}\widetilde{\overline{\Pi}}_{T-t}*X\right)\left(t\right), \quad t\in(0,T)\setminus N^\omega,\ \omega\in\Omega_0.$$

Moreover, since $L\left(\{0\}\right)\geq 0$, by (4.65) we immediately have

$$\mathrm{Re}\,\psi\left(T-t\right)L\left(\{0\}\right)X_t\leq\overline{\psi}\left(T-t\right)L\left(\{0\}\right)X_t, \quad t\in(0,T)\setminus N^\omega,\ \omega\in\Omega_0.$$

The other random addend $\int_0^t \mathrm{Re}\, f\left(T-s\right)X_s\,\mathrm{d}s$ appears in both the expressions for $\mathrm{Re}\left(V_t^T\right)$ and $\overline{V}_t^T$, so it does not need to be discussed.

As for the deterministic terms, we observe that, by Hölder's inequality,

$$\begin{aligned}&\left|\int_0^{T-t} g_0\left(T-s\right)\left(\mathrm{Re}\left(\mathcal{R}\left(s,\psi\left(s\right)\right)\right)-\overline{\mathcal{R}}\left(s,\overline{\psi}\left(s\right)\right)\right)\mathrm{d}s\right|\\ &\qquad\leq\left\|g_0\right\|_{L^2([0,T])}\left\|\mathrm{Re}\left(\mathcal{R}\left(\cdot,\psi\left(\cdot\right)\right)\right)-\overline{\mathcal{R}}\left(\cdot,\overline{\psi}\left(\cdot\right)\right)\right\|_{L^2([0,T])},\end{aligned}$$

for any $t\in(0,T)$. Hence, calling $C_1=\left\|g_0\right\|_{L^2([0,T])}\left\|\mathrm{Re}\left(\mathcal{R}\left(\cdot,\psi\right)\right)-\overline{\mathcal{R}}\left(\cdot,\psi\right)\right\|_{L^2([0,T])}$, we have

$$\int_0^{T-t} g_0\left(T-s\right)\mathrm{Re}\left(\mathcal{R}\left(s,\psi\left(s\right)\right)\right)\mathrm{d}s\leq C_1+\int_0^{T-t} g_0\left(T-s\right)\overline{\mathcal{R}}\left(s,\overline{\psi}\left(s\right)\right)\mathrm{d}s, \quad t\in\left(0,T\right).$$

Furthermore, recalling the continuity of ψ, $\overline{\psi}$ and g_0, we call $C_2=\max_{t\in[0,T]}\left\{\left|\overline{\psi}-\mathrm{Re}\,\psi\right|\left(T-t\right)g_0\left(t\right)\right\}$, so that we have

$$-\mathrm{Re}\,\psi\left(T-t\right)L\left(\{0\}\right)g_0\left(t\right)\leq L\left(\{0\}\right)C_2-\overline{\psi}\left(T-t\right)L\left(\{0\}\right)g_0\left(t\right), \quad t\in\left(0,T\right).$$

Finally, looking at (4.66) we compute

$$\left(\mathrm{d}\left(\widetilde{\widetilde{\Pi}}_{T-t}-\operatorname{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*1\right)(t)\leq-\int_{[0,T-t]}\left[\overline{\psi}-\operatorname{Re}\psi\right](T-t-s)\left[L\left(t+\mathrm{d}s\right)-L\left(\mathrm{d}s\right)\right]$$
$$\leq 2\max_{t\in[0,T]}\left|\overline{\psi}\left(t\right)-\operatorname{Re}\psi\left(t\right)\right|L\left([0,T]\right)=C_3,\quad t\in(0,T)\,.$$

Hence exploiting the continuity of the input curve we conclude that

$$\left|\left(\mathrm{d}\left(\widetilde{\widetilde{\Pi}}_{T-t}-\operatorname{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*g_0\right)(t)\right|\leq C_3\max_{t\in[0,T]}\left|g_0\left(t\right)\right|,$$

which in turn implies

$$-\left(\mathrm{d}\left(\operatorname{Re}\left(\widetilde{\Pi}_{T-t}\right)\right)*g_0\right)(t)\leq C_3\max_{t\in[0,T]}\left|g_0\left(t\right)\right|-\left(\mathrm{d}\widetilde{\widetilde{\Pi}}_{T-t}*g_0\right)(t),\quad t\in(0,T)\,.$$

Combining all these results we deduce that

$$\operatorname{Re}\left(V_t^T\left(\omega\right)\right)\leq C_1+L\left(\{0\}\right)C_2+C_3\max_{t\in[0,T]}\left|g_0\left(t\right)\right|+\overline{V}_t^T\left(\omega\right),\quad t\in(0,T)\setminus N^{\omega},\,\omega\in\Omega_0. \tag{4.68}$$

Since N^{ω} is a null set, its complementary $(N^{\omega})^c=(0,T)\setminus N^{\omega}$ is dense in $[0,T]$. Recalling the regularity for the trajectories of the processes $\operatorname{Re}\left(V^T\right)$ and $\overline{V}^T$, we can assume that for every $\omega\in\Omega_0$ both the functions $\operatorname{Re}\left(V_\cdot^T\left(\omega\right)\right)$ and $\overline{V}_\cdot^T\left(\omega\right)$ are càdlàg in $[0,T]$ and left–continuous in T. Accordingly, we pass to the limit –from the right in $[0,T)$ and from the left in $T-$ to deduce, from (4.68), that

$$\operatorname{Re}\left(V_t^T\left(\omega\right)\right)\leq C_1+L\left(\{0\}\right)C_2+C_3\max_{t\in[0,T]}\left|g_0\left(t\right)\right|+\overline{V}_t^T\left(\omega\right),\quad t\in[0,T]\,,\ \omega\in\Omega_0,$$

i.e., (4.67) holds choosing $C=\exp\left\{C_1+L\left(\{0\}\right)C_2+C_3\max_{t\in[0,T]}\left|g_0\left(t\right)\right|\right\}$.

The second statement of the theorem follows from [112, Lemma 1.4], as $\left(\exp\left\{\overline{V}^T\right\}\right)_{t\in[0,T]}$ is a real–valued martingale. Thus, the proof is complete. ∎

Combining Theorem 4.14 with Theorem 4.5 (see (4.29)) we deduce the following result about weak uniqueness for (4.61).

Corollary 4.15. *The weak solution X of* (4.61) *is unique in law in $L^2_{\mathrm{loc}}\left(\mathbb{R}_+\right)$, that is: if $Y=\left(Y_t\right)_{t\geq0}$ is another predictable process (defined on a possibly different stochastic basis) such that $Y\geq0$, $\mathbb{P}\otimes\mathrm{d}t-$a.e., which satisfies* (4.61)*, then the laws of X and Y on the spaces $L^2\left([0,T]\right),\,T>0$, are the same.*

Proof. Fix $T>0$ and consider another weak solution Y of (4.61). We assume that X and Y are defined on the same stochastic basis to keep notation simple. The paths of Y are in $L^2\left([0,T]\right),\,\mathbb{P}-$a.s., by Lemma 4.12. We want to show that

$$\mathbb{E}\left[\exp\left\{i\int_0^T f\left(s\right)X_s\,\mathrm{d}s\right\}\right]=\mathbb{E}\left[\exp\left\{i\int_0^T f\left(s\right)Y_s\,\mathrm{d}s\right\}\right],\quad f\in L^2\left([0,T]\right). \tag{4.69}$$

First, we verify the previous equation for $f\in C\left([0,T]\right)$. Denoting by $\widetilde{f}\left(s\right)=if\left(T-s\right),\,s\in[0,T]\,,$ by Theorem 4.5 and Theorem 4.14 we have

$$\mathbb{E}\left[\exp\left\{i\int_0^T f\left(s\right)X_s\,\mathrm{d}s\right\}\right]=\mathbb{E}\left[\exp\left\{\int_0^T\widetilde{f}\left(T-s\right)X_s\,\mathrm{d}s\right\}\right]=\mathbb{E}\left[\exp\left\{V_0^T\right\}\right]$$
$$=\mathbb{E}\left[\exp\left\{\int_0^T\widetilde{f}\left(T-s\right)Y_s\,\mathrm{d}s\right\}\right]=\mathbb{E}\left[\exp\left\{i\int_0^T f\left(s\right)Y_s\,\mathrm{d}s\right\}\right],$$

where we use the fact that V_0^T in (4.22) does not depend on the solution process, but only on the solution of the Riccati–Volterra equation. Therefore (4.69) holds for continuous functions. Since $C([0,T])$ is dense in $L^2([0,T])$, Hölder's inequality allows to carry out a dominated convergence argument that yields (4.69) for all $f \in L^2([0,T])$. Hence, the laws of X and Y are the same on the space $L^2([0,T])$ by, for instance, [66, Proposition 2.5, Chapter 2]. This completes the proof. ■

Appendix 4.A The forward process

Given a kernel $K \in L^2_{\text{loc}}\left(\mathbb{R}_+;\mathbb{R}^{m\times d}\right)$, we want to find an expression for the forward process

$$\mathbb{E}\left[X_T \middle| \mathcal{F}_t\right], \quad 0 \le t \le T,$$

for almost every $T \in \mathbb{R}_+$.

If $b \equiv 0$, then (4.6) implies

$$X_T = g_0(T) + \left(K * \mathrm{d}\widetilde{Z}\right)_T = g_0(T) + \int_0^T K(T-s)\,\mathrm{d}\widetilde{Z}_s, \quad \mathbb{P}-\text{a.s., for a.e. } T \in \mathbb{R}_+.$$

By the martingale property ensured by (4.7) we immediately infer that, for almost every $T \in \mathbb{R}_+$,

$$\mathbb{E}\left[X_T \middle| \mathcal{F}_t\right] = g_0(T) + \int_0^t K(T-s)\,\mathrm{d}\widetilde{Z}_s, \quad \mathbb{P}-\text{a.s., } t \in [0,T]. \tag{4.70}$$

If $b \neq 0$, then we consider $m = d$ and introduce the resolvent of the second kind R_B associated with $-KB$. Note that $R_B \in L^2_{\text{loc}}\left(\mathbb{R}_+;\mathbb{R}^{d\times d}\right)$ by [96, Theorem 3.5, Chapter 2]. Convolving (4.6) with R_B and [96, Theorem 2.2 (viii), Chapter 2] yield

$$(R_B * X)(T) = (R_B * g_0)(T) + ((R_B * K) * b(X))(T) + \left(R_B * \left(K * \mathrm{d}\widetilde{Z}\right)\right)(T), \quad \text{for a.e. } T \in \mathbb{R}_+,\ \mathbb{P}-\text{a.s.}$$

The associativity of the stochastic convolution proved in Lemma 4.3 (with $\rho = R_B$) and the joint measurability of the processes involved let us rewrite this equality as follows:

$$\begin{aligned}(R_B * X)(T) = (R_B * g_0)(T) + ((R_B * K) * b_0)(T) + ((R_B * KB) * X)(T) \\ + \left((R_B * K) * \mathrm{d}\widetilde{Z}\right)_T, \quad \mathbb{P}-\text{a.s., for a.e. } T \in \mathbb{R}_+.\end{aligned} \tag{4.71}$$

From the resolvent identity (see the footnote [2]) we have $R_B * KB = KB + R_B$ a.e. in $\mathbb{R}_+$, so we rewrite Equation (4.71) as follows

$$\begin{aligned}0 = (R_B * g_0)(T) + ((R_B * K) * b_0)(T) + (KB * X)(T) + \left((R_B * K) * \mathrm{d}\widetilde{Z}\right)_T, \\ \mathbb{P}-\text{a.s., for a.e. } T \in \mathbb{R}_+.\end{aligned} \tag{4.72}$$

Consider the *canonical resolvent* $E_B = K - R_B * K$; subtracting (4.72) from (4.6) we have

$$X_T = (g_0 - (R_B * g_0))(T) + (E_B * b_0)(T) + \left(E_B * \mathrm{d}\widetilde{Z}\right)_T, \quad \mathbb{P}-\text{a.s., for a.e. } T \in \mathbb{R}_+.$$

Hence by the martingale property guaranteed by (4.7) we are able to find an expression for the forward process $\mathbb{E}\left[X_T \middle| \mathcal{F}_t\right]$, namely for almost every $T \in \mathbb{R}_+$, for every $t \in [0,T]$ it holds

$$\mathbb{E}\left[X_T \middle| \mathcal{F}_t\right] = (g_0(T) - (R_B * g_0)(T)) + (E_B * b_0)(T) + \int_0^t E_B(T-s)\,\mathrm{d}\widetilde{Z}_s, \quad \mathbb{P}-\text{a.s.} \tag{4.73}$$

Finally, notice that (4.73) reduces to (4.70) as $b \equiv 0$. Indeed, since $E_0 = K$ a.e. in $\mathbb{R}_+$ as $R_0 = 0 \left(\in \mathbb{R}^{d\times d}\right)$, combining (4.6) with Lemma 4.2 (see (4.9)) we have

$$X_T = g_0\left(T\right) + \left(E_0 * \mathrm{d}\widetilde{Z}\right)_T, \quad \mathbb{P}-\text{a.s., for a.e. } T \in \mathbb{R}_+, \tag{4.74}$$

and the assertion follows by the martingale property.

Remark 4.8. *Equation* (4.73) *with* $t = 0$ *implies that* $\mathbb{E}\left[X_T\right] = \left(g_0 - \left(R_B * g_0\right)\right)\left(T\right) + \left(E_B * b_0\right)\left(T\right)$ *for a.e.* $T \in \mathbb{R}_+$. *This result can be confirmed with a direct method. Specifically, by* (4.7) *and Tonelli's theorem the function* $\mathbb{E}\left[|X_\cdot|\right] \in L^1_{\mathrm{loc}}\left(\mathbb{R}_+;\mathbb{R}\right)$. *Hence taking expectations in* (4.2) *we obtain, by Fubini's theorem,*

$$\mathbb{E}\left[X_T\right] = \left(g_0 + K * b_0\right)\left(T\right) + \left(KB * \mathbb{E}\left[X_\cdot\right]\right)\left(T\right), \quad \textit{for a.e. } T \in \mathbb{R}_+,$$

i.e., $\mathbb{E}\left[X_\cdot\right] + \left(\left(-KB\right) * \mathbb{E}\left[X_\cdot\right]\right) = g_0 + K * b_0$ *a.e. in* $\mathbb{R}_+$. *By the variation of constants formula [96, Theorem* 3.5, *Chapter* 2*] we conclude*

$$\mathbb{E}\left[X_T\right] = \left(g_0 - \left(R_B * g_0\right) + \left(E_B * b_0\right)\right)\left(T\right), \quad \textit{for a.e. } T \in \mathbb{R}_+,$$

as desired.

Appendix 4.B On the 1−dimensional deterministic Riccati–Volterra equation

Here we focus on the Riccati–Volterra equation used in Section 4.4, i.e., (4.19) with

$$\mathscr{R}\left(t,u\right) = f\left(t\right) + bu + \frac{c}{2}u^2 + \int_{\mathbb{R}_+}\left(e^{u\xi} - 1 - u\xi\right)\nu\left(\mathrm{d}\xi\right), \quad \left(t,u\right) \in \mathbb{R}_+ \times \mathbb{C}_-, \tag{4.75}$$

where $f \in C\left(\mathbb{R}_+;\mathbb{C}_-\right)$. Throughout the section, we require Hypothesis 4.2 on the kernel K.

4.B.1 Existence of a global solution

It is easy to argue that (4.19) admits a continuous, noncontinuable solution ψ, with $\operatorname{Re}\psi \le 0$, defined on the maximal interval $[0, T_{\max})$ (see [2, Theorem 2.5, Step 1]). We are concerned with showing that $T_{\max} = \infty$, i.e., that ψ does not explode in finite time (cfr. [96, Theorem 1.1, Chapter 12]).
Fix a generic $T \in (0, T_{\max})$; taking real and imaginary parts in (4.19) and (4.75) we have, on the interval $[0, T]$,

$$\begin{aligned}\operatorname{Re}\psi = K * \Bigg[&\operatorname{Re} f + b\operatorname{Re}\psi + \frac{c}{2}\left(|\operatorname{Re}\psi|^2 - |\operatorname{Im}\psi|^2\right) \\ &+ \int_{\mathbb{R}_+}\left(\cos\left(\operatorname{Im}\psi\cdot\xi\right)e^{\operatorname{Re}\psi\cdot\xi} - 1 - \operatorname{Re}\psi\cdot\xi\right)\nu\left(\mathrm{d}\xi\right)\Bigg],\end{aligned} \tag{4.76}$$

$$\operatorname{Im}\psi = K * \left[\operatorname{Im} f + b\operatorname{Im}\psi + c\operatorname{Re}\psi\operatorname{Im}\psi + \int_{\mathbb{R}_+}\left(\sin\left(\operatorname{Im}\psi\cdot\xi\right)e^{\operatorname{Re}\psi\cdot\xi} - \operatorname{Im}\psi\cdot\xi\right)\nu\left(\mathrm{d}\xi\right)\right]. \tag{4.77}$$

First we study the imaginary part. In particular, we consider the function $h\colon \mathbb{R}_- \times \mathbb{R} \to \mathbb{R}$ defined as follows

$$h\left(x,y\right) = \begin{cases}\frac{1}{y}\int_{\mathbb{R}_+}\left(\sin\left(y\,\xi\right) - y\,\xi\right)e^{x\cdot\xi}\nu\left(\mathrm{d}\xi\right), & y \neq 0\\ 0, & y = 0\end{cases}, \quad x \le 0.$$

Note that h is continuous and non–positive in its domain. By construction

$$y \cdot h(x,y) = \int_{\mathbb{R}_+} (\sin(y\,\xi) - y\,\xi)\, e^{x\cdot\xi} \nu(\mathrm{d}\xi), \quad (x,y) \in \mathbb{R}_- \times \mathbb{R}.$$

Hence we can use this function to rewrite (4.77) in the following form

$$\operatorname{Im}\psi = K * \left[\operatorname{Im} f + b \operatorname{Im}\psi + c \operatorname{Re}\psi \operatorname{Im}\psi + \left(\int_{\mathbb{R}_+} \xi \left(e^{\operatorname{Re}\psi\cdot\xi} - 1\right) \nu(\mathrm{d}\xi) \right) \operatorname{Im}\psi + h(\operatorname{Re}\psi, \operatorname{Im}\psi) \operatorname{Im}\psi \right],$$

which holds on $[0,T]$. Consider the unique, continuous, non–negative solution on $[0,T]$ of the linear equation

$$g = K * \left[|\operatorname{Im} f| + b\,g + \left(c \operatorname{Re}\psi + \int_{\mathbb{R}_+} \xi \left(e^{\operatorname{Re}\psi\cdot\xi} - 1\right) \nu(\mathrm{d}\xi) + h(\operatorname{Re}\psi, \operatorname{Im}\psi) \right) g \right].$$

By Hypothesis 4.2, we can invoke [5, Theorem C.1] to deduce that $|\operatorname{Im}\psi| \le g$ on $[0,T]$. Next we introduce u, the unique, continuous solution of the linear equation

$$u = K * [|\operatorname{Im} f| + b\,u].$$

Notice that u is defined on $\mathbb{R}_+$, and that $g \le u$ on $[0,T]$ (again by [5, Theorem C.1]), as in this interval one has

$$c \operatorname{Re}\psi + \int_{\mathbb{R}_+} \xi \left(e^{\operatorname{Re}\psi\cdot\xi} - 1\right) \nu(\mathrm{d}\xi) + h(\operatorname{Re}\psi, \operatorname{Im}\psi) \le 0.$$

Therefore we have obtained the bound

$$|\operatorname{Im}\psi(t)| \le u(t), \quad 0 \le t \le T. \tag{4.78}$$

Secondly, Equation (4.76) ensures that $\operatorname{Re}\psi$ satisfies

$$\begin{aligned}\operatorname{Re}\psi = K * \Bigg[& \operatorname{Re} f + b \operatorname{Re}\psi + \frac{c}{2}\left(|\operatorname{Re}\psi|^2 - |\operatorname{Im}\psi|^2\right) + \int_{\mathbb{R}_+} \left(e^{\operatorname{Re}\psi\cdot\xi} - 1 - \operatorname{Re}\psi\cdot\xi\right) \nu(\mathrm{d}\xi) \\ & - \left| \int_{\mathbb{R}_+} e^{\operatorname{Re}\psi\cdot\xi} (\cos(\operatorname{Im}\psi\cdot\xi) - 1)\, \nu(\mathrm{d}\xi) \right| \Bigg]\end{aligned}$$

on $[0,T]$. Since $|\cos(x) - 1| = 1 - \cos(x) \le x^2/2$, $x \in \mathbb{R}$, we have (also recalling (4.78))

$$\left| \int_{\mathbb{R}_+} e^{\operatorname{Re}\psi\cdot\xi} (\cos(\operatorname{Im}\psi\cdot\xi) - 1)\, \nu(\mathrm{d}\xi) \right| \le \frac{1}{2}\left(\int_{\mathbb{R}_+} |\xi|^2 \nu(\mathrm{d}\xi) \right) |\operatorname{Im}\psi|^2 \le \frac{1}{2}\left(\int_{\mathbb{R}_+} |\xi|^2 \nu(\mathrm{d}\xi) \right) u^2, \tag{4.79}$$

which holds on $[0,T]$. This suggests to introduce the linear equation

$$l = K * \left[\operatorname{Re} f + b\,l - \left(\frac{c}{2} + \frac{1}{2} \int_{\mathbb{R}_+} |\xi|^2 \nu(\mathrm{d}\xi) \right) u^2 \right],$$

which has a unique, continuous, non–positive solution l defined on the whole $\mathbb{R}_+$. At this point, observe that the difference $\operatorname{Re}\psi - l$ satisfies the linear equation

$$\begin{aligned}\chi = K * \Bigg[& b\,\chi + \frac{c}{2} |\operatorname{Re}\psi|^2 + \frac{c}{2}\left(u^2 - |\operatorname{Im}\psi|^2\right) + \int_{\mathbb{R}_+} \left(e^{\operatorname{Re}\psi\cdot\xi} - 1 - \operatorname{Re}\psi\cdot\xi\right) \nu(\mathrm{d}\xi) \\ & + \left(\frac{1}{2}\left(\int_{\mathbb{R}_+} |\xi|^2 \nu(\mathrm{d}\xi) \right) u^2 - \left| \int_{\mathbb{R}_+} e^{\operatorname{Re}\psi\cdot\xi} (\cos(\operatorname{Im}\psi\cdot\xi) - 1)\, \nu(\mathrm{d}\xi) \right| \right) \Bigg].\end{aligned}$$

It admits a unique, continuous solution on $[0,T]$ which is non–negative by (4.78), (4.79) and the fact that $e^x-1-x\geq 0$, $x\in\mathbb{R}$. Since $T\in(0,T_{\max})$ was chosen arbitrarily, we conclude that

$$l(t)\leq \operatorname{Re}\psi(t)\leq 0 \quad \text{and} \quad |\operatorname{Im}\psi(t)|\leq u(t), \quad 0\leq t<T_{\max}.$$

Now recalling that l and u are continuous in $\mathbb{R}_+$, so they are bounded in every compact interval, we conclude that $T_{\max}=\infty$, as desired.

Finally we notice that if f takes values in $\mathbb{R}_-$, then from (4.78) we deduce that any solution of (4.19) is real–valued, as well. In particular, $\bar{\psi}$ in (4.63) is $\mathbb{R}_-$–valued.

4.B.2 A comparison result

The goal of this appendix is to prove the inequality (4.65) in Theorem 4.13 (ii), which is of utmost importance for the argument in Section 4.4. Precisely, we want to show that

$$\operatorname{Re}\psi(t)\leq\bar{\psi}(t), \quad t\geq 0,$$

where $\psi\in C(\mathbb{R}_+;\mathbb{C}_-)$ and $\bar{\psi}\in C(\mathbb{R}_+;\mathbb{R}_-)$ satisfy (4.19) and (4.63), respectively. Direct computations based on the definitions in (4.75) and (4.64) show that, for every $u\in\mathbb{C}_-$ and $t\geq 0$,

$$\begin{aligned}\operatorname{Re}(\mathcal{R}(t,u))=&\operatorname{Re}f(t)+b\operatorname{Re}(u)+\frac{c}{2}\left(|\operatorname{Re}(u)|^2-|\operatorname{Im}(u)|^2\right)\\&+\int_{\mathbb{R}_+}\left(\cos(\operatorname{Im}(u)\xi)\,e^{\operatorname{Re}(u)\xi}-1-\operatorname{Re}(u)\xi\right)\nu(\mathrm{d}\xi)\\ \leq&\operatorname{Re}f(t)+b\operatorname{Re}(u)+\frac{c}{2}|\operatorname{Re}(u)|^2+\int_{\mathbb{R}_+}\left(e^{\operatorname{Re}(u)\xi}-1-\operatorname{Re}(u)\xi\right)\nu(\mathrm{d}\xi)=\bar{\mathcal{R}}(t,\operatorname{Re}(u)).\end{aligned}$$

Summarizing,

$$\operatorname{Re}(\mathcal{R}(t,u))\leq\bar{\mathcal{R}}(t,\operatorname{Re}(u)), \quad u\in\mathbb{C}_-,\ t\geq 0.$$

Then taking the real parts in (4.19) and recalling that –under Hypothesis 4.2– the kernel K is nonnegative on $(0,\infty)$, we obtain

$$\operatorname{Re}(\psi(t))\leq\int_0^t K(t-s)\,\bar{\mathcal{R}}(s,\operatorname{Re}(\psi(s)))\,\mathrm{d}s, \quad t\geq 0.$$

Therefore we can introduce a nonnegative function $\gamma\colon\mathbb{R}_+\to\mathbb{R}_+$ defined by the relation

$$\operatorname{Re}(\psi(t))=-\gamma(t)+\int_0^t K(t-s)\,\bar{\mathcal{R}}(s,\operatorname{Re}(\psi(s)))\,\mathrm{d}s, \quad t\geq 0; \tag{4.80}$$

we immediately note that, using (4.19), one can rewrite γ as follows

$$\gamma(t)=\int_0^t K(t-s)\left(\bar{\mathcal{R}}(s,\operatorname{Re}\psi(s))-\operatorname{Re}(\mathcal{R}(s,\psi(s)))\right)\mathrm{d}s, \quad t\geq 0. \tag{4.81}$$

For a generic map $g\colon\mathbb{R}_+\to\mathbb{R}$ consider the condition

$$\Delta_h g-(\Delta_h K*L)(0)\,g-\mathrm{d}(\Delta_h K*L)*g\geq 0, \quad h\geq 0; \tag{4.82}$$

we denote by $\mathcal{G}_K = \{g\colon \mathbb{R}_+ \to \mathbb{R}$ s.t. g is continuous, satisfies (4.82) and $g(0) \geq 0\}$ the set of admissible curves, see [5, Condition B.3], and also [2, Equations (2.14)-(2.15)]. By [5, Remark B.6] and (4.81) we infer that $\gamma \in \mathcal{G}_K$.

At this point we subtract (4.80) from (4.63) to deduce, calling $\delta = \overline{\psi} - \operatorname{Re}\psi$, that

$$\delta(t) = \gamma(t) + \int_0^t K(t-s)\left(\overline{\mathcal{R}}\left(s, \overline{\psi}(s)\right) - \overline{\mathcal{R}}\left(s, \operatorname{Re}(\psi(s))\right)\right) \mathrm{d}s, \quad t \geq 0. \tag{4.83}$$

We then need to study the increments of $\overline{\mathcal{R}}$ in the second variable. Namely, fix $u_1, u_2 \in \mathbb{R}_-$ and use the definition (4.64) to write

$$\begin{aligned}\overline{\mathcal{R}}(t,u_1) - \overline{\mathcal{R}}(t,u_2) &= b(u_1-u_2) + \frac{c}{2}\left(u_1^2 - u_2^2\right) + \int_{\mathbb{R}_+}\left(e^{u_1\xi} - e^{u_2\xi} - (u_1-u_2)\xi\right)\nu(\mathrm{d}\xi)\\ &= \left[b + \frac{c}{2}(u_1+u_2)\right](u_1-u_2) + \int_{\mathbb{R}_+}\left(e^{u_1\xi} - e^{u_2\xi} - (u_1-u_2)\xi\right)\nu(\mathrm{d}\xi), \quad t \geq 0.\end{aligned}$$

Hence substituting $\overline{\psi}$ and $\operatorname{Re}\psi$ to u_1 and u_2, respectively, we have

$$\overline{\mathcal{R}}\left(t, \overline{\psi}(t)\right) - \overline{\mathcal{R}}\left(t, \operatorname{Re}(\psi(t))\right) = \underbrace{\left[b + \frac{c}{2}\left(\overline{\psi}(t) + \operatorname{Re}(\psi(t))\right)\right]}_{=z(t)}\delta(t) + \underbrace{\int_{\mathbb{R}_+}\left(e^{\overline{\psi}(t)\xi} - e^{\operatorname{Re}\psi(t)\xi} - \delta(t)\xi\right)\nu(\mathrm{d}\xi)}_{=w(t)}$$

for $t \geq 0$. Going back to (4.83),

$$\delta(t) = \gamma(t) + \int_0^t K(t-s)\left(z(s)\delta(s) + w(s)\right)\mathrm{d}s, \quad t \geq 0. \tag{4.84}$$

We aim to apply [5, Theorem C.1] in order to conclude $\delta \geq 0$ in $\mathbb{R}_+$.

- In the continuous case the integral in $\nu(\mathrm{d}\xi)$, i.e., the function w, simply disappears, hence the application of [5, Theorem C.1] is straightforward.
- In the jump case we need to deal with such an integral. Observe that the function w has opposite sign with respect to δ, so there is no hope of applying [5, Theorem C.1] unless we modify its expression. Fortunately this can be done using the mean value theorem, in combination with simple real–analysis arguments.

  First, for every $\xi > 0$ we define $f_\xi(u) = e^{\xi u}$, $u \in \mathbb{R}$, so $f_\xi'(u) = \xi e^{\xi u}$. Observe that the derivative f_ξ' is continuous and strictly increasing in $\mathbb{R}$, hence its inverse $h_\xi = \left(f_\xi'\right)^{-1}$ is continuous on $(0,\infty)$, as well. By the mean value theorem, for every $u_1, u_2 \in \mathbb{R}$ there exists $c_\xi \in [u_1 \wedge u_2, u_1 \vee u_2]$ such that

  $$f_\xi(u_2) - f_\xi(u_1) = f_\xi'(c_\xi)(u_2 - u_1).$$

  In particular $c_\xi \in (u_1 \wedge u_2, u_1 \vee u_2)$ when $u_1 \neq u_2$.
  Secondly, we consider the functions $\overline{\psi}$ and $\operatorname{Re}\psi$, and we can say that for every $t \in \mathbb{R}_+$ there exists $c_\xi(t) \in \left[\overline{\psi}(t) \wedge \operatorname{Re}\psi(t), \overline{\psi}(t) \vee \operatorname{Re}\psi(t)\right]$ (in the interior of such interval whenever $\overline{\psi} \neq \operatorname{Re}\psi$, i.e., whenever $\delta \neq 0$) such that

  $$e^{\xi\overline{\psi}(t)} - e^{\xi\operatorname{Re}\psi(t)} = f_\xi'(c_\xi(t))\left(\overline{\psi}(t) - \operatorname{Re}\psi(t)\right) = \xi e^{\xi c_\xi(t)}\delta(t). \tag{4.85}$$

By the axiom of choice we construct the function $c\colon \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_-$ defined by

$$c(\xi,t) = c_\xi(t), \quad \xi > 0,\, t \geq 0,$$

and $c(0,t) = 0$, $t \geq 0$. Note that the codomain of $c(\cdot,\cdot)$ is $\mathbb{R}_-$ since both $\bar{\psi}$ and $\operatorname{Re}\psi$ take values there. Recalling the definition of w and using (4.85) we can write

$$w(t) = \int_{\mathbb{R}_+} \left[e^{\bar{\psi}(t)\xi} - e^{\operatorname{Re}(\psi(t))\xi} - \delta(t)\xi\right] \nu(\mathrm{d}\xi) = \underbrace{\left(\int_{\mathbb{R}_+} \xi\left[e^{\xi c(\xi,t)} - 1\right]\nu(\mathrm{d}\xi)\right)}_{=\bar{w}(t)} \delta(t), \quad t \geq 0.$$

Now we have to prove that $\bar{w}$ is continuous on $\mathbb{R}_+$. The first step is to show the continuity of the function $c(\xi,\cdot)$ in $\mathbb{R}_+$ for every fixed $\xi \in \mathbb{R}_+$. It is of course trivial for $\xi = 0$, so we just focus on $\xi > 0$. If $\bar{t} \in (0,\infty)$ such that $\delta(\bar{t}) > 0$, then we can find $\epsilon > 0$ such that $\delta > 0$ in $(\bar{t}-\epsilon, \bar{t}+\epsilon)$. Hence we use (4.85) to prove that

$$c(\xi,t) = h_\xi\left(\frac{e^{\xi\bar{\psi}(t)} - e^{\xi\operatorname{Re}\psi(t)}}{\delta(t)}\right), \quad t \in (\bar{t}-\epsilon, \bar{t}+\epsilon),$$

recalling that $h_\xi = \left(f'_\xi\right)^{-1}$. So $c(\xi,\cdot)$ is continuous in the points $\bar{t} \in (0,\infty)$ where $\delta(\bar{t}) > 0$. An analogous reasoning shows the continuity in the points where $\delta < 0$. Consider now $\bar{t} \in \mathbb{R}_+$ a zero for δ, i.e., $\delta(\bar{t}) = 0$. For every sequence $(t_n)_n \subset \mathbb{R}_+$ such that $t_n \to \bar{t}$ as $n \to \infty$ one has, by construction,

$$\bar{\psi}(t_n) \wedge \operatorname{Re}\psi(t_n) \leq c(\xi,t_n) \leq \bar{\psi}(t_n) \vee \operatorname{Re}\psi(t_n), \quad n \in \mathbb{N}.$$

Therefore an application of the squeeze theorem gives

$$\lim_{n\to\infty} c(\xi,t_n) = \operatorname{Re}\psi(\bar{t}) = \bar{\psi}(\bar{t}) = c(\xi,\bar{t}).$$

At this point we deduce the continuity of the function $\bar{w}$ using the dominated convergence theorem. Indeed, take $t \in \mathbb{R}_+$, a sequence $t_n \to t$, and define $g_{(n)}(\xi) = \xi\left[e^{\xi c\left(\xi,t_{(n)}\right)} - 1\right]$, $\xi \in \mathbb{R}_+$. Then $g_n \to g$ pointwise in $\mathbb{R}_+$ by the continuity of $c(\xi,\cdot)$ and, for a certain $C > 0$ s.t. $t_n \leq C$, $n \in \mathbb{N}$ (which exists since $(t_n)_n$ is bounded), we have

$$\begin{aligned}\left|\xi\left[e^{\xi c(\xi,t_n)} - 1\right]\right| &= \xi\left[1 - e^{\xi c(\xi,t_n)}\right] \leq \xi\left[1 - e^{\xi \min_{0\leq s\leq C} c(\xi,s)}\right] \\ &\leq \xi\left[1 - e^{\xi \min_{0\leq s\leq C}\left\{\left(\operatorname{Re}\psi\wedge\bar{\psi}\right)(s)\right\}}\right] \in L^1(\mathrm{d}\nu), \quad n \in \mathbb{N}.\end{aligned}$$

Therefore we can rewrite (4.84) as follows

$$\delta(t) = \gamma(t) + \int_0^t K(t-s)\left(z(s) + \bar{w}(s)\right)\delta(s)\,\mathrm{d}s, \quad t \geq 0,$$

and we invoke [5, Theorem C.1] to assert that $\delta \geq 0$, i.e., that (4.65) holds.

Chapter 5

The rough Hawkes Heston stochastic volatility model

In this chapter, we propose a stochastic volatility model, called the *rough Hawkes Heston model*, where the spot variance is described by an affine stochastic Volterra equation of convolution type with jumps, see Chapter 4. As a result, our model incorporates both rough volatility and jump clustering phenomena. The affine framework of the model enables to efficiently price options on the underlying and the corresponding volatility index via Fourier inversion techniques. In Section 5.5, we calibrate a parsimonious specification of the model characterized by a power kernel and an exponential law for the jumps using S&P 500 and VIX options data. The results demonstrate that the rough Hawkes Heston model is able to jointly replicate the implied volatility smiles for both S&P 500 and VIX options with remarkable accuracy.

5.1 The model

We study a stochastic volatility model where the spot variance $\sigma^2 = (\sigma_t^2)_{t\geq 0}$ is a predictable process with trajectories in $L^2_{\text{loc}}(\mathbb{R}_+)$. It is defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ where the filtration $\mathbb{F}$ satisfies the usual conditions.

We consider, throughout our study, a kernel K that satisfies the next requirement, see [2, 3, 8, 37].

Hypothesis 5.1. *The kernel $K \in L^2_{\text{loc}}(\mathbb{R}_+)$ is nonnegative, nonincreasing, not identically zero and continuously differentiable on $(0, \infty)$. Furthermore, its resolvent of the first kind L is nonnegative and nonincreasing, i.e., $s \mapsto L([s, s+t])$ is nonincreasing for every $t \geq 0$.*

Note that Hypothesis 5.1 coincides with Hypothesis 4.2 in Chapter 4. We recall that, given a kernel $K \in L^1_{\text{loc}}(\mathbb{R}_+; \mathbb{R}^{d\times d})$, an $\mathbb{R}^{d\times d}-$valued measure L is called its (measure) resolvent of the first kind if $L * K = K * L = I$, where $I \in \mathbb{R}^{d\times d}$ is the identity matrix. The resolvent of the first kind does not always exist, but if it does then it is unique, see [96, Theorem 5.2, Chapter 5]. Under Hypothesis 5.1, the existence of the resolvent of the first kind is ensured by [96, Theorem 5.5, Chapter 5].

Let g_0 be a function representing the initial spot variance curve. A parametric form of g_0 will be specified for the application of the model (see Section 5.5). At this point, however, we only make the following assumption.

Hypothesis 5.2. *g_0 is continuous and nondecreasing, with $g_0(0) \geq 0$.*

Fix $b \in \mathbb{R}$, $c > 0$, an $\mathbb{F}-$Brownian motion $W_2 = (W_{2,t})_{t \geq 0}$ and a nonnegative measure ν on $\mathbb{R}_+$ such that $\nu(\{0\}) = 0$ and that $\int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z) < \infty$. We assume that the spot variance σ^2 is a $\mathbb{Q} \otimes \mathrm{d}t-$a.e. nonnegative predictable process satisfying

$$\sigma^2 = g_0 + K * \mathrm{d}Z, \quad \mathbb{Q} \otimes \mathrm{d}t - \text{a.e.}, \tag{5.1}$$

where Z is the following semimartingale having jump measure $\mu(\mathrm{d}t, \mathrm{d}z)$ and compensator $\sigma_t^2 \mathrm{d}t \otimes \nu(\mathrm{d}z)$:

$$\mathrm{d}Z_t = b\, \sigma_t^2 \mathrm{d}t + \sqrt{c}\, \sigma_t\, \mathrm{d}W_{2,t} + \int_{\mathbb{R}_+} z\, \left(\mu\left(\mathrm{d}t, \mathrm{d}z\right) - \sigma_t^2 \mathrm{d}t \otimes \nu\left(\mathrm{d}z\right)\right), \quad Z_0 = 0.$$

Therefore, the instantaneous variance σ^2 satisfies a stochastic affine Volterra equation of convolution type with jumps. From now on, we denote by $\tilde{\mu}(\mathrm{d}t, \mathrm{d}z) = \mu(\mathrm{d}t, \mathrm{d}z) - \sigma_t^2 \mathrm{d}t \otimes \nu(\mathrm{d}z)$ the compensated jump measure of Z. Since the intensity of the jumps of σ^2 is proportional to σ^2 itself, the spot variance is a Hawkes-type process, which is coherent with other models that incorporate endogeneity of financial markets such as [26, 46, 79, 93, 113]. In the sequel, we denote by $\widetilde{Z} = (\widetilde{Z}_t)_{t \geq 0}$ the process $\mathrm{d}\widetilde{Z}_t = \sqrt{c}\, \sigma_t\, \mathrm{d}W_{2,t} + \int_{\mathbb{R}_+} z\, \tilde{\mu}(\mathrm{d}t, \mathrm{d}z)$, $t \geq 0$, with starting condition $\widetilde{Z}_0 = 0$. Notice that $\widetilde{Z}$ is a square-integrable martingale by [37, Lemma 1], see also Lemma 4.1 in Chapter 4.

According to [37, Appendix A], see Appendix 4.A of Chapter 4,

$$\sigma^2 = g_0 - R_{-bK} * g_0 + E_{b,K} * \mathrm{d}\widetilde{Z}, \quad \mathbb{Q} \otimes \mathrm{d}t - \text{a.e.}, \tag{5.2}$$

where R_{-bK} is the resolvent of the second kind of $-bK$ and $E_{b,K}$ is the canonical resolvent of K with parameter b. We recall that the resolvent of the second kind R_K for a kernel $K \in L^1_{\text{loc}}(\mathbb{R}_+)$ is the unique solution $R_K \in L^1_{\text{loc}}(\mathbb{R}_+)$ of the two equations $K * R_K = R_K * K = K - R_K$, see [96, Theorem 3.1, Chapter 2] and the subsequent definition. The canonical resolvent $E_{\lambda,K}$ of K with parameter λ is defined by $E_{\lambda,K} = -\lambda^{-1} R_{-\lambda K}$ for $\lambda \neq 0$, whereas $E_{0,K} = K$.

Remark 5.1. *If we assume that K and the shifted kernels $K(\cdot + 1/n)$, $n \in \mathbb{N}$, satisfy Hypothesis 5.1, then under Hypothesis 5.2 the (weak) existence of the spot variance process σ^2, satisfying* (5.1)*, is ensured by [2, Theorem 2.13] and [37, Lemma 12] (see also Lemma 4.12). Assuming weak existence, weak uniqueness is established in [37, Corollary* 15*] under Hypothesis 5.1, see Corollary 4.15 in Chapter 4. We refer to [3] and [37] for more information about stochastic Volterra equations and stochastic convolution for processes with jumps.*

A useful tool for the development of the theory is the adjusted forward process, which we now define as in Equation (4.25). For every $t \geq 0$, it is denoted by $(g_t(s))_{s>t}$ and it is a jointly measurable process on $\Omega \times (t, \infty)$ such that

$$g_t\left(s\right) = g_0\left(s\right) + \int_0^t K\left(s - r\right) \mathrm{d}Z_r, \quad \mathbb{Q} - \text{a.s.}, \; s > t. \tag{5.3}$$

Thanks to [135, Theorem 46] and the fact that $\mathbb{F}$ satisfies the usual conditions, we can consider $g_t(\cdot)$ to be $\mathcal{F}_t \otimes \mathcal{B}(t, \infty)-$measurable.
Analogous arguments provide a version of the conditional expectation process $\mathbb{E}[\sigma^2 | \mathcal{F}_t] = (\mathbb{E}[\sigma_s^2 | \mathcal{F}_t])_{s>t}$ which is $\mathcal{F}_t \otimes \mathcal{B}(t, \infty)-$measurable. In particular, from (5.2) (cfr. (4.73))

$$\mathbb{E}\left[\sigma_s^2 \middle| \mathcal{F}_t\right] = g_0(s) - (R_{-bK} * g_0)(s) + \int_0^t E_{b,K}\left(s - r\right) \mathrm{d}\widetilde{Z}_r, \quad \mathbb{Q} - \text{a.s.}, \; s > t. \tag{5.4}$$

We now prescribe the dynamics of the log returns process $X = (X_t)_{t\geq 0}$ as follows:

$$\mathrm{d}X_t = -\left(\frac{1}{2} + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 + \Lambda z\right) \nu\left(\mathrm{d}z\right)\right) \sigma_t^2 \,\mathrm{d}t + \sigma_t \left(\sqrt{1-\rho^2}\,\mathrm{d}W_{1,t} + \rho\,\mathrm{d}W_{2,t}\right) - \Lambda \int_{\mathbb{R}_+} z\, \tilde{\mu}\left(\mathrm{d}t, \mathrm{d}z\right), \quad X_0 = 0, \quad (5.5)$$

where $\rho \in [-1, 1]$ is a correlation parameter, $W_1 = (W_{1,t})_{t\geq 0}$ is an $\mathbb{F}-$Brownian motion independent from W_2 and $\Lambda \geq 0$ is a leverage parameter forcing common jumps for volatility and underlying with opposite signs. This is coherent with empirical findings in [171], stylized features studied in [52], and the financial/econometric literature with jumps, e.g. [16, 17, 19, 26, 60, 154, 166]. We have assumed, for the sake of readability and without loss of generality, that interest rates and dividends are zero. The price process of the underlying asset will be $S = (S_t)_{t\geq 0} = (S_0 e^{X_t})_{t\geq 0}$, where $S_0 > 0$ represents the initial price. An application of Itô's formula shows that S is a local martingale. Indeed,

$$\begin{aligned}\frac{\mathrm{d}S_t}{S_{t-}} &= -\left(\frac{1}{2} + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 + \Lambda z\right) \nu\left(\mathrm{d}z\right)\right) \sigma_t^2 \mathrm{d}t + \sigma_t \left(\sqrt{1-\rho^2}\,\mathrm{d}W_{1,t} + \rho\,\mathrm{d}W_{2,t}\right) \\ &\quad - \Lambda \int_{\mathbb{R}_+} z\, \tilde{\mu}\left(\mathrm{d}t, \mathrm{d}z\right) + \frac{1}{2}\sigma_t^2 \,\mathrm{d}t + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 + \Lambda z\right) \mu\left(\mathrm{d}t, \mathrm{d}z\right) \\ &= \sigma_t \left(\sqrt{1-\rho^2}\,\mathrm{d}W_{1,t} + \rho\,\mathrm{d}W_{2,t}\right) + \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1\right) \tilde{\mu}\left(\mathrm{d}t, \mathrm{d}z\right) =: \mathrm{d}N_t,\end{aligned}$$

where $N = (N_t)_{t\geq 0}$ is a local martingale with $N_0 = 0$. In particular, since S starts at S_0, it follows that $S = S_0 \mathcal{E}(N)$, where $\mathcal{E}$ denotes the Doléans-Dade exponential. In the next section, see Corollary 5.4, we will improve on this result by showing that, for every $T > 0$, the restriction of S to $[0, T]$ is a true martingale.

5.2 The Fourier–Laplace transform of the log returns

In this section we study, for a fixed $T \geq 0$, the conditional Fourier-Laplace transform of X_T, $\mathbb{E}[e^{wX_T} | \mathcal{F}_t]$, $t \in [0, T]$. Here $w \in \mathbb{C}$ is subject to suitable conditions that will be specified in the sequel. In particular, we want to find a formula that allow us to compute the prices of options written on the underlying asset using Fourier-inversion techniques [74, 75, 81, 95]. We will adopt the following notation: for $z \in \mathbb{C}$ we denote by $\operatorname{Re} z$ and $\operatorname{Im} z$ the real and imaginary parts of z, respectively. We let $\mathbb{C}_+$ [resp., $\mathbb{C}_-$] be the set of complex numbers with nonnegative [resp., nonpositive] real part.

Let us define the mapping $\mathcal{R} \colon \mathbb{C}_+ \times \mathbb{C}_- \to \mathbb{C}$ by

$$\mathcal{R}\left(u, v\right) = \frac{1}{2}\left(u^2 - u\right) + \left(b + \rho\sqrt{c}\, u\right) v + \frac{c}{2} v^2 + \int_{\mathbb{R}_+} \left[e^{(v - \Lambda u)z} - u\left(e^{-\Lambda z} - 1\right) - 1 - vz\right] \nu\left(\mathrm{d}z\right), \quad (5.6)$$

for every $(u, v) \in \mathbb{C}_+ \times \mathbb{C}_-$. For the development of the theory we need the following result about deterministic Riccati-Volterra equations, whose proof is postponed to Appendix 5.A.

Theorem 5.1. *Suppose that K satisfies Hypothesis 5.1 and $w \in \mathbb{C}$ is such that $\operatorname{Re} w \in [0, 1]$.*

(i) *There exists a unique continuous solution $\psi_w \colon \mathbb{R}_+ \to \mathbb{C}_-$ of the Riccati-Volterra equation*

$$\psi_w\left(t\right) = \int_0^t K\left(t - s\right) \mathcal{R}\left(w, \psi_w\left(s\right)\right) \mathrm{d}s = \left(K * \mathcal{R}\left(w, \psi_w\left(\cdot\right)\right)\right)\left(t\right), \quad t \geq 0. \quad (5.7)$$

In particular, $\psi_{\operatorname{Re} w}$ is $\mathbb{R}_-$-valued.

(ii) *The following inequalities hold:*

$$\operatorname{Re}\psi_w(t) \leq \psi_{\operatorname{Re} w}(t) \leq 0, \quad t \geq 0. \tag{5.8}$$

We also need the next preparatory lemma, which can be proven similarly to [2, Lemma 6.1].

Lemma 5.2. *Let f_1, f_2, $f_3\colon [0,T] \to \mathbb{R}$ be bounded measurable functions such that $f_3 \leq 0$ in $[0,T]$. Then, the Doléans-Dade exponential*

$$\mathcal{E}\left(\int_0^t f_1(s)\,\sigma_s\,\mathrm{d}W_{1,s} + \int_0^t f_2(s)\,\sigma_s\,\mathrm{d}W_{2,s} + \int_0^t \int_{\mathbb{R}_+} \left(e^{f_3(s)z} - 1\right)\tilde{\mu}(\mathrm{d}s,\mathrm{d}z)\right), \quad t \in [0,T]$$

is a martingale.

We are now ready to state the main result of this section. We introduce for every $\varepsilon \in \mathbb{R}$ the shift operator Δ_ε, which, given $I \subset \mathbb{R}$ and a function $f\colon I \to \mathbb{C}$, assigns the function $\Delta_\varepsilon f\colon I - \varepsilon \to \mathbb{C}$ defined by $\Delta_\varepsilon f(t) = f(t+\varepsilon)$, $t \in I - \varepsilon$.

Theorem 5.3. *Suppose that K satisfies Hypothesis 5.1 and that the resolvent of the first kind L is the sum of a locally integrable function and a point mass at 0. Moreover, suppose that the total variation bound*

$$\sup_{\varepsilon \in (0,\overline{T}]} \|\Delta_\varepsilon K * L\|_{TV([0,\overline{T}])} < \infty$$

holds for all $\overline{T} > 0$. Then, under Hypothesis 5.2, for every $w \in \mathbb{C}$ such that $\operatorname{Re} w \in [0,1]$,

$$\mathbb{E}\left[\exp\{wX_T\}\,\middle|\,\mathcal{F}_t\right] = \exp\left\{\widetilde{V}_t(w,T)\right\}, \quad \mathbb{Q}-a.s.,\ t \in [0,T], \tag{5.9}$$

where $\widetilde{V}_t(w,T) = wX_t + \int_t^T \mathcal{R}(w,\psi_w(T-s))g_t(s)\mathrm{d}s$, $t \in [0,T]$.

Proof. Let $w \in \mathbb{C}$ be such that $\operatorname{Re} w \in [0,1]$. Define the càdlàg, adapted, $\mathbb{C}-$valued semimartingale $(V_t(w,T))_{t\in[0,T]}$ by

$$\begin{aligned} V_t(w,T) &= V_0(w,T) + wX_t + \int_0^t \psi_w(T-s)\,\mathrm{d}\widetilde{Z}_s \\ &\quad - \int_0^t \Bigg(\frac{1}{2}\left(w^2 - w\right) + \rho\sqrt{c}\,w\,\psi_w(T-s) + \frac{c}{2}\psi_w(T-s)^2 \\ &\quad + \int_{\mathbb{R}_+}\left(e^{(-\Lambda w + \psi_w(T-s))z} - w\left(e^{-\Lambda z} - 1\right) - 1 - \psi_w(T-s)\,z\right)\nu(\mathrm{d}z)\Bigg)\sigma_s^2\,\mathrm{d}s, \end{aligned} \tag{5.10}$$

$$V_0(w,T) = \int_0^T \mathcal{R}(w,\psi_w(T-s))\,g_0(s)\,\mathrm{d}s. \tag{5.11}$$

The same arguments as in the proof of [37, Theorem 5] (see Theorem 4.5 in Chapter 4), which essentially rely on the stochastic Fubini's theorem (see, e.g., [153, Theorem 65, Chapter IV]), allow us to prove that

$$V_t(w,T) = \widetilde{V}_t(w,T), \quad \mathbb{Q}-\text{a.s.},\ t \in [0,T]. \tag{5.12}$$

We now define $H(w,T) = (H_t(w,T))_{t\in[0,T]} = (\exp\{V_t(w,T)\})_{t\in[0,T]}$. By Itô's formula and the dynamics in (5.5) and (5.10) we have, omitting (w,T) for sake of readability,

$$\begin{aligned}\frac{\mathrm{d}H_t}{H_{t-}} &= \Bigg[w\,\mathrm{d}X_t - \Bigg(\frac{c}{2}\,\psi_w\left(T-t\right)^2 + \int_{\mathbb{R}_+}\left(e^{(-\Lambda w+\psi_w(T-t))z} - 1 - w\left(e^{-\Lambda z}-1\right) - \psi_w\left(T-t\right)z\right)\nu\left(\mathrm{d}z\right)\\ &\quad + \frac{1}{2}\left(w^2-w\right) + \rho\sqrt{c}\,w\,\psi_w\left(T-t\right)\Bigg)\sigma_t^2\,\mathrm{d}t + \psi_w\left(T-t\right)\mathrm{d}\widetilde{Z}_t\Bigg] + \frac{1}{2}\left(c\,\psi_w\left(T-t\right)^2 + w^2\right)\sigma_t^2\,\mathrm{d}t\\ &\quad + \rho\sqrt{c}\,w\,\psi_w\left(T-t\right)\sigma_t^2\,\mathrm{d}t + \int_{\mathbb{R}_+}\left(e^{(-\Lambda w+\psi_w(T-t))z} - 1 - \left(-\Lambda w + \psi_w\left(T-t\right)\right)z\right)\mu\left(\mathrm{d}t,\mathrm{d}z\right)\\ &= \left[\sigma_t\left(w\sqrt{1-\rho^2}\,\mathrm{d}W_{1,t} + \left(w\rho + \sqrt{c}\,\psi_w\left(T-t\right)\right)\mathrm{d}W_{2,t}\right) + \int_{\mathbb{R}_+}\left(e^{(-\Lambda w+\psi_w(T-t))z} - 1\right)\tilde{\mu}\left(\mathrm{d}t,\mathrm{d}z\right)\right],\end{aligned}$$

with $H_0 = \exp(V_0)$. We define $N(w,T) = (N_t(w,T))_{t\in[0,T]}$ by $N_0(w,T) = 0$ and

$$\begin{aligned}\mathrm{d}N_t\left(w,T\right) &= \sigma_t\left(w\sqrt{1-\rho^2}\,\mathrm{d}W_{1,t} + \left(w\rho + \sqrt{c}\,\psi_w\left(T-t\right)\right)\mathrm{d}W_{2,t}\right)\\ &\quad + \int_{\mathbb{R}_+}\left(e^{(-\Lambda w+\psi_w(T-t))z} - 1\right)\tilde{\mu}\left(\mathrm{d}t,\mathrm{d}z\right).\end{aligned}$$

Then $N(w,T)$ is a local martingale and the previous computations show that, omitting again (w,T), $H = \exp\{V_0\}\mathcal{E}(N)$ up to evanescence, where $\mathcal{E}$ denotes the Doléans-Dade exponential. Therefore $H(w,T)$ is a local martingale. If it is indeed a true martingale, then (5.9) directly follows from (5.12) noting also that $\widetilde{V}_T(w,T) = wX_T$.

In order to argue the martingale property of $H(w,T)$, first we observe that by Lemma 5.2 the real-valued process $H(\mathrm{Re}\,w,T) = (H_t(\mathrm{Re}\,w,T))_{t\in[0,T]} = (\exp\{V_t(\mathrm{Re}\,w,T)\})_{t\in[0,T]}$ is a true martingale. Secondly, we invoke [37, Corollary 11], see Corollary 4.11 in Chapter 4, to obtain the following alternative expression for $V(w,T)$ (an analogous one holds for $V(\mathrm{Re}\,w,T)$)

$$\begin{aligned}V_t\left(w,T\right) = wX_t + \int_0^{T-t}\mathcal{R}\left(w,\psi_w\left(s\right)\right)g_0\left(T-s\right)\mathrm{d}s + \psi_w\left(T-t\right)L\left(\{0\}\right)\left(\sigma^2-g_0\right)\left(t\right)\\ + \left(\mathrm{d}\Pi_{T-t} * \left(\sigma^2-g_0\right)\right)\left(t\right), \quad \text{for a.e. } t\in\left(0,T\right), \mathbb{Q}-\text{a.s.,} \quad (5.13)\end{aligned}$$

where for every $\varepsilon > 0$, $\Pi_\varepsilon(t) = \int_0^\varepsilon \mathcal{R}(w,\psi_w(s))(\Delta_{\varepsilon-s}K * L)(t)\mathrm{d}s$, $t\geq 0$, is a locally absolutely continuous function. The application of this result is legitimate because the procedure carried out in [37] (see also Chapter 4) to infer (5.13) only depends on (5.1), (5.7) and the boundedness on compact intervals of $\mathbb{R}_+$ of $\mathcal{R}(w,\psi_w(\cdot))$, and does not rely on the expression of $\mathcal{R}$. A similar argument together with (5.8) and Hypothesis 5.2 allows us to parallel the comparison method in the proof of [37, Theorem 14], see Theorem 4.14 in Chapter 4, to conclude that there is a constant $C > 0$ such that

$$\left|H_t\left(w,T\right)\right| = \left|\exp\left\{V_t\left(w,T\right)\right\}\right| = \exp\left\{\mathrm{Re}\,V_t\left(w,T\right)\right\} \leq C\exp\left\{V_t\left(\mathrm{Re}\,w,T\right)\right\} = CH_t\left(\mathrm{Re}\,w,T\right),$$

for $t\in[0,T]$, $\mathbb{Q}-$a.s. At this point it is sufficient to invoke [112, Lemma 1.4] to claim that $H(w,T)$ is a true martingale, hence the proof is complete. ■

From the previous theorem we deduce the martingale property of our price process S with a direct approach (it can also be obtained by Lemma 5.2).

Corollary 5.4. *Under the hypotheses of Theorem 5.3, the price process $S = (S_t)_{t\in[0,T]}$ is a martingale.*

Proof. The computations at the end of Section 5.1 show that the stock price S is a nonnegative local martingale, hence it is a supermartingale. In order for it to be a martingale, it is sufficient to show that $\mathbb{E}[S_T] = S_0$. By (5.9) in Theorem 5.3 with $w = 1$ we have

$$\mathbb{E}\left[S_T\right] = S_0 \exp\left\{\int_0^T \mathcal{R}\left(1, \psi_1\left(T-s\right)\right) g_0\left(s\right) \mathrm{d}s\right\}.$$

From (5.6)-(5.7), we observe that $\psi_1 \equiv 0$ in $\mathbb{R}_+$. This implies that $\mathcal{R}(1, \psi_1(\cdot)) = 0$ in $\mathbb{R}_+$, which concludes the proof. ■

Equation (5.9) in Theorem 5.3 gives a semi-explicit expression to compute the Fourier-Laplace transform Ψ^{X_T} of X_T in a suitable region of $\mathbb{C}$, namely

$$\Psi^{X_T}\left(w\right) = \exp\left\{\int_0^T \mathcal{R}\left(w, \psi_w\left(T-s\right)\right) g_0\left(s\right) \mathrm{d}s\right\}, \quad w \in \mathbb{C} \text{ such that } \operatorname{Re} w \in [0,1]. \tag{5.14}$$

As shown in the following proposition, whose proof is in Appendix 5.B, we can use Ψ^{X_T} to price options with maturity T on the underlying asset S via Fourier-inversion techniques.

Proposition 5.5. *Fix a log strike $k > 0$. Then, under the hypotheses of Theorem 5.3, the price $C_S(k,T)$ of a call option on the underlying asset S with log strike k and maturity T is*

$$C_S\left(k,T\right) = S_0 - \frac{1}{\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}_+} \operatorname{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\right]\frac{1}{\frac{1}{4}+\lambda^2}\mathrm{d}\lambda, \tag{5.15}$$

and the price $P_S(k,T)$ of a put option with the same log strike, maturity and underlying is

$$P_S\left(k,T\right) = e^k - \frac{1}{\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}_+} \operatorname{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\right]\frac{1}{\frac{1}{4}+\lambda^2}\mathrm{d}\lambda. \tag{5.16}$$

Remark 5.2. *The expression in* (5.15) *coincides with [130, Formula* (3.11)*], but we have to independently prove it (see Appendix 5.B). Indeed, in [130] the author obtains* (5.15) *starting from the inversion of the generalized Fourier transform of the payoff function $w(x) = (e^x - e^k)^+$, $x \in \mathbb{R}$, of a call option with log strike k (here x represents the log price). Namely, for $x \in \mathbb{R}$,*

$$w\left(x\right) = -\frac{1}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} \frac{e^{k(iz+1)}}{z^2 - iz} e^{-izx}\mathrm{d}z, \quad z_i > 1.$$

If we were to follow the same approach here, then we would find a problem: we only have proved that Ψ^{X_T} is defined for complex numbers with real part in $[0,1]$. Therefore, in the previous expression, we would need $z_i \in [0,1]$, which is a contradiction. This setback cannot be immediately fixed by considering put options and then applying the put-call parity formula, because again the intersection between the complex strip ($z_i < 0$), where the Fourier transform for the payoff function is defined, and the strip where $\Psi^{X_T}(-i\,\cdot)$ is available is empty. We refer to [165, Section 4] for a survey of pricing based on Fourier-inversion techniques.

5.3 The Fourier–Laplace transform of VIX^2

In this section the underlying asset S represents the SPX index. Then, according to the CBOE VIX white paper and [68], the theoretical value of VIX=$(\mathrm{VIX}_T)_{T\geq 0}$ is

$$\mathrm{VIX}_T = \sqrt{\left(-\frac{2}{\delta}\mathbb{E}\left[X_{T+\delta} - X_T \middle| \mathcal{F}_T\right]\right)^+} \times 100, \quad T \geq 0. \tag{5.17}$$

Here $\delta = \frac{1}{12}$ and represents 30 days, the time to expiration of the log contracts involved in the computation of the index. Note that in (5.17), the positive part has been inserted to guarantee the good definition of the random variable VIX_T in the whole space Ω, however the radicand is nonnegative $\mathbb{Q}-$a.s., as we are about to show.

We first derive, in the following theorem, an expression for $\mathbb{E}[X_{T+\delta} - X_T|\mathcal{F}_T]$, $T \geq 0$, in terms of the adjusted forward process at time T, $g_T(\cdot)$.

Theorem 5.6. *The log contract satisfies an infinite dimensional affine relation with respect to the adjusted forward process. More specifically,*

$$\mathbb{E}\left[X_{T+\delta} - X_T \middle| \mathcal{F}_T\right] = c_1 \int_T^{T+\delta} \left(1 + b\left(E_{b,K} * 1\right)\left(T + \delta - s\right)\right) g_T\left(s\right) \mathrm{d}s \leq 0, \quad \mathbb{Q} - a.s., \tag{5.18}$$

where $c_1 = -(\frac{1}{2} + \int_{\mathbb{R}_+}(e^{-\Lambda z} - 1 + \Lambda z)\nu(\mathrm{d}z))$.

Proof. By (5.5) and the martingale property of the local martingale part of the expression (see [37, Lemma 1] or Lemma 4.1 in Chapter 4), we have

$$\mathbb{E}\left[X_{T+\delta} - X_T \middle| \mathcal{F}_T\right] = c_1 \int_T^{T+\delta} \mathbb{E}\left[\sigma_s^2 \middle| \mathcal{F}_T\right] \mathrm{d}s, \quad \mathbb{Q} - \text{a.s.}$$

Recalling that $\sigma^2 \geq 0$, $\mathbb{Q} \otimes \mathrm{d}t-$a.e., we infer that $\mathbb{E}[\sigma_s^2|\mathcal{F}_T] \geq 0$ for a.e. $s > T$, $\mathbb{Q}-$a.s., hence the value of a log contract at time T is nonpositive $\mathbb{Q}-$a.s.

By (5.2), (5.4), the stochastic Fubini's theorem – whose application is guaranteed by [37, Lemma 1], see also Lemma 4.1 – and a suitable change of variables, we infer that, $\mathbb{Q}-$a.s.

$$\begin{aligned}
c_1^{-1}\mathbb{E}\left[X_{T+\delta} - X_T \middle| \mathcal{F}_T\right] &= \int_0^{T+\delta} f_0\left(s\right) \mathrm{d}s - \int_0^T \sigma_s^2 \,\mathrm{d}s + \int_0^{T+\delta} \left(\int_0^T 1_{\{r\leq s\}} E_{b,K}\left(s - r\right) \mathrm{d}\widetilde{Z}_r\right) \mathrm{d}s \\
&= \int_0^{T+\delta} f_0\left(s\right) \mathrm{d}s - \int_0^T \sigma_s^2 \,\mathrm{d}s + \int_0^T \left(E_{b,K} * 1\right)\left(T + \delta - r\right) \mathrm{d}\widetilde{Z}_r \\
&= \int_0^{T+\delta} f_0\left(s\right) \mathrm{d}s - \int_0^T \left(1 + b\left(E_{b,K} * 1\right)\right)\left(T + \delta - s\right) \sigma_s^2 \,\mathrm{d}s + \int_0^T \left(E_{b,K} * 1\right)\left(T + \delta - r\right) \mathrm{d}Z_r,
\end{aligned} \tag{5.19}$$

where $f_0 = g_0 - R_{-bK} * g_0$. Notice that $E_{b,K} * 1$ is the unique, continuous (nonnegative) solution of the linear Volterra equation $\chi = K * (1 + b\chi)$. Then, another application of stochastic Fubini's theorem yields, $\mathbb{Q}-$a.s.,

$$\begin{aligned}
\int_0^T \left(E_{b,K} * 1\right)\left(T + \delta - r\right) \mathrm{d}Z_r &= \int_0^T \left(\int_r^{T+\delta} K\left(s - r\right)\left(1 + b\left(E_{b,K} * 1\right)\left(T + \delta - s\right)\right) \mathrm{d}s\right) \mathrm{d}Z_r \\
&= \int_0^{T+\delta} \left(1 + b\left(E_{b,K} * 1\right)\left(T + \delta - s\right)\right) \left(\int_0^T 1_{\{r\leq s\}} K\left(s - r\right) \mathrm{d}Z_r\right) \mathrm{d}s.
\end{aligned}$$

To conclude, we observe that by [96, Theorem 2.2 (viii), Chapter 2]

$$-\left(\left(R_{-bK} * g_0\right) * 1\right)(T+\delta) = b\left(\left(E_{b,K} * 1\right) * g_0\right)(T+\delta),$$

and plugging the previous two equalities in (5.19), together with (5.1), (5.3), we obtain the relation in (5.18). ■

We deduce the following corollary showing an affine relation between the square of the VIX index and the adjusted forward process.

Corollary 5.7. *The square of* VIX *satisfies an infinite dimensional affine relation with respect to the adjusted forward process. More specifically*

$$\begin{aligned} \mathrm{VIX}_T^2 &= -10^4 \frac{2}{\delta} \mathbb{E}\left[X_{T+\delta} - X_T \middle| \mathcal{F}_T\right], \quad \mathbb{Q}-a.s. \\ &= -10^4 \frac{2}{\delta} c_1 \int_T^{T+\delta} \left(1 + b\left(E_{b,K} * 1\right)(T+\delta-s)\right) g_T(s)\, \mathrm{d}s, \quad \mathbb{Q}-a.s., \end{aligned} \tag{5.20}$$

where $c_1 = -(\frac{1}{2} + \int_{\mathbb{R}_+} (e^{-\Lambda z} - 1 + \Lambda z)\nu(\mathrm{d}z))$.

Remark 5.3. *Our framework also allows us to obtain an explicit infinite dimensional affine relation between the variance swaps and the adjusted forward process. Specifically, the variance swap rate is*

$$\frac{1}{\delta}\mathbb{E}\left[[X,X]_{T+\delta} - [X,X]_T \middle| \mathcal{F}_T\right] = \frac{c_2}{\delta} \int_T^{T+\delta} \left(1 + b\left(E_{b,K} * 1\right)(T+\delta-s)\right) g_T(s)\, \mathrm{d}s, \quad \mathbb{Q}-a.s., \tag{5.21}$$

where $c_2 = 1 + \Lambda^2 \int_{\mathbb{R}_+} |z|^2 \nu(\mathrm{d}z)$*. Note that for* $\Lambda = 0$ *we have* $c_2 = -2c_1$*, hence in this case log contracts and variance swaps coincide up to the factor* $-2/\delta$ *(see* (5.18)*-*(5.21)*). Therefore, when there are no jumps in the dynamics of the underlying, by* (5.20) *we recover the fact that* VIX^2 *is a variance swap. Moreover, observe that the relation in* (5.21) *is an extension of [118, Lemma 4.4] in the classical affine setting. We refer to [54, 68, 132] for more details regarding the distinction between variance swaps and* VIX^2.

We are now interested in finding the conditional Fourier-Laplace transform of VIX_T^2. Before addressing this question, we need some technical intermediate steps. We first recall the following functional space as defined in [4], see also (4.82) and the subsequent phrase.

$$\begin{aligned} \mathcal{G}_K = \{ g \colon \mathbb{R}_+ \to \mathbb{R} \text{ continuous} :\ & g(0) \geq 0 \text{ and} \\ & \Delta_\varepsilon g - \left(\Delta_\varepsilon K * L\right)(0)\, g - \mathrm{d}\left(\Delta_\varepsilon K * L\right) * g \geq 0,\ \varepsilon \geq 0 \}. \end{aligned} \tag{5.22}$$

Lemma 5.8. *Suppose that* K *satisfies Hypothesis 5.1. Define the function* $h \colon \mathbb{R}_+ \to \mathbb{R}$ *by*

$$h(t) = -10^4 \frac{2}{\delta} c_1 \left[1 + b\left(E_{b,K} * 1\right)(\delta - t)\right] 1_{\{t \leq \delta\}}, \quad t \geq 0.$$

Then h *is a continuous nonnegative function on* $[0,\delta)$ *and* $t \mapsto \int_{\mathbb{R}_+} h(s)K(s+t)\mathrm{d}s$ *belongs to* $\mathcal{G}_K$.

Proof. The first step is to show that $1 + b(E_{b,K} * 1) \geq 0$ in $\mathbb{R}_+$, which implies that h is also nonnegative. This can be deduced from the fact that this function is the unique, continuous solution in $\mathbb{R}_+$ of the Volterra equation $\chi = 1 + bK * \chi$, which is nonnegative by [5, Theorem C.1]. Secondly, h has compact support, and under Hypothesis 5.1 for every $\varepsilon \geq 0$ the function $\Delta_\varepsilon K * L$ is right-continuous nondecreasing in $\mathbb{R}_+$ and (see the proof of [8, Lemma 2.6])

$$\Delta_\varepsilon K = (\Delta_\varepsilon K * L)(0)\, K + \mathrm{d}(\Delta_\varepsilon K * L) * K, \quad \mathrm{d}t - \text{a.e. in } \mathbb{R}_+.$$

As a consequence, for every $t \geq 0$

$$\begin{aligned}\Delta_\varepsilon K(s+t) &= (\Delta_\varepsilon K * L)(0)\, K(s+t) + (\mathrm{d}(\Delta_\varepsilon K * L) * K)(s+t) \\ &\geq (\Delta_\varepsilon K * L)(0)\, K(s+t) + \int_0^t K(s+t-u)\, \mathrm{d}(\Delta_\varepsilon K * L)(u), \quad \text{for a.e. } s \in [0,\delta].\end{aligned}$$

This implies, by Tonelli's theorem, that $t \mapsto \int_{\mathbb{R}_+} h(s)K(s+t)\mathrm{d}s$ belongs to $\mathcal{G}_K$. ■

We now define, for every $w \in \mathbb{C}_-$, the function $h_w(t) = w \cdot h(t)$, $t \geq 0$, and consider the Riccati-Volterra equation

$$\phi_w = \int_0^\infty h_w(s)\, K(s+\cdot)\, \mathrm{d}s + K * (G(\phi_w(\cdot))), \tag{5.23}$$

where

$$G(u) = bu + \frac{c}{2}u^2 + \int_{\mathbb{R}_+} (e^{uz} - 1 - uz)\, \nu(\mathrm{d}z), \quad u \in \mathbb{C}_-. \tag{5.24}$$

Lemma 5.9. *Suppose that K satisfies Hypothesis 5.1. For every $w \in \mathbb{C}_-$, there exists a unique continuous solution $\phi_w \colon \mathbb{R}_+ \to \mathbb{C}_-$ to (5.23). Moreover,*

$$\operatorname{Re} \phi_w(t) \leq \phi_{\operatorname{Re} w}(t), \quad t \geq 0. \tag{5.25}$$

Proof. Having in mind [5, Theorem C.1], the existence of a global solution of (5.23) can be deduced as in [37, Theorem 13] (see Theorem 4.13 in Chapter 4), whereas the uniqueness of such ϕ_w is obtained with a procedure analogous to the proof of Theorem 5.1, see *Step III* with $\Lambda = 0$ in Appendix 5.A. Moreover, again by analogy with [37, Theorem 13 (ii)], see Theorem 4.13 (ii), the comparison result (5.25) holds. ■

Before stating the theorem that provides the conditional Fourier-Laplace transform of VIX^2_T, we define

$$\Phi_w(t,s) = h_w(s-t)\, 1_{\{s \geq t\}} + G(\phi_w(t-s))\, 1_{\{s<t\}}, \quad t,s \geq 0. \tag{5.26}$$

Theorem 5.10. *Assume the same hypotheses as in Theorem 5.3. Then, for every $w \in \mathbb{C}_-$,*

$$\mathbb{E}\left[\exp\left\{w \cdot \mathrm{VIX}^2_T\right\} \middle| \mathcal{F}_t\right] = \exp\left\{\widetilde{U}_t(w,T)\right\}, \quad \mathbb{Q} - a.s.,\ t \in [0,T], \tag{5.27}$$

where $\widetilde{U}_t(w,T) = \int_t^\infty \Phi_w(T,s) g_t(s)\mathrm{d}s$, $t \in [0,T]$.

Proof. Fix $w \in \mathbb{C}_-$. First of all, notice that by the definition of h_w and (5.20)

$$\begin{aligned}\widetilde{U}_T(w,T) = \int_T^\infty h_w(s-T)\, g_T(s)\, \mathrm{d}s &= -10^4 \frac{2}{\delta} c_1 w \int_T^{T+\delta} (1 + b(E_{b,K} * 1)(T+\delta-s))\, g_T(s)\, \mathrm{d}s \\ &= w \cdot \mathrm{VIX}^2_T, \quad \mathbb{Q} - \text{a.s.}\end{aligned} \tag{5.28}$$

We introduce the process (cfr. (4.31))

$$\bar{g}_T(s) = \begin{cases} \sigma_s^2, & s \in [0,T], \\ g_T(s), & s > T. \end{cases}$$

Note that by (5.1) and (5.3), $\bar{g}_T(\cdot)$ is a joint measurable modification of $g_0 + \int_0^T 1_{\{r\le\cdot\}} K(\cdot - r)\mathrm{d}Z_r$. For every $t \in [0,T]$, the stochastic Fubini's theorem, (5.23), (5.26), and suitable changes of variables, yield

$$\begin{aligned}\int_0^\infty \Phi_w(T,s)\left(\bar{g}_t(s) - g_0(s)\right)\mathrm{d}s &= \int_0^\infty \Phi_w(T,s)\left(\int_0^t 1_{\{u\le s\}} K(s-u)\,\mathrm{d}Z_u\right)\mathrm{d}s \\ &= \int_0^t \left(\int_0^\infty h_w(s)\,K(s+T-u)\,\mathrm{d}s + \int_0^{T-u} K(s)\,G\left(\phi_w(T-u-s)\right)\mathrm{d}s\right)\mathrm{d}Z_u \\ &= \int_0^t \phi_w(T-u)\,\mathrm{d}Z_u, \quad \mathbb{Q}-\text{a.s.}\end{aligned} \tag{5.29}$$

Moreover, by (5.26), the following equality holds:

$$\begin{aligned}\int_0^t \Phi_w(T,s)\,\sigma_s^2\,\mathrm{d}s &= \int_0^t h_w(s-T)\,1_{\{s\ge T\}}\sigma_s^2\,\mathrm{d}s + \int_0^t G\left(\phi_w(T-s)\right)\sigma_s^2\,\mathrm{d}s \\ &= \int_0^t G\left(\phi_w(T-s)\right)\sigma_s^2\,\mathrm{d}s.\end{aligned} \tag{5.30}$$

Recalling the definition of $\widetilde{U}_t(w,T)$, we combine (5.29) and (5.30) to write

$$\begin{aligned}\widetilde{U}_t(w,T) &= \int_t^\infty \Phi_w(T,s)\,g_0(s)\,\mathrm{d}s + \int_0^\infty \Phi_w(T,s)\left(\bar{g}_t(s) - g_0(s)\right)\mathrm{d}s - \int_0^t \Phi_w(T,s)\left(\sigma_s^2 - g_0(s)\right)\mathrm{d}s \\ &= \int_0^\infty \Phi_w(T,s)\,g_0(s)\,\mathrm{d}s + \int_0^t \phi_w(T-u)\,\mathrm{d}Z_u - \int_0^t G\left(\phi_w(T-s)\right)\sigma_s^2\,\mathrm{d}s, \quad \mathbb{Q}-\text{a.s.}\end{aligned} \tag{5.31}$$

In the sequel we denote by $U(w,T) = (U_t(w,T))_{t\in[0,T]}$ the càdlàg process defined by the rightmost side of (5.31). An application of Itô's formula together with (5.24) shows that $E(w,T) = (\exp\{U_t(w,T)\})_{t\in[0,T]}$ is a local martingale, namely $E(w,T) = \exp\{\int_0^\infty \Phi_w(T-s)g_0(s)\mathrm{d}s\}\mathcal{E}(\widetilde{N}(w,T))$, where $\mathcal{E}$ denotes the Doléans-Dade exponential and $\widetilde{N}(w,T) = (\widetilde{N}_t(w,T))_{t\in[0,T]}$ is defined by

$$\mathrm{d}\widetilde{N}_t(w,T) = \sqrt{c}\,\phi_w(T-t)\,\sigma_t\,\mathrm{d}W_{2,t} + \int_{\mathbb{R}_+}\left(e^{\phi_w(T-t)z} - 1\right)\tilde{\mu}(\mathrm{d}t,\mathrm{d}z), \quad \widetilde{N}_0(w,T) = 0.$$

If $E(w,T)$ is a true martingale, then (5.27) follows from (5.28) and (5.31). As in the proof of Theorem 5.3, we search for an expression of $U(w,T)$ which is affine on the past trajectory of σ^2. However, we cannot directly invoke [37, Theorem 10], see also Theorem 4.10 in Chapter 4, due to the different structure of the Riccati-Volterra equation in (5.23) and of the process $U(w,T)$ itself. Fortunately, we can adapt the procedure in the proof of [37, Theorem 10] or Theorem 4.10. Specifically, thanks to the local boundedness of $\Phi_w(T,\cdot)$ (see (5.26)), $\mathbb{Q}-$a.s.,

$$U_t(w,T) = \int_t^{T+\delta} \Phi_w(T,s)\,g_0(s)\,\mathrm{d}s + \phi_w(T-t)\,Z_t + \left(\pi_{T+\delta-t} * \left(\sigma^2 - g_0\right)\right)(t), \quad \text{for a.e. } t \in (0,T).$$

Here the functions

$$\pi_{T+\delta-t}(u) = \int_0^{T+\delta-t} \Phi_w(T, T+\delta-s)((\Delta_{T+\delta-t-s}K)' * L)(u)\mathrm{d}s, \quad t \in (0,T),$$

are well defined for almost every $u \in \mathbb{R}_+$ and belong to $L^1_{\mathrm{loc}}(\mathbb{R}_+)$. At this point, for every $t \in (0,T)$ we introduce the locally absolutely continuous function

$$\begin{aligned}\widetilde{\Pi}_{T+\delta-t}(u) &= \int_0^u \pi_{T+\delta-t}(s)\,\mathrm{d}s + \phi_w(T-t)\,L([0,u]) \\ &= \int_0^{T+\delta-t} \Phi_w(T, T+\delta-s)\left((\Delta_{T+\delta-t-s}K) * L\right)(u)\,\mathrm{d}s, \quad u \geq 0,\end{aligned}$$

where the second equality is due to (5.23) and a suitable change of variables. Therefore, also recalling (5.1), the previous formula for $U(w,T)$ can be rewritten as, $\mathbb{Q}$−a.s., for a.e. $t \in (0,T)$,

$$U_t(w,T) = \int_t^{T+\delta} \Phi_w(T,s)\,g_0(s)\,\mathrm{d}s + \left(\mathrm{d}\widetilde{\Pi}_{T+\delta-t} * \left(\sigma^2 - g_0\right)\right)(t) + \phi_w(T-t)\,L(\{0\})\left(\sigma^2 - g_0\right)(t),$$

which is an affine expression in terms of the past trajectories of σ^2. Now by Lemma 5.2 the real-valued process $E(\mathrm{Re}\,w,T) = (\exp\{U_t(\mathrm{Re}\,w,T)\})_{t\in[0,T]}$ is a true martingale. Thus, thanks to (5.25), we can parallel the comparison argument in the proof of [37, Theorem 14], see also Theorem 4.14 in Chapter 4 to deduce that

$$|\exp\{U_t(w,T)\}| = \exp\{\mathrm{Re}\,U_t(w,T)\} \leq C\exp\{U_t(\mathrm{Re}\,w,T)\}, \quad t \in [0,T], \mathbb{Q}-\text{a.s.},$$

for some constant $C > 0$. An application of [112, Lemma 1.4] completes the proof. ■

5.3.1 VIX put options and futures prices

Theorem 5.10 provides a semi-explicit formula for the Fourier-Laplace transform λ_T of VIX_T^2 in $\mathbb{C}_-$, namely

$$\begin{aligned}\lambda_T(w) = \mathbb{E}\left[\exp\left\{w\cdot \mathrm{VIX}_T^2\right\}\right] &= \exp\left\{\int_0^\infty \Phi_w(T,s)\,g_0(s)\,\mathrm{d}s\right\} \\ &= \exp\left\{\int_0^\delta h_w(s)\,g_0(s+T)\,\mathrm{d}s + \left(g_0 * G(\phi_w(\cdot))\right)(T)\right\}, \quad w \in \mathbb{C}_-. \quad (5.32)\end{aligned}$$

This allows us to price put options written on VIX with the Fourier-inversion technique for the bilateral Laplace transform shown in [47]. More specifically, for a log strike $k \in \mathbb{R}$, the payoff function of such options defined on the whole real line is $w(x) = (e^k - \sqrt{x^+})^+$, $x \in \mathbb{R}$, where x^+ represents VIX^2. Then, denoting by $P(k,T)$ the price of a put option with maturity T (and log strike k) we have (cfr. [47, Equations (7.6)-(7.8)])

$$\begin{aligned}P(k,T) = \mathbb{E}\left[\left(e^k - \mathrm{VIX}_T\right)^+\right] &= -\frac{1}{4\sqrt{\pi}i}\int_{z_r-i\infty}^{z_r+i\infty} \frac{\mathrm{erf}\left(e^k\sqrt{z}\right)}{z^{3/2}}\lambda_T(z)\,\mathrm{d}z \\ &= -\frac{1}{4\sqrt{\pi}}\int_{\mathbb{R}} \mathrm{Re}\left[\frac{\mathrm{erf}\left(e^k\sqrt{z_r+iu}\right)}{(z_r+iu)^{3/2}}\lambda_T(z_r+iu)\right]\mathrm{d}u \\ &= -\frac{1}{2\sqrt{\pi}}\int_{\mathbb{R}_+} \mathrm{Re}\left[\frac{\mathrm{erf}\left(e^k\sqrt{z_r+iu}\right)}{(z_r+iu)^{3/2}}\lambda_T(z_r+iu)\right]\mathrm{d}u, \quad z_r < 0. \quad (5.33)\end{aligned}$$

Here erf represents the error function $\operatorname{erf} z = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} \mathrm{d}t$, $z \in \mathbb{C}$, and for $z \in \mathbb{C}$ and $a \geq 0$, we consider the power $z^a = \Lambda^a e^{ia\theta}$, where $z = \Lambda e^{i\theta}$ with $\Lambda \geq 0$, $\theta \in (-\pi, \pi]$. In particular, we write $\sqrt{z} = z^{1/2}$. The last equality in (5.33) is due to the fact that the integrand is even. Indeed, this follows from the well-known symmetry relation $\operatorname{erf} \bar{z} = \overline{\operatorname{erf} z}$, $z \in \mathbb{C}$, as well as the identities (for $u \neq 0$)

$$\operatorname{Re}\left(\sqrt{z_r + iu}\right) = \sqrt{\frac{z_r + \sqrt{z_r^2 + u^2}}{2}}, \qquad \operatorname{Im}\left(\sqrt{z_r + iu}\right) = \operatorname{sgn}(u) \sqrt{\frac{-z_r + \sqrt{z_r^2 + u^2}}{2}}.$$

Moreover, we can use λ_T to determine $\mathbb{E}[\mathrm{VIX}_T]$, i.e., the futures price of VIX at time T. In order to do this, notice that for every $x \geq 0$ the function $(\sqrt{\pi s})^{-1}(e^{-xs} - 1) + \sqrt{x}\operatorname{erf}(\sqrt{sx})$, $s > 0$, is an antiderivative of $(2\sqrt{\pi})^{-1}(1 - e^{-xs})s^{-3/2}$, $s > 0$. From this relation we deduce the following integral representation for the square-root function

$$\sqrt{x^+} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-sx^+}}{s^{\frac{3}{2}}} \mathrm{d}s, \quad x \in \mathbb{R}.$$

An application of Tonelli's theorem yields

$$\mathbb{E}\left[\mathrm{VIX}_T\right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \lambda_T(-s)}{s^{\frac{3}{2}}} \mathrm{d}s. \tag{5.34}$$

Remark 5.4. *Taking $w = it$, $t \in \mathbb{R}$, in* (5.14) *and* (5.32)*, we obtain semi-explicit expressions for the characteristic functions of X_T and* VIX_T^2*. As a consequence, option pricing on the underlying and the corresponding volatility index can be performed via standard characteristic function inversion algorithms different from the formulae in* (5.15), (5.16) *and* (5.33)*. An analogous argument applies to the* VIX *futures prices in* (5.34)*. Among such Fourier-based techniques, the* COS *method (see [80]) could provide a computationally efficient alternative. However, utilizing the* COS *method would require an analysis of the truncation domain using the cumulants of X_T and* VIX_T^2*. Considering that an optimal implementation of the model* (5.1)-(5.5) *is not the main objective of the present chapter (see also Remark 5.5 and Section 5.5), we postpone the improvement of efficiency, possibly using the* COS *method itself, to future research specifically focusing on the numerical aspects of the rough Hawkes Heston model.*

5.4 Numerical approximation of the model

According to the formulae in (5.15)-(5.16) and (5.33), in order to price options on S and VIX with maturity T, one needs to compute $\Psi^{X_T}(w_1)$ and $\lambda_T(w_2)$, where w_1, w_2 belong to appropriate regions of $\mathbb{C}$. In addition, the values $\lambda_T(-s)$, $s \geq 0$, are also necessary to determine the futures price of VIX at time T. Consequently, looking at the expressions of these Fourier-Laplace transforms in (5.14) and (5.32), the solutions of the Riccati-Volterra equations (5.7) and (5.23), i.e., ψ_{w_1}, ϕ_{w_2} and ϕ_{-s}, have to be approximated on the interval $[0, T]$. Among the available numerical methods to approximate them we choose the multi-factor scheme suggested in [5]. Another possibility would be to use the Adams scheme [71, 72], hybrid schemes as in [45], or an adaptation of the multi-factor hybrid approach in [160].

The multi-factor scheme consists in approximating the kernel K with a weighted sum of exponentials, namely with functions K_n, $n \in \mathbb{N}$, of the form

$$K_n(t) = \sum_{j=1}^n m_{j,n} e^{-x_{j,n} t}, \quad t \geq 0, \tag{5.35}$$

where $m_{j,n}, x_{j,n} > 0$, $j = 1, \dots, n$. In what follows, we write $\mathbf{m} = \{m_{j,n} | \, j = 1, \dots, n, \, n \in \mathbb{N}\}$ and $\mathbf{x} = \{x_{j,n} | \, j = 1, \dots, n, \, n \in \mathbb{N}\}$. Notice that K_n, $n \in \mathbb{N}$, is completely monotone on $(0, \infty)$, meaning that it is nonnegative and infinitely differentiable on this interval, with nonpositive [resp., nonnegative] odd [resp., even] $k-$derivative, $k \in \mathbb{N}$. More details about this approximation and the idea behind it can be found in Remark 5.5 below and in the references therein.

Given $n \in \mathbb{N}$ and $w \in \mathbb{C}$ such that $\operatorname{Re} w \in [0,1]$, we now introduce the Riccati-Volterra equation

$$\psi_{w,n}(t) = \int_0^t K_n(t-s)\, \mathcal{R}(w, \psi_{w,n}(s))\, \mathrm{d}s = (K_n * \mathcal{R}(w, \psi_{w,n}(\cdot)))(t), \quad t \geq 0. \tag{5.36}$$

Note that the existence and uniqueness of $\psi_{w,n}$ is guaranteed by Theorem 5.1 (i), because K_n satisfies Hypothesis 5.1. The advantage in considering (5.36) instead of (5.7) is that its solution $\psi_{w,n}$ can be obtained by numerically solving a system of integral equations with standard methods. More precisely, $\psi_{w,n}(t) = \sum_{j=1}^n m_{j,n} \psi^{(j)}_{w,n}(t)$ for every $t \geq 0$, where

$$\psi^{(j)}_{w,n}(t) = e^{-x_{j,n} t} \int_0^t e^{x_{j,n} s} \mathcal{R}\left(w, \sum_{k=1}^n m_{k,n} \psi^{(k)}_{w,n}(s)\right) \mathrm{d}s, \quad j = 1, \dots, n.$$

Analogously, for every $n \in \mathbb{N}$ and $w \in \mathbb{C}_-$, we consider the Riccati-Volterra equation

$$\phi_{w,n}(t) = \int_0^\infty h_w(s)\, K_n(s+t)\, \mathrm{d}s + (K_n * (G(\phi_{w,n}(\cdot))))(t), \quad t \geq 0. \tag{5.37}$$

We have that $\phi_{w,n}(t) = \sum_{j=1}^n m_{j,n} \phi^{(j)}_{w,n}(t)$, $t \geq 0$, with

$$\phi^{(j)}_{w,n}(t) = e^{-x_{j,n} t} \left(\int_0^\infty h_w(s)\, e^{-x_{j,n} s} \mathrm{d}s + \int_0^t e^{x_{j,n} s} G\left(\sum_{k=1}^n m_{k,n} \phi^{(k)}_{w,n}(s) \right) \mathrm{d}s \right), \quad j = 1, \dots, n.$$

The following theorem offers an estimate on the uniform distance on $[0,T]$ between ψ_w and $\psi_{w,n}$, as well as between ϕ_w and $\phi_{w,n}$. In the former case, it generalizes [5, Theorem 4.1] to our framework with jumps. Its proof, which we postpone to Appendix 5.C, relies on results related to Riccati-Volterra equations which are proved in Appendix 5.A.

Theorem 5.11. *Assume that K satisfies Hypothesis 5.1. Let $T > 0$ and denote by $E_{\lambda,n}$ the canonical resolvent of K_n with parameter $\lambda \in \mathbb{R}$, $n \in \mathbb{N}$.*

(i) *Suppose that $\int_0^T |E_{b+\rho^+\sqrt{c},n}(s)| \mathrm{d}s \leq \widetilde{C}$ for every $n \in \mathbb{N}$, where $\widetilde{C} = \widetilde{C}(\rho, b, \mathbf{m}, \mathbf{x}, T) > 0$. Then there exists a constant $C = C(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ such that, for every $w \in \mathbb{C}$ with $\operatorname{Re} w \in [0,1]$ and $n \in \mathbb{N}$,*

$$\sup_{t \in [0,T]} |\psi_w(t) - \psi_{w,n}(t)| \leq C \left(1 + |\operatorname{Im} w|^6\right) \int_0^T E_{C\left(1+|\operatorname{Im} w|^2\right),K}(s)\, \mathrm{d}s \times \int_0^T |K_n(s) - K(s)|\, \mathrm{d}s. \tag{5.38}$$

In addition, if $b < 0$ and $\rho < 0$ then the constant C does not depend on $\mathbf{m}$ or $\mathbf{x}$, and the dependence on T is via $\|K\|_{L^1([0,T])}$.

(ii) *Suppose that* $\int_0^{T\vee\delta} |E_{b^+,n}(s)| \mathrm{d}s \le \widetilde{C}$ *for every* $n \in \mathbb{N}$, *where* $\widetilde{C} = \widetilde{C}(b, \mathbf{m}, \mathbf{x}, T, \delta) > 0$. *Then there exists a constant* $C = C(b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T, \delta) > 0$ *such that, for every* $w \in \mathbb{C}_-$ *and* $n \in \mathbb{N}$,

$$\sup_{t\in[0,T]} |\phi_w(t) - \phi_{w,n}(t)| \le C\left(1 + |w|^6\right) \int_0^T E_{C(1+|w|^2),K}(s)\,\mathrm{d}s \int_0^{T\vee\delta} |K_n(s) - K(s)|\,\mathrm{d}s. \quad (5.39)$$

Remark 5.5. *When the kernel* K *is completely monotone, a standard way to determine* $\mathbf{m}$ *and* $\mathbf{x}$ *in* (5.35) *relies on the Bernstein-Widder theorem (see, e.g., [96, Theorem* 2.5*, Chapter* 5*]), according to which there exists a nonnegative measure* μ *on* $\mathbb{R}_+$ *such that* $K(t) = \int_{\mathbb{R}_+} e^{-xt}\mu(\mathrm{d}x)$, $t > 0$. *Approximating* μ *with a weighted sum of Dirac measures gives* K_n. *More specifically, for a fixed* $n \in \mathbb{N}$, *it is customary to take a strictly increasing sequence of nonnegative numbers* $(\rho_{j,n})_{j=0,\ldots,n}$, *and then choose, for every* $j = 1, \ldots, n$,

$$m_{j,n} = \int_{\rho_{j-1,n}}^{\rho_{j,n}} \mu(\mathrm{d}y), \qquad x_{j,n} = m_{j,n}^{-1} \int_{\rho_{j-1,n}}^{\rho_{j,n}} y\,\mu(\mathrm{d}y). \quad (5.40)$$

We mention that in some instances (most notably when K *is the fractional kernel, see for instance [1, Lemma A.3] and [5, Proposition 3.3]) it is possible to show the convergence* $K_n \to K$ *in* $L^2_{\mathrm{loc}}(\mathbb{R}_+)$. *Thanks to [96, Theorem* 3.1*, Chapter* 2*], this ensures the validity of the hypotheses required in both points of Theorem 5.11, and therefore the convergence of the multi-factor scheme.*

We remark that (5.40) *is not an optimal choice for the exponential approximation* (5.35) *when it comes to efficiency. In this regard, better solutions can be found in [160, Section 5], where an extensive numerical analysis with comparisons among different approximation methods is presented for the rough fractional kernel. It is also important to mention the recent work in [20], which provides good low-factor approximations using higher-order quadrature rules for the choice of the nodes and weights. For the numerical application we use the nodes and weights in* (5.40) *because, as explained in the following section, this is sufficient for our calibration purposes.*

5.5 Calibration

We have shown that Fourier-based methods can be applied to the rough Hawkes Heston model in order to price options on the underlying and the corresponding volatility index. Based on these techniques, in this section we calibrate a parsimonious specification of the rough Hawkes Heston model to S&P 500 and VIX options data on May 19, 2017. This is the same data set as in [91]. In Table 5.3 [resp., Table 5.4] we report the strikes and maturities of the SPX options [resp., VIX options] considered for the calibration. Our objective is to minimize the relative RMSE (root-mean-square error) between market and theoretical implied volatilities of both SPX and VIX options. More precisely, denoting by $\mathbf{\Theta}$ the parameters of the model, the goal of the calibration procedure is to determine

$$\arg\min_{\mathbf{\Theta}} \sqrt{\sum_{i,j}\left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i,K_j) - \sigma^{\mathbf{\Theta}}_{\mathrm{SPX}}(T_i,K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i,K_j)}\right)^2 + \sum_{i,j}\left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i,K_j) - \sigma^{\mathbf{\Theta}}_{\mathrm{VIX}}(T_i,K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i,K_j)}\right)^2}. \quad (5.41)$$

The implied volatilities of VIX options are computed with respect to market futures. As we will explain below, after calibration, model futures approximate well market futures (see Figure 5.7).

As it is customary in rough volatility models, for our parametrization we choose a power kernel of the form $K(t) = t^{\alpha-1}/\Gamma(\alpha)$, $\alpha \in (1/2, 1]$. We then approximate it with the sum of exponentials in

(5.35), considering $n = 200$. The weights $m_{j,200}$ and mean reversion terms $x_{j,200}$, $j = 1, \ldots, 200$, are computed according to (5.40), where $\rho_{j,200} = r^{j-100}$, $j = 0, \ldots, 200$, $r > 1$, is the geometric partition. In this setting, the integrals in (5.40) can be computed explicitly, see [1, Equation (3.3)]. In order to select r we follow [1, Remark 3.2], namely

$$\bar{r} = \arg\min_{r>1} \|K - K_{200}\|_{L^2([0,0.091])} .$$

We refer to [49, Equation (5.2)] for an explicit expression of the functional to minimize. Here we take the L^2-norm in $[0, 0.091]$ because $T = 0.091$ is the longest maturity that we consider for the calibration. However, the sensitivity of the model with respect to the upper bound of the interval is negligible. In the sequel, we denote the multi-factor kernel corresponding to $\bar{r}$ by $\bar{K}_{200}$.
By analogy with the rough Heston model introduced and studied in [78, 79], we consider an initial input curve g_0 of the form

$$g_0(t) = \sigma_0^2 + \beta \int_0^t K(s)\, \mathrm{d}s = \sigma_0^2 + \frac{\beta}{\Gamma(\alpha+1)} t^{\alpha}, \quad t \geq 0, \tag{5.42}$$

where σ_0^2, $\beta \geq 0$. Note that the structure of g_0 in (5.42) is quite restrictive, but it allows to keep the model parsimonious with a small number of parameters. Indeed, the initial variance curve in (5.42) is specified only by σ_0^2 and β. Our choice is also justified by the fact that we focus on short time-to-maturity options, so the lack of flexibility for g_0 is not a drawback in our application. More general forms of g_0 or expressions extracted from the replication formula for the log-contract as in [7, Equation (5.1)] can be used to calibrate VIX smiles for longer times-to-maturity. In our numerical illustration, we consider the kernel $\bar{K}_{200}$ and g_0 as in (5.42). These choices guarantee the well-posedness, in the weak sense, of (5.1), because $\bar{K}_{200}$ is completely monotone and g_0 satisfies Hypothesis 5.2 (see Remark 5.1). For the law of the jumps, to keep the number of parameters low we choose an exponential distribution with rate 1, $\nu(\mathrm{d}z) = \exp(-z)\,\mathrm{d}z$. Our parsimonious specification of the model has therefore – other than the two parameters (β, σ_0^2) related to g_0 – five evolution-related parameters $(\alpha, \rho, b, c, \Lambda)$. Like in [91], we concentrate on short maturities for which, as pointed out in [99], "VIX derivatives are most liquid and the joint calibration is most difficult." The resulting calibrated parameters are reported in Table 5.1.

α	ρ	b	c	Λ	β	σ_0^2
0.506	-0.737	-2.008	0.156	0.242	0.048	0.0074

Table 5.1: Calibrated parameters for the rough Hawkes Heston model.

Starting from the values in Table 5.1, we then minimize the following functional of $\mathbf{\Theta}$, which takes into account the relative number of SPX/VIX options in the sample considered for the calibration:

$$c_1 \sqrt{\sum_{i,j} \left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i, K_j) - \sigma^{\mathbf{\Theta}}_{\mathrm{SPX}}(T_i, K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}(T_i, K_j)} \right)^2} + c_2 \sqrt{\sum_{i,j} \left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i, K_j) - \sigma^{\mathbf{\Theta}}_{\mathrm{VIX}}(T_i, K_j)}{\sigma^{\mathrm{mkt}}_{\mathrm{VIX}}(T_i, K_j)} \right)^2},$$
$$\text{where } c_1 = 71.4\%, \quad c_2 = 28.6\%.$$

However, no significant changes in the parameters have to be reported.

We observe that the value of α in Table 5.1 is very close to its lower bound limit 0.5. This is coherent with previous estimates in the rough volatility literature, see for instance [12, 21, 24, 76, 87, 90, 91].

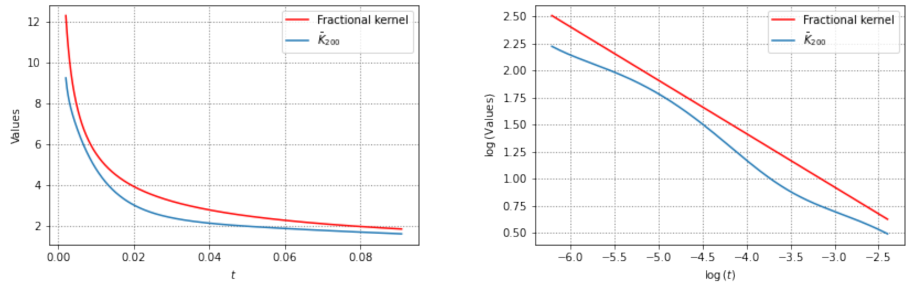

Figure 5.1: On the left, the approximation of the fractional kernel with the multi-factor kernel $\bar{K}_{200}$ in the domain $[\frac{1}{500}, 0.091]$. Here $\alpha = 0.506$. On the right, the corresponding log-log plot.

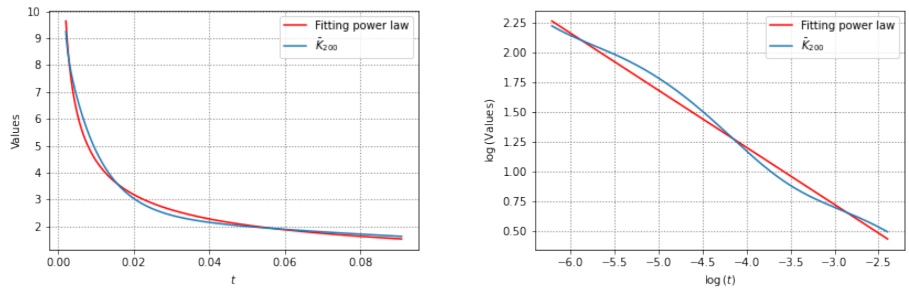

Figure 5.2: On the left, the power-law kernel $\widetilde{K}$ (see (5.43)) fitting $\bar{K}_{200}$, with $\alpha = 0.506$, in the domain $[\frac{1}{500}, 0.091]$. On the right, the corresponding log-log plot.

On the other hand, such a small value of α causes the approximation of the fractional kernel K via the sum of exponentials $\bar{K}_{200}$ to deteriorate, despite the high number of factors considered ($n = 200$). This is shown in Figure 5.1. Nonetheless, $\bar{K}_{200}$ still approximates well the singular kernel

$$\widetilde{K}(t) = 0.485\, t^{0.519-1}, \quad t > 0, \tag{5.43}$$

as shown in Figure 5.2 and is able to reproduce a power law for the ATM implied volatility skew as illustrated in Figure 5.12. In this regard, more efficient exponential representations such as those proposed in [20, 160] could give a better approximation of the fractional kernel, while reducing the number of factors n in (5.35), with a significant improvement in computational cost. This kind of numerical analysis is, however, beyond the scope of this chapter. Furthermore, we remark that the recent study [6], comparing different types of kernels (fractional, log-modulated, shifted and exponential) in a Gaussian Volterra driven setup, concludes that a conventional one factor, Markovian exponential kernel is able to jointly fit the SPX/VIX smiles outperforming its rough and non-rough path-dependent counterparts. Since the theoretical study of the previous sections covers a wide range of kernels (e.g., completely monotone), an interesting question is whether a similar behavior is exhibited in our framework. This important question could be addressed in futures studies of efficient approximations and variants of the rough Hawkes Heston model.

The estimation of the correlation parameter ρ is also in line with empirical estimates, e.g. [52], and what is commonly known as the *leverage effect* [60, 76, 133]. We notice that for the joint calibration we can keep the vol-of-vol parameter c small because an important part of the volatility fluctuation is captured by the *self-exciting* jumps controlled by the parameters α and Λ. This responds to the issue, raised in [99], that the "very large negative skew of short-term SPX options, which in continuous models implies a very large volatility of volatility, seems inconsistent with the comparatively low levels of VIX implied volatilities."

The calibrated implied volatility smiles for S&P 500 and VIX options are plotted in Figures 5.4 and 5.6, respectively. We zoom the calibration of the S&P 500 options at-the-money in Figure 5.5. Figure 5.7 focuses on the VIX term structure, which we do not include in the functional (5.41) used for the calibration. We remark that the term structure of the model is not flexible enough to perfectly reproduce the shape of the market VIX futures, mainly due to a change of convexity. However, the maximal relative distance between market and model data is $\sim 0.5\%$, which is decent considering that we have implicitly assumed a "flat" initial volatility curve, see (5.42). Furthermore, Figure 5.8 shows the implied volatility smile for S&P 500 options with expiration $T = 0.179$ (July 21st, 2017): one month after the last, calibrated maturity $T = 0.091$. Despite the fact that $T = 0.179$ is not included in (5.41), the model is able to replicate market implied volatility smiles with an accuracy comparable to the other maturities considered for S&P 500 options (cfr. Figure 5.4). Overall, these graphs show that the model fits remarkably well both S&P 500 and VIX implied volatilities.

The shape of the smile at-the-money for S&P 500 options is well-captured and the distance to the bid-ask corridor – across the maturities – is at most of one bid-ask spread. For the two shortest maturities, most of the model implied volatilities at-the-money are actually inside the bid-ask corridor. The fit is not perfect for very negative log-moneyness. This is also seen – possibly to a less extent – in the quadratic rough Heston model [91]. We conjecture that, at the cost of increasing the complexity of the model, even better results could be obtained if we replace the exponential law for the jumps by a law with Pareto tails as suggested in [52, 114] and the references therein. Regarding the VIX implied volatilities, we observe that – even for options deep out-of-the-money – the model implied volatilities stay almost systematically within the bid-ask corridor, whether it is calculated using call or for put options.

5.5.1 Calibration with the rough Heston model

The model (5.1)-(5.5) proposed in this chapter is an extension of the rough Heston model (see [78, 79]), obtained by adding a jump component with intensity proportional to the spot variance σ^2. In particular, the rough Heston model can be recovered from (5.1)-(5.5) by setting $\nu(\mathrm{d}z) = 0$. In Introduction and Section 5.1, we justify the presence of self-exciting jumps common to the underlying and the volatility based on empirical evidence of jump-clustering phenomena and endogeneity of financial markets. The purpose of this part (see also Subsection 5.6.1) is to investigate the impact of jumps with numerical experiments. To do this, we perform a calibration exercise with the rough Heston model using the same dataset as before (cfr. Section 5.5) and considering the same functional (5.41) to minimize. The calibrated parameters are reported in Table 5.2.

α	ρ	b	c	β	σ_0^2
0.516	-0.781	-1.988	0.166	0.050	0.0096

Table 5.2: Calibrated parameters for the rough Heston model.

The resulting implied volatility smiles of SPX/VIX options for two selected maturities ($T = 0.032$, $T = 0.091$) are shown in Figure 5.9. Here we see that the rough Heston model is able to reproduce quite well the level and shape of VIX options smiles, although we might argue that, for $T = 0.091$, it produces a curve which is too concave at-the-money and too flat out-of-the-money. However, the rough Heston model struggles to capture the out-of-the-money skew of S&P 500 options, especially for $T = 0.032$.

In an attempt to understand whether this effect is due to the VIX options in (5.41), we run another calibration exercise with the aim of minimizing the following functional in $\mathbf{\Theta}$:

$$\sqrt{\sum_{i,j}\left(\frac{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}\left(T_i,K_j\right)-\sigma^{\mathbf{\Theta}}_{\mathrm{SPX}}\left(T_i,K_j\right)}{\sigma^{\mathrm{mkt}}_{\mathrm{SPX}}\left(T_i,K_j\right)}\right)^2}. \tag{5.44}$$

The strikes and maturities used in this example are listed in Tables 5.3-5.5. We perform the same optimization for the rough Hawkes Heston model and display the outcomes in Figure 5.10. These results confirm that the rough Heston model does not produce a correct shape for the left-tails of SPX implied volatility smiles, and that in this aspect it is consistently outperformed by the extension (5.1)-(5.5) proposed in this chapter.

On the basis of these experiments, we conclude that the introduction of a jump component in the rough Heston model is significant and does not cause redundancy. In particular, the jumps allow to better reproduce the skew of out-of-the-money S&P 500 options, especially for short maturities.

5.6 Sensitivities of the implied volatilities

In this section we study the sensitivity of the implied volatilities of S&P 500 and VIX options to the parameters of the rough Hawkes Heston model. Starting from the calibrated parameters presented in Table 5.1, we analyze the impact of a change in the evolution-related parameters $(\alpha, \rho, b, c, \Lambda)$ and the initial curve parameters (β, σ_0^2) on the implied volatilities for the shortest maturity, and for the shortest and longest maturities, respectively.

We begin with the sensitivity with respect to the parameter $\alpha \in (0.5, 1]$, which as we will see plays a crucial role in our model. We can observe in Figure 5.11 – as is the case for other rough volatility models – that modifications of the parameter α change the ATM skew of the implied volatility of S&P 500 options. A good convexity and ATM skew, for the maturities considered in the calibration, can be obtained with very low values of the parameter α, confirming the findings in the rough volatility literature. To elucidate the influence of the parameter α on the ATM skews, we plot in Figure 5.12 the log-log plots of ATM skews as a function of maturity, for the calibrated parameters and different values of α. We observe that a perfect power decay, for the given maturities, is captured by $\alpha = 0.506$, but not by higher values of α. For $\alpha = 0.506$, the linear fit is almost perfect with a -0.597 power decay and an unquestionable coefficient of determination $R^2 = 0.99905$. It is important to mention at this point the recent works [67, 100] which point out that the linear fit is no longer optimal when considering a larger range of maturities. Our findings for the maturities considered in the calibration are coherent with the results in the rough volatility literature, e.g. [21, 90], indicating a power law, as T goes to zero, for the ATM skew as a function of maturity given approximately by $T^{-\frac{1}{2}}$. For other values of α, the linear fit is also observed for the shortest maturities. In Figure 5.12, we plot the estimated power decay for the short maturities as a function of α. This plot shows that the relationship between the power decay and α is approximately linear for the short maturities.

More importantly, within the joint calibration framework, the parameter α has a big impact on the level and shape of implied volatilities of VIX options. This is confirmed by Figure 5.11. In particular, the difference in level between the implied volatilities of VIX options for $\alpha = 0.506$ and $\alpha = 0.6$ is similar to the one between $\alpha = 0.6$ and $\alpha = 0.9$. As α decreases the implied volatilities shift downwards. This feature is fundamental to bring down the VIX implied volatilities maintaining the correct skew for SPX implied volatilities, explaining therefore the shift mentioned in [98, 99]. We ratify therefore –

within the affine framework – the relevance of rough non-Markovian volatility to jointly calibrate SPX and VIX smiles.

We now analyze the dependency of the implied volatilities with respect to the other parameters. Figure 5.13 shows the sensitivities with respect to the evolution-related parameters (b, c, ρ, Λ). We notice that – unless we zoom at-the-money – the sensitivity of the SPX smiles with respect to (b, c, Λ) is relatively small. The main effect of an increment in the reverting speed $-b$ is a shift slightly downwards of the SPX implied volatility and a more pronounced upward shift and a reduction of the concavity on the VIX implied volatility. The impact of the volatility of volatility c is similar for SPX options, with a slight change of concavity, and a more pronounced and less symmetric effect on the level and concavity of implied volatility of VIX options. As usual, the correlation parameter ρ plays a big role by moving the minimum value to the left ($\rho < 0$) or to the right ($\rho > 0$). Obviously, the VIX smiles do not depend on the correlation ρ. The effect of the (jump) leverage Λ is relatively small on SPX implied volatilities but fundamental on the VIX implied volatilities. For SPX implied volatilities, the impact of Λ could be reduced to a rotation with the at-the-money value as pivot. The parameter Λ also controls the level of VIX implied volatility out-of-the-money. As Λ increases this level goes down, achieving the correct shift for the calibrated parameter. This effect is similar to the one observed for the vol-of-vol c, but the sensitivity is larger, and it allows us to keep a low value of c for the joint calibration. This explains, the importance in our model of self-exciting jumps in opposite directions for the underlying and volatility.

We now turn to the parameters (β, σ_0^2) of the initial curve $g_0(t) = \sigma_0^2 + \beta \int_0^t K(s)\,\mathrm{d}s,\ t \geq 0$ (see (5.42)). Figure 5.14 shows the SPX and VIX implied volatility sensitivities for the shortest and longest maturity. The impact of both parameters is similar for SPX and VIX options. When σ_0^2 or β increase the SPX implied volatilities move up and to the right, while the VIX implied volatilities move down and the concavity increases.

5.6.1 Comparison with the rough Heston model

We now continue the discussion started in Subsection 5.5.1 regarding the relevance of jumps in the implementation of the rough Hawkes Heston model. Contrary to Subsection 5.5.1, here we do not focus on the calibration to a particular dataset, but we take a more general point of view. More precisely, we are interested in understanding how the jumps affect the implied volatility curves of SPX/VIX options keeping all the other parameters constant. To do this, we take the values in Table 5.1 for the rough Hawkes Heston model and simply remove the jump component by setting $\nu(\mathrm{d}z) = 0$, recovering then the rough Heston model. Figure 5.3 clearly shows that the jumps have a significant impact on implied volatility smiles. In fact, they change the shape of S&P 500 curves and the level of VIX curves. Hence we conclude that the model (5.1)-(5.5) suggested in this chapter is a parsimonious extension of the rough Heston model which does provide a considerably richer framework. Speaking of the joint calibration, the level and shape of SPX/VIX implied volatility curves constitute two main issues to reconcile in order to successfully tackle the problem, see Introduction. Since the self-exciting jumps affect them both, it appears that the rough Hawkes Heston model has an important advantage over its continuous counterpart (rough Heston), which is also coherent with the experiments in Subsection 5.5.1.

5.7 Conclusion

We develop and study a new stochastic volatility model named the rough Hawkes Heston model. It is a tractable affine Volterra model with rough volatility and volatility jumps that cluster and that have the opposite direction but occur at the same time as the jumps of the underlying prices. This

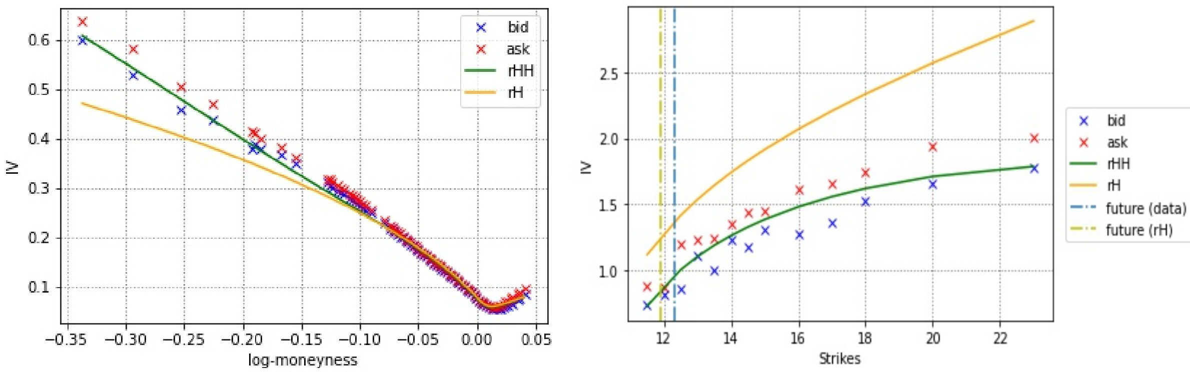

Figure 5.3: Effect of the jump component on the implied volatility of SPX (left) and VIX (right) options. The rough Hawkes Heston (rHH) is in green and the rough Heston (rH) in orange. The blue and red crosses are respectively the bid and ask of market implied volatilities for $T = 0.032$.

model shares many features with other existing models, mainly the Heston [104], Barndorff-Nielsen and Shephard [17], and rough Heston [79] models. It takes advantage of the low regularity and memory features of rough volatility models, the large fluctuation of jumps, the clusters of Hawkes processes and the explicit Fourier-Laplace transform of the affine setup. By combining the modeling advantages of these approaches, it is able to better capture the joint dynamics of underlying prices and their volatility index in a tractable fashion. The addition of a singular kernel in the dynamics of the volatility, together with jumps, incorporates not only the rough volatility feature but also a jump-clustering component. The presence of common jumps in the underlying and the volatility in opposite directions is coherent with previous studies such as [171]. Moreover, the introduction of jumps that cluster – as in [26] – is in accordance with empirical findings, e.g. [52, 53]. Similar to [17, 79, 104], the rough Hawkes Heston model is parsimonious with only five evolution-related parameters, and it belongs to the class of affine Volterra models [8, 37], which allows efficient Fourier-based techniques for pricing.

The parameter α describing the power kernel in the volatility dynamics controls – as in the rough Heston model – the underlying implied volatilities ATM skews for short maturities. Our calibration example indicates that this value is close to 0.5, which agrees with previous estimates in the rough volatility literature [21, 90]. This is not, however, the only role played by the parameter α in our setup, because the power kernel also affects the jump-clustering feature of the model. As a consequence, the parameter α plays a crucial role in controlling the level of VIX implied volatilities. Together with the jump-leverage parameter Λ, the power kernel allows us to bring down the VIX implied volatilities maintaining the correct skew for SPX implied volatilities, consequently capturing the shift mentioned in [98, 99]. This confirms the relevance, in our affine framework, of rough volatility and clustering jumps to model simultaneously the S&P 500 and VIX dynamics.

The affine relation between variance swap rates and forward variance – which generalizes the affine relation between variance swap rates and spot variance in the classical framework [118] – is a by-product of our affine Volterra framework. This affine relation has been confirmed empirically in [132].

To conclude, the rough Hawkes Heston model is able – in a tractable and parsimonious fashion – to jointly calibrate S&P 500 and VIX options. The parsimonious character of our model is an advantage compared to other models that jointly calibrate SPX/VIX options with either a large number of parameters [54, 101] or based on martingale transport considerations [99]. The affine character of the rough Hawkes Heston model allows fast pricing using Fourier-techniques, instead of Monte Carlo or machine learning methods as those used for instance in [91, 161]. Moreover, all the parameters in our model have a financial interpretation, and a complete sensitivity analysis shows that they are not redundant since each of them controls a different feature of the S&P 500 and VIX volatility smiles.

Data Availability Statement

Market data was purchased from the CBOE website `https://datashop.cboe.com/`.

Appendix 5.A Proof of Theorem 5.1

In this appendix we prove Theorem 5.1 regarding the Riccati-Volterra equation (5.6)-(5.7) used to study the Fourier-Laplace transform of the log returns $(X_t)_{t\geq 0}$. We use the following notation: given u, $v \in \mathbb{C}$, let $[u, v]$ be the segment in $\mathbb{C}$ having u and v as endpoints, i.e. $[u, v] = \{z \in \mathbb{C} : z = (1-t)u + tv,\ t \in [0, 1]\}$, and denote by $u \vee v = \operatorname{Re} u \vee \operatorname{Re} v + i \operatorname{Im} u \vee \operatorname{Im} v$.

Proof. Fix $w \in \mathbb{C}$ with $\operatorname{Re} w \in [0, 1]$.

(i) The proof of this point is divided into three steps. In the first step, we show the existence of a noncontinuable solution ψ_w of (5.7). In the second step, we prove that ψ_w does not explode in finite time, i.e., that it is global solution. To conclude, in the third and last step, we prove the uniqueness of ψ_w.

Step I. Let us compute from (5.6), for every $v \in \mathbb{C}_-$,

$$\begin{aligned}\operatorname{Re} \mathcal{R}(w, v) &= \frac{1}{2}\left(|\operatorname{Re} w|^2 - \operatorname{Re} w\right) + \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\operatorname{Re} v \\ &+ \frac{c}{2}|\operatorname{Re} v|^2 - \frac{1}{2}\left(|\operatorname{Im} w|^2 + c\,|\operatorname{Im} v|^2 + 2\rho\sqrt{c}\operatorname{Im} w \operatorname{Im} v\right) \\ &+ \int_{\mathbb{R}_+}\left[e^{(\operatorname{Re} v - \Lambda \operatorname{Re} w)z}\cos\left((\operatorname{Im} v - \Lambda \operatorname{Im} w)\, z\right) - \operatorname{Re} w\left(e^{-\Lambda z} - 1\right) - 1 - \operatorname{Re} v z\right]\nu(\mathrm{d}z). \end{aligned} \tag{5.45}$$

Since $|\rho| \leq 1$ we have $|\rho\sqrt{c}\operatorname{Im} w \operatorname{Im} v| \leq \sqrt{c}|\operatorname{Im} w||\operatorname{Im} v|$, which implies

$$-\frac{1}{2}\left(|\operatorname{Im} w|^2 + c\,|\operatorname{Im} v|^2 + 2\rho\sqrt{c}\operatorname{Im} w \operatorname{Im} v\right) \leq -\frac{1}{2}\left(|\operatorname{Im} w| - \sqrt{c}\,|\operatorname{Im} v|\right)^2 \leq 0. \tag{5.46}$$

Recalling that $\operatorname{Re} w \in [0, 1]$, we then obtain

$$\begin{aligned}\operatorname{Re} \mathcal{R}(w, v) &\leq \left(b + \rho\sqrt{c}\operatorname{Re} w\right)\operatorname{Re} v + \frac{c}{2}|\operatorname{Re} v|^2 + \int_{\mathbb{R}_+}\left[e^{-\Lambda \operatorname{Re} w z} - \operatorname{Re} w\left(e^{-\Lambda z} - 1\right) - 1\right]\nu(\mathrm{d}z) \\ &\qquad + \int_{\mathbb{R}_+}\left[e^{(\operatorname{Re} v - \Lambda \operatorname{Re} w)z} - e^{-\Lambda \operatorname{Re} w z} - \operatorname{Re} v z\right]\nu(\mathrm{d}z) \\ &\leq \left(b + \rho\sqrt{c}\operatorname{Re} w + \int_{\mathbb{R}_+} z\left(e^{-\Lambda \operatorname{Re} w z} - 1\right)\nu(\mathrm{d}z)\right)\operatorname{Re} v + \frac{c}{2}|\operatorname{Re} v|^2 \\ &\qquad + \int_{\mathbb{R}_+} e^{-\Lambda \operatorname{Re} w z}\left(e^{\operatorname{Re} v z} - 1 - \operatorname{Re} v z\right)\nu(\mathrm{d}z), \end{aligned} \tag{5.47}$$

where for the second inequality we use

$$e^{-\Lambda \operatorname{Re} w z} - \operatorname{Re} w\left(e^{-\Lambda z} - 1\right) - 1 \leq 0, \quad z \geq 0. \tag{5.48}$$

Let $h\colon \mathbb{R}_+ \times \mathbb{R}_- \to \mathbb{R}_-$ be the continuous function defined by

$$h(x,y) = \begin{cases} \frac{1}{y}\int_{\mathbb{R}_+} e^{-\Lambda xz}\left(e^{yz}-1-yz\right)\nu(\mathrm{d}z), & y<0 \\ 0, & y=0 \end{cases}, \quad x \geq 0,$$

and note that $y \cdot h(x,y) = \int_{\mathbb{R}_+} e^{-\Lambda xz}(e^{yz}-1-yz)\nu(\mathrm{d}z)$. At this point, we can use (5.47) to show that

$$\mathrm{Re}\ \mathcal{R}(w,v) \leq \left(C_w + \frac{c}{2}\mathrm{Re}\, v + h\left(\mathrm{Re}\, w, \mathrm{Re}\, v\right)\right)\mathrm{Re}\, v, \quad v \in \mathbb{C}_-, \tag{5.49}$$

where $C_w = b + \rho\sqrt{c}\,\mathrm{Re}\, w + \int_{\mathbb{R}_+} z(e^{-\Lambda \mathrm{Re}\, w z}-1)\nu(\mathrm{d}z)$.

We now introduce the function $\widetilde{\mathcal{R}}_w\colon \mathbb{C} \to \mathbb{C}$ given by

$$\widetilde{\mathcal{R}}_w(v) = \mathcal{R}\left(w, -\mathrm{Re}\, v^- + i\mathrm{Im}\, v\right) + C_w \mathrm{Re}\, v^+, \quad v \in \mathbb{C}.$$

Observe that, by construction (see also (5.49))

$$\mathrm{Re}\ \widetilde{\mathcal{R}}_w(v) \leq \left(C_w - \frac{c}{2}\mathrm{Re}\, v^- + h\left(\mathrm{Re}\, w, -\mathrm{Re}\, v^-\right)\right)\mathrm{Re}\, v, \quad v \in \mathbb{C}.$$

Since $\widetilde{\mathcal{R}}_w$ is continuous, we can invoke [96, Therorem 1.1, Chapter 12] to assert the existence of a continuous, noncontinuable solution $\psi_w\colon [0,T_{\max}) \to \mathbb{C}$ of the equation

$$\chi = K * \widetilde{\mathcal{R}}_w\left(\chi(\cdot)\right), \quad t \in [0,T_{\max}), \tag{5.50}$$

for some $T_{\max} \in (0,\infty]$. If we can show that $\mathrm{Re}\,\psi_w \leq 0$ in $[0,T_{\max})$, then we conclude that ψ_w is indeed a noncontinuable solution of (5.7), as well. To this end, consider the continuous function $\zeta(t) = C_w - \frac{c}{2}\mathrm{Re}\,\psi_w(t)^- + h(\mathrm{Re}\, w, -\mathrm{Re}\,\psi_w(t)^-)$ defined for $t \in [0,T_{\max})$. Taking the real part in (5.50), for every $T \in (0,T_{\max})$, we obtain

$$\mathrm{Re}\,\psi_w(t) = -\gamma_T(t) + \int_0^t K(t-s)\,\zeta(s)\,\mathrm{Re}\,\psi_w(s)\,\mathrm{d}s, \quad t \in [0,T],$$

where $\gamma_T(t) = \int_0^t K(t-s)1_{\{s\leq T\}}(\zeta(s)\mathrm{Re}\,\psi_w(s) - \mathrm{Re}\ \widetilde{\mathcal{R}}_w(\psi_w(s)))\mathrm{d}s$. By [5, Remark B.6] $\gamma_T \in \mathcal{G}_K$ (recall (5.22)), and we can invoke [5, Theorem C.1] to infer that $\mathrm{Re}\,\psi_w \leq 0$ in $[0,T]$. Given that T was arbitrary, such an inequality holds in the whole interval $[0,T_{\max})$, completing the first step of the proof.

Step II. Our goal here is to show that $T_{\max} = \infty$. Let us fix again a generic $T \in (0,T_{\max})$. Taking the imaginary part in (5.6) and (5.7) we have, on the interval $[0,T]$,

$$\begin{aligned}\mathrm{Im}\,\psi_w = K * \Bigg[&\left(\mathrm{Re}\, w - \frac{1}{2}\right)\mathrm{Im}\, w + \left(b+\rho\sqrt{c}\,\mathrm{Re}\, w\right)\mathrm{Im}\,\psi_w + \rho\sqrt{c}\,\mathrm{Im}\, w\,\mathrm{Re}\,\psi_w + c\,\mathrm{Re}\,\psi_w\,\mathrm{Im}\,\psi_w \\ &+ \int_{\mathbb{R}_+}\left(e^{\mathrm{Re}(\psi_w-\Lambda w)\cdot z}\sin\left(\mathrm{Im}\left(\psi_w-\Lambda w\right)\cdot z\right) - \mathrm{Im}\, w\left(e^{-\Lambda z}-1\right) - \mathrm{Im}\,\psi_w\cdot z\right)\nu(\mathrm{d}z)\Bigg].\end{aligned} \tag{5.51}$$

Consider the function $d\colon \mathbb{R}_- \times \mathbb{R} \to \mathbb{R}$ defined as follows

$$d(x,y) = \begin{cases} \frac{1}{y}\int_{\mathbb{R}_+} e^{xz}\left(\sin(y\,z) - y\,z\right)\nu(\mathrm{d}z), & y \neq 0 \\ 0, & y=0 \end{cases}, \quad x \leq 0.$$

Note that d is continuous and nonpositive in its domain. Moreover, by construction

$$y \cdot d(x,y) = \int_{\mathbb{R}_+} e^{xz} \left(\sin(y\,z) - y\,z\right) \nu(\mathrm{d}z), \quad (x,y) \in \mathbb{R}_- \times \mathbb{R}.$$

To shorten the notation we define $\widetilde{\psi_w} = \psi_w - \Lambda w$. Using the function d we rewrite (5.51) as

$$\begin{aligned}
&\operatorname{Im}\psi_w + \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w \\
&= \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + K * \Bigg[\left(\operatorname{Re} w - \frac{1}{2} - \int_{\mathbb{R}_+} \left(e^{-\Lambda z} - 1 + \Lambda z\right) \nu(\mathrm{d}z) - \frac{\rho^+}{\sqrt{c}} \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \right) \operatorname{Im} w \\
&+ \left(-\rho^-\sqrt{c}\operatorname{Re}\psi_w - \left(\Lambda + \frac{\rho^+}{\sqrt{c}}\right) \int_{\mathbb{R}_+} z \left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z} - 1\right) \nu(\mathrm{d}z) - \left(\Lambda + \frac{\rho^+}{\sqrt{c}}\right) d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right) \right) \operatorname{Im} w \\
&+ \left(\left(b + \rho\sqrt{c}\operatorname{Re} w\right) + c\operatorname{Re}\psi_w + \int_{\mathbb{R}_+} z \left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z} - 1\right) \nu(\mathrm{d}z) + d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right) \right) \left(\operatorname{Im}\psi_w + \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w \right) \Bigg] \\
&=: \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w \\
&\quad + K * \left[\left(C_1 - \frac{\rho^+}{\sqrt{c}} \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \right) \operatorname{Im} w + f_1(\cdot)\operatorname{Im} w + \left(b + \rho\sqrt{c}\operatorname{Re} w + f_2(\cdot)\right) \left(\operatorname{Im}\psi_w + \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w \right) \right],
\end{aligned}$$

which holds on $[0,T]$. In particular, note that $f_1 \geq 0$ and $f_2 \leq 0$ in $[0,T]$. We want to find a continuous function $u\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $|\operatorname{Im}\psi_w| \leq u$ on $[0,T]$. To do this, we argue by cases on $\operatorname{Im} w$. In the following, we denote $\widetilde{\Lambda} = \max\{\rho^- c^{-1/2}, \Lambda\}$. All the claims regarding the sign of solutions to linear Volterra equations are justified by [5, Theorem C.1].

If $\operatorname{Im} w \geq 0$, then we can consider the unique, nonnegative, continuous solution $l_1\colon [0,T] \to \mathbb{R}_+$ of the linear equation

$$l_1 = \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + K * \left[\left| C_1 - \frac{\rho^+}{\sqrt{c}} \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \right| \operatorname{Im} w + \left(\left(b + \rho\sqrt{c}\operatorname{Re} w\right) + f_2 \right) l_1 \right].$$

Since the function $\operatorname{Im}\psi_w + \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + l_1$ satisfies – in $[0,T]$ – the linear equation

$$\chi = 2\frac{\rho^+}{\sqrt{c}}\operatorname{Im} w + K * \left[2\left(C_1 - \frac{\rho^+}{\sqrt{c}} \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \right)^+ \operatorname{Im} w + f_1 \operatorname{Im} w + \left(\left(b + \rho\sqrt{c}\operatorname{Re} w\right) + f_2 \right) \chi \right],$$

we deduce that $\operatorname{Im}\psi_w \geq -l_1 - \frac{\rho^+}{\sqrt{c}}\operatorname{Im} w$ on $[0,T]$. Next, we introduce the unique, nonnegative, continuous solution $\overline{l_1}\colon \mathbb{R}_+ \to \mathbb{R}_+$ of the linear equation

$$\overline{l_1} = \frac{\rho^+}{\sqrt{c}}\left|\operatorname{Im} w\right| + K * \left[\left| C_1 - \frac{\rho^+}{\sqrt{c}} \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \right| \left|\operatorname{Im} w\right| + \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \overline{l_1} \right] \tag{5.52}$$

and observe that $\overline{l_1} - l_1 \geq 0$ on $[0,T]$, because $\overline{l_1} - l_1$ solves on $[0,T]$

$$\chi = K * \left[-f_2\, l_1 + \left(b + \rho\sqrt{c}\operatorname{Re} w\right) \chi \right].$$

Hence, $\operatorname{Im}\psi_w \geq -\overline{l_1} - \frac{\rho^+}{\sqrt{c}}|\operatorname{Im} w|$ on $[0,T]$. We now focus on the upper bound. Observe that

$$\begin{aligned}\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w &= -\widetilde{\Lambda}\operatorname{Im} w \\ &+ K * \Bigg[\Big(C_1 + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\widetilde{\Lambda}\Big)\operatorname{Im} w + \big(b+\rho\sqrt{c}\operatorname{Re} w + f_2\big)\Big(\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w\Big) \\ &+ \bigg(\Big(\widetilde{\Lambda}c+\rho\sqrt{c}\Big)\operatorname{Re}\psi_w + \Big(\widetilde{\Lambda}-\Lambda\Big)\Big(\int_{\mathbb{R}_+} z\left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z}-1\right)\nu(\mathrm{d}z) + d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right)\Big)\bigg)\operatorname{Im} w\Bigg].\end{aligned}$$

We then take the unique, nonnegative, continuous solution $u_1\colon [0,T]\to\mathbb{R}_+$ of the linear equation

$$u_1 = \widetilde{\Lambda}\operatorname{Im} w + K * \Big[\Big|C_1 + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\widetilde{\Lambda}\Big|\operatorname{Im} w + \big(b+\rho\sqrt{c}\operatorname{Re} w + f_2\big)u_1\Big].$$

We infer that $u_1 - (\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w) \geq 0$ since $\widetilde{\Lambda}c+\rho\sqrt{c}$, $\widetilde{\Lambda}-\Lambda \geq 0$, and $u_1 - (\operatorname{Im}\psi_w - \widetilde{\Lambda}\operatorname{Im} w)$ satisfies (on $[0,T]$)

$$\begin{aligned}\chi &= 2\widetilde{\Lambda}\operatorname{Im} w + K * \Bigg[2\Big(C_1 + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\widetilde{\Lambda}\Big)^-\operatorname{Im} w + \big(b+\rho\sqrt{c}\operatorname{Re} w + f_2\big)\chi \\ &- \bigg(\Big(\widetilde{\Lambda}c+\rho\sqrt{c}\Big)\operatorname{Re}\psi_w + \Big(\widetilde{\Lambda}-\Lambda\Big)\Big(\int_{\mathbb{R}_+} z\left(e^{\operatorname{Re}\widetilde{\psi_w}\cdot z}-1\right)\nu(\mathrm{d}z) + d\left(\operatorname{Re}\widetilde{\psi_w}, \operatorname{Im}\widetilde{\psi_w}\right)\Big)\bigg)\operatorname{Im} w\Bigg].\end{aligned}$$

To end, we introduce the unique, nonnegative, continuous solution $\overline{u_1}\colon \mathbb{R}_+\to\mathbb{R}_+$ of the linear equation

$$\overline{u_1} = \widetilde{\Lambda}\,|\operatorname{Im} w| + K * \Big[\Big|C_1 + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\widetilde{\Lambda}\Big|\,|\operatorname{Im} w| + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\overline{u_1}\Big], \tag{5.53}$$

and since $\overline{u_1} - u_1$ satisfies the linear equation $\chi = K * [-f_2\, u_1 + (b+\rho\sqrt{c}\operatorname{Re} w)\chi]$ on $[0,T]$, we conclude that $\overline{u_1} \geq u_1$ on the same interval. Therefore, $\operatorname{Im}\psi_w \leq \overline{u_1} + \widetilde{\Lambda}\operatorname{Im} w$ on $[0,T]$.

In the case $\operatorname{Im} w \leq 0$ the argument is analogous, but the upper and lower bounds are inverted. Specifically, with the same steps as the ones just carried out, we have $-\overline{u_1} - \widetilde{\Lambda}|\operatorname{Im} w| \leq \operatorname{Im}\psi_w \leq \overline{l_1} + \frac{\rho^+}{\sqrt{c}}|\operatorname{Im} w|$ on $[0,T]$.

Therefore, defining the continuous function $u\colon \mathbb{R}_+\to\mathbb{R}_+$ by $u = \overline{l_1} + \overline{u_1} + (\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}})|\operatorname{Im} w|$, we have

$$|\operatorname{Im}\psi_w(t)| \leq u(t), \quad 0\leq t\leq T. \tag{5.54}$$

Taking the real part in (5.7) and using (5.45) we deduce that

$$\begin{aligned}\operatorname{Re}\psi_w = K * \Bigg[&\frac{1}{2}\Big(|\operatorname{Re} w|^2 - \operatorname{Re} w\Big) + \big(b+\rho\sqrt{c}\operatorname{Re} w\big)\operatorname{Re}\psi_w + \frac{c}{2}|\operatorname{Re}\psi_w|^2 \\ &- \frac{1}{2}\Big(|\operatorname{Im} w|^2 + c\,|\operatorname{Im}\psi_w|^2 + 2\rho\sqrt{c}\operatorname{Im} w\operatorname{Im}\psi_w\Big) - \left|\int_{\mathbb{R}_+} e^{\operatorname{Re}\widetilde{\psi_w}\cdot z}\Big(\cos\Big(\operatorname{Im}\widetilde{\psi_w}\cdot z\Big)-1\Big)\nu(\mathrm{d}z)\right| \\ &+ \int_{\mathbb{R}_+}\Big(e^{\operatorname{Re}\psi_w\cdot z}\big(e^{-\Lambda\operatorname{Re} wz}-1\big) - \operatorname{Re} w\big(e^{-\Lambda z}-1\big)\Big)\nu(\mathrm{d}z) + \int_{\mathbb{R}_+}\Big(e^{\operatorname{Re}\psi_w\cdot z} - 1 - \operatorname{Re}\psi_w\cdot z\Big)\nu(\mathrm{d}z)\Bigg]\end{aligned}$$

on $[0,T]$. Since $|\cos(x)-1| = 1-\cos(x) \le x^2/2$, $x \in \mathbb{R}$, by (5.54) we have

$$\left|\int_{\mathbb{R}_+} e^{\mathrm{Re}\,\widetilde{\psi_w}\cdot z}\left(\cos\left(\mathrm{Im}\,\widetilde{\psi_w}\cdot z\right)-1\right)\nu(\mathrm{d}z)\right| \le \frac{1}{2}\left(\int_{\mathbb{R}_+}|z|^2\,\nu(\mathrm{d}z)\right)\left|\mathrm{Im}\,\widetilde{\psi_w}\right|^2$$
$$\le \left(\int_{\mathbb{R}_+}|z|^2\,\nu(\mathrm{d}z)\right)\left(u^2+\Lambda^2\,|\mathrm{Im}\,w|^2\right), \quad \text{on } [0,T]. \tag{5.55}$$

Moreover, notice that by (5.54), since $|\rho|\le 1$

$$\frac{1}{2}\left||\mathrm{Im}\,w|^2 + c\,|\mathrm{Im}\,\psi_w|^2 + 2\rho\sqrt{c}\,\mathrm{Im}\,w\mathrm{Im}\,\psi_w\right| \le \frac{1}{2}\left(|\mathrm{Im}\,w| + \sqrt{c}\,|\mathrm{Im}\,\psi_w|\right)^2 \le |\mathrm{Im}\,w|^2 + cu^2. \tag{5.56}$$

These facts coupled with (5.48) suggest to consider the linear equation

$$l = K * \left[\frac{1}{2}\left(|\mathrm{Re}\,w|^2 - \mathrm{Re}\,w - 2\,|\mathrm{Im}\,w|^2\right) + \int_{\mathbb{R}_+}\left(e^{-\Lambda \mathrm{Re}\,wz} - 1 - \mathrm{Re}\,w\left(e^{-\Lambda z}-1\right)\right)\nu(\mathrm{d}z) - c\,u^2\right.$$
$$\left. - \left(\int_{\mathbb{R}_+}|z|^2\,\nu(\mathrm{d}z)\right)\left(u^2+\Lambda^2\,|\mathrm{Im}\,w|^2\right) + \left(b+\rho\sqrt{c}\,\mathrm{Re}\,w\right)l\right], \tag{5.57}$$

which has a unique, continuous, nonpositive solution l defined on the whole $\mathbb{R}_+$. At this point, observe that the difference $\mathrm{Re}\,\psi_w - l$ satisfies the linear equation

$$\chi = K * \left[\left(b+\rho\sqrt{c}\,\mathrm{Re}\,w\right)\chi + \frac{c}{2}\,|\mathrm{Re}\,\psi_w|^2\right.$$
$$+ \left(|\mathrm{Im}\,w|^2 + cu^2 - \frac{1}{2}\left(|\mathrm{Im}\,w|^2 + c\,|\mathrm{Im}\,\psi_w|^2 + 2\rho\sqrt{c}\,\mathrm{Im}\,w\mathrm{Im}\,\psi_w\right)\right)$$
$$+ \int_{\mathbb{R}_+}\left(e^{\mathrm{Re}\,\psi_w\cdot z} - 1 - \mathrm{Re}\,\psi_w\cdot z\right)\nu(\mathrm{d}z) + \int_{\mathbb{R}_+}\left(e^{\mathrm{Re}\,\psi_w\cdot z}-1\right)\left(e^{-\Lambda \mathrm{Re}\,wz}-1\right)\nu(\mathrm{d}z)$$
$$\left. + \left(\left(\int_{\mathbb{R}_+}|z|^2\,\nu(\mathrm{d}z)\right)\left(u^2+\Lambda^2\,|\mathrm{Im}\,w|^2\right) - \left|\int_{\mathbb{R}_+} e^{\mathrm{Re}\,\widetilde{\psi_w}\cdot z}\left(\cos\left(\mathrm{Im}\,\widetilde{\psi_w}\cdot z\right)-1\right)\nu(\mathrm{d}z)\right|\right)\right].$$

It admits a unique, continuous solution on $[0,T]$ which is nonnegative by (5.55), (5.56) and the fact that $e^x - 1 - x \ge 0$, $x\in\mathbb{R}$. Since $T\in(0,T_{\max})$ was chosen arbitrarily, we infer that

$$l(t) \le \mathrm{Re}\,\psi_w(t) \le 0 \quad \text{and} \quad |\mathrm{Im}\,\psi_w(t)| \le u(t), \quad 0 \le t < T_{\max}.$$

Recalling that l and u are continuous on $\mathbb{R}_+$, and in particular bounded on every compact interval, we conclude that $T_{\max}=\infty$, as desired.

<u>Step III.</u> Consider two global solutions ψ_w, ψ'_w of (5.7), and let $\delta = \psi_w - \psi'_w$ and $\tilde{\delta} = \psi'_w \vee \psi_w$. Then, for every $t\ge 0$,

$$\delta(t) = \int_0^t K(t-s)\left[\left(b+\rho\sqrt{c}\,w + \frac{c}{2}\left(\psi_w+\psi'_w\right)(s) + \int_{\mathbb{R}_+} z\left(e^{\left(-\Lambda w+\tilde{\delta}(s)\right)z}-1\right)\nu(\mathrm{d}z)\right)\delta(s)\right.$$
$$\left. + \int_{\mathbb{R}_+} e^{\left(-\Lambda w+\tilde{\delta}(s)\right)z}\left(e^{\left(\psi_w-\tilde{\delta}\right)(s)z} - e^{\left(\psi'_w-\tilde{\delta}\right)(s)z} - \delta(s)\,z\right)\nu(\mathrm{d}z)\right]\mathrm{d}s. \tag{5.58}$$

We introduce the function $k_w\colon \mathbb{C}_- \times \mathbb{C}_- \to \mathbb{C}$ defined for $(u,v) \in \mathbb{C}_- \times \mathbb{C}_-$ by

$$k_w(u,v) = \begin{cases} \frac{1}{v-u}\int_{\mathbb{R}_+} e^{(-\Lambda w + u\vee v)z}\left(e^{(v-u\vee v)z} - e^{(u-u\vee v)z} - (v-u)z\right)\nu(\mathrm{d}z), & u \neq v \\ 0, & \text{otherwise} \end{cases}. \tag{5.59}$$

We claim that k_w is continuous on its domain. This is a consequence of an application of the mean value theorem to the functions $f_z(u) = e^{uz} - uz$, $u \in \mathbb{C}_-$, with the parameter $z \in \mathbb{R}_+$. Indeed, using the inequality $|1-\cos x| \le x^2$, $x \in \mathbb{R}$,

$$\begin{aligned} |f_z(v) - f_z(u)| &\le z \sup_{\xi\in[u,v]} \left|e^{\xi z} - 1\right| |v-u| \\ &\le z \sup_{\xi\in[u,v]} \left(\left|e^{\operatorname{Re}\xi\cdot z} - 1\right| + \sqrt{2}e^{\frac{1}{2}\operatorname{Re}\xi\cdot z}\left(1-\cos\left(\operatorname{Im}\xi\cdot z\right)\right)^{\frac{1}{2}}\right)|v-u| \\ &\le z\left(\left(1 - e^{(\operatorname{Re}u\wedge\operatorname{Re}v)z}\right) + \sqrt{2}\left(|\operatorname{Im}u| \vee |\operatorname{Im}v|\right)|z|\right)|v-u|, \quad u, v \in \mathbb{C}_-, z \in \mathbb{R}_+. \end{aligned} \tag{5.60}$$

Consequently, the continuity of k_w follows from

$$|f_z(v - u\vee v) - f_z(u - u\vee v)| \le |z|^2\left(1+\sqrt{2}\right)|v-u|^2, \quad u, v \in \mathbb{C}_-, z \in \mathbb{R}_+. \tag{5.61}$$

Coming back to (5.58) we have (on $\mathbb{R}_+$)

$$\begin{aligned} \delta = K * \Bigg[\Bigg(& b + \rho\sqrt{c}\,w + \frac{c}{2}\left(\psi_w + \psi_w'\right)(\cdot) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda w + \tilde{\delta}(\cdot))z} - 1\right)\nu(\mathrm{d}z) \\ &+ k_w\left(\psi_w'(\cdot), \psi_w(\cdot)\right)\Bigg)\delta\Bigg], \end{aligned} \tag{5.62}$$

which is a linear equation admitting the zero function as its unique solution. Hence $\psi_w' = \psi_w$ on $\mathbb{R}_+$, completing the proof of this step.

The fact that $\psi_{\operatorname{Re}w}$ is $\mathbb{R}_-$−valued follows from (5.54), because in this case $u \equiv 0$. This concludes the proof of the statement in (i).

(ii) From (5.45) and (5.46) we deduce that $\operatorname{Re}\mathscr{R}(w,v) \le \mathscr{R}(\operatorname{Re}w, \operatorname{Re}v)$ for every $v \in \mathbb{C}_-$. Taking the real part in (5.7) and recalling that – under Hypothesis 5.1 – the kernel K is nonnegative on $(0,\infty)$ we obtain

$$\operatorname{Re}\psi_w(t) \le \int_0^t K(t-s)\,\mathscr{R}\left(\operatorname{Re}w, \operatorname{Re}\psi_w(s)\right)\mathrm{d}s, \quad t \ge 0.$$

We can then introduce a nonnegative function $\widetilde{\gamma}\colon \mathbb{R}_+ \to \mathbb{R}_+$ defined by the relation

$$\operatorname{Re}\psi_w(t) = -\widetilde{\gamma}(t) + \int_0^t K(t-s)\,\mathscr{R}(\operatorname{Re}w, \operatorname{Re}\psi_w(s))\,\mathrm{d}s, \quad t \ge 0. \tag{5.63}$$

Using (5.7), one can rewrite $\widetilde{\gamma}$ as

$$\widetilde{\gamma}(t) = \int_0^t K(t-s)\left(\mathscr{R}\left(\operatorname{Re}w, \operatorname{Re}\psi_w(s)\right) - \operatorname{Re}\mathscr{R}(w, \psi_w(s))\right)\mathrm{d}s, \quad t \ge 0.$$

Thus $\widetilde{\gamma} \in \mathcal{G}_K$ by [5, Remark B.6]. At this point we subtract (5.63) from (5.7) (with $\operatorname{Re}w$ instead of w) to deduce that $\delta = \psi_{\operatorname{Re}w} - \operatorname{Re}\psi_w$ satisfies

$$\delta(t) = \widetilde{\gamma}(t) + \int_0^t K(t-s)\left(\mathscr{R}\left(\operatorname{Re}w, \psi_{\operatorname{Re}w}(s)\right) - \mathscr{R}\left(\operatorname{Re}w, \operatorname{Re}\psi_w(s)\right)\right)\mathrm{d}s, \quad t \ge 0. \tag{5.64}$$

If we denote by $\tilde{\delta} = \operatorname{Re}\psi_w \vee \psi_{\operatorname{Re} w}$, we then need to study (on $\mathbb{R}_+$)

$$\begin{aligned}
&\mathscr{R}\left(\operatorname{Re} w, \psi_{\operatorname{Re} w}\right) - \mathscr{R}\left(\operatorname{Re} w, \operatorname{Re}\psi_w\right) \\
&\quad = \left(b + \rho\sqrt{c}\operatorname{Re} w + \frac{c}{2}\left(\operatorname{Re}\psi_w + \psi_{\operatorname{Re} w}\right) + \int_{\mathbb{R}_+} z\left(e^{\left(-\Lambda \operatorname{Re} w + \tilde{\delta}\right)z} - 1\right)\nu\left(\mathrm{d}z\right)\right)\delta \\
&\qquad + \int_{\mathbb{R}_+} e^{\left(-\Lambda \operatorname{Re} w + \tilde{\delta}\right)z}\left(e^{\left(\psi_{\operatorname{Re} w} - \tilde{\delta}\right)z} - e^{\left(\operatorname{Re}\psi_w - \tilde{\delta}\right)z} - \delta z\right)\nu\left(\mathrm{d}z\right) \\
&\quad =: \left(w_1\left(\cdot\right) + k_{\operatorname{Re} w}\left(\operatorname{Re}\psi_w\left(\cdot\right), \psi_{\operatorname{Re} w}\left(\cdot\right)\right)\right)\delta,
\end{aligned}$$

with $k_{\operatorname{Re} w}$ as in (5.59). Going back to (5.64),

$$\delta\left(t\right) = \widetilde{\gamma}\left(t\right) + \int_0^t K\left(t-s\right)\left(w_1\left(s\right) + k_{\operatorname{Re} w}\left(\operatorname{Re}\psi_w\left(s\right), \psi_{\operatorname{Re} w}\left(s\right)\right)\right)\delta\left(s\right)\mathrm{d}s, \quad t \geq 0.$$

We can now apply [5, Theorem C.1] in order to conclude that $\delta \geq 0$ on $\mathbb{R}_+$. This yields (5.8) and concludes the proof of (ii). ∎

Appendix 5.B Proof of Proposition 5.5

This section is devoted to the proof of Proposition 5.5, a result which allows to price options on the underlying asset S with maturity $T > 0$.

Proof. Let us define the function $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f\left(m\right) = \mathbb{E}\left[e^{X_T} - \left(e^{X_T} - e^m\right)^+\right]e^{-\frac{1}{2}m} = \mathbb{E}\left[e^{X_T}1_{\{X_T \leq m\}} + e^m 1_{\{m < X_T\}}\right]e^{-\frac{1}{2}m}, \quad m \in \mathbb{R}. \tag{5.65}$$

Denote by μ_T the probability distribution of X_T on $\mathbb{R}$ and note that $f \in L^1(\mathbb{R})$, because, thanks to Tonelli's theorem,

$$\int_{\mathbb{R}} e^{-\frac{1}{2}m}\left[\int_{\mathbb{R}}\left(e^x 1_{\{x \leq m\}} + e^m 1_{\{m < x\}}\right)\mu_T\left(\mathrm{d}x\right)\right]\mathrm{d}m = 4\int_{\mathbb{R}} e^{\frac{1}{2}x}\mu_T\left(\mathrm{d}x\right) = 4\mathbb{E}\left[e^{\frac{1}{2}X_T}\right] < \infty. \tag{5.66}$$

Therefore we can compute the Fourier transform of f as follows

$$\begin{aligned}
\hat{f}\left(\lambda\right) &= \int_{\mathbb{R}} e^{\left(-\frac{1}{2}+i\lambda\right)m}\left[\int_{\mathbb{R}}\left(e^x 1_{\{x \leq m\}} + e^m 1_{\{m < x\}}\right)\mu_T\left(\mathrm{d}x\right)\right]\mathrm{d}m \\
&= \int_{\mathbb{R}}\left[e^x\int_x^\infty e^{\left(-\frac{1}{2}+i\lambda\right)m}\mathrm{d}m + \int_{-\infty}^x e^{\left(\frac{1}{2}+i\lambda\right)m}\mathrm{d}m\right]\mu_T\left(\mathrm{d}x\right) = \frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right), \quad \lambda \in \mathbb{R},
\end{aligned}$$

where in the second equality we are allowed to use Fubini's theorem by (5.66).

Since $|\Psi^{X_T}(\frac{1}{2}+i\lambda)| \leq \mathbb{E}[e^{\frac{1}{2}X_T}] < \infty$ and, by dominated convergence, f in continuous on $\mathbb{R}$, we invoke the Fourier inversion theorem, see for instance [163, Theorem 9.11], to obtain

$$f\left(m\right) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-im\lambda}\frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\mathrm{d}\lambda, \quad m \in \mathbb{R}. \tag{5.67}$$

Combining (5.65) and (5.67) and recalling Corollary 5.4 we deduce that

$$\mathbb{E}\left[\left(e^{X_T} - e^m\right)^+\right] = 1 - \frac{1}{2\pi}\int_{\mathbb{R}} e^{\left(\frac{1}{2}-i\lambda\right)m}\frac{1}{\frac{1}{4}+\lambda^2}\Psi^{X_T}\left(\frac{1}{2}+i\lambda\right)\mathrm{d}\lambda, \quad m \in \mathbb{R}. \tag{5.68}$$

Now, for every $k \in \mathbb{R}$, we can determine the price $C_S(k,T)$ of a call option written on S with log strike k and maturity T. Indeed, taking $m = k - \log(S_0)$ in (5.68) we have

$$C_S(k,T) = \mathbb{E}\left[\left(S_T - e^k\right)^+\right] = S_0 - \frac{1}{2\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}} e^{i\lambda(\log(S_0)-k)}\frac{1}{\frac14+\lambda^2}\Psi^{X_T}\left(\frac12 + i\lambda\right)\mathrm{d}\lambda$$
$$= S_0 - \frac{1}{\pi}\sqrt{S_0 e^k}\int_{\mathbb{R}_+}\operatorname{Re}\left[e^{i\lambda(\log(S_0)-k)}\Psi^{X_T}\left(\frac12 + i\lambda\right)\right]\frac{1}{\frac14+\lambda^2}\mathrm{d}\lambda,$$

which coincides with (5.15). The expression (5.16) for the price $P_S(k,T)$ of a put option with the same underlying, log strike and maturity as before, follows from (5.15), Corollary 5.4, and the put-call parity formula. This completes the proof. ■

Appendix 5.C Proof of Theorem 5.11

This section is devoted to the proof of Theorem 5.11, a result providing estimates for the multi-factor approximation of the Riccati-Volterra equations appearing in the Fourier-Laplace transform of the log returns and VIX^2.

Proof. Fix $T>0$. We first prove Point (i). Take $w \in \mathbb{C}$ such that $\operatorname{Re} w \in [0,1]$ and $n \in \mathbb{N}$, and observe that $|\psi_{w,n}| \le \overline{l_{1,n}} + \overline{u_{1,n}} - l_n + (\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}})|\operatorname{Im} w|$ on $\mathbb{R}_+$. Here $\widetilde{\Lambda} = \max\{\rho^- c^{-1/2},\, \Lambda\}$ and $\overline{l_{1,n}}$ [resp., $\overline{u_{1,n}}$, l_n] is the unique, continuous solution of (5.52) [resp., (5.53), (5.57)] in Appendix 5.A with K_n instead of K. [5, Corollary C.4] guarantees the existence of a positive constant $C_1 = C_1(\rho,b,c,\Lambda,\nu)$ such that

$$\overline{l_{1,n}}(t) + \overline{u_{1,n}}(t) + \left(\widetilde{\Lambda} + \frac{\rho^+}{\sqrt{c}}\right)|\operatorname{Im} w| \le C_1\left(1 + \int_0^T \left|E_{b+\rho^+\sqrt{c},n}(s)\right|\mathrm{d}s\right)|\operatorname{Im} w|, \quad t \in [0,T].$$

Then, recalling the hypothesis of boundedness for $(\int_0^T |E_{b+\rho^+\sqrt{c},n}(s)|\mathrm{d}s)_n$ and using (5.57), another application of [5, Corollary C.4] provides the existence of a constant $C_2 = C_2(\rho,b,c,\Lambda,\nu,\mathbf{m},\mathbf{x},T) > 0$ such that $|l_n(t)| \le C_2(1+|\operatorname{Im} w|^2)$, $t\in[0,T]$. This implies, given that $n \in \mathbb{N}$ is arbitrary, that

$$\sup_{n\in\mathbb{N}}\sup_{t\in[0,T]}|\psi_{w,n}(t)| \le C_3\left(1+|\operatorname{Im} w|^2\right), \quad \text{for some } C_3 = C_3(\rho,b,c,\Lambda,\nu,\mathbf{m},\mathbf{x},T) > 0. \tag{5.69}$$

Since the same argument works for ψ_w, without loss of generality, we assume that the upper bound in (5.69) holds also for ψ_w. Now, from (5.7) and (5.36) we have (on $\mathbb{R}_+$)

$$\psi_w - \psi_{w,n} = (K-K_n) * \mathcal{R}(w,\psi_{w,n}(\cdot)) + K * (\mathcal{R}(w,\psi_w(\cdot)) - \mathcal{R}(w,\psi_{w,n}(\cdot))), \quad n \in \mathbb{N}.$$

For every $v \in \mathbb{C}_-$, recalling the inequality $e^x - 1 - x \le x^2/2$, $x \le 0$, and thanks to the computations in Appendix 5.A (see (5.60))

$$\left|\int_{\mathbb{R}_+}\left[e^{(v-\Lambda w)z} - w\left(e^{-\Lambda z}-1\right) - 1 - vz\right]\nu(\mathrm{d}z)\right|$$
$$\le 4\sqrt{2}\left[\frac{\Lambda^2}{2}(1+|\operatorname{Im} w|) + |v|^2 + \Lambda^2\left(1+|\operatorname{Im} w|^2\right)\right]\int_{\mathbb{R}_+}|z|^2\,\nu(\mathrm{d}z).$$

Then by (5.69) and (5.6) we deduce that there exists a constant $C_4 = C_4(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ such that

$$\sup_{t\in[0,T]} \left|\left(\left(K - K_n\right) * \left(\mathcal{R}\left(w, \psi_{w,n}\left(\cdot\right)\right)\right)\right)(t)\right| \leq C_4 \left(1 + |\operatorname{Im} w|^4\right) \int_0^T \left|K_n\left(s\right) - K\left(s\right)\right| \mathrm{d}s, \quad n \in \mathbb{N}. \quad (5.70)$$

In what follows, we denote by $h_n = (K - K_n) * \mathcal{R}(w, \psi_{w,n}(\cdot))$, i.e., the function that we have just bounded. Next, computations analogous to those carried out to obtain the Volterra equation (5.62) in Appendix 5.A, allow us to write (on $\mathbb{R}_+$)

$$\mathcal{R}\left(w, \psi_w\right) - \mathcal{R}\left(w, \psi_{w,n}\right) = \Bigg(b + \rho\sqrt{c}\, w + \frac{c}{2}\left(\psi_w + \psi_{w,n}\right) + \int_{\mathbb{R}_+} z\left(e^{(-\Lambda w + \psi_{w,n}\vee\psi_w)z} - 1\right)\nu\left(\mathrm{d}z\right) + k_w\left(\psi_{w,n}, \psi_w\right)\Bigg)\left(\psi_w - \psi_{w,n}\right),$$

where k_w is the continuous function in (5.59). Therefore, since $|k_w(u,v)| \leq (1+\sqrt{2})(\int_{\mathbb{R}_+}|z|^2\nu(\mathrm{d}z))|v-u|$ for every $u, v \in \mathbb{C}_-$ (see (5.61)) and recalling (5.69)-(5.70), an application of [5, Corollary C.4] yields

$$\sup_{t\in[0,T]} \left|\psi_w\left(t\right) - \psi_{w,n}\left(t\right) - h_n\left(t\right)\right| \leq C_5 \left(1 + |\operatorname{Im} w|^6\right) \frac{\int_0^T E_{b^+ + \rho^+\sqrt{c} + c_\nu C_3\left(1+|\operatorname{Im} w|^2\right), K}\left(s\right)\mathrm{d}s}{\int_0^T \left|E_{b^+ + \rho^+\sqrt{c}, K}\left(s\right)\right|\mathrm{d}s} \times \int_0^T \left|K_n\left(s\right) - K\left(s\right)\right|\mathrm{d}s, \quad n \in \mathbb{N}. \quad (5.71)$$

for some $C_5 = C_5(\rho, b, c, \Lambda, \nu, \mathbf{m}, \mathbf{x}, T) > 0$ and where $c_\nu = 2(1+\sqrt{2})(\int_{\mathbb{R}_+}|z|^2\nu(\mathrm{d}z))$. Notice that by [96, Proposition 8.1, Chapter 9] and Hypothesis 5.1, $E_{b^+ + \rho^+\sqrt{c} + c_\nu C_3(1+|\operatorname{Im} w|^2), K} \geq 0$. Consequently, thanks to [5, Theorem C.1, Remark B.6], $E_{b^+ + \rho^+\sqrt{c}, K} \leq E_{b^+ + \rho^+\sqrt{c} + c_\nu C_3(1+|\operatorname{Im} w|^2), K}$ a.e. in $\mathbb{R}_+$. Hence the ratio in (5.71) is greater or equal to 1. Combining (5.71) with (5.70) yields (5.38).

In order to prove the final remark about the independence of the constant C in (5.38) with respect to $\mathbf{m}$ and $\mathbf{x}$, note that in the previous argument such a dependence is only due to $\widetilde{C}$, the positive constant given by the hypothesis controlling the sequence $(\int_0^T |E_{b+\rho^+\sqrt{c},n}(s)|\mathrm{d}s)_n$. When $b < 0$, the kernels $-bK_n$ inherit the property of complete monotonicity from K_n. If in addition $\rho < 0$, we can use [96, Theorem 3.1, Chapter 5] to infer that $\int_0^T |E_{b+\rho^+\sqrt{c},n}(s)|\mathrm{d}s = \int_0^T |E_{b,n}(s)|\mathrm{d}s \leq |b|^{-1}$ for every $n \in \mathbb{N}$, and $\int_0^T |E_{b^+ + \rho^+\sqrt{c}, K}(s)|\mathrm{d}s = \|K\|_{L^1([0,T])}$. In particular, in this case C depends on T only via the L^1-norm of K in $[0,T]$ (see (5.69)-(5.71)).

The proof of Point (ii) follows by an analogous argument. In this case we use the estimates in [37, Appendix B.1] (see Appendix 4.B.1 of Chapter 4) and the fact that $\int_0^\delta K_n(s)\mathrm{d}s \leq \int_0^{T\vee\delta} E_{b^+,n}(s)\mathrm{d}s \leq \widetilde{C}$, $n \in \mathbb{N}$. We also combine [5, Corollary C.4], the comparison result for linear Volterra equations in [23, Theorem 2], and the inequality

$$\int_0^\delta h\left(s\right) K_n\left(s + t\right)\mathrm{d}s \leq \int_0^\delta h\left(s\right) K_n\left(s\right)\mathrm{d}s, \quad t \geq 0,$$

which holds also for K by Hypothesis 5.1. ■

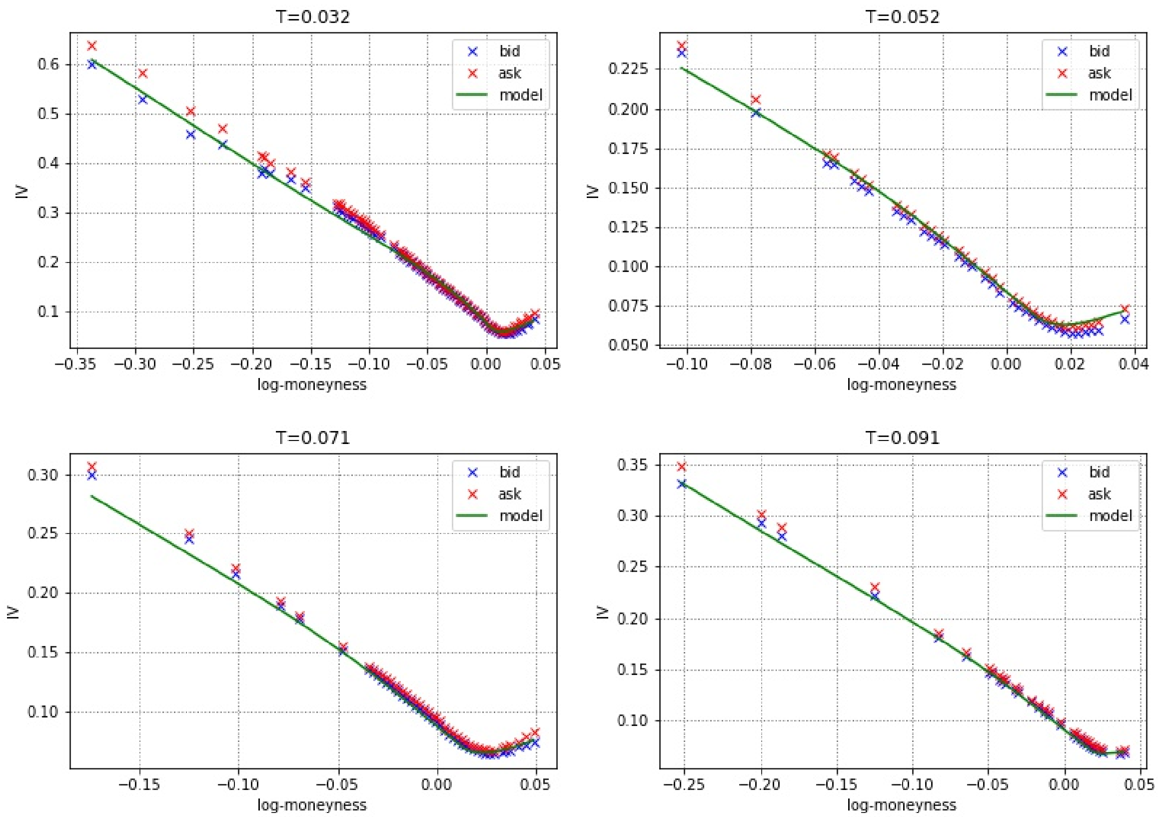

Figure 5.4: Calibrated implied volatility of SPX options on May 19, 2017, using the parameters in Table 5.1. The blue and red crosses are respectively the bid and ask of market implied volatilities. The implied volatility smiles from the model are in green. The abscissa is in log-moneyness and T is time to expiry in years.

Maturity	**Strikes**
May 31, 2017	1700, 1775, 1850, 1900, 1965, 1970, 1980, 2015, 2040, 2095, 2100, 2105, 2115, 2120, 2125, 2135, 2140, 2145, 2150, 2155, 2160, 2165, 2175, 2200, 2210, 2215, 2220, 2225, 2230, 2240, 2245, 2255, 2260, 2265, 2270, 2275, 2280, 2285, 2290, 2295, 2300, 2305, 2310, 2315, 2320, 2325, 2330, 2335, 2340, 2345, 2350, 2355, 2360, 2365, 2370, 2375, 2380, 2385, 2390, 2395, 2400, 2405, 2410, 2415, 2420, 2425, 2430, 2435, 2440, 2445, 2450, 2455, 2465, 2470, 2480.
June 7, 2017	2150, 2200, 2250, 2255, 2270, 2275, 2280, 2300, 2305, 2310, 2320, 2325, 2330, 2335, 2405, 2410, 2415, 2420, 2425, 2430, 2435, 2440, 2445, 2450, 2470.
June 14, 2017	2000, 2100, 2150, 2200, 2220, 2270, 2300, 2305, 2310, 2315, 2320, 2415, 2420, 2425, 2430, 2435, 2440, 2445, 2450, 2460, 2465, 2470, 2480, 2490, 2500.
June 21, 2017	1850, 1950, 1975, 2100, 2190, 2230, 2265, 2270, 2280, 2285, 2290, 2425, 2430, 2435, 2440, 2470, 2475.

Table 5.3: SPX options data on May 19, 2017, considered for the calibration. The listed maturities correspond to $T = 0.032, 0.052, 0.071, 0.091$, respectively.

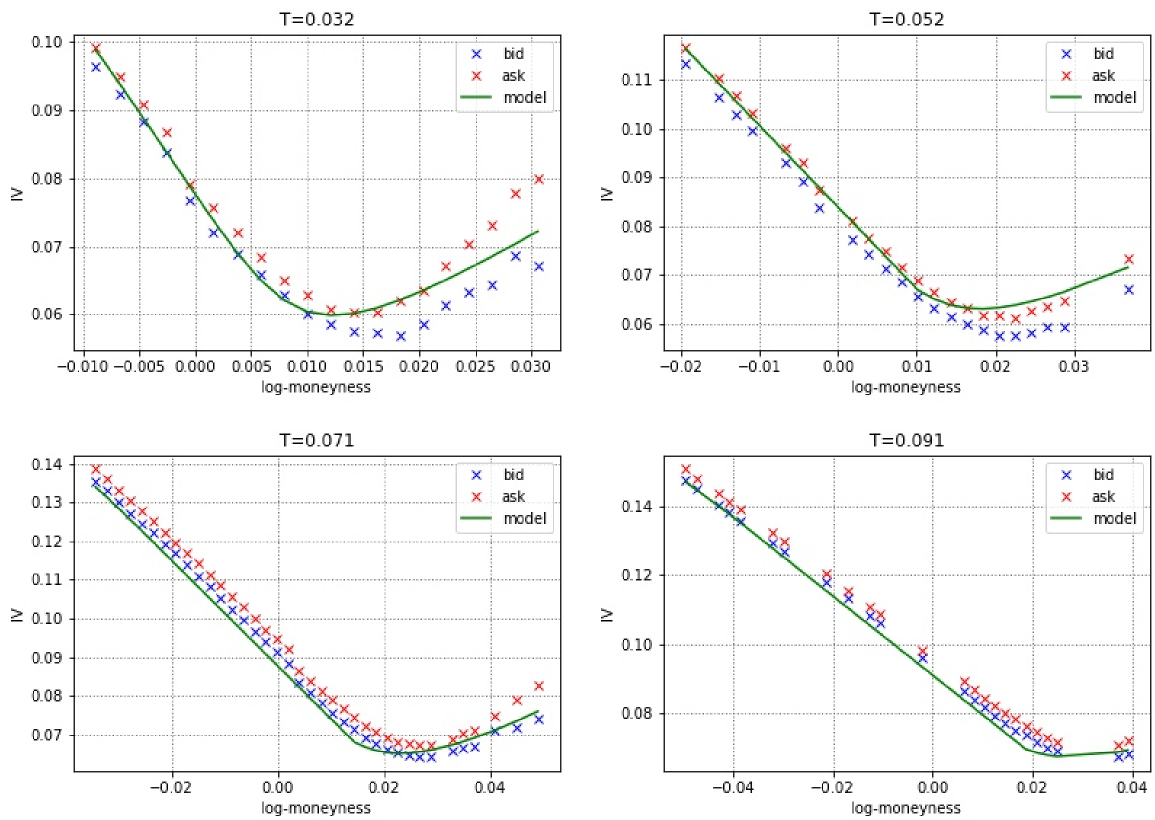

Figure 5.5: Zoom at-the-money of the calibrated implied volatility of SPX options on May 19, 2017, using the parameters in Table 5.1.

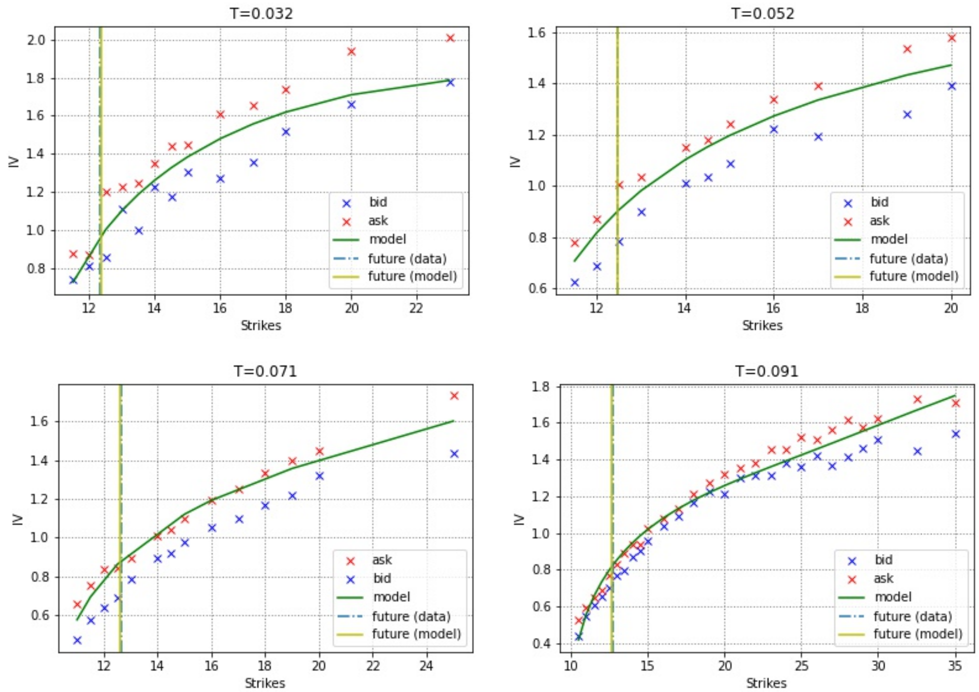

Figure 5.6: Calibrated implied volatility of VIX options on May 19, 2017, using the parameters in Table 5.1. The blue and red crosses are respectively the bid and ask of market implied volatilities. The implied volatility smiles from the model are in green. The abscissa is in strikes and T is time to expiry in years.

Maturity	Strikes
May 31, 2017	11.5, 12.5, 13, 13.5, 14, 14.5, 15, 16, 17, 18, 20, 23.
June 7, 2017	11.5, 12, 12.5, 13, 14, 14.5, 15, 16, 17, 19, 20.
June 14, 2017	11, 11.5, 12.5, 15 , 16, 19, 25.
June 21, 2017	10.5, 11, 11.5, 12, 12.5, 13, 13.5, 14, 14.5, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32.5, 35.

Table 5.4: VIX options data on May 19, 2017, considered for the calibration. The listed maturities correspond to $T = 0.032, 0.052, 0.071, 0.091$, respectively.

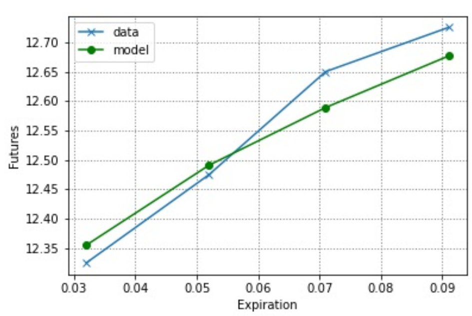

Figure 5.7: VIX term structure.

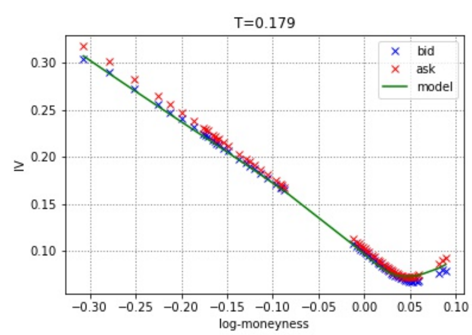

Figure 5.8: Implied volatility of SPX options on May 19, 2017, with time to expiry $T = 0.179$ years.

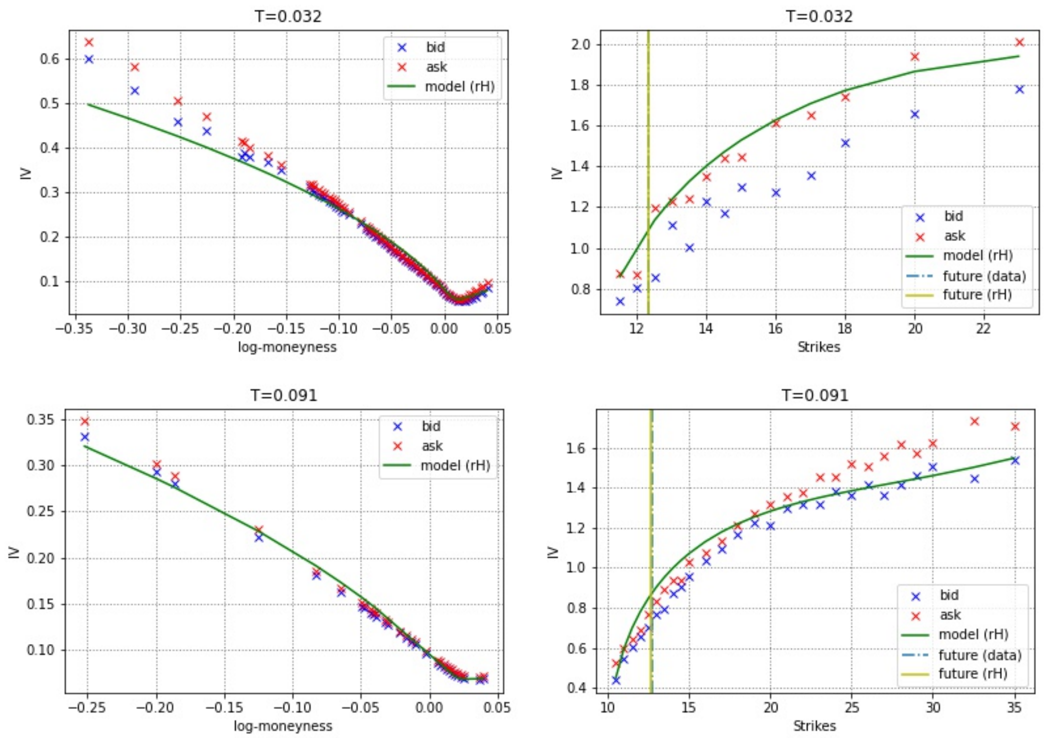

Figure 5.9: Calibrated implied volatility of SPX (left) and VIX (right) options on May 19, 2017, using the rough Heston model. The parameters are reported in Table 5.2. The abscissa is in log-moneyness for SPX options and in strikes for VIX options. The time to expiry in years is $T = 0.032$ in the first line and $T = 0.091$ in the second line.

Maturity	**Strikes**
July 21, 2017	1750, 1800, 1850, 1900, 1925, 1950, 1975, 1995, 2000, 2005, 2015, 2020, 2025, 2030, 2040, 2050, 2075, 2090, 2100, 2110, 2125, 2140, 2160, 2170, 2175, 2180, 2185, 2190, 2350, 2360, 2365, 2370, 2375, 2380, 2385, 2390, 2400, 2405, 2415, 2420, 2425, 2430, 2435, 2440, 2445, 2450, 2455, 2460, 2465, 2470, 2475, 2480, 2485, 2490, 2495, 2500, 2505, 2510, 2515, 2520, 2525, 2580, 2590, 2600.

Table 5.5: SPX options data on May 19, 2017, considered for the calibration of (5.44) in addition to Table 5.3. The maturity corresponds to $T = 0.179$.

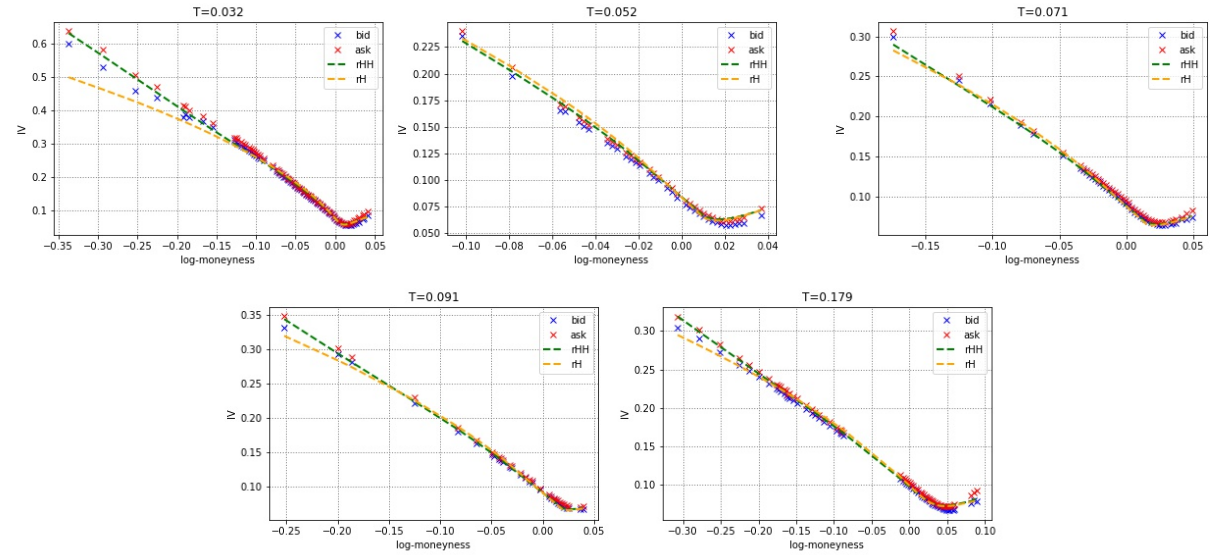

Figure 5.10: SPX implied volatility smiles obtained by minimizing the functional in (5.44) for the rough Heston model (orange) and the rough Hawkes Heston model (green). The blue and red crosses are respectively the bid and ask of market implied volatilities on May 19, 2017. The abscissa is in log-moneyness and T is the time to expiry in years,

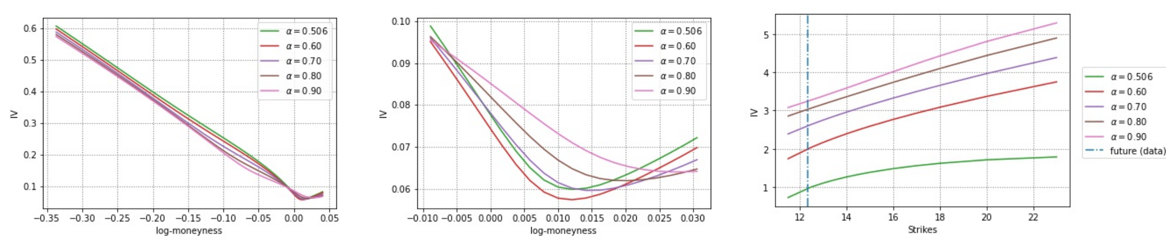

Figure 5.11: Sensitivity of implied volatility for SPX (left and center) and VIX (right) options with respect to the kernel power α for the shortest maturity.

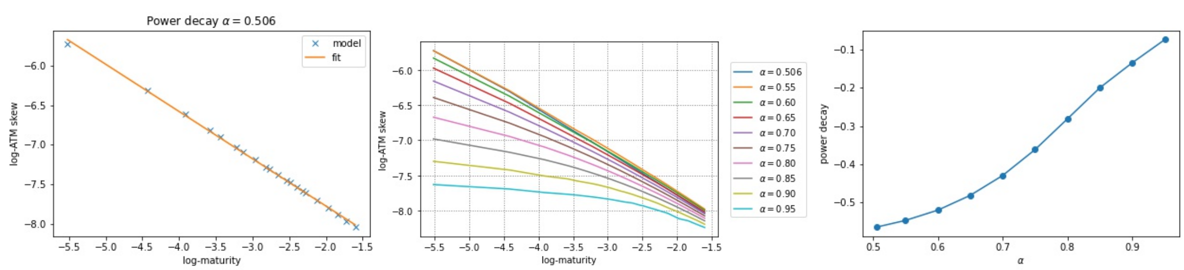

Figure 5.12: Power decay of the ATM volatility skew. On the left, the log-log plot of ATM volatility skew for the calibrated parameters of Table 5.1. At the center, the log-log plot of ATM volatility skew for different values of α; the other parameters are as in Table 5.1. On the right, the fitted power decay of the ATM volatility skew as function of α; the power decay is estimated using the five shortest maturities, i.e. $\log(T) \in [-5.5, -3.5]$.

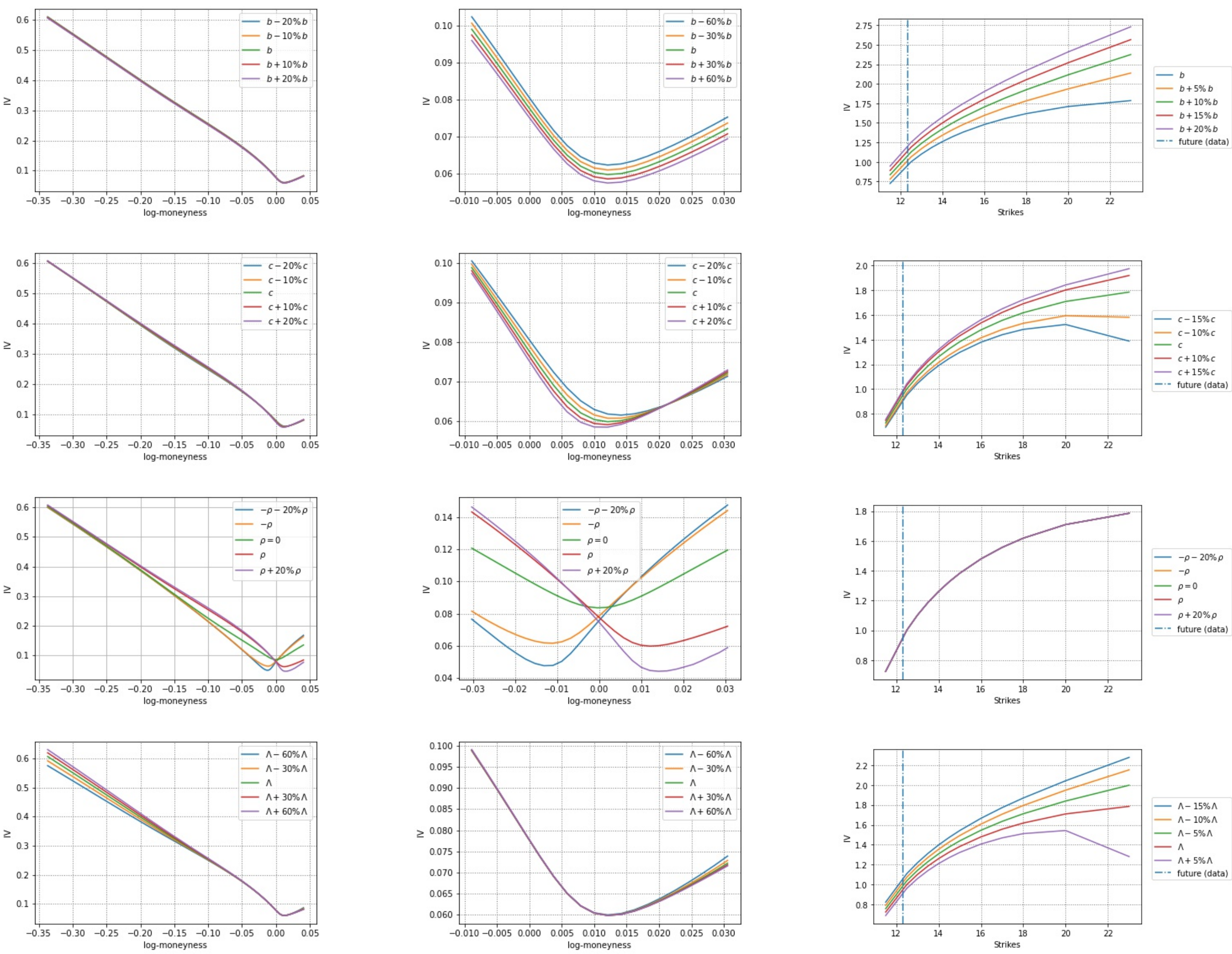

Figure 5.13: Sensitivity of implied volatility for SPX (left, center) and VIX (right) options for the shortest maturity with respect to: the mean reversion speed parameter b (first line), the volatility of volatility c (second line), the correlation ρ (third line), and the jump-leverage Λ (fourth line).

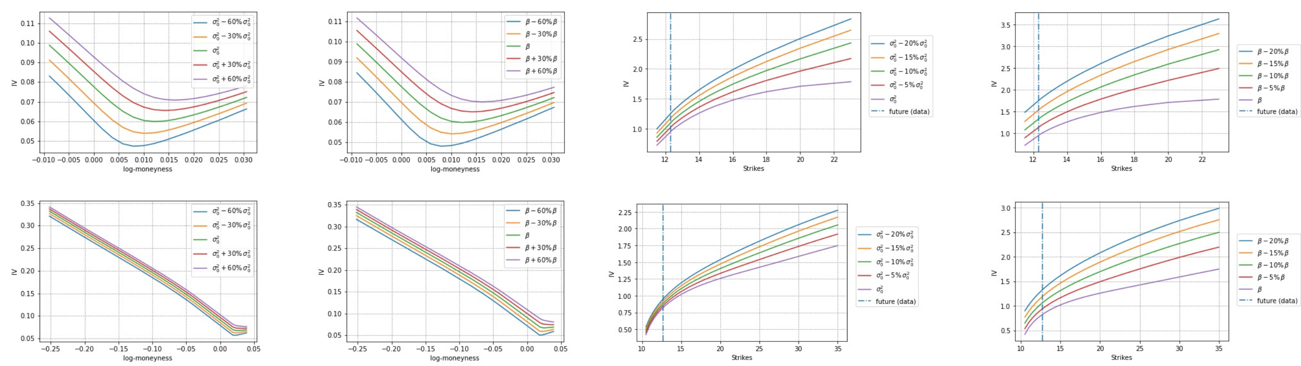

Figure 5.14: Sensitivity of implied volatility for SPX (left, left-center) and VIX (right-center, right) options for the shortest (first line) and longest maturity (second line) with respect to the initial spot variance curve, i.e. intercept σ_0^2, and proportional coefficient β.

Part III

Chapter 6

On the Kolmogorov equation associated with Volterra equations and Fractional Brownian Motion

In this chapter we study a theoretical connection between the subjects treated in Parts I and II. More precisely, we analyze a particular class of SPDEs which represent an abstract, infinite–dimensional reformulation of stochastic Volterra equations in $\mathbb{R}^d$ driven by additive, fractional Brownian motions of Riemann–Liouville type. These SPDEs require an original extension of the drift operator and its Fréchet differentials. We prove that the SPDEs generate a Markov stochastic flow which is twice Fréchet differentiable with respect to the initial data. This stochastic flow is then employed to solve, in the classical sense of infinite–dimensional calculus, the corresponding path–dependent Kolmogorov equation. Notably, we associate a time–dependent infinitesimal generator with the fBm. In Section 6.4, we show some obstructions in the analysis of the mild formulation of the Kolmogorov equation for SPDEs driven by the same infinite–dimensional noise. This problem, which is relevant to the theory of regularization-by-noise, remains open for future research

6.1 Infinite–dimensional reformulations for Volterra SDEs

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a complete filtered probability space, with expectation denoted by $\mathbb{E}$, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions. Fix $d \in \mathbb{N}$ and consider an $\mathbb{R}^d-$valued standard Brownian motion $W = (W_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. In what follows, we denote by $k_2\colon (0,\infty) \to (0,\infty)$ the fractional kernel which controls the noise in the Volterra SDE (I.23), namely

$$k_2(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1}, \quad t > 0, \text{ for some } \alpha \in \left(\frac{1}{2}, 1\right). \tag{6.1}$$

As already mentioned in Introduction, we note that the arguments and results of this chapter continue to hold even when $\alpha \in [1, \frac{3}{2})$, i.e., when the fBM governing (I.23) has Hurst parameter in $[1/2, 1)$, see also Remark 6.4.

Fix $T > 0$. Suppose that the measurable vector field $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies, for some $L > 0$,

$$|b\,(t,x)| \leq L\,(1 + |x|)\,, \qquad |b\,(t,x) - b\,(t,y)| \leq L\,|x-y|\,,$$

for every $t \in [0,T]$ and $x, y \in \mathbb{R}^d$. By (strong) solution of (I.23) we mean a continuous adapted process satisfying the identity for every $t \in [0,T]$, $\mathbb{P}-$a.s. Existence and pathwise uniqueness of strong solutions of (I.23) have been studied in literature under additional requirements on k_1, see, e.g., Equation (2.5) and Theorem 3.3 in [8].

Let H be the Hilbert space $L^2\left(0,T;\mathbb{R}^d\right)$ and denote by $\langle\cdot,\cdot\rangle_H$ the usual inner product. Denoting by $\mathcal{L}(\mathbb{R}^d;H)$ the space of linear and bounded operators from $\mathbb{R}^d$ to H, define $\sigma\colon [0,T] \to \mathcal{L}(\mathbb{R}^d;H)$ by

$$[\sigma(t)\,x](\xi) = k_2(\xi - t)\,1_{\{t<\xi\}}x, \quad x \in \mathbb{R}^d,\ t,\xi \in [0,T]. \tag{6.2}$$

For every $q \geq 2$, we denote by $\mathcal{H}^q$ the space $L^q\big(\Omega;H\big)$, endowed with the usual norm $\|\cdot\|_{\mathcal{H}^q}$, and by $\mathcal{H}_t^q \subset \mathcal{H}^q$ the subspace of $\mathcal{F}_t-$measurable functions, $t \in [0,T]$. Notice that

$$\|\sigma(t)\|_{\mathrm{HS}}^2 \leq d\,\|k_2\|_2^2, \quad t \in [0,T],$$

where $\|\cdot\|_{\mathrm{HS}}$ represents the Hilbert–Schmidt norm and $\|\cdot\|_2$ the norm in $L^2(0,T;\mathbb{R})$. As a consequence, since $\int_0^T \|\sigma(s)\|_{\mathrm{HS}}^2\,\mathrm{d}s < \infty$, we can construct the stochastic integral

$$\Sigma_{s,t} = \int_s^t \sigma(r)\,\mathrm{d}W_r \in \mathcal{H}_t^q, \quad 0 \leq s \leq t \leq T. \tag{6.3}$$

By [66, Theorem 4.36], there exists a constant $C_{d,q} > 0$ such that

$$\|\Sigma_{s,t}\|_{\mathcal{H}^q} \leq C_{d,q}\,\|k_2\|_2\,\sqrt{t-s}, \quad 0 \leq s \leq t \leq T. \tag{6.4}$$

Let Λ be the space $C\left([0,T];\mathbb{R}^d\right)$, and define $B : [0,T] \times \Lambda \to H$ by

$$[B(t,w)](\xi) = k_1(\xi - t)\,1_{\{t<\xi\}}b(t,w(t)), \quad t,\xi \in [0,T]. \tag{6.5}$$

In the sequel, a stochastic process taking values in H will be denoted by, e.g., $(w_t)_{t\in[0,T]}$, namely with the time variable as a subscript. Then, for a fixed $t_0 \in [0,T]$, w_{t_0} is a random function, denoted by $w_{t_0}(\xi)$, $\xi \in [0,T]$.

In the following proposition, we show that it is possible to construct a solution to (I.24), i.e., an $\mathbb{F}-$adapted process with values in H satisfying (I.24) $\mathbb{P}-$a.s., for every $t \in [0,T]$, using a solution of (I.23).

Proposition 6.1. *Let $X = (X_t)_{t\in[0,T]}$ be a solution of (I.23). For every $t \in [0,T]$, define the $\mathbb{R}^d-$valued stochastic process $\theta_t = (\theta_t(\xi))_{\xi\in[t,T]}$ by*

$$\theta_t(\xi) = x_0 + \int_0^t k_1(\xi - s)\,b(s,X_s)\,\mathrm{d}s + \int_0^t k_2(\xi - s)\,\mathrm{d}W_s, \quad \xi \in [t,T].$$

Define the $H-$valued stochastic process $(w_t)_{t\in[0,T]}$ by setting, for each $t \in [0,T]$,

$$w_t(\xi) = \begin{cases} X_\xi, & \xi \leq t, \\ \theta_t(\xi), & \xi > t. \end{cases} \tag{6.6}$$

Then $(w_t)_{t\in[0,T]}$ is a solution of (I.24) *with $\phi \in H$ being the function identically equal to x_0.*

Proof. Fix $t \in [0,T]$. Note that, by the Kolmogorov–Chentsov continuity criterion, there exists a continuous version of the stochastic process $(\int_0^t k_2(\xi - s)\,\mathrm{d}W_s)_{\xi\in[t,T]}$. Hence, also employing the dominated convergence theorem, we deduce that the process θ_t has continuous trajectories $\theta_t(\cdot)$ in $[t,T]$. It follows that w_t defined in (6.6) takes values in H.
In addition, by [66, Proposition 3.18], we observe that w_t is an $\mathcal{F}_t-$measurable random variable, because X is continuous and $\mathbb{F}-$adapted, $\theta_t(\cdot)$ is continuous and $\theta_t(\xi)$ is $\mathcal{F}_t-$measurable for every $\xi \in [t,T]$. Thus, the $H-$valued stochastic process $(w_t)_{t\in[0,T]}$ is $\mathbb{F}-$adapted.

We now want to prove that w_t satisfies (I.24). By (I.23) and the definition of θ_t, we have, $\mathbb{P}-$a.s.,

$$\begin{aligned} w_t\left(\xi\right) &= X_\xi 1_{\{\xi\le t\}} + \theta_t(\xi)1_{\{\xi>t\}} = x_0 + \int_0^{t\wedge\xi} k_1\left(\xi-s\right)b\left(s,X_s\right)\mathrm{d}s + \int_0^{t\wedge\xi} k_2\left(\xi-s\right)\mathrm{d}W_s \\ &= x_0 + \int_0^t k_1\left(\xi-s\right)1_{\{\xi>s\}}b\left(s,X_s\right)\mathrm{d}s + \int_0^t k_2\left(\xi-s\right)1_{\{\xi>s\}}\mathrm{d}W_s, \quad \xi\in[0,T]. \end{aligned} \tag{6.7}$$

We focus on the integral in $\mathrm{d}W$, with the aim of understanding its relation with $\Sigma_{0,t} = \int_0^t \sigma(s)\mathrm{d}W_s$, see (6.3). By (6.2) and [66, Proposition 4.30],

$$\left\langle \int_0^t \sigma\left(s\right)\mathrm{d}W_s, h\right\rangle_H = \int_0^t \left(\int_0^T k_2(\xi-s)1_{\{\xi>s\}}h(\xi)\,\mathrm{d}\xi\right)^\top \mathrm{d}W_s, \quad \mathbb{P}-\text{a.s., for every } h\in H.$$

Moreover, an application of the stochastic Fubini's theorem yields

$$\begin{aligned} \left\langle \int_0^t k_2(\cdot-s)1_{\{\cdot>s\}}\mathrm{d}W_s,\, h\right\rangle_H &= \int_0^T \left(\int_0^t k_2(\xi-s)1_{\{\xi>s\}}\,\mathrm{d}W_s\right)^\top h(\xi)\,\mathrm{d}\xi \\ &= \int_0^t \left(\int_0^T k_2(\xi-s)1_{\{\xi>s\}}h(\xi)\,\mathrm{d}\xi\right)^\top \mathrm{d}W_s, \quad \mathbb{P}-\text{a.s., for every } h\in H. \end{aligned}$$

Considering that H is separable, combining the two previous equations we deduce that

$$\left\langle \int_0^t \sigma\left(s\right)\mathrm{d}W_s, h\right\rangle_H = \left\langle \int_0^t k_2(\cdot-s)1_{\{\cdot>s\}}\mathrm{d}W_s,\, h\right\rangle_H, \quad h\in H,\, \mathbb{P}-\text{a.s.},$$

which in turn implies that

$$\left(\int_0^t \sigma\left(s\right)\mathrm{d}W_s\right)(\xi) = \int_0^t k_2(\xi-s)1_{\{\xi>s\}}\mathrm{d}W_s, \quad \text{for a.e. } \xi\in[0,T],\, \mathbb{P}-\text{a.s.} \tag{6.8}$$

Going back to (6.7), recalling the definition of B in (6.5) and denoting by $\phi\in H$ the function identically equal to x_0, by the standard properties of Bochner's integral we conclude that

$$w_t = \phi + \int_0^t B\left(s,w_s\right)\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W_s, \quad \mathbb{P}-\text{a.s.}$$

Therefore $(w_t)_{t\in[0,T]}$ satisfies (I.24), completing the proof. ■

The previous proposition gives us the classical infinite–dimensional reformulation of the Volterra SDE (I.23), quoted by Equation (I.24) in Introduction. However, for the procedure carried out in Section 6.2, it turns out that a second reformulation is more convenient.

Proposition 6.2. *Let* $(X_t)_{t\in[0,T]}$ *be a solution of* (I.23) *and* $\phi \in H$ *be the function identically equal to* x_0. *Let* $\theta_t(\xi)$ *and* $w_t(\xi)$ *be defined as in Proposition 6.1. Then, for every* $t \in [0,T]$, *the following identity holds:*

$$w_t = \phi + \int_0^t B(s, w_t)\, \mathrm{d}s + \int_0^t \sigma(s)\, \mathrm{d}W_s, \quad \mathbb{P} - a.s. \tag{6.9}$$

Proof. Observing that, for a.e. $\xi \in [0,T]$,

$$\int_0^t k_1(\xi - s)\, 1_{\{\xi > s\}} b(s, X_s)\, \mathrm{d}s = \int_0^t B(s, w_t)(\xi)\, \mathrm{d}s = \left[\int_0^t B(s, w_t)\, \mathrm{d}s\right](\xi),$$

the proof is the same as the one of Proposition 6.1, ■

Motivated by the infinite–dimensional reformulation of Proposition 6.2, in Section 6.2 we focus on studying Equation (6.9). Our aim is to investigate the property of its solutions and the associated Kolmogorov equation, which is the subject of Section 6.3. However, the implementation of this plan is challenging, due to the particular structure of the drift function $B\colon [0,T] \times \Lambda \to H$. More precisely, the issue with the expression of B in (6.5) is that it is meaningful only for continuous functions, as it involves a punctual evaluation. Consequently, unlike the classical case, the functional space Λ in the domain of B is different from the arrival Hilbert space H. This requires an abstract formulation of the problem that, to the best of our knowledge, is not covered by the existing literature.

6.2 Abstract formulation and differentiability of the stochastic flow

In this section, we introduce and study an abstract formulation for the equation (6.9), with a particular attention devoted to the differentiability of its solution with respect to the initial data, see Subsections 6.2.1-6.2.2. In our reasoning, we introduce an extension of the drift operator B, denoted by $\overline{B}$, which is a characterizing and original feature of the approach that we propose.

For every k, $p \in \mathbb{N}$, we denote by $\|\cdot\|_p$ the usual norm on the Banach space $L^p\big(0,T;\mathbb{R}^k\big)$. We denote by

$$H_\square \text{ the Hilbert space } L^2\big((0,T)\times(0,T);\mathbb{R}^d\big) \text{ endowed with the norm } \|\cdot\|_{2,\square}.$$

Recall $H = L^2\big(0,T;\mathbb{R}^d\big)$ and $\Lambda = C\left([0,T];\mathbb{R}^d\right)$. For every $w \in \Lambda$, we consider a map $B(w): [0,T] \times [0,T] \to \mathbb{R}^d$ subject to the next requirement.

Hypothesis 6.1. *The function* $B\colon \Lambda \to H_\square$ *satisfies*

$$\|B(w_1)\|_{2,\square} \le C_0(1 + \|w_1\|_2), \qquad \|B(w_1) - B(w_2)\|_{2,\square} \le C_0 \|w_1 - w_2\|_2, \tag{6.10}$$

for every w_1, $w_2 \in \Lambda$, *for some constant* $C_0 = C_0(d,T) > 0$.
Moreover, given $w \in \Lambda$ *and* $0 < t \le T$, *for a.e.* $r \in (0,t)$ *the function* $B(w)(r,\cdot) \in H$ *is of Volterra–type, namely* $B(w)(r,\xi) = 0$ *for a.e.* $\xi \in (0,r)$, *and depends on* w *only via its restriction* $w|_{(0,t)}$ *to* $(0,t)$.

In the sequel, we are going to progressively introduce stricter hypotheses on the drift map B (see, in particular, Hypotheses 6.2-6.3), which will allow us to prove the main result on the Kolmogorov equation, see Theorem 6.9 in Section 6.3. In Example 6.1, we show a function B, obtained by choosing b in (6.5) with an affine structure, that satisfies these requirements.

Under Hypothesis 6.1, we can invoke the theorem of extension of uniformly continuous functions to uniquely define a continuous map $\overline{B}\colon H \to H_\square$ such that $\overline{B}\big|_\Lambda = B$. Note that $\overline{B}$ satisfies (6.10) for every $w_1, w_2 \in H$. Given $w \in H$ and $r \in (0,T)$, we are going to write $\overline{B}(r,w) = \overline{B}(w)(r,\cdot) \in H$: these maps are well defined for a.e. $r \in (0,T)$.
For a fixed $0 < t \le T$, we remark that also $\overline{B}(r,w)$ is of Volterra–type in the sense of Hypothesis 6.1 for a.e. $r \in (0,t)$, and that it depends on w only via $w|_{(0,t)}$. For these reasons, in the sequel we will refer to Hypothesis 6.1 while talking about $\overline{B}$.

Recall the spaces $\mathcal{H}^q = L^q(\Omega; H)$, $q \ge 2$, and the subspaces $\mathcal{H}^q_t \subset \mathcal{H}^q$ of $\mathcal{F}_t$−measurable functions introduced in Section 6.1, as well as the random variables $\Sigma_{s,t} \in \mathcal{H}^q_t$ in (6.3). For every $0 \le s \le t \le T$ and $\phi \in \mathcal{H}^q$, we are interested in the equation

$$w = \phi + \int_s^t \overline{B}(r,w)\,\mathrm{d}r + \int_s^t \sigma(r)\,\mathrm{d}W_r, \tag{6.11}$$

whose well–posedness in $\mathcal{H}^q$ is given by the next result.

Theorem 6.3. *Under Hypothesis 6.1, for every $q \ge 2$, $\phi \in \mathcal{H}^q$ and $s,t \in [0,T]$, with $s \le t$, there exists a unique solution $w^{s,\phi}_t \in \mathcal{H}^q$ of (6.11). In particular, if $\phi \in \mathcal{H}^q_s$ then $w^{s,\phi}_t \in \mathcal{H}^q_t$.*

Furthermore, the following cocycle property holds in $\mathcal{H}^q$:

$$w^{s,\phi}_t = w^{u,w^{s,\phi}_u}_t, \quad 0 \le s < u < t \le T,\ \phi \in \mathcal{H}^q. \tag{6.12}$$

Proof. Fix $q \ge 2$, $0 \le s \le t \le T$ and $\phi \in \mathcal{H}^q_s$. Consider $N = N(d,T) \in \mathbb{N}$ so big that $C_0\sqrt{T/N} < 1$, where $C_0 = C_0(d,T)$ is the constant in (6.10). Let us introduce an equispaced partition $\{t_k\}_{k=0}^N$ of $[s,t]$ where $t_0 = s$ and $t_N = t$: its mesh $\Delta \le T/N$. Define the mapping $\Gamma^{t_1}_{t_0}\colon \mathcal{H}^q_{t_1} \to \mathcal{H}^q_{t_1}$ by

$$\Gamma^{t_1}_{t_0} w = \phi + \int_{t_0}^{t_1} \overline{B}(r,w)\,\mathrm{d}r + \int_{t_0}^{t_1} \sigma(r)\,\mathrm{d}W_r, \quad w \in \mathcal{H}^q_{t_1}. \tag{6.13}$$

Under Hypothesis 6.1, $\Gamma^{t_1}_{t_0}$ is well defined. Indeed, for every $w \in \mathcal{H}^q_{t_1}$,

$$\begin{aligned}\left\|\Gamma^{t_1}_{t_0} w\right\|^q_{\mathcal{H}^q} &= \mathbb{E}\left[\left\|\Gamma^{t_1}_{t_0} w\right\|^q_2\right] \le 3^{q-1}\mathbb{E}\left[\|\phi\|^q_2 + \left(\int_{t_0}^{t_1}\left(\int_0^T \left|\overline{B}(w)(r,\xi)\right|^2 \mathrm{d}\xi\right)^{\frac{1}{2}} \mathrm{d}r\right)^q + \left\|\int_{t_0}^{t_1}\sigma(r)\,\mathrm{d}W_r\right\|^q_2\right] \\ &\le 3^{q-1}\mathbb{E}\left[\|\phi\|^q_2 + C_0^q \Delta^{\frac{q}{2}}\left(1 + \|w\|_2\right)^q + \left\|\int_{t_0}^{t_1}\sigma(r)\,\mathrm{d}W_r\right\|^q_2\right] < \infty,\end{aligned}$$

where we use Bochner's theorem in the first inequality and the first bound in (6.10), coupled with Jensen's inequality, in the second one. Analogously, using the second inequality in (6.10), we write

$$\begin{aligned}\left\|\Gamma^{t_1}_{t_0} w_1 - \Gamma^{t_1}_{t_0} w_2\right\|_{\mathcal{H}^q} &\le \mathbb{E}\left[\left(\int_{t_0}^{t_1}\left(\int_0^T \left|\overline{B}(w_1) - \overline{B}(w_2)\right|^2(r,\xi)\,\mathrm{d}\xi\right)^{\frac{1}{2}}\mathrm{d}r\right)^q\right]^{\frac{1}{q}} \\ &\le C_0\sqrt{\Delta}\,\|w_1 - w_2\|_{\mathcal{H}^q}, \quad w_1, w_2 \in \mathcal{H}^q_{t_1}.\end{aligned} \tag{6.14}$$

Hence, for our choice of $N \in \mathbb{N}$, the map $\Gamma^{t_1}_{t_0}$ is a contraction in $\mathcal{H}^q_{t_1}$, whose unique fixed point is $\overline{w}_1$. Noting that $\overline{w}_1$ is the unique solution of (6.11) with t_1 instead of t, we denote it by $w^{s,\phi}_{t_1}$.

Since the relation between constants in (6.14), which is necessary to make $\Gamma_{t_0}^{t_1}$ a contraction, does not depend on the initial condition, under Hypothesis 6.1 the previous argument can be iterated to construct the solution $w_t^{s,\phi}$ of (6.11). More precisely, define the map $\Gamma_{t_1}^{t_2}\colon \mathcal{H}_{t_2}^q \to \mathcal{H}_{t_2}^q$ by

$$\Gamma_{t_1}^{t_2} w = \overline{w}_1 + \int_{t_1}^{t_2} \overline{B}\,(r,w)\,\mathrm{d}r + \int_{t_1}^{t_2} \sigma\,(r)\,\mathrm{d}W_r, \quad w \in \mathcal{H}_{t_2}^q.$$

Computations similar to those above show that $\Gamma_{t_1}^{t_2}$ is well defined. Moreover,

$$\left\|\Gamma_{t_1}^{t_2} w_1 - \Gamma_{t_1}^{t_2} w_2\right\|_{\mathcal{H}^q} \le C_0\sqrt{\Delta}\,\|w_1 - w_2\|_{\mathcal{H}^q}\,, \quad w_1,\, w_2 \in \mathcal{H}_{t_2}^q.$$

Thus, $\Gamma_{t_1}^{t_2}$ is a contraction in $\mathcal{H}_{t_2}^q$, whose unique fixed point is $\overline{w}_2 = w_{t_2}^{t_1, w_{t_1}^{s,\phi}}$. Now, by the Volterra–type property of $\overline{B}$ and σ, together with the standard features of the Bochner's and stochastic integrals (see (6.8)), we infer that

$$\left(\int_{t_1}^{t_2} \overline{B}\,(r,\overline{w}_2)\,\mathrm{d}r\right)(\xi) = \left(\int_{t_1}^{t_2} \sigma\,(r)\,\mathrm{d}W_r\right)(\xi) = 0, \quad \text{for a.e. } \xi \in (0,t_1),\ \mathbb{P}-\text{a.s.}, \tag{6.15}$$

whence

$$\overline{w}_2\big|_{(0,t_1)} = \overline{w}_1\big|_{(0,t_1)}, \quad \mathbb{P}-\text{a.s.}$$

Furthermore, $\mathbb{P}-$a.s., for a.e. $r \in (s,t_1)$, $\overline{B}\,(r,\overline{w}_1)$ depends on $\overline{w}_1$ only via $\overline{w}_1\big|_{(0,r)}$, which yields

$$\overline{B}\,(r,\overline{w}_1) = \overline{B}\,(r,\overline{w}_2)\,, \quad \text{for a.e. } r \in (s,t_1),\ \mathbb{P}-\text{a.s.} \tag{6.16}$$

Therefore, recalling (6.13),

$$\begin{aligned}\overline{w}_2 = \phi + \int_s^{t_1} \overline{B}\,(r,\overline{w}_1)\,\mathrm{d}r + \int_{t_1}^{t_2} \overline{B}\,(r,\overline{w}_2)\,\mathrm{d}r + \int_s^{t_2} \sigma\,(r)\,\mathrm{d}W_r \\ = \phi + \int_s^{t_2} \overline{B}\,(r,\overline{w}_2)\,\mathrm{d}r + \int_s^{t_2} \sigma\,(r)\,\mathrm{d}W_r.\end{aligned} \tag{6.17}$$

This shows that $\overline{w}_2$ is a solution of (6.11) with t_2 instead of t.
To prove that $\overline{w}_2$ is in fact the unique solution of this equation, we consider another random variable $\widetilde{w} \in \mathcal{H}_{t_2}^q$ satisfying (6.17). Then, relying on the same properties of $\overline{B}$ and σ as those used above, we deduce that

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\left(\phi + \int_s^{t_1} \overline{B}\left(r, 1_{(0,t_1)}\widetilde{w}\right)\mathrm{d}r + \int_s^{t_1} \sigma\,(r)\,\mathrm{d}W_r\right). \tag{6.18}$$

Moreover, we observe that also $1_{(0,t_1)}\overline{w}_1 \in \mathcal{H}^q$ satisfies (6.18). Therefore, using Bochner's theorem and Jensen's inequality, by Hypothesis 6.1 we can compute

$$\begin{aligned}\left\|1_{(0,t_1)}\left(\overline{w}_1 - \widetilde{w}\right)\right\|_{\mathcal{H}^q}^q &\le \mathbb{E}\left[\left\|\int_s^{t_1}\left(\overline{B}\left(r, 1_{(0,t_1)}\overline{w}_1\right) - \overline{B}\left(r, 1_{(0,t_1)}\widetilde{w}\right)\right)\mathrm{d}r\right\|_2^q\right] \\ &\le \mathbb{E}\left[\left(\int_s^{t_1}\left\|\overline{B}\left(r, 1_{(0,t_1)}\overline{w}_1\right) - \overline{B}\left(r, 1_{(0,t_1)}\widetilde{w}\right)\right\|_2 \mathrm{d}r\right)^q\right] \\ &\le \Delta^{\frac{q}{2}}\mathbb{E}\left[\left\|\overline{B}\left(1_{(0,t_1)}\overline{w}_1\right) - \overline{B}\left(1_{(0,t_1)}\widetilde{w}\right)\right\|_{2,\square}^q\right] \le \Delta^{\frac{q}{2}} C_0^q \left\|1_{(0,t_1)}\left(\overline{w}_1 - \widetilde{w}\right)\right\|_{\mathcal{H}^q}^q,\end{aligned}$$

which allow us to conclude, recalling that $\sqrt{\Delta}C_0 < 1$,

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\overline{w}_1, \quad \mathbb{P}-\text{a.s.}$$

Going back to (6.17), by (6.13) and the previous equality we have, $\mathbb{P}-$a.s.,

$$\widetilde{w} = \phi + \int_s^{t_1} \overline{B}\,(r, \overline{w}_1)\,\mathrm{d}r + \int_s^{t_1} \sigma\,(r)\,\mathrm{d}W_r + \int_{t_1}^{t_2} \overline{B}\,(r, \widetilde{w})\,\mathrm{d}r + \Sigma_{t_1,t_2} = \overline{w}_1 + \int_{t_1}^{t_2} \overline{B}\,(r, \widetilde{w})\,\mathrm{d}r + \Sigma_{t_1,t_2}.$$

It follows that $\widetilde{w}$ is a fixed point of the map $\Gamma_{t_1}^{t_2}$ in $\mathcal{H}_{t_2}^q$: by uniqueness, we obtain $\widetilde{w} = \overline{w}_2$. Hence $\overline{w}_2$ is the unique solution of (6.11) with t_2 instead of t, which we denote by $w_{t_2}^{s,\phi}$.

This argument by steps can be repeated to cover the whole interval $[s,t]$. In this way, we obtain the unique solution $w_t^{s,\phi}$of (6.11) in $\mathcal{H}_t^q$. The same procedure also works when the initial condition $\phi \in \mathcal{H}^q$, i.e., when ϕ is not necessarily $\mathcal{F}_s-$measurable. In such a case, it provides a unique solution $w_t^{s,\phi} \in \mathcal{H}^q$.

The cocycle property in (6.12) follows by a similar reasoning. Indeed, if we fix $u \in (s,t)$, then by the Volterra–type property of $\overline{B}$ and σ (cfr. (6.15)) we have

$$w_t^{u,w_u^{s,\phi}}\Big|_{(0,u)} = w_u^{s,\phi}\Big|_{(0,u)}, \quad \mathbb{P}-\text{a.s.} \tag{6.19}$$

Invoking again Hypothesis 6.1 as in (6.16),

$$\begin{aligned} w_t^{u,w_u^{s,\phi}} &= \phi + \int_s^u \overline{B}\left(r, w_u^{s,\phi}\right)\mathrm{d}r + \int_u^t \overline{B}\left(r, w_t^{u,w_u^{s,\phi}}\right)\mathrm{d}r + \int_s^t \sigma\,(r)\,\mathrm{d}W_r \\ &= \phi + \int_s^t \overline{B}\left(r, w_t^{u,w_u^{s,\phi}}\right)\mathrm{d}r + \int_s^t \sigma\,(r)\,\mathrm{d}W_r, \end{aligned}$$

hence the equality in (6.12) is inferred by the uniqueness of the solution of (6.11). The proof is now complete. ■

Remark 6.1. *The cocycle property in* (6.12) *(see also* (6.19)*) yields* $w_t^{s,\phi}(\xi) = w_u^{s,\phi}(\xi)$ *for a.e.* $\xi \in (0,u)$, $\mathbb{P}-$*a.s., for every* $0 \le s \le u \le t \le T$ *and* $\phi \in \mathcal{H}^q$, $q \ge 2$.

Remark 6.2. *For every* $p \in (2,(1-\alpha)^{-1})$, *the fractional kernel* k_2 *in* (6.1) *belongs to the space* $L^p\big(0,T;\mathbb{R}\big)$.

As a consequence, according to [144, Lemma 8.27*, Theorem* 8.29*], the stochastic integral* $\Sigma_{s,t}$ *in* (6.3) *belongs to the space*

$$\mathcal{L}_t^p = \left(L_t^p\,(\Omega; L^p)\,, \|\cdot\|_{\mathcal{L}^p}\right), \qquad \text{where} \qquad L^p = L^p\big(0,T;\mathbb{R}^d\big).$$

As before, the subscript t in the previous expression indicates a space of $\mathcal{F}_t-$*measurable random variables. Moreover, the following inequality holds (cfr.* (6.4)*):*

$$\|\Sigma_{s,t}\|_{\mathcal{L}^p} \le C_{d,p}\,\|k_2\|_p\,\sqrt{t-s}, \quad \text{for some } C_{d,p} > 0. \tag{6.20}$$

We denote by

$L_\square^p$ *the Banach space* $L^p\big(\,(0,T)\times(0,T)\,;\mathbb{R}^d\big)$, *endowed with the norm* $\|\cdot\|_{p,\square}$.

In addition to Hypothesis 6.1, suppose that $B\colon \Lambda \to L^p_\square$ *and that it satisfies*

$$\left\|B\left(w_1\right)\right\|_{p,\square} \leq C_{0,p}\left(1+\left\|w_1\right\|_p\right), \qquad \left\|B\left(w_1\right)-B\left(w_2\right)\right\|_{p,\square} \leq C_{0,p}\left\|w_1-w_2\right\|_p, \tag{6.21}$$

for every w_1, $w_2 \in \Lambda$, *for some constant* $C_{0,p} = C_{0,p}(d,T) > 0$. *Note that* $\overline{B}\colon H \to H_\square$ *satisfies* (6.21) *for every* w_1, $w_2 \in L^p$.

In this framework, one can argue as in the proof of Theorem 6.3 to infer that, for every $\phi \in \mathcal{L}^p_s$, *there exists a unique solution* $w^{s,\phi}_t$ *of* (6.11) *belonging to the space* $\mathcal{L}^p_t$.

The following corollary to Theorem 6.3 gives a Lipschitz–type dependence of the solution $w^{s,\phi}_t$ of (6.11) on the initial condition ϕ, which combined with (6.12) allows to prove the $\mathbb{F}-$Markov property of the process $(w^{s,\phi}_t)_{t\in[s,T]}$.

Corollary 6.4. *Let* $q \geq 2$. *Under Hypothesis 6.1, there exists a constant* $C_1 = C_1\left(d,q,T\right) > 0$ *such that, for every* $0 \leq s < t \leq T$,

$$\left\|w^{s,\phi}_t - w^{s,\psi}_t\right\|_{\mathcal{H}^q} \leq C_1 \left\|\phi-\psi\right\|_{\mathcal{H}^q}, \quad \phi,\, \psi \in \mathcal{H}^q. \tag{6.22}$$

In addition, for all $s \in [0,T]$ *and* $\phi \in \mathcal{H}^q_s$, *the process* $(w^{s,\phi}_t)_{t\in[s,T]}$ *is* $\mathbb{F}-$*Markov, and*

$$\mathbb{E}\left[\Phi\left(w^{s,\phi}_u\right)\middle|\mathcal{F}_t\right] = \mathbb{E}\left[\Phi\left(w^{t,\psi}_u\right)\right]\Big|_{\psi=w^{s,\phi}_t}, \quad \mathbb{P}-a.s.,\ s\leq t\leq u\leq T,\ \Phi\in\mathcal{B}_b(H), \tag{6.23}$$

where $\mathcal{B}_b(H)$ *denotes the space of bounded Borel measurable functions from* H *to* $\mathbb{R}$.

Proof. Fix $q \geq 2$, $0 \leq s < t \leq T$ and consider $N = N\left(d,T\right) \in \mathbb{N}$ so big that $2C_0\sqrt{T/N} < 2^{1/q}$, where $C_0 = C_0\left(d,T\right)$ is the constant in (6.10). Moreover, take an equispaced partition $\{t_k\}_{k=0}^N$ of $[s,t]$ where $t_0 = s$ and $t_N = t$. By (6.10)-(6.11), for every ϕ, $\psi \in \mathcal{H}^q$,

$$\begin{aligned}\left\|w^{s,\phi}_{t_1} - w^{s,\psi}_{t_1}\right\|_2^q &\leq 2^{q-1}\left\|\phi-\psi\right\|_2^q + 2^{q-1}\left(\frac{T}{N}\right)^{\frac{q}{2}}\left\|\overline{B}\left(w^{s,\phi}_{t_1}\right)-\overline{B}\left(w^{s,\psi}_{t_1}\right)\right\|_{2,\square}^q \\ &\leq 2^{q-1}\left\|\phi-\psi\right\|_2^q + 2^{q-1}C_0^q\left(\frac{T}{N}\right)^{\frac{q}{2}}\left\|w^{s,\phi}_{t_1} - w^{s,\psi}_{t_1}\right\|_2^q, \quad \mathbb{P}\text{–a.s.},\end{aligned}$$

hence

$$\left\|w^{s,\phi}_{t_1} - w^{s,\psi}_{t_1}\right\|_2^q \leq 2^{q-1}\left(1-2^{q-1}C_0^q\left(\frac{T}{N}\right)^{\frac{q}{2}}\right)^{-1}\left\|\phi-\psi\right\|_2^q, \quad \mathbb{P}-\text{a.s.}$$

Thus, by the cocycle property in (6.12), for every ϕ, $\psi \in \mathcal{H}^q$,

$$\begin{aligned}\left\|w^{s,\phi}_t - w^{s,\psi}_t\right\|_2^q &= \left\|w_{t_N}^{t_{N-1},w^{s,\phi}_{t_{N-1}}} - w_{t_N}^{t_{N-1},w^{s,\psi}_{t_{N-1}}}\right\|_2^q \\ &\leq 2^{q-1}\left(1-2^{q-1}C_0^q\left(\frac{T}{N}\right)^{\frac{q}{2}}\right)^{-1}\left\|w_{t_{N-1}}^{t_{N-2},w^{s,\phi}_{t_{N-2}}} - w_{t_{N-1}}^{t_{N-2},w^{s,\psi}_{t_{N-2}}}\right\|_2^q \\ &\leq 2^{N(q-1)}\left(1-2^{q-1}C_0^q\left(\frac{T}{N}\right)^{\frac{q}{2}}\right)^{-N}\left\|\phi-\psi\right\|_2^q, \quad \mathbb{P}\text{–a.s.},\end{aligned}$$

which shows (6.22) upon taking expectations and $q-$th root, as desired.

The Markov property of the process $(w_t^{s,\phi})_{t\in[s,T]}$, $\phi \in \mathcal{H}_s^q$, is a consequence of (6.23). In turn, the equality in (6.23) can be readily obtained by paralleling the monotone class argument in [66, Theorem 9.14], which essentially relies on the cocycle property in (6.12) and the Lipschitz–continuous dependence in (6.22). Thus, the proof is complete. ■

6.2.1 First–order differentiability in the initial data

In this subsection we focus on deterministic initial conditions for (6.11), i.e., $\phi \in H$. From now on, we denote the Hilbert space $\mathcal{H}^2 = L^2(\Omega; H)$ simply by $\mathcal{H}$.

In order to study the first–order Fréchet differentiability of $w_t^{s,\phi}$ in H, we require hypotheses on B which are stronger than Hypothesis 6.1. In fact, we need some conditions on the Frechét differentiability of B in the normed space $(\Lambda, \|\cdot\|_2)$. In the sequel, we write Λ_2 for $(\Lambda, \|\cdot\|_2)$ to have a compact notation.

Hypothesis 6.2. *The map $B\colon \Lambda \to H_\square$ satisfies Hypothesis 6.1. Moreover, B is Λ_2-Fréchet differentiable, and there exists a constant $C_0 = C_0\,(d,T) > 0$ such that*

$$\|DB\,(w_1)\,(w_2)\|_{2,\square} \le C_0\,\|w_2\|_2\,, \quad w_1, w_2 \in \Lambda, \tag{6.24}$$

and

$$\|DB\,(w_1) - DB\,(w_2)\|_{\mathcal{L}(\Lambda_2;H_\square)} \le C_0\,\|w_1 - w_2\|_2^\gamma\,, \quad w_1, w_2 \in \Lambda, \text{ for some } \gamma \in (0,1]\,. \tag{6.25}$$

Without loss of generality, we assume the constant C_0 in (6.24)-(6.25) to be the same as the one in (6.10).
Under Hypothesis 6.2, precisely by (6.24) and the theorem of extension of uniformly continuous functions, for every $w_1 \in \Lambda$ it is possible to extend $DB\,(w_1) \in \mathcal{L}\,(\Lambda; H_\square)$ to an operator $\overline{DB}\,(w_1) \in \mathcal{L}\,(H; H_\square)$ satisfying (6.24) for all $w_2 \in H$. Moreover, by (6.25),

$$\left\|\overline{DB}\,(w_1) - \overline{DB}\,(w_2)\right\|_{\mathcal{L}(H;H_\square)} = \|DB\,(w_1) - DB\,(w_2)\|_{\mathcal{L}(\Lambda_2;H_\square)} \le C_0\,\|w_1 - w_2\|_2^\gamma\,, \quad w_1, w_2 \in \Lambda, \tag{6.26}$$

hence we can extend (without changing the notation)

$$\overline{DB}\colon H \to \mathcal{L}\,(H; H_\square), \text{ with } \overline{DB} \text{ satisfying (6.24)-(6.26) for every } w_1, w_2 \in H. \tag{6.27}$$

We want to show that $\overline{B}$ is $H-$Fréchet differentiable, with $D\overline{B} = \overline{DB}$. By Taylor's formula applied on B, recalling that $\overline{B}\big|_\Lambda = B$ and $\overline{DB}(w_1)\big|_\Lambda = DB(w_1)$, $w_1 \in \Lambda$, we write

$$\begin{aligned} &\overline{B}\,(w_2) - \overline{B}\,(w_1) - \overline{DB}\,(w_1)\,(w_2 - w_1) = r\,(w_1, w_2)\,, \quad w_1, w_2 \in \Lambda, \text{ where} \\ &r\,(x,y) = \int_0^1 \left(\overline{DB}\,(x + h\,(y-x)) - \overline{DB}\,(x)\right)(y-x)\,\mathrm{d}h, \quad x,\, y \in H. \end{aligned} \tag{6.28}$$

Note that $r\colon H \times H \to H_\square$ is continuous. Indeed, for every $x, y \in H$ and every sequence $H \times H \ni (x_n, y_n) \to (x,y)$ as $n \to \infty$, by Bochner's theorem and (6.27), after some algebraic computations we deduce that

$$\begin{aligned} \|r\,(x_n,y_n) - r\,(x,y)\|_{2,\square} \le{}& \frac{C_0}{\gamma+1}\,\|y_n - x_n\|_2^\gamma\,\|y_n - y + x - x_n\|_2 \\ &+ C_0\left(\frac{1}{\gamma+1}\,\|y_n - y + x - x_n\|_2^\gamma + 2\,\|x - x_n\|_2^\gamma\right)\|y - x\|_2 \to 0, \quad \text{as } n \to \infty. \end{aligned}$$

It then follows from the continuity of $\overline{B}$ in H and (6.27) that (6.28) holds for every $w_1, w_2 \in H$. Moreover, since by (6.27) $\|r(x,y)\|_{2,\square} \le C_0(\gamma+1)^{-1}\|y-x\|_2^{1+\gamma}$, x, $y \in H$, we conclude that

$$\overline{B}(w_2) - \overline{B}(w_1) - \overline{DB}(w_1)(w_2 - w_1) = \mathrm{o}(\|w_2 - w_1\|_2), \quad w_1, w_2 \in H.$$

Therefore $\overline{B}$ is $H-$Fréchet differentiable, with $D\overline{B} = \overline{DB}$.

We also notice that, for every w_1, $w_2 \in H$ and $0 < t \le T$,

$$D\overline{B}(w_1)(r, w_2) \coloneqq [D\overline{B}(w_1)(w_2)](r, \cdot) \in H \text{ is of Volterra–type, for a.e. } r \in (0,t), \tag{6.29}$$

and that

$$D\overline{B}(w_1)(r, w_2) \text{ depends on } w_i \text{ only via } w_i\big|_{(0,t)},\ i = 1,2, \text{ for a.e. } r \in (0,t): \tag{6.30}$$

these two properties are inherited from $\overline{B}$, see Hypothesis 6.1.

The next result shows that, under Hypothesis 6.2, the solution $w_t^{s,\phi}$ of (6.11), considered as a map from H to $\mathcal{H}$, is $H-$Fréchet differentiable.

Theorem 6.5. *Under Hypothesis 6.2, for every $0 \le s \le t \le T$, the mapping $w_t^{s,\cdot} \in C^{1+\gamma}(H;\mathcal{H})$. In particular, for every ϕ, $\psi \in H$, $Dw_t^{s,\phi}\psi$ is the unique solution in $\mathcal{H}$ of the following equation:*

$$Dw_t^{s,\phi}\psi = \psi + \int_s^t D\overline{B}\left(w_t^{s,\phi}\right)\left(r, Dw_t^{s,\phi}\psi\right)\mathrm{d}r. \tag{6.31}$$

Furthermore, there exists a constant $C_2 = C_2(d,T) > 0$ such that, for every ϕ, ψ, $\eta \in H$, $\mathbb{P}-$a.s.,

$$\left\|Dw_t^{s,\phi}\eta\right\|_2 \le C_2\|\eta\|_2, \qquad \left\|Dw_t^{s,\phi}\eta - Dw_t^{s,\psi}\eta\right\|_2 \le C_2\left\|w_t^{s,\phi} - w_t^{s,\psi}\right\|_2^{\gamma}\|\eta\|_2. \tag{6.32}$$

Proof. Fix $0 \le s \le t \le T$ and $\phi \in H$. Firstly, we prove the well–posedness in $\mathcal{H}$ of the equation

$$w = \psi + \int_s^t D\overline{B}\left(w_t^{s,\phi}\right)(r, w)\,\mathrm{d}r, \quad \psi \in H. \tag{6.33}$$

Consider $N = N(d,T) \in \mathbb{N}$ so big that $C_0\sqrt{T/N} < 1$, where $C_0 = C_0(d,T)$ is the constant in Hypotheses 6.1-6.2. In addition, take an equispaced partition $\{t_k\}_{k=0}^N$ of $[s,t]$ where $t_0 = s$ and $t_N = t$: its mesh $\Delta \le T/N$. By (6.27) (see also (6.24)) and Bochner's theorem, the following estimate holds:

$$\begin{aligned}\left\|\int_{t_0}^{t_1} D\overline{B}\left(w_t^{s,\phi}\right)(w_1 - w_2)(r,\cdot)\,\mathrm{d}r\right\|_{\mathcal{H}} &\le \sqrt{\Delta}\mathbb{E}\left[\int_{t_0}^{t_1}\int_0^T \left|D\overline{B}\left(w_t^{s,\phi}\right)(w_1 - w_2)\right|^2(r,\xi)\,\mathrm{d}\xi\,\mathrm{d}r\right]^{\frac{1}{2}} \\ &\le \sqrt{\Delta}\mathbb{E}\left[\left\|D\overline{B}\left(w_t^{s,\phi}\right)(w_1 - w_2)\right\|_{2,\square}^2\right]^{\frac{1}{2}} \le C_0\sqrt{\Delta}\|w_1 - w_2\|_{\mathcal{H}}, \quad w_1, w_2 \in \mathcal{H}.\end{aligned} \tag{6.34}$$

Thus, employing a fixed point argument as in the proof of Theorem 6.3, we deduce the existence of a unique solution $\overline{w}_1^{\psi} \in \mathcal{H}$ of (6.33) with t_1 instead of t, for every $\psi \in H$.

We claim that the operator $Dw_{t_1}^{s,\phi}\colon H \to \mathcal{H}$ defined by $Dw_{t_1}^{s,\phi}\psi = \overline{w}_1^{\psi}$, $\psi \in H$, is the Fréchet differential of $w_{t_1}^{s,\phi}$. Indeed, the linearity of $Dw_{t_1}^{s,\phi}$ is straightforward, while the continuity is ensured by the following computation, which can be argued from (6.33) similarly to (6.34):

$$\left\|Dw_{t_1}^{s,\phi}\psi\right\|_2 \le \left(1 - C_0\sqrt{T/N}\right)^{-1}\|\psi\|_2, \quad \mathbb{P}-\text{a.s.}, \psi \in H. \tag{6.35}$$

Moreover, recalling (6.11)-(6.31),

$$\begin{aligned}\left\|w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}-Dw_{t_1}^{s,\phi}h\right\|_{\mathcal{H}} &\leq \sqrt{\Delta}\mathbb{E}\left[\left\|\overline{B}\left(w_{t_1}^{s,\phi+h}\right)-\overline{B}\left(w_{t_1}^{s,\phi}\right)-D\overline{B}\left(w_{t_1}^{s,\phi}\right)Dw_{t_1}^{s,\phi}h\right\|_{2,\square}^2\right]^{\frac{1}{2}}\\
&\leq \sqrt{T/N}\left(\mathbb{E}\left[\left\|D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}-Dw_{t_1}^{s,\phi}h\right)\right\|_{2,\square}^2\right]^{\frac{1}{2}}\right.\\
&\quad\left.+\mathbb{E}\left[\left\|\int_0^1\left(D\overline{B}\left(w_{t_1}^{s,\phi}+u\left(w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}\right)\right)-D\overline{B}\left(w_{t_1}^{s,\phi}\right)\right)\left(w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}\right)\mathrm{d}u\right\|_{2,\square}^2\right]^{\frac{1}{2}}\right)\\
&\leq \sqrt{T/N}C_0\left(\left\|w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}-Dw_{t_1}^{s,\phi}h\right\|_{\mathcal{H}}+\mathbb{E}\left[\left\|w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}\right\|_2^{2(1+\gamma)}\right]^{\frac{1}{2}}\right),\quad h\in H, \qquad (6.36)\end{aligned}$$

where we apply Taylor's formula on $\overline{B}$ for the second inequality and (6.27) together with Bochner's theorem for the third. Notice that $H\subset\mathcal{H}^q$ for every $q\geq 2$. Therefore, by Corollary 6.4 with $q=2(1+\gamma)$, from (6.36) we infer that

$$\left\|w_{t_1}^{s,\phi+h}-w_{t_1}^{s,\phi}-Dw_{t_1}^{s,\phi}h\right\|_{\mathcal{H}}\leq\sqrt{T/N}C_0C_1^{1+\gamma}\left(1-\sqrt{T/N}C_0\right)^{-1}\|h\|_2^{1+\gamma}=\mathrm{o}\left(\|h\|_2\right),\quad h\in H, \quad (6.37)$$

for some constant $C_1=C_1(\gamma,d,T)>0$. This shows that $Dw_{t_1}^{s,\phi}$ is the Fréchet differential of $w_{t_1}^{s,\phi}$, as desired.

Next, consider

$$w=Dw_{t_1}^{s,\phi}\psi+\int_{t_1}^{t_2}D\overline{B}\left(w_{t_2}^{s,\phi}\right)(r,w)\,\mathrm{d}r,\quad \psi\in H: \qquad (6.38)$$

the well–posedness of this equation in $\mathcal{H}$ can be obtained via a fixed–point argument as in the above step. We denote by $\overline{w}_2^{\psi}\in\mathcal{H}$, $\psi\in H$, the unique solution of (6.38).

We argue that $\overline{w}_2^{\psi}$ is the unique solution of (6.33) with t_2 instead of t, for every $\psi\in H$. By the Volterra–type property of $D\overline{B}$ in (6.29) and (6.38) we have, $\mathbb{P}-$a.s.,

$$\overline{w}_2^{\psi}\Big|_{(0,t_1)}=Dw_{t_1}^{s,\phi}\psi\Big|_{(0,t_1)}.$$

Furthermore, thanks to the relation $w_{t_2}^{s,\phi}=w_{t_2}^{t_1,w_{t_1}^{s,\phi}}$ in (6.12) and the properties of $\overline{B}$ under Hypothesis 6.1 we can write, $\mathbb{P}-$a.s.,

$$w_{t_2}^{s,\phi}\Big|_{(0,t_1)}=w_{t_1}^{s,\phi}\Big|_{(0,t_1)}, \qquad (6.39)$$

see Remark 6.1. Consequently, by the property of $D\overline{B}$ in (6.30) and recalling that $Dw_{t_1}^{s,\phi}\psi$ satisfies (6.33) with t_1 instead of t, from (6.38) we conclude that, $\mathbb{P}-$a.s.,

$$\begin{aligned}\overline{w}_2^{\psi}=\psi+\int_s^{t_1}D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(r,Dw_{t_1}^{s,\phi}\psi\right)\mathrm{d}r+\int_{t_1}^{t_2}D\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(r,\overline{w}_2^{\psi}\right)\mathrm{d}r&\\
=\psi+\int_s^{t_2}D\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(r,\overline{w}_2^{\psi}\right)\mathrm{d}r.& \qquad (6.40)\end{aligned}$$

Hence $\overline{w}_2^{\psi}$ solves (6.33) with t replaced by t_2; to prove that it is in fact the unique solution, we consider another random variable $\widetilde{w} \in \mathcal{H}$ satisfying (6.40). Then, by (6.29)-(6.30),

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\left(\psi + \int_s^{t_1} D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(r, 1_{(0,t_1)}\widetilde{w}\right) \mathrm{d}r\right). \tag{6.41}$$

We observe that also $1_{(0,t_1)}\overline{w}_1^{\psi} \in \mathcal{H}$ satisfies (6.41). Therefore, using Bochner's theorem and Jensen's inequality, by (6.27) we can compute

$$\begin{aligned}\left\|1_{(0,t_1)}\left(\overline{w}_1^{\psi} - \widetilde{w}\right)\right\|_{\mathcal{H}}^2 &\le \mathbb{E}\left[\left(\int_s^{t_1} \left\|D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(r, 1_{(0,t_1)}\left(\overline{w}_1^{\psi} - \widetilde{w}\right)\right)\right\|_2 \mathrm{d}r\right)^2\right] \\ &\le \Delta\mathbb{E}\left[\left\|D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(1_{(0,t_1)}\left(\overline{w}_1^{\psi} - \widetilde{w}\right)\right)\right\|_{2,\square}^2\right] \le \Delta C_0^2 \left\|1_{(0,t_1)}\left(\overline{w}_1 - \widetilde{w}\right)\right\|_{\mathcal{H}}^2, \end{aligned} \tag{6.42}$$

which allow us to conclude, recalling that $\sqrt{\Delta}C_0 < 1$,

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\overline{w}_1^{\psi}, \quad \mathbb{P}-\text{a.s.}$$

Going back to (6.40), by (6.33) and the previous equality we have, $\mathbb{P}-$a.s.,

$$\widetilde{w} = \psi + \int_s^{t_1} D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(r, \overline{w}_1^{\psi}\right) \mathrm{d}r + \int_{t_1}^{t_2} D\overline{B}\left(w_{t_2}^{s,\phi}\right)(r, \widetilde{w}) \,\mathrm{d}r = \overline{w}_1^{\psi} + \int_{t_1}^{t_2} D\overline{B}\left(w_{t_2}^{s,\phi}\right)(r, \widetilde{w}) \,\mathrm{d}r.$$

It follows that $\widetilde{w}$ satisfies (6.38): by uniqueness, we obtain $\widetilde{w} = \overline{w}_2^{\psi}$. Hence $\overline{w}_2^{\psi}$ is the unique solution of (6.33) in $\mathcal{H}$ with t_2 instead of t.

We define the operator $Dw_{t_2}^{s,\phi} \colon H \to \mathcal{H}$ by $Dw_{t_2}^{s,\phi}\psi = \overline{w}_2^{\psi}$, $\psi \in H$, and claim that it is the Fréchet differential of $w_{t_2}^{s,\phi}$. To see this, note that the linearity of $Dw_{t_2}^{s,\phi}$ is a consequence of the well–posedness of (6.40). As for the continuity, it is ensured by the following computations, where we use (6.27)-(6.35)-(6.38):

$$\begin{aligned}\left\|Dw_{t_2}^{s,\phi}\psi\right\|_2 &\le \left\|Dw_{t_1}^{s,\phi}\psi\right\|_2 + \int_{t_1}^{t_2} \left\|D\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(r, Dw_{t_2}^{s,\phi}\psi\right)\right\|_2 \mathrm{d}r \\ &\le \left(1 - C_0\sqrt{T/N}\right)^{-1} \|\psi\|_2 + \sqrt{\Delta}C_0 \left\|Dw_{t_2}^{s,\phi}\psi\right\|_2, \quad \mathbb{P}-\text{a.s.}, \ \psi \in H,\end{aligned}$$

whence

$$\left\|Dw_{t_2}^{s,\phi}\psi\right\|_2 \le \left(1 - C_0\sqrt{T/N}\right)^{-2} \|\psi\|_2, \quad \mathbb{P}-\text{a.s.}, \ \psi \in H. \tag{6.43}$$

Moreover, by the cocycle property in (6.12) and reasoning as in (6.36), by (6.11)-(6.38) we obtain, for some constant $c > 0$,

$$\begin{aligned}\left\|w_{t_2}^{s,\phi+h} - w_{t_2}^{s,\phi} - Dw_{t_2}^{s,\phi}h\right\|_{\mathcal{H}} &= \left\|w_{t_2}^{t_1, w_{t_1}^{s,\phi+h}} - w_{t_2}^{t_1, w_{t_1}^{s,\phi}} - Dw_{t_1}^{s,\phi}h - \int_{t_1}^{t_2} D\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(r, Dw_{t_2}^{s,\phi}h\right) \mathrm{d}r\right\|_{\mathcal{H}} \\ &\le \left\|w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} - Dw_{t_1}^{s,\phi}h\right\|_{\mathcal{H}} + \left\|\int_{t_1}^{t_2}\left(\overline{B}\left(w_{t_2}^{s,\phi+h}\right) - \overline{B}\left(w_{t_2}^{s,\phi}\right) - D\overline{B}\left(w_{t_2}^{s,\phi}\right)Dw_{t_2}^{s,\phi}h\right)(r,\cdot)\,\mathrm{d}r\right\|_{\mathcal{H}} \\ &\le c\,\|h\|_2^{1+\gamma} = \mathrm{o}\left(\|h\|_2\right), \quad h \in H,\end{aligned} \tag{6.44}$$

where we also employ (6.37) in the last inequality. This shows that $Dw^{s,\phi}_{t_2}$ is the Fréchet differential of $w^{s,\phi}_{t_2}$, as desired.

Repeating this argument $N-$times, we deduce that $Dw^{s,\phi}_t\colon H\to\mathcal{H}$ defined by $Dw^{s,\phi}_t\psi=\overline{w}^{\psi}_N$, where $\overline{w}^{\psi}_N$ is the unique solution of (6.33) in $\mathcal{H}$, for every $\psi\in H$, is the Fréchet differential of $w^{s,\phi}_t$. In particular, the first bound in (6.32) is true, because (cfr. (6.35)-(6.43))

$$\left\|Dw^{s,\phi}_t\psi\right\|_2\le\left(1-C_0\sqrt{T/N}\right)^{-N}\|\psi\|_2=:\overline{C}\,\|\psi\|_2\,,\quad\mathbb{P}-\text{a.s.},\ \phi,\psi\in H. \tag{6.45}$$

As regards the second inequality in (6.32), by (6.27), (6.31) and (6.45) we have, for every ϕ, ψ, $\eta\in H$, $\mathbb{P}-$a.s.,

$$\begin{aligned}\left\|Dw^{s,\phi}_{t_1}\eta-Dw^{s,\psi}_{t_1}\eta\right\|_2&=\left\|\int_s^{t_1}\left(D\overline{B}\left(w^{s,\phi}_t\right)Dw^{s,\phi}_{t_1}\eta-D\overline{B}\left(w^{s,\psi}_t\right)Dw^{s,\psi}_{t_1}\eta\right)(r,\cdot)\,\mathrm{d}r\right\|_2\\&\le\sqrt{\Delta}\left(\left\|D\overline{B}\left(w^{s,\phi}_t\right)\left(Dw^{s,\phi}_{t_1}\eta-Dw^{s,\psi}_{t_1}\eta\right)\right\|_{2,\square}+\left\|\left(D\overline{B}\left(w^{s,\phi}_t\right)-D\overline{B}\left(w^{s,\psi}_t\right)\right)Dw^{s,\psi}_{t_1}\eta\right\|_{2,\square}\right)\\&\le C_0\sqrt{T/N}\left(\left\|Dw^{s,\phi}_{t_1}\eta-Dw^{s,\psi}_{t_1}\eta\right\|_2+\overline{C}\left\|w^{s,\phi}_t-w^{s,\psi}_t\right\|_2^{\gamma}\|\eta\|_2\right),\end{aligned}$$

where in the first equality we also use (6.30) and (6.39) with t instead of t_2. It follows that

$$\left\|Dw^{s,\phi}_{t_1}\eta-Dw^{s,\psi}_{t_1}\eta\right\|_2\le\left(1-C_0\sqrt{T/N}\right)^{-1}C_0\overline{C}\sqrt{T/N}\left\|w^{s,\phi}_t-w^{s,\psi}_t\right\|_2^{\gamma}\|\eta\|_2\,.$$

By (6.38), we sequentially iterate this computation to obtain the second bound in (6.32) with

$$C_2=\max\{\overline{C},NC_0\overline{C}^2\sqrt{T/N}\}.$$

At this point, taking expectations and using Corollary 6.4 with $q=2\gamma$ (recall that $H\subset\mathcal{H}^q$), by Jensen's inequality we infer that, for some constant $C>0$,

$$\begin{aligned}\left\|Dw^{s,\phi}_t-Dw^{s,\psi}_t\right\|_{\mathcal{L}(H;\mathcal{H})}&=\sup_{\|\eta\|_2\le1}\mathbb{E}\left[\left\|Dw^{s,\phi}_t\eta-Dw^{s,\psi}_t\eta\right\|_2^2\right]^{\frac12}\le C_2\mathbb{E}\left[\left\|w^{s,\phi}_t-w^{s,\psi}_t\right\|_2^{2\gamma}\right]^{\frac12}\\&\le C\,\|\phi-\psi\|_2^{\gamma}\,,\quad\phi,\,\psi\in H.\end{aligned}$$

This shows that $Dw^{s,\cdot}_t\in C^{\gamma}\left(H;\mathcal{L}(H;\mathcal{H})\right)$, completing the proof. ■

6.2.2 Second–order differentiability in the initial data

Recalling the normed space $\Lambda_2=\left(\Lambda,\|\cdot\|_2\right)$, in the sequel we identify $\mathcal{L}(\Lambda_2;\mathcal{L}(\Lambda_2;H_\square))$ with the space $\mathcal{L}(\Lambda_2,\Lambda_2;H_\square)$ of bilinear forms from $\Lambda_2\times\Lambda_2$ to $H_\square$ in the usual way.
For the purpose of investigating the second–order Fréchet differential in H of $w^{s,\phi}_t$, we need to require another condition on B.

Hypothesis 6.3. *The map $B\colon\Lambda\to H_\square$ satisfies Hypothesis 6.2. Moreover, B is twice Λ_2-Fréchet differentiable, and there exists a constant $C_0=C_0\,(d,T)>0$ such that*

$$\left\|D^2B\,(w_1)\,(w_2,w_3)\right\|_{2,\square}\le C_0\,\|w_2\|_2\,\|w_3\|_2\,,\quad w_1,w_2,w_3\in\Lambda, \tag{6.46}$$

and

$$\left\|D^2B\,(w_1)-D^2B\,(w_2)\right\|_{\mathcal{L}(\Lambda_2,\Lambda_2;H_\square)}\le C_0\,\|w_1-w_2\|_2^{\beta}\,,\quad w_1,w_2\in\Lambda,\ \textit{for some }\beta\in(0,1]\,. \tag{6.47}$$

Once again, we can assume that the constant C_0 in (6.46)-(6.47) is the same as the one in (6.10) and (6.24)-(6.25).
By (6.46), we invoke the theorem of extension of uniformly continuous functions to extend, for every $w_1, w_2 \in \Lambda$, the map $D^2B(w_1)(w_2,\cdot) \in \mathcal{L}(\Lambda_2; H_\square)$ to an operator $\overline{D^2B}(w_1)(w_2,\cdot) \in \mathcal{L}(H; H_\square)$ satisfying (6.46) for all $w_3 \in H$. It follows that, by linearity,

$$\left\|\overline{D^2B}(w_1)(w_2) - \overline{D^2B}(w_1)(w_3)\right\|_{\mathcal{L}(H;H_\square)} = \left\|\overline{D^2B}(w_1)(w_2 - w_3)\right\|_{\mathcal{L}(H;H_\square)} \leq C_0 \left\|w_2 - w_3\right\|_2, \quad w_1, w_2, w_3 \in \Lambda,$$

hence we can extend (without changing notation) $\overline{D^2B}(w_1) \in \mathcal{L}(H, H; H_\square)$, for all $w_1 \in \Lambda$. At this point, by (6.47) we infer that, for every $w_1, w_2 \in \Lambda$,

$$\left\|\overline{D^2B}(w_1) - \overline{D^2B}(w_2)\right\|_{\mathcal{L}(H,H;H_\square)} = \left\|\overline{D^2B}(w_1) - \overline{D^2B}(w_2)\right\|_{\mathcal{L}(\Lambda_2,\Lambda_2;H_\square)} \leq C_0 \left\|w_1 - w_2\right\|_2^{\beta}, \tag{6.48}$$

whence, via another extension, from now on we consider

$$\overline{D^2B}\colon H \to \mathcal{L}(H, H; H_\square) \text{ satisfying (6.46)-(6.48) for every } w_i \in H,\ i = 1, 2, 3. \tag{6.49}$$

We want to show that $\overline{B}$ is twice $H-$Fréchet differentiable, with $D^2\overline{B} = \overline{D^2B}$. By Taylor's formula applied to DB,

$$\left(D\overline{B}(w_2) - D\overline{B}(w_1) - \overline{D^2B}(w_1)(w_2 - w_1)\right) w_3 = r(w_1, w_2, w_3), \quad w_1, w_2, w_3 \in \Lambda, \text{ where} \tag{6.50}$$

$$r(x, y, z) = \left(\int_0^1 \left(\overline{D^2B}(x + h(y - x)) - \overline{D^2B}(x)\right)(y - x)\,\mathrm{d}h\right) z, \quad x,\ y,\ z \in H.$$

We note that $r\colon H \times H \times H \to H_\square$ is continuous. Indeed, for every $x, y, z \in H$ and every sequence $((x_n, y_n, z_n))_n \subset H \times H \times H$ such that $(x_n, y_n, z_n) \to (x, y, z)$ as $n \to \infty$, with some algebraic computations we obtain, by (6.49),

$$\begin{aligned}&\left\|r(x_n, y_n, z_n) - r(x, y, z)\right\|_{2,\square} \leq 2C_0 \left\|y_n - x_n\right\|_2 \left\|z_n - z\right\|_2 \\ &+ C_0 \left\|z\right\|_2 \left(2\left\|y_n - x_n + x - y\right\|_2 + \left(\frac{1}{\beta + 1}\left\|y_n - y + x - x_n\right\|_2^{\beta} + 2\left\|x_n - x\right\|_2^{\beta}\right)\left\|y - x\right\|_2\right) \underset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

It then follows from the continuity of $D\overline{B}$ in H and (6.49) that (6.50) holds for every $w_1, w_2, w_3 \in H$. Moreover, observing that, by (6.49), $\left\|r(x, y, \cdot)\right\|_{\mathcal{L}(H;H_\square)} \leq C_0(\beta + 1)^{-1}\left\|y - x\right\|_2^{1+\beta}$, $x, y \in H$, we conclude that

$$D\overline{B}(w_2) - D\overline{B}(w_1) - \overline{D^2B}(w_1)(w_2 - w_1) = \mathrm{o}\left(\left\|w_2 - w_1\right\|_2\right), \quad w_1, w_2 \in H.$$

Therefore $\overline{B}$ is twice $H-$Fréchet differentiable, with $D^2\overline{B} = \overline{D^2B}$.

We also note that, for every w_1, w_2, $w_3 \in H$ and $0 < t \leq T$,

$$D^2\overline{B}(w_1)(w_2, w_3)(r, \cdot) \in H \text{ is of Volterra–type, for a.e. } r \in (0, t), \tag{6.51}$$

and that

$$D^2\overline{B}(w_1)(w_2, w_3)(r, \cdot) \text{ depends on } w_i \text{ only via } w_i\big|_{(0,t)},\ i = 1, 2, 3, \text{ for a.e. } r \in (0, t)\text{:} \tag{6.52}$$

these properties are inherited from $D\overline{B}$ (cfr. (6.29)-(6.30) in the discussion following Hypothesis 6.2).

In conclusion, we notice that, by (6.49) (see also (6.46)),

$$\left\|D^2\overline{B}\right\|_\infty = \sup_{w\in H}\left\|D^2\overline{B}(w)\right\|_{\mathcal{L}(H,H;H_\square)} \le C_0. \tag{6.53}$$

As a consequence, by the mean value theorem we deduce that (6.26) (see also (6.27)) holds with $\gamma = 1$, i.e., under Hypothesis 6.3 the map $D\overline{B}\colon H \to \mathcal{L}(H;H_\square)$ is globally Lipschitz–continuous. Since $D\overline{B}$ is also bounded (see (6.24)-(6.27)), in what follows we suppose, without loss of generality, that

$$\text{under Hypothesis 6.3, } D\overline{B}\colon H\to\mathcal{L}(H;H_\square) \text{ satisfies (6.27) with } \gamma=\beta. \tag{6.54}$$

The next result shows that, in the framework of this subsection, the solution $w_t^{s,\phi}$ of (6.11), considered as a map from H to $\mathcal{H}$, is twice $H-$Fréchet differentiable.

Theorem 6.6. *Under Hypothesis 6.3, for every $0 \le s \le t \le T$, the mapping $w_t^{s,\cdot} \in C^{2+\beta}(H;\mathcal{H})$. In particular, for every ϕ, ψ, $\eta \in H$, $D^2w_t^{s,\phi}(\psi,\eta)$ is the unique solution in $\mathcal{H}$ of the following equation:*

$$D^2w_t^{s,\phi}(\psi,\eta) = \int_s^t \left(D^2\overline{B}\left(w_t^{s,\phi}\right)\left(Dw_t^{s,\phi}\psi, Dw_t^{s,\phi}\eta\right) + D\overline{B}\left(w_t^{s,\phi}\right)D^2w_t^{s,\phi}(\psi,\eta)\right)(r,\cdot)\,\mathrm{d}r. \tag{6.55}$$

Furthermore, there exists a constant $C_3 = C_3(d,T) > 0$ such that, for every $\phi,\psi,\eta,\theta \in H$, $\mathbb{P}-$a.s.,

$$\left\|D^2w_t^{s,\phi}(\eta,\theta)\right\|_2 \le C_3\|\eta\|_2\|\theta\|_2\,, \quad \left\|\left(D^2w_t^{s,\phi} - D^2w_t^{s,\psi}\right)(\eta,\theta)\right\|_2 \le C_3\left\|w_t^{s,\phi}-w_t^{s,\psi}\right\|_2^\beta \|\eta\|_2\|\theta\|_2\,. \tag{6.56}$$

Proof. Fix $0\le s\le t\le T$ and $\phi\in H$. We first want to prove the well–posedness in $\mathcal{H}$ of the equation

$$w = \int_s^t \left(D^2\overline{B}\left(w_t^{s,\phi}\right)\left(Dw_t^{s,\phi}\psi, Dw_t^{s,\phi}\eta\right) + D\overline{B}\left(w_t^{s,\phi}\right)w\right)(r,\cdot)\,\mathrm{d}r, \quad \psi,\eta\in H. \tag{6.57}$$

Consider $N = N(d,T) \in \mathbb{N}$ so big that $C_0\sqrt{T/N} < 1$, where $C_0 = C_0(d,T)$ is the constant in Hypotheses 6.1-6.2-6.3. In addition, take an equispaced partition $\{t_k\}_{k=0}^N$ of $[s,t]$ where $t_0 = s$ and $t_N = t$: its mesh $\Delta \le T/N$. Under Hypothesis 6.3, the bound in (6.34) holds and allows to employ a fixed point argument as in the proof of Theorem 6.3 (see also Theorem 6.5) to deduce the existence of a unique solution $\overline{w}_1^{\psi,\eta}\in\mathcal{H}$ of (6.57) with t_1 instead of t, for every $\psi,\eta\in H$.

We claim that the operator $D^2w_{t_1}^{s,\phi}\colon H\times H\to\mathcal{H}$ defined by $D^2w_{t_1}^{s,\phi}(\psi,\eta) = \overline{w}_1^{\psi,\eta}$, $\psi,\eta\in H$, is the second–order Fréchet differential of $w_{t_1}^{s,\phi}$. Indeed, considering that $Dw_{t_1}^{s,\phi}\in\mathcal{L}(H;\mathcal{H})$, $D\overline{B}(w_{t_1}^{s,\phi})\in\mathcal{L}(H;H_\square)$ and $D^2\overline{B}(w_{t_1}^{s,\phi})\in\mathcal{L}(H,H;H_\square)$, the fact that $D^2w_{t_1}^{s,\phi}$ is bilinear directly follows from (6.57). As for the boundedness, by (6.27)-(6.49) (see also (6.24)-(6.46)) and (6.32) we can compute, applying Bochner's theorem to (6.55), for some constant $C_2 = C_2(d,T) > 0$,

$$\begin{aligned}\left\|D^2w_{t_1}^{s,\phi}(\psi,\eta)\right\|_2 &\le C_0\sqrt{\Delta}\left(\left\|Dw_{t_1}^{s,\phi}\psi\right\|_2\left\|Dw_{t_1}^{s,\phi}\eta\right\|_2 + \left\|D^2w_{t_1}^{s,\phi}(\psi,\eta)\right\|_2\right)\\ &\le C_0\sqrt{T/N}\left(C_2\|\psi\|_2\|\eta\|_2 + \left\|D^2w_{t_1}^{s,\phi}(\psi,\eta)\right\|_2\right), \quad \mathbb{P}-\text{a.s.}, \psi,\eta\in H.\end{aligned} \tag{6.58}$$

Hence

$$\left\|D^2w_{t_1}^{s,\phi}(\psi,\eta)\right\|_2 \le \left(1 - C_0\sqrt{T/N}\right)^{-1} C_0C_2\sqrt{T/N}\,\|\psi\|_2\|\eta\|_2\,, \quad \mathbb{P}-\text{a.s.}, \psi,\eta\in H. \tag{6.59}$$

We now observe that, by Taylor's formula applied to $D\overline{B}$ (cfr. (6.50)), from (6.31)-(6.55) we have, for every $h \in H$,

$$\left\| Dw_{t_1}^{s,\phi+h} - Dw_{t_1}^{s,\phi} - D^2 w_{t_1}^{s,\phi} h \right\|_{\mathcal{L}(H;\mathcal{H})} \leq \mathbf{I}_1 + \mathbf{II}_1 + \mathbf{III}_1 + \mathbf{IV}_1, \tag{6.60}$$

where we set

$$\mathbf{I}_1 = \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| \int_s^{t_1} D\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(Dw_{t_1}^{s,\phi+h}\eta - Dw_{t_1}^{s,\phi}\eta - D^2 w_{t_1}^{s,\phi}(h,\eta)\right)(r,\cdot)\,\mathrm{d}r \right\|_2^2 \right]^{\frac{1}{2}},$$

$$\mathbf{II}_1 = \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| \int_s^{t_1} \left(D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} - Dw_{t_1}^{s,\phi}h, Dw_{t_1}^{s,\phi}\eta\right)\right)(r,\cdot)\,\mathrm{d}r \right\|_2^2 \right]^{\frac{1}{2}},$$

$$\mathbf{III}_1 = \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| \int_s^{t_1} \left(D\overline{B}\left(w_{t_1}^{s,\phi+h}\right) - D\overline{B}\left(w_{t_1}^{s,\phi}\right)\right)\left(Dw_{t_1}^{s,\phi+h}\eta - Dw_{t_1}^{s,\phi}\eta\right)(r,\cdot)\,\mathrm{d}r \right\|_2^2 \right]^{\frac{1}{2}},$$

$$\mathbf{IV}_1 = \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\Bigg[\Bigg\| \int_s^{t_1} \Bigg(\int_0^1 \Bigg(D^2\overline{B}\left(w_{t_1}^{s,\phi} + v\left(w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi}\right)\right)$$
$$- D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\Bigg)\left(w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi}\right)\mathrm{d}v\Bigg) Dw_{t_1}^{s,\phi}\eta\;(r,\cdot)\,\mathrm{d}r \Bigg\|_2^2 \Bigg]^{\frac{1}{2}}.$$

By (6.27) (see, in particular, (6.24))

$$|\mathbf{I}_1| \leq C_0\sqrt{T/N} \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| Dw_{t_1}^{s,\phi+h}\eta - Dw_{t_1}^{s,\phi}\eta - D^2 w_{t_1}^{s,\phi}(h,\eta) \right\|_2^2 \right]^{\frac{1}{2}}$$
$$= C_0\sqrt{T/N} \left\| Dw_{t_1}^{s,\phi+h} - Dw_{t_1}^{s,\phi} - D^2 w_{t_1}^{s,\phi} h \right\|_{\mathcal{L}(H;\mathcal{H})}.$$

Moreover, considering (6.26)-(6.32) (see also (6.54)) and Corollary 6.4, which we can apply with $q = 2(1+\beta)$ because $\phi, h \in H \subset \mathcal{H}^q$ (see also), for some $C_1 = C_1(\beta, d, T) > 0$ we can write

$$|\mathbf{III}_1| \leq \sqrt{\Delta} \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| \left(D\overline{B}\left(w_{t_1}^{s,\phi+h}\right) - D\overline{B}\left(w_{t_1}^{s,\phi}\right)\right)\left(Dw_{t_1}^{s,\phi+h}\eta - Dw_{t_1}^{s,\phi}\eta\right) \right\|_{2,\square}^2 \right]^{\frac{1}{2}}$$
$$\leq \left\| D^2\overline{B} \right\|_\infty C_2\sqrt{T/N} \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} \right\|_2^{2(1+\beta)} \|\eta\|_2^2 \right]^{\frac{1}{2}} \leq C_0 C_1^{1+\beta} C_2 \sqrt{T/N}\, \|h\|_2^{1+\beta},$$

where we also use the mean value theorem on $D\overline{B}$ and (6.53). As for $\mathbf{II}_1$, by (6.32)-(6.49) we compute

$$|\mathbf{II}_1| \leq \sqrt{\Delta} \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} - Dw_{t_1}^{s,\phi}h, Dw_{t_1}^{s,\phi}\eta\right) \right\|_{2,\square}^2 \right]^{\frac{1}{2}}$$
$$\leq C_0 C_2 \sqrt{\Delta} \sup_{\|\eta\|_2 \leq 1} \mathbb{E}\left[\left\| w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} - Dw_{t_1}^{s,\phi}h \right\|_2^2 \|\eta\|_2^2 \right]^{\frac{1}{2}}$$
$$\leq C_0 C_2 \sqrt{T/N} \left\| w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi} - Dw_{t_1}^{s,\phi}h \right\|_{\mathcal{H}} = \mathrm{o}\left(\|h\|_2\right).$$

Finally, again by (6.32)-(6.49) (see also (6.48)) and Corollary 6.4, employed with $q = 2(1+\beta)$, we have

$$|\mathbf{IV}_1| \leq \sqrt{\Delta}\mathbb{E}\left[\left(\int_0^1 \left\|D^2\overline{B}\left(w_{t_1}^{s,\phi} + v\left(w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi}\right)\right) - D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\right\|_{\mathcal{L}(H,H;H_\square)} \mathrm{d}v\right)^2 \times \left\|w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi}\right\|_2^2 \left\|Dw_{t_1}^{s,\phi}\eta\right\|_2^2\right]^{\frac{1}{2}}$$

$$\leq C_0C_2\sqrt{T/N} \sup_{\|\eta\|_2\leq 1} \mathbb{E}\left[\left\|w_{t_1}^{s,\phi+h} - w_{t_1}^{s,\phi}\right\|_2^{2(1+\beta)} \|\eta\|_2^2\right]^{\frac{1}{2}} \leq C_0C_1^{1+\beta}C_2\sqrt{T/N}\,\|h\|_2^{1+\beta}.$$

Going back to (6.60), we conclude that

$$\left\|Dw_{t_1}^{s,\phi+h} - Dw_{t_1}^{s,\phi} - D^2w_{t_1}^{s,\phi}h\right\|_{\mathcal{L}(H;\mathcal{H})} \leq \left(1 - C_0\sqrt{T/N}\right)^{-1}(\mathbf{II}_1 + \mathbf{III}_1 + \mathbf{IV}_1) = \mathrm{o}\left(\|h\|_2\right), \quad h \in H. \tag{6.61}$$

This shows that $D^2w_{t_1}^{s,\phi}$ is the second–order Fréchet differential of $w_{t_1}^{s,\phi}$, as desired.

Next, consider

$$w = D^2w_{t_1}^{s,\phi}(\psi,\eta) + \int_{t_1}^{t_2}\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)w\right)(r,\cdot)\,\mathrm{d}r, \quad \psi,\eta\in H. \tag{6.62}$$

Arguing as in the previous step, we infer the well–posedness of this equation in $\mathcal{H}$: we denote by $\overline{w}_2^{\psi,\eta} \in \mathcal{H}$ its unique solution, for every ψ, $\eta \in H$.

Given ψ, $\eta \in H$, we now show that $\overline{w}_2^{\psi,\eta}$ is the unique solution of (6.57) with t_2 instead of t. By the Volterra–type property of $D^2\overline{B}$ [resp., $D\overline{B}$] in (6.51) [resp., (6.29)] and (6.62) we have, $\mathbb{P}-$a.s.,

$$\overline{w}_2^{\psi,\eta}\Big|_{(0,t_1)} = D^2w_{t_1}^{s,\phi}(\psi,\eta)\Big|_{(0,t_1)}.$$

Moreover, since $Dw_{t_2}^{s,\phi}\psi$ satisfies (6.38), we infer that, $\mathbb{P}-$a.s.,

$$Dw_{t_2}^{s,\phi}\psi\Big|_{(0,t_1)} = Dw_{t_1}^{s,\phi}\psi\Big|_{(0,t_1)},$$

with an analogous result holding for η. Consequently, recalling also (6.39) and Remark 6.1, by the property of $D^2\overline{B}$ [resp., $D\overline{B}$] in (6.52) [resp., (6.30)], from (6.62) we obtain, $\mathbb{P}-$a.s.,

$$\begin{aligned}\overline{w}_2^{\psi,\eta} &= \int_s^{t_1}\left(D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(Dw_{t_1}^{s,\phi}\psi, Dw_{t_1}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_1}^{s,\phi}\right)D^2w_{t_1}^{s,\phi}(\psi,\eta)\right)(r,\cdot)\,\mathrm{d}r \\ &\quad + \int_{t_1}^{t_2}\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)\overline{w}_2^{\psi,\eta}\right)(r,\cdot)\,\mathrm{d}r \\ &= \int_s^{t_2}\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)\overline{w}_2^{\psi,\eta}\right)(r,\cdot)\,\mathrm{d}r,\end{aligned} \tag{6.63}$$

where we also use the fact that $D^2w_{t_1}^{s,\phi}(\psi,\eta)$ solves (6.57) with t_1 instead of t. Hence $\overline{w}_2^{\psi,\eta}$ solves (6.57) with t replaced by t_2. In order to prove that it is in fact the unique solution of this equation, we consider another random variable $\widetilde{w} \in \mathcal{H}$ satisfying (6.63). Then, by (6.51)-(6.52),

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\left(\int_s^{t_1}\left(D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(Dw_{t_1}^{s,\phi}\psi, Dw_{t_1}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_1}^{s,\phi}\right)1_{(0,t_1)}\widetilde{w}\right)(r,\cdot)\,\mathrm{d}r\right). \tag{6.64}$$

We observe that also $1_{(0,t_1)}\overline{w}_1^{\psi,\eta} \in \mathcal{H}$ satisfies (6.64). Therefore we can perform the same computations as in (6.42) to deduce that

$$1_{(0,t_1)}\widetilde{w} = 1_{(0,t_1)}\overline{w}_1^{\psi,\eta}, \quad \mathbb{P}-\text{a.s.}$$

Going back to (6.63), by the previous equality we have, $\mathbb{P}$–a.s.,

$$\begin{aligned}\widetilde{w} = \int_s^{t_1} &\left(D^2\overline{B}\left(w_{t_1}^{s,\phi}\right)\left(Dw_{t_1}^{s,\phi}\psi, Dw_{t_1}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_1}^{s,\phi}\right)\widetilde{w}\right)(r,\cdot)\,\mathrm{d}r \\ &+ \int_{t_1}^{t_2}\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)\widetilde{w}\right)(r,\cdot)\,\mathrm{d}r \\ = \overline{w}_1^{\psi,\eta} + \int_{t_1}^{t_2} &\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)\widetilde{w}\right)(r,\cdot)\,\mathrm{d}r.\end{aligned}$$

It follows that $\widetilde{w}$ satisfies (6.62): by uniqueness, we obtain $\widetilde{w} = \overline{w}_2^{\psi,\eta}$. Hence $\overline{w}_2^{\psi,\eta}$ is the unique solution of (6.57) in $\mathcal{H}$ with t_2 instead of t.

We define the operator $D^2w_{t_2}^{s,\phi}\colon H \times H \to \mathcal{H}$ by $D^2w_{t_2}^{s,\phi}(\psi,\eta) = \overline{w}_2^{\psi,\eta}$, $\psi,\eta \in H$, and claim that it is the second–order Fréchet differential of $w_{t_2}^{s,\phi}$. Indeed, as we have argued for $D^2w_{t_1}^{s,\phi}$, the map $D^2w_{t_2}^{s,\phi}$ is bilinear thanks to the the well–posedness of (6.63). As for the boundedness, arguing as in (6.58), by (6.59)-(6.62) we can write, for every ψ, $\eta \in H$, $\mathbb{P}-$a.s.,

$$\begin{aligned}\left\|Dw_{t_2}^{s,\phi}(\psi,\eta)\right\|_2 \le{}& \left\|D^2w_{t_1}^{s,\phi}(\psi,\eta)\right\|_2 \\ &+ \int_{t_1}^{t_2}\left\|\left(D^2\overline{B}\left(w_{t_2}^{s,\phi}\right)\left(Dw_{t_2}^{s,\phi}\psi, Dw_{t_2}^{s,\phi}\eta\right) + D\overline{B}\left(w_{t_2}^{s,\phi}\right)Dw_{t_2}^{s,\phi}(\psi,\eta)\right)(r,\cdot)\right\|_2\mathrm{d}r \\ \le{}& C_0C_2\left(\left(1 - C_0\sqrt{T/N}\right)^{-1}\sqrt{T/N} + \sqrt{\Delta}\right)\|\psi\|_2\|\eta\|_2 + \sqrt{\Delta}C_0\left\|D^2w_{t_2}^{s,\phi}(\psi,\eta)\right\|_2,\end{aligned}$$

whence

$$\left\|Dw_{t_2}^{s,\phi}(\psi,\eta)\right\|_2 \le 2C_0C_2\left(1 - C_0\sqrt{T/N}\right)^{-2}\sqrt{T/N}\,\|\psi\|_2\|\eta\|_2, \quad \mathbb{P}-\text{a.s.}, \psi, \eta \in H. \tag{6.65}$$

Moreover, combining (6.38) with (6.62), we can argue as in (6.60) to infer that

$$\left\|Dw_{t_2}^{s,\phi+h} - Dw_{t_2}^{s,\phi} - D^2w_{t_2}^{s,\phi}h\right\|_{\mathcal{L}(H;\mathcal{H})} = \mathrm{o}\left(\|h\|_2\right), \quad h \in H,$$

which shows that $D^2w_{t_2}^{s,\phi}$ is the second–order Fréchet differential of $w_{t_2}^{s,\phi}$, as desired.

This reasoning can be repeated $N-$times to deduce that the operator $D^2w_t^{s,\phi}\colon H \times H \to \mathcal{H}$ defined by $D^2w_t^{s,\phi}(\psi,\eta) = \overline{w}_N^{\psi,\eta}$, where $\overline{w}_N^{\psi,\eta}$ is the unique solution of (6.57) in $\mathcal{H}$, for every ψ, $\eta \in H$, is the second–order Fréchet differential of $w_t^{s,\phi}$. In particular, the first bound in (6.56) is true, because (cfr. (6.59)-(6.65))

$$\begin{aligned}&\left\|D^2w_t^{s,\phi}(\psi,\eta)\right\|_2 \le NC_0C_2\left(1 - C_0\sqrt{T/N}\right)^{-N}\sqrt{T/N}\,\|\psi\|_2\|\eta\|_2 =: \widetilde{C}\,\|\psi\|_2\|\eta\|_2, \\ &\mathbb{P}-\text{a.s.}, \phi,\psi,\eta \in H.\end{aligned} \tag{6.66}$$

As for the second inequality in (6.56), by (6.27), (6.32), (6.49), (6.54), (6.55) and (6.66) we compute, for every $\phi, \psi, \eta, \theta \in H$, $\mathbb{P}-$a.s.,

$$\begin{aligned}
&\left\|D^2 w_{t_1}^{s,\phi}(\eta,\theta) - D^2 w_{t_1}^{s,\psi}(\eta,\theta)\right\|_2 \\
&= \left\| \int_s^{t_1} \left(D^2 \overline{B}\left(w_t^{s,\phi}\right)\left(Dw_t^{s,\phi}\eta, Dw_t^{s,\phi}\theta\right) - D^2 \overline{B}\left(w_t^{s,\psi}\right)\left(Dw_t^{s,\psi}\eta, Dw_t^{s,\psi}\theta\right) \right.\right. \\
&\qquad \left.\left. + D\overline{B}\left(w_t^{s,\phi}\right) D^2 w_{t_1}^{s,\phi}(\eta,\theta) - D\overline{B}\left(w_t^{s,\psi}\right) D^2 w_{t_1}^{s,\psi}(\eta,\theta) \right)(r,\cdot)\,\mathrm{d}r \right\|_2 \\
&\leq \sqrt{\Delta}\Big(\left\| \left(D^2\overline{B}\left(w_t^{s,\phi}\right) - D^2\overline{B}\left(w_t^{s,\psi}\right)\right)\left(Dw_t^{s,\phi}\eta, Dw_t^{s,\phi}\theta\right)\right\|_{2,\square} \\
&\quad + \left\|D^2\overline{B}\left(w_t^{s,\psi}\right)\left(\left(Dw_t^{s,\phi} - Dw_t^{s,\psi}\right)\eta, Dw_t^{s,\phi}\theta\right)\right\|_{2,\square} + \left\|D^2\overline{B}\left(w_t^{s,\psi}\right)\left(Dw_t^{s,\psi}\eta, \left(Dw_t^{s,\phi} - Dw_t^{s,\psi}\right)\theta\right)\right\|_{2,\square} \\
&\quad + \left\|\left(D\overline{B}\left(w_t^{s,\phi}\right) - D\overline{B}\left(w_t^{s,\psi}\right)\right) D^2 w_{t_1}^{s,\phi}(\eta,\theta)\right\|_{2,\square} + \left\|D\overline{B}\left(w_t^{s,\psi}\right)\left(D^2 w_{t_1}^{s,\phi}(\eta,\theta) - D^2 w_{t_1}^{s,\psi}(\eta,\theta)\right)\right\|_{2,\square}\Big) \\
&\leq C_0\sqrt{T/N}\Big(\left(\widetilde{C} + 3C_2\right)\left\|w_t^{s,\phi} - w_t^{s,\psi}\right\|_2^\beta \|\eta\|_2 \|\theta\|_2 + \left\|D^2 w_{t_1}^{s,\phi}(\eta,\theta) - D^2 w_{t_1}^{s,\psi}(\eta,\theta)\right\|_2 \Big),
\end{aligned}$$

whence

$$\left\|D^2 w_{t_1}^{s,\phi}(\eta,\theta) - D^2 w_{t_1}^{s,\psi}(\eta,\theta)\right\|_2 \leq \left(1 - C_0\sqrt{T/N}\right)^{-1} C_0\left(\widetilde{C} + 3C_2\right)\sqrt{T/N}\left\|w_t^{s,\phi} - w_t^{s,\psi}\right\|_2^\beta \|\eta\|_2\|\theta\|_2 .$$

By (6.62), we sequentially iterate this computation to obtain the second inequality in (6.56) with

$$C_3 = \max\{\widetilde{C}, N\left(1 - C_0\sqrt{T/N}\right)^{-N} C_0(\widetilde{C} + 3C_2)\sqrt{T/N}\}.$$

Thus, taking expectations and using Corollary 6.4 with $q = 2$, by Jensen's inequality we deduce that, for some constant $c > 0$,

$$\begin{aligned}
\left\|D^2 w_t^{s,\phi} - D^2 w_t^{s,\psi}\right\|_{\mathcal{L}(H,H;\mathcal{H})} &= \sup_{\|\eta\|_2, \|\theta\|_2 \leq 1} \mathbb{E}\left[\left\|D^2 w_t^{s,\phi}(\eta,\theta) - D^2 w_t^{s,\psi}(\eta,\theta)\right\|_2^2\right]^{\frac{1}{2}} \\
&\leq C_3 \mathbb{E}\left[\left\|w_t^{s,\phi} - w_t^{s,\psi}\right\|_2^{2\beta}\right]^{\frac{1}{2}} \leq c\,\|\phi - \psi\|_2^\beta, \quad \phi, \psi \in H.
\end{aligned}$$

This shows that $D^2 w_t^{s,\cdot} \in C^\beta(H; \mathcal{L}(H,H;\mathcal{H}))$, completing the proof. ■

6.3 The Kolmogorov equation

Recall the definition of the map $\sigma\colon [0,T] \to \mathcal{L}(\mathbb{R}^d; H)$ in (6.2). Given $u\colon [0,T] \times H \to \mathbb{R}$ and a terminal condition $\Phi\colon H \to \mathbb{R}$, in this section we investigate the following *Kolmogorov backward equation* in integral form:

$$\begin{aligned}
u(t,\phi) = \Phi(\phi) &+ \int_t^T \langle \nabla u(r,\phi), B(r,\phi)\rangle_H\,\mathrm{d}r + \frac{1}{2}\int_t^T \mathrm{Tr}\left(D^2 u(r,\phi)\,\sigma(r)\,\sigma(r)^*\right)\mathrm{d}r, \\
&t \in [0,T],\ \phi \in \Lambda.
\end{aligned} \tag{6.67}$$

Our aim is to find a solution of (6.67) via the random variables $w_T^{t,\phi} \in \mathcal{H}$ satisfying (6.11) for every $t \in [0,T]$ and $\phi \in H$. This is done in Theorem 6.9, for which we need a couple of preparatory results.

Lemma 6.7. *There exists a constant* $C_{\alpha,d} > 0$ *such that*

$$\left\| \int_s^t (\sigma(t) - \sigma(r))\,\mathrm{d}W_r \right\|_{\mathcal{H}} \le C_{\alpha,d} |t-s|^{\alpha}, \quad 0 \le s \le t \le T. \tag{6.68}$$

Proof. Fix $0 \le s \le t \le T$ and denote by $(e_k)_{k=1,\ldots,d}$ the canonical basis of $\mathbb{R}^d$. Using straightforward substitutions, by (6.2) we compute, for every $k = 1, \ldots, d$,

$$\begin{aligned} \|(\sigma(t) - \sigma(r))\, e_k\|_2^2 &= \int_0^T \left| k_2(\xi - t)\, 1_{\{\xi > t\}} - k_2(\xi - r)\, 1_{\{\xi > r\}} \right|^2 \mathrm{d}\xi \\ &= \int_0^{t-r} |k_2(\xi)|^2 \,\mathrm{d}\xi + \int_0^{T-t} |k_2(\xi + t - r) - k_2(\xi)|^2 \,\mathrm{d}\xi, \quad r \in [s,t]. \end{aligned} \tag{6.69}$$

Recalling that (see (6.1)) $k_2(u) = \frac{1}{\Gamma(\alpha)} u^{\alpha-1}$, $\alpha \in (1/2, 1)$, $u > 0$, for every $r \in [s,t]$ we have

$$\int_0^{t-r} |k_2(\xi)|^2 \,\mathrm{d}\xi = \frac{1}{(\Gamma(\alpha))^2 (2\alpha - 1)} |t-r|^{2\alpha-1},$$

and

$$\int_0^{T-t} |k_2(\xi + t - r) - k_2(\xi)|^2 \,\mathrm{d}\xi \le \frac{1}{(\Gamma(\alpha))^2} \left(\int_0^{\infty} \left((\xi+1)^{\alpha-1} - \xi^{\alpha-1} \right)^2 \mathrm{d}\xi \right) |t-r|^{2\alpha-1}.$$

Therefore the discussion at the end of Page 98 in [66] ensures that (6.68) holds with

$$C_{\alpha,d} = \frac{\sqrt{d}}{\Gamma(\alpha)} \left(\frac{1}{2\alpha} \right)^{\frac{1}{2}} \left(\frac{1}{2\alpha - 1} + \int_0^{\infty} \left((\xi+1)^{\alpha-1} - \xi^{\alpha-1} \right)^2 \mathrm{d}\xi \right)^{\frac{1}{2}},$$

completing the proof. ■

The following lemma analyzes some properties of the solution $w_t^{s,\phi} \in \mathcal{L}_t^p$ of (6.11) in the framework of Remark 6.2. Recall that $\mathcal{L}_t^p = L_t^p(\Omega; L^p)$, where $L^p = L^p\big(0,T;\mathbb{R}^d\big)$, and that $L_\square^p = L^p\big((0,T) \times (0,T); \mathbb{R}^d\big)$.

Lemma 6.8. *Suppose that* $B\colon \Lambda \to L_\square^p$ *satisfies Hypothesis 6.1 and* (6.21)*, for some* $p \in \big[2, (1-\alpha)^{-1}\big)$*. Then there exists a constant* $C_{1,p} = C_{1,p}(\alpha, d, T) > 0$ *such that*

$$\left\| w_t^{s,\phi} \right\|_{\mathcal{L}^p} \le C_{1,p} \left(1 + \|\phi\|_p \right), \quad 0 \le s \le t \le T,\ \phi \in L^p. \tag{6.70}$$

Furthermore, for every $\phi \in L^p$*, there is a constant* $C_{\phi,p} = C_{\phi,p}(\alpha, d, T) > 0$ *such that*

$$\left\| w_t^{s,\phi} - \phi \right\|_{\mathcal{L}^p} \le C_{\phi,p} \sqrt{t-s}, \quad 0 \le s \le t \le T. \tag{6.71}$$

When $p = 2$, the assumptions of Lemma 6.8 reduce to Hypothesis 6.1 and $\|\cdot\|_{\mathcal{L}^p} = \|\cdot\|_{\mathcal{H}}$.

Proof. Fix $0 \le s \le t \le T$ and $\phi \in L^p$. Recall that, under the hypotheses of the lemma, the unique solution $w_t^{s,\phi} \in \mathcal{H}$ of (6.11) belongs to the space $\mathcal{L}_t^p$, see Remark 6.2.
Consider $N = N(d,p,T) \in \mathbb{N}$ so big that $C_{0,p}(2T/N)^{1-\frac{1}{p}} < 1$, where $C_{0,p}$ is the constant appearing in

(6.21). Take an equispaced partition $\{t_k\}_{k=0}^N$ of $[s,t]$ with $t_0 = s$ and $t_N = t$: its mesh $\Delta \le T/N$. By (6.11)-(6.21) we have, using Bochner's theorem and Jensen's inequality,

$$\left\| w_{t_1}^{s,\phi} \right\|_{\mathcal{L}^p} \le \|\phi\|_p + C_{0,p}(2\Delta)^{1-\frac{1}{p}} \left(1 + \left\| w_{t_1}^{s,\phi} \right\|_{\mathcal{L}^p}\right) + \left\| \int_s^{t_1} \sigma(r)\,\mathrm{d}W_r \right\|_{\mathcal{L}^p},$$

which in turn implies, by (6.20), for some constant $c = c(d,p,T) > 0$,

$$\left\| w_{t_1}^{s,\phi} \right\|_{\mathcal{L}^p} \le \left(1 - C_{0,p}\,(2T/N)^{1-\frac{1}{p}}\right)^{-1} \left(\|\phi\|_p + c\,\|k_2\|_p + C_{0,p}\,(2T/N)^{1-\frac{1}{p}}\right).$$

At this point, invoking $N-$times the cocycle property in (6.12) we obtain (6.70).

As for (6.71), using (6.20)-(6.21) we compute, for some constant $C = C(d,p) > 0$, recalling the notation $\Sigma_{s,t}$ introduced in (6.3),

$$\begin{aligned}
\left\| w_t^{s,\phi} - \phi \right\|_{\mathcal{L}^p} &\le \mathbb{E}\left[\left(\int_s^t \left\| \overline{B}\left(r, w_t^{s,\phi}\right)\right\|_p \mathrm{d}r\right)^p\right]^{\frac{1}{p}} + \|\Sigma_{s,t}\|_{\mathcal{L}^p} \\
&\le (t-s)^{1-\frac{1}{p}}\,\mathbb{E}\left[\int_s^t \mathrm{d}r \int_0^T \left|\overline{B}\left(w_t^{s,\phi}\right)\right|^p (r,\xi)\,\mathrm{d}\xi\right]^{\frac{1}{p}} + C\,\|k_2\|_p \sqrt{t-s} \\
&\le \sqrt{t-s}\left(C\,\|k_2\|_p + 2^{1-\frac{1}{p}} T^{\frac{1}{2}-\frac{1}{p}} C_{0,p}\left(1 + \left\| w_t^{s,\phi}\right\|_{\mathcal{L}^p}\right)\right).
\end{aligned}$$

Thus, by (6.70) the proof is complete. ■

We are now ready to prove the main result of the chapter, which shows the connection between the solution $w_T^{t,\phi}$, $t \in [0,T]$, $\phi \in H$, of (6.11) and the backward Kolmogorov equation in integral form (6.67).

Theorem 6.9. *Suppose that $B\colon \Lambda \to L^p_\square$ satisfies Hypothesis 6.3 and* (6.21)*, for some $p \in \left(2, (1-\alpha)^{-1}\right)$. In addition, let the function $r \mapsto B(r,\phi)$ belong to $C\big([0,T];H\big)$, for every $\phi \in \Lambda$. Fix $\Phi \in C_b^{2+\beta}(H)$ and define the map $u\colon\ [0,T] \times H \to \mathbb{R}$ by*

$$u(t,\phi) = \mathbb{E}\left[\Phi\left(w_T^{t,\phi}\right)\right], \quad t \in [0,T],\ \phi \in H, \tag{6.72}$$

where $w_T^{t,\phi} \in \mathcal{H}$ is the unique solution of (6.11)*. Then $u \in L^\infty\big(0,T; C_b^{2+\beta}(H)\big) \cap C([0,T] \times H; \mathbb{R})$ and solves the Kolmogorov backward equation in integral form* (6.67).

Proof. The fact that the function u defined in (6.72) belongs to $L^\infty\big(0,T; C_b^{2+\beta}(H)\big) \cap C([0,T] \times H; \mathbb{R})$ is one of the results contained in Lemma 6.11 (see Appendix 6.A). Consequently, here we only focus on proving that u solves (6.67).

Fix $0 \le s < t \le T$ and $\phi \in \Lambda$. Since $\Lambda \subset \mathcal{H}_s^q$, $q \ge 2$, we can use (6.23) in Corollary 6.4 to write

$$u(s,\phi) = \mathbb{E}\left[\mathbb{E}\left[\Phi\left(w_T^{s,\phi}\right)\middle|\mathcal{F}_t\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\Phi\left(w_T^{t,\psi}\right)\right]\Big|_{\psi = w_t^{s,\phi}}\right] = \mathbb{E}\left[u\left(t, w_t^{s,\phi}\right)\right]. \tag{6.73}$$

Taylor's formula applied to the mapping $u(t,\cdot) \in C_b^{2+\beta}(H)$ yields, denoting by $h = w_t^{s,\phi} - \phi \in \mathcal{H}$,

$$u\left(t, w_t^{s,\phi}\right) - u(t,\phi) = \langle \nabla u(t,\phi), h\rangle_H + \frac{1}{2}\left\langle D^2 u(t,\phi)\,h, h\right\rangle_H + r_{u(t,\cdot)}\left(\phi, w_t^{s,\phi}\right), \quad \text{where}$$

$$r_{u(t,\cdot)}(x,y) = \int_0^1 (1-r)\left\langle \left(D^2 u(t, x + r(y-x)) - D^2 u(t,x)\right)(y-x), y-x\right\rangle_H \mathrm{d}r, \quad x,\, y \in H. \tag{6.74}$$

To keep the notation simple, in this proof we denote by $\overline{B}_{s,t}(w_t^{s,\phi}) = \int_s^t \overline{B}(r, w_t^{s,\phi})\,\mathrm{d}r \in \mathcal{H}$. Using the expression in (6.11) for $h = w_t^{s,\phi} - \phi$ and noticing that $\mathbb{E}[\Sigma_{s,t}] = 0 \in H$ by [66, Proposition 4.28], we take expectations in the previous chain of equalities to obtain, from (6.73),

$$\begin{aligned} u(s,\phi) - u(t,\phi) &= \left\langle \nabla u(t,\phi), \mathbb{E}\left[\int_s^t \overline{B}\left(r, w_t^{s,\phi}\right)\mathrm{d}r\right]\right\rangle_H \\ &\quad + \frac{1}{2}\mathbb{E}\left[\left\langle D^2 u(t,\phi)\left(\overline{B}_{s,t}\left(w_t^{s,\phi}\right) + \Sigma_{s,t}\right), \overline{B}_{s,t}\left(w_t^{s,\phi}\right) + \Sigma_{s,t}\right\rangle_H\right] + \mathbb{E}\left[r_{u(t,\cdot)}\left(\phi, w_t^{s,\phi}\right)\right]. \end{aligned} \tag{6.75}$$

For all $N \in \mathbb{N}$, consider an equispaced partition $\{t_k^{(N)}\}_{k=0}^N$ of $[s,T]$ with mesh Δ_N, where $t_0^{(N)} = s$ and $t_N^{(N)} = T$. By (6.75), we have

$$\begin{aligned} u(s,\phi) - \Phi(\phi) &= \sum_{k=1}^N \left(u\left(t_{k-1}^{(N)}, \phi\right) - u\left(t_k^{(N)}, \phi\right)\right) \\ &= \sum_{k=1}^N \left\langle \nabla u\left(t_k^{(N)}, \phi\right), \mathbb{E}\left[\overline{B}_{t_{k-1}^{(N)}, t_k^{(N)}}\left(w_{t_k^{(N)}}^{t_{k-1}^{(N)},\phi}\right)\right]\right\rangle_H \\ &\quad + \frac{1}{2}\sum_{k=1}^N \mathbb{E}\left[\left\langle D^2 u\left(t_k^{(N)}, \phi\right)\left(\overline{B}_{t_{k-1}^{(N)}, t_k^{(N)}}\left(w_{t_k^{(N)}}^{t_{k-1}^{(N)},\phi}\right) + \Sigma_{t_{k-1}^{(N)}, t_k^{(N)}}\right), \overline{B}_{t_{k-1}^{(N)}, t_k^{(N)}}\left(w_{t_k^{(N)}}^{t_{k-1}^{(N)},\phi}\right) + \Sigma_{t_{k-1}^{(N)}, t_k^{(N)}}\right\rangle_H\right] \\ &\quad + \sum_{k=1}^N \mathbb{E}\left[r_{u\left(t_k^{(N)},\cdot\right)}\left(\phi, w_{t_k^{(N)}}^{t_{k-1}^{(N)},\phi}\right)\right] =: \mathbf{I}^N + \mathbf{II}^N + \mathbf{III}^N. \end{aligned} \tag{6.76}$$

In the sequel, we omit the superscript N from the points of the partition to ease notation, i.e., we write t_k for $t_k^{(N)}$. Firstly, we analyze $\mathbf{I}^N$, which we decompose using the properties of the Bochner's integral as follows:

$$\begin{aligned} \mathbf{I}^N &= \sum_{k=1}^N \langle \nabla u(t_k,\phi), B(t_k,\phi)\rangle_H (t_k - t_{k-1}) \\ &\quad + \sum_{k=1}^N \mathbb{E}\left[\int_{t_{k-1}}^{t_k} \left\langle \nabla u(t_k,\phi), \overline{B}\left(r, w_{t_k}^{t_{k-1},\phi}\right) - B(r,\phi)\right\rangle_H \mathrm{d}r\right] \\ &\quad + \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \langle \nabla u(t_k,\phi), B(r,\phi) - B(t_k,\phi)\rangle_H \,\mathrm{d}r =: \mathbf{I}_1^N + \mathbf{I}_2^N + \mathbf{I}_3^N. \end{aligned}$$

Note that $\mathbf{I}_1^N \to \int_s^T \langle \nabla u(r,\phi), B(r,\phi)\rangle_H \mathrm{d}r$ as $N \to \infty$ by Lemma 6.11 in Appendix 6.A. Next, Jensen's inequality, (6.21), (6.71) and the continuous immersion

$$L^p\big((t_{k-1}, t_k) \times (0,T); \mathbb{R}^d\big) \hookrightarrow L^2\big((t_{k-1}, t_k) \times (0,T); \mathbb{R}^d\big)$$

yield, for some constant $C_{\phi,p} = C_{\phi,p}(\alpha, d, T) > 0$,

$$\begin{aligned}
\left|\mathbf{I}_2^N\right| &\leq \|\nabla u\|_\infty \sqrt{\Delta_N} \sum_{k=1}^N \mathbb{E}\left[\left(\int_{t_{k-1}}^{t_k} \mathrm{d}r \int_0^T \left|\overline{B}\left(w_{t_k}^{t_{k-1},\phi}\right) - B(\phi)\right|^2 (r,\xi)\,\mathrm{d}\xi\right)^{\frac{1}{2}}\right] \\
&\leq T^{\frac{1}{2}-\frac{1}{p}} (\Delta_N)^{1-\frac{1}{p}} \|\nabla u\|_\infty \sum_{k=1}^N \mathbb{E}\left[\left\|\overline{B}\left(w_{t_k}^{t_{k-1},\phi}\right) - B(\phi)\right\|_{p,\square}\right] \\
&\leq T^{\frac{1}{2}-\frac{1}{p}} C_{0,p} (\Delta_N)^{1-\frac{1}{p}} \|\nabla u\|_\infty \sum_{k=1}^N \mathbb{E}\left[\left\|w_{t_k}^{t_{k-1},\phi} - \phi\right\|_p\right] \leq T^{\frac{3}{2}-\frac{1}{p}} C_{0,p} C_{\phi,p} \|\nabla u\|_\infty (\Delta_N)^{\frac{1}{2}-\frac{1}{p}} \underset{N\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Here, we set $\|\nabla u\|_\infty = \sup_{t\in[0,T]} \sup_{\phi\in H} \|\nabla u(t,\phi)\|_2$. Regarding $\mathbf{I}_3^N$, we define the modulus of continuity of the map $B(\cdot,\phi)\colon [0,T] \to H$ by

$$\mathfrak{w}\left(B(\cdot,\phi),\delta\right) = \sup_{|u-v|\leq\delta} . \left\|B(u,\phi) - B(v,\phi)\right\|_2, \quad \delta > 0.$$

Since, by hypothesis, $B(\cdot,\phi)$ is continuous on the compact $[0,T]$, it is also uniformly continuous, hence we infer that $\left|\mathbf{I}_3^N\right| \leq T\|\nabla u\|_\infty \mathfrak{w}\left(B(\cdot,\phi),\Delta_N\right) \underset{N\to\infty}{\longrightarrow} 0$. Therefore, we have just shown that

$$\lim_{N\to\infty} \mathbf{I}^N = \int_s^T \langle \nabla u(r,\phi), B(r,\phi)\rangle_H \,\mathrm{d}r. \tag{6.77}$$

Now we investigate $\mathbf{II}^N$, which we split as follows:

$$\begin{aligned}
2\mathbf{II}^N =& \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2u(t_k,\phi)\,\overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right), \overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right)\right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2u(t_k,\phi)\,\overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right), \Sigma_{t_{k-1},t_k}\right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2u(t_k,\phi)\,\Sigma_{t_{k-1},t_k}, \overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right)\right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2u(t_k,\phi)\,\Sigma_{t_{k-1},t_k}, \Sigma_{t_{k-1},t_k}\right\rangle_H\right] =: \mathbf{II}_1^N + \mathbf{II}_2^N + \mathbf{II}_3^N + \mathbf{II}_4^N.
\end{aligned}$$

Let us set $\left\|D^2u\right\|_\infty = \sup_{t\in[0,T]} \sup_{\phi\in H} \left\|D^2u(t,\phi)\right\|_{\mathcal{L}(H;H)}$. By (6.21)-(6.70), arguing similarly to $\mathbf{I}_2^N$ we have, for some $c > 0$,

$$\begin{aligned}
\left|\mathbf{II}_1^N\right| &\leq \left\|D^2u\right\|_\infty \sum_{k=1}^N \mathbb{E}\left[\left\|\overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right)\right\|_2^2\right] \\
&\leq T^{1-\frac{2}{p}} \Delta_N^{1-\frac{2}{p}} \left\|D^2u\right\|_\infty \sum_{k=1}^N \mathbb{E}\left[\left\|\overline{B}\left(w_{t_k}^{t_{k-1},\phi}\right)\right\|_{p,\square}^2\right](t_k - t_{k-1}) \leq c\Delta_N^{1-\frac{2}{p}} \left\|D^2u\right\|_\infty \left(1 + \|\phi\|_p^2\right).
\end{aligned}$$

Moreover, by Hölder's inequality and (6.4), for some $\tilde{c} > 0$,

$$\left|\mathbf{II}_2^N\right| \leq \left\|D^2 u\right\|_\infty \sum_{k=1}^N \left\|\overline{B}_{t_{k-1},t_k}\left(w_{t_k}^{t_{k-1},\phi}\right)\right\|_{\mathcal{H}} \left\|\Sigma_{t_{k-1},t_k}\right\|_{\mathcal{H}} \leq \left\|D^2 u\right\|_\infty \widetilde{c}\, T^{\frac{3}{2}-\frac{1}{p}} \left\|k_2\right\|_2 \Delta_N^{\frac{1}{2}-\frac{1}{p}} \left(1+\|\phi\|_p\right).$$

Since the second bound holds for $\mathbf{II}_3^N$, too, we see that $\mathbf{II}_i^N \to 0$ as $N \to \infty$, $i = 1, 2, 3$.
As for $\mathbf{II}_4^N$, we write it as the following sum:

$$\begin{aligned}
\mathbf{II}_4^N =& \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2 u\left(t_k,\phi\right) \int_{t_{k-1}}^{t_k} \sigma\left(t_k\right) \mathrm{d}W_r, \int_{t_{k-1}}^{t_k} \sigma\left(t_k\right) \mathrm{d}W_r \right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2 u\left(t_k,\phi\right) \int_{t_{k-1}}^{t_k} \left(\sigma(r)-\sigma\left(t_k\right)\right) \mathrm{d}W_r, \int_{t_{k-1}}^{t_k} \sigma\left(t_k\right) \mathrm{d}W_r \right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2 u\left(t_k,\phi\right) \int_{t_{k-1}}^{t_k} \sigma\left(t_k\right) \mathrm{d}W_r, \int_{t_{k-1}}^{t_k} \left(\sigma(r)-\sigma\left(t_k\right)\right) \mathrm{d}W_r \right\rangle_H\right] \\
&+ \sum_{k=1}^N \mathbb{E}\left[\left\langle D^2 u\left(t_k,\phi\right) \int_{t_{k-1}}^{t_k} \left(\sigma(r)-\sigma\left(t_k\right)\right) \mathrm{d}W_r, \int_{t_{k-1}}^{t_k} \left(\sigma(r)-\sigma\left(t_k\right)\right) \mathrm{d}W_r \right\rangle_H\right] \\
=:& \mathbf{II}_{4,1}^N + \mathbf{II}_{4,2}^N + \mathbf{II}_{4,3}^N + \mathbf{II}_{4,4}^N.
\end{aligned}$$

By [66, Proposition 4.30], we have, for every $k = 1, \ldots, N$,

$$D^2 u\left(t_k,\phi\right) \int_{t_{k-1}}^{t_k} \sigma\left(t_k\right) \mathrm{d}W_r = \int_{t_{k-1}}^{t_k} D^2 u\left(t_k,\phi\right) \sigma\left(t_k\right) \mathrm{d}W_r, \quad \mathbb{P}-\text{a.s.},$$

whence, by [66, Corollary 4.29] and Lemma 6.11,

$$\mathbf{II}_{4,1}^N = \sum_{k=1}^N \operatorname{Tr}\left(D^2 u\left(t_k,\phi\right) \sigma\left(t_k\right) \sigma\left(t_k\right)^*\right)\left(t_k - t_{k-1}\right) \underset{N\to\infty}{\longrightarrow} \int_s^T \operatorname{Tr}\left(D^2 u(r,\phi)\, \sigma(r)\, \sigma(r)^*\right) \mathrm{d}r.$$

Furthermore, Hölder's inequality, (6.68) in Lemma 6.7 and [66, Proposition 4.20] yield, for $i = 2, 3$, for some constants $c_1, c_2 > 0$,

$$\left|\mathbf{II}_{4,i}^N\right| \leq c_1 \left\|k_2\right\|_2 \left\|D^2 u\right\|_\infty \sqrt{\Delta_N} \sum_{k=1}^N \left\|\int_{t_{k-1}}^{t_k} \left(\sigma(r)-\sigma\left(t_k\right)\right) \mathrm{d}W_r\right\|_{\mathcal{H}} \leq T c_2 \left\|k_2\right\|_2 \left\|D^2 u\right\|_\infty \Delta_N^{\alpha-\frac{1}{2}} \underset{N\to\infty}{\longrightarrow} 0.$$

Analogous estimates show that $\mathbf{II}_{4,4}^N \to 0$ as $N \to \infty$, as well. Thus,

$$\lim_{N\to\infty} \mathbf{II}^N = \frac{1}{2} \int_s^T \operatorname{Tr}\left(D^2 u(r,\phi)\, \sigma(r)\, \sigma(r)^*\right) \mathrm{d}r. \tag{6.78}$$

At last we study the remainder term $\mathbf{III}^N$ in (6.76). To do this, we employ the fact that the map $D^2 u(t,\cdot) : H \to \mathcal{L}(H;H)$ is $\beta-$Hölder continuous uniformly in time, see (6.91) in Lemma 6.11. We choose $\tilde{\beta} \in (0, \beta)$ such that $2 + \tilde{\beta} < p$; by the expression of $r_{u(t_k,\cdot)}$ in (6.74) we deduce that

$$\begin{aligned}
\left|\mathbf{III}^N\right| &\le \sum_{k=1}^{N} \int_0^1 \mathbb{E}\left[\left\| D^2 u\left(t_k, \phi + r\left(w_{t_k}^{t_{k-1},\phi} - \phi\right) - D^2 u\left(t_k, \phi\right)\right)\right\|_{\mathcal{L}(H;H)} \left\| w_{t_k}^{t_{k-1},\phi} - \phi\right\|_2^2\right] \mathrm{d}r \\
&\le C \sum_{k=1}^{N} \mathbb{E}\left[\left\| w_{t_k}^{t_{k-1},\phi} - \phi\right\|_2^{2+\tilde{\beta}}\right] \le C T^{\left(\frac{1}{2}-\frac{1}{p}\right)(2+\tilde{\beta})} \sum_{k=1}^{N} \mathbb{E}\left[\left\| w_{t_k}^{t_{k-1},\phi} - \phi\right\|_p^{2+\tilde{\beta}}\right] \\
&\le C \sum_{k=1}^{N} \left(t_k - t_{k-1}\right)^{1+\frac{\tilde{\beta}}{2}} \underset{N\to\infty}{\longrightarrow} 0, \qquad (6.79)
\end{aligned}$$

where in the last passage we use Lemma 6.8 and Jensen's inequality. Here $C > 0$ is a constant allowed to change from line to line. Combining (6.77), (6.78), (6.79) in (6.76), we obtain

$$u\left(s, \phi\right) - \Phi\left(\phi\right) = \int_s^T \left\langle \nabla u\left(r, \phi\right), B\left(r, \phi\right)\right\rangle_H \mathrm{d}r + \frac{1}{2}\int_s^T \mathrm{Tr}\left(D^2 u\left(r, \phi\right) \sigma\left(r\right) \sigma\left(r\right)^*\right) \mathrm{d}r,$$

i.e., (6.67). Thus, the proof is complete. ■

Remark 6.3. *Under the hypotheses of Theorem 6.9, for every $\phi \in \Lambda$ the function $u\left(\cdot, \phi\right) : [0, T] \to \mathbb{R}$ defined in* (6.72) *is absolutely continuous on $[0, T]$, because the integrands on the right–hand side of* (6.67) *are bounded on $[0, T]$. Thus, the fundamental theorem of calculus shows that $u\colon [0, T] \times H \to \mathbb{R}$ satisfies the following* Kolmogorov backward equation *in differential form:*

$$\begin{cases} \partial_t u\left(t, \phi\right) + \left\langle \nabla u\left(t, \phi\right), B\left(t, \phi\right)\right\rangle_H + \frac{1}{2}\mathrm{Tr}\left(D^2 u\left(t, \phi\right) \sigma\left(t\right) \sigma\left(t\right)^*\right) = 0, & \text{for a.e. } t \in (0, T), \ \phi \in \Lambda, \\ u\left(T, \phi\right) = \Phi\left(\phi\right), \quad \phi \in H. \end{cases}$$

Remark 6.4. *All the arguments and computations leading to Theorem 6.9 continue to hold when the power α of the kernel k_2 in* (6.1) *varies in $[1, \frac{3}{2})$, i.e., k_2 is the continuous kernel in $\mathbb{R}_+$ given by*

$$k_2(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1}, \quad t \ge 0, \ \text{for some } \alpha \in \left[1, \frac{3}{2}\right).$$

We have however decided to present the theory in the case $\alpha \in (\frac{1}{2}, 1)$ to emphasize the fact that our approach is able to handle rough kernels with explosions at $t = 0$.

Example 6.1. Given two continuous maps $A\colon [0, T] \to \mathbb{R}^{d\times d}$ and $b\colon [0, T] \to \mathbb{R}^d$, define $B\colon \Lambda \to H_\square$ by (cfr. (6.5))

$$B(w)\colon [0, T] \times [0, T] \to \mathbb{R}^d \quad \text{such that}$$
$$B(w)(t, \xi) = 1_{\{\xi>t\}} k_2(\xi - t) \left(A(t)w(t) + b(t)\right), \ t, \xi \in [0, T], \qquad (6.80)$$

for every $w \in \Lambda$. We now show that B satisfies all the hypotheses of Theorem 6.9.

For every $t \in (0, T]$ and $r \in (0, t)$, from the definition in (6.80) it is immediate to see that $B(w)(r, \xi) = 0$, $\xi \in (0, r)$, and that $B(w)(r, \cdot)$ depends on w only via $w\big|_{(0,t)}$. Denote by $\|A\|_\infty = \sup_{t\in[0,T]} |A(t)|$ and by $\|b\|_\infty = \sup_{t\in[0,T]} |b(t)|$, where $|A(t)|$ is the operator norm in $\mathbb{R}^{d\times d}$. Computing, for every w_1, $w_2 \in \Lambda$,

$$\begin{aligned}
\|B(w_1)\|_\square^2 &\le \int_0^T \left(\int_0^T |k_2(\xi - t)|^2 1_{\{\xi>t\}} \left(\|b\|_\infty + \|A\|_\infty |w_1(t)|\right)^2 \mathrm{d}\xi\right) \mathrm{d}t \\
&\le 2T \max\left\{\|b\|_\infty^2, \|A\|_\infty^2\right\} \|k_2\|_2^2 \left(1 + \|w_1\|_2^2\right),
\end{aligned}$$

and

$$\|B(w_2) - B(w_1)\|_{\square}^2 \leq \|A\|_\infty^2 \int_0^T \left(\int_0^T |k_2(\xi - t)|^2 \, 1_{\{\xi > t\}} \, |w_2(t) - w_1(t)|^2 \, \mathrm{d}\xi \right) \mathrm{d}t$$
$$\leq \|A\|_\infty^2 \, \|k_2\|_2^2 \, \|w_2 - w_1\|_2^2 \,,$$

we deduce that Hypothesis 6.1 is satisfied. Since the previous computations can be repeated for every $p \in (2, (1-\alpha)^{-1})$, then condition (6.21) in Remark 6.2 is verified, as well.
As for Hypothesis 6.2, evidently the operator $DB(w_1) \in \mathcal{L}\big(\Lambda_2; H_\square\big)$ defined by

$$[DB(w_1)(w_2)](t, \xi) = 1_{\{\xi > t\}} k_2(\xi - t) A(t) w_2(t), \quad t, \xi \in [0, T], \, w_2 \in \Lambda_2, \tag{6.81}$$

is the Λ_2-Fréchet differential of B in w_1, for any $w_1 \in \Lambda$. Indeed,

$$B(w_1 + h) - B(w_1) - DB(w_1)(h) = 0, \quad w_1, h \in \Lambda.$$

Moreover, from (6.81) we have, for every $w_1, w_2 \in \Lambda$,

$$\|DB(w_1)(w_2)\|_\square \leq \|A\|_\infty \, \|k_2\|_2 \, \|w_2\|_2 \,, \qquad \|DB(w_1) - DB(w_2)\|_{\mathcal{L}(\Lambda_2; H_\square)} = 0,$$

which in particular gives (6.25) with $\gamma = 1$.
The requirements of Hypothesis 6.3 are trivially satisfied (with $\beta = 1$) because, given the affine structure of this example, $D^2 B(w_1) = 0 \in \mathcal{L}(\Lambda_2, \Lambda_2; H_\square)$, $w_1 \in \Lambda$.
In conclusion, for every $w \in \Lambda$, the map $t \mapsto B(t, w) = B(w)(t, \cdot)$ is continuous from $[0, T]$ to H. Indeed, denoting by $\tilde{b}(t)$ the $\mathbb{R}^d-$valued continuous function $A(t)w(t) + b(t)$, by (6.69) and the two following equations we have, for any $r, t \in [0, T]$,

$$\|B(t, w) - B(r, w)\|_2^2 = \int_0^T \left| k_2(\xi - t) \, 1_{\{\xi > t\}} \tilde{b}(t) - k_2(\xi - r) \, 1_{\{\xi > r\}} \tilde{b}(r) \right|^2 \mathrm{d}\xi$$
$$\leq 2 \big\|\tilde{b}\big\|_\infty^2 \int_0^T \left| k_2(\xi - t) \, 1_{\{\xi > t\}} - k_2(\xi - r) \, 1_{\{\xi > r\}} \right|^2 \mathrm{d}\xi + 2 \, \|k_2\|_2^2 \left| \tilde{b}(t) - \tilde{b}(r) \right|^2$$
$$\leq L \left(|t - r|^{2\alpha - 1} + \left| \tilde{b}(t) - \tilde{b}(r) \right|^2 \right),$$

for some constant $L > 0$.

6.4 The mild Kolmogorov equation

A classical approach to the study of the Kolmogorov equation is its mild formulation, see for example [65, Section 6.5] and [66, Section 9.5]. Contrary to the strategy adopted in the previous section, where we have constructed a solution to (6.67) via a stochastic equation (cfr. Theorem 6.9), for the mild Kolmogorov equation we look for a *direct* solution. With the term *direct*, we mean a solution which is determined by a fixed point argument, hence which does not rely on the underlying stochastic PDE.

In this section, we first present a formal reasoning leading to the mild form of (6.67), see (6.85). After that, in Subsection 6.4.1 we explain some difficulties in proving the well–posedness of such a mild formulation, which are essentially due to the structure of the noise. Since it not the purpose of this section to present a general theory with abstract hypotheses, we limit ourselves to observe that the mild Kolmogorov equation cannot be solved for a class of interesting drifts b using common techniques (cfr. Lemma 6.10). Finally, in Subsection 6.4.2, we highlight the theoretical importance of the mild Kolmogorov equation. In particular, we sketch a procedure –relying on the mild form– typically used to prove uniqueness in law for a stochastic PDE under weak regularity requirements on the coefficients. We only mention that studying the relation between the transition semigroup of an SDE and the corresponding mild Kolmogorov equation can also be used for numerical applications, as recently investigated by [85] in the Brownian case and [35] (see also Chapter 2) in the case of isotropic, stable Lévy processes.

Let $\mathcal{C} = C_b(H;\mathbb{R})$ and consider the backward Kolmogorov equation in differential form, formally written as

$$\begin{cases} \partial_s v(s,x) + \langle b(s,x), \nabla v(s,x)\rangle_H + \frac{1}{2}\mathrm{Tr}\left(D^2 v(s,x)\,\sigma(s)\,\sigma(s)^\star\right) = 0, & s\in[0,T),\ x\in H,\\ v(T,x) = \phi(x), \quad \phi\in\mathcal{C}. \end{cases} \tag{6.82}$$

Here, H and σ are those of the previous sections (see, in particular, (6.2)), whereas the drift $b\colon [0,T]\times H \to H$ is a bounded measurable map which could be non–smooth.
We reformulate (6.82) in order to study it in the space $\mathcal{C}$. Let $u(t,x) \coloneqq v(T-t,x)$: u solves the forward equation

$$\begin{cases} \partial_t u(t,x) = \mathcal{A}_{T-t} u(t,x) + \langle b(T-t,x), \nabla u(t,x)\rangle_H, & t\in(0,T],\ x\in H,\\ u(0,x) = \phi(x), \qquad \phi\in\mathcal{C}, \end{cases} \tag{6.83}$$

where we set

$$\mathcal{A}_{T-t} f(x) = \frac{1}{2}\mathrm{Tr}\left(D^2 f(x)\,\sigma(T-t)\,\sigma(T-t)^*\right).$$

Fix $s\in[0,T]$. For every $t\in[s,T]$, we define the linear evolution operator $R_T(t,s):\mathcal{C}\to\mathcal{C}$ by

$$\left(R_T(t,s)\phi\right)(x) = \mathbb{E}\left[\phi\left(x + \int_s^t \sigma(T-r)\,\mathrm{d}W_r\right)\right], \quad x\in H,\ \phi\in\mathcal{C},$$

where W is an $\mathbb{R}^d-$valued, standard Brownian motion as the one introduced in Section 6.1. Consider the auxiliary equation

$$\begin{cases} \partial_t z(t,x) = \mathcal{A}_{T-t} z(t,x), & t\in(s,T],\ x\in H,\\ z(s,x) = \phi(x), \quad \phi\in\mathcal{C}; \end{cases} \tag{6.84}$$

if $\phi\in C_b^{2+\beta}(H)$, then Theorem 6.9 and Remark 6.3 imply that the function $(R_T(t,s)\phi)(x)$ solves this Cauchy problem for almost every $t\in(s,T)$, for every $x\in\Lambda$. At this point, we can introduce the mild formulation of the Kolmogorov equation (6.67):

$$u(t,x) = \left(R_T(t,0)\phi\right)(x) + \int_0^t \left(R_T(t,s)\langle b(T-s,\cdot), \nabla u(s,\cdot)\rangle_H\right)(x)\,\mathrm{d}s, \quad \phi\in\mathcal{C}. \tag{6.85}$$

Note that, heuristically speaking, (6.85) corresponds to the Kolmogorov equation (6.83). Indeed, if $u(t,x)$ solves (6.85), then a formal application of Leibnitz integral rule and (6.84) yield

$$\begin{aligned} \partial_t u(t,\cdot) &= \partial_t R_T(t,0)\phi + R_T(t,t)\langle b(T-t,\cdot), \nabla u(t,\cdot)\rangle_H + \int_0^t \partial_t R_T(t,s)\langle b(T-s,\cdot), \nabla u(s,\cdot)\rangle_H\,\mathrm{d}s \\ &= \mathcal{A}_{T-t} R_T(t,0)\phi + \langle b(T-t,\cdot), \nabla u(t,\cdot)\rangle_H + \int_0^t \mathcal{A}_{T-t} R_T(t,s)\langle b(T-s,\cdot), \nabla u(s,\cdot)\rangle_H\,\mathrm{d}s \\ &= \mathcal{A}_{T-t} u + \langle b(T-t,\cdot), \nabla u(t,\cdot)\rangle_H. \end{aligned}$$

As we have already mentioned, the aim is to prove directly, i.e., by a fixed point argument not relying on a stochastic equation, that (6.85) admits a solution of class, e.g., $C([0,T];\mathcal{C})$. In this regards, the regularity properties of the evolution operator $R_T(t,s)$ are paramount, hence we now discuss them. According to [66, Proposition 4.28], the $H-$valued random variable $\int_s^t \sigma(T-r)\,\mathrm{d}W_r$ is Gaussian, centered, with covariance operator

$$Q_T(t,s) = \int_s^t \sigma(T-r)\,\sigma(T-r)^*\,\mathrm{d}r = \int_{T-t}^{T-s} \sigma(\tau)\,\sigma(\tau)^*\,\mathrm{d}\tau. \tag{6.86}$$

This covariance operator is not trivial as it would be in the case of constant σ. In fact, in such a case it would be easy to see that $R_T(t,s)\,\phi$, $\phi \in \mathcal{C}$, is differentiable in the direction σ (and only in this direction). In our framework with a time–varying σ, the question of the directions of differentiability of $R_T(t,s)\,\phi$, $\phi \in \mathcal{C}$, is much more complex. Nevertheless, it has to be addressed, because the directional differentiability of $R_T(t,s)\,\phi$ is essential to solve directly (6.85). This may be seen in various ways, one of which is the change of variable

$$\theta_T(t,x) = \langle b(T-t,x), \nabla u(t,x)\rangle_H,$$

that leads to the study of the equation

$$\theta_T(t,x) = \langle b(T-t,x), \nabla(R_T(t,0)\,\phi)(x)\rangle_H + \int_0^t \langle b(T-t,x), \nabla(R_T(t,s)\,\theta_T(s,\cdot))(x)\rangle_H\,\mathrm{d}s, \quad \phi \in \mathcal{C}. \tag{6.87}$$

If we can prove that, for some $C, \epsilon > 0$,

$$\sup_{x\in H} |\langle b(T-t,x), \nabla(R_T(t,s)\,\psi)(x)\rangle_H| \le \frac{C}{|t-s|^{1-\epsilon}} \|\psi\|_\infty, \quad 0 \le s < t \le T,\ \psi \in \mathcal{C}, \tag{6.88}$$

then we may try to set up a fixed point argument for the θ_T-equation (6.87) in a suitable space of bounded, measurable functions. This would in turn give a solution for equation (6.85) by simply setting

$$u(t,x) = (R_T(t,0)\,\phi)(x) + \int_0^t (R_T(t,s)\,\theta_T(s,\cdot))(x)\,\mathrm{d}s.$$

6.4.1 The gradient estimate

Using the Gaussian structure of the $H-$valued random variable

$$Z_T(t,s) = \int_s^t \sigma(T-r)\,\mathrm{d}W_r, \quad 0 \le s < t \le T,$$

and denoting by $Q_T(t,s)^{-1}$ the pseudo–inverse of $Q_T(t,s)$, one can prove –via the Cameron Martin formula (see, e.g., [66, Theorem 2.23])– that, for every $\psi \in \mathcal{C}$,

$$\langle b(T-t,x), \nabla(R_T(t,s)\,\psi)(x)\rangle_H = \mathbb{E}\left[\left\langle Q_T(t,s)^{-1}\, b(T-t,x), Z_T(t,s)\right\rangle_H \psi(x+Z_T(t,s))\right],$$

if

$$b(T-t,x) \in \mathrm{Range}\,(Q_T(t,s)).$$

This is not the most general condition to obtain the existence of such directional derivative. Indeed, we could split $Q_T(t,s)^{-1}$ and use the fact that $Q_T(t,s)^{-1/2} Z_T(t,s)$ has good properties, which reduces the problem to investigating $b(T-t,x) \in \text{Range}\left(Q_T(t,s)^{1/2}\right)$. However, handling the square root is even more difficult and thus, for the time being, we analyze the more restrictive condition.

When the previous holds, arguing as in (6.4), for some $c>0$ we have

$$\begin{aligned}\sup_{x\in H} \left|\left\langle b(T-t,x), \nabla\left(R_T(t,s)\psi\right)(x)\right\rangle_H\right| &\le \|\psi\|_\infty \sup_{x\in H} \mathbb{E}\left[\left|\left\langle Q_T(t,s)^{-1} b(T-t,x), Z_T(t,s)\right\rangle_H\right|\right] \\ &\le c\,\|\psi\|_\infty \|k_2\|_2 (t-s)^{1/2} \sup_{x\in H}\left\|Q_T(t,s)^{-1} b(T-t,x)\right\|_2.\end{aligned}$$

Therefore a sufficient condition for the gradient estimate (6.88) is

$$\sup_{x\in H}\left\|Q_T(t,s)^{-1} b(T-t,x)\right\|_2 \le \frac{C}{|t-s|^{\frac{3}{2}-\epsilon}}, \quad 0\le s<t\le T, \text{ for some } C>0.$$

For a general b, standing the potentially very strong degeneracy of $Q_T(t,s)$, we do not see any hope to prove the gradient estimate (6.88). A particular case that, a priori, may look promising, is when the Volterra drift is of the same kind as the noise part, namely (cfr. (6.2))

$$[b(t,x)](\xi) = \bar\beta(x)\, k_2(\xi-t)\, 1_{\{t<\xi\}} = [\sigma(t)\bar\beta(x)](\xi), \quad \xi\in[0,T], \text{ for some } \bar\beta \in \mathcal{B}_b\left(H;\mathbb{R}^d\right).$$

In this case, since $b(T-t,x) = \sigma(T-t)\bar\beta(x)$, we need to prove that

$$\sigma(T-t)\, e_k \in \text{Range}\left(Q_T(t,s)\right), \quad k=1,\dots,d, \tag{6.89}$$

and that

$$\left\|Q_T(t,s)^{-1}\sigma(T-t)\, e_k\right\|_2 \le \frac{C}{|t-s|^{\frac{3}{2}-\epsilon}}, \quad 0\le s<t\le T,\ k=1,\dots,d, \text{ for some } C>0,$$

where $(e_k)_{k=1,\dots,d}$ is the canonical basis of $\mathbb{R}^d$. Recalling that, by (6.86), $Q_T(t,s) = \int_{T-t}^{T-s}\sigma(\tau)\sigma(\tau)^*\,\mathrm{d}\tau$, apparently we could think that (6.89) is true. But it is not, as the necessary condition given by the next lemma shows.

Lemma 6.10. *Let $0\le s<t\le T$ and suppose that $f\in \text{Range}\left(Q_T(t,s)\right)\subset H$. Then $f=g$ almost everywhere in $(0,T)$, where $g\colon(0,T)\to\mathbb{R}^d$ is a continuous function such that $g=0$ in $(0,T-t)$.*

Proof. Fix $0\le s<t\le T$. Consider $f\in\text{Range}\left(Q_T(t,s)\right)$, so that there exists $v\in H$ such that, by (6.86), $f=\int_{T-t}^{T-s}\sigma(\tau)\sigma(\tau)^* v\,\mathrm{d}\tau$. In particular, for every $k=1,\dots,d$, denoting by $\cdot$ the scalar product in $\mathbb{R}^d$, by the standard properties of Bochner's integral we obtain

$$f\cdot e_k = \left(\int_{T-t}^{T-s}\left(\sigma(\tau)^* v\right)\ k_2(\cdot-\tau)\,1_{\{\cdot>\tau\}}\mathrm{d}\tau\right)\cdot e_k = \int_{T-t}^{T-s}\left\langle\sigma(\tau)\,e_k, v\right\rangle_H k_2(\cdot-\tau)\,1_{\{\cdot>\tau\}}\mathrm{d}\tau.$$

Furthermore, recalling (6.1), for a.e. $\xi\in(0,T)$ we have

$$(f\cdot e_k)(\xi) = \frac{1}{\Gamma(\alpha)}\int_{T-t}^{T-s} 1_{\{\tau<\xi\}}\left\langle\sigma(\tau)\,e_k, v\right\rangle_H (\xi-\tau)^{\alpha-1}\,\mathrm{d}\tau.$$

We denote by g_k the function appearing on the right–hand side of the previous equation, i.e.,

$$g_k(\xi) = \frac{1}{\Gamma(\alpha)} \int_{T-t}^{T-s} 1_{\{\tau<\xi\}} \langle \sigma(\tau) e_k, v \rangle_H (\xi - \tau)^{\alpha-1} \, \mathrm{d}\tau, \quad \xi \in (0,T).$$

We want to show the continuity of g_k on the interval $[T-t,T)$: this ensures that g_k is continuous on the whole $(0,T)$, since trivially $g_k = 0$ on $(0,T-t]$. We first write

$$g_k(\xi) = \int_0^{\xi} 1_{\{\tau>T-t\}} \langle \sigma(\tau) e_k, v \rangle_H (\xi - \tau)^{\alpha-1} \, \mathrm{d}\tau, \quad \xi \in [T-t, T-s],$$

and notice that, as $\sigma(\cdot)e_k \in C([0,T];H)$ (see (6.69) in the proof of Lemma 6.7), the mapping $\langle \sigma(\cdot)e_k, v\rangle_H$ is continuous on $[0,T]$. Therefore we invoke [96, Theorem 2.2 (i), Chapter 2] to conclude that g_k is continuous on $[T-t,T-s]$. Secondly, since

$$g_k(\xi) = \int_{T-t}^{T-s} \langle \sigma(\tau) e_k, v \rangle_H (\xi - \tau)^{\alpha-1} \, \mathrm{d}\tau, \quad \xi \in [T-s, T),$$

the continuity of g_k on $[T-s,T)$ can be inferred employing the dominated convergence theorem. Thus, g_k is continuous on $(0,T)$. This shows that the components $f \cdot e_k$, $k = 1, \dots, d$, of the function $f\colon [0,T] \to \mathbb{R}^d$ are almost everywhere equal on $(0,T)$ to continuous functions g_k, which completes the proof. ■

Remark 6.5. *Lemma 6.10 prevents us from choosing another interesting drift $b(t,x)$, namely*

$$[b(t,x)](\xi) = \bar{\beta}(x) 1_{(t,T)}(\xi), \quad \xi \in [0,T], \text{ for some } \bar{\beta} \in \mathcal{B}_b(H;\mathbb{R}^d).$$

6.4.2 Concerning regularization by noise via the Kolmogorov equation

Among the interests of the Kolmogorov equation, there is the theory of regularization by noise: both in finite and infinite dimensions, it has been shown that a sufficiently regular solution to the Kolmogorov equation allows to prove suitable uniqueness results for the underlying stochastic differential equation (see examples in [62, 63, 82, 169, 173]). In contrast with Sections 6.2-6.3, one deals with a stochastic PDE

$$\mathrm{d}X_t = b(t,X_t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W_t, \quad X_0 = x \in H, \tag{6.90}$$

which –a priori– is not well posed, because $b\colon [0,T] \times H \to H$ is subject to weak regularity assumptions not including Lipschitz continuity. The aim is to prove the uniqueness in law of a mild solution to (6.90). A typical approach to achieve this takes the following steps:

1. Write the Kolmogorov equation in mild form (6.85) associated with (6.90) and prove the existence of solutions by a fixed point argument.

2. Possibly after a regularization procedure (see an example in infinite dimensions in [82, Theorem 2.9, Section 2.3.3]), apply Itô formula to $u(T-t, X_t)$, where u solves (6.85) and X_t is any solution of (6.90), prove that the local martingale term is a martingale and obtain an expression for

$$\mathbb{E}[\phi(X_t)].$$

   In this way, one deduces that two solutions have the same marginals. A control on the gradient of u, like the one discussed in Subsection 6.4.1, may help in this step to prove that the local martingale term is a martingale.

3. Apply specific arguments (see [169], [172]) to obtain uniqueness in law.

Under suitable assumptions on b which guarantee the well–posedness of (6.85) (whence Step 1 follows), the details of Steps 2-3 will be the subject of a future research.

Appendix 6.A Regularity of the solution (6.72) of the Kolmogorov equation

In this appendix, we present an auxiliary lemma, namely Lemma 6.11, containing regularity results about the solution $u\colon [0,T]\times H \to \mathbb{R}$ of the Kolmogorov backward equation (6.67) defined in (6.72). Such a lemma plays a key role in the proof of Theorem 6.9.

Lemma 6.11. *Suppose that* $\Phi \in C_b^{2+\beta}(H)$ *and that Hypothesis 6.3 holds. Then, the map* $u\colon [0,T]\times H \to \mathbb{R}$ *defined in* (6.72) *belongs to* $L^\infty\big(0,T; C_b^{2+\beta}(H)\big) \cap C([0,T]\times H;\mathbb{R})$. *In particular, there exists a constant* $C_{d,T,\beta,\Phi} > 0$ *such that*

$$\left\|D^2 u(t,\phi) - D^2 u(t,\psi)\right\|_{\mathcal{L}(H;H)} \le C_{d,T,\beta,\Phi} \left\|\phi-\psi\right\|_2^\beta, \quad \phi,\psi\in H,\, t\in[0,T]. \tag{6.91}$$

Furthermore, the map $(t,\phi,\psi)\mapsto \langle \nabla u(t,\phi),\psi\rangle_H$ *[resp.,* $(t,\phi,\psi,\eta)\mapsto \langle D^2u(t,\phi)\psi,\eta\rangle_H$*] is continuous in* $[0,T]\times H\times H$ *[resp.,* $[0,T]\times H\times H\times H$*].*

Proof. We start off by proving that $u\in C([0,T]\times H;\mathbb{R})$. Consider $t\in[0,T]$, $\phi\in H$ and two sequences $(t_n)_n\subset[0,T]$ and $(\phi_n)_n\subset H$ such that $t_n\to t$ and $\phi_n\to\phi$ as $n\to\infty$. Since $\nabla\Phi\colon H\to H$ is bounded, by the mean value theorem we compute, recalling the definition of u in (6.72),

$$\begin{aligned}|u(t_n,\phi_n)-u(t,\phi)| &\le \mathbb{E}\left[\left|\Phi\left(w_T^{t_n,\phi_n}\right)-\Phi\left(w_T^{t_n,\phi}\right)\right|\right] + \mathbb{E}\left[\left|\Phi\left(w_T^{t_n,\phi}\right)-\Phi\left(w_T^{t,\phi}\right)\right|\right]\\ &\le \|\nabla\Phi\|_\infty\left(\left\|w_T^{t_n,\phi_n}-w_T^{t_n,\phi}\right\|_{\mathcal{H}} + \left\|w_T^{t_n,\phi}-w_T^{t,\phi}\right\|_{\mathcal{H}}\right).\end{aligned} \tag{6.92}$$

By (6.22) in Corollary 6.4, we infer that $\lim_{n\to\infty}\left\|w_T^{t_n,\phi_n}-w_T^{t_n,\phi}\right\|_{\mathcal{H}}=0$. As for $\left\|w_T^{t_n,\phi}-w_T^{t,\phi}\right\|_{\mathcal{H}}$, we first assume that $t_n>t$. Then, by the flow property in (6.12) and Corollary 6.4 we have, for some constants $c_1,c_2>0$ which might depend on ϕ,

$$\left\|w_T^{t_n,\phi}-w_T^{t,\phi}\right\|_{\mathcal{H}} = \left\|w_T^{t_n,\phi}-w_T^{t_n,w_{t_n}^{t,\phi}}\right\|_{\mathcal{H}} \le c_1\left\|w_{t_n}^{t,\phi}-\phi\right\|_{\mathcal{H}} \le c_2\sqrt{|t_n-t|},$$

where the last inequality is due to Lemma 6.8, see (6.71). An analogous argument shows that the previous bound holds even in the case $t_n\le t$, therefore $\lim_{n\to\infty}\left\|w_T^{t_n,\phi}-w_T^{t,\phi}\right\|_{\mathcal{H}}=0$. Going back to (6.92), we conclude that $\lim_{n\to\infty}|u(t_n,\phi_n)-u(t,\phi)|=0$, hence $u\colon[0,T]\times H\to\mathbb{R}$ is continuous, as desired.

We now prove that $u\in L^\infty\big(0,T;C_b^{2+\beta}(H)\big)$. Since $\Phi\in C_b^{2+\beta}(H)$, there exists a constant $C_\Phi>0$ such that

$$\left\|D^2\Phi(\phi)-D^2\Phi(\psi)\right\|_{\mathcal{L}(H;H)} \le C_\Phi\left\|\phi-\psi\right\|_2^\beta, \quad \phi,\psi\in H. \tag{6.93}$$

Obviously, from the boundedness of Φ we have $\|u\|_\infty=\sup_{t\in[0,T]}\sup_{\phi\in H}|u(t,\phi)|<\infty$. First, we want to show that, for every $t\in[0,T]$, $u(t,\cdot)\in C_b^1(H)$, with

$$\langle\nabla u(t,\phi),\psi\rangle_H = \mathbb{E}\left[\left\langle\nabla\Phi\left(w_T^{t,\phi}\right), Dw_T^{t,\phi}\psi\right\rangle_H\right], \quad \phi,\psi\in H. \tag{6.94}$$

To see this, by Taylor's formula applied to Φ we compute, for every ϕ, $h \in H$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\Phi\left(w_T^{t,\phi+h}\right) - \Phi\left(w_T^{t,\phi}\right) - \left\langle \nabla\Phi\left(w_T^{t,\phi}\right), Dw_T^{t,\phi}h\right\rangle_H\right|\right] \\
&\leq \|\nabla\Phi\|_\infty \mathbb{E}\left[\left\|w_T^{t,\phi+h} - w_T^{t,\phi} - Dw_T^{t,\phi}h\right\|_2\right] \\
&\qquad + \mathbb{E}\left[\left|\int_0^1 \left\langle \nabla\Phi\left(w_T^{t,\phi} + r\left(w_T^{t,\phi+h} - w_T^{t,\phi}\right)\right) - \nabla\Phi\left(w_T^{t,\phi}\right), w_T^{t,\phi+h} - w_T^{t,\phi}\right\rangle_H \mathrm{d}r\right|\right] \\
&\leq \|\nabla\Phi\|_\infty \left\|w_T^{t,\phi+h} - w_T^{t,\phi} - Dw_T^{t,\phi}h\right\|_{\mathcal{H}} + \left\|D^2\Phi\right\|_\infty \left\|w_T^{t,\phi+h} - w_T^{t,\phi}\right\|_{\mathcal{H}}^2 = \mathrm{o}\left(\|h\|_2\right). \qquad (6.95)
\end{aligned}$$

Here, for the second inequality we use the Lipschitz continuity of the map $\nabla\Phi\colon H \to H$ –guaranteed by the mean value theorem– and for the third equality we invoke Corollary 6.4 and Theorem 6.5. This shows (6.94), from which we deduce the continuity of the function $\nabla u(t,\cdot)\colon H \to H$. In particular, by (6.32), there exists a constant $C_1 = C_1(d,T)$ such that $\|\nabla u\|_\infty \leq C_1 \|\nabla\Phi\|_\infty$.
We also note that, arguing as in (6.95) and thanks to the estimates of $\left\|w_T^{t,\phi+h} - w_T^{t,\phi} - Dw_T^{t,\phi}h\right\|_{\mathcal{H}}$ in the proof of Theorem 6.5 (see, for instance, (6.37)-(6.44)), for every $M > 0$ we have

$$\sup_{t\in[0,T]} \sup_{\|\phi\|_2,\|\psi\|_2\leq M} \mathbb{E}\left[\left|\Phi\left(w_T^{t,\phi+h\psi}\right) - \Phi\left(w_T^{t,\phi}\right) - h\left\langle \nabla\Phi\left(w_T^{t,\phi}\right), Dw_T^{t,\phi}\psi\right\rangle_H\right|\right] = \mathrm{o}\left(h\right), \quad h \in \mathbb{R}. \qquad (6.96)$$

which gives the continuity of the map $(t,\phi,\psi) \mapsto \langle \nabla u(t,\phi), \psi\rangle_H$ in $[0,T]\times H\times H$ as $u \in C([0,T]\times H;\mathbb{R})$.

Secondly, we claim that $u\left(t,\cdot\right)$ is twice Fréchet differentiable in H, with

$$\begin{aligned}
&\left\langle D^2u\left(t,\phi\right)\psi, \eta\right\rangle_H = \mathbb{E}\left[\left\langle D^2\Phi\left(w_T^{t,\phi}\right) Dw_T^{t,\phi}\psi, Dw_T^{t,\phi}\eta\right\rangle_H + \left\langle \nabla\Phi\left(w_T^{t,\phi}\right), D^2w_T^{t,\phi}\left(\psi,\eta\right)\right\rangle_H\right], \\
&\qquad \phi,\psi,\eta \in H. \qquad (6.97)
\end{aligned}$$

Indeed, recalling (6.94), an application of Taylor's formula on $\nabla\Phi$ yields

$$\begin{aligned}
&\left|\left\langle \nabla u\left(t,\phi+h\right) - \nabla u\left(t,\phi\right) - D^2u\left(t,\phi\right)h, \psi\right\rangle_H\right| \\
&\quad = \left|\mathbb{E}\left[\left\langle \nabla\Phi\left(w_T^{t,\phi+h}\right), Dw_T^{t,\phi+h}\psi\right\rangle_H - \left\langle \nabla\Phi\left(w_T^{t,\phi}\right), Dw_T^{t,\phi}\psi\right\rangle_H\right.\right. \\
&\qquad\qquad \left.\left. - \left\langle D^2\Phi\left(w_T^{t,\phi}\right) Dw_T^{t,\phi}h, Dw_T^{t,\phi}\psi\right\rangle_H - \left\langle \nabla\Phi\left(w_T^{t,\phi}\right), D^2w_T^{t,\phi}\left(h,\psi\right)\right\rangle_H\right]\right| \\
&\quad \leq \mathbb{E}\left[\left|\left\langle D^2\Phi\left(w_T^{t,\phi}\right)\left(w_T^{t,\phi+h} - w_T^{t,\phi} - Dw_T^{t,\phi}h\right), Dw_T^{t,\phi}\psi\right\rangle_H\right|\right] \\
&\qquad + \mathbb{E}\left[\left|\left\langle \nabla\Phi\left(w_T^{t,\phi}\right), Dw_T^{t,\phi+h}\psi - Dw_T^{t,\phi}\psi - D^2w_T^{t,\phi}\left(h,\psi\right)\right\rangle_H\right|\right] \\
&\qquad + \mathbb{E}\left[\left|\left\langle \nabla\Phi\left(w_T^{t,\phi+h}\right) - \nabla\Phi\left(w_T^{t,\phi}\right), \left(Dw_T^{t,\phi+h} - Dw_T^{t,\phi}\right)\psi\right\rangle_H\right|\right] + R_\Phi\left(\phi,\psi,h\right) \\
&\quad =: \left(\mathbf{I}_1 + \mathbf{II}_1 + \mathbf{III}_1 + R_\Phi\right)\left(\phi,\psi,h\right), \qquad (6.98)
\end{aligned}$$

for every $\phi,\psi,h \in H$. Here, we denote by

$$\begin{aligned}
&R_\Phi\left(\phi,\psi,h\right) \\
&\quad = \mathbb{E}\left[\left|\left\langle \int_0^1 \left(D^2\Phi\left(w_T^{t,\phi} + r\left(w_T^{t,\phi+h} - w_T^{t,\phi}\right)\right) - D^2\Phi\left(w_T^{t,\phi}\right)\right)\left(w_T^{t,\phi+h} - w_T^{t,\phi}\right)\mathrm{d}r, Dw_T^{t,\phi}\psi\right\rangle_H\right|\right].
\end{aligned}$$

Using (6.32), (6.93) and Corollary 6.4, for some constant $c_3 > 0$ we compute

$$\begin{aligned}R_\Phi(\phi,\psi,h) &\le C_\Phi C_1 \mathbb{E}\left[\left\|w_T^{t,\phi+h}-w_T^{t,\phi}\right\|_2^{1+\beta}\right]\|\psi\|_2 \le C_\Phi C_1 \left\|w_T^{t,\phi+h}-w_T^{t,\phi}\right\|_{\mathcal{H}}^{1+\beta}\|\psi\|_2\\ &\le c_3\|\psi\|_2\|h\|_2^{1+\beta}, \quad \phi,\psi,h\in H,\end{aligned}$$

where we also employ Jensen's inequality noticing that $1+\beta\le 2$. Next,

$$|\mathbf{I}_1(\phi,\psi,h)| \le C_1\left\|D^2\Phi\right\|_\infty\|\psi\|_2\left\|w_T^{t,\phi+h}-w_T^{t,\phi}-Dw_T^{t,\phi}h\right\|_{\mathcal{H}}, \quad \phi,\psi,h\in H,$$

and

$$|\mathbf{II}_1(\phi,\psi,h)| \le \|\nabla\Phi\|_\infty\|\psi\|_2\left\|Dw_T^{t,\phi+h}-Dw_T^{t,\phi}-D^2w_T^{t,\phi}(h,\cdot)\right\|_{\mathcal{L}(H;\mathcal{H})}, \quad \phi,\psi,h\in H.$$

Finally, by Corollary 6.4 and (6.32) (recall that, under Hypothesis 6.3, we take $\gamma=\beta$ in (6.27), see (6.54))

$$|\mathbf{III}_1(\phi,\psi,h)| \le C_1\left\|D^2\Phi\right\|_\infty\|\psi\|_2\,\mathbb{E}\left[\left\|w_T^{t,\phi+h}-w_T^{t,\phi}\right\|_2^{1+\beta}\right] \le \tilde{c}\left\|D^2\Phi\right\|_\infty\|\psi\|_2\|h\|_2^{1+\beta}, \quad \phi,\psi,h\in H,$$

for some $\tilde{c}>0$. Going back to (6.98), by Theorem 6.6, the previous estimates let us write, for some constant $C>0$,

$$\begin{aligned}\left\|\nabla u(t,\phi+h)-\nabla u(t,\phi)-D^2u(t,\phi)h\right\|_2 &= \sup_{\|\psi\|_2\le 1}\left|\left\langle \nabla u(t,\phi+h)-\nabla u(t,\phi)-D^2u(t,\phi)h,\psi\right\rangle_H\right|\\ &\le C\left(\left\|w_T^{t,\phi+h}-w_T^{t,\phi}-Dw_T^{t,\phi}h\right\|_{\mathcal{H}}+\left\|Dw_T^{t,\phi+h}-Dw_T^{t,\phi}-D^2w_T^{t,\phi}(h,\cdot)\right\|_{\mathcal{L}(H;\mathcal{H})}+\|h\|_2^{1+\beta}\right)\\ &= \mathrm{o}\left(\|h\|_2\right), \quad \phi,h\in H,\end{aligned} \tag{6.99}$$

which proves (6.97). In particular, by (6.32)-(6.56), there is a constant $C_2=C_2(d,T)>0$ such that

$$\left\|D^2u\right\|_\infty \le C_2\left(\left\|D^2\Phi\right\|_\infty+\|\nabla\Phi\|_\infty\right).$$

In addition, arguing as in (6.99) (see also (6.96)) and thanks to the estimates of $\big\|Dw_T^{t,\phi+h}-Dw_T^{t,\phi}-D^2w_T^{t,\phi}(h,\cdot)\big\|_{\mathcal{L}(H;\mathcal{H})}$ in the proof of Theorem 6.6 (see, for instance, (6.61)), for every $M>0$ we have

$$\sup_{t\in[0,T]}\ \sup_{\|\phi\|_2,\|\psi\|_2,\|\eta\|_2\le M}\left|\left\langle \nabla u(t,\phi+h\psi)-\nabla u(t,\phi)-hD^2u(t,\phi)\psi,\eta\right\rangle_H\right| = \mathrm{o}(h), \quad h\in\mathbb{R}.$$

Since we have proved that $\langle\nabla u(t,\phi),\psi\rangle_H$ is continuous in $[0,T]\times H\times H$, the previous equation ensures that the map $(t,\phi,\psi,\eta)\mapsto\langle D^2u(t,\phi)\psi,\eta\rangle_H$ is continuous in $[0,T]\times H\times H\times H$, as desired.

In conclusion, we prove that $u(t,\cdot) \in C_b^{2+\beta}(H)$. From (6.97), for every $\phi_1, \phi_2 \in H$,

$$\begin{aligned}
&\left\langle \left(D^2 u(t,\phi_1) - D^2 u(t,\phi_2)\right)\psi, \eta\right\rangle_H \\
&= \mathbb{E}\left[\left\langle \left(D^2\Phi\left(w_T^{t,\phi_1}\right) - D^2\Phi\left(w_T^{t,\phi_2}\right)\right) Dw_T^{t,\phi_1}\psi, Dw_T^{t,\phi_1}\eta\right\rangle_H\right] \\
&\quad + \mathbb{E}\left[\left\langle D^2\Phi\left(w_T^{t,\phi_2}\right)\left(Dw_T^{t,\phi_1} - Dw_T^{t,\phi_2}\right)\psi, Dw_T^{t,\phi_1}\eta\right\rangle_H\right] \\
&\quad + \mathbb{E}\left[\left\langle D^2\Phi\left(w_T^{t,\phi_2}\right) Dw_T^{t,\phi_2}\psi, \left(Dw_T^{t,\phi_1} - Dw_T^{t,\phi_2}\right)\eta\right\rangle_H\right] \\
&\quad + \mathbb{E}\left[\left\langle \nabla\Phi\left(w_T^{t,\phi_1}\right) - \nabla\Phi\left(w_T^{t,\phi_2}\right), D^2 w_T^{t,\phi_1}(\psi,\eta)\right\rangle_H\right] \\
&\quad + \mathbb{E}\left[\left\langle \nabla\Phi\left(w_T^{t,\phi_2}\right), \left(D^2 w_T^{t,\phi_1} - D^2 w_T^{t,\phi_2}\right)(\psi,\eta)\right\rangle_H\right] \\
&=: \left(\mathbf{I}_2 + \mathbf{II}_2 + \mathbf{III}_2 + \mathbf{IV}_2 + \mathbf{V}_2\right)(\phi_1,\phi_2,\psi,\eta), \quad \psi,\eta \in H.
\end{aligned}$$

To keep notation the short, in what follows we consider arbitrary $\psi, \eta \in H$, we do not write $(\phi_1, \phi_2, \psi, \eta)$ and we denote by $c = c(d, T, \beta) > 0$ a constant that might change from line to line. Observe that, by (6.32)-(6.93), Corollary 6.4 and Jensen's inequality,

$$|\mathbf{I}_2| \leq c\, C_\Phi \|\psi\|_2 \|\eta\|_2 \|\phi_1 - \phi_2\|_2^\beta.$$

Moreover, by (6.32) (see also (6.54)),

$$|\mathbf{II}_2| \leq c \left\|D^2\Phi\right\|_\infty \|\psi\|_2 \|\eta\|_2 \|\phi_1 - \phi_2\|_2^\beta.$$

An analogous estimate holds for $|\mathbf{III}_2|$, too. As for the remaining addends, by (6.56) we have

$$|\mathbf{IV}_2| \leq c \left\|D^2\Phi\right\|_\infty \|\psi\|_2 \|\eta\|_2 \|\phi_1 - \phi_2\|_2,$$

and

$$|\mathbf{V}_2| \leq c \|\nabla\Phi\|_\infty \|\psi\|_2 \|\eta\|_2 \|\phi_1 - \phi_2\|_2^\beta.$$

Thus, the function $D^2 u(t,\cdot) : H \to \mathcal{L}(H; H)$ is $\beta-$Hölder continuous uniformly in time and the proof is complete. ■

References

[1] Abi Jaber, E. (2019). Lifting the Heston model. *Quantitative Finance, 19*(12), 1995–2013.

[2] Abi Jaber, E. (2021). Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels. *Bernoulli, 27*(3), 1583–1615.

[3] Abi Jaber, E., Cuchiero, C., Larsson, M., & Pulido, S. (2021). A weak solution theory for stochastic Volterra equations of convolution type. *The Annals of Applied Probability, 31*(6), 2924–2952.

[4] Abi Jaber, E., & El Euch, O. (2019a). Markovian structure of the Volterra Heston model. *Statistics & Probability Letters, 149*, 63–72.

[5] Abi Jaber, E., & El Euch, O. (2019b). Multi-factor approximation of rough volatility models. *SIAM Journal on Financial Mathematics, 10*(2), 309–349.

[6] Abi Jaber, E., Illand, C., & Li, S. (2022a). Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *Preprint arXiv: 2212.08297.*

[7] Abi Jaber, E., Illand, C., & Li, S. (2022b). The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates SPX & VIX smiles. *Preprint arXiv: 2212.10917.*

[8] Abi Jaber, E., Larsson, M., & Pulido, S. (2019). Affine Volterra processes. *The Annals of Applied Probability, 29*(5), 3155–3200.

[9] Alos, E., Garcia-Lorite, D., & Gonzalez, A. M. (2022). On smile properties of volatility derivatives: Understanding the VIX skew. *SIAM Journal on Financial Mathematics, 13*(1), 32–69.

[10] Alos, E., Gatheral, J., & Radoičić, R. (2020). Exponentiation of conditional expectations under stochastic volatility. *Quantitative Finance, 20*(1), 13–27.

[11] Alos, E., León, J. A., & Vives, J. (2007). On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics, 11*(4), 571–589.

[12] Alos, E., & Shiraya, K. (2019). Estimating the Hurst parameter from short term volatility swaps: a Malliavin calculus approach. *Finance and Stochastics, 23*(2), 423–447.

[13] Arnold, L., & Imkeller, P. (1996). Stratonovich calculus with spatial parameters and anticipative problems in multiplicative ergodic theory. *Stochastic Processes and their Applications, 62*(1), 19–54.

[14] Avellaneda, M., & Papanicolaou, A. (2018). Statistics of VIX Futures and Applications to Trading Volatility Exchange-Traded Products. *Journal of Investment Strategies, 7*(2), 1–33.

[15] Barczy, M., Li, Z., & Pap, G. (2015). Yamada-Watanabe results for stochastic differential equations with jumps. *International Journal of Stochastic Analysis, 2015.*

[16] Barndorff-Nielsen, O., & Shephard, N. (2001a). Modelling by Lévy processes for financial econometrics. In *Lévy Processes*, Birkhauser, 283–318.

[17] Barndorff-Nielsen, O., & Shephard, N. (2001b). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistics Society: Series B (Statistical Methodology), 63*(2), 167–241.

[18] Barrasso, A., & Russo, F. (2022). Gâteaux type path–dependent PDEs and BSDEs with Gaussian forward processes. *Stochastics and Dynamics, 22*(01), 2250007.

[19] Bates, D. (1996). Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *The Review of Financial Studies, 9*(1), 69–107.

[20] Bayer, C., & Breneis, S. (2023). Markovian approximations of stochastic Volterra equations with the fractional kernel. *Quantitative Finance, 23*(1), 53–70.

[21] Bayer, C., Friz, P., & Gatheral, J. (2016). Pricing under rough volatility. *Quantitative Finance, 16*(6), 887–904.

[22] Beck, C., Becker, S., Grohs, P., Jaafari, N., & Jentzen, A. (2021). Solving the Kolmogorov PDE by means of deep learning. *Journal of Scientific Computing, 88*(3), 1–28.

[23] Beesack, P. R. (1969). Comparison theorems and integral inequalities for Volterra integral equations. *Proceedings of the American Mathematical Society, 20*(1), 61–66.

[24] Bennedsen, M., Lunde, A., & Pakkanen, M. S. (2021). Decoupling the short-and long-term behavior of stochastic volatility. *Journal of Financial Econometrics.*

[25] Benth, F. E., Benth, J. S., & Koekebakker, S. (2008). *Stochastic modelling of electricity and related markets* (Vol. 11). World Scientific.

[26] Bernis, G., Brignone, R., Scotti, S., & Sgarra, C. (2021). A Gamma Ornstein-Uhlenbeck model driven by a Hawkes process. *Mathematics and Financial Economics, 15*(4), 747–773.

[27] Bertsekas, D., & Shreve, S. (1978). *Stochastic optimal control: The discrete–time case.* Mathematics in Science and Engineering.

[28] Bezandry, P. H., & Fernique, X. (1990). Analyse de fonctions aleatoires peu regulieres sur [0, 1]. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique, 310*(10), 745–750.

[29] Bezandry, P. H., & Fernique, X. (1992). Sur la propriété de la limite centrale dans $\mathcal{D}[0,1]$. In *Annales de l'IHP Probabilités et statistiques, 28*(1), 31–46.

[30] Bichteler, K. (2002). *Stochastic integration with jumps* (No. 89). Cambridge University Press.

[31] Billingsley, P. (2013). *Convergence of probability measures.* John Wiley & Sons.

[32] Bogachev, V. I. (2007). *Measure theory* (Vol. 1). Springer Science & Business Media.

[33] Bogdan, K., Stós, A., & Sztonyk, P. (2003). Harnack inequality for stable processes on d-sets. *Studia Math, 158*(2), 163–198.

[34] Bondi, A. (2022a). Smoothing effect and Derivative formulas for Ornstein–Uhlenbeck processes driven by subordinated cylindrical Brownian noises. *Journal of Functional Analysis, 283*(10), 109660.

[35] Bondi, A. (2022b). Probability computation for high-dimensional semilinear SDEs driven by isotropic $\alpha-$stable processes via mild Kolmogorov equations. To appear on *Electronic Journal of Probability.*

[36] Bondi, A., & Flandoli, F. (2023). On the Kolmogorov equation associated with Volterra equations and Fractional Brownian Motion. *Preprint arXiv: 2309.13597.*

[37] Bondi, A., Livieri, G., & Pulido, S. (2022). Affine Volterra processes with jumps. *Preprint arXiv: 2203.06400.*

[38] Bondi, A., & Priola, E. (2023). A sharp càdlàg property for jump diffusions and dynamic programming principle. *Preprint arXiv: 2307.16871.*

[39] Bondi, A., Pulido, S., & Scotti, S. (2022). The rough Hawkes Heston stochastic volatility model. *Preprint arXiv: 2210.12393.*

[40] Bouchard, B., & Touzi, N. (2011). Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization, 49*(3), 948–962.

[41] Brignone, R., & Sgarra, C. (2019). Asian options pricing in Hawkes-type jump-diffusion models. *Annals of Finance, 16*(1), 101–119.

[42] Brown, L. D., & Purves, R. (1973). Measurable selections of extrema. *The annals of statistics*, 902–912.

[43] Brzeźniak, Z., Liu, W., & Zhu, J. (2014). Strong solutions for SPDE with locally monotone coefficients driven by Lévy noise. *Nonlinear Analysis: Real World Applications, 17,* 283–310.

[44] Brzeźniak, Z., & Zabczyk, J. (2010). Regularity of Ornstein-Uhlenbeck processes driven by a Lévy white noise. *Potential Analysis, 32*(2), 153–188.

[45] Callegaro, G., Grasselli, M., & Pages, G. (2021). Fast hybrid schemes for fractional Riccati equations (rough is not so tough). *Mathematics of Operations Research, 46*(1), 221–254.

[46] Callegaro, G., Mazzoran, A., & Sgarra, C. (2022). A self-exciting modeling framework for forward prices in power markets. *Applied Stochastic Models in Business and Industry, 38*(1), 27–48.

[47] Carr, P., & Lee, R. (2009). Robust Replication of Volatility Derivatives. *Available at* `https://math.uchicago.edu/%7Erl/rrvd.pdf`

[48] Chen, N., & Majda, A. J. (2018). Efficient statistically accurate algorithms for the Fokker–Planck equation in large dimensions. *Journal of Computational Physics, 354,* 242–268.

[49] Chevalier, E., Pulido, S., & Zúñiga, E. (2022). American options in the Volterra Heston model. *SIAM Journal on Financial Mathematics, 13*(2), 426–458.

[50] Cohen, S. N., & Elliott, R. J. (2015). *Stochastic calculus and applications* (Vol. 2). New York: Birkhäuser.

[51] Comte, F., & Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical finance, 8*(4), 291–323.

[52] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance, 1*(2), 223–236.

[53] Cont, R. (2011). Statistical Modeling of High Frequency Financial Data: Facts, Models and Challenges. *IEEE Signal Processing, 28*(5), 16–25.

[54] Cont, R., & Kokholm, T. (2013). A Consistent Pricing Model for Index Options and Volatility Derivatives. *Mathematical Finance, 23*(2), 248–274.

[55] Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–408.

[56] Cuchiero, C., Filipović, D., Mayerhofer, E., & Teichmann, J. (2011). Affine processes on positive semidefinite matrices. *The Annals of Applied Probability, 21*(2), 397–463.

[57] Cuchiero, C., Gazzani, G., & Svaluto-Ferro, S. (2022). Signature-based models: theory and calibration. *Preprint arXiv: 2207.13136.*

[58] Cuchiero, C., & Teichmann, J. (2019). Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions in Economics and Finance, 42*(2), 407–448.

[59] Cuchiero, C., & Teichmann, J. (2020). Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *Journal of evolution equations, 20*(4), 1301–1348.

[60] Curato, I. V., & Sanfelici, S. (2015). Measuring the Leverage Effect in a High Frequency Trading Framework. *Handbook of High Frequency Trading*, 425–446.

[61] Da Prato, G. (2014). *Introduction to stochastic analysis and Malliavin calculus* (Vol. 13). Springer.

[62] Da Prato, G., Flandoli, F., Priola, E., & Röckner, M. (2013). Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift. *The Annals of Probability, 41*(5), 3306–3344.

[63] Da Prato, G., Flandoli, F., Röckner, M., & Veretennikov, A. Y. (2016). Strong uniqueness for SDEs in Hilbert spaces with nonregular drift. *The Annals of Probability, 44*(3), 1985–2023.

[64] Da Prato, G., & Zabczyk, J. (1996). *Ergodicity for infinite dimensional systems* (Vol. 229). Cambridge University Press.

[65] Da Prato, G., & Zabczyk, J. (2002). *Second order partial differential equations in Hilbert spaces* (Vol. 293). Cambridge University Press.

[66] Da Prato, G., & Zabczyk, J. (2014). *Stochastic equations in infinite dimensions.* Cambridge University Press.

[67] Delemotte, J., De Marco, S., & Segonne, F. (2023). Yet another analysis of the SP500 at-the-money skew: crossover of different power-law behaviours. *Available at SSRN 4428407.*

[68] Demeterfi, K., Derman, E., Kamal, M., & Zou, J. (1999). More than you ever wanted to know about volatility swaps. *Goldman Sachs Quantitative Strategies Research Notes.*

[69] Deng, C. S., & Schilling, R. L. (2017). Harnack inequalities for SDEs driven by time-changed fractional Brownian motions. *Electronic Journal of Probability, 22*, 1–23.

[70] Deng, C. S., Schilling, R. L., & Song, Y. H. (2017). Subgeometric rates of convergence for Markov processes under subordination. *Advances in Applied Probability, 49*(1), 162–181.

[71] Diethelm, K., Ford, N. J., & Freed, A. D. (2002). A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics, 29*(1), 3–22.

[72] Diethelm, K., Ford, N. J., & Freed, A. D. (2004). Detailed error analysis for a fractional Adams method. *Numerical algorithms, 36*(1), 31–52.

[73] Dotsis, G., Psychoyios, D., & Skiadopoulos, G. (2007). An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking & Finance, 31*(12), 3584–3603.

[74] Duffie, D., Filipović, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability, 13*(3), 984–1053.

[75] Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*(6), 1343–1376.

[76] El Euch, O., Fukasawa, M., & Rosenbaum, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance and Stochastics, 22*(2), 241–280.

[77] El Euch, O., Gatheral, J., Radoičić, R., & Rosenbaum, M. (2020). The Zumbach effect under rough Heston. *Quantitative Finance, 20*(2), 235–241.

[78] El Euch, O., & Rosenbaum, M. (2018). Perfect hedging in rough Heston models. *The Annals of Applied Probability, 28*(6), 3813–3856.

[79] El Euch, O., & Rosenbaum, M. (2019). The characteristic function of rough Heston models. *Mathematical Finance, 29*(1), 3–38.

[80] Fang, F., & Oosterlee, C. W. (2009). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing, 31*(2), 826–848.

[81] Filipović, D. (2001). A general characterization of one factor affine term structure models. *Finance and Stochastics, 5*(3), 389–412.

[82] Flandoli, F. (2011). *Random Perturbation of PDEs and Fluid Dynamic Models: École d'été de Probabilités de Saint-Flour XL–2010* (Vol. 2015). Springer Science & Business Media.

[83] Flandoli, F., & Luo, D. (2020). Convergence of transport noise to Ornstein-Uhlenbeck for 2D Euler equations under the enstrophy measure. *The Annals of Probability, 48*(1), 264–295.

[84] Flandoli, F., Luo, D., & Ricci, C. (2021). A numerical approach to Kolmogorov equation in high dimension based on Gaussian analysis. *Journal of Mathematical Analysis and Applications, 493*(1), 124505.

[85] Flandoli, F., Luo, D., & Ricci, C. (2023). Numerical computation of probabilities for nonlinear SDEs in high dimension using Kolmogorov equation. *Applied Mathematics and Computation, 436*, 127520.

[86] Flandoli, F., & Zanco, G. (2016). An infinite-dimensional approach to path-dependent Kolmogorov equations. *The Annals of Probability, 44*(4), 2643–2693.

[87] Fukasawa, M. (2017). Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance, 17*(2), 189–198.

[88] Galeati, L., & Harang, F. A. (2022). Regularization of multiplicative SDEs through additive noise. *The Annals of Applied Probability, 32*(5), 3930–3963.

[89] Galeati, L., Harang, F. A., & Mayorcas, A. (2023). Distribution dependent SDEs driven by additive fractional Brownian motion. *Probability Theory and Related Fields, 185*(1–2), 251–309.

[90] Gatheral, J., Jaisson, T., & Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance, 18*(6), 933–949.

[91] Gatheral, J., Jusselin, P., & Rosenbaum, M. (2020). The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem. *Risk, cutting edge*, `https://www.risk.net/cutting-edge/banking/7530461/the-quadratic-rough-heston-model-and-the-joint-sp-500vix-smile-calibration-problem`.

[92] Gatheral, J., & Keller-Ressel, M. (2019). Affine forward variance models. *Finance and Stochastics, 23*(3), 501–533.

[93] Gonzato, L., & Sgarra, C. (2021). Self-exciting jumps in the oil market: Bayesian estimation and dynamic hedging. *Energy Economics, 99*, 105279.

[94] Goutte, S., Ismail, A., & Pham, H. (2017). Regime-switching stochastic volatility model: estimation and calibration to VIX options. *Applied Mathematical Finance, 24*(1), 38–75.

[95] Grasselli, M. (2016). The 4/2 stochastic volatility model: a unified approach for the Heston and the 3/2 model. *Mathematical Finance, 27*(4), 1013–1034.

[96] Gripenberg, G., Londen, S. O., & Staffans, O. (1990). *Volterra integral and functional equations* (No. 34). Cambridge University Press.

[97] Grzelak, L.A. (2022). On Randomization of Affine Diffusion Processes with Application to Pricing of Options on VIX and S&P 500. *Preprint arXiv: 2208.12518.*

[98] Guyon, J. (2020a). Inversion of convex ordering in the VIX market. *Quantitative Finance, 20*(10), 1597–1623.

[99] Guyon, J. (2020b). The joint S&P 500/VIX smile calibration puzzle solved. *Risk, April.*

[100] Guyon, J., & El Amrani, M. (2022). Does the Term-Structure of Equity At-the-Money Skew Really Follow a Power Law? *Available at SSRN 4174538.*

[101] Guyon, J., & Lekeufack, J. (2022). Volatility Is (Mostly) Path-Dependent. *Available at SSRN 4174589.*

[102] Guyon, J., & Mustapha, S. (2022). Neural Joint S&P 500/VIX Smile Calibration. *Available at SSRN 4309576.*

[103] Harang, F. A., & Ling, C. (2022). Regularity of Local Times Associated with Volterra–Lévy Processes and Path–Wise Regularization of Stochastic Differential Equations. *Journal of Theoretical Probability, 35*(3), 1706–1735.

[104] Heston, S. L. (1993). A Closed-Form Solution for Option with Stochastic Volatility with Applications to Bond and Currency Options. *The review of financial studies, 6*(2), 327–343.

[105] Horst, U., & Xu, W. (2022). The microstructure of stochastic volatility models with self-exciting jump dynamics. *Preprint arXiv: 1911.12969.*

[106] Ikeda, N., & Watanabe, S. (2014). *Stochastic differential equations and diffusion processes.* Elsevier.

[107] Imkeller, P., & Scheutzow, M. (1999). On the spatial asymptotic behavior of stochastic flows in Euclidean space. *The Annals of Probability, 27*(1), 109–129.

[108] Ishikawa, Y. (2004). Optimal control problem associated with jump processes. *Applied Mathematics and Optimization, 50*, 21–65.

[109] Jacod, J. (2006). *Calcul stochastique et problemes de martingales* (Vol. 714). Springer.

[110] Jacod, J., & Shiryaev, A. (2013). *Limit theorems for stochastic processes* (Vol. 288). Springer Science & Business Media.

[111] Jakubowski, A. (1986). On the Skorokhod topology. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (Vol. 22, No. 3, pp. 263–285). Institut Henri Poincaré.

[112] Jarrow, R. A. (2018). *Continuous-time asset pricing theory.* Springer.

[113] Jiao, Y., Ma, C., & Scotti, S. (2017). Alpha-CIR model with branching processes in sovereign interest rate modeling. *Finance and Stochastics, 21*(3), 789–813.

[114] Jiao, Y., Ma, C., Scotti, S., & Zhou, C. (2021). The Alpha-Heston stochastic volatility model. *Mathematical Finance, 31*(3), 943–978.

[115] Kager, G., & Scheutzow, M. (1997). Generation of one-sided random dynamical systems by stochastic differential equations. *Electronic Journal of Probability, 2*(8), 1–17.

[116] Kallenberg, O. (1997). *Foundations of modern probability* (Vol. 2). New York: Springer.

[117] Kallenberg, O. (2017). *Random measures, theory and applications* (Vol. 1). Cham: Springer International Publishing.

[118] Kallsen, J., Muhle-Karbe, J., & Voss, M. (2011). Pricing options on variance in affine stochastic volatility models. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 21*(4), 627–641.

[119] Karatzas, I., & Shreve, S. E. (1998). Brownian motion. In *Brownian Motion and Stochastic Calculus* (pp. 47–127). Springer, New York, NY.

[120] Keller-Ressel, M. (2011). Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 21*(1), 73–98.

[121] Keller-Ressel, M., Larsson, M., & Pulido, S. (2018). Affine rough models. *Preprint arXiv: 1812.08486.*

[122] Khan, L. A. (1986). Separability in function spaces. *Journal of mathematical analysis and applications, 113*(1), 88–92.

[123] Krylov, N. V. (2008). *Controlled diffusion processes* (Vol. 14). Springer Science & Business Media.

[124] Kunita, H. (1990). *Stochastic flows and stochastic differential equations* (Vol. 24). Cambridge University Press.

[125] Kunita, H. (2004). Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis* (pp. 305–373). Birkhäuser Boston.

[126] Kunita, H. (2019). *Stochastic Flows and Jump–Diffusions* (Vol. 92). Springer.

[127] Kusuoka, S. (2010). Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions. *Kyoto Journal of Mathematics, 50*(3), 491–520.

[128] Leahy, J. M., & Mikulevičius, R. (2016). On classical solutions of linear stochastic integro-differential equations. *Stochastics and Partial Differential Equations: Analysis and Computations, 4*, 535–591.

[129] Lépingle, D., & Mémin, J. (1978). Sur l'intégrabilité uniforme des martingales exponentielles. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 42*(3), 175–203.

[130] Lewis, A. L. (2001). A simple option formula for general jump-diffusion and other exponential Lévy processes. *Available at SSRN 282110.*

[131] Livieri, G., Mouti, S., Pallavicini, A., & Rosenbaum, M. (2018). Rough volatility: evidence from option prices. *IISE transactions, 50*(9), 767–776.

[132] Mancino, M. E., Scotti, S., & Toscano, G. (2020). Is the Variance Swap Rate Affine in the Spot Variance? Evidence from S&P500 Data. *Applied Mathematical Finance, 27*(4), 288–316.

[133] Mancino, M. E., & Toscano, G. (2022). Rate-efficient asymptotic normality for the Fourier estimator of the leverage process. *Statistics and Its Interface, 15*(1), 73–89.

[134] Marinelli, C., & Röckner, M. (2014). On maximal inequalities for purely discontinuous martingales in infinite dimensions. In *Séminaire de Probabilités XLVI* (pp. 293–315). Springer, Cham.

[135] Meyer, P. A. (1966). *Probability and potentials* (Vol. 1318). Blaisdell Publishing Company.

[136] Mytnik, L., & Salisbury, T. S. (2015). Uniqueness for Volterra-type stochastic integral equations. *Preprint arXiv: 1502.05513.*

[137] Natanson, I. P. (2016). *Theory of functions of a real variable.* Courier Dover Publications.

[138] Nicolato, E., Pisani, C., & Sloth, D. (2017). The impact of jump distributions on the implied volatility of variance. *SIAM Journal on Financial Mathematics, 8*(1), 28–53.

[139] Nualart, D., & Ouknine, Y. (2002). Regularization of differential equations by fractional noise. *Stochastic Processes and their Applications, 102*(1), 103–116.

[140] Nualart, D., & Ouknine, Y. (2003). Stochastic differential equations with additive fractional noise and locally unbounded drift. *Stochastic inequalities and applications* (pp. 353–365). Birkhäuser Basel.

[141] Pacati, C., Pompa, G., & Renò, R. (2018). Smiling twice: the Heston++ model. *Journal of Banking & Finance, 96*, 185–206.

[142] Papanicolaou, A. (2022). Consistent time-homogeneous modeling of SPX and VIX derivatives. *Mathematical Finance, 32*(3), 907–940.

[143] Papanicolaou, A., & Sircar, R. (2014). A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance, 14*(10), 1811–1827.

[144] Peszat, S., & Zabczyk, J. (2007). *Stochastic partial differential equations with Lévy noise: An evolution equation approach* (Vol. 113). Cambridge University Press.

[145] Pham, H. (2009). *Continuous–time stochastic control and optimization with financial applications* (Vol. 61). Springer Science & Business Media.

[146] Pham, H., & Wei, X. (2017). Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics. *SIAM Journal on Control and Optimization, 55*(2), 1069–1101.

[147] Pollard, D. (2012). *Convergence of stochastic processes.* Springer Science & Business Media.

[148] Pragarauskas, H. (1978). Control of the solution of a stochastic equation with discontinuous trajectories. *Lithuanian Mathematical Journal, 18*(1), 100–114.

[149] Priola, E. (2018). Davie's type uniqueness for a class of SDEs with jumps. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (Vol. 54, No. 2, pp. 694–725). Institut Henri Poincaré.

[150] Priola, E. (2020). On Davie's uniqueness for some degenerate SDEs. *Theory of Probability and Mathematical Statistics, 103*, 41–58.

[151] Priola, E., & Zabczyk, J. (2011). Structural properties of semilinear SPDEs driven by cylindrical stable processes. *Probability theory and related fields, 149*(1-2), 97–137.

[152] PROTTER, P. E. (1985). Volterra equations driven by semimartingales. *The Annals of Probability, 13*(2), 519–530.

[153] PROTTER, P. E. (2005). *Stochastic integration and differential equations*. Springer, Berlin, Heidelberg.

[154] RAFFAELLI, I., SCOTTI, S., & TOSCANO, G. (2022). Hawkes-driven stochastic volatility models: goodness-of-fit testing of alternative intensity specifications with S&P500 data. *Annals of Operations Research*, 1–19.

[155] RECCHIONI, M. C., IORI, G., TEDESCHI, G., & OUELLETTE, M. S. (2021). The complete Gaussian kernel in the multi-factor Heston model: Option pricing and implied volatility applications. *European Journal of Operational Research, 293*(1), 336–360.

[156] RHOADS, R. (2011). *Trading VIX Derivatives: Trading and Hedging Strategies Using VIX Futures, Options, and Exchange-Traded Notes.* Wiley.

[157] RIEDLE, M. (2015). Ornstein-Uhlenbeck processes driven by cylindrical Lévy processes. *Potential Analysis, 42*(4), 809–838.

[158] RIEDLE, M. (2018). Stable cylindrical Lévy processes and the stochastic Cauchy problem. *Electronic Communications in Probability, 23*, 1–12.

[159] RØMER, S. E. (2022a). Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. *Quantitative Finance, 22*(10), 1805–1838.

[160] RØMER, S. E. (2022b). Hybrid multifactor scheme for stochastic Volterra equations with completely monotone kernels. *Available at SSRN 3706253.*

[161] ROSENBAUM, M., & ZHANG, J. (2021). Deep calibration of the quadratic rough Heston model. *Risk, cutting-hedge*, `https://www.risk.net/cutting-edge/7954656/deep-calibration-of-the-quadratic-rough-heston-model`.

[162] ROSESTOLATO, M. (2018). A note on stochastic Fubini's theorem and stochastic convolution. *Preprint arXiv: 1606.06340v2.*

[163] RUDIN, W. (1987). *Real and complex analysis* (3rd ed.). McGraw-Hill International Editions: Mathematics series.

[164] SATO, K.-I. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press.

[165] SCHMELZLE, M. (2010). Option pricing formulae using Fourier transform: Theory and application. *Preprint*, `https://pfadintegral.com/docs/Schmelzle2010%20Fourier%20Pricing.pdf`.

[166] SEPP, A. (2008). Pricing options on realized variance in the Heston model with jumps in returns and volatility. *Journal of Computational Finance, 11*, 33–70.

[167] SITU, R. (2006). *Theory of stochastic differential equations with jumps and applications: mathematical and analytical techniques with applications to engineering.* Springer Science & Business Media.

[168] Song, Z., & Xiu, D. (2012). A tale of two option markets: State-price densities implied from S&P 500 and VIX option prices. *Unpublished working paper. Federal Reserve Board and University of Chicago.*

[169] Stroock, D. W., & Varadhan, S. S. (1997). *Multidimensional diffusion processes* (Vol. 233). Springer Science & Business Media.

[170] Święch, A., Zabczyk, J. (2016). Integro–PDE in Hilbert spaces: existence of viscosity solutions. *Potential Analysis, 45,* 703–736.

[171] Todorov, V., & Tauchen, G. (2011). Volatility jumps. *Journal of Business & Economic Statistics, 29*(3), 356–371.

[172] Trevisan, D. (2016). Well–posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electronic Journal of Probability, 21*(22), 1–41.

[173] Veretennikov, A. Y. (1981). On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mathematics of the USSR–Sbornik, 39*(3), 387.

[174] Viens, F., & Zhang, J. (2019). A martingale approach for fractional Brownian motions and related path dependent PDEs. *The Annals of Applied Probability, 29*(6), 3489–3540.

[175] Wang, F. Y., Xu, L., & Zhang, X. (2015). Gradient estimates for SDEs driven by multiplicative Lévy noise. *Journal of Functional Analysis, 269*(10), 3195–3219.

[176] Xia, W. (2020). Power-type derivatives for rough volatility with jumps. *Preprint arXiv: 2008.10184.*

[177] Yong, J., & Zhou, X. Y. (1999). *Stochastic controls: Hamiltonian systems and HJB equations* (Vol. 43). Springer Science & Business Media.

[178] Zhang, X. (2013). Derivative formulas and gradient estimates for SDEs driven by α-stable processes. *Stochastic Processes and their Applications, 123*(4), 1213–1228.