DISSERTATION | DOCTORAL THESIS

Titel | Title

Signature of càdlàg rough paths:
universal properties and applications in finance

verfasst von | submitted by

Dott.ssa Dott.ssa mag. Francesca Primavera

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 405
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Mathematik
Betreut von | Supervisor:	Univ.-Prof. Dipl.-Ing. Dr. Christa Cuchiero, Privatdoz. Dr sc. ETH Sara Papariello-Svaluto-Ferro, MSc ETH

Abstract

The theory of rough paths was introduced in Lyons (1998) as a way to understand and overcome the lack of continuity of the Itô-map, which describes the dependence of the solution of a stochastic differential equation (SDE) on the driving signal. Lyons observed that by enhancing the driving process with a finite number of higher-order objects, one can give a meaning in a *pathwise* sense to differential equations driven by *continuous* rough signals, such as Brownian motion. These enhanced or lifted paths are called rough paths. Furthermore, by equipping the space of rough paths with certain variation metrics, the Itô-map becomes continuous. These seminal ideas found many applications, in particular in stochastic analysis, providing a pathwise perspective and stability results to the theory of SDEs.

One important object in this theory is the *signature of a path,* whose original idea goes back to Chen (1957, 1977) but later on, it has been taken up in the context of rough paths and introduced for more general *rougher* signals. The signature can be viewed as an enhancement of a (rough) path through iterated integrals. Its importance is particularly due to its role as a well-suited feature map capturing essential path characteristics. Indeed, in various respects, linear functionals of the signature act similarly to polynomials on path space and can be viewed as a canonical selection of basis functions. This explains why these signature based methods become more and more popular in machine learning, econometrics and mathematical finance.

While in its early development, the theory of rough paths primarily focused on continuous paths, advancements by Friz and Shekhar (2017) have extended this framework to include càdlàg paths and introduced a notion of signature that encompasses also paths with jumps.

Building on this definition, one main goal of this thesis is to rigorously establish the role of the signature as a linear regression basis for path functionals in the càdlàg setting and to investigate the scope of its applicability from both theoretical and practical perspectives.

As a first step we prove a *universal approximation theorem* for continuous (with respect to the J_1-Skorokhod topology) *non-anticipative path functionals* in terms of linear functionals of the signature of càdlàg rough paths.

Then, as an important application, we define a new class of universal signature models based on an augmented Lévy process, which we call *Lévy-type signature models.* The approach that we follow consists in parameterizing the model itself or its semimartingale characteristics as linear functionals of the signature of an augmented Lévy process.

We show that these models extend the continuous signature models for asset prices as proposed so far, while still preserving universality and tractability properties.

Leveraging the universal approximation properties of the signature process, in the second part of the thesis we broaden the scope and introduce a more general class of jump-diffusions, which we call *Sig-jump-diffusions*, whose defining property is that the process semimartingale characteristics are parameterized by linear combinations of potentially infinitely many components of their own signature. For these processes, which includes the class of so-called *holomorphic jump-diffusions* introduced in the thesis, a significant extension of polynomial processes, we show that the expected signature, as well as the expected value of holomorphic functions of the process' marginals, can be computed via duality methods for stochastic processes. In particular, these quantities admit an analytic representation in terms of a solution of an (extended tensor algebra or sequence-valued) infinite-dimensional linear differential equation.

In the last part of the thesis, we derive a functional Itô-formula for non-anticipative maps of rough paths building on the approximating properties of the signature and as a byproduct of this result we show that sufficiently regular non-anticipative path functionals admit a functional Taylor expansion.

Kurzfassung

Die Theorie der rauen Pfade wurde von Lyons (1998) eingeführt, um die mangelnde Stetigkeit der Itô-Abbildung, d.h. die Abhängigkeit der Lösung einer stochastischen Differentialgleichung (SDE) vom Antriebssignal, zu verstehen und zu überwinden. Lyons stellte fest, dass man Differentialgleichungen, die von *kontinuierlichen* rauen Signalen wie der Brownschen Bewegung angetrieben werden, durch die Erweiterung des Antriebsprozesses mit einer endlichen Anzahl von Objekten höherer Ordnung eine Bedeutung in einem *pfadweisen* Sinne geben kann. Diese erweiterten oder "gelifteten" Pfade werden als raue Pfade bezeichnet. Darüber hinaus wird die Itô-Abbildung stetig, indem der Raum der rauen Pfade mit bestimmten Variationsmetriken ausgestattet wird. Diese bahnbrechenden Ideen fanden viele Anwendungen, insbesondere in der Stochastischen Analysis durch eine pfadweise Perspektive auf SDEs und entsprechende Stabilitätsresultate.

Ein wichtiges Objekt dieser Theorie ist die *Signatur eines Pfades*, deren ursprüngliche Idee auf Chen (1957, 1977) zurückgeht, die aber später im Zusammenhang mit rauen Pfaden aufgegriffen und für allgemeinere *rauere* Signale eingeführt wurde. Die Signatur kann als Erweiterung eines (rauen) Pfades durch iterierte Integrale betrachtet werden. Ihre Bedeutung liegt insbesondere in ihrer Rolle als gut geeignete Feature-Abbildung, die die wesentlichen Pfadeigenschaften erfasst. Tatsächlich verhalten sich lineare Funktionale der Signatur in vielerlei Hinsicht ähnlich wie Polynome im Pfadraum und können als kanonische Auswahl von Basisfunktionen betrachtet werden. Dies erklärt, warum Signatur-Methoden in den Bereichen, Machine Learning, Ökonometrie und Finanzmathematik immer beliebter werden.

Während sich die Theorie der rauen Pfade in ihrer frühen Entwicklung hauptsächlich auf stetige Pfade konzentrierte, haben neuere Fortschritte von Friz and Shekhar (2017) diesen Rahmen erweitert, um auch càdlàg Pfade zu umfassen, und einen Begriff der Signatur eingeführt, der auch für Pfade mit Sprüngen gilt.

Aufbauend auf dieser Definition der Signatur ist es ein Hauptziel dieser Arbeit, die Rolle der Signatur als lineare Regressionsbasis für Pfadfunktionale im càdlàg -Setting rigoros zu klären und den Umfang ihrer Anwendbarkeit sowohl aus theoretischer als auch aus praktischer Sicht zu untersuchen.

Als ersten Schritt beweisen wir ein *universelles Approximationstheorem* für (bezüglich der J_1-Skorokhod-Topologie) stetige *nicht-antizipative Pfadfunktionale* durch linearer Funktionale der Signatur von rauen càdlàg Pfaden.

Dann definieren wir als wichtige Anwendung eine neue Klasse universeller Signaturmodelle basierend auf einem erweiterten Lévy-Prozess, die wir *Lévy-artige Signaturmodelle*

nennen. Der Ansatz, den wir verfolgen, besteht darin, das Modell selbst oder seine Semimartingal-Charakteristiken als lineare Funktionale der Signatur eines erweiterten Lévy-Prozesses zu parametrisieren. Wir zeigen, dass diese Modelle die bisher vorgeschlagenen stetigen Signaturmodelle für Aktienpreise erweitern und dabei die Universalitätseigenschaften sowie gute analytische Handhabbarkeit beibehalten.

Unter Ausnutzung der universellen Approximationseigenschaften des Signaturprozesses erweitern wir im zweiten Teil der Arbeit den Anwendungsbereich und führen eine allgemeinere Klasse von Sprungdiffusionen ein, die wir *Sig-Sprungdiffusion* nennen und deren definierende Eigenschaft darin besteht, dass die Semimartingal-Charakteristiken durch lineare Kombinationen von potenziell unendlich vielen Komponenten ihrer eigenen Signatur parametrisiert werden. Für diese Prozesse, die insbesondere die hier eingeführte Klasse sogenannter *holomorpher Prozesse* enthält, eine bedeutende Erweiterung von polynomialen Prozessen, zeigen wir, dass die erwartete Signatur sowie der Erwartungswert holomorpher Funktionen der Prozessmarginalen über Dualitätsmethoden berechnet werden können. Insbesondere lassen sich diese Größen analytisch in Form einer Lösung einer (erweiterten Tensor Algebra- bzw. folgenwertigen) unendlichdimensionalen linearen Differentialgleichung darstellen.

Im letzten Teil der Arbeit leiten wir eine funktionale Itô-Formel für nicht-antizipative Abbildungen von rauen Pfaden her, die auf den approximierenden Eigenschaften der Signatur aufbaut. Als Nebenprodukt dieses Resultats zeigen wir eine funktionale Taylor-Entwicklung für ausreichend reguläre nicht-antizipative Pfadfunktionale.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported and contributed to the completion of this thesis.

First and foremost, I am profoundly grateful to my advisors, Christa Cuchiero and Sara Svaluto-Ferro, whose invaluable guidance, expertise, and encouragement have been instrumental throughout this research. Their unwavering support and insightful feedback have significantly shaped this work.

My thanks also go to Xin Guo for inviting me to the University of Berkeley, California, to conduct two months of intensive research. Her exceptional kindness and expert advice significantly enriched this experience.

I also want to thank Fred Espen Benth and Peter Friz for accepting the task of reviewing my thesis.

A special thanks go to my colleagues and friends at the Department of Statistics and Operations Research of the University of Vienna for all the coffees and the aperos, and particularly to Janka, Luca, Camilla, Florian, Yuan, Guido, Benedict, and Stefan.

Finally, I would like to express my gratitude towards my parents for their support throughout all these years and to Giangiuseppe for all his patience and constant encouragement.

Contents

Introduction

The theory of rough paths was introduced in Lyons (1998) as a way to understand and overcome the lack of continuity of the Itô-map, which describes the dependence of the solution of a stochastic differential equation on the driving signal. Lyons observed that by enhancing the driving process with a finite number of higher-order objects, one can give a meaning in a *pathwise* sense to differential equations driven by *continuous* rough signals, such as Brownian motion. These enhanced or lifted paths are called rough paths. Furthermore, equipping the space of rough paths with certain variation metrics, the Itô-map becomes continuous. These seminal ideas found many applications, in particular in stochastic analysis, providing a pathwise perspective and stability results to the theory of SDEs, McKean–Vlasov equations, as well as SPDEs (see, e.g., Cass and Lyons (2015), Friz and Hairer (2014)). For instance, SDEs can be solved in a pathwise way with continuous dependence on the driving semimartingale by choosing a priori a higher-order enhancement of the semimartingale driver.

One important object in this theory is the *signature of a path,* whose original idea goes back to Chen (1957, 1977) (see also Magnus (1954), Föllmer (1981), Sussmann (1988)), but later on, it has been taken up in the context of rough paths and introduced for more general *rougher* signals.

The signature $\mathbb{X}$ of a *continuous weakly geometric*[1] p-rough path $\mathbf{X}$, for $p \geqslant 1$, defined on the interval $[0,1]$, can be understood as an enhancement of the path with canonically defined higher-order iterated integrals, i.e.,

$$\mathbb{X}_t := \Big(1, \int_0^t d\mathbf{X}_{t_1}, \int_0^t \int_0^{t_1} d\mathbf{X}_{t_2} \otimes d\mathbf{X}_{t_1}, \dots, \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} d\mathbf{X}_{t_n} \otimes \cdots \otimes d\mathbf{X}_{t_1}, \dots \Big)$$

for $t \in [0,1]$. By "canonically," we mean that the signature is the unique path that projects onto $\mathbf{X}$ while maintaining specific algebraic and regularity conditions, the latter expressed in terms of certain variation metrics. This is precisely the essence of the fundamental *Lyons' extension theorem* appeared in Lyons (1998).

A signature, or more precisely the set of *linear functionals of the signature*, owes its importance in particular to the fact that it serves as a well-suited feature map capturing essential path characteristics. Indeed, in various respects, linear functionals of the signature act similarly to polynomials on path space and can be viewed as a canonical selection of basis functions. For instance, as a consequence of the Stone-Weierstrass

[1] The notion of the signature can be introduced for any continuous rough path (see Theorem 2.21 in Lyons (1998)). Since the main results of the present thesis concern the subset of *weakly geometric* rough paths, we focus on this specific set.

theorem, continuous functionals on (rough) path space, when restricted to compact sets, can be approximated by linear functionals of the signature, that is, by linear combinations of the above introduced iterated integrals (see Kalsi et al. (2020), Bayer et al. (2023)).

This explains why signature methods have been widely applied to treat streamed data in the context of machine learning and have recently also entered the field of stochastic modeling, and in particular mathematical finance, see e.g. Bühler et al. (2020); Kalsi et al. (2020); Arribas et al. (2021); Lyons et al. (2020); Liao et al. (2024); Bayer et al. (2023) and references therein. Indeed, they allow for data-driven modeling approaches, while stylized facts or first principles from mathematical finance can still be easily guaranteed.

Recent advancements by Friz and Shekhar (2017), Friz and Zhang (2017), and Chevyrev and Friz (2019) have extended the rough paths framework to include càdlàg paths, establishing a rigorous pathwise integration theory for this broader class of driving signals. Notably, Friz and Shekhar (2017) addresses the question of whether an analog of the notion of signature or, more precisely, an analog of Lyons' extension theorem can be established in the presence of jumps. To answer this question, Friz and Shekhar (2017) build on earlier ideas from Marcus (1978, 1981), which involve introducing an "artificial" additional time interval at each jump time of $\mathbf{X}$ and appropriately interpolating the state before and after the jump. In this work, we refer to this specific procedure as *Marcus transformation of a weakly geometric càdlàg rough path*. Combining these concepts with the theory of continuous paths leads to the formulation of the *minimal jumps extension theorem*, which asserts the following fact.

Every weakly geometric càdlàg p-rough path $\mathbf{X}$, for $p \geqslant 1$, admits a canonically defined extension path $\mathbb{X}$ up to a *minimal jump constraint*:

$$\Delta \mathbb{X}_t = \exp(\log^{([p])}(\Delta \mathbf{X}_t)), \quad \text{for all } t \in [0,1], \tag{0.0.1}$$

for exp and $\log^{([p])}$ denoting the exponential and the logarithm map on the extended and truncated tensor algebra, respectively[2]. The constraint (0.0.1) means that the jumps of the higher-order components of $\mathbb{X}$ are given as products of the jumps of $\mathbf{X}$ and combinations thereof. This minimal jump extension then constitutes the *signature in the càdlàg setting*. To properly introduce this notion with all details, we recall all relevant concepts from càdlàg rough path theory in Chapter 1.

Taking up the above definition of the signature, the first goal of this thesis is to make mathematically rigorous the role of the signature components as linear regression basis for path functionals in the càdlàg setting. This is mainly motivated by our second line of research: the study of signature-based models for finance that allow for the inclusion of jumps. As the stochastic analysis setting is our primary motivation, we focus on the signature of weakly geometric càdlàg p-rough paths for $p \in [1, 3)$, as it is what matters the most in this context.

To answer the first research question, in Chapter 2, which is based on Cuchiero

[2] We refer to Section 1.1, equations (1.1.3) and (1.1.2), for the precise definitions of exponential and logarithm maps, respectively.

et al. (2024d), we prove a *universal approximation theorem* (UAT) for continuous (with respect to the J_1-Skorokhod topology) *non-anticipative path functionals* in terms of linear functionals of the signature of càdlàg rough paths.

This is achieved by proceeding in two steps. First, we show that continuous functionals of càdlàg rough paths over the time interval $[0,1]$ can be uniformly approximated on compact sets of paths by linear functionals of the time-extended signature evaluated at the final time 1. Secondly, we extend this by considering functionals of the stopped paths, stopped at all deterministic times $t \in [0,1]$, which can thus be associated with non-anticipative path functionals, and prove that such functionals can be approximated uniformly over time and on compact sets of paths (see Theorem 2.2.13).

These results fill a gap in the literature and advance the established understanding of the signature of continuous paths to the càdlàg setting. Furthermore, the ingredients of the proof differ notably in technical complexity compared to the continuous case, stemming from the distinctive properties of the Skorokhod topology, with respect to which projection maps are in general not continuous (see Remark 2.2.14(ii)).

As previously stated, our principal motivation to treat this problem comes from signature-based models for finance that allow for the inclusion of jumps. Indeed, all signature models for asset prices that have been proposed so far (see Arribas et al. (2021); Cuchiero et al. (2023a, 2024b)) build on stochastic processes with continuous trajectories.

To extend these results to processes with càdlàg trajectories, we rely on the previously established UAT and leverage the close relationship between stochastic analysis and rough paths theory by considering the signature of càdlàg semimartingales. Specifically, in the second part of Chapter 2, we define a universal class of signature models for asset prices based on an augmented Lévy process X, which we call *Lévy-type signature models*. This marks the first application of the signature of càdlàg semimartingales as a signature-based asset price model in Mathematical Finance.

Our approach consists of parameterizing the model itself or its characteristics as linear functionals of the signature of the augmented process X, which we call *market's primary underlying process*. To analyze the tractability properties of these models, we first show that the truncated signature process of a generic multivariate Lévy process is a polynomial process (see Cuchiero et al. (2012); Filipović and Larsson (2020)), resulting in an easily computable formula (based on a matrix exponential) for the expectation of linear functionals of the signature's process $\mathbb{X}$.

This, in turn, allows us to derive explicit formulas for the expected signature of the price process in terms of the expected signature of X. Consequently, pricing sig-payoffs – i.e., linear functionals of the signature of the time-extended price process – becomes highly efficient. Not only that, but also hedging problems can be traced back to the signature of the primary Lévy process X. Indeed, we show that the quadratic hedging problem can be solved by computing the explicit form of the mean-variance hedging strategy in terms of the signature $\mathbb{X}$.

A key step to obtain all these results is to rewrite the price process itself as a linear functional of $\mathbb{X}$, a reformulation that we shall call *sig-model representation*. Given the presence of jumps, this step is not at all straightforward and is achieved through a meticulous analysis of each of the components of $\mathbb{X}$ (see Corollary 2.3.14 and, in particular, Theorem 2.3.17).

In Chapter 3, which is based on Cuchiero et al. (2024f), we explore a broader class of processes called *Sig-jump-diffusions*. In this class, the process' characteristics are parameterized by *linear combinations of potentially infinitely many components of their own signature*. This approach represents a significant generalization compared to the previous framework, which involves only a finite number of linear combinations. Moreover, the jump-diffusion's characteristics can depend on its own signature and not necessarily on the one of an exogenously given process. The chapter's primary contribution is the derivation of an explicit formula for the *expected signature* of these processes using duality theory in this stochastic context.

Historically, duality theory for Markov processes with respect to a duality function goes back to several contributions in the early fifties, e.g., Karlin and McGregor (1957), where classifications of birth and death processes are considered. Since then, it has been extended in several directions (see, e.g., Ethier and Kurtz (1986)) and applied in the areas of interacting particle systems, queuing theory, and population genetics.

The concept of dual processes can be formalized as follows: let $T > 0$ be some finite time-horizon and consider two time-homogeneous Markov processes $(X_t)_{t\in[0,T]}$ and $(U_t)_{t\in[0,T]}$ with respective state spaces S and U. Then X and U are *dual* with respect to some measurable function $H : S \times U \to \mathbb{R}$ if for all $x \in S$, $u \in U$ and $t \in [0, T]$

$$\mathbb{E}_x[H(X_t, u)] = \mathbb{E}_u[H(x, U_t)], \tag{0.0.2}$$

holds, and both the left and right-hand sides are well-defined. Here, $\mathbb{E}_x$ denotes the expected value for the Markov process X starting at $X_0 = x$ and similarly for U. Modulo technical conditions, the dual relation (0.0.2) holds if and only if the generators denoted by $\mathcal{A}$ and $\mathcal{B}$ satisfy

$$\mathcal{A}H(\cdot, u)(x) = \mathcal{B}H(x, \cdot)(u), \quad \text{for all } x \in S, u \in U, \tag{0.0.3}$$

(see e.g., Jansen and Kurt (2014)).

In the context of Sig-jump-diffusions, duality relations occur between the signature process and the solutions of an extended tensor algebra-valued linear ODE, providing an analytic representation of the expected signature of Sig-jump-diffusions, i.e.,

$$\mathbb{E}_{\mathbb{x}}[H(\mathbb{X}_t, \mathbf{u})] = H(\mathbb{x}, \mathbf{c}(t)),$$

for $t \in [0, T]$. Here $\mathbb{X}_0 = \mathbb{x}$, the (deterministic) dual process $(\mathbf{c}(t))_{t\in[0,T]}$ is a solution of a linear ODE in the extended tensor algebra with initial value $\mathbf{u}$, and $\mathbf{u}$ is such that the evaluation of the map $H(\cdot, \mathbf{u})$ at $\mathbb{X}_t$ is expressed as a convergent series of (possibly infinitely many) linear combinations of the components of $\mathbb{X}_t$ (see Theorem 3.3.8). These relations can also be recognized from the perspective of extended tensor algebra-valued polynomial processes, recently introduced in Cuchiero et al. (2023c). Indeed, our results extend the findings of the latter paper to the non-continuous setting.

One of the main difficulties in the analysis of this problem consists in characterizing the elements $\mathbf{u}$ for which the series of linear combinations of signature components $H(\mathbb{X}_t, \mathbf{u})$ is convergent as well as establishing integrability conditions for such quantities. This is already a delicate problem in the diffusion case. With the addition of a jump component, this becomes even more complex, see Section 3.3.

A class of processes for which explicit conditions can be derived consists of the class of polynomial processes, which are the object of interest in the last part of the chapter. We show that a large class of polynomial jump-diffusions can be embedded in the Sig-jump-diffusions framework. Additionally, their expected signature can be computed by solving a *finite-dimensional system of linear ODEs*, extending, therefore, the results derived in Chapter 2 for the class of Lévy processes.

Recognizing that any real-valued entire function (in the sense of Definition 4.D.1, restricted to $\mathbb{R}^d$)) of a generic process can also be interpreted as a convergent power series of (certain) linear combinations of its signature components (see Remarks 4.3.2 and 4.3.7(iv)) and utilizing complex analysis techniques, we can also derive precise convergence results and integrability conditions for a *a larger class of processes beyond polynomial ones.*

This is the topic of Chapter 4, which is based on Cuchiero et al. (2024e). Here, the main goal is to define and specify a class of processes X, which we call *holomorphic jump-diffusions*, such that for all $t \in [0, T]$,

$$\mathbb{E}_x[H(X_t, \mathbf{u})] = H(x, \mathbf{c}(t)). \tag{0.0.4}$$

Here $H(\cdot, \mathbf{u})$ is (the restriction to $\mathbb{R}^d$ of) a holomorphic function on $\mathbb{C}^d$, $\mathbf{u} = (\mathbf{u}_\alpha)_{\alpha \in \mathbb{N}_0^d}$ are the coefficients determining some of its power series representation and $(\mathbf{c}(t))_{t \in [0,T]}$ is a solution of a sequence-valued infinite dimensional linear ODE with initial condition $\mathbf{u}$. More precisely, we define *holomorphic jump-diffusions* as solutions to the martingale problem specified by the (extended) generator $\mathcal{A}$, which maps some subset of the holomorphic functions to holomorphic functions.

The fact that we consider *jump-diffusions* instead of just continuous processes makes things again rather involved. Indeed, already characterizing these subsets is due to the appearance of jumps not at all straightforward (see Theorem 4.3.6).

Under technical conditions involving, in particular, the existence of solutions to the sequence-valued linear ODEs and certain moment conditions, we then establish – what we call – the holomorphic formula (0.0.4). To apply this result, we establish further sufficient conditions which are verified for Lévy processes, affine processes, and certain non-polynomial jump diffusions on compact sets (see Section 4.3.5), which is one of the main contributions of this chapter.

The final research strand of this thesis is the formulation of a *functional (rough) Itô-formula* and *Taylor expansion* for non-anticipative maps of weakly geometric rough paths. It is indeed a natural question to wonder – alongside Stone-Weierstrass-type results – whether sufficiently regular non-anticipative path functionals admit a Taylor expansion in terms of signature components and coefficients that have an interpretation as derivatives.

We delve into these topics in Chapter 5, which is based on Cuchiero et al. (2024c). Here, relying again on rough path theory, we first prove the functional Itô-formula, and then by iteration, we derive a functional Taylor expansion. The significant advantages of adopting a (pathwise) rough paths perspective are twofold. Firstly, by examining functionals on the space of weakly geometric rough paths with specific variation metrics, we include a broad range of path functionals that were excluded in the seminal works by Dupire (2019) and Cont and Fournie (2013). Their functional Itô-formulas are developed within a probabilistic/Föllmer-type pathwise framework, where path spaces are equipped with variants of the uniform topology. These topologies are often too weak to ensure the continuity of many functionals, such as iterated integrals (see, e.g., the example in equation (5.1.1)), which are effectively incorporated into our framework.

Secondly, the stability results of rough paths theory allow the application of a new *proof technique based on the approximation properties of linear functionals of the signature*. Indeed, our approach consists of establishing an Itô-formula for these specific functionals and then extending it to more general maps via density arguments.

To elaborate further, this result is derived following the subsequent steps.

Inspired by the paper Chevyrev and Friz (2019) and in particular by the construction of the signature path described in the proof of the minimal jump extension theorem, we introduce the class of so-called *non-anticipative Marcus canonical* path functionals. Roughly speaking, these are all those maps that are invariant under Marcus transformations.

Then, adopting the Lie group valued paths point of view on the space of weakly geometric rough paths, we introduce a notion of *vertical derivatives for Marcus canonical paths functionals*, which is adapted both to the group structure of the paths and the Marcus property of the maps considered (see Definition 5.2.4). While this concept is inspired by the (Euclidean) version introduced by Dupire (2019), it differs significantly in one crucial aspect. Given the Marcus property of the functionals considered here, our vertical derivative can be interpreted as a derivative involving a *delayed vertical perturbation* of the original path. This aspect is fundamental to our definition, as it establishes a differential calculus for path functionals that permits *non-commutative derivative orders* (see Remark 5.2.23), a key aspect of the present setup (see Remarks 5.2.29 and 5.5.8(ii)).

In particular, with these notions of derivatives, a functional Itô-formula for linear functionals of the signature follows then just by the definition of the signature of weakly geometric càdlàg rough paths (see Lemma 5.2.28).

Relying on that, we then prove a *UAT for regular vertically differentiable maps of weakly geometric càdlàg rough path* via linear functionals of their signature, which holds both at the level of the functionals and their derivatives (see Theorem 5.3.15). Given the non-linear structure of the (Lie-type) vertical derivatives, deriving this result has proven to be a very delicate task, involving an intricate application of Nachbin-type theorems (see Nachbin (1949)), as well as some key ideas from Lie group theory.

Finally, a combination of the UAT for vertically differentiable path functionals and rough paths-type interpolation arguments yield the *(rough) functional Itô-formula* (see Theorems 5.4.1 and 5.4.4). The *functional Taylor expansion* follows then via an iterative application of the latter result (see Theorems 5.5.1 and 5.5.5).

1 Signature of weakly geometric càdlàg rough paths

This section introduces the concept of the signature of weakly geometric càdlàg rough paths. We start by establishing the algebraic framework intrinsic to rough path theory and revisiting the notion of weakly geometric càdlàg rough paths and rough integration for these paths. Next, we delve into the notion of signature. Since the signature for càdlàg paths builds upon the established framework for continuous paths, we first introduce this notion for continuous paths and then transition to the càdlàg setting. We conclude the section with a discussion on the signature of càdlàg semimartingales.

1.1 Algebraic setting

1.1.1 The tensor algebra

Fix $d \in \mathbb{N}$ and let $\mathbb{R}^d$ be the Euclidean space. The extended tensor algebra over $\mathbb{R}^d$ is defined by

$$T((\mathbb{R}^d)) := \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

where $(\mathbb{R}^d)^{\otimes n}$ denotes the n-fold tensor product of $\mathbb{R}^d$ with the convention $(\mathbb{R}^d)^{\otimes 0} := \mathbb{R}$. We equip $T((\mathbb{R}^d))$ with the standard addition $+$, tensor multiplication $\otimes$ and scalar multiplication. For $N \in \mathbb{N}$, the truncated tensor algebra is defined by

$$T^N(\mathbb{R}^d) := \bigoplus_{n=0}^{N} (\mathbb{R}^d)^{\otimes n}.$$

Elements of the extended and truncated tensor algebra bold face, e.g., $\mathbf{u} = (\mathbf{u}^{(n)})_{n=0}^{\infty} \in T((\mathbb{R}^d))$. Elements of the truncated tensor algebra instead, in boldface, e.g., $\mathbf{u} = (\mathbf{u}^{(n)})_{n=0}^{N} \in T^N(\mathbb{R}^d)$. Let $\pi_n : T((\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes n}$ be the map such that for $\mathbf{u} \in T((\mathbb{R}^d))$, $\pi_n(\mathbf{u}) = \mathbf{u}^{(n)}$, and $\pi_{\leqslant N} : T((\mathbb{R}^d)) \to T^N(\mathbb{R}^d)$ be such that for $\mathbf{u} \in T((\mathbb{R}^d))$, $\pi_{\leqslant N}(\mathbf{u}) = \mathbf{u} = (\mathbf{u}^{(n)})_{n=0}^{N}$. For $c \in \mathbb{R}$, set

$$T_c^N(\mathbb{R}^d) := \{\mathbf{u} \in T^N(\mathbb{R}^d) \colon \mathbf{u}^{(0)} = c\}.$$

The space $T_1^N(\mathbb{R}^d)$ is a Lie group under the tensor multiplication $\otimes$, truncated beyond level N. The neutral element with respect to $\otimes$ is $\mathbf{1} := (1, 0, \dots, 0) \in T_1^N(\mathbb{R}^d)$. Moreover, for any $\mathbf{u} = (\mathbf{1} + \mathbf{b}) \in T_1^N(\mathbb{R}^d)$, with $\mathbf{b} \in T_0^N(\mathbb{R}^d)$, its inverse is given by

$$\mathbf{u}^{-1} = \sum_{k=0}^{N} (-1)^k \mathbf{b}^{\otimes k}. \tag{1.1.1}$$

The exponential and logarithm maps are defined as follows:

$$\begin{aligned} \exp^{(N)} : T_0^N(\mathbb{R}^d) &\to T_1^N(\mathbb{R}^d) & \log^{(N)} : T_1^N(\mathbb{R}^d) &\to T_0^N(\mathbb{R}^d) \\ \mathbf{b} &\mapsto \mathbf{1} + \sum_{k=1}^{N} \frac{\mathbf{b}^{\otimes k}}{k!}, & \mathbf{1} + \mathbf{b} &\mapsto \sum_{k=1}^{N} (-1)^{k+1} \frac{\mathbf{b}^{\otimes k}}{k}, \end{aligned} \tag{1.1.2}$$

where the tensor multiplication is again always truncated beyond level N. We furthermore introduce the (non-truncated) exponential map, which is given by

$$\exp(\mathbf{x}) := 1 + \sum_{k=1}^{\infty} \frac{\mathbf{x}^{\otimes k}}{k!} \in T((\mathbb{R}^{d+1})), \tag{1.1.3}$$

for each $\mathbf{x} \in T((\mathbb{R}^{d+1}))$ such that $\pi_0(\mathbf{x}) = 0$.

We refer to Chapter 7 in Friz and Victoir (2010) for more details on these algebraic aspects.

1.1.2 Free nilpotent groups

We introduce the so-called *free step-N nilpotent Lie groups* over $\mathbb{R}^d$. Let $\mathfrak{g}^N(\mathbb{R}^d)$ be the *free step-N nilpotent Lie algebra* over $\mathbb{R}^d$, i.e.,

$$\mathfrak{g}^N(\mathbb{R}^d) := \{0\} \oplus \mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d] \oplus \cdots \oplus \underbrace{[\mathbb{R}^d, [\mathbb{R}^d, \ldots [\mathbb{R}^d, \mathbb{R}^d]]]}_{(N-1) \text{ brackets}} \subseteq T_0^N(\mathbb{R}^d), \tag{1.1.4}$$

where, for $\mathbf{u} \in T_0^M(\mathbb{R}^d)$, $1 \leqslant M \leqslant N-1$, $\mathbb{b}^{(1)} \in \mathbb{R}^d$, $[\mathbb{b}^{(1)}, \mathbf{u}] := \mathbb{b}^{(1)} \otimes \mathbf{u} - \mathbf{u} \otimes \mathbb{b}^{(1)}$.

The image of $\mathfrak{g}^N(\mathbb{R}^d)$ through the exponential map is a subgroup of $T_1^N(\mathbb{R}^d)$ with respect to $\otimes$. It is called *free step-N nilpotent Lie group* and is denoted by

$$G^N(\mathbb{R}^d) := \exp^{(N)}(\mathfrak{g}^N(\mathbb{R}^d)). \tag{1.1.5}$$

We equip it with the so-called Carnot-Caratheodory (CC) norm $\|\cdot\|_{CC}$ (see Definition and Theorem 7.32 in Friz and Victoir (2010)) and the induced (left-invariant) metric, denoted by d_{CC} (see Definition 7.41 in Friz and Victoir (2010)). Finally, we introduce the set of so-called *group-like elements*, defined via

$$G((\mathbb{R}^d)) := \{\mathbf{x} \in T((\mathbb{R}^d)) \mid \pi_{\leqslant N}(\mathbf{x}) \in G^N(\mathbb{R}^d) \text{ for all } N\}. \tag{1.1.6}$$

We refer to Chapter 7 in Friz and Victoir (2010) for more details on the specific group $G^N(\mathbb{R}^d)$ (see also Section 2 in Lyons et al. (2007)) and to Bonfiglioli et al. (2007) and Schmeding (2022) for a more general treatment of Lie groups.

1.1.3 Linear functionals on the extended tensor algebra

Let $I = (i_1, \dots, i_n)$ be a multi-index with entries in $\{1, \dots, d\}$. Denoting by $\epsilon_1, \dots, \epsilon_d$ the canonical basis of $\mathbb{R}^d$, we use the notation $|I| := n$ and $\epsilon_I := \epsilon_{i_1} \otimes \epsilon_{i_2} \otimes \cdots \otimes \epsilon_{i_n}$. Observe that $(\epsilon_I)_I$ is the canonical orthonormal basis of $(\mathbb{R}^d)^{\otimes n}$. Furthermore, we denote by $\epsilon_\varnothing$ the basis element of $(\mathbb{R}^d)^{\otimes 0}$ and set $|\varnothing| := 0$. We also set $I' := (i_1, \dots, i_{n-1})$ for $n > 1$, $I' = \varnothing$ for $n = 1$, $I'' := (I')'$ for $n > 1$, and use the convention $\epsilon_{I''} = 0$ for $n = 1$.

Given $\mathbf{x} \in T((\mathbb{R}^d))$, we write $\mathbf{x}_I := \langle \mathbf{x}, \epsilon_I \rangle$ and for each $\mathbf{u} \in T(\mathbb{R}^d)$ we set

$$\langle \mathbf{u}, \mathbf{x} \rangle := \sum_{|I| \geqslant 0} \mathbf{u}_I \mathbf{x}_I \in \mathbb{R}.$$

We define the set of *linear functionals* on $T((\mathbb{R}^d))$ as follows

$$\mathscr{L}_{\mathbb{R}^d} := \{\mathbb{x} \mapsto \langle \mathbf{u}, \mathbb{x} \rangle \colon \mathbf{u} \in T(\mathbb{R}^d)\}.$$

For $k \in \mathbb{N}$, we denote by $(\,\cdot\,)^{(k)} : T((\mathbb{R}^d)) \to (T((\mathbb{R}^d)))^{\otimes k}$ the shifts given by

$$\mathbf{u}_I^{(k)} := \sum_{|J| \geqslant 0} \mathbf{u}_{JI} \epsilon_J, \tag{1.1.7}$$

for each $|I| = k$, where JI denotes the concatenation of the multi indices J and I. For notational convenience, we also write $\mathbf{u}^{(0)} := \mathbf{u}$. Note that setting $\mathcal{V}^d := \big(T((\mathbb{R}^d))\big)^{\otimes d}$

$$(\mathbf{u}^{(0)}, \mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots) \in T((\mathcal{V}^d)).$$

For two multi-indices $I \in \{1, \dots, d\}^{|I|}$, $J \in \{1, \dots, d\}^{|J|}$, and two indices of length 1, $a, b \in \{1, \dots, d\}$, the shuffle product $\sqcup\!\sqcup$ is defined recursively by

$$\begin{aligned} &I \sqcup\!\sqcup \varnothing = \varnothing \sqcup\!\sqcup I = I, \\ &(I, a) \sqcup\!\sqcup (J, b) = ((I \sqcup\!\sqcup (J, b)), a) + (((I, a) \sqcup\!\sqcup J), b), \end{aligned}$$

where (I, a) denotes the concatenation of multi-indices.

Via the shuffle product every polynomial on the set of group-like elements $G((\mathbb{R}^d))$ (see equation (1.1.6)) admits a linear representative. More precisely, for $\mathbf{x} \in G((\mathbb{R}^d))$ and two multi-indices $I \in \{1, \dots, d\}^{|I|}$, $J \in \{1, \dots, d\}^{|J|}$, it holds that

$$\langle \epsilon_I, \mathbf{x} \rangle \langle \epsilon_J, \mathbf{x} \rangle = \langle \epsilon_I \sqcup\!\sqcup \epsilon_J, \mathbf{x} \rangle \tag{1.1.8}$$

where $\epsilon_I \sqcup\!\sqcup \epsilon_J := \sum_{k=1}^K \epsilon_{I_k}$ where K, I_k for $k = 1, \dots, K$ are determined via $I \sqcup\!\sqcup J = \sum_{k=1}^K I_k$. Moreover, for each $\mathbf{x} \in G((\mathbb{R}^d))$ and each multi-index $|I| \geqslant 1$ it holds

$$\sum_{J \colon J^{ord} = I^{ord}} \langle \epsilon_J, \mathbf{x} \rangle = \frac{1}{I!} \langle \epsilon_{i_1}, \mathbf{x} \rangle \cdots \langle \epsilon_{i_{|I|}}, \mathbf{x} \rangle, \tag{1.1.9}$$

where I^{ord} is the multi-index obtained by sorting the entries of I and $I! := \prod_{k=1}^d I(k)!$ for $I(k) := \sum_{j=1}^{|I|} 1_{\{i_j = k\}}$.

Finally, for $\mathbf{x} \in T((\mathbb{R}^d))$, we say that $\mathbf{x}$ is *symmetric* if $\mathbf{x}_I = \mathbf{x}_{I^{\mathrm{ord}}}$ for every $|I| \geqslant 0$.

1.2 Weakly geometric càdlàg rough paths and rough integration

1.2.1 Space of paths and two-parameter functions

Throughout, we denote by $C([0,1],E)$ and $D([0,1],E)$ the space of continuous and càdlàg maps (paths), respectively, from the interval $[0,1]$ into a metric space (E,d) equipped with metric d. For $t \in (0,1]$, we denote a partition of $[0,t]$ by $\pi_{[0,t]} = \{0 = t_0 < t_1 < \cdots < t_k = t\}$, and write $\sum_{t_i \in \pi_{[0,t]}}$ for the summation over all points in $\pi_{[0,t]}$. The mesh size of $\pi_{[0,t]}$ is given by $|\pi_{[0,t]}| := \max\{t_{i+1} - t_i \ : \ i = 0, \ldots, k-1\}$. For $p > 0$, we define the p-variation of a path $X \in D([0,1],E)$ by

$$\|X\|_{p\text{-}var} := \sup_{\pi_{[0,1]} \subset [0,1]} \left(\sum_{t_i \in \pi_{[0,t]}} d(X_{t_i}, X_{t_{i+1}})^p \right)^{\frac{1}{p}}. \tag{1.2.1}$$

If X takes values in a vector space, for $s,t \in [0,1]$, $s \leqslant t$, we use the shortcut $X_{s,t} := X_t - X_s$ and denote the jumps by $\Delta X_t := \lim_{s \nearrow t} X_{s,t}$. The space of continuous and càdlàg paths of finite p-variation are denoted respectively by $C^p([0,1],E)$ and $D^p([0,1],E)$. We endow these spaces with the q-variation pseudometric, for $q \geqslant p$, defined via

$$d_q(X,Y) := \sup_{\pi_{[0,1]} \subset [0,1]} \left(\sum_{t_i \in \pi_{[0,1]}} d(X_{t_i,t_{i+1}}, Y_{t_i,t_{i+1}})^q \right)^{\frac{1}{q}}, \tag{1.2.2}$$

for all $X, Y \in D^p([0,1],E)$. Additionally, we also consider two-parameter functions $A : \Delta_1 \to V$, where $(V, \|\cdot\|)$ is a normed vector space and $\Delta_1 := \{(s,t) \in [0,1] \times [0,1] \ : \ s \leqslant t\}$. As well as for paths, the notion of p-variation is valid for such functions and defined as follows

$$\|A\|_{p\text{-}var} := \sup_{\pi_{[0,1]} \subset [0,1]} \left(\sum_{t_i \in \pi_{[0,1]}} \|A_{t_i,t_{i+1}}\|^p \right)^{\frac{1}{p}}.$$

Moreover, for $q \geqslant p$, we similarly set $d_q(A,A') := \|A - A'\|_{q\text{-}var}$ for all $A, A' : \Delta_1 \to V$ for which $\|A\|, \|A'\| < \infty$.

We stress that if X is a path with values in a vector space V, that is $X \in D([0,1],V)$, then $X_{s,t}$ denotes the increment $X_t - X_s$. Instead, if A is a two-parameter function defined on Δ_1, $A_{s,t}$ denotes the evaluation of A at the pair of times $(s,t) \in \Delta_1$.

Finally, we call a map $\phi : [0,1] \to [0,1]$ a *time reparametrization* if it is increasing and bijective. We say that a path $X \in D([0,1],E)$ is a time reparametrization of some $Y \in D([0,1],E)$ if $Y = X_\phi$, for some ϕ time reparametrization.

1.2.2 Weakly geometric càdlàg rough paths

Let $C([0,1], G^N(\mathbb{R}^d))$ and $D([0,1], G^N(\mathbb{R}^d))$ be the space of continuous and càdlàg maps (paths), respectively, from the interval $[0,1]$ into $(G^N(\mathbb{R}^d), d_{CC})$.

For $\mathbf{X} \in D([0,1], G^N(\mathbb{R}^d))$, $s,t \in [0,1]$, $s \leqslant t$, the path increments are defined via

$$\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t \tag{1.2.3}$$

and the jumps by $\Delta \mathbf{X}_t := \lim_{s \nearrow t} \mathbf{X}_{s,t}$. For $p > 0$, we denote by $[p]$ its entire part, and by $C^p([0,1], G^N(\mathbb{R}^d))$ and $D^p([0,1], G^N(\mathbb{R}^d))$ the subspaces of $C([0,1], G^N(\mathbb{R}^d))$ and $D([0,1], G^N(\mathbb{R}^d))$, respectively, consisting of paths of finite p-variation, computed as in (1.2.1) with respect to the d_{CC} metric. We endow these spaces with the q-variation pseudometric, for $q \geqslant p$, defined as in (1.2.2).

Càdlàg paths of finite p-variation with values in the specific group $G^{[p]}(\mathbb{R}^d)$ are said weakly geometric càdlàg p-rough path. We formalize this notion in the definition below.

In this work, we restrict the presentation only to paths of finite p-variation for $p \in [1,3)$, as this is what matters the most in stochastic analysis settings, which is our primary motivation. However, we remark that it naturally extends to paths of arbitrary low regularity.

Definition 1.2.1. Let $p \in [1,3)$. A càdlàg path $\mathbf{X} : [0,1] \to G^{[p]}(\mathbb{R}^d)$ is said to be a *weakly geometric càdlàg p-rough path over $\mathbb{R}^d$* if

$$\|\mathbf{X}\|_{p\text{-}var} < \infty.$$

We denote the space of such paths by $D^p([0,1], G^{[p]}(\mathbb{R}^d))$. For $p \geqslant 2$, a weakly geometric càdlàg p-rough path is said to be *Marcus-like* if for all $t \in [0,1]$,

$$\log^{(2)}(\Delta \mathbf{X}_t) \in \{0\} \oplus \mathbb{R}^d \oplus \{0\}.$$

Remark 1.2.2.
(i) Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and notice that by definition for all $t \in [0,1]$,

$$\mathbf{X}_t := \begin{cases} \exp^{(1)}(X_t), & \text{if } p \in [1,2), \\ \exp^{(2)}(X_t, A_t), & \text{if } p \in [2,3), \end{cases}$$

for $\exp^{(1)}$, $\exp^{(2)}$ given in equation (1.1.2) and

$$X : [0,1] \to \mathbb{R}^d, \qquad A : [0,1] \to [\mathbb{R}^d, \mathbb{R}^d]$$

càdlàg paths.

(ii) For $p \in [1,2)$, a weakly geometric càdlàg p-rough path $\mathbf{X}$ is nothing else than a $\mathbb{R}^d$-valued càdlàg path with finite p-variation. Indeed, $G^1(\mathbb{R}^d) = \{1\} \oplus \mathbb{R}^d$ and the CC distance on $G^1(\mathbb{R}^d)$ is by definition equal to the Euclidean one on $\mathbb{R}^d$.

(iii) Let $p \in [2,3)$, and $\mathbf{X}$ a weakly geometric càdlàg p-rough path. The following conditions hold true.

a) Since the topology induced by the CC distance on $G^2(\mathbb{R}^d)$ coincides with the Euclidean one on $\mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$ (see Corollary 7.16 in Friz and Victoir (2010)), the map

$$[0,1] \ni t \longmapsto (X_{0,t}, \mathbb{X}^{(2)}_{0,t}) := (\pi_1(\mathbf{X}_{0,t}), \pi_2(\mathbf{X}_{0,t})) \in \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$$

is càdlàg.

b) Setting for $(s,t) \in \Delta_1$,

$$
\begin{aligned}
(1, X_{s,t}) &:= (1, \pi_1(\mathbf{X}_{s,t})) && \text{if } p \in [1,2), \\
(1, X_{s,t}, \mathbb{X}^{(2)}_{s,t}) &:= (1, \pi_1(\mathbf{X}_{s,t}), \pi_2(\mathbf{X}_{s,t})) && \text{if } p \in [2,3),
\end{aligned}
$$

by definition of path increments (see equation (1.2.3)), $X_{s,t} = X_t - X_s$ and

$$\mathbb{X}^{(2)}_{s,t} = \mathbb{X}^{(2)}_{0,t} - \mathbb{X}^{(2)}_{0,s} - X_{0,s} \otimes X_{s,t}.$$

This implies that while $X_{s,t}$ corresponds to the increment of the càdlàg path $X_{0,\cdot}$ between s and t, $\mathbb{X}^{(2)}_{s,t}$ is instead the evaluation at $(s,t) \in \Delta_1$ of a two parameter function $\mathbb{X}^{(2)} : \Delta_1 \to (\mathbb{R}^d)^{\otimes 2}$.

c) By definition of $G^2(\mathbb{R}^2)$-valued path, it also holds that

$$Sym(\mathbb{X}^{(2)}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t},$$

where $Sym(\mathbb{X}^{(2)}_{s,t})$ denotes the symmetric part of the $\mathbb{X}^{(2)}_{s,t}$.

d) By Theorem 7.44 in Friz and Victoir (2010),

$$\|X\|_{p\text{-}var} + \|\mathbb{X}^{(2)}\|^{1/2}_{p/2\text{-}var} < \infty,$$

where $\|\mathbb{X}^{(2)}\|_{p/2-var}$ denotes the $p/2$-variation of the two-parameter function $\mathbb{X}^{(2)}$.

Based on the preceding discussion, we can conclude that for a weakly geometric càdlàg p-rough path $\mathbf{X}$, it is always feasible to specify a pair $(X, \mathbb{X}^{(2)})$ that meets the criteria (iii)a, (iii)b, (iii)c, (iii)d. One can show that the converse is also true, leading to two equivalent definitions of weakly geometric càdlàg p-rough path.

(iv) For $p \in [2,3)$, *Marcus-like* means that for all $t \in [0,1]$,

$$\Delta \mathbf{X}_t = \lim_{s \nearrow t} (1, X_{s,t}, \mathbb{X}^{(2)}_{s,t}) = (1, \Delta X_t, \frac{1}{2}(\Delta X_t)^{\otimes 2}).$$

Assumption 1.2.3. *Throughout, unless explicitly mentioned, we let* $p \in [1,3)$.

1.2.3 Rough integration

This section reviews the fundamental concepts of rough integration theory in the càdlàg setting, initially explored in Friz and Shekhar (2017) and further developed in Friz and Zhang (2017). First, we set up the essential ingredients necessary for introducing the main concepts. Then, we discuss Young integration theory, which is sufficient for dealing with paths of finite p-variation for $p \in [1, 2)$. Finally, we advance to the more general level 2 rough integration theory, which handles paths with lower regularity. Throughout, we write $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)$ for the space of linear maps from $\mathbb{R}^d$ into $\mathbb{R}^m$, for some $m \in \mathbb{N}$.

The following definition is crucial for understanding the convergence of the (compensated) Riemann sums determining the Young and the rough integral.

Definition 1.2.4. For $\Xi : \Delta_1 \to \mathbb{R}^m$ and a partition $\pi_{[0,1]}$ of the interval $[0, 1]$, we say that the quantity $\sum_{s_i \in \pi_{[0,1]}} \Xi_{s_i, s_{i-1}}$ converges to $K \in \mathbb{R}^d$ in the

(i) Mesh Riemann-Stieltjes (MRS) sense: if for any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $\pi_{[0,1]}$ with $|\pi_{[0,1]}| < \delta$, it holds $|\sum_{s_i \in \pi_{[0,1]}} \Xi_{s_i, s_{i-1}} - K| \leqslant \varepsilon$.

(ii) Refinement Riemann-Stieltjes (RRS) sense: if for any $\varepsilon > 0$ there exists a partition $\pi^{\varepsilon}_{[0,1]}$ such that for any refinement $\pi_{[0,1]}$ of $\pi^{\varepsilon}_{[0,1]}$, it holds that $|\sum_{s_i \in \pi_{[0,1]}} \Xi_{s_i, s_{i-1}} - K| \leqslant \varepsilon$.

Remark 1.2.5. For more details on these kinds of convergence, we refer to Friz and Zhang (2017) and Dudley et al. (1998) and the discussion therein. We just recall that if MRS and RRS limits exist, they are unique and that MRS implies RRS convergence, with the same limit.

Young integration Let $p \in [1, 2)$. Recall from Remark 1.2.2(iii) that for $(s, t) \in \Delta_1$ and $\mathbf{X} \in D^p([0, 1], G^1(\mathbb{R}^d))$, we set $\mathbf{X}_{s,t} = (1, X_{s,t})$.

Proposition 1.2.6. (*Proposition* 2.4 in Friz and Zhang (2017)). Let $\mathbf{X} \in D^p([0, 1], G^1(\mathbb{R}^d))$ and $Y \in D^{p'}([0, 1], \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m))$, for some p' such that $\frac{1}{p} + \frac{1}{p'} > 1$. Then, for all $t \in [0, 1]$, the limit

$$\int_0^t Y_{s^-} d\mathbf{X}_s := \lim_{(MRS)|\pi_{[0,t]}| \to 0} \sum_{s_i \in \pi_{[0,t]}} Y_{s_i} X_{s_i, s_{i+1}} \tag{1.2.4}$$

exists. We call the limit in (1.2.4) the *Young integral* of Y with respect to $\mathbf{X}$.

Remark 1.2.7. (i) By definition of MRS convergence (see Definition 1.2.4), the limit (1.2.4) exists along every sequence of partition $\pi_{[0,t]}$ of the interval $[0, t]$ with mesh size $|\pi_{[0,t]}|$ tending to zero and it does not depend on the choice of the sequence of partition.

(ii) If $p = 1$, the path $\mathbf{X}$ has bounded variation, and in this case the integral in (1.2.4) coincides with the Lebesgue-Stieltjes integral $\int_{(0,t]} Y_s \nu_{\mathbf{X}}(ds)$, for $\nu_{\mathbf{X}}([a, b]) = \mathbf{X}(b) - \mathbf{X}(a)$, $(a, b) \in [0, t] \times [0, t]$, $a \leqslant b$ (see Proposition 1.9 in Friz and Zhang (2017)).

Corollary 1.2.8. *(see Proposition 2.4 in Friz and Zhang (2017)) Let* $\mathbf{X} \in D^p([0, 1], G^1(\mathbb{R}^d))$. *For any* $Y \in D^{p'}([0, 1], \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)$, *for some* p' *such that* $\frac{1}{p} + \frac{1}{p'} > 1$, $\int_0^{\cdot} Y_{s^-} d\mathbf{X}_s \in D^p([0, 1], \mathbb{R}^m)$.

Level 2 **rough integration** We mainly follow the presentation given in Allan et al. (2023a). Let $p \in [2,3)$ and $p' \geqslant p$ be such that

$$\frac{2}{p} + \frac{1}{p'} > 1. \tag{1.2.5}$$

Define $r \geqslant 1$ by the relation

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{p'}. \tag{1.2.6}$$

Notice that $1 \leqslant \frac{p}{2} \leqslant r < p \leqslant p' < \infty$.

Definition 1.2.9. Let $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$. A pair of càdlàg paths (Y, Y') is called a *controlled rough path* (with respect to $X := \pi_1(\mathbf{X})$) if $Y \in D^p([0,1], \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m))$, and for some p', r satisfying (1.2.5), (1.2.6), it holds that $Y' \in D^{p'}([0,1], \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m))$, and $R : \Delta_1 \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^m)$, defined by $R_{s,t} := Y_{s,t} - Y'_s X_{s,t}$, for $(s,t) \in \Delta_1$, has finite r-variation. We denote the space of such controlled paths by $\mathcal{V}^{p',r}_{\mathbf{X}}$.

Remark 1.2.10. Notice that a pair of controlled rough paths (Y, Y') as in Definition 1.2.9 is controlled with respect to $X := \pi_1(\mathbf{X})$. Here $\mathbb{X}^{(2)} := \pi_2(\mathbf{X})$ does not play any role. Nevertheless, with some abuse of notation, we denote the space of such paths by $\mathcal{V}^{p',r}_{\mathbf{X}}$.

Recall from Remark 1.2.2(iii) that for $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$ and $(s,t) \in \Delta_1$, we set $\mathbf{X}_{s,t} = (1, X_{s,t}, \mathbb{X}^{(2)}_{s,t})$.

Proposition 1.2.11. (*Proposition* 2.4 in Allan et al. (2023b)) Let $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$ and $(Y, Y') \in \mathcal{V}^{p',r}_{\mathbf{X}}$, for some $p' \geqslant p$ such that $\frac{2}{p} + \frac{1}{p'} > 1$. Then, for all $t \in [0,1]$, the limit

$$\int_0^t Y_{s^-} d\mathbf{X}_s := \lim_{(MRS)|\pi_{[0,t]}| \to 0} \sum_{s_i \in \pi_{[0,t]}} \big(Y_{s_i} X_{s_i, s_{i+1}} + Y'_{s_i} \mathbb{X}^{(2)}_{s_i, s_{i+1}}\big), \tag{1.2.7}$$

exists. We call the limit in (1.2.7) the *rough integral* of (Y, Y') with respect to $\mathbf{X}$.

Remark 1.2.12. (i) The observation in Remark 1.2.7(i) applies also here.

(ii) Recall that $\mathcal{L}(\mathbb{R}^d; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^m))$ can be identified with the space $\mathcal{L}(\mathbb{R}^{d \times d}; \mathbb{R}^m)$. This allows us to make sense of the product $Y'\mathbb{X}^{(2)}$ in equation (1.2.7), which then takes values in $\mathbb{R}^m$.

Corollary 1.2.13. *(Remark 2.6 in Allan et al. (2023b)) Let* $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$*. For any* $(Y, Y') \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$*, with* $p' > p$ *such that* $\frac{1}{p'} + \frac{2}{p} > 1$*, the pair* $(\int_0^\cdot Y_{s^-} d\widehat{\mathbf{X}}_s, Y) \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$.

Notation 1.2.14. Throught the work, whenever the involved paths are continuous, we write $\int_0^\cdot Y_s d\mathbf{X}_s$ in place of (1.2.4) or (1.2.7).

Consistency between Young and rough integration We conclude the section with a discussion on the consistency between the Young and the rough integral. To this end, fix $X \in D^1([0,1], G^1(\mathbb{R}^d))$ and for all $t \in [0,1]$, set

$$\mathbf{X}_t := \left(1, X_t - X_0, \int_0^t X_{0,s^-} \otimes dX_s + \frac{1}{2} \sum_{0<s\leqslant t} (\Delta X_s)^{\otimes 2}\right), \tag{1.2.8}$$

where $\int_0^\cdot X_{0,s^-} \otimes dX_s$ denotes the Young integral of $X_{0,\cdot}$ with respect to X. By Corollary 1.2.8 and the geometric properties of the Young integral (see e.g., Proposition 2.4 in Friz and Zhang (2017)), we get that $\mathbf{X} \in D^1([0,1], G^2(\mathbb{R}^d))$. Therefore, since for any $p > 1$, $D^1([0,1], G^2(\mathbb{R}^d)) \subseteq D^p([0,1], G^2(\mathbb{R}^d))$, $\mathbf{X}$ can be interpreted as a weakly geometric càdlàg p-rough path, for any $p \in [2,3)$.

Lemma 1.2.15. Let $p \in [2,3)$ and $\mathbf{X} \in D^p([0,1]; G^2(\mathbb{R}^d))$. Assume that $\mathbf{X}$ is defined as in equation (1.2.8), for some $X \in D^1([0,1], G^1(\mathbb{R}^d))$, and let $(Y, Y') \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$, with $p' > p$ such that $\frac{1}{p'} + \frac{2}{p} > 1$. Then,

$$\int_0^t Y_{s^-} d\mathbf{X}_s = \int_0^t Y_{s^-} dX_s + \frac{1}{2} \sum_{0<s\leqslant t} Y'_{s^-} (\Delta X_s)^{\otimes 2}, \tag{1.2.9}$$

where the integral on the LHS denotes the rough integral of (Y, Y') with respect to $\mathbf{X}$, while the integrals on the RHS are Young integrals of Y with respect to X, and Y' with respect to the path $[0,1] \ni t \mapsto \sum_{0<s\leqslant t} (\Delta X_s)^{\otimes 2}$, respectively.

Proof. Notice that the path $[0,1] \ni t \mapsto \sum_{0<s\leqslant t} (\Delta X_s)^{\otimes 2}$ is of finite variation, implying that the second integral on the RHS of equation (1.2.9) is well defined. Set for $(s,t) \in \Delta_1$

$$\mathbb{Y}_{s,t} := \int_s^t X_{s,r^-} \otimes dX_r.$$

Fix $u \in (0,1]$. By the definitions of rough and Young integral (see equations (1.2.7), (1.2.4), respectively), we need to show that

$$\lim_{(MRS)|\pi_{[0,u]}|\to 0} \sum_{s_i \in \pi_{[0,u]}} Y'_{s_i} \mathbb{Y}_{s_i, s_{i+1}} = 0.$$

The claim follows by Remark 1.2.5 and a similar reasoning as in Theorem 35 in Friz and Shekhar (2017). $\square$

1.2.4 Extended weakly geometric càdlàg rough paths

In this section, we introduce the space of extended weakly geometric càdlàg rough paths. Roughly speaking this space consists of all the weakly geometric p-rough paths with values in $G^{[p]}(\mathbb{R}^{d+1})$ obtained through the addition of an auxiliary path component of finite variation to some $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$. When $p \in [1,2)$, this extension is quite straightforward. When $p \in [2,3)$ instead, care must be taken to ensure consistency between Young and rough integration (see Remark 1.2.18(iii)), and to preserve the group-valued constraint. For the explanation of the index 0 used here we refer to Notation 1.2.17 below.

Definition 1.2.16. Fix $Z \in D^1([0,1], \mathbb{R})$ and $\widehat{\mathbf{X}} \in D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$.

(i) If $p \in [1,2)$, we say that $\widehat{\mathbf{X}}$ is a Z-extended weakly geometric càdlàg p-rough path if for all $i = 1, \dots, d$,

$$\langle \epsilon_i, \widehat{\mathbf{X}} \rangle = \langle \epsilon_i, \mathbf{X} \rangle, \quad \text{and} \quad \langle \epsilon_0, \widehat{\mathbf{X}} \rangle = Z,$$

for some $\mathbf{X} \in D^p([0,1], G^1(\mathbb{R}^d))$.

(ii) If $p \in [2,3)$, we say that $\widehat{\mathbf{X}}$ is a Z-extended weakly geometric càdlàg p-rough path if for all $i, j = 1, \dots, d$,

$$\langle \epsilon_i, \widehat{\mathbf{X}} \rangle = \langle \epsilon_i, \mathbf{X} \rangle, \qquad \langle \epsilon_{ji}, \widehat{\mathbf{X}} \rangle = \langle \epsilon_{ji}, \mathbf{X} \rangle,$$

and for all $j = 0, \dots, d$,

$$\langle \epsilon_0, \widehat{\mathbf{X}} \rangle = Z, \qquad \langle \epsilon_{j0}, \widehat{\mathbf{X}} \rangle = \int_0^{\cdot} \langle \epsilon_j, \widehat{\mathbf{X}}_{0,s-} \rangle dZ_s + \frac{1}{2} \sum_{0<u\leqslant \cdot} \Delta \langle \epsilon_j, \widehat{\mathbf{X}}_u \rangle \Delta Z_u. \quad (1.2.10)$$

for some $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$, and the integral in (1.2.10) is of Young type.

Notation 1.2.17. We use the index 0 to denote the additional auxiliary component Z of $\widehat{\mathbf{X}}$. Unless explicitly mentioned, this notation will be used throughout the work.

Remark 1.2.18. (i) In Definition 1.2.16, we use the $\hat{\cdot}$ notation to denote a Z-extended weakly geometric càdlàg p-rough path $\widehat{\mathbf{X}}$. Even though the specific path considered to define the extension might be different from case to case, this notation will be used throughout the work, and the differences will be emphasized as needed.

(ii) Let $p \in [2,3)$. Notice that by definition any Z-extended weakly geometric càdlàg p-rough path $\widehat{\mathbf{X}}$ lies in $G^2(\mathbb{R}^{d+1})$. Therefore, by the shuffle property (1.1.8), for all $i = 0, 1, \dots, d$, $t \in [0,1]$,

$$\langle \epsilon_{(0,i)}, \widehat{\mathbf{X}}_t \rangle = Z_t \langle \epsilon_i, \widehat{\mathbf{X}}_t \rangle - \langle \epsilon_{(i,0)}, \widehat{\mathbf{X}}_t \rangle.$$

Moreover, since $Z \in D^1([0,1], \mathbb{R})$ and $\langle \epsilon_i, \mathbf{X} \rangle \in D^p([0,1], \mathbb{R})$, the Young integral in (1.2.10) is well defined and for all $t \in [0,1]$,

$$\sum_{0<u\leqslant t} |\Delta \langle \epsilon_i, \widehat{\mathbf{X}}_u \rangle \Delta Z_u| \leqslant \left(\sum_{0<u\leqslant t} |\Delta \langle \epsilon_i, \widehat{\mathbf{X}}_u \rangle|^3 \right)^{1/3} \left(\sum_{0<u\leqslant t} |\Delta Z_u|^{\frac{3}{2}} \right)^{2/3} < \infty.$$

This implies that the path defined in equation (1.2.10) is of finite variation (see also Corollary 1.2.8) and the definition is well-posed.

(iii) Let $\widehat{\mathbf{X}}$ be a Z-extended weakly geometric càdlàg p-rough path, for $p \in [2,3)$. For some $p' \geqslant p$ with $\frac{2}{p} + \frac{1}{p'} > 1$, let $(Y, Y') \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$ be such that for all $(s,t) \in \Delta_1$,

$$\begin{aligned} Y_s \pi_1(\mathbf{X}_{s,t}) &= \langle \epsilon_0, Y_s \rangle Z_{s,t}, \\ Y'_s \pi_2(\mathbf{X}_{s,t}) &= \langle \epsilon_{i0}, Y'_s \rangle \langle \epsilon_{i0}, \mathbf{X}_{s,t} \rangle, \end{aligned}$$

for some $i = 1, \ldots, d$. Then, a slight adaptation of the reasoning in the proof of Lemma 1.2.15 to the present setting yields that for all $t \in (0,1]$,

$$\int_0^t Y_{s^-} d\widehat{\mathbf{X}}_s = \int_0^t Y_{s^-} dZ_s + \frac{1}{2} \sum_{0 < u \leqslant t} Y'_{s^-} \Delta \langle \epsilon_i, \widehat{\mathbf{X}}_u \rangle \Delta Z_u,$$

where the integrals on the RHS are of Young type.

(iv) One can easily see that given $Z \in D^1([0,1], \mathbb{R})$ and $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$, it is always possible to build a Z-extended weakly geometric p-rough path $\widehat{\mathbf{X}}$ as in Definition 1.2.16. Trivially, the reverse is also true.

1.3 Signature of weakly geometric càdlàg rough paths

In this section, we introduce the notion of the signature of weakly geometric càdlàg rough paths. This notion was first studied for continuous paths in Lyons (1998). In this paper Lyons proves that every weakly geometric continuous p-rough path (starting from $\mathbf{1} \in G^{[p]}(\mathbb{R}^d)$) admits an extension to a path of finite p-variation with values in $G^N(\mathbb{R}^d)$, for $N > [p]$. This extension is unique, and even more, it allows us to determine a bijection from the space of continuous weakly geometric p-rough paths to the space of continuous paths with values in $G^N(\mathbb{R}^d)$, for $N > [p]$ with finite p-variation. This result, known as *Lyon's extension theorem* (Theorem 2.2.1 in Lyons (1998), Theorem 3.7 in Lyons et al. (2007)), forms the foundational basis for the definition of the signature of a continuous rough path. Indeed, the signature is defined as the unique path with values in $G((\mathbb{R}^d))$ (introduced in equation (1.1.6)) whose projection to $G^N(\mathbb{R}^d)$ are given by the Lyons's extension theorem.

In Friz and Shekhar (2017), the concept of signature has been extended to càdlàg paths. Here, the notion of weakly geometric càdlàg rough paths and rough integration for these paths has been first introduced. Then, building on the theory for continuous paths and some earlier ideas by Marcus (1978, 1981), an analog of Lyons' extension theorem in the càdlàg setting is also provided. However, unlike in the continuous setting, the uniqueness of the extension is lost and recovered only up to an additional algebraic constraint on the jumps of the extended path. The extension theorem, which is known as *minimal jump extension theorem* (Theorem 20 in Friz and Shekhar (2017)) provides an injective (but not surjective) map from the space of weakly geometric càdlàg p-rough paths to the space of càdlàg paths with values in $G^N(\mathbb{R}^d)$, for $N > [p]$ with finite p-variation.

The section is structured as follows. We first introduce the notion of the signature of continuous weakly geometric p-rough paths. Then we present the signature of càdlàg paths.

To this end, we shall recall and elaborate on the Marcus transformation of a càdlàg path to a continuous path (as introduced in Chevyrev (2017)), which is the fundamental idea on which the notion of signature for càdlàg paths is based.

1.3.1 Signature of continuous rough paths

We follow the presentation in Friz and Victoir (2010) as it is in line with the group-valued paths point of view for weakly geometric rough paths that we use in Section 1.2.

Theorem 1.3.1. *(Lyon's extension theorem, see e.g., Theorem 9.5 in Friz and Victoir (2010)) Let $p \in [1,3)$ and $\mathbb{N} \ni N > [p]$. Every $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$ admits a unique extension to a continuous path $\mathbb{X}^N : [0,1] \to G^N(\mathbb{R}^d)$, such that $\mathbb{X}^N$ starts from $\mathbf{1} \in G^N(\mathbb{R}^d)$ and is of finite p-variation with respect to the CC metric d_{CC} on $G^N(\mathbb{R}^d)$.*

Remark 1.3.2. A key additional result proved in Theorem 9.5 in Friz and Victoir (2010) is that the map

$$\begin{aligned} S^N(\mathbf{X}) : C^p_o([0,1], G^{[p]}(\mathbb{R}^d)) &\mapsto C^p_o([0,1], G^N(\mathbb{R}^d)) \\ \mathbf{X} &\longmapsto S^N(\mathbf{X}) := \mathbb{X}^N \end{aligned} \tag{1.3.1}$$

is a bijection. Here, we denote by $C^p_o([0,1], G^{[p]}(\mathbb{R}^d))$ and $C^p_o([0,1], G^N(\mathbb{R}^d))$ the subsets of paths in $C^p([0,1], G^{[p]}(\mathbb{R}^d))$ and $C^p([0,1], G^N(\mathbb{R}^d))$, respectively, which start from the identity of the group $\mathbf{1}$.

Finally, we present the notion of the signature of a continuous weakly geometric rough path.

Definition 1.3.3. Let $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$. The *signature* of $\mathbf{X}$ is defined as the unique path

$$\mathbb{X} : [0,1] \to G((\mathbb{R}^d)),$$

such that for all $\mathbb{N} \ni N > [p]$, $\pi_{\leqslant N}(\mathbb{X}) = \mathbb{X}^N$, where $\mathbb{X}^N$ denotes the unique extension path of $\mathbf{X}$ in $G^N(\mathbb{R}^d)$ provided by Theorem 1.3.1.

One can also prove that the extension path provided in Theorem 1.3.1 can be computed by solving a linear rough differential equation (RDE).

Corollary 1.3.4. *Let $\mathbb{N} \ni N > [p]$, $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$. The unique extension path $\mathbb{X}^N$ of $\mathbf{X}$ with values in $G^N(\mathbb{R}^d)$ provided by Theorem 1.3.1 satisfies the following linear RDE*

$$d\mathbb{X}^N = \mathbb{X}^N \otimes d\mathbf{X}, \qquad \mathbb{X}^N_0 = \mathbf{1} \in G^N(\mathbb{R}^d), \tag{1.3.2}$$

which admits a unique solution, whose explicit form reads

$$\mathbb{X}_t^N = 1 + \int_0^t \mathbb{X}_s^N \otimes d\mathbf{X}_s, \qquad t \in [0,1]. \tag{1.3.3}$$

Here the integral is understood as a rough integral.

We refer to Lemma 2.10 in Lyons et al. (2007), Exercise 4.6 in Friz and Hairer (2014), and Section 5 in Chevyrev (2024) for the proof and more details on the statement of Corollary 1.3.4.

Remark 1.3.5. (i) Since for every $N > [p]$, equation (1.3.2) admits a unique global solution (see Section 10.7 in Friz and Victoir (2010)), the signature $\mathbb{X}$ of a path $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d)$ can also be defined without ambiguity as the unique solution of the following linear RDE in the extended tensor algebra:

$$d\mathbb{X} = \mathbb{X} \otimes d\mathbf{X}, \qquad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d)).$$

(ii) By Corollary 1.3.4, when $p \in [1,2)$, the signature $\mathbb{X}$ of a path $\mathbf{X} \in C^p([0,1], G^1(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^1(\mathbb{R}^d)$ is the continuous $G((\mathbb{R}^d))$-valued path given by the sequence of all its iterated Young integrals, i.e., for all $t \in [0,1]$,

$$\mathbb{X}_t := \Big(1, \int_0^t d\mathbf{X}_{t_1}, \int_0^t \int_0^{t_1} d\mathbf{X}_{t_2} \otimes d\mathbf{X}_{t_1}, \dots, \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} d\mathbf{X}_{t_n} \otimes \dots \otimes d\mathbf{X}_{t_1}, \dots \Big).$$

Notice that for all $i \in \mathbb{N}$,

$$\int_0^t \int_0^{t_1} \dots \int_0^{t_{i-1}} d\mathbf{X}_{t_i} \otimes \dots \otimes d\mathbf{X}_{t_1} \in (\mathbb{R}^d)^{\otimes i}.$$

Finally, we conclude this section by highlighting some important properties of the signature path. Even though these results are well-known, we include the proofs for clarity of exposition.

Proposition 1.3.6. Let $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d)$, and for $N > [p]$, let S^N be the extension map introduced in equation (1.3.1). Then,

(i) $S(\mathbf{X})^N_{\phi(t)} = S^N(\mathbf{X}_\phi)_t$, for all $t \in [0,1]$ and ϕ time reparametrization;

(ii) $S^N(\mathbf{X})_{s,t} = S^N(\mathbf{X})_{s,u} \otimes S^N(\mathbf{X})_{u,t}$, for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$.

Proof. (i): For $p = 1$ we refer to the proof of Proposition 7.10 in Friz and Victoir (2010). For the remaining cases, the assertion follows by Corollary 1.3.4, the interpretation of an RDE as an integral equation, and the property of the Young and rough integral, for which it holds that for all $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$ and $t \in [0,1]$

$$\int_0^{\phi(t)} Y_s d\mathbf{X}_s = \int_0^t Y_{\phi(s)} d\mathbf{X}_{\phi(s)},$$

for suitably chosen paths Y (or (Y, Y')) such that the integrals are well-defined.

(ii) It directly follows from the definition of path increments (see equation (1.2.3)) for paths with values in the Lie group $G^N(\mathbb{R}^d)$. $\square$

Remark 1.3.7. We refer to properties (i) and (ii) in Proposition 1.3.6 as *invariance under reparametrization* and *Chen's relation*.

1.3.2 Marcus transformation of a weakly geometric càdlàg rough path

In this section, we recall and elaborate on the Marcus transformation of weakly geometric càdlàg p-rough paths for $p \in [1,3)$, introduced in Chevyrev (2017), (see also Section 2.3 in Chevyrev and Friz (2019)). This is an operation that associates to every càdlàg path a continuous one obtained by introducing an additional time interval at each jump time and connecting the states before and after the jump via the so-called *log-linear path-function* denoted by $\ell\ell$. Let us start by listing all the objects required for the precise definition.

(i) Fix $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$, and let $(t_k)_{k\in\mathbb{N}}$ denote the sequence of its jumps times.

(ii) Fix a sequence $R := (r_k)_k$ such that for all $k \in \mathbb{N}$, $r_k > 0$ and $\Sigma_R := \sum_{k=1}^{\infty} r_k < \infty$, and define for all $t \in [0,1]$

$$\tau_{\mathbf{X},R}(t) := t + \sum_{k=1}^{\infty} r_k 1_{\{t_k \leqslant t\}}. \tag{1.3.4}$$

Notice that $\tau_{\mathbf{X},R}$ is an increasing càdlàg function from $[0,1]$ with values in $[0, 1+\Sigma_R]$ whose sequence of jumps times coincide with the sequence of jumps times of $\mathbf{X}$.

(iii) Consider the *log-linear path function* $\ell\ell$ defined as follows:

$$\begin{aligned} \ell\ell : G^{[p]}(\mathbb{R}^d) \times G^{[p]}(\mathbb{R}^d) &\to C([0,1], G^{[p]}(\mathbb{R}^d)) \\ (\mathbf{x}, \mathbf{y}) &\mapsto \Big(s \mapsto \mathbf{x} \otimes \exp^{([p])}(s \log^{([p])}(\mathbf{x}^{-1} \otimes \mathbf{y}))\Big). \end{aligned} \tag{1.3.5}$$

Notice that for all $(\mathbf{x}, \mathbf{y}) \in G^{[p]}(\mathbb{R}^d) \times G^{[p]}(\mathbb{R}^d)$, $\ell\ell(\mathbf{x}, \mathbf{y})(0) = \mathbf{x}$ and $\ell\ell(\mathbf{x}, \mathbf{y})(1) = \mathbf{y}$.

(iv) Define $\mathbf{Y} \in C([0, 1+\Sigma_R], G^{[p]}(\mathbb{R}^d))$ as follows: for all $s \in [0, 1+\Sigma_R]$,

$$\mathbf{Y}_s := \begin{cases} \mathbf{X}_t, & \text{if } (1), \\ \mathbf{X}_{t_k^-} \otimes \exp^{([p])}\left(\frac{(s-\tau_{\mathbf{X},R}(t_k^-))}{r_k} \log^{([p])}(\Delta \mathbf{X}_{t_k})\right) & \text{if } (2), \end{cases}$$

where,

(1) : $s = \tau_{\mathbf{X},R}(t)$ for some $t \in [0,1]$;

(2) : $s \in [\tau_{\mathbf{X},R}(t_k^-), \tau_{\mathbf{X},R}(t_k)]$, $k \in \mathbb{N}$.

Definition 1.3.8. Let $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$. Fix a sequence R as in (ii) and let $\tau_{\mathbf{X},R}$ be the increasing càdlàg function built through R as in equation (1.3.4). Let $\mathbf{Y}$ be the continuous path defined via $\tau_{\mathbf{X},R}$ and the log-linear path function $\ell\ell$ as in (iv), and let ψ_R denote an increasing bijection from $[0,1]$ to $[0, 1+\Sigma_R]$. The *Marcus-transformed path of* $\mathbf{X}$ associated to the pair (R, ψ_R) is the continuous path $\widetilde{\mathbf{X}} \in C([0,1], G^{[p]}(\mathbb{R}^d))$ given by

$$\widetilde{\mathbf{X}}_\cdot := \mathbf{Y}_{\psi_R(\cdot)}.$$

Assumption 1.3.9. *From now on, whenever we refer to the Marcus-transformed path of a càdlàg path* $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$ *associated to a pair* (R, ψ_R)*, we implicitly assume that* R *is a sequence satisfying condition (ii) and* ψ_R *is an increasing bijection from* $[0,1]$ *to* $[0, 1+\Sigma_R]$.

Remark 1.3.10. Fix $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$.

(i) If $\mathbf{X} \in C([0,1], G^{[p]}(\mathbb{R}^d))$, for any sequence R as in (ii), the map $\tau_{\mathbf{X},R}$ in (1.3.4) is the identity map, i.e., $\tau_{\mathbf{X},R}(t) = t$ for all $t \in [0,1]$, and $\mathbf{Y} = \mathbf{X}$, for $\mathbf{Y}$ as in (iv). In this case, we establish the convention that ψ_R is an increasing bijection from $[0,1]$ to $[0,1]$ and thus $\widetilde{\mathbf{X}}$ is simply a time-reparametrization of $\mathbf{X}$.

(ii) Let $\widetilde{\mathbf{X}}$ be the Marcus-transformed path of $\mathbf{X}$ associated to some pair (R, ψ_R). Observe that we can recover $\mathbf{X}$ from $\widetilde{\mathbf{X}}$ via

$$\mathbf{X}_\cdot := \widetilde{\mathbf{X}}_{\psi_R^{-1}(\tau_{\mathbf{X},R}(\cdot))}.$$

(iii) The Marcus-transformed paths of $\mathbf{X}$ associated with two different pairs (R, ψ_R) and $(\widetilde{R}, \psi_{\widetilde{R}})$, are simply time-reparametrizations of one another. More precisely, for notational convenience let us denote by $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ the transformed paths associated to (R, ψ_R) and $(\widetilde{R}, \psi_{\widetilde{R}})$, respectively. Then one can show that there exists a time reparametrization ϕ such that $\mathbf{Z}_\phi = \widetilde{\mathbf{Z}}$ and for all $t \in [0,1]$,

$$\phi(\psi_R^{-1}(\tau_{\mathbf{X},R}(t))) = \psi_{\widetilde{R}}^{-1}(\tau_{\mathbf{X},\widetilde{R}}(t)).$$

1.3.3 Signature of càdlàg rough paths

The concept of the signature of a càdlàg rough path (introduced in Theorem 20 in Friz and Shekhar (2017)) builds upon the established framework for continuous paths. Roughly speaking, to compute the signature of a càdlàg rough path $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1}$, the initial step involves transforming $\mathbf{X}$ into a continuous path via the Marcus transformation detailed in Section 1.3.2. This transformation yields a transformed path $\widetilde{\mathbf{X}}$ (with respect to some pair (R, ψ_R)), which is a continuous weakly geometric p-rough path by construction, for which the existence (and uniqueness) of the signature path is guaranteed by Lyons's extension theorem (Theorem 1.3.1). The signature of the original weakly geometric càdlàg rough path $\mathbf{X}$ is defined then as the unique $G((\mathbb{R}^d))$-valued path $\mathbb{X}$ such that the projection paths over $G^N(\mathbb{R}^d)$, denoted by $\mathbb{X}^N$, are the càdlàg paths given by

$$\mathbb{X}^N_{\cdot} := S^N(\widetilde{\mathbf{X}})_{\psi_R^{-1}(\tau_{\mathbf{X},R}(\cdot))}, \tag{1.3.6}$$

for S^N denoting the extension map in equation (1.3.1) and $\tau_{\mathbf{X},R}$ the càdlàg map defined in equation (1.3.4).

Before presenting the precise statement of Theorem 20 in Friz and Shekhar (2017), we provide some comments on this first sketch about the construction of the signature path.

Remark 1.3.11. (i) The definition of the signature $\mathbb{X}$ of a càdlàg rough path is independent of the particular Marcus transformed path $\widetilde{\mathbf{X}}$ associated with some specific pair (R, ψ_R). This follows by Proposition 1.3.6 and Remark 1.3.10(iii).

(ii) By definition of the Marcus transformation, one gets that for all t_k, $k \in \mathbb{N}$, jumps time of $\mathbf{X}$, and $s \in [\tau_{\mathbf{X},R}(t_k^-), \tau_{\mathbf{X},R}(t_k)]$,

$$\widetilde{\mathbf{X}}_{\psi_R^{-1}(s)} = \widetilde{\mathbf{X}}_{\psi_R^{-1}(\tau_{\mathbf{X},R}(t_k^-))} \otimes \exp^{([p])}\left(\frac{(s - \tau_{\mathbf{X},R}(t_k^-))}{r_k} \log^{([p])}(\Delta \mathbf{X}_{t_k})\right).$$

As a consequence, by Lyons' extension theorem and Proposition 1.3.6(ii), for all $N > [p]$,

$$\Delta \mathbb{X}^N_{t_k} := S^N(\widetilde{\mathbf{X}})^{-1}_{\psi_R^{-1}(\tau_{\mathbf{X},R}(t_k^-))} \otimes S^N(\widetilde{\mathbf{X}})_{\psi_R^{-1}(\tau_{\mathbf{X},R}(t_k))} = \exp^{(N)}\left(\log^{([p])}(\Delta \mathbf{X}_{t_k})\right),$$

which explains the choice of the name *minimial jump extension theorem* for Lyon's equivalent theorem in the càdlàg setting.

As a result of these considerations, the presentation of the next theorem follows smoothly.

Theorem 1.3.12. *(Minimal jump extension theorem, see Theorem 20 in Friz and Shekhar (2017)) Let $\mathbb{N} \ni N > [p]$. Every $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$ admits a unique extension to a càdlàg path $\mathbb{X}^N : [0,1] \to G^N(\mathbb{R}^d)$, such that $\mathbb{X}^N$ starts from $\mathbf{1} \in G^N(\mathbb{R}^d)$, is of finite p-variation with respect to the CC metric d_{CC} on $G^N(\mathbb{R}^d)$, and satisfies the following condition:*

$$\log^{(N)}(\Delta \mathbb{X}^N_t) = \log^{([p])}(\Delta \mathbf{X}_t) \quad \text{for all} \quad t \in [0,1]. \tag{1.3.7}$$

$\mathbb{X}^N$ *is called* minimal jump extension *of* $\mathbf{X}$ *in* $G^N(\mathbb{R}^d)$.

Remark 1.3.13. (i) Observe that the left-hand side of (1.3.7) lies in $T_0^N(\mathbb{R}^d)$ while the right-hand side lies in $T_0^{[p]}(\mathbb{R}^d)$. This slight abuse of notation has to be interpreted as follows: $\langle \epsilon_I, \log^{(N)}(\Delta \mathbb{X}^N_t) \rangle = 0$ for all $t \in [0,1]$, $|I| > 1$ if $\mathbf{X}$ is Marcus-like, $|I| > [p]$ otherwise.

(ii) Notice that if in particular $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$, the statement of Theorem 1.3.12 coincides with the statement of Theorem 1.3.1.

(iii) In the càdlàg setting, uniqueness of the extension is obtained provided that the additional algebraic constraint on the jumps (1.3.7) is satisfied. The form of this constraint is due to the use of the log-linear path functional (see equation (1.3.5)) to interpolate the jumps (see Remark 1.3.11(ii)). The choice of another interpolation procedure would have resulted in a different notion of the signature path.

(iv) Simple examples show that without the constraint imposed by condition (1.3.7), the uniqueness of the extension is lost (see, for instance, Example 18 in Friz and Shekhar (2017)). This is due to the fact that there are nontrivial pure jump paths of finite p-variation with $p < 1$. As a consequence, contrary to the restriction to the set of continuous paths (see Remark 1.3.2), the map

$$\begin{aligned} S^N(\mathbf{X}) : D^p_o([0,1], G^{[p]}(\mathbb{R}^d)) &\mapsto D^p_o([0,1], G^N(\mathbb{R}^d)) \\ \mathbf{X} &\longmapsto S^N(\mathbf{X}) := \mathbb{X}^N, \end{aligned} \tag{1.3.8}$$

for $N > [p]$ and $\mathbb{X}^N$ denoting the minimal jump extension of $\mathbf{X}$, is injective but not surjective. Here, we denote by $D^p_o([0,1], G^{[p]}(\mathbb{R}^d))$ and $D^p_o([0,1], G^N(\mathbb{R}^d))$ the subsets of paths in $D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and $D^p([0,1], G^N(\mathbb{R}^d))$, respectively, which start from the identity of the group $\mathbf{1}$.

Finally, the signature of a weakly geometric càdlàg rough path is defined as the unique path extension provided by Theorem 1.3.12.

Definition 1.3.14. Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$. The *signature* of $\mathbf{X}$ is defined as the unique path

$$\mathbb{X} : [0,1] \to G((\mathbb{R}^d)),$$

such that for all $\mathbb{N} \ni N > [p]$, $\pi_{\leqslant N}(\mathbb{X}) = \mathbb{X}^N$, where $\mathbb{X}^N$ denotes the unique extension path of $\mathbf{X}$ in $G^N(\mathbb{R}^d)$ provided by Theorem 1.3.12.

Notation 1.3.15. From now on, given $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$, we refer to $\mathbb{X}$ as *signature* of $\mathbf{X}$ and to $\mathbb{X}^N$ as *truncated signature* of order N of $\mathbf{X}$.

Remark 1.3.16. (i) If in particular $\mathbf{X} \in C^{[p]}[0,1], G^{[p]}(\mathbb{R}^d))$, Definition 1.3.14 and Definition 1.3.3 coincide.

(ii) In this càdlàg literature, it is frequent to refer to the signature as *Marcus signature*. This is because the construction if this object relies on the original Marcus' idea (see Marcus (1978, 1981)) of turning a càdlàg $\mathbb{R}^d$-valued path into a continuous one, by introducing an additional time interval at each jump time and replacing the jumps by a straight line which connects the states before and after the jump. In the current càdlàg rough path setting, where the state space is $G^{[p]}(\mathbb{R}^d)$, the jumps of $\mathbf{X}$ are connected by means of the log-linear path functional which is in fact the analog of the straight line in the Lie group $G^{[p]}(\mathbb{R}^d)$.

As for continuous paths, one can prove that the truncated signature of a weakly geometric càdlàg rough path can be computed by solving a (linear) Marcus-type RDE. This notion of RDE has been introduced in Chevyrev and Friz (2019), where once again, the authors rely on the theory of RDE driven by continuous paths to establish a new solution theory for RDE driven by weakly geometric càdlàg rough paths. Before treating the special case of the equation solved by the signature, we provide a summary of this solution concept.

A Marcus-type RDE driven by some $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ is denoted by

$$dY_t = V(Y_t) \diamond d\mathbf{X}_t, \tag{1.3.9}$$

for some given suitable vector field V. A solution of such RDE is defined as the $\psi_R^{-1}(\tau_{\mathbf{X},R}(\cdot))$-time changed of the solution of the RDE driven by the continuous Marcus transformed path $\widetilde{\mathbf{X}}$ (associated to some pair (R, ψ_R)). More precisely, let $\widetilde{Y}$ be the solution of the RDE

$$d\widetilde{Y}_t = V(\widetilde{Y}_t) d\widetilde{\mathbf{X}}_t.$$

The solution $\mathbf{Y}$ to the Marcus-type RDE (1.3.9) is then defined as $Y := \widetilde{Y}_{\psi_R^{-1}(\tau_{\mathbf{X},R}(\cdot))}$ (see Definition 2.1 and Remark 3.4 in Chevyrev and Friz (2019)). One can show that Y does not depend on the specific pair (R, ψ_R) (see Remark 3.2 in Chevyrev and Friz (2019)). Moreover, by exploiting the specific parametrization of the log-linear path functional $\ell\ell$ employed to build the transformed path $\widetilde{\mathbf{X}}$ (see equation (1.3.5)), the original Marcus-type RDE (1.3.9) can be rewritten as an explicit integral equation driven by $\mathbf{X}$ (see Definition 37 in Friz and Shekhar (2017)).

Now, recall from the discussion at the beginning of this section that by construction, the truncated signature path of $\mathbf{X}$, denoted by $\mathbb{X}^N$, satisfies

$$\mathbb{X}^N_\cdot := S^N(\widetilde{\mathbf{X}})_{\psi_R^{-1}(\tau_{\mathbf{X},R}(\cdot))}.$$

Moreover, by Corollary 1.3.4 $S^N(\widetilde{\mathbf{X}})$ solves the linear RDE (1.3.2). Hence, the following corollary follows directly. We refer to Corollary 39 in Friz and Shekhar (2017) for more details.

Corollary 1.3.17. *Let $\mathbb{N} \ni N > [p]$, $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d))$. The unique extension path $\mathbb{X}^N$ of $\mathbf{X}$ with values in $G^N(\mathbb{R}^d)$ provided by Theorem 1.3.12 satisfies the following linear Marcus-type RDE*

$$d\mathbb{X}^N = \mathbb{X}^N \otimes \diamond d\mathbf{X}, \qquad \mathbb{X}^N_0 = \mathbf{1} \in G^N(\mathbb{R}^d), \tag{1.3.10}$$

which admits a unique solution, whose explicit form can be written as

$$\mathbb{X}^N_t = 1 + \int_0^t \mathbb{X}^N_{s-} \otimes d\mathbf{X}_s + \sum_{0<s\leqslant t} \mathbb{X}^N_{s-} \otimes \big(\exp^{(N)}(\log^{([p])}(\Delta\mathbf{X}_s)) - \Delta\mathbf{X}_s \big). \tag{1.3.11}$$

The integral in (1.3.11) is understood as a rough integral and the summation term is well-defined as an absolutely summable series.

Remark 1.3.18. (i) If in particular $\mathbf{X} \in C^{[p]}([0,1], G^{[p]}(\mathbb{R}^d))$, equations (1.3.10) and (1.3.2) coincide.

(ii) For $N > [p]$, $\mathbb{X}^N$ is the unique solution of the Marcus-type RDE (1.3.10), (uniqueness follows from the uniqueness results for continuous linear RDEs (see Section 10.7 in Friz and Victoir (2010))). Therefore, given a weakly geometric càdlàg p-rough path $\mathbf{X}$, its signature can also be defined without ambiguity as the unique solution of the following Marcus-type RDE in the extended tensor algebra:

$$d\mathbb{X} = \mathbb{X} \otimes \diamond d\mathbf{X}, \qquad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d)).$$

(iii) Solutions of Marcus-type RDEs are also called *geometric* solutions. This notion has to be distinguished with the concept of *forward* solution of a RDE, introduced in Friz and Zhang (2017).

We remark furthermore that the same type of (geometric) equation has also been considered in a semimartingale context by Kurtz et al. (1995) (see also Section 1.4).

Finally, the invariance under reparametrization property and the validity of Chen's relation described in Proposition 1.3.6 for continuous paths, remain valid also in the càdlàg setting.

Corollary 1.3.19. *Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ with $\mathbf{X}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^d)$, and for $N > [p]$, let S^N be the extension map introduced in equation (1.3.8). Then,*

(i) $S(\mathbf{X})^N_{\phi(t)} = S^N(\mathbf{X}_\phi)_t$, for all $t \in [0,1]$ and ϕ time reparametrization;

(ii) $S^N(\mathbf{X})_{s,t} = S^N(\mathbf{X})_{s,u} \otimes S^N(\mathbf{X})_{u,t}$, for all $0 \leqslant s \leqslant u \leqslant t \leqslant 1$.

Proof. (i) see Proposition 5.2.11 and Section 5.2.4.

(ii): it directly follows the construction of the signature of a càdlàg paths combined with Proposition 1.3.6. □

Assumption 1.3.20. *Throughout, we assume that all the weakly geometric càdlàg rough paths start at $\mathbf{1} \in G^{[p]}(\mathbb{R}^d)$, i.e., $\mathbf{X}_0 = \mathbf{1}$. For notational convenience, with a slight abuse of notation, we denote the space of such paths by $D^p([0,1], G^{[p]}(\mathbb{R}^d))$ (or $C^p([0,1], G^{[p]}(\mathbb{R}^d))$ for the subset of continuous paths) instead of $D^p_o([0,1], G^{[p]}(\mathbb{R}^d))$ (or $C^p_o([0,1], G^{[p]}(\mathbb{R}^d))$).*

Linear projections of the signature path We conclude the section with a more in-depth description of the signature path. To this end, we project along some linear functions on the extended tensor algebra the explicit form of the (unique) solution of the Marcus-type RDE (1.3.10), which is in fact the signature path. Fix $\mathbf{u} \in T(\mathbb{R}^d)$ and recall the shifts introduced in equation (1.1.7). Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and denote by $\mathbb{X}$ its signature. Then, a projection of equation (1.3.11) along $\mathbf{u}$ combined with Lemma 2.9 in Friz and Zhang (2017) yields that

(i) if $p \in [1,2)$,

$$\begin{aligned}\langle \mathbf{u}, \mathbb{X}_\cdot\rangle =& \langle \mathbf{u}, \mathbb{X}_0\rangle + \int_0^\cdot \langle \mathbf{u}^{(1)}, \mathbb{X}_{s-}\rangle d\mathbf{X}_s \\ &+ \sum_{0<s\leqslant \cdot} \langle \mathbf{u}, \mathbb{X}_s\rangle - \langle \mathbf{u}, \mathbb{X}_{s-}\rangle - \langle \mathbf{u}^{(1)}, \mathbb{X}_{s-}\rangle \Delta \mathbf{X}_s,\end{aligned} \tag{1.3.12}$$

where the integral term is as a Young integral of $\langle \mathbf{u}^{(1)}, \mathbb{X}_\cdot\rangle \in D^p([0,1], \mathbb{R}^{d+1})$ with respect to $\mathbf{X}$;

(ii) if $p \in [2,3)$,

$$\begin{aligned}\langle \mathbf{u}, \mathbb{X}_\cdot\rangle =& \langle \mathbf{u}, \mathbb{X}_0\rangle + \int_0^\cdot \langle \mathbf{u}^{(1)}, \mathbb{X}_{s-}\rangle d\mathbf{X}_s \\ &+ \sum_{0<s\leqslant \cdot} \langle \mathbf{u}, \mathbb{X}_s\rangle - \langle \mathbf{u}, \mathbb{X}_{s-}\rangle - \langle \mathbf{u}^{(1)}, \mathbb{X}_{s-}\rangle \Delta X_s - \langle \mathbf{u}^{(2)}, \mathbb{X}_{s-}\rangle \Delta \mathbb{X}^{(2)}_s,\end{aligned} \tag{1.3.13}$$

where the integral term is as a rough integral of the controlled rough path $\big(\langle \mathbf{u}^{(1)}, \mathbb{X}_\cdot\rangle, \langle \mathbf{u}^{(2)}, \mathbb{X}_\cdot\rangle\big) \in \mathcal{V}^{p,\frac{p}{2}}_{\mathbf{X}}$ with respect to $\mathbf{X}$, and for all $t \in (0,1]$, we set $(1, \Delta X_t, \Delta \mathbb{X}^{(2)}_t) := \Delta \mathbf{X}_t$.

1.4 Càdlàg semimartingales as rough paths

Càdlàg semimartingales fit well into the theory of càdlàg rough paths. Indeed, every semimartingale admits a canonical lift which is a.s. a Marcus-like càdlàg p-rough path for any $p > 2$. For the proof of the subsequent proposition, we refer to Friz and Shekhar (2017) and Chevyrev and Friz (2019).

Proposition 1.4.1. Let $p \in (2,3)$, $X := (X_t)_{t\in[0,1]}$ be a $\mathbb{R}^d$-valued càdlàg semimartingale on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$, and let $([X,X]^c_t)_{t\in[0,1]}$ be its $(\mathbb{R}^d)^{\otimes 2}$-valued continuous quadratic variation. Set

$$\mathbb{X}^{(2)}_{0,t} := \int_0^t X_{0,r-} \otimes dX_r + \frac{1}{2}[X,X]^c_t + \frac{1}{2} \sum_{0<u\leqslant t} \Delta X_u \otimes \Delta X_u, \qquad t \in [0,1],$$

where the integral is understood as Itô-integral. Define for all $t \in [0,1]$,

$$\mathbf{X}_t := (1, X_t - X_0, \mathbb{X}^{(2)}_{0,t}). \tag{1.4.1}$$

Then $\mathbf{X}(\omega) \in D^p([0,1]; G^2(\mathbb{R}^d))$ a.s. We call $\mathbf{X}$ Marcus lift of X.

In the following, unless explicitly mentioned, whenever we consider a semimartingale X, we always assume that X is defined on some filtered probability space $(\Omega.\mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$.

Remark 1.4.2. Let $\mathbf{X}$ be the Marcus lift of a semimartingale X. Recall Remark 1.2.2(i) and notice that for all $t \in [0,1]$, $\mathbf{X}_t = \exp^{(2)}(X_t - X_0, A_t)$, where, $A : [0,1] \to [\mathbb{R}^d, \mathbb{R}^d]$ is càdlàg path defined via

$$\langle \epsilon_{ij}, A_t \rangle := \frac{1}{2}\left(\int_0^t X^i_{0,s} dX^j_s - \int_0^t X^j_{0,s} dX^i_s \right),$$

for all $t \in [0,1]$ and each $i < j$.

The signature of semimartingales can be computed by adopting stochastic integration methods instead of resorting to rough integrals. Proposition 4.16 in Chevyrev and Friz (2019) accounts for this. For completeness, we report an adaptation of this statement to our specific context.

Proposition 1.4.3. Let X be an $\mathbb{R}^d$-valued semimartingale and $\mathbf{X}$ its Marcus lift. It holds that the Marcus-type RDE

$$d\mathbb{X} = \mathbb{X} \otimes \diamond d\mathbf{X}, \qquad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d))$$

coincides a.s. with the Marcus SDE (see Kurtz et al. (1995))

$$d\mathbb{X} = \mathbb{X} \otimes \diamond dX, \qquad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d)). \tag{1.4.2}$$

Existence and uniqueness of the solution of the Marcus SDE (1.4.2) are provided by Theorem 3.2 in Kurtz et al. (1995). The explicit form of equation (1.4.2) is given by

$$\mathbb{X}_t = 1 + \int_0^t \mathbb{X}_{s^-} \otimes dX_s + \frac{1}{2}\int_0^t \mathbb{X}_{s^-} \otimes d[X,X]^c_s + \sum_{0<s\leqslant t} \mathbb{X}_{s^-} \otimes \big(\exp(\Delta X_s) - \Delta X_s - 1\big),$$

where $\exp(\Delta X_s) = \sum_{k=0}^\infty \frac{1}{k!}(\Delta X_s)^{\otimes k}$. This means that $\langle \epsilon_\varnothing, \mathbb{X}_t \rangle = 1$ and for each multi-index I with entries in $\{1, \dots, d\}$, we get

$$\begin{aligned}\langle \epsilon_I, \mathbb{X}_t \rangle = &\int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s^-} \rangle dX^{i_{|I|}}_s + \frac{1}{2}\int_0^t \langle \epsilon_{I''}, \mathbb{X}_{s^-} \rangle d[X^{i_{|I|-1}}, X^{i_{|I|}}]^c_s 1_{\{|I|>1\}} \\ &+ \sum_{0<s\leqslant t} \sum_{\epsilon_{I_1}\otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} \langle \epsilon_{I_1}, \mathbb{X}_{s^-} \rangle \langle \epsilon_{I_2}, (\Delta X_s)^{\otimes |I_2|} \rangle 1_{\{|I_2|>1\}}.\end{aligned}$$

Observe that here we consider Itô type integrals.

Remark 1.4.4. (i) If X is a $\mathbb{R}$-valued càdlàg semimartingale, by Itô's formula we get that

$$\mathbb{X}_t = (1, X_t - X_0, \frac{(X_t - X_0)^2}{2}, \dots, \frac{(X_t - X_0)^n}{n!}, \dots) \in G((\mathbb{R})).$$

(ii) If X is a continuous $\mathbb{R}^d$-valued semimartingale, then $\mathbb{X}$ coincides with the Stratonovich signature, see e.g., Chapter 17.2 in Friz and Victoir (2010).

2 Universal approximation theorems for continuous functions of càdlàg paths and Lévy-type signature models

The present chapter is based on Cuchiero et al. (2024d), a joint work with Christa Cuchiero and Sara Svaluto-Ferro.

2.1 Introduction

As explained in the introduction on the present thesis, methods based on the signature of a path represent a non-parametric way for extracting characteristic features from time series data which is essential in machine learning tasks. This explains why these techniques are more and more applied in econometrics and mathematical finance, see e.g., Bühler et al. (2020), Kalsi et al. (2020), Arribas et al. (2021), Lyons et al. (2020), Liao et al. (2024), Bayer et al. (2023), Cuchiero et al. (2022), Min and Hu (2021), Akyildirim et al. (2022) and the references therein. Indeed, signature-based methods allow for data-driven modeling approaches, while stylized facts or first principles from mathematical finance can still be easily guaranteed.

The choice of the signature as feature map capturing the specific characteristics of the path, can be explained by a *universal approximation theorem* (UAT) according to which continuous (with respect to certain variation distances) functionals of *continuous* paths can be approximated on compact sets of paths by linear functionals of the time-extended signature. This result is however only proved for continuous paths and therefore leaves open the question of approximating continuous functionals of the more general set of *càdlàg* paths, which have particular relevance when it comes to financial modeling. Based on advances on the signature of càdlàg paths by Friz and Shekhar (2017) and using the Skorokhod-J_1-topology on (Lie group valued) càdlàg paths, we here present a UAT that solves this question. Firstly, we prove that continuous path functionals over the time interval $[0, 1]$ can be uniformly approximated on compact sets of paths by linear functionals of the time-extended signature evaluated at the final time 1. Secondly, we extend this result by considering functionals of the stopped paths, stopped at all deterministic times $t \in [0, 1]$, which can thus be associated with *non-anticipative* path-functionals (see Remark 2.2.14). We then prove that the coefficients of the linear functionals do not depend on $t \in [0, 1]$ but can be chosen uniformly in time. This is subject of Theorem 2.2.13, our main result in the first part of the chapter.

Our principal motivation to treat this problem comes from signature-based models for finance that allow for an inclusion of jumps. Indeed, all signature models for asset prices that have been proposed so far, (see Arribas et al. (2021), Cuchiero et al. (2023a, 2024b)), build on stochastic processes with continuous trajectories. In order to extend them to processes with càdlàg trajectories and to show universality properties among classical jump-diffusion models in finance, the UAT proved in Section 2.2 is important, as will become apparent from our definition of *Lévy-type signature models* below.

The approach that we follow consists in parameterizing the model itself or its characteristics as linear functionals of the signature of an augmented Lévy process X that we call market's primary underlying process. More precisely, $\mathbb{X}$ denotes the signature of the $d_X + 2$ dimensional Lévy process $(X_t)_{t\geqslant 0}$, whose dynamics under some risk neutral measure $\mathbb{Q}$ are given by

$$X_t := \left(t, W_t, \int_0^t \int_{\mathbb{R}} x(\mu-\nu)(ds,dx), \int_0^t \int_{\mathbb{R}} x^2 \mu(ds,dx), \dots, \int_0^t \int_{\mathbb{R}} x^{d_X} \mu(ds,dx)\right),$$

where μ denotes a homogeneous Poisson random measure with compensator ν and W a one-dimensional Brownian motion. As introduced in Friz and Shekhar (2017) for general semimartingales, we define the signature $\mathbb{X}$ of the Lévy process X, as solution of the Marcus stochastic differential equation (see Kurtz et al. (1995) for more details) in the extended tensor algebra $T((\mathbb{R}^d))$ over $\mathbb{R}^d$,

$$d\mathbb{X} = \sum_{i=1}^{d} \mathbb{X} \otimes \diamond dX^i, \qquad \mathbb{X}_0 = (1,0,0,\dots) \in G((\mathbb{R}^d)).$$

We refer to Section 1.4 for the precise definitions of the involved quantities. We then introduce *Lévy-type signature models* under some risk neutral measure $\mathbb{Q}$ via

$$\begin{aligned} S(\boldsymbol{\ell})_t = S_0 &+ \int_0^t \Bigg(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-}\rangle\Bigg) dW_s \\ &+ \int_0^t \int_{\mathbb{R}} \Bigg(\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{s-}\rangle\Bigg) x(\mu-\nu)(ds,dx), \end{aligned} \tag{2.1.1}$$

for some parameters $\ell_W^J, \ell_\nu^J \in \mathbb{R}$. Here $\langle \epsilon_J, \mathbb{X}_t\rangle$ denotes the component $\mathbb{X}_t$ corresponding to the J-th basis element in $T((\mathbb{R}^d))$.

As already indicated above, due to the linear characteristics in $\mathbb{X}$, these models can be considered as *universal*, since by our càdlàg versions of the UAT the characteristics of a Lévy driven SDE can be approximated arbitrarily well, if they are continuous (with respect to the Skorokhod-J_1-topology) path functionals of the (enhanced) Lévy process X. This applies in particular to jump-diffusions as considered e.g. in Section III.2c in Jacod and Shiryaev (1987).

To analyze tractability properties of models of the form (2.1.1), we first show that the truncated signature process of a generic multivariate Lévy process is a *polynomial process* *on the truncated tensor algebra* and derive its expected value via the so-called moment formula, thus solving a linear ODE (see Cuchiero et al. (2012), Filipović and Larsson (2020)). Specifically, we prove that under moment assumptions on the Lévy measure, the expected signature of a Lévy process assumes the Lévy-Khintchine form described in Friz and Shekhar (2017).

This in turn can be used to get explicit formulas of the expected signature of S in terms of the expected signature of X, which makes pricing of so-called sig-payoffs, which are nothing else than linear functionals of the signature of $(t, S_t)_{t \geqslant 0}$ (see Lyons et al. (2020)), very efficient. A key step to obtain this result is to rewrite $S(\boldsymbol{\ell})_t$ itself as linear functionals of $\mathbb{X}_t$, a reformulation that we shall call *sig-model representation* (see Corollary 2.3.14). Not only pricing of sig-payoffs (see Section 2.3.3) can be traced back to the signature of X, but also the quadratic hedging problem can be solved by computing the explicit form of the mean variance hedging strategy in terms of $\mathbb{X}$ (see Section 2.3.4).

Note that by our UAT sig-payoffs approximate all J_1-continuous (possibly path-dependent) payoffs arbitrary well on compact set of paths, so that their prices and hedges can again be used as proxies. The corresponding coefficients of the sig-payoffs can be obtained by linear regression methods using typical price trajectories, e.g. generated by several appropriate models with jumps. For calibration purposes this procedure is usually not accurate enough (comparable with polynomial approximations, see also the Appendix in Cuchiero et al. (2023a)), but the obtained prices and hedges can for instance be used for variance reduction techniques. Let us remark here that one crucial advantage of signature based models, in particular the current Lévy-type signature model, is that calibration via Monte Carlo methods is highly efficient. Indeed, Lévy-type signature models share the same advantage as the continuous model of Cuchiero et al. (2023a), since the calibration task can be split into an offline sampling and a standard optimization. This is again due to the sig-model representation which allows to precompute all signature samples of $\mathbb{X}$ offline and thus to avoid an (Euler) simulation of trajectories in each optimization step. This adds further importance to the sig-model representation and suggests an interpretation of the compensator ν as hyperparameter (similarly to optimization tasks in the context of machine learning). Note that standard Monte Carlo methods can then further be enhanced by variance reduction using sig-payoffs as explained above.

Finally, in Section 2.3.5 we discuss equivalent measure changes that allow to keep the tractable structure of Lévy-type signature models also under the physical measure $\mathbb{P}$.

Let us finally remark that another strand of research related to finance where (càdlàg) rough paths play a significant role is robust and model-free finance, see for example Perkowski and Prömel (2016), Allan et al. (2023a) and in particular Allan et al. (2023b) for càdlàg rough paths foundations for robust finance. Indeed, in this context the theory of rough integration allows one to go beyond classical Itô-integrals and to overcome issues associated with null sets that are inherent in stochastic models, when e.g. dealing with volatility uncertainty. This in turn opens the door to a worst case analysis and a robustification of financial models and strategies. Note that we here do not pursue this direction but interpret all integrals in our financial model setup as Itô-integrals.

The remainder of the chapter is structured as follows. Section 2.2 is dedicated to the universal approximation theorem. In Section 2.3 we introduce Lévy-type signature models, discuss their universality properties and derive pricing and hedging formulas based on polynomial technology. All the technical proofs are given in the Appendix.

Throughout, unless explicitly mentioned, we let $p \in [1, 3)$. Moreover, we refer to Chapter 1 for the notation and main results concerning weakly geometric càdlàg rough paths and their signature.

2.2 Universal approximation theorem

In this section, we present a universal approximation theorems (UAT) for continuous functionals of weakly geometric càdlàg rough paths. We start by specifying the topology and the subspace of paths that will be considered.

2.2.1 The subspace of time-extended paths and the topology

Definition 2.2.1. Let (E, d) be a metric space and $D([0, 1]; E)$ the space of càdlàg paths on it. Denote by Λ the set of all strictly increasing bijections of $[0, 1]$ to itself. The *Skorokhod J_1-metric* on $D([0, 1]; E)$ is defined via

$$\sigma_\infty(X, Y) := \inf_{\lambda\in\Lambda}\{|\lambda| \vee \sup_{s\in[0,1]} d(X_{\lambda(s)}, Y_s)\}, \qquad X, Y \in D([0, 1]; E),$$

where, $|\lambda| := \sup_{s\in[0,1]} |\lambda(s) - s|$.

Remark 2.2.2. We endow the space of càdlàg paths considered throughout the chapter with the J_1-topology, induced by the metric σ_∞. In particular, if

$$(E, d) = (G^N(\mathbb{R}^d), d_{CC}),$$

$N \in \mathbb{N}$, the Skorokhod J_1-metric on $D([0, 1]; G^N(\mathbb{R}^d))$ reads

$$\sigma_\infty(\mathbf{X}, \mathbf{Y}) := \inf_{\lambda\in\Lambda}\{|\lambda| \vee \sup_{s\in[0,1]} d_{CC}(\mathbf{X}_{\lambda(s)}, \mathbf{Y}_\mathbf{s})\}, \qquad \mathbf{X}, \mathbf{Y} \in D([0, 1]; G^N(\mathbb{R}^d)).$$

Recall the notion of extended weakly geometric càdlàg rough path introduced given in Definition 1.2.16.

Definition 2.2.3. Let $p \in [1, 3)$ and $\mathrm{Id} : [0, 1] \to \mathbb{R}$ be the path such that $\mathrm{Id}_t := t$, for all $t \in [0, 1]$. The subspace of *time-extended weakly geometric càdlàg p-rough paths* is defined as the set of all the Id-extended weakly geometric càdlàg rough paths. We denote the space of such paths by $\widehat{D}^p([0, 1]; G^{[p]}(\mathbb{R}^{d+1}))$ and the subset of time-extended weakly geometric continuous p-rough path by $\widehat{C}^p([0, 1], G^{[p]}(\mathbb{R}^{d+1}))$.

Notation 2.2.4. In this chapter we use index -1 to denote the time component of a time-extended weakly geometric càdlàg rough paths.

Remark 2.2.5. (i) By Assumption 1.3.20, for all $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1}))$, it holds that $\widehat{\mathbf{X}}_0 = \mathbf{1} \in G^{[p]}(\mathbb{R}^{d+1})$.

(ii) The consideration of the set of time-extended weakly geometric rough paths is a key choice for showing condition (iii) in the proof of Proposition 2.2.6 (see also Corollary 2.2.10). For this purpose, the specific definition of the time-extended paths as in Definition 1.2.16 is essential as it allows to recover consistency with the classical Young integral (see Remark 1.2.18 (iii)).

2.2.2 UAT for continuous path functionals

We start by proving that continuous path functionals can be uniformly approximated on compact sets by linear functionals of the time-extended signature evaluated at the final time. Recall that the involved topological space has been defined in Definition 2.2.3.

Proposition 2.2.6. Let $p \in [1,3)$, $K \subset \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1}))$ a compact subset, which is bounded with respect to the p-variation norm, and $F : \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$ a continuous functional. For each $\widehat{\mathbf{X}} \in K$, denote by $\widehat{\mathbb{X}}$ its signature. Then, for every $\varepsilon > 0$ there exists a linear functional $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that

$$\sup_{\widehat{\mathbf{X}} \in K} |F(\widehat{\mathbf{X}}) - L(\widehat{\mathbb{X}}_1)| \leqslant \varepsilon.$$

The proof of Proposition 2.2.6 can be found in Appendix 2.A.1.

Remark 2.2.7. (i) On the subset of continuous paths the J_1-metric corresponds to the sup one defined via

$$d_\infty(\mathbf{X}, \mathbf{Y}) := \sup_{s \in [0,1]} d_{CC}(\mathbf{X}_s, \mathbf{Y}_\mathbf{s}), \qquad \mathbf{X}, \mathbf{Y} \in C([0,1]; G^{[p]}(\mathbb{R}^{d+1})).$$

Therefore, Theorem 2.2.6 states in particular that d_∞-continuous functionals of *continuous paths* can be approximated, uniformly on compact sets, by linear functionals of the time-extended signature evaluated at the final point. For different versions of approximations theorems for continuous functionals of *continuous* paths we refer to Levin et al. (2013), Kiraly and Oberhauser (2019), Lyons et al. (2020).

(ii) Requiring the set K to be bounded with respect to the p-variation norm is not redundant. For instance, on the subset of continuous paths the sup distance d_∞ is dominated by the p-variation one. This implies that not every J_1-compact set is bounded with respect to the p-variation norm.

(iii) Proposition 2.2.6 can be reformulated with respect to the rough paths variant of the Skorokhod M_1-topology introduced in Section 3.2 in Chevyrev and Friz (2019).

Indeed, by Theorem 3.13 in Chevyrev and Friz (2019) one can see that the set A given by (2.A.1) is a linear subspace of the continuous functions from K to $\mathbb{R}$, when on K the Skorokhod M_1-topology is considered. Notice however that neither version of these approximation results implies the other, as in general, the two Skorokhod topologies are not comparable. See Proposition 3.10 and Remark 3.6 in Chevyrev and Friz (2019) for more details.

(iv) Notice that the statement of Proposition 2.2.6 could be reformulated in terms of functionals on $D^p([0,1]; G^{[p]}(\mathbb{R}^d))$. Indeed, one just needs to notice that the map that associates to every weakly geometric càdlàg rough paths $\mathbf{X} \in D([0,1]; G^{[p]}(\mathbb{R}^d))$ its (unique) time-extended path $\widehat{\mathbf{X}}$ (see Section 1.2.4) is continuous with respect to the J_1-topology on both input and output space. From this and the reasoning in proof of Proposition 2.2.6, we can conclude that continuous functionals of weakly geometric càdlàg rough paths can be uniformly approximated on compact sets by linear functionals of the signature of the relative time-extended paths.

Next, for illustrative purposes, we also provide a simple adaptation of Proposition 2.2.6 to the semimartingales setting. Recall that for a semimartingale X its Marcus lift $\mathbf{X}$ is computed as in Proposition 1.4.1 and its signature $\mathbb{X}$ is the solution of the Marcus SDE (1.4.2).

Corollary 2.2.8. *Suppose that hypothesis of Proposition 2.2.6 hold for $p \in (2,3)$. Let $X = (X_t)_{t\in[0,1]}$ be an $\mathbb{R}^d$-valued semimartingale, set $\widehat{X} := (t, X_t)_{t\in[0,1]}$ and denote by $\widehat{\mathbf{X}}$ its Marcus lift and by $\widehat{\mathbb{X}}$ its signature. Then, for every $\varepsilon > 0$ there exists a linear functional $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that*

$$|F(\widehat{\mathbf{X}}(\omega)) - L(\widehat{\mathbb{X}}_1(\omega))| 1_{\{\widehat{\mathbf{X}}(\omega)\in K\}} \leqslant \varepsilon$$

for almost every $\omega \in \Omega$.

Proof. Since the Itô-integral with respect to semimartingales of finite variation coincides a.s. with the Young integral, it holds that $\widehat{\mathbf{X}}(\omega) \in \widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+1}))$ a.s. The claim follows from Proposition 2.2.6. □

Remark 2.2.9. Note that the continuous functional F in Corollary 2.2.8 can of course be a functional that only involves X and not its Marcus lift.

As a direct result of the proof of Proposition 2.2.6, the following statement concerning *uniqueness of the signature of time-extended weakly geometric càdlàg rough paths* holds true.

Corollary 2.2.10. *Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1}))$ and denote by $\widehat{\mathbb{X}}_1$ its signature evaluated at the final point 1. Then, $\widehat{\mathbb{X}}_1$ uniquely determines $\widehat{\mathbf{X}}$.*

Proof. This result follows directly from the proof of condition (iii) in Proposition 2.2.6. The reasoning therein does not depend on the set K, nor on the metric σ_∞, showing that a time-extended weakly geometric càdlàg p-rough path is completely characterized by its signature evaluated at the final time 1. $\square$

Remark 2.2.11. Let $\widehat{X} := (t, X_t)_{t\in[0,1]}$ be a time-extended càdlàg semimartingale with $\widehat{X}_0 = 0$. By Corollary 2.2.10 we get that a.s. $\widehat{\mathbb{X}}_1(\omega)$ uniquely determines $X(\omega)$.

We now present the main result of the first part of the chapter: an extended version of Proposition 2.2.6 which leads to an approximation of paths functionals not only on compact sets of paths but also uniformly in time. For its formulation, we first need to introduce the notion of a *stopped weakly geometric càdlàg rough path.*

Definition 2.2.12. Let $p \in [1,3)$. Given $\mathbf{X} \in D^p([0,1]; G^{[p]}(\mathbb{R}^d))$ and $t \in [0,1]$, we define the *stopped weakly geometric càdlàg rough path of* $\mathbf{X}$ stopped at time t as the càdlàg path $\mathbf{X}^t : [0,1] \to G^{[p]}(\mathbb{R}^d)$ given by

$$\mathbf{X}^t_u := \mathbf{X}_u 1_{\{u<t\}} + \mathbf{X}_t 1_{\{u\geqslant t\}}.$$

The proof of the following theorem can be found in Appendix 2.A.2.

Theorem 2.2.13. *Let $p \in [1,3)$, $K \subset \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1}))$ a compact subset, bounded with respect to the p-variation norm, and $F : D^p([0,1]; G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$ a continuous path functional. For each $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1]; G^{[p]}(\mathbb{R}^{d+1}))$, denote by $\widehat{\mathbb{X}}$ its signature. Then, for every $\varepsilon > 0$ there exists a linear functional $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that*

$$\sup_{t\in[0,1]} \sup_{\widehat{\mathbf{X}}\in K} |F(\widehat{\mathbf{X}}^t) - L(\widehat{\mathbb{X}}_t)| \leqslant \varepsilon.$$

Remark 2.2.14.

(i) Note that by Definition 2.2.12, for each $\mathbf{X} \in D^p([0,1]; G^{[p]}(\mathbb{R}^d))$ the map $\mathbf{X} \mapsto F(\mathbf{X}^t)$ just depends on $\mathbf{X}$ through its values $(\mathbf{X}_s)_{s\in[0,t]}$ on $[0,t]$. Therefore, for a given path functional $F : D^p([0,1]; G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$, it is always possible to associate a so-called non-anticipative path functional

$$\tilde{F} : [0,1] \times D^p([0,1]; G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$$

(see e.g., Dupire (2019), Cont et al. (2016); Cont and Fournie (2013); Ananova and Cont (2017)) such that for all $t \in [0,1]$, $\mathbf{X} \in D^p([0,1]; G^{[p]}(\mathbb{R}^d))$,

$$\tilde{F}(t, \mathbf{X}) := F(\mathbf{X}^t).$$

(ii) A similar result in the setting of *continuous* rough paths has been formulated in Kalsi et al. (2020), where, however, different techniques are used. Indeed, in contrast to our approach, Kalsi et al. (2020) consider non-anticipative functionals on the space of continuous rough paths equipped with an appropriate topology (fully clarified in Bayer et al. (2023)). Here, a similar reasoning could not be applied because the J_1-topology does not behave well under the restriction of the interval $[0, 1]$. This implies that the projection map φ introduced in Lemma A.1 in Bayer et al. (2023) is not continuous, and therefore the corresponding arguments do not hold.

Similarly as in Corollary 2.2.8, in the semimartingales setting, Theorem 2.2.13 reads as follows.

Corollary 2.2.15. *Assume that the hypothesis of Theorem 2.2.13 hold for $p \in (2, 3)$. Let $X = (X_t)_{t\in[0,1]}$ be an $\mathbb{R}^d$-valued semimartingale, set $\widehat{X} := (t, X_t)_{t\in[0,1]}$ and denote by $\widehat{\mathbf{X}}$ its Marcus lift and by $\widehat{\mathbb{X}}$ its signature. Then, for every $\varepsilon > 0$ there exists a linear functional $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that*

$$\sup_{t\in[0,1]} |F(\widehat{\mathbf{X}}^t(\omega)) - L(\widehat{\mathbb{X}}_t(\omega))| 1_{\{\widehat{\mathbf{X}}(\omega)\in K\}} \leqslant \varepsilon$$

for almost every $\omega \in \Omega$.

Finally, we conclude this section with some illustrative examples of path functionals for which Theorem 2.2.13 holds true.

Example 2.2.16. (i) For $|I| \leqslant 2$, consider $F : D([0, 1]; G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ given by

$$F(\mathbf{X}) := \sup_{s\leqslant 1} |\langle \epsilon_I, \mathbf{X}_s\rangle|$$

and note that for all $t \in [0, 1]$, $F(\mathbf{X}^t) = \sup_{s\leqslant t} |\langle \epsilon_I, \mathbf{X}_s\rangle|$. Furthermore, by Proposition VI.2.4 in Jacod and Shiryaev (1987), F is continuous with respect to the Skorokhod J_1-topology at all paths $\mathbf{X}$ that do not admit a jump at time 1. Let K be as in Theorem 2.2.13 and assume that all elements in K do not jump at 1. By Theorem 2.2.13, for every $\varepsilon > 0$ there is a map $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that

$$\sup_{t\in[0,1]} \sup_{\widehat{\mathbf{X}}\in K} |F(\widehat{\mathbf{X}}^t) - L(\widehat{\mathbb{X}}_t)| = \sup_{t\in[0,1]} \sup_{\widehat{\mathbf{X}}\in K} | \sup_{s\leqslant t} |\langle \epsilon_I, \widehat{\mathbf{X}}_s\rangle| - L(\widehat{\mathbb{X}}_t)| \leqslant \varepsilon.$$

(ii) Let Y be the solution of the Marcus-type RDE

$$dY_s = V(Y_s) \diamond d\mathbf{X}_s, \quad Y_0 = y_0, \tag{2.2.1}$$

for some suitable vector field V and let $\mathbf{X}$ be a weakly geometric càdlàg p-rough paths, for $p \in [2, 3)$. Consider the path functional $F : D^p([0, 1]; G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ given by

$$F(\mathbf{X}) := Y_1.$$

For each $t \in [0, 1]$, denote by $\widehat{Y}_1^t$ the solution map evaluated at time 1 of the Marcus-type RDE (2.2.1) driven by the stopped time-extended path $\widehat{\mathbf{X}}^t \in D^p([0, 1]; G^2(\mathbb{R}^{d+1}))$.

Observe that $F(\widehat{\mathbf{X}}^t) = \widehat{Y}_1^t = \widehat{Y}_t$. Moreover, by Proposition 3.18 in Chevyrev and Friz (2019) and a similar argument as in the proof of Proposition 2.2.6, F is continuous with respect to the Skorokhod J_1-topology on sets of bounded p-variation norm. Thus, for each $\varepsilon > 0$ and K satisfying the conditions of Theorem 2.2.13, there exists $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that

$$\sup_{t\in[0,1]} \sup_{\widehat{\mathbf{X}}\in K} |F(\widehat{\mathbf{X}}^t) - L(\widehat{\mathbb{X}}_t)| = \sup_{t\in[0,1]} \sup_{\widehat{\mathbf{X}}\in K} |\widehat{Y}_t - L(\widehat{\mathbb{X}}_t)| \leqslant \varepsilon.$$

Therefore, solutions of Marcus-type RDEs can be approximated uniformly in time by linear functionals of the signature of the driving signal.

As a side note, observe that by Theorem 5.3 in Friz and Zhang (2017), the same type of argument applies to solutions of RDEs of *forward-type*. Notice in particular that this concept of RDE is in line with the Itô-theory of càdlàg semimartingales (see e.g. Proposition 6.9 in Friz and Zhang (2017)).

2.3 Lévy-type signature models

2.3.1 The market's primary process and the model

In this section we introduce the class of Lévy-type signature models. The key object is a multidimensional process, which will have the role of encoding all the randomness of the model. Because of this, it will be called the *market's primary process*.

For its definition, we consider a standard Brownian motion W and a homogeneous Poisson random measure μ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$, whose filtration $(\mathcal{F}_t)_{t\in[0,T]}$, is the augmented one generated by W and μ. We suppose that μ and W are independent and that μ has intensity measure

$$\nu(dt, dx) = dt \times F(dx).$$

Here, dt denotes the Lebesgue measure on $([0,T], \mathcal{B}([0,T]))$ and $F(dx)$ a σ-finite measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfying $F(\{0\}) = 0$ and

$$\int_{\mathbb{R}} |x|^k F(dx) < \infty \qquad \text{for all } k \geqslant 2. \tag{2.3.1}$$

Definition 2.3.1. Fix $d_X \geqslant 2$. The *market's primary process* is the $\mathbb{R}^{d_X+2}$-valued process X given by

$$X_t := \left(t, W_t, \int_0^t \int_{\mathbb{R}} x(\mu - \nu)(ds, dx), \int_0^t \int_{\mathbb{R}} x^2 \mu(ds, dx), \dots, \int_0^t \int_{\mathbb{R}} x^{d_X} \mu(ds, dx)\right).$$

Observe that we do not use the $\widehat{\cdot}$ notation, even though the first component of X is t. Moreover, from now on we will use the following numbering for the components of the

$\mathbb{R}^{d_X+2}$-valued market's primary process X

$$\begin{aligned}
X_t^{-1} &= t,\\
X_t^{0} &= W_t,\\
X_t^{1} &= \int_0^t \int_{\mathbb{R}} x(\mu-\nu)(ds,dx),\\
X_t^{k} &= \int_0^t \int_{\mathbb{R}} x^k \mu(ds,dx), \qquad \text{for } 2 \leqslant k \leqslant d_X.
\end{aligned}$$

The next proposition states that the market's primary process is a Lévy process.

Proposition 2.3.2. The $\mathbb{R}^{d_X+2}$-valued market's primary process X introduced in Definition 2.3.1 is a Lévy process and its triplet (b^X, C^X, F^X), associated with the "truncation function" (see Definition II.2.3 and Corollary II.2.38 in Jacod and Shiryaev (1987)) $\chi(y) = y$, is given by

$$b^X = (1, 0, 0, \int_{\mathbb{R}} x^2 F(dx), \ldots, \int_{\mathbb{R}} x^{d_X} F(dx)),$$

$C^X_{i,j} = 0$ if $i \neq 0$ or $j \neq 0$, $C^X_{0,0} = 1$, and $F^X(dy) = h_* F(dx)$, where

$$h(x) := (0, 0, x, x^2, x^3, \ldots, x^{d_X}).$$

Proof. Observe that X is the sum of the two independent $\mathbb{R}^{d_X+2}$-valued processes Y and Z given by $Y_t = (t, W_t, 0, \ldots, 0)$ and

$$Z_t = \Big(0, 0, \int_0^t \int_{\mathbb{R}} x(\mu-\nu)(ds,dx), \int_0^t \int_{\mathbb{R}} x^2 \mu(ds,dx), \ldots, \int_0^t \int_{\mathbb{R}} x^{d_X} \mu(ds,dx)\Big).$$

Using the same numbering for the components, we get that Y is a Lévy process with triplet $(b^Y, C^Y, 0)$, where $b^Y = (1, 0, \ldots, 0)$, $C^Y_{i,j} = 0$ for $i \neq 0$ or $j \neq 0$ and $C^Y_{0,0} = 1$.

Next, we verify that Z is a well-defined Lévy process. Let M be the random measure on $\mathbb{R}_+ \times \mathbb{R}^{d_X+2}$, given by $M := H_*\mu$ for $H(t,x) := (t, h(x))$. Since μ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$, and H is a measurable map, M is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^{d_X+2}$ with intensity

$$\nu^Z(dt, dy) = H_*(dt \times F(dx)) = dt \times (h_* F(dx)).$$

Moreover, since $(h_*F)(\{0\}) = F(\{0\}) = 0$, and

$$\int_{\mathbb{R}^{d_X+2}} (1 \wedge \|y\|^2) h_* F(dy) = \int_{\mathbb{R}} (1 \wedge \|h(x)\|^2) F(dx) < \infty,$$

we also have that h_*F is a Lévy measure on $\mathbb{R}^{d_X+2}$. Using that

$$Z = b^Z t + \int_0^t \int_{\mathbb{R}^{d_X+2}} y(M - \nu^Z)(ds, dy),$$

for $b^Z = (0,0,0,\int_{\mathbb{R}} x^2 F(dx),\dots,\int_{\mathbb{R}} x^{d_X} F(dx))$, we can conclude that Z is a well-defined Lévy process, whose triplet with respect to $\chi(y) = y$ is given by $(b^Z, 0, h_* F))$. The claim follows. $\square$

Remark 2.3.3. Observe that by condition (2.3.1), the $\mathbb{R}^{d_X+2}$-valued Lévy process X introduced in Definition 2.3.1 has finite k-moments, for all $k \in \mathbb{N}$.

Notation 2.3.4. Throughout the rest of the chapter, for a multi-index $I = (i_1,\dots,i_n)$ we write $\Sigma(I) := i_1 + i_2 + \cdots + i_n$.

Remember that for Since X is a Lévy process, following the discussion of Section 1.4, we can compute its signature $\mathbb{X}$ by solving the $T((\mathbb{R}^{d_X+2}))$-valued Marcus SDE given by

$$d\mathbb{X} = \mathbb{X} \otimes \diamond dX, \qquad \mathbb{X}_0 = (1,0,0,\dots) \in T((\mathbb{R}^{d_X+2})),$$

whose component-wise version is given by

$$\begin{aligned}\langle \epsilon_I, \mathbb{X}_t\rangle = &\int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s-}\rangle dX_s^{i_{|I|}} + 1_{\{i_{|I|-1}=i_{|I|}=0\}} \frac{1}{2}\int_0^t \langle \epsilon_{I''}, \mathbb{X}_{s-}\rangle ds \\ &+ \sum_{0<s\leqslant t} \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} \langle \epsilon_{I_1}, \mathbb{X}_{s-}\rangle \Delta X_s^{\Sigma(I_2)} 1_{\{I_2 \in \{1,\dots,d_X\}^{|I_2|}\}},\end{aligned} \tag{2.3.2}$$

for each multi-index I with entries in $\{-1,0,1,\dots,d_X\}$.

We are now interested in computing the expected signature of $\mathbb{X}$, i.e., $\mathbb{E}[\langle \epsilon_I, \mathbb{X}\rangle]$ for each multi-index I. One can prove (see e.g. Friz and Shekhar (2017)), that the expected signature of an $\mathbb{R}^d$-valued Lévy process is finite whenever all the moments of the Lévy measure are finite. Moreover, it can be computed analytically using the method in Friz and Shekhar (2017) or by recognizing that $\mathbb{X}^N$ is a polynomial process on $T^N(\mathbb{R}^d)$ for each $N \in \mathbb{N}$, and applying the so-called moment formula (see Cuchiero et al. (2012), Filipović and Larsson (2020)). Recall that for each $\mathbf{x} \in T((\mathbb{R}^d))$ such that $\pi_0(\mathbf{x}) = 0$, we use the notation

$$\exp(\mathbf{x}) := 1 + \sum_{k=1}^{\infty} \frac{\mathbf{x}^{\otimes k}}{k!}. \tag{2.3.3}$$

The proof of the following result via polynomial technology can be found in Appendix 2.B.1.

Proposition 2.3.5. Fix $d \in \mathbb{N}$. Let $b^L \in \mathbb{R}^d$, $C^L \in (\mathbb{R}^d)^{\otimes 2}$ a positive definite $d \times d$ symmetric matrix and F^L a σ-finite measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that $F^L(\{0\}) = 0$ and

$$\int_{\mathbb{R}^d} \|x\|^k F^L(dx) < \infty \quad \text{for all } k \geqslant 2. \tag{2.3.4}$$

Let L be an $\mathbb{R}^d$-valued Lévy process with triplet (b^L, C^L, F^L) with respect to the “truncation function” $\chi(x) = x$. For each $N \in \mathbb{N}$, the truncated signature $\mathbb{L}^N$ of L is a

$T^N(\mathbb{R}^d)$-valued polynomial process. This in particular implies that all the moments of $\mathbb{L}^N$ exist and are finite. Moreover,

$$\mathbb{E}[\mathbb{L}_t] = \exp\left(\left(b^L + \frac{C^L}{2} + \int_{\mathbb{R}^d} (\exp(x) - 1 - x) F^L(dx)\right) t\right). \tag{2.3.5}$$

Remark 2.3.6. Let us restate expression (2.3.5) in a more explicit form. Observe that since

$$\int_{\mathbb{R}^d} (\exp(x) - 1 - x) F^L(dx) = \left(0, 0, \frac{1}{2}\int_{\mathbb{R}^d} x^{\otimes 2} F^L(dx), \ldots, \frac{1}{k!}\int_{\mathbb{R}^d} x^{\otimes k} F^L(dx), \ldots\right),$$

(2.3.5) can be written as $\mathbb{E}[\mathbb{L}_t] = \exp(Qt)$ for

$$Q := \left(0, b^L, \frac{1}{2}\left(C^L + \int_{\mathbb{R}^d} x^{\otimes 2} F^L(dx)\right), \ldots, \frac{1}{k!}\int_{\mathbb{R}^d} x^{\otimes k} F^L(dx), \ldots\right).$$

Applying once again the definition of the exponential function (2.3.3), we obtain that for $|I| > 0$ it holds

$$\mathbb{E}[\langle \epsilon_I, \mathbb{L}_t\rangle] = \sum_{k=1}^{n} \frac{t^k}{k!} \sum_{\epsilon_{I_1}\otimes\cdots\otimes\epsilon_{I_k} = \epsilon_I} \langle \epsilon_{I_1}, Q\rangle \cdots \langle \epsilon_{I_k}, Q\rangle,$$

where $\langle \epsilon_\emptyset, Q\rangle = 0$, $\langle \epsilon_{j_1}, Q\rangle = b^L_{j_1}$, $\langle \epsilon_{j_1,j_2}, Q\rangle = \frac{1}{2}(C^L_{j_1,j_2} + \int_{\mathbb{R}^d} x_{j_1}x_{j_2} F^L(dx))$, and

$$\langle \epsilon_J, Q\rangle = \frac{1}{k!}\int_{\mathbb{R}^d} x_{j_1}\cdots x_{j_m} F^L(dx)$$

for each $J = (j_1, \ldots, j_m)$ for $m \geqslant 3$. For example, setting $I = (1,3)$ we get that

$$\begin{aligned}\mathbb{E}[\langle \epsilon_{(1,3)}, \mathbb{L}_t\rangle] &= t\langle \epsilon_{(1,3)}, Q\rangle + \frac{t^2}{2}\langle \epsilon_1, Q\rangle\langle \epsilon_3, Q\rangle \\ &= \frac{t}{2}\left(C^L_{1,3} + \int_{\mathbb{R}^d} x_1 x_3 F^L(dx)\right) + \frac{t^2}{2} b^L_1 b^L_3.\end{aligned}$$

Remark 2.3.7. Observe that following the argument of Lemma A.2 in Cuchiero et al. (2023a), we can use the i.i.d. increments of a Lévy process to get

$$\mathbb{E}[\langle \epsilon_I, \mathbb{L}_{s+t}\rangle | \mathcal{F}_s] = \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2} = \epsilon_I} \langle \epsilon_{I_1}, \mathbb{L}_s\rangle \mathbb{E}[\langle \epsilon_{I_2}, \mathbb{L}_t\rangle].$$

Using this representation, we can see that the formula obtained in Proposition 2.3.5 also provides an explicit representation of the conditional moments of $\mathbb{L}$.

We are now ready to define a Lévy-type signature model, which will represent the price process of an arbitrary financial asset under an equivalent martingale measure $\mathbb{Q}$.

Definition 2.3.8. Let X be the market's primary process introduced in Definition 2.3.1 and $\mathbb{X}$ its signature. We define a *Lévy-type signature model* as follows

$$S(\boldsymbol{\ell})_t = S_0 + \int_0^t \Big(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-} \rangle \Big) dW_s + \int_0^t \int_{\mathbb{R}} \Big(\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{s-} \rangle \Big) x(\mu - \nu)(ds, dx), \tag{2.3.6}$$

where $\ell_W^J, \ell_\nu^J \in \mathbb{R}$ and J are multi-indices of length at most $n \in \mathbb{N}$ whose indices are in $\{-1, 0, 1, \ldots, d\}$ with $d \leqslant d_X$.

Remark 2.3.9. (i) The motivation for this type of model comes from Theorem 2.2.13 proved in Section 2.2. Indeed, note that any continuous path functional of the enhanced market's primary underlying process can be approximated by linear functionals of its signature. The characteristics of $S(\boldsymbol{\ell})$ can thus be interpreted as approximations of continuous path functionals F of $\mathbf{X}$, with $\mathbf{X}$ denoting the Marcus lift of X as computed in Proposition 1.4.1.

(ii) The no-arbitrage condition is satisfied. In fact, the process $S(\boldsymbol{\ell})$ is a square integrable martingale. This is a direct consequence of Proposition 2.3.2, Remark 2.3.3, Proposition 2.3.5, and the shuffle product formula according to which every polynomial on the signature admits a linear representation.

(iii) The price process $S(\boldsymbol{\ell})$ is a càdlàg semimartingale whose characteristics, with respect to the "truncation function" $\chi(x) = x$, are given by

$$B_t^S = 0, \quad C_t^S = \int_0^t \Big(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-} \rangle \Big)^2 ds, \quad \nu^S(dt \times dx) = K_t(dx) \times dt, \tag{2.3.7}$$

where

$$K_t(A) = \int_{\mathbb{R}} 1_{A\setminus\{0\}} \Big(\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{t-} \rangle x \Big) F(dx).$$

(iv) Notice that the jumps of the price process $S(\boldsymbol{\ell})$ are proportional to the jumps of the (signature of) the market's primary process X. The advantage of this specification lies in the possibility of writing linear functionals of the signature of $S(\boldsymbol{\ell})$ in terms of the signature of X (see Theorem 2.3.17 and the implications in terms of pricing and hedging stated in Corollary 2.3.20 and Theorem 2.3.22). A more general jump size specification could also be achieved by considering in (2.3.6) more than one pure jump driving signal specified in terms of another homogeneous Poisson random measure $\tilde{\mu}$. The aforementioned tractability property can then be recovered provided that in the definition of the market's primary process X, the moments of $\tilde{\mu}$ and cross moments of μ and $\tilde{\mu}$ are considered as well.

2.3.2 Sig-model representation

The purpose of this section is to achieve an alternative representation of a Lévy-type signature model, namely to express $S(\boldsymbol{\ell})$ itself as linear functional of the signature of the market's primary process. We will see that this new representation, that will be called sig-model representation, is more suitable for addressing the problem of pricing, and model calibration to both time-series data and option prices. Indeed, the sig-model representation is fully in line with the representation of Cuchiero et al. (2023a) in the simpler continuous case, so that all the arguments from there concerning tractability for pricing and calibration apply. We start by representing Itô integrals as linear functionals of the signature of X.

Definition 2.3.10. For each multi-index I with indices in $\{-1,\dots,d\}$ and each $j \in \{-1,\dots,d\}$, we consider the following multi-index transformation

$$
\begin{aligned}
(\epsilon_I;\epsilon_j)^\sim := \epsilon_I \otimes \epsilon_j - \frac{1}{2}\epsilon_I \otimes \epsilon_{-1} 1_{\{i_{|I|}=j=0\}} \\
+ \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|)(\epsilon_{I_1} \otimes \epsilon_{\Sigma(I_2)+j}) 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}},
\end{aligned}
$$

where

$$
\alpha(|I_2|) := \sum_{k=1}^{|I_2|} (-1)^k \alpha(|I_2|,k),
$$

and denoting by $\mathcal{I}(r,k)$ the set of all multi-indices $J \in \mathbb{N}^k$ such that $\Sigma(J) = r$,

$$
\alpha(r,k) := \sum_{J\in\mathcal{I}(r,k)} \prod_{i=1}^{k} \frac{1}{(j_i+1)!}.
$$

Remark 2.3.11. Fix a multi-index I with entries in $\{-1,\dots,d\}$ and $j \in \{-1,\dots,d\}$ and observe that $(\epsilon_I;\epsilon_j)^\sim = \sum_{r=1}^{R} \beta_r \epsilon_{I_r}$, for some $R \in \mathbb{N}$, $\beta_r \in \mathbb{R}$ and multi-indices I_r with indices in $\{-1,0,\dots,\tilde{d_X}\}$, for some $\tilde{d_X} \in \mathbb{N}$. Notice that in principle $\tilde{d_X}$ may not satisfy the condition $\tilde{d_X} \leqslant d_X$, leading to a difficulty in interpreting the expression $\langle(\epsilon_I;\epsilon_j)^\sim, \mathbb{X}_t\rangle$. Therefore, whenever this transformation is used in the rest of the chapter, we provide sufficient conditions for it to be well-defined.

The proof of the next proposition can be found in Appendix 2.B.2.

Proposition 2.3.12. Fix $n \in \mathbb{N}$, $I \in \{-1,0,1,\dots,d\}^n$ and $j \in \{-1,0,1,\dots,d\}$. Denote by I_1, I_2 be the unique (possibly empty) multi-indices such that $\epsilon_I = \epsilon_{I_1} \otimes \epsilon_{I_2}$ with $i_{|I_1|} \in \{0,-1\}$ and $I_2 \in \{1,\dots,d\}^{|I_2|}$. If either $j \in \{-1,0\}$ or $j \in \{1,\dots,d\}$ and $\Sigma(I_2)+j \leqslant d_X$, we get that

$$
\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j = \langle(\epsilon_I;\epsilon_j)^\sim, \mathbb{X}_t\rangle.
$$

Remark 2.3.13. Fix $j \geqslant 1$ and observe that by Proposition 2.3.12 it holds

$$\begin{aligned}\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j &= \langle (\epsilon_I; \epsilon_j)^\sim, \mathbb{X}_t\rangle \\ &= \langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t\rangle + \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|)\langle \epsilon_{I_1} \otimes \epsilon_{\Sigma(I_2)+j}, \mathbb{X}_t\rangle 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}},\end{aligned}$$

for some coefficients $\alpha(|I_2|) \in \mathbb{R}$. Since

$$\max_{|I_2|\leqslant n} \Sigma(I_2) + j = nd + j,$$

we can conclude that $nd + j \leqslant d_X$ is a necessary and sufficient condition for being able to write $\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j$ as a linear functional of the signature of X for each multi-index $|I| \leqslant n$.

The sig-model representation is a direct application of Proposition 2.3.12. We thus omit the proof of the next corollary.

Corollary 2.3.14. *If $nd + 1 \leqslant d_X$, then the Lévy-type signature model defined in (2.3.6) admits the following sig-model representation*

$$S(\boldsymbol{\ell})_t = S_0 + \sum_{|J|\leqslant n} \big(\ell_W^J \langle (\epsilon_J; \epsilon_0)^\sim, \mathbb{X}_t\rangle + \ell_\nu^J \langle (\epsilon_J; \epsilon_1)^\sim, \mathbb{X}_t\rangle\big). \tag{2.3.8}$$

Remark 2.3.15. (i) Notice that accordingly to Proposition 2.3.12, the condition given by $nd + 1 \leqslant d_X$ could be relaxed if some coefficients ℓ_ν^J are equal to zero.

(ii) Note that thanks to the sig-model representation, $S(\boldsymbol{\ell})_t$ itself can be thought as an approximation of some continuous path functional of the enhanced market's underlying process.

(iii) Moreover, it allows to reduce the calibration to time-series data to a linear regression problem in the coefficients ℓ similarly as in Section 4.1 of Cuchiero et al. (2023a). Indeed, suppose that a trajectory of the market's primary process is observable, which means that we can extract from time-series asset price data the trajectory of the "market Brownian motion" W and the "market Poisson random measure" μ, together with its compensator. Note here, that X is specified under $\mathbb{Q}$, which means that the $\mathbb{Q}$-Brownian motion and the $\mathbb{Q}$-Poisson random measure already account for the market price of risk. In other words we extract by means of quadratic variation estimators similarly as in Example 4.1 of Cuchiero et al. (2023a) a $\mathbb{P}$-Brownian with drift (corresponding to the $\mathbb{Q}$-Brownian motion) and a $\mathbb{P}$-Poisson random measure with potentially a different intensity ($\mathbb{Q}$-Poisson random measure). Once X is extracted, its signature can be computed along the observed trajectory. Using a simple linear regression, the coefficients ℓ_W^J can then be matched with the market's continuous quadratic variation since the continuous quadratic variation of $S(\boldsymbol{\ell})$ is given by (2.3.7). By means of (2.3.8), the remaining parameters ℓ_ν^J can then be estimated by performing a linear regression directly on the price data.

(iv) For pricing and calibrating to option data the sig-model representation also has a crucial advantage. Indeed, Monte-Carlo prices of any (possibly path-dependent) option on $S(\boldsymbol{\ell})$ can be easily obtained by computing all signature samples of $\mathbb{X}$ offline. Calibration to option data then reduces to a simple optimization task to infer ℓ_W^J and ℓ_ν^J from market data. In contrast to classical models or when directly working with (2.3.6), the sig-model representation does not involve any Euler discretization in each optimization step, which is a crucial advantage. We refer to Cuchiero et al. (2023a, 2024b) where this advantage has also been exploited.

Example 2.3.16. In Remark 2.3.13 we have explained why in order to achieve the sig-model representation one has to include some moments of the jump-measure μ in the definition of the primary process X (see Definition 2.3.1). We now want to show that this necessary condition is no longer needed if the last component of X consists of a standard Poisson process.

In this context ν satisfies

$$\nu(dt, dx) = dt \times \lambda\delta_1(dx),$$

for some $\lambda > 0$ where δ_1 denotes the Dirac measure. Set then

$$X_t : t \mapsto \left(t, W_t, \int_0^t \int_{\mathbb{R}} x\mu(ds, dx)\right), \tag{2.3.9}$$

and observe that for all $k \geqslant 2$ and $t \in [0, 1]$

$$\int_0^t \int_{\mathbb{R}} x^k \mu(ds, dx) = \int_0^t \int_{\mathbb{R}} x\mu(ds, dx) = \langle \epsilon_1, X_t \rangle.$$

Notice that to simplify the notation we do not compensate the sum of jumps in the jump-component of X. As a consequence, for $j = 1$ the multi-index transformation introduced in Definition 2.3.10 simplifies to

$$(\epsilon_I; \epsilon_1)^{\sim} := \epsilon_I \otimes \epsilon_1 + \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \alpha(|I_2|)(\epsilon_{I_1} \otimes \epsilon_1) 1_{\{I_2 \in \{1, \ldots, 1\}^{|I_2|}\}}.$$

Observe in particular that the linear combination describing $(\epsilon_I; \epsilon_1)^{\sim}$ just involves indices with values in $\{-1, 0, 1\}$. The pairing with $\mathbb{X}_t$, where X is as in (2.3.9), is thus well-defined. By Proposition 2.3.12 we can conclude that the Lévy-type signature model

$$\begin{aligned} S(\boldsymbol{\ell})_t = S_0 &+ \int_0^t \Bigg(\sum_{|J| \leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-} \rangle \Bigg) dW_s \\ &+ \int_0^t \int_{\mathbb{R}} \Bigg(\sum_{|J| \leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{s-} \rangle \Bigg) x(\mu - \nu)(ds, dx), \end{aligned} \tag{2.3.10}$$

with $\ell_W^J, \ell_\nu^J \in \mathbb{R}$ and J multi-index with values in $\{-1,0,1\}$ can be rewritten as

$$S(\boldsymbol{\ell})_t = S_0 + \sum_{|J|\leqslant n} \ell_W^J \langle (\epsilon_J; \epsilon_0)^{\sim}, \mathbb{X}_t\rangle + \sum_{|J|\leqslant n} \ell_\nu^J \langle (\epsilon_J; \epsilon_1)^{\sim} - \lambda(\epsilon_J; \epsilon_{-1})^{\sim}, \mathbb{X}_t\rangle.$$

It is important to notice that the improvement of tractability achieved in this setting comes with a loss in generality. In particular, if a jump occurs at time t its size has to be

$$\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle,$$

and is thus known in t^-.

2.3.3 Pricing of sig-payoffs

In this section, we address the problem of pricing a contingent claim in Lévy-type signature models. As an application of Proposition 2.2.6, we consider contingent claims whose payoff is represented by a linear functional of the signature of the price process extended with time (extended price process henceforth), i.e., so-called sig-payoffs, already analyzed in Lyons et al. (2020), Arribas et al. (2021), Cuchiero et al. (2023a) and the references therein. In order to give a clear treatment of the problem, we start by computing the signature of the time-extended price process

$$\widehat{S}(\boldsymbol{\ell})_t := (t, S(\boldsymbol{\ell})_t), \tag{2.3.11}$$

whose components will be denoted by

$$\widehat{S}(\boldsymbol{\ell})_t^{-1} = t, \qquad \widehat{S}(\boldsymbol{\ell})_t^{1} = S(\boldsymbol{\ell})_t.$$

As $S(\boldsymbol{\ell})$ is given by (2.3.6), the extended price process $\widehat{S}(\boldsymbol{\ell})$ is an $\mathbb{R}^2$-valued càdlàg semimartingale. Hence, its signature $\widehat{\mathbb{S}}(\boldsymbol{\ell})$ can be computed following Proposition 1.4.2, and thus

$$\begin{aligned}\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t\rangle = &\int_0^t \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle d\widehat{S}(\boldsymbol{\ell})_s^{i_{|I|}} + 1_{\{i_{|I|}=i_{|I|-1}=1\}} \frac{1}{2}\int_0^t \langle \epsilon_{I''}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle d[S(\boldsymbol{\ell}), S(\boldsymbol{\ell})]_s^c \\ &+ \sum_{0<s\leqslant t} \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} \langle \epsilon_{I_1}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle \Delta S(\boldsymbol{\ell})_s^{|I_2|} 1_{\{I_2=(1,\dots,1)\}},\end{aligned} \tag{2.3.12}$$

for each I with entries in $\{-1,1\}$.

Now, we show how to express the signature of $\widehat{S}(\boldsymbol{\ell})$ in terms of the signature of the underlying process X. This result is again based on the multi-index transformation introduced in Definition 2.3.10. Before making the argument precise, let us make the following observations and introduce some useful notation.

Consider the price process defined as in equation (2.3.6). Let $M := \sum_{i=0}^{n}(d+2)^i$ denote the number of possible multi-indices with length at most n and indices in $\{-1,0,1,\dots,d\}$. Observe that we are considering also the empty index $\varnothing$. Define

$$\{\epsilon_{J^h} : 1 \leqslant h \leqslant M\},$$

to be an ordered version of $\{\epsilon_J : |J| \leqslant n\}$. For instance, we may consider the order with respect to the length first and alphabetically next. Then,

$$S(\boldsymbol{\ell})_t = S_0 + \sum_{h=1}^{M} \ell_h^W \langle (\epsilon_{J^h}; \epsilon_0)^\sim, \mathbb{X}_t \rangle + \sum_{h=1}^{M} \ell_h^\mu \langle (\epsilon_{J^h}; \epsilon_1)^\sim, \mathbb{X}_t \rangle.$$

For $\alpha = (\alpha^1, \alpha^2, \dots, \alpha^M) \in \mathbb{N}^M$ we write

$$\epsilon^{\sqcup\!\sqcup\alpha} := \frac{1}{\alpha!}(\sqcup\!\sqcup^{\alpha^1}\epsilon_{J^1}) \sqcup\!\sqcup (\sqcup\!\sqcup^{\alpha^2}\epsilon_{J^2}) \sqcup\!\sqcup \cdots \sqcup\!\sqcup (\sqcup\!\sqcup^{\alpha^M}\epsilon_{J^M}),$$

and for $\ell_{1,\dots,M} := (\ell_1, \ell_2, \dots, \ell_M) \in \mathbb{R}^M$, we also set $\ell_{1,\dots,M}^\alpha := \ell_1^{\alpha^1} \cdots \ell_M^{\alpha^M}$.

We are now ready to state the announced result. It will become clear later that this is the key result to solve analytically the pricing problem for sig-payoffs. The proof can be found in Appendix 2.B.3, where a discussion on the assumption on the lower bound on d_X is also provided.

Theorem 2.3.17. *Let $\widehat{S}(\boldsymbol{\ell})$ be the time-extended price process defined in (2.3.11) and $\widehat{\mathbb{S}}(\boldsymbol{\ell})$ its signature. Let I be a multi-index with indices in $\{-1,1\}$. If $d_X \geqslant |I|(dn+1)$, then there is a linear combination of multi-indices $U_I(\boldsymbol{\ell})$ with indices in $\{-1,0,1,\dots,d_X\}$, such that the following equality holds*

$$\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle = \langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle. \tag{2.3.13}$$

Furthermore, $\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle$ is a polynomial of degree $|I|$ in $(\ell_{1,\dots,M}^W, \ell_{1,\dots,M}^\nu) \in \mathbb{R}^{2M}$, and $U_I(\boldsymbol{\ell})$ can be computed recursively on the length $|I|$ of I setting $\epsilon_{U_\varnothing(\boldsymbol{\ell})} = \epsilon_\varnothing$ and

$$\begin{aligned}
\epsilon_{U_I(\boldsymbol{\ell})} =& (\epsilon_{U_{I'}(\boldsymbol{\ell})}; \epsilon_{-1})^\sim 1_{\{i_{|I|}=-1\}} + \sum_{\alpha\in\mathbb{N}^M, \Sigma(\alpha)=1} (\ell_{1,\dots,M}^W)^\alpha (\epsilon_{U_{I'}(\boldsymbol{\ell})} \sqcup\!\sqcup \epsilon^{\sqcup\!\sqcup\alpha}; \epsilon_0)^\sim 1_{\{i_{|I|}=1\}} \\
&+ \sum_{\alpha\in\mathbb{N}^M, \Sigma(\alpha)=1} (\ell_{1,\dots,M}^\nu)^\alpha (\epsilon_{U_{I'}(\boldsymbol{\ell})} \sqcup\!\sqcup \epsilon^{\sqcup\!\sqcup\alpha}; \epsilon_1)^\sim 1_{\{i_{|I|}=1\}} \\
&+ \sum_{\alpha\in\mathbb{N}^M, \Sigma(\alpha)=2} (\ell_{1,\dots,M}^W)^\alpha (\epsilon_{U_{I''}(\boldsymbol{\ell})} \sqcup\!\sqcup \epsilon^{\sqcup\!\sqcup\alpha}; \epsilon_{-1})^\sim 1_{\{|I|>1\}} 1_{\{(i_{|I|-1}, i_{|I|})=(1,1)\}} \\
&+ \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \sum_{\alpha\in\mathbb{N}^M, \Sigma(\alpha)=|I_2|} (\ell_{1,\dots,M}^\nu)^\alpha (\epsilon_{U_{I_1}(\boldsymbol{\ell})} \sqcup\!\sqcup \epsilon^{\sqcup\!\sqcup\alpha}; \epsilon_{|I_2|})^\sim 1_{\{|I_2|>1\}} 1_{\{I_2=(1,\dots,1)\}}.
\end{aligned}$$

Remark 2.3.18. Consider the setting of Example 2.3.16. From an inspection of the proof of Theorem 2.3.17, it can be deduced that even in this case we do not need to consider higher moments of the jump measure in the definition of the market's primary process. This means that for every multi-index I with indices in $\{-1, 1\}$, there always exists a linear combination of multi-indices $U_I(\boldsymbol{\ell})$ with entries in $\{-1, 0, 1\}$, such that

$$\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle = \langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle. \tag{2.3.14}$$

Here $S(\boldsymbol{\ell})$ and X denote the processes defined in equation (2.3.10) and (2.3.9) respectively. More precisely, the representation in (2.3.14) holds with $(\epsilon_I; \epsilon_1)^{\sim}$ replaced by $(\epsilon_I; \epsilon_1)^{\sim} - \lambda(\epsilon_I; \epsilon_{-1})^{\sim}$ and $(\epsilon_I; \epsilon_k)^{\sim}$ replaced by $(\epsilon_I; \epsilon_1)^{\sim}$ for each $k \geqslant 2$.

Finally, we recall the notion of sig-payoffs as introduced in Lyons et al. (2020) and address the problem of computing their arbitrage free prices.

Definition 2.3.19. Let $\widehat{S}(\boldsymbol{\ell})$ be the time-extended price process defined in (2.3.11) and $\widehat{\mathbb{S}}(\boldsymbol{\ell})$ its signature. We define a *sig-payoff* with maturity $T > 0$ as the payoff C^S that pays to the holder of the claim an amount equal to

$$C^S = \sum_{|I| \leqslant m} h^I \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle, \tag{2.3.15}$$

where, $h^I \in \mathbb{R}$ and I are multi-indices of length at most $m \in \mathbb{N}$ with indices in $\{-1, 1\}$.

Observe that by Theorem 2.3.17, Proposition 2.3.2, and Proposition 2.3.5 the random variable C^S given in (2.3.15) is square integrable.

Applying Theorem 2.3.17, we are now ready to provide a pricing formula for sig-payoffs.

Corollary 2.3.20. *Let*

$$C^S = \sum_{|I| \leqslant m} h^I \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle$$

be a sig-payoff with maturity $T > 0$. If $d_X \geqslant m(nd + 1)$, then, its price can be expressed by

$$\sum_{|I| \leqslant m} h^I \mathbb{E}_{\mathbb{Q}}[\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle] = \sum_{|I| \leqslant m} h^I \mathbb{E}_{\mathbb{Q}}[\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_T \rangle], \tag{2.3.16}$$

where $U_I(\boldsymbol{\ell})$ is as in Theorem 2.3.17. This in particular implies that (2.3.16) is a polynomial of degree at most m in $(\ell^W_{1,\dots,M}, \ell^\nu_{1,\dots,M}) \in \mathbb{R}^{2M}$.

Proof. The result follows from Theorem 2.3.17. □

Remark 2.3.21. Note that, thanks to Proposition 2.3.5, the expectations $\mathbb{E}_{\mathbb{Q}}[\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_T \rangle]$, for $|I| \leqslant m$, can be computed analytically. Hence, after their computation, the calibration of the model to option prices corresponding to sig-payoffs reduces to a polynomial optimization problem in $(\ell^W_{1,\dots,M}, \ell^\nu_{1,\dots,M}) \in \mathbb{R}^{2M}$. For usual calibration to call and put options the above described Monte-Carlo approach can be combined with variance reduction techniques based on (2.3.16). For more details about these topics in the continuous setting see e.g. Lyons et al. (2020) and Cuchiero et al. (2023a).

2.3.4 Hedging of sig-payoffs

In the Lévy-type signature model the market is not complete. One risk management criterion is to minimize the hedging error in a mean square sense, i.e., to use a quadratic loss function. In this context losses and gains are treated in a symmetric manner and the criterion to be minimized can be either the squared hedging error at maturity or measured locally in time (see e.g. Schweizer (1991, 1990, 1995, 2001), Pham (2000) and the references therein for an extensive analysis of these problems). In what follows, we focus on the first approach which leads to the construction of the so-called mean variance hedging strategy. It specifically consists in solving the optimization problem

$$\mathbb{E}\bigg[\Big(C - v - \int_0^T \theta_s dS_s\Big)^2\bigg], \tag{2.3.17}$$

over all initial endowments $v \in \mathbb{R}$ and all hedging strategies θ. Here, C denotes the payoff at time T of a European contingent claim and S the price process of the considered risky asset. When the price process S is a martingale (thus working under a risk neutral measure $\mathbb{Q}$), this problem has been explicitly solved in Föllmer and Sondermann (1986). In particular, the authors showed that, given any contingent claim with maturity $T > 0$ and square integrable payoff C, there exists a unique solution $\varphi^* = (v^*, \theta^*)$ to the optimization problem (2.3.17) which is computed as follows. Let V denote the martingale generated by C, that is $V_t := \mathbb{E}[C|\mathcal{F}_t]$, $t \leqslant T$. Then, the optimal initial endowment and the optimal hedging strategy, are given by

$$v^* = V_0, \qquad \theta^* = \frac{d[V,S]^{pred}}{d[S]^{pred}},$$

where $[\,\cdot\,]^{pred}$ and $[\cdot,\cdot]^{pred}$ stand for predictable quadratic variation and covariation, respectively.

We now adapt these results to our setting, working under $\mathbb{Q}$ where the price process is a martingale. For a discussion why measuring the risk under $\mathbb{Q}$ can be reasonable we refer to Kallsen and Pauwels (2010). The proof of the subsequent theorem can be found in Appendix 2.B.4.

Theorem 2.3.22. *Fix a multi-index I and let $C^S = \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle$ be a sig-payoff with maturity $T > 0$. Assume that $d_X \geqslant |I|(nd+1)$. Then, the unique solution $\varphi^* = (v^*, \theta^*)$ to the optimization problem (2.3.17) with respect to C^S is given by*

$$\begin{aligned} v^* &= \mathbb{E}[\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle] \\ \theta^*_t &= \frac{\sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}\otimes\epsilon_{J_3} = \epsilon_{U_I(\boldsymbol{\ell})}} \sum_{|H|\leqslant n} \langle \epsilon_{J_1 \sqcup\!\sqcup H}, \mathbb{X}_{t-}\rangle \gamma_2(J_2, H, \boldsymbol{\ell}) \mathbb{E}[\langle \epsilon_{J_3}, \mathbb{X}_{T-t}\rangle]}{\Big(\sum_{|H|\leqslant n} \ell^J_W \langle \epsilon_H, \mathbb{X}_{t-}\rangle\Big)^2 + \Big(\sum_{|H|\leqslant n} \ell^H_\nu \langle \epsilon_H, \mathbb{X}_{t-}\rangle\Big)^2 \int_{\mathbb{R}} x^2 F(dx)}, \end{aligned}$$

where $\gamma_2(J_2, H, \boldsymbol{\ell}) = \gamma_W(J_2)\ell^H_W + \gamma_\nu(J_2)\ell^H_\nu \int_{\mathbb{R}} x^{\Sigma(J_2)+1} F(dx)$ for

$$\gamma_W(J_2) = 1_{\{J_2 = (0)\}} \quad \textit{and} \quad \gamma_\nu(J_2) = \frac{1}{|J_2|!} 1_{\{|J_2|\geqslant 1\}} 1_{\{J_2 \in \{1,\dots,d_X\}^{|J_2|}\}}.$$

Remark 2.3.23. We report here some further observations about Theorem 2.3.22.

(i) By linearity the results of Theorem 2.3.22 extend to $C^S = \sum_{|I|\leqslant m} h^I \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle$.

(ii) Remark 2.3.6 provides explicit formulas for $\mathbb{E}[\langle \epsilon_J, \mathbb{X}_T\rangle]$ and, as a consequence of Theorem 2.3.17, also for $v^* = \mathbb{E}[\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_T\rangle]$.

(iii) The above result gives a fully explicit formula for the mean variance strategy that hinges solely on the option and the model's parameters as well as the signature of the market's primary process X.

(iv) If $i_{|I|-1} = -1$ and $i_{|I|} = 1$ by (2.3.12) we have that

$$\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t\rangle = \int_0^t \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle dS(\boldsymbol{\ell})_s,$$

which is a true martingale. This yields $V_t = \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t\rangle$ and hence

$$\theta_t^* = \frac{d[V, S(\boldsymbol{\ell})]_t^{pred}}{d[S(\boldsymbol{\ell})]_t^{pred}} = \frac{\langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{t-}\rangle d[S(\boldsymbol{\ell})]_t^{pred}}{d[S(\boldsymbol{\ell})]_t^{pred}} = \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{t-}\rangle.$$

(v) A contingent claim H^S is attainable if it almost surely admits the representation

$$H^S = \mathbb{E}[H^S] + \int_0^T \xi_s^* dS(\boldsymbol{\ell})_s.$$

In the case of a sig-payoff $H^S = \sum_{|I|\leqslant m} h^I \langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle$ this is for instance the case if for each I such that $h^I \neq 0$ it holds $i_{|I|-1} = -1$ and $i_{|I|} = 1$. In this case we indeed have that

$$H^S = h^{\varnothing} + \sum_{1\leqslant |I|\leqslant m} h^I \int_0^T \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle dS(\boldsymbol{\ell})_s,$$

and thus in particular $\mathbb{E}[H^S] = h_{\varnothing}$ and $\xi_s^* = \sum_{1\leqslant |I|\leqslant m} h^I \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-}\rangle$.

2.3.5 Measure transformations for Lévy-type signature models

In the previous sections, we worked with Lévy-type signature models under some martingale measure $\mathbb{Q}$. Precisely, we started with a market's primary process X and its signature, and then defined a Lévy-type signature model as the process obtained by integrating, in Itô-sense, linear functionals of the signature of X with respect to its martingale components.

The goal of this section is to specify the dynamics of such a Lévy-type signature model under the real world measure $\mathbb{P}$. In particular, we will investigate conditions guaranteeing that a Lévy-type signature model under $\mathbb{Q}$ remains a Lévy-type signature model also under $\mathbb{P}$, meaning that its characteristics are expressed in terms of the signature of a primary $\mathbb{P}$-Lévy process whose components include the $\mathbb{P}$-driving signals. As a result, explicit formulas for the computation of the moments of the price process also under the measure $\mathbb{P}$ can be provided. This can then be exploited for parameter estimation under $\mathbb{P}$ using the generalized methods of moments, which goes back to Hansen (1982) and which has been applied in the context of Lévy models and polynomial processes, e.g., by Kallsen and Muhle-Karbe (2011) and Zhou (2003) respectively.

First, recall that the equivalence of the measures $\mathbb{Q}$ and $\mathbb{P}$ implies the existence of a positive $\mathbb{Q}$-martingale Z, with $\mathbb{E}_{\mathbb{Q}}[Z] = 1$, such that

$$d\mathbb{P} = Zd\mathbb{Q}. \tag{2.3.18}$$

Moreover, since we are dealing with the augmented filtration generated by W and μ, the martingale Z is the solution of the following SDE

$$dZ_t = Z_{t-}\left(f(t)dW_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^{g(t,x)} - 1)(\mu - \nu^{\mathbb{Q}})(dt, dx)\right), \qquad Z_0 = 1, \tag{2.3.19}$$

where $f(t)$ and $g(t, y)$ are predictable processes such that Z is a true martingale (see e.g. Theorem 2.1 in Kunita (2004)). We here indicate the $\mathbb{Q}$ dependence of the driving signals $W^{\mathbb{Q}}$ and $\mu - \nu^{\mathbb{Q}}$. In such a case, we have that (see e.g. Theorem 2.3 in Kunita (2004))

$$W^{\mathbb{P}} = W^{\mathbb{Q}} - \int_0^{\cdot} f(s)ds$$

is a standard $\mathbb{P}$-Brownian motion and the compensator of the jump measure μ is

$$\nu^{\mathbb{P}}(dt, dx) = e^{g(t,x)}\nu^{\mathbb{Q}}(dt, dx) = e^{g(t,x)}F(dx)dt.$$

Moreover, the dynamics under $\mathbb{P}$ of the $\mathbf{Q}$-Lévy model defined in (2.3.6) are given by

$$\begin{aligned} dS(\boldsymbol{\ell})_t =& \left(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle f(t) + \sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle \int_{\mathbb{R}} y(e^{g(t,x)} - 1)F(dx)\right) dt \\ &+ \left(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle\right) dW_t^{\mathbb{P}} \\ &+ \int_{\mathbb{R}} \left(\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle\right) x(\mu - \nu^{\mathbb{P}})(dt, dx). \end{aligned} \tag{2.3.20}$$

Therefore, its characteristics are expressed in terms of linear functionals of the signature of the process X, which is in general not a $\mathbb{P}$-Lévy process and whose components do not include the $\mathbb{P}$-driving signals.

In order to recover some tractability also under $\mathbb{P}$, we are interested in Lévy-type signature models which satisfy (2.3.13) also for some $\mathbb{P}$-Lévy process Y which fulfills the hypothesis of Proposition 2.3.5.

Definition 2.3.24. A process $S(\boldsymbol{\ell})$ is called $\mathbb{P}$-$\mathbf{Q}$-Lévy-type signature model if it satisfies (2.3.6) and admits the representation

$$S(\boldsymbol{\ell})_t = S_0 + \int_0^t \Big(\sum_{|J|\leqslant n} \bar{\ell}_t^J \langle \epsilon_J, \mathbb{Y}_{s-} \rangle \Big) ds + \int_0^t \Big(\sum_{|J|\leqslant n} \bar{\ell}_W^J \langle \epsilon_J, \mathbb{Y}_{s-} \rangle \Big) d\overline{W}_s$$
$$+ \int_0^t \int_{\mathbb{R}} \Big(\sum_{|J|\leqslant n} \bar{\ell}_\nu^J \langle \epsilon_J, \mathbb{Y}_{s-} \rangle \Big) x(\overline{\mu} - \overline{\nu})(ds, dx),$$

for some $\bar{\ell}_t^J, \bar{\ell}_W^J, \bar{\ell}_\nu^J \in \mathbb{R}$ and some Lévy process Y of the form

$$Y_t := \Big(t, \overline{W}_t, \int_0^t \int_{\mathbb{R}} x(\overline{\mu} - \overline{\nu})(ds, dx), \int_0^t \int_{\mathbb{R}} x^2 \overline{\mu}(ds, dx), \dots, \int_0^t \int_{\mathbb{R}} x^{d_X} \overline{\mu}(ds, dx) \Big),$$

for some d_X large enough, $\mathbb{P}$-Brownian motion $\overline{W}$, and homogeneous $\mathbb{P}$-Poisson random measure $\overline{\mu}$ with compensator $\overline{\nu}$, such that Y has finite moments of all orders.

The proof of the following proposition can be found in Appendix 2.B.5

Proposition 2.3.25. Let $S(\boldsymbol{\ell})$ be a Lévy-type signature model under $\mathbf{Q}$ and fix $\mathbb{P}$ according to (2.3.18) and (2.3.19), for $g(t,x) = g(x)$ a deterministic function and

$$f(t) = \sum_{|I|\leqslant p} f^I \langle \epsilon_I, \mathbb{X}_{t-} \rangle, \tag{2.3.21}$$

with $f^I \in \mathbb{R}$, $p \in \mathbb{N}$, I multi-indices with indices in the set $\{-1, 0, 1, 2, \dots, d_X\} \backslash \{0, 1\}$. If

$$\int_{\mathbb{R}} (e^{g(x)} - 1)^2 F(dx) < \infty \quad \text{and} \quad \mathbb{E}_{\mathbf{Q}} \Big[\exp\Big(\frac{1}{2} \int_0^T f(s)^2 ds \Big) \Big] < \infty, \tag{2.3.22}$$

then $S(\boldsymbol{\ell})$ is a $\mathbb{P}$-$\mathbf{Q}$-Lévy-type signature model with respect to the process Y given by

$$Y_t := \Big(t, W_t^{\mathbb{P}}, \int_0^t \int_{\mathbb{R}} x(\mu - \nu^{\mathbb{P}})(ds, dx), \int_0^t \int_{\mathbb{R}} x^2 \mu(ds, dx), \dots, \int_0^t \int_{\mathbb{R}} x^{d_X} \mu(ds, dx) \Big). \tag{2.3.23}$$

Remark 2.3.26. Since tractability under $\mathbf{Q}$ is a key feature for the results of the chapter, the process $S(\boldsymbol{\ell})$ in Proposition 2.3.25 is first specified through its dynamics under $\mathbf{Q}$. Applying the above measure change procedure in the reverse direction one can however also start with a process first specified under $\mathbb{P}$.

2.A Proofs of Section 2.2

For notational convenience, in the proofs of Section 2.2 we focus only on the set of time-extended weakly geometric càdlàg p-rough paths, for $p \in [2, 3)$. However, as pointed out in Remark 2.A.1 (i) below, the exact same reasoning can also be applied when working with time-extended weakly geometric càdlàg p-rough paths, for p in $[1, 2)$.

2.A.1 Proof of Proposition 2.2.6

This result follows by the Stone-Weierstrass theorem applied to the set A given by

$$A := \operatorname{span}\{L : K \to \mathbb{R}; \widehat{\mathbf{X}} \mapsto \langle \epsilon_I, \widehat{\mathbb{X}}_1 \rangle : |I| \geqslant 0\}. \tag{2.A.1}$$

Therefore, we must prove that A satisfies the following conditions:

(i) it is a linear subspace of the space of continuous functions from K to $\mathbb{R}$,

(ii) it is a subalgebra and contains a non-zero constant function,

(iii) it separates points.

(i): Let $N \geqslant 3$ and denote by $\widehat{\mathbb{X}}^N$ the solution of the Marcus-type RDE given by (1.3.10). By Proposition 3.18(ii) in Chevyrev and Friz (2019), the solution map of equation (1.3.10)

$$\begin{aligned}(K, \sigma_\infty) &\to (D([0,1]; T^N(\mathbb{R}^{d+1})), \sigma_\infty) \\ \widehat{\mathbf{X}} &\mapsto \widehat{\mathbb{X}}^N\end{aligned} \tag{2.A.2}$$

is continuous on K for every $N \geqslant 3$, where the J_1-metric on $D([0,1]; T^N(\mathbb{R}^{d+1}))$ is computed via the distance ρ induced by the norm on $T^N(\mathbb{R}^{d+1})$ defined via $|\mathbf{a}|_{T^N(\mathbb{R}^{d+1})} := \max_{n=0,\dots,N} |\mathrm{a}^{(n)}|_{(\mathbb{R}^{d+1})^{\otimes n}}$, for any $\mathbf{a} \in T^N(\mathbb{R}^{d+1})$.

As already explained in Section 1.3.3, equation (1.3.10) is, in fact, an example of a Marcus-type RDE (see Chevyrev and Friz (2019)) with linear vector fields. Therefore, by applying the result stated in Theorem 10.53 of Friz and Victoir (2010), one can see that all the assumptions of Proposition 3.18 in Chevyrev and Friz (2019) are satisfied.

Next, observe that for a generic metric space (E, d), the evaluation map

$$\begin{aligned}(D([0,1]; E), \sigma_\infty) &\to (E, d) \\ X &\mapsto X_1\end{aligned}$$

is continuous. More precisely, let $X \in D([0,1]; E)$ and let $(X^n)_n \subset D([0,1]; E)$ be a sequence such that $X^n \to X$ as $n \to \infty$ with respect to the J_1-topology. Then, there exists $(\lambda^n)_n \subset \Lambda$ such that $|\lambda^n| \to 0$ and $\sup_{s \in [0,1]} d(X^n_{\lambda^n(s)}, X_s) \to 0$ as $n \to \infty$ (see e.g. Chapter 3 in Billingsley (1999)), and in particular $d(X^n_{\lambda^n(1)}, X_1) = d(X^n_1, X_1) \to 0$, as $n \to \infty$.

In the present setting this result yields that continuity of (2.A.2) implies continuity of

$$\begin{aligned}(K, \sigma_\infty) &\to (T^N(\mathbb{R}^{d+1}), \rho) \\ \widehat{\mathbf{X}} &\mapsto \widehat{\mathbb{X}}^N_1.\end{aligned}$$

Since the map $\widehat{\mathbb{X}}^N_1 \mapsto \langle \epsilon_I, \widehat{\mathbb{X}}^N_1 \rangle$ is continuous for each I, the claim follows.

(ii): Since $\widehat{\mathbb{X}}_1$ is a group-like element, the shuffle property holds. Therefore, A is a subalgebra. Moreover, it also contains the non-zero constant function $\widehat{\mathbf{X}} \mapsto \langle \epsilon_\varnothing, \widehat{\mathbb{X}}_1 \rangle \equiv 1$.

(iii): As a final step, we need to show that the set A separates points. Let us consider $\widehat{\mathbf{X}}, \widehat{\mathbf{Y}} \in K$, with $\widehat{\mathbf{X}} \neq \widehat{\mathbf{Y}}$. We prove that there exists $n \in \mathbb{N}$ and $i, k, j = -1, 1, \dots, d$ such that

$$\begin{aligned}&\langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1} + ((\epsilon_j \otimes \epsilon_k) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle \\ &\quad \neq \langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1} + ((\epsilon_j \otimes \epsilon_k) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle.\end{aligned}$$

We reason by contradiction. Assume that for all $n \in \mathbb{N}$ and all $i, k, j \in \{-1, 1, \dots, d\}$

$$\begin{aligned}&\langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1} + ((\epsilon_j \otimes \epsilon_k) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle \\ &\quad = \langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1} + ((\epsilon_j \otimes \epsilon_k) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle.\end{aligned} \tag{2.A.3}$$

In particular, for $k = j = -1$, by Remark 1.2.18 (iii), for all $n \in \mathbb{N}$ and $i = -1, 1, \dots, d$ we get

$$\langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle = \langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle.$$

By Remark 1.2.18 (iii), $\langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle = \int_0^1 \langle \epsilon_i, \widehat{\mathbf{X}}_{s-} \rangle \frac{s^n}{n!} ds$, and

$$\langle (\epsilon_i \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle = \int_0^1 \langle \epsilon_i, \widehat{\mathbf{Y}}_{s-} \rangle \frac{s^n}{n!} ds.$$

By Corollary 4.24 in Brezis (2011) and the càglàd property of the paths

$$s \mapsto \langle \epsilon_i, \widehat{\mathbf{X}}_{s-} \rangle, \qquad s \mapsto \langle \epsilon_i, \widehat{\mathbf{Y}}_{s-} \rangle,$$

both starting in 0 by Assumption 1.3.20, we then deduce that $\langle \epsilon_i, \widehat{\mathbf{X}} \rangle = \langle \epsilon_i, \widehat{\mathbf{Y}} \rangle$.

Similarly, choosing $i = -1$, equation (2.A.3) yields that

$$\langle ((\epsilon_k \otimes \epsilon_j) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle = \langle ((\epsilon_k \otimes \epsilon_j) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle$$

for all $n \in \mathbb{N}$ and all $j, k \in \{-1, 1, \dots, d\}$. Once again, by Remark 1.2.18 (iii),

$$\langle ((\epsilon_k \otimes \epsilon_j) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{X}}_1 \rangle = \int_0^1 \langle \epsilon_{(k,j)}, \widehat{\mathbf{X}}_{s-} \rangle \frac{s^n}{n!} ds,$$

and

$$\langle ((\epsilon_k \otimes \epsilon_j) \sqcup\!\sqcup\, \epsilon_{-1}^{\otimes n}) \otimes \epsilon_{-1}, \widehat{\mathbb{Y}}_1 \rangle = \int_0^1 \langle \epsilon_{(k,j)}, \widehat{\mathbf{Y}}_{s-} \rangle \frac{s^n}{n!} ds.$$

The same reasoning as above yields then that $\langle \epsilon_{(k,j)}, \widehat{\mathbf{X}} \rangle = \langle \epsilon_{(k,j)}, \widehat{\mathbf{Y}} \rangle$ and thus $\widehat{\mathbf{X}} = \widehat{\mathbf{Y}}$.

Remark 2.A.1. (i) We presented the proof only for $p \in [2, 3)$. However, an analogous and simpler result can be given for continuous functionals of weakly geometric càdlàg p-rough path, with $p \in [1, 2)$. Indeed, the only difference in this setting concerns the interpretation of the Marcus RDE (1.3.10) as a differential equation where the integral in (1.3.11) is of Young type.

(ii) On the subset of continuous paths the continuity of the solution map (2.A.2) can be deduced by standard theorems in rough paths theory. We refer to e.g. Theorem 10.26 and Corollary 10.28 in Friz and Victoir (2010), where continuity with respect to the variation pseudometric d_p (see Section 1.2.1) is also provided.

2.A.2 Proof of Theorem 2.2.13

We start by summarizing the reasoning used in the proof of Theorem 2.2.13. Recall that we analyze here only the case of $p \in [2,3)$ but as explained in Remark 2.A.1 (i), the proof can be easily adapted to $p \in [1,2)$.

Consider the set of stopped weakly geometric rough paths given by

$$H := \left\{ \hat{\mathbf{X}}^t : t \in [0,1], \hat{\mathbf{X}} \in K \right\}.$$

We need to show that continuous path functionals can be uniformly approximated on H by linear functionals of the signature evaluated at the deterministic time of stopping $t \in [0,1]$. To this end, we rely on Proposition 2.2.6. In particular, in order to recover all the conditions needed to apply Proposition 2.2.6, we consider the auxiliary subset of $\widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+2}))$ given by

$$\widehat{H} := \left\{ \widehat{\hat{\mathbf{X}}^t} : t \in [0,1], \hat{\mathbf{X}} \in K \right\},$$

consisting of all weakly geometric càdlàg rough paths with values in $G^2(\mathbb{R}^{d+2})$ obtained as time extension of paths in H. Observe that is a well-defined subset of $\widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+2}))$.

Next, we first prove that the set $\widehat{H}$ is relatively compact in $D([0,1]; G^2(\mathbb{R}^{d+2}))$ (Step 1). Then we show that its closure is bounded with respect to the p-variation norm and that it is a subset of the time-extended weakly geometric càdlàg rough paths $\widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+2}))$ (Step 2). After introducing an auxiliary path functional and an application of Proposition 2.2.6 (Step 3), we then just need to conclude that the obtained linear functional is of the desired form (Step 4).

Remark 2.A.2. Recall that $\widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+2})) \subset D([0,1]; G^2(\mathbb{R}^{d+2}))$ and that the metric space $(G^2(\mathbb{R}^{d+2}), d_{CC})$ is a Polish space where closed bounded sets are compact (see Corollary 7.50 in Friz and Victoir (2010)). Therefore, a set $B \subset G^2(\mathbb{R}^{d+2})$ such that $\sup_{\mathbf{x} \in B} \|\mathbf{x}\|_{CC} < \infty$ has compact closure.

Furthermore, let ρ denote the distance on $G^2(\mathbb{R}^{d+2})$ induced by the norm on $T^2(\mathbb{R}^{d+2})$ defined via $|\mathbf{a}|_{T^2(\mathbb{R}^{d+2})} := \max_{n=0,\dots,2} |\mathrm{a}^{(n)}|_{(\mathbb{R}^{d+2})^{\otimes n}}$, for any $\mathbf{a} \in T^2(\mathbb{R}^{d+2})$. By Proposition 7.49 in Friz and Victoir (2010), the map

$$id : (G^2(\mathbb{R}^{d+2}), d_{CC}) \to (G^2(\mathbb{R}^{d+2}), \rho)$$

is Lipschitz on bounded sets.

Notation 2.A.3. Throughout the proof, for all $t \in [0,1]$, $\hat{\mathbf{X}} \in K$, we set $\check{\mathbf{X}}^t := \widehat{\hat{\mathbf{X}}^t}$ and denote by $\check{\mathbb{X}}^t$ its signature.

Step 1: We show that the set $\widehat{H}$ is relatively compact in $D([0,1];G^2(\mathbb{R}^{d+2}))$. By Theorem 12.3 in Billingsley (1999), a set $A \subset D([0,1];G^2(\mathbb{R}^{d+2}))$ is relatively compact if and only if

$$\{\mathbf{Z}_s \colon s \in [0,1], \mathbf{Z} \in A\} \tag{2.A.4}$$

is relatively compact in $(G^2(\mathbb{R}^{d+2}), d_{CC})$ and

$$\lim_{\delta \to 0} \sup_{\mathbf{Z} \in A} \left(\inf_{\{t_i\}_\delta} \max_{1 \leqslant i \leqslant \nu} \sup_{s,u \in [t_{i-1}, t_i)} d_{CC}(\mathbf{Z}_s, \mathbf{Z}_u) \right) = 0 \tag{2.A.5}$$

where $\{t_i\}_\delta$ denotes a partition $0 = t_0 < t_1 < \cdots < t_\nu = 1$ of $[0,1]$, satisfying

$$\min_{1 \leqslant i \leqslant \nu} (t_i - t_{i-1}) > \delta.$$

First, notice that since K is compact in $\widehat{D}^p([0,1];G^2(\mathbb{R}^{d+1}))$ and each element of K is already time-extended, the set

$$\widehat{H}^1 := \left\{ \check{\mathbf{X}}^1 : \widehat{\mathbf{X}} \in K \right\} \subset \widehat{H},$$

is compact in $D([0,1];G^2(\mathbb{R}^{d+2}))$. By Remark 2.A.2 it thus holds that

$$\sup_{\widehat{\mathbf{X}} \in K} \sup_{s \in [0,1]} \|\check{\mathbf{X}}^1_s\|_{CC} < \infty.$$

By definition of $\widehat{H}$ and since $\|\check{\mathbf{X}}^t_{t,s}\|_{CC} = (s-t)$ for each $s \in [t,1]$, we can then compute

$$\begin{aligned}
\sup_{\widehat{\mathbf{Y}} \in \widehat{H}} \sup_{s \in [0,1]} \|\widehat{\mathbf{Y}}_s\|_{CC} &= \sup_{t \in [0,1], \widehat{\mathbf{X}} \in K} \sup_{s \in [0,1]} \|\check{\mathbf{X}}^t_s\|_{CC} \\
&\leqslant \sup_{t \in [0,1], \widehat{\mathbf{X}} \in K} \max \left(\sup_{s \in [0,t]} \|\check{\mathbf{X}}^1_s\|_{CC}, \sup_{s \in [t,1]} (\|\check{\mathbf{X}}^1_t\|_{CC} + \|\check{\mathbf{X}}^t_{t,s}\|_{CC}) \right) < \infty.
\end{aligned}$$

This shows that (2.A.4) is satisfied by $\widehat{H}$. Similarly, observe that $d_{CC}(\check{\mathbf{X}}^t_s, \check{\mathbf{X}}^t_u) = (u-s)$ for each $t \leqslant s \leqslant u$ and

$$d_{CC}(\check{\mathbf{X}}^t_s, \check{\mathbf{X}}^t_u) \leqslant d_{CC}(\check{\mathbf{X}}^1_s, \check{\mathbf{X}}^1_t) + (u-t)$$

for each $s \leqslant t \leqslant u$. We can then compute

$$\begin{aligned}
&\lim_{\delta \to 0} \sup_{\widehat{\mathbf{Y}} \in \widehat{H}} \left(\inf_{\{t_i\}_\delta} \max_{1 \leqslant i \leqslant \nu} \sup_{s,u \in [t_{i-1}, t_i)} d_{CC}(\widehat{\mathbf{Y}}_s, \widehat{\mathbf{Y}}_u) \right) \\
&\quad = \lim_{\delta \to 0} \sup_{t \in [0,1], \widehat{\mathbf{X}} \in K} \left(\inf_{\{t_i\}_\delta} \max_{1 \leqslant i \leqslant \nu} \sup_{s,u \in [t_{i-1}, t_i)} d_{CC}(\check{\mathbf{X}}^t_s, \check{\mathbf{X}}^t_u) \right) = 0,
\end{aligned}$$

proving that $\widehat{H}$ satisfies (2.A.5) too and is thus relatively compact.

Step 2: Let $cl(\widehat{H})$ be the closure of $\widehat{H}$ in $D([0,1];G^2(\mathbb{R}^{d+2}))$. We now show that $cl(\widehat{H})$ is bounded with respect to the p-variation norm and that it is a subset of the time-extended weakly geometric rough paths over $\mathbb{R}^{d+2}$, i.e.,

$$cl(\widehat{H}) \subset \widehat{D}^p([0,1];G^2(\mathbb{R}^{d+2})).$$

Let $(\widehat{\mathbf{Y}}^n)_n \subseteq \widehat{H}$ be a sequence and $\mathbf{Y} \in D([0,1];G^2(\mathbb{R}^{d+2}))$ be such that $\widehat{\mathbf{Y}}^n \to \mathbf{Y}$ in the J_1-topology. By definition, there exists a sequence of time-reparametrization $(\lambda_n)_n$ such that

$$\lim_{n\to\infty} \sup_{s\in[0,1]} |\lambda_n(s) - s| = 0, \quad \text{and} \quad \lim_{n\to\infty} \sup_{s\in[0,1]} d_{CC}(\widehat{\mathbf{Y}}^n_{\lambda_n(s)}, \mathbf{Y}_s) = 0. \tag{2.A.6}$$

Let then $\mathcal{D}$ be a fixed partition of $[0,1]$. Observe that since by assumption K is bounded with respect to $\|\cdot\|_{p-var}$, $\widehat{H}^1$ and thus $\widehat{H}$ are bounded as well. We can thus compute

$$\begin{aligned}\sum_{t_i\in\mathcal{D}} d_{CC}(\mathbf{Y}_{t_i}, \mathbf{Y}_{t_{i+1}})^p &= \lim_{n\to\infty} \sum_{t_i\in\mathcal{D}} d_{CC}\Big(\widehat{\mathbf{Y}}^n_{\lambda_n(t_i)}, \widehat{\mathbf{Y}}^n_{\lambda_n(t_{i+1})}\Big)^p \leqslant \limsup_{n\to\infty} \|\widehat{\mathbf{Y}}^n_{\lambda_n(\cdot)}\|^p_{p-var} \\ &= \limsup_{n\to\infty} \|\widehat{\mathbf{Y}}^n\|^p_{p-var} \leqslant \sup_{n\in\mathbb{N}} \|\widehat{\mathbf{Y}}^n\|^p_{p-var},\end{aligned}$$

where the second equality follows by the invariance of the p-variation norm with respect to time-reparametrization. Taking the sup over all the partition we get that $cl(\widehat{H})$ is bounded with respect to the p-variation norm.

Next, note that $\lambda_n(0) = 0$ for all $n \in \mathbb{N}$. Therefore, it holds that

$$\widehat{\mathbf{Y}}^n_{\lambda_n(0)} = \widehat{\mathbf{Y}}^n_0 \to \mathbf{Y}_0 = \mathbf{1} \in G^2(\mathbb{R}^{d+2}).$$

Moreover, by

$$\lim_{n\to\infty} \sup_{s\in[0,1]} \big(\big|\|\widehat{\mathbf{Y}}^n_{\lambda_n(s)}\|_{CC} - \|\mathbf{Y}_s\|_{CC}\big) \leqslant \lim_{n\to\infty} \sup_{s\in[0,1]} d_{CC}(\widehat{\mathbf{Y}}^n_{\lambda_n(s)}, \mathbf{Y}_s) = 0,$$

and the fact that $\mathbf{Y}$ is càdlàg we deduce that the set

$$\{\mathbf{Y}_s \colon s \in [0,1], n \in \mathbb{N}\} \cup \{\widehat{\mathbf{Y}}^n_{\lambda_n(s)} \colon s \in [0,1], n \in \mathbb{N}\}$$

is bounded. By Remark 2.A.2 we thus get

$$\lim_{n\to\infty} \sup_{s\in[0,1]} \rho(\widehat{\mathbf{Y}}^n_{\lambda_n(s)}, \mathbf{Y}_s) = 0,$$

implying in particular that for all $|I| \leqslant 2$

$$\lim_{n\to\infty} \sup_{s\in[0,1]} |\langle \epsilon_I, \widehat{\mathbf{Y}}^n_{\lambda_n(s)} - \mathbf{Y}_s\rangle| = 0.$$

Since for all $n \in \mathbb{N}$, $s \in [0,1]$, $\langle \epsilon_{-1}, \widehat{\mathbf{Y}}^n_{\lambda_n(s)} \rangle = \lambda_n(s)$, by condition (2.A.6), $\langle \epsilon_{-1}, \mathbf{Y}_s \rangle = s$ for all $s \in [0,1]$. Furthermore, since continuity points of $\widehat{\mathbf{Y}}^n$ have Lebesgue measure 1, for all $i \in \{-1, 1, \ldots, d+1\}$ we have

$$\langle \epsilon_i \otimes \epsilon_{-1}, \widehat{\mathbf{Y}}^n_{\lambda_n(s)} \rangle = \int_0^{\lambda_n(s)} \langle \epsilon_i, \widehat{\mathbf{Y}}^n_{u^-} \rangle du = \int_0^{\lambda_n(s)} \langle \epsilon_i, \widehat{\mathbf{Y}}^n_u \rangle du = \int_0^1 \langle \epsilon_i, \widehat{\mathbf{Y}}^n_u \rangle 1_{\{u \leqslant \lambda_n(s)\}} du.$$

Finally, observe that for fixed $s \in [0,1]$ and for all continuity points $u \in [0,1] \backslash \{s\}$ of $\mathbf{Y}$,

$$\lim_{n \to \infty} \langle \epsilon_i, \widehat{\mathbf{Y}}^n_u \rangle 1_{\{u \leqslant \lambda_n(s)\}} = \langle \epsilon_i, \mathbf{Y}_u \rangle 1_{\{u \leqslant s\}}.$$

Thus, the dominated convergence theorem yields

$$\lim_{n \to \infty} \langle \epsilon_i \otimes \epsilon_{-1}, \widehat{\mathbf{Y}}^n_{\lambda_n(s)} \rangle = \langle \epsilon_i \otimes \epsilon_{-1}, \mathbf{Y}_s \rangle = \int_0^s \langle \epsilon_i, \mathbf{Y}_u \rangle du,$$

and the claim follows.

Step 3: Set $\widehat{F} : \widehat{D}^p([0,1]; G^2(\mathbb{R}^{d+2})) \to \mathbb{R}$ as $\widehat{F}(\widehat{\mathbf{Z}}) := F(\mathbf{Z})$ where

$$\mathbf{Z} \in D^p([0,1]; G^2(\mathbb{R}^{d+1}))$$

denotes the weakly geometric rough path obtained by removing the time-extension of $\widehat{\mathbf{Z}}$ (see Remark 1.2.18 (iv)). Observe that by Notation 2.A.3, if in particular $\widehat{\mathbf{Z}} \in \widehat{H}$, then $\widehat{\mathbf{Z}} = \breve{\mathbf{X}}^t$ and $\mathbf{Z} = \widehat{\mathbf{X}}^t$ for some $t \in [0,1]$ and $\widehat{\mathbf{X}} \in K$. Since $\widehat{F}$ is a continuous path functional, an application of Proposition 2.2.6 to the compact set $cl(\widehat{H})$ yields the existence of a linear functional $\tilde{L} \in \mathscr{L}_{\mathbb{R}^{d+2}}$ such that

$$\sup_{t \in [0,1]} \sup_{\widehat{\mathbf{X}} \in K} |\widehat{F}(\breve{\mathbf{X}}^t) - \tilde{L}(\breve{\mathbb{X}}^t_1)| \leqslant \sup_{\widehat{\mathbf{Y}} \in cl(\widehat{H})} |\widehat{F}(\widehat{\mathbf{Y}}) - \tilde{L}(\widehat{\mathbb{Y}}_1)| \leqslant \varepsilon.$$

By definition of the path functional $\widehat{F}$ and Notation 2.A.3, we then conclude that

$$\sup_{t \in [0,1]} \sup_{\widehat{\mathbf{X}} \in K} |F(\widehat{\mathbf{X}}^t) - \tilde{L}(\breve{\mathbb{X}}^t_1)| \leqslant \varepsilon.$$

Step 4: We show that for every linear functional $\tilde{L} \in \mathscr{L}_{\mathbb{R}^{d+2}}$, there exists another linear functional $L \in \mathscr{L}_{\mathbb{R}^{d+1}}$ such that for all $t \in [0,1]$ and $\widehat{\mathbf{X}} \in K$

$$\tilde{L}(\breve{\mathbb{X}}^t_1) = L(\widehat{\mathbb{X}}_t),$$

where $\widehat{\mathbb{X}}_t$ denotes the signature at time t of the path $\widehat{\mathbf{X}}$. Observe that since each component of $\breve{\mathbb{X}}^t_{t,1}$ is a polynomial in t, for each multi-index I there is a map $L_I \in \mathscr{L}_{\mathbb{R}^{d+2}}$ such that

$$\langle \epsilon_I, \breve{\mathbb{X}}^t_{t,1} \rangle = L_I(\breve{\mathbb{X}}^1_t),$$

where $\check{\mathbb{X}}^1_t$ denotes the signature at time t of the path $\check{\mathbf{X}}^1 \in H^1$. This in particular implies that for each multi-index I

$$\langle \epsilon_I, \check{\mathbb{X}}^t_1 \rangle = \langle \epsilon_I, \check{\mathbb{X}}^1_t \otimes \check{\mathbb{X}}^t_{t,1} \rangle = \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \langle \epsilon_{I_1}, \check{\mathbb{X}}^1_t \rangle \langle \epsilon_{I_2}, \check{\mathbb{X}}^t_{t,1} \rangle$$
$$= \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \langle \epsilon_{I_1}, \check{\mathbb{X}}^1_t \rangle L_{I_2}(\check{\mathbb{X}}^1_t) =: h_I(\check{\mathbb{X}}^1_t)$$

for some $h_I \in \mathscr{L}_{\mathbb{R}^{d+2}}$. Since each $\widehat{\mathbf{X}}$ in K is already time-extended, the claim follows. $\square$

Remark 2.A.4. Observe that Step 4 would not hold true if in the statement of Theorem 2.2.13 we considered K to be a compact subset of the set of weakly geometric rough paths which are not necessarily time-extended, i.e., $K \subset D^p([0,1]; G^2(\mathbb{R}^{d+1}))$.

2.B Proofs of Section 2.3

2.B.1 Proof of Proposition 2.3.5

Fix $d \in \mathbb{N}$ and let L be a $\mathbb{R}^d$-valued Lévy process with triplet (b^L, C^L, F^L) with respect to the "truncation function" $\chi(x) = x$. For notational convenience we write (b, C, F) for (b^L, C^L, F^L). We prove that (i) the truncated signature $\mathbb{L}^N$ of L is a polynomial process in the truncated tensor algebra $T^N(\mathbb{R}^d)$ and (ii) the expected signature satisfies (2.3.5).

(i) First, recall that $\mathbb{L}$ is the solution of a Marcus SDE (1.4.2). Therefore, for each multi-index I with entries in $\{1, \dots, d\}$ we have

$$\langle \epsilon_I, \mathbb{L}_t \rangle = \int_0^t \langle \epsilon_{I'}, \mathbb{L}_{s-} \rangle dL^{i_{|I|}}_s + C_{i_{|I|-1}, i_{|I|}} \frac{1}{2} \int_0^t \langle \epsilon_{I''}, \mathbb{L}_{s-} \rangle ds$$
$$+ \sum_{0<s\leq t} \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} \langle \epsilon_{I_1}, \mathbb{L}_{s-} \rangle \langle \epsilon_{I_2}, (\Delta L_s)^{\otimes |I_2|} \rangle.$$

Thus, $\mathbb{L}^N$ is a $T^N(\mathbb{R}^d)$-valued semimartingale. Moreover, by condition (2.3.4), *a.s.*

$$\sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} |\langle \epsilon_{I_1}, \mathbb{L}_{s-} \rangle| \int_{\mathbb{R}^d} |\langle \epsilon_{I_2}, x^{\otimes |I_2|} \rangle| F(dx) < \infty.$$

Therefore, the process $\int_0^\cdot \int_{\mathbb{R}^d} \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} \langle \epsilon_{I_1}, \mathbb{L}_{s-} \rangle \langle \epsilon_{I_2}, x^{\otimes |I_2|} \rangle F(dx) ds$ is of locally integrable variation and by Theorem II.1.8 in Jacod and Shiryaev (1987),

$$\int_0^\cdot \int_{\mathbb{R}^d} \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} \langle \epsilon_{I_1}, \mathbb{L}_{s-} \rangle \langle \epsilon_{I_2}, x^{\otimes |I_2|} \rangle (\mu(dx, ds) - F(dx) ds)$$

is a local martingale. We now prove that $\mathbb{L}^N$ is a jump-diffusion in the sense of Definition III.2.18 in Jacod and Shiryaev (1987). In order to simplify the notation, for each $\mathbf{y} \in T^N(\mathbb{R}^d)$ and each multi-index I, we set $\mathbf{y}_I := \langle \epsilon_I, \mathbf{y} \rangle$ and

$$\delta(\mathbf{y}, x)_I := \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} \mathbf{y}_{I_1} \langle \epsilon_{I_2}, x^{\otimes |I_2|} \rangle 1_{\{|I_2|>0\}}, \qquad \gamma(\mathbf{y}, x)_I := \delta(\mathbf{y}, x)_I - \mathbf{y}_{I'} x_{i_{|I|}}.$$

An application of Itô's formula yields that for any bounded C^2 function on $T^N(\mathbb{R}^d)$

$$f(\mathbb{L}_t^N) - f(\mathbb{L}_0^N) - \int_0^t \mathcal{A} f(\mathbb{L}_s^N) ds$$

is a local martingale, where $\mathcal{A} : C_b^2(T^N(\mathbb{R}^d)) \to C(T^N(\mathbb{R}^d))$ is such that for any $f \in C_b^2(T^N(\mathbb{R}^d))$ and $\mathbf{y} \in T^N(\mathbb{R}^d)$

$$\begin{aligned}
\mathcal{A} f(\mathbf{y}) = &\sum_{1 \leqslant |I| \leqslant N} \frac{\partial f(\mathbf{y})}{\partial \mathbf{y}_I} \mathbf{y}_{I'} b_{i_{|I|}} + \sum_{2 \leqslant |I| \leqslant N} \frac{\partial f(\mathbf{y})}{\partial \mathbf{y}_I} \left(\frac{1}{2} \mathbf{y}_{I''} C_{i_{|I|-1}, i_{|I|}} + \int_{\mathbb{R}^d} \gamma(\mathbf{y}, x)_I F(dx) \right) \\
&+ \frac{1}{2} \sum_{1 \leqslant |I|, |J| \leqslant N} \frac{\partial^2 f(\mathbf{y})}{\partial \mathbf{y}_I \partial \mathbf{y}_J} \mathbf{y}_{I'} \mathbf{y}_{J'} C_{i_{|I|}, j_{|J|}} \\
&+ \int_{\mathbb{R}^d} \left(f(\mathbf{y} + \delta(\mathbf{y}, x)) - f(\mathbf{y}) - \sum_{1 \leqslant |I| \leqslant N} \frac{\partial f(\mathbf{y})}{\partial \mathbf{y}_I} \delta(\mathbf{y}, x)_I \right) F(dx).
\end{aligned} \tag{2.B.1}$$

The operator $\mathcal{A}$ is a jump-diffusion operator whose coefficients

$$(\beta, \zeta, K) := ((\beta_I)_{|I| \leqslant N}, (\zeta_{I,J})_{|I|,|J| \leqslant N}, K)$$

are given by

$$\begin{aligned}
\beta_I(\mathbf{y}) &:= \mathbf{y}_{I'} b_{i_{|I|}} 1_{|I| \geqslant 1} + \left(\frac{1}{2} \mathbf{y}_{I''} C_{i_{|I|-1}, i_{|I|}} + \int_{\mathbb{R}^d} \gamma(\mathbf{y}, x)_I F(dx) \right) 1_{|I| \geqslant 2}, \\
\zeta_{I,J}(\mathbf{y}) &:= \mathbf{y}_{I'} \mathbf{y}_{J'} C_{i_{|I|}, j_{|J|}} 1_{|I|, |J| \geqslant 1}, \\
K(\mathbf{y}, A) &:= \int_{\mathbb{R}^d} 1_{A \setminus \{0\}}(\delta(\mathbf{y}, x)) F(dx).
\end{aligned}$$

Therefore, by Theorem II.2.42 in Jacod and Shiryaev (1987), $\mathbb{L}^N$ is a jump-diffusion on $T^N(\mathbb{R}^d)$ whose characteristics, with respect to the "truncation function" $\chi(\mathbf{y}) = \mathbf{y}$, are given by

$$\int_0^t \beta(\mathbb{L}_s^N) ds, \quad \int_0^t \zeta(\mathbb{L}_s^N) ds \quad \text{and} \quad K(\mathbb{L}_{t-}^N, dx) \times dt.$$

Notice also that $K(\mathbf{y}, \{0\}) = F(\{0\}) = 0$ for all $\mathbf{y} \in T^N(\mathbb{R}^d)$, and, by condition (2.3.4),

$$\int_{T^N(\mathbb{R}^d)} \|\boldsymbol{\xi}\| \wedge \|\boldsymbol{\xi}\|^2 K(\mathbf{y}, d\boldsymbol{\xi}) = \int_{\mathbb{R}^d} \|\delta(\mathbf{y}, x)\| \wedge \|\delta(\mathbf{y}, x)\|^2 F(dx) < \infty.$$

Finally, we show that $\mathbb{L}^N$ is a polynomial jump diffusion on $T^N(\mathbb{R}^d)$. Once again by condition (2.3.4),

$$\int_{T^N(\mathbb{R}^d)} \|\boldsymbol{\xi}\|^k K(\mathbf{y}, d\boldsymbol{\xi}) = \int_{\mathbb{R}^d} \|\delta(\mathbf{y}, x)\|^k F(dx) < \infty,$$

for all $\mathbf{y} \in T^N(\mathbb{R}^d)$ and $k \geqslant 2$. Moreover, for every set of multi-indices $I_1, \ldots, I_k$ it holds

$$\int_{T^N(\mathbb{R}^d)} \boldsymbol{\xi}_{I_1} \cdots \boldsymbol{\xi}_{I_k} K(\mathbf{y}, d\boldsymbol{\xi}) = \int_{\mathbb{R}^d} \delta(\mathbf{y}, x)_{I_1} \cdots \delta(\mathbf{y}, x)_{I_k} F(dx)$$

is a polynomial of degree at most k on $T^N(\mathbb{R}^d)$. Since each component of β and ζ is polynomial on $T^N(\mathbb{R}^d)$ respectively of degree 1 and 2, by Lemma 2.2 in Filipović and Larsson (2020) we can conclude that $\mathbb{L}^N$ is a polynomial jump-diffusion on $T^N(\mathbb{R}^d)$.

(ii) We compute now $\mathbb{E}[\langle \epsilon_I, \mathbb{L}_t^N \rangle]$ for a fixed multi-index I with $|I| \leqslant N$. To this end we apply the so-called moment formula (see for instance Theorem 2.5 in Filipović and Larsson (2020)). Observe that the projection function

$$\begin{aligned} \langle \epsilon_I, \cdot \rangle : T^N(\mathbb{R}^d) &\longrightarrow \mathbb{R} \\ \mathbf{y} &\longmapsto \langle \epsilon_I, \mathbf{y} \rangle \end{aligned}$$

is a polynomial of degree 1 on $T^N(\mathbb{R}^d)$. Since we already proved in (i) that $\mathbb{L}^N$ is a polynomial process, we also know that the corresponding generator $\mathcal{A}$ is mapping the space of polynomials of degree at most one on $T^N(\mathbb{R}^d)$ to itself. More precisely, it follows from (2.B.1) that for each multi-index J we have

$$\mathcal{A}\langle \epsilon_J, \cdot \rangle = \sum_{\epsilon_{J_1} \otimes \epsilon_{J_2} = \epsilon_J} \langle \epsilon_{J_1}, Q \rangle \langle \epsilon_{J_2}, \cdot \rangle,$$

where $Q := (0, b, \frac{1}{2}(C + \int_{\mathbb{R}^d} x^{\otimes 2} F(dx)), \ldots, \frac{1}{k!} \int_{\mathbb{R}^d} x^{\otimes k} F(dx), \ldots)$. This can equivalently be written as

$$\mathcal{A}\langle \epsilon_J, \cdot \rangle(\mathbf{y}) = \langle \epsilon_J, Q \otimes \mathbf{y} \rangle$$

By Theorem 2.5 in Filipović and Larsson (2020) we can then conclude that

$$\mathbb{E}[\langle \epsilon_I, \mathbb{L}_t^N \rangle] = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathcal{A}^k \langle \epsilon_I, \cdot \rangle(\epsilon_\varnothing) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \langle \epsilon_I, Q^{\otimes k} \otimes \epsilon_\varnothing \rangle = \langle \epsilon_I, \exp(tQ) \otimes \epsilon_\varnothing \rangle.$$

2.B.2 Proof of Proposition 2.3.12

The claim is clear for $I = \varnothing$. If $I \neq \varnothing$, recall from equation (2.3.2) that

$$\begin{aligned} \langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t \rangle = &\int_0^t \langle \epsilon_I, \mathbb{X}_{s-} \rangle dX_s^j + 1_{\{i_{|I|} = j = 0\}} \frac{1}{2} \int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s-} \rangle ds \\ &+ \sum_{0 < s \leqslant t} \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{(|I_2| + 1)!} 1_{\{|I_2| > 0\}} \langle \epsilon_{I_1}, \mathbb{X}_{s-} \rangle \Delta X_s^{\Sigma(I_2) + j} 1_{\{j > 0\}} 1_{\{I_2 \in \{1, \ldots, d\}^{|I_2|}\}}. \end{aligned}$$

Let $\mathcal{I}(r,k)$ be the set of all multi-indices $J \in \mathbb{N}^k$ such that $\Sigma(J) = r$. Set

$$\alpha(r,k) := \sum_{J\in\mathcal{I}(r,k)} \prod_{i=1}^{k} \frac{1}{(j_i+1)!}$$

and note that

$$\begin{aligned}\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j &= \langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t\rangle - 1_{\{i_{|I|}=j=0\}} \frac{1}{2} \int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s-}\rangle ds \\ &\quad - \sum_{0<s\leqslant t} \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|,1)\langle \epsilon_{I_1}, \mathbb{X}_{s-}\rangle \Delta X_s^{\Sigma(I_2)+j} 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}} \\ &= \langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t\rangle - 1_{\{i_{|I|}=j=0\}} \frac{1}{2} \int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s-}\rangle ds \\ &\quad - \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|,1) \int_0^t \langle \epsilon_{I_1}, \mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_2)+j} 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}.\end{aligned}$$

In particular, observe that $\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j$ reads

$$\langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t\rangle - \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|,1) \int_0^t \langle \epsilon_{I_1}, \mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_2)+j} 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}},$$

for each $j > 0$. We now prove by induction that for each $r \in \mathbb{N}_0$ it holds

$$\begin{aligned}&\int_0^t \langle \epsilon_I, \mathbb{X}_{s-}\rangle dX_s^j - \langle \epsilon_I \otimes \epsilon_j, \mathbb{X}_t\rangle + 1_{\{i_{|I|}=j=0\}} \frac{1}{2} \int_0^t \langle \epsilon_{I'}, \mathbb{X}_{s-}\rangle ds \\ &= \sum_{k=1}^{r} (-1)^k \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|,k)\langle \epsilon_{I_1} \otimes \epsilon_{\Sigma(I_2)+j}, \mathbb{X}_t\rangle 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}} \\ &\qquad + (-1)^{r+1} \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \alpha(|I_2|,r+1) \int_0^t \langle \epsilon_{I_1}, \mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_2)+j} 1_{\{j>0\}} 1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}.\end{aligned}$$

We already showed that the claim holds for $r = 0$. Supposing that it is true for $r - 1$, we get

$$\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I}\alpha(|I_2|,r)\int_0^t\langle\epsilon_{I_1},\mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_2)+j}1_{\{j>0\}}1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}$$
$$\begin{aligned}
&=\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I}\alpha(|I_2|,r)\langle\epsilon_{I_1}\otimes\epsilon_{\Sigma(I_2)+j},\mathbb{X}_t\rangle 1_{\{j>0\}}1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}\\
&\quad-\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I}\alpha(|I_2|,r)\sum_{\epsilon_{I_{11}}\otimes\epsilon_{I_{12}}=\epsilon_{I_1}}\alpha(|I_{12}|,1)\\
&\quad\times\int_0^t\langle\epsilon_{I_{11}},\mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_{12})+\Sigma(I_2)+j}1_{\{I_{12}\in\{1,\dots,d\}^{|I_{12}|}\}}1_{\{j>0\}}1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}\\
&\overset{(\star)}{=}\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I}\alpha(|I_2|,r)\langle\epsilon_{I_1}\otimes\epsilon_{\Sigma(I_2)+j},\mathbb{X}_t\rangle 1_{\{j>0\}}1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}\\
&\quad-\sum_{\epsilon_{I_{11}}\otimes\epsilon_{I_{122}}=\epsilon_I}\alpha(|I_{122}|,r+1)\int_0^t\langle\epsilon_{I_{11}},\mathbb{X}_{s-}\rangle dX_s^{\Sigma(I_{122})+j}1_{\{j>0\}}1_{\{I_{122}\in\{1,\dots,d\}^{|I_{122}|}\}},
\end{aligned}$$

where in $(\star)$ we set $\epsilon_{I_{122}}=\epsilon_{I_{12}}\otimes\epsilon_{I_2}$. Since $\alpha(r,k)=0$ for $k>r$, we can conclude that

$$\begin{aligned}
&\int_0^t\langle\epsilon_I,\mathbb{X}_{s-}\rangle dX_s^j-\langle\epsilon_I\otimes\epsilon_j,\mathbb{X}_t\rangle+1_{\{i_{|I|}=j=0\}}\frac{1}{2}\int_0^t\langle\epsilon_{I'},\mathbb{X}_{s-}\rangle ds\\
&\quad=\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I}\sum_{k=1}^{|I_2|}(-1)^k\alpha(|I_2|,k)\langle\epsilon_{I_1}\otimes\epsilon_{\Sigma(I_2)+j},\mathbb{X}_t\rangle 1_{\{j>0\}}1_{\{I_2\in\{1,\dots,d\}^{|I_2|}\}}.
\end{aligned}$$

2.B.3 Proof of Theorem 2.3.17

Let $k\in\mathbb{N}$ and note that $(\sum_{h=1}^M\ell_h\langle\epsilon_{J^h},\mathbb{X}_{s-}\rangle)^k$ is a polynomial of degree k in $(\ell_h)_{1\leqslant h\leqslant M}$. Precisely by the multinomial theorem, for each $\alpha=(\alpha^1,\alpha^2,\dots,\alpha^M)\in\mathbb{N}^M$ it holds

$$\Big(\sum_{h=1}^M\ell_h\langle\epsilon_{J^h},\mathbb{X}_{s-}\rangle\Big)^k=k!\sum_{\alpha\in\mathbb{N}^M,\Sigma(\alpha)=k}(\ell_{1,\dots,M})^\alpha\langle\epsilon^{\sqcup\!\sqcup\alpha},\mathbb{X}_{s-}\rangle. \tag{2.B.2}$$

Setting $U_I(\boldsymbol{\ell})$ as in the statement of the proposition, we prove $\langle\epsilon_I,\widehat{\mathbb{S}}(\boldsymbol{\ell})_t\rangle=\langle\epsilon_{U_I(\boldsymbol{\ell})},\mathbb{X}_t\rangle$ proceeding by induction on $m=|I|$. Since $\langle\epsilon_\varnothing,\widehat{\mathbb{S}}(\boldsymbol{\ell})_t\rangle=1=\langle\epsilon_\varnothing,\mathbb{X}_t\rangle$ the claim follow for $m=0$. Next, suppose that the claim holds for $|I|=m-1$ and fix $I=(i_1,\dots,i_m)$.

By equation (2.3.12), $\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle$ can be written as

$$
\begin{aligned}
&1_{\{i_{|I|}=-1\}} \int_0^t \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-} \rangle ds + 1_{\{i_{|I|}=1\}} \int_0^t \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-} \rangle \Big(\sum_{h=1}^M \ell_h^W \langle \epsilon_{J^h}, \mathbb{X}_{s-} \rangle \Big) dW_s \\
&\quad + 1_{\{i_{|I|}=1\}} \int_0^t \int_{\mathbb{R}} \langle \epsilon_{I'}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-} \rangle \Big(\sum_{h=1}^M \ell_h^\mu \langle \epsilon_{J^h}, \mathbb{X}_{s-} \rangle \Big) x(\mu - \nu)(ds, dx) \\
&\quad + 1_{\{(i_{|I|-1}, i_{|I|})=(1,1)\}} \frac{1}{2} \int_0^t \langle \epsilon_{I''}, \widehat{S}(\boldsymbol{\ell})_{s-} \rangle \Big(\sum_{h=1}^M \ell_h^W \langle \epsilon_{J^h}, \mathbb{X}_{s-} \rangle \Big)^2 ds \\
&\quad + \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} 1_{\{I_2=(1,\ldots,1)\}} \\
&\qquad \times \int_0^t \int_{\mathbb{R}} \langle \epsilon_{I_1}, \widehat{\mathbb{S}}(\boldsymbol{\ell})_{s-} \rangle \Big(\sum_{h=1}^M \ell_h^\mu \langle \epsilon_{J^h}, \mathbb{X}_{s-} \rangle \Big)^{|I_2|} x^{|I_2|} \mu(ds, dx).
\end{aligned}
$$

Therefore, as a direct application of Proposition 2.3.12 and the induction hypothesis, with the notation of (2.B.2) we can write $\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle$ as

$$
\begin{aligned}
&\langle (\epsilon_{U_{I'}(\boldsymbol{\ell})}; \epsilon_{-1})^\sim, \mathbb{X}_t \rangle 1_{\{i_{|I|}=-1\}} \\
&\quad + \sum_{\alpha \in \mathbb{N}^M, \Sigma(\alpha)=1} (\ell_{1,\ldots,M}^W)^\alpha \langle (\epsilon_{U_{I'}(\boldsymbol{\ell})} ⧢ \epsilon^{⧢\alpha}; \epsilon_0)^\sim, \mathbb{X}_t \rangle 1_{\{i_{|I|}=1\}} \\
&\quad + \sum_{\alpha \in \mathbb{N}^M, \Sigma(\alpha)=1} (\ell_{1,\ldots,M}^\nu)^\alpha \langle (\epsilon_{U_{I'}(\boldsymbol{\ell})} ⧢ \epsilon^{⧢\alpha}; \epsilon_1)^\sim, \mathbb{X}_t \rangle 1_{\{i_{|I|}=1\}} \\
&\quad + \sum_{\alpha \in \mathbb{N}^M, \Sigma(\alpha)=2} (\ell_{1,\ldots,M}^W)^\alpha \langle (\epsilon_{U_{I''}(\boldsymbol{\ell})} ⧢ \epsilon^{⧢\alpha}; \epsilon_{-1})^\sim, \mathbb{X}_t \rangle 1_{\{|I|>1\}} 1_{\{(i_{|I|-1}, i_{|I|})=(1,1)\}} \\
&\quad + \sum_{\epsilon_{I_1} \otimes \epsilon_{I_2} = \epsilon_I} \sum_{\alpha \in \mathbb{N}^M, \Sigma(\alpha)=|I_2|} (\ell_{1,\ldots,M}^\nu)^\alpha \langle (\epsilon_{U_{I_1}(\boldsymbol{\ell})} ⧢ \epsilon^{⧢\alpha}; \epsilon_{|I_2|})^\sim, \mathbb{X}_t \rangle 1_{\{|I_2|>1\}} 1_{\{I_2=(1,\ldots,1)\}} \\
&= \langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle.
\end{aligned}
$$

Note that $\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle$ is a polynomial of degree $|I|$ in $(\ell_{1,\ldots,M}^W, \ell_{1,\ldots,M}^\nu) \in \mathbb{R}^{2M}$ by construction.

Remark 2.B.1. (i) Assuming that $d_X \geqslant |I|(nd+1)$ is needed to guarantee that

$$
\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle = \langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_t \rangle \tag{2.B.3}
$$

is well-defined. This condition is verified if and only if the entries of every multi-index appearing in $U_I(\boldsymbol{\ell})$ is bounded by d_X. More precisely, the claim is clear for $|I| = 0$. For $|I| = 1$ it follows by noticing that by Remark 2.3.13 the expression $\langle (\epsilon_J; \epsilon_1)^\sim, \mathbb{X}_t \rangle$ is well-defined for each $|J| \leqslant n$ if and only if $nd + 1 \leqslant d_X$. Well definiteness of $\langle (\epsilon^{⧢\alpha}; \epsilon_{|I|})^\sim, \mathbb{X} \rangle$ for each $|I| \geqslant 2$ and α with $\Sigma(\alpha) = |I|$ is a necessary condition for well definiteness of (2.B.3). This includes in particular the case

$$
\epsilon^{⧢\alpha} = \frac{1}{|I|!} ⧢^{|I|} (\epsilon_d^{\otimes n}) = C_{n,|I|} \epsilon_d^{\otimes(n|I|)}
$$

for some $C_{n,|I|} \in \mathbb{R}$.

Proceeding as in Remark 2.3.13 we can conclude that this is the case if and only if $nd|I| + |I| \leqslant d_X$. Notice that the condition $d_X \geqslant |I|(nd+1)$ could be optimized if some coefficients ℓ_ν^J introduced in the definition of the Lévy-type signature model $S(\boldsymbol{\ell})$ are equal to 0.

(ii) Observe that a key factor in deriving the representation in equation (2.B.3) is that the jumps of the Lévy-type signature model $S(\boldsymbol{\ell})$ are proportional to the jumps of the (signature of the) market's primary process X.

2.B.4 Proof of Theorem 2.3.22

Since

$$S(\boldsymbol{\ell})_t = \int_0^t \Big(\sum_{|H|\leqslant n} \ell_W^H \langle \epsilon_H, \mathbb{X}_{s-} \rangle \Big) dW_s + \int_0^t \int_{\mathbb{R}} \Big(\sum_{|H|\leqslant n} \ell_\nu^H \langle \epsilon_H, \mathbb{X}_{s-} \rangle x \Big) (\mu - \nu)(ds, dx) \tag{2.B.4}$$

we already know that

$$d[S(\boldsymbol{\ell})]_t^{pred} = \int_0^t \Big(\sum_{|H|\leqslant n} \ell_W^J \langle \epsilon_H, \mathbb{X}_{s-} \rangle \Big)^2 + \Big(\sum_{|H|\leqslant n} \ell_\nu^H \langle \epsilon_H, \mathbb{X}_{s-} \rangle \Big)^2 \int_{\mathbb{R}} x^2 F(dx) ds.$$

Moreover, using that X is a Lévy process by Proposition 2.3.2 and Remark 2.3.7 we also get that

$$\mathbb{E}[\langle \epsilon_I, \widehat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle | \mathcal{F}_t] = \mathbb{E}[\langle \epsilon_{U_I(\boldsymbol{\ell})}, \mathbb{X}_T \rangle | \mathcal{F}_t] = \sum_{\epsilon_J \otimes \epsilon_H = \epsilon_{U_I(\boldsymbol{\ell})}} \langle \epsilon_J, \mathbb{X}_t \rangle \mathbb{E}[\langle \epsilon_H, \mathbb{X}_{T-t} \rangle].$$

Observe that in the first step we used the result of Theorem 2.3.17. Since for each $|H| \geqslant 1$ by Remark 2.3.6 it holds

$$\mathbb{E}[\langle \epsilon_H, \mathbb{X}_{T-t} \rangle] = \sum_{k=1}^{|H|} \frac{(T-t)^k}{k!} \sum_{\epsilon_{H_1} \otimes \cdots \otimes \epsilon_{H_k} = \epsilon_H} \langle \epsilon_{H_1}, Q \rangle \cdots \langle \epsilon_{H_k}, Q \rangle,$$

which is a continuous process of finite variation, by the product rule we know that $t \mapsto \langle \epsilon_J, \mathbb{X}_t \rangle \mathbb{E}[\langle \epsilon_H, \mathbb{X}_{T-t} \rangle]$ can be decomposed as the sum of a continuous process of finite variation and $\int_0^\cdot \mathbb{E}[\langle \epsilon_H, \mathbb{X}_{T-s} \rangle] d\langle \epsilon_J, \mathbb{X}_s \rangle$. Observe also that for each J_1 and J_2 such that $\Sigma(J_2) \geqslant 2$, it holds

$$\begin{aligned}\sum_{0<s\leqslant t} \langle \epsilon_{J_1}, \mathbb{X}_{s-} \rangle \Delta X_s^{\Sigma(J_2)} &= \int_0^t \int_{\mathbb{R}} \langle \epsilon_{J_1}, \mathbb{X}_{s-} \rangle x^{\Sigma(J_2)} (\mu - \nu)(ds, dx) \\ &\quad + \int_0^t \langle \epsilon_{J_1}, \mathbb{X}_{s-} \rangle ds \int_{\mathbb{R}} x^{\Sigma(J_2)} F(dx).\end{aligned}$$

Representation (2.3.2) thus yields that $\langle\epsilon_J,\mathbb{X}\rangle$ can be decomposed into a continuous process of finite variation and a square integrable martingale M^J given by

$$\begin{aligned}
M_t^J &= \int_0^t \langle\epsilon_{J'},\mathbb{X}_{s-}\rangle dW_s 1_{\{j_{|J|}=0\}} + \int_0^t\int_{\mathbb{R}} \langle\epsilon_{J'},\mathbb{X}_{s-}\rangle x^{j_{|J|}}(\mu-\nu)(ds,dx)1_{\{j_{|J|}>0\}} \\
&\quad + \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}=\epsilon_J} \frac{1}{|J_2|!}1_{\{|J_2|>1\}}1_{\{J_2\in\{1,\dots,d_X\}^{|J_2|}\}}\int_0^t\int_{\mathbb{R}}\langle\epsilon_{J_1},\mathbb{X}_{s-}\rangle x^{\Sigma(J_2)}(\mu-\nu)(ds,dx) \\
&= \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}=\epsilon_J}\gamma_W(J_2)\int_0^t\langle\epsilon_{J_1},\mathbb{X}_{s-}\rangle dW_s + \gamma_\nu(J_2)\int_0^t\int_{\mathbb{R}}\langle\epsilon_{J_1},\mathbb{X}_{s-}\rangle x^{\Sigma(J_2)}(\mu-\nu)(ds,dx),
\end{aligned}$$

for $\gamma_W(J_2)=1_{\{J_2=(0)\}}$ and $\gamma_\nu(J_2)=\frac{1}{|J_2|!}1_{\{|J_2|\geqslant 1\}}1_{\{J_2\in\{1,\dots,d_X\}^{|J_2|}\}}$. Define

$$V_t := \mathbb{E}[\langle\epsilon_I,\widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle|\mathcal{F}_t]$$

for each $t\in[0,T]$. Writing J_3 in place of H and using that V is a true martingale by definition, we can thus conclude that

$$\begin{aligned}
V_t &= V_0 + \sum_{\epsilon_J\otimes\epsilon_{J_3}=\epsilon_{U_I(\boldsymbol{\ell})}}\int_0^t \mathbb{E}[\langle\epsilon_{J_3},\mathbb{X}_{T-s}\rangle]dM_s^J \\
&= V_0 + \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}\otimes\epsilon_{J_3}=\epsilon_{U_I(\boldsymbol{\ell})}} \\
&\qquad \int_0^t\langle\epsilon_{J_1},\mathbb{X}_{s-}\rangle\mathbb{E}[\langle\epsilon_{J_3},\mathbb{X}_{T-s}\rangle](\gamma_W(J_2)dW_s+\gamma_\nu(J_2)x^{\Sigma(J_2)}(\mu-\nu)(ds,dx)),
\end{aligned}$$

for $V_0=\mathbb{E}[\langle\epsilon_I,\widehat{\mathbb{S}}(\boldsymbol{\ell})_T\rangle]$. Using (2.B.4) we can then compute

$$d[V,S(\boldsymbol{\ell})]_t^{pred} = \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}\otimes\epsilon_{J_3}=\epsilon_{U_I(\boldsymbol{\ell})}}\sum_{|H|\leqslant n}\langle\epsilon_{J_1},\mathbb{X}_{t-}\rangle\gamma_2(J_2,H,\boldsymbol{\ell})\mathbb{E}[\langle\epsilon_{J_3},\mathbb{X}_{T-t}\rangle]\langle\epsilon_H,\mathbb{X}_{t-}\rangle dt,$$

where $\gamma_2(J_2,H,\boldsymbol{\ell})=\gamma_W(J_2)\ell_W^H+\gamma_\nu(J_2)\ell_\nu^H\int_{\mathbb{R}}x^{\Sigma(J_2)+1}F(dx)$. The claim follows.

2.B.5 Proof of Proposition 2.3.25

First, observe that by Theorem 9 in Protter and Shimbo (2006), the pair $(f(t),g(x))$ is such that the solution of the SDE (2.3.19) is a true martingale.

Since g is deterministic and independent of time, μ is a homogeneous $\mathbb{P}$-Poisson random measure (see e.g. Theorem II.4.8 in Jacod and Shiryaev (1987)). Moreover, $e^{g(x)}F(dx)$ is a Lévy measure such that $\int_{|x|>1}|x|^k e^{g(x)}F(dx)<\infty$, for all $k\in\mathbb{N}$. Thus, proceeding as in the proof of Proposition 2.3.2, we conclude that the process Y defined in (2.3.23) is a Lévy process with finite moments of all orders.

By equation (2.3.20), we know that the $\mathbb{P}$-characteristics of $S(\boldsymbol{\ell})$ with respect to the "truncation function" $\chi(x) = x$ are given by

$$\begin{aligned}
B_t^S &= \int_0^t \Bigg(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-}\rangle f(s) + \sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{s-}\rangle \int_{\mathbb{R}} x(e^{g(x)} - 1) F(dx) \Bigg) ds,\\
C_t^S &= \int_0^t \Bigg(\sum_{|J|\leqslant n} \ell_W^J \langle \epsilon_J, \mathbb{X}_{s-}\rangle \Bigg)^2 ds,\\
\nu^S(dt \times dx) &= K_t^{\mathbb{P}}(dx) \times dt,
\end{aligned}$$

where $K_t^{\mathbb{P}}(A) = \int_{\mathbb{R}} 1_{A\setminus\{0\}}\big(\sum_{|J|\leqslant n} \ell_\nu^J \langle \epsilon_J, \mathbb{X}_{t-}\rangle x\big) e^{g(x)} F(dx)$. Since by condition (2.3.22) we know that $\int_{\mathbb{R}} y(e^{g(x)} - 1)F(dx) < \infty$, and

$$f(t) = \sum_{|I|\leqslant p} f^I \langle \epsilon_I, \mathbb{X}_{t-}\rangle = \sum_{|I|\leqslant p} f^I \langle \epsilon_I, \mathbb{Y}_{t-}\rangle$$

it suffices to show that for each $|J| \leqslant n$ it holds $\langle \epsilon_J, \mathbb{X}_t\rangle = \langle \epsilon_J^{\mathbb{P}}, \mathbb{Y}_t\rangle$ for some linear combination of multi-indices $\epsilon_J^{\mathbb{P}}$. We proceed by induction. For $I = \varnothing$ the result is clear. Fix $k \leqslant n$, J such that $|J| = k$, and suppose that the claim holds for each $|J| < k$. Then equation (2.3.2) yields

$$\begin{aligned}
\langle \epsilon_J, \mathbb{X}_t\rangle = &\int_0^t \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle dX_s^{j_k} + 1_{\{j_{k-1}=j_k=0\}} \frac{1}{2} \int_0^t \langle \epsilon_{J''}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle ds\\
&+ \sum_{0<s\leqslant t} \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}=\epsilon_J} \frac{1}{|J_2|!} 1_{\{|J_2|>1\}} \langle \epsilon_{J_1}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle \Delta X_s^{\Sigma(J_2)} 1_{\{J_2 \in \{1,\dots,d_X\}^{|J_2|}\}}.
\end{aligned}$$

Note that since $\Sigma(J_2) > 2$ we have that $\Delta X_s^{\Sigma(J_2)} = \Delta Y_s^{\Sigma(J_2)}$. Moreover,

$$\begin{aligned}
\int_0^t \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle dW_s^{\mathbb{Q}} &= \int_0^t \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle dW_s^{\mathbb{P}} + f(s) \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle ds\\
&= \int_0^t \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle dW_s^{\mathbb{P}} + \sum_{|I|\leqslant p} f^I \langle \epsilon_{J'}^{\mathbb{P}} \sqcup \epsilon_I, \mathbb{Y}_{s-}\rangle ds
\end{aligned}$$

and $\int_0^t \int_{\mathbb{R}} \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle x(\mu - \nu^{\mathbb{Q}})(ds, dx)$ can be written as

$$\int_0^t \int_{\mathbb{R}} \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle x(\mu - \nu^{\mathbb{P}})(ds, dx) + \langle \epsilon_{J'}^{\mathbb{P}}, \mathbb{Y}_{s-}\rangle \int_{\mathbb{R}} x(e^{g(x)} - 1)F(dx)ds.$$

Thus, by Proposition 2.3.12 we can write $\langle \epsilon_J, \mathbb{X}_t^{\mathbb{P}}\rangle$ as

$$\begin{aligned}
&\langle (\epsilon_{J'}^{\mathbb{P}}; \epsilon_{j_k})^\sim, \mathbb{Y}_t\rangle + 1_{\{j_{k-1}=j_k=0\}} \frac{1}{2} \langle \epsilon_{J''}^{\mathbb{P}} \otimes \epsilon_{-1}, \mathbb{Y}_t\rangle\\
&+ \sum_{\epsilon_{J_1}\otimes\epsilon_{J_2}=\epsilon_J} \frac{1}{|J_2|!} 1_{\{|J_2|>1\}} \langle (\epsilon_{J_1}^{\mathbb{P}}; \epsilon_{\Sigma(J_2)})^\sim, \mathbb{Y}_t\rangle 1_{\{J_2\in\{1,\dots,d_X\}^{|J_2|}\}}\\
&+ 1_{\{j_k=0\}} \sum_{|I|\leqslant p} f^I \langle (\epsilon_{J'}^{\mathbb{P}} \sqcup \epsilon_I) \otimes \epsilon_0, \mathbb{Y}_t\rangle + 1_{\{j_k=1\}} \langle \epsilon_{J'}^{\mathbb{P}} \otimes \epsilon_0, \mathbb{Y}_t\rangle \int_{\mathbb{R}} x(e^{g(x)} - 1)F(dx),
\end{aligned}$$

concluding the proof.

Remark 2.B.2. For notational convenience, the function f in (2.3.21) has been defined via multi-indices I with indices in the set $\{-1, 0, 1, 2, \ldots, d_X\} \backslash \{0, 1\}$. However, observe that the result of Proposition 2.3.25 still holds if we consider the larger set of multi-indices with indices in $\{-1, 0, 1, 2, \ldots, d_X\} \backslash \{0\}$.

3 Sig-jump-diffusions

The present chapter is based on Cuchiero et al. (2024f), a joint work with Christa Cuchiero and Sara Svaluto-Ferro.

3.1 Introduction

The goal of this chapter is to introduce a general class of jump-diffusions based on the notion of signature. At the core of our analysis lie the universal approximation properties of the signature process, one of the main subjects of the thesis studied in detail in Chapter 2. According to this type of result, any continuous path functional can be approximated arbitrarily well by linear functions of the signature process (see Theorem 2.2.6 and the subsequent discussion for the precise statement). Relying on this, we define the so-called *Sig-jump-diffusion* as processes whose characteristics can be written as a linear combination of *potentially infinitely many components* of their own signature (see Definition 3.2.3). This general framework includes all the continuous signature models for asset pricing that have been proposed so far (see e.g., Arribas et al. (2021); Cuchiero et al. (2023a, 2024b)) where only diffusive drivers and finite sums of signature components are considered, as well as the Lévy-type signature modes presented in Chapter 2. Moreover, based on duality considerations for stochastic processes discussed in the introduction, we show that expected values of potentially infinitely many linear combinations of the signature components of such Sig-jump-diffusions can be computed analytically by solving certain linear ODEs in the extended tensor algebra (see Theorem 3.3.8). This extends to the non-continuous setting the recently established results of Cuchiero et al. (2023c). The technical difficulties already encountered in the diffusion case remain, in particular, due to convergence questions related to the infinite series of signature components. Additionally, the presence of a jump term complicates the analysis. Indeed, in order to establish duality results as in Theorem 3.3.8, one has to deal with a series of converging series involving signature's components (see equation (3.3.5)), making the task of establishing convergence considerably more complex.

We generally tackle these issues by considering the closure of an operator which is well-defined on the set of coefficients of linear functions of the signature (see Definition 3.3.6). In some specific cases, we can then establish precise convergence results. For instance, this is the case for the class of polynomial processes (see Remark 3.3.11(i)) as well as for the subclass of so-called *holomorphic jump-diffusions* analyzed in detail in Chapter 4. In particular, in Proposition 3.3.10 we show that a large class of polynomial jump-diffusions can be embedded in the Sig-jump-diffusions framework and that their truncated signature is again a polynomial process, resulting in an easily computable formula (based on a matrix exponential) for the expectation of linear functions of the signature process.

For related literature, e.g., Friz et al. (2022b); Jaber and Lin (2024) and the connections to the present work, we refer to Section 4.1.1.

The remainder of the chapter is structured as follows. In Section 3.2, we establish the proper notation for dealing with convergent series of the signature components and introduce the class of Sig-jump-diffusions. In Section 3.3 we present the main representation results (see Theorem 3.3.7 and Theorem 3.3.8). The last part of the section is dedicated to the analysis of the specific class of polynomial processes.

We refer to Chapter 1 for the notation and main results concerning weakly geometric càdlàg rough paths and their signature.

3.2 Sig-jump-diffusions

Let $G((\mathbb{R}^d))$ denote the set of group-like elements introduced in equation (1.1.6) and fix $S \subseteq \mathbb{R}^d$. Define

$$\mathcal{S} := \{\mathbf{x} \in G((\mathbb{R}^d)) : \langle \epsilon_I, \mathbf{x} \rangle \in S \text{ for all } |I| = 1\}$$

The set $\mathcal{S}$ is endowed with the sigma-algebra induced by the product sigma-algebra on $T((\mathbb{R}^d))$. Set

$$\mathcal{S}^* := \{\mathbf{u} \in T((\mathbb{R}^d)) + iT((\mathbb{R}^d)) \ : \ |\mathbf{u}|_{\mathbf{x}} < \infty \text{ for all } \ \mathbf{x} \in \mathcal{S}\}, \tag{3.2.1}$$

where $|\mathbf{u}|_{\mathbf{x}} := \sum_{n=0}^{\infty} |\sum_{|I|=n} \mathbf{u}_I \mathbf{x}_I|$. For $\mathbf{u} \in \mathcal{S}^*$ we write $\mathbf{u}^{\leqslant N} := \pi_{\leqslant N}(\mathbf{u}) \in T^N(\mathbb{R}^d)$, for each $N \in \mathbb{N}$, and

$$\langle \mathbf{u}, \mathbf{x} \rangle := \sum_{n=0}^{\infty} \sum_{|I|=n} \mathbf{u}_I \mathbf{x}_I, \tag{3.2.2}$$

for each $\mathbf{x} \in \mathcal{S}$. We now introduce the set of suitable coefficients of the processes which will be studied in this chapter.

Definition 3.2.1. We call *Sig-coefficients* any triplet (b, a, K) satisfying the following properties.

- $b : \mathcal{S} \to \mathbb{R}^d : b(\mathbf{x})_I = \langle \mathbf{b}^I, \mathbf{x} \rangle$, for $\mathbf{b}^I \in \mathcal{S}^*$ and $|I| = 1$. Note that $\mathbf{b} \in \mathcal{S}^{*d}$.
- $a : \mathcal{S} \to (\mathbb{R}^d)^{\otimes 2} : a(\mathbf{x})_I = \langle \mathbf{a}^I, \mathbf{x} \rangle$, for $\mathbf{a}^I \in \mathcal{S}^*$ and $|I| = 2$. Note that $\mathbf{a} \in (\mathcal{S}^{*d})^{\otimes 2}$. Assume that $a(\mathbf{x}) \in \mathbb{S}^d_+$ for all $\mathbf{x} \in \mathcal{S}$.
- K is a kernel from $\mathcal{S}$ to $\mathbb{R}^d$ such that for all $\mathbf{x} \in \mathcal{S}$, $K(\mathbf{x}, \{0\}) = 0$ and for all $|I| \geqslant 2$

$$\int_{\mathbb{R}^d} \xi_{i_1} \cdots \xi_{i_{|I|}} \ K(\mathbf{x}, d\xi) = \langle \mathbf{k}^I, \mathbf{x} \rangle \in (\mathbb{R}^d)^{\otimes n}, \tag{3.2.3}$$

for some $\mathbf{k}^I \in \mathcal{S}^*$. Note that $\int_{\mathbb{R}^d} \xi^{\otimes n} \, K(\mathbf{x}, d\xi) = \langle \mathbf{k}^n, \mathbf{x} \rangle \in (\mathbb{R}^d)^{\otimes n}$ for some $\mathbf{k}^n \in (\mathcal{S}^{*d})^{\otimes n}$.

Remark 3.2.2. Notice that (3.2.3) implies that for all $n \geqslant 2$, $\mathbf{x} \in \mathcal{S}$,

$$\int_{\mathbb{R}^d} (\|\xi\| \wedge \|\xi\|^2) K(\mathbf{x}, d\xi) < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} \|\xi\|^n K(\mathbf{x}, d\xi) < \infty.$$

Definition 3.2.3. Fix some Sig-coefficents (b, a, K). Let $X = (X_t)_{t \in [0,T]}$ be an S-valued semimartingale on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ such that for all bounded $f \in C^2(\mathbb{R}^d)$ the process $s \mapsto b(\mathbb{X}_{s-}), a(\mathbb{X}_{s-}), \int_{\mathbb{R}^d} g(\xi) K(\mathbb{X}_{s-}, d\xi)$ is left-continuous for $g = 1_A$ and A being a Borel set in $\mathbb{R}^d$, and $g(\xi) = \xi^{\otimes n}$ for $n \geqslant 2$. Suppose that the process $N^f := (N_t^f)_{t \in [0,T]}$ given by

$$N_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{A} f(\mathbb{X}_{s-}) ds, \quad t \in [0, T]$$

defines a local martingale, where for each $s \in [0, T]$

$$\begin{aligned} \mathcal{A} f(\mathbb{X}_{s-}) := \nabla f(X_{s-})^\top b(\mathbb{X}_{s-}) + \frac{1}{2} \mathrm{Tr}(a(\mathbb{X}_{s-}) \nabla^2 f(X_{s-})) \\ + \int_{\mathbb{R}^d} f(X_{s-} + \xi) - f(X_{s-}) - \nabla f(X_{s-})^\top \xi \; K(\mathbb{X}_{s-}, d\xi), \end{aligned}$$

and $\mathbb{X} := (\mathbb{X}_t)_{t \in [0,T]}$ denotes the signature of X. We call X *Sig-jump-diffusion*.

Remark 3.2.4. Notice that by Remark 3.2.2, Proposition II.2.29 and Theorem II.2.42 in Jacod and Shiryaev (1987), X is a special semimartingale with characteristics with respect to the truncation functional $\chi(\xi) = \xi$,

$$\left(\int_0^\cdot b(\mathbb{X}_{s-}) ds, \int_0^\cdot a(\mathbb{X}_{s-}) ds, K(\mathbb{X}_{s-}, d\xi) ds \right).$$

It is important to note that X is not necessarily a jump-diffusion in sense of Definition II.2.18 in Jacod and Shiryaev (1987). Modulo technicalities due to the infinite dimensionality, this is however the case for the signature process $\mathbb{X}$ (see Lemma 3.3.1).

3.3 Expected signature of Sig-jump-diffusions

As a first step, we study the dynamics of the signature of X.

Lemma 3.3.1. Let $X = (X_t)_{t\in[0,T]}$ be a Sig-jump-diffusion with Sig-coefficients (b, a, K) and fix a multi-index I. Then it holds that for all $t \in [0, T]$,

$$\begin{aligned}\langle\epsilon_I, \mathbb{X}_t\rangle = \int_0^t \Big(&\langle\epsilon_{I'}, \mathbb{X}_{s-}\rangle\langle\mathbf{b}^{i_{|I|}}, \mathbb{X}_{s-}\rangle 1_{|I|>0} + \frac{1}{2}\langle\epsilon_{I''}, \mathbb{X}_{s-}\rangle\langle\mathbf{a}^{i_{|I|-1}, i_{|I|}}, \mathbb{X}_{s-}\rangle 1_{|I|>1} \\ &+ \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}}\langle\epsilon_{I_1}, \mathbb{X}_{s-}\rangle \int_{\mathbb{R}^d}\langle\epsilon_{I_2}, \xi^{\otimes|I_2|}\rangle K(\mathbb{X}_{s-}, d\xi)\Big) ds \\ &+ \text{local martingale.}\end{aligned} \tag{3.3.1}$$

Proof. First note that by Remark 3.2.4 the characteristics of X are given by

$$\int_0^\cdot b(\mathbb{X}_{s-})ds, \int_0^\cdot a(\mathbb{X}_{s-})ds, K(\mathbb{X}_{s-}, d\xi)ds.$$

Next, fix $t \in (0, T]$ and recall that by definition $\mathbb{X}$ satisfied $\langle\epsilon_\varnothing, \mathbb{X}_t\rangle = 1$ and

$$\begin{aligned}\langle\epsilon_I, \mathbb{X}_t\rangle = &\int_0^t \langle\epsilon_{I'}, \mathbb{X}_{s-}\rangle dX_s^{i_{|I|}} 1_{|I|>0} + \frac{1}{2}\int_0^t \langle\epsilon_{I''}, \mathbb{X}_{s-}\rangle d[X^{i_{|I|-1}}, X^{i_{|I|}}]_s^c 1_{|I|>1} \\ &+ \sum_{0<s\leqslant t}\sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}}\langle\epsilon_{I_1}, \mathbb{X}_{s-}\rangle\langle\epsilon_{I_2}, (\Delta X_s)^{\otimes|I_2|}\rangle,\end{aligned} \tag{3.3.2}$$

for each $|I| > 0$ (see Section 1.4). Since $s \mapsto \int_{\mathbb{R}^d} \xi^{\otimes n} K(\mathbb{X}_{s-}, d\xi)$ is left-continuous for $n \geqslant 2$ we have that

$$\int_0^T \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>1\}} |\langle\epsilon_{I_1}, \mathbb{X}_{t-}\rangle| \int_{\mathbb{R}^d} |\langle\epsilon_{I_2}, \xi^{\otimes|I_2|}\rangle| K(\mathbb{X}_{t-}, d\xi)dt < \infty \quad a.s.$$

By Theorem II.1.8 in Jacod and Shiryaev (1987), the compensator of second line of (3.3.2) is given by

$$\int_0^\cdot \int_{\mathbb{R}^d} \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} 1_{\{|I_2|>0\}}\langle\epsilon_{I_1}, \mathbb{X}_{s-}\rangle\langle\epsilon_{I_2}, \xi^{\otimes|I_2|}\rangle K(\mathbb{X}_{t-}, d\xi)ds.$$

The claim follows by definition of characteristics. □

This result can be reformulated in a more compact way with the following notation.

Definition 3.3.2. Define

$$\mathbf{Q} := \Big(0, \mathbf{b}, \frac{1}{2}(\mathbf{a} + \mathbf{k}^2), \frac{1}{3!}\mathbf{k}^3, \dots, \frac{1}{n!}\mathbf{k}^n, \dots\Big) \in T((\mathcal{S}^{*d})).$$

For each $\mathbf{u} \in \mathcal{S}^*$ set

$$\mathbf{u}^{(k)} \dot{\sqcup\!\sqcup}\ \mathbf{Q} := \sum_{|I|=k} \mathbf{u}_I^{(k)} \sqcup\!\sqcup\ \mathbf{Q}^I \in \mathcal{S}^*.$$

Remark 3.3.3. For $\mathbf{x} \in \mathcal{S}$ set

$$Q(\mathbf{x}) := \left(0, b(\mathbf{x}), \frac{1}{2}(a(\mathbf{x}) + \int_{\mathbb{R}^d} \xi^{\otimes 2} K(\mathbf{x}, d\xi)), \dots, \frac{1}{k!}\int_{\mathbb{R}^d} \xi^{\otimes k} K(\mathbf{x}, d\xi), \dots\right) \tag{3.3.3}$$

and notice that for each I it holds $Q(\mathbf{x})_I = \langle \mathbf{Q}^I, \mathbf{x}\rangle$.

Corollary 3.3.4. *Let $X = (X_t)_{t\in[0,T]}$ be a Sig-jump-diffusion with Sig-coefficients (b, a, K) and fix a multi-index I. Then it holds that for all $t \in [0,T]$,*

$$\langle \mathbf{u}, \mathbb{X}_t\rangle = \langle \mathbf{u}, \epsilon_\varnothing\rangle + \int_0^t \langle L(\mathbf{u}), \mathbb{X}_s\rangle ds + M_t, \tag{3.3.4}$$

where M is a local martingale with $M_0 = 0$ and

$$L(\mathbf{u}) := \sum_{k=0}^{\infty} \mathbf{u}^{(k)} \mathbin{\dot{⧢}} \mathbf{Q}, \tag{3.3.5}$$

for all $\mathbf{u} \in T(\mathbb{R}^d)$.

Proof. Observe that for each I by Lemma 3.3.1 it holds $L(\epsilon_I) = \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \epsilon_{I_1} ⧢ \mathbf{Q}^{I_2}$. The result follows by linearity. $\square$

Remark 3.3.5. Notice that in (3.3.5) the sum is finite and thus well defined. In particular for each I it holds $\langle L(\epsilon_I), \mathbf{x}\rangle = \langle \epsilon_I, \mathbf{x} \otimes Q(\mathbf{x})\rangle$, where for each $\mathbf{x} \in \mathcal{S}$, $Q(\mathbf{x})$ is given in equation (3.3.3).

We now aim to extend (3.3.4) to some elements $\mathbf{u}$ of $T((\mathbb{R}^d))\backslash T(\mathbb{R}^d)$. As a first step, we extend the operator L to a subset of its pointwise closure.

Definition 3.3.6. For each $\mathbf{u} \in \mathcal{S}^*$ we write $\mathbf{u} \in \mathcal{D}(\overline{L})$ and

$$\langle \overline{L}(\mathbf{u}), \mathbf{x}\rangle := \lim_{N\to\infty} \langle L(\mathbf{u}^{\leqslant N}), \mathbf{x}\rangle$$

whenever the limit is well defined for each $\mathbf{x} \in \mathcal{S}$.

Theorem 3.3.7. *Let $X = (X_t)_{t\in[0,T]}$ be a Sig-jump-diffusion with Sig-coefficients (b, a, K) and denote by $\mathbb{X}$ its signature. Fix $\mathcal{U} \subseteq \mathcal{D}(\overline{L})$ and for each $\mathbf{u} \in \mathcal{U}$, define $V^1 := (V_t^2)_{t\in[0,T]}$ and $V^2 := (V_t^2)_{t\in[0,T]}$ via*

$$V_t^1 := \sup_{N\in\mathbb{N}} |\langle \mathbf{u}^{\leqslant N}, \mathbb{X}_t\rangle|, \qquad V_t^2 := \sup_{N\in\mathbb{N}} |\langle L(\mathbf{u}^{\leqslant N}), \mathbb{X}_t\rangle|,$$

for all $t \in [0,T]$. Suppose that $(\sup_{s\leqslant t} V_t^1)_{t\in[0,T]}$ is locally integrable and V^2 is a càdlàg process. Assume that the process $t \mapsto \langle \mathbf{u}, \mathbb{X}_t\rangle$ is càdlàg . Then the process $M := (M_t)_{\in[0,T]}$ defined via

$$M_t := \langle \mathbf{u}, \mathbb{X}_t\rangle - \langle \mathbf{u}, \epsilon_\varnothing\rangle - \int_0^t \langle \overline{L}(\mathbf{u}), \mathbb{X}_s\rangle ds, \tag{3.3.6}$$

is a local martingale with $M_0 = 0$ for each $\mathbf{u} \in \mathcal{U}$.

Proof. Fix $\mathbf{u} \in \mathcal{U}$. By Corollary 3.3.4, for all $N \in \mathbb{N}$ the process $M^N := (M_t^N)_{t\in[0,T]}$ given by

$$M_t^N := \langle \mathbf{u}^{\leqslant N}, \mathbb{X}_t \rangle - \langle \mathbf{u}^{\leqslant N}, \epsilon_\varnothing \rangle - \int_0^t \langle L(\mathbf{u}^{\leqslant N}), \mathbb{X}_u \rangle \, du, \tag{3.3.7}$$

is a local martingale. We show that $\sup_{N\in\mathbb{N}} \sup_{s\leqslant T} |M_{s\wedge\tau_n}^N|$ is integrable for $\tau_n := \tau_n^1 \wedge \tau_n^2$ where $(\tau_n^1)_{n\in\mathbb{N}}$ is a localizing sequence for the locally integrable process V^1 and

$$\tau_n^2 := \inf\{t > 0 \colon \sup_{N\in\mathbb{N}} |\langle L(\mathbf{u}^{\leqslant N}), \mathbb{X}_t \rangle| > n\}.$$

This implies that $M_{\cdot\wedge\tau_n}^N$ is a true martingale for each $N \in \mathbb{N}$ and by dominated convergence theorem the same holds for $M_{\cdot\wedge\tau_n} := \lim_{N\to\infty} M_{\cdot\wedge\tau_n}^N$. The claim then follows by noting that since $\mathbf{u} \in \mathcal{S}^*$, $\lim_{N\to\infty} \langle \mathbf{u}^{\leqslant N}, \mathbb{X}_t \rangle = \langle \mathbf{u}, \mathbb{X}_t \rangle$ and a further application of the dominated convergence theorem yields

$$\lim_{N\to\infty} \int_0^{t\wedge\tau_n} \langle L(\mathbf{u}^{\leqslant N}), \mathbb{X}_s \rangle \, ds = \int_0^{t\wedge\tau_n} \langle \overline{L}(\mathbf{u}), \mathbb{X}_s \rangle \, ds.$$

To see that $\sup_{N\in\mathbb{N}} \sup_{s\leqslant T} |M_{s\wedge\tau_n}^N|$ is integrable it suffices to note that

$$\sup_{s\leqslant T} |M_{s\wedge\tau_n}^N| \leqslant C + V_{T\wedge\tau_n}^1 + \int_0^{T\wedge\tau_n} V_u^2 du.$$

□

Next, we show that some expected quantities of Sig-jump-diffusions have a tractable representation. For a similar result concerning continuous processes, we refer to Theorem 4.29 in Cuchiero et al. (2023c).

Theorem 3.3.8. *Let $X = (X_t)_{t\in[0,T]}$ be a Sig-jump-diffusion with sig-coefficients (b, a, K), and $\mathcal{U} \subseteq \mathcal{D}(\bar{L})$ so that all the conditions of Theorem 3.3.7 are verified. Fix $\mathbf{u} \in \mathcal{U}$ and assume that there exists a $\mathcal{U}$-valued solution $(\mathbf{c}(t))_{t\in[0,T]}$ of the linear ODE*

$$\langle \mathbf{c}(t), \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{x} \rangle + \int_0^t \langle \overline{L}(\mathbf{c}(s)), \mathbf{x} \rangle ds, \tag{3.3.8}$$

for all $\mathbf{x} \in \mathcal{S}$. If furthermore

$$\mathbb{E}[\sup_{s,t\leqslant T} |\langle \mathbf{c}(s), \mathbb{X}_t \rangle|] < \infty \qquad \text{and} \qquad \mathbb{E}[\sup_{s,t\leqslant T} |\langle \overline{L}(\mathbf{c}(s)), \mathbb{X}_t \rangle|] < \infty,$$

then it holds that

$$\mathbb{E}[\langle \mathbf{u}, \mathbb{X}_T \rangle] = \langle \mathbf{c}(T), \epsilon_\varnothing \rangle.$$

Proof. The proof follows by Lemma A.1 in Cuchiero et al. (2023c). □

Finally, we focus on the class of polynomial processes as defined in Filipović and Larsson (2020). In particular we consider the class of jump-diffusions whose coefficients $(\bar{b}, \bar{a}, \bar{K})$ satisfies

$$\bar{b}_i \in \mathrm{Pol}_1(\mathbb{R}^d), \quad \bar{a}_{i_1 i_2} \in \mathrm{Pol}_2(\mathbb{R}^d), \quad \text{and} \quad \bar{m}_{i_1 \cdots i_n} := \int_{\mathbb{R}^d} \xi_{i_1} \cdots \xi_{i_n} \bar{K}(\cdot, d\xi) \in \mathrm{Pol}_n(\mathbb{R}^d), \tag{3.3.9}$$

for all $n \geqslant 2$. Note that the argument of each of these maps is the process itself and not its signature. This class is included in the class of Sig-jump-diffusion, as we show in the following lemma. To simplify the notation for a sufficiently differentiable map $f : \mathbb{R}^d \to \mathbb{R}$ we set

$$f^{(\varnothing)}(x) := f(x) \qquad \text{and} \qquad f^{(J)}(x) := \partial_{x_{j_1}} \cdots \partial_{x_{j_n}} f(x),$$

for $J = (j_1, \ldots, j_n)$.

Lemma 3.3.9. $X = (X_t)_{t \in [0,T]}$ is an S-valued polynomial jump-diffusion satisfying (3.3.9) if and only if X is a Sig-jump-diffusion whose Sig-coefficients (b, a, K) satisfy the following conditions.

- $b : \mathcal{S} \to \mathbb{R}^d : b(\mathbf{x})_I = \langle \mathbf{b}^I, \mathbf{x} \rangle$, for $|I| = 1$ where $\mathbf{b}^I \in T^1(\mathbb{R}^d)$ is given by

$$\mathbf{b}^I_J = \bar{b}^{(J)}_{i_1}(X_0) \qquad \text{for each } |J| \leqslant 1.$$

- $a : \mathcal{S} \to (\mathbb{R}^d)^{\otimes 2} : a(\mathbf{x})_I = \langle \mathbf{a}^I, \mathbf{x} \rangle$, for $|I| = 2$ where $\mathbf{a}^I \in T^2(\mathbb{R}^d)$ is given by

$$\mathbf{a}^I_J = \bar{a}^{(J)}_{i_1 i_2}(X_0) \qquad \text{for each } |J| \leqslant 2.$$

- For all $|I| \geqslant 2$ we have that

$$m_I(\mathbf{x}) := \int_{\mathbb{R}^d} \xi_{i_1} \cdots \xi_{i_{|I|}} K(\mathbf{x}, d\xi) = \langle \mathbf{k}^I, \mathbf{x} \rangle,$$

  where $\mathbf{k}^I \in T^{|I|}(\mathbb{R}^d)$ is given by

$$\mathbf{k}^I_J = \bar{m}^{(J)}_{i_1 \cdots i_{|I|}}(X_0) \qquad \text{for each } |J| \leqslant |I|.$$

Proof. The proof follows by performing a Taylor expansion of the maps in (3.3.9), the definition of $\mathcal{S}$ and equation (1.1.9). $\square$

To conclude, we show that for a polynomial process and any initial condition $\mathbf{u} \in T(\mathbb{R}^d)$, the infinite-dimensional ODE (3.3.8) reduces to a finite-dimensional system of linear ODEs.

Proposition 3.3.10. Let $X = (X_t)_{t\in[0,T]}$ be an S-valued polynomial jump-diffusion with coefficients $(\bar{b}, \bar{a}, \bar{K})$. Then, for each $N \in \mathbb{N}$, the truncated signature $\mathbb{X}^N$ of X is a $T^N(\mathbb{R}^d)$-valued polynomial process. This in particular implies that all the moments of $\mathbb{X}^N$ exist and are finite. Moreover,

$$\mathbb{E}[\mathbb{X}_t^N] = \exp(tL)\mathbb{X}_0^N \tag{3.3.10}$$

where $\exp(tL)$ is the exponential of the finite-dimensional linear operator given by the restriction of tL to $T^N(\mathbb{R}^d)$.

Proof. First note that by Lemma 1.1.9 for each $\mathbf{x} \in \mathcal{S}$ and $|I| \leqslant 2$ it holds that $a(\mathbf{x})_I = \bar{a}_{i_1 i_2}(\pi_1(\mathbf{x}) + X_0)$, and for each $|I| \geqslant 2$ it holds that $m(\mathbf{x})_I = \bar{m}_{i_1\cdots i_{|I|}}(\pi_1(\mathbf{x}) + X_0)$. This in particular implies both maps can be seen as polynomials of degree at most $|I|$ in $\pi_1(\mathbf{x})$.

The proof of the polynomial property of $\mathbb{X}^N$ follows the proof of Proposition 2.3. In particular, we get that $\mathbb{X}^N$ is a jump-diffusion whose coefficients

$$(\beta, \zeta, \mathcal{K}) := ((\beta_I)_{|I|\leqslant N}, (\zeta_{I,J})_{|I|,|J|\leqslant N}, \mathcal{K})$$

are given by

$$\begin{aligned}
\beta_I(\mathbf{x}^N) &:= \langle \epsilon_{I'} \sqcup \mathbf{b}^{(i_{|I|})}, \mathbf{x}^N \rangle 1_{|I|>0} + \frac{1}{2}\langle \epsilon_{I''} \sqcup \mathbf{a}^{(i_{|I|-1}, i_{|I|})}, \mathbf{x}^N \rangle 1_{|I|\geqslant 2} \\
&\quad + \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} 1_{|I_2|>1} \frac{1}{|I_2|!} \langle \epsilon_{I_1} \sqcup \mathbf{k}^{I_2}, \mathbf{x}^N \rangle, \\
\zeta_{IJ}(\mathbf{x}^N) &:= \langle \epsilon_{I'}, \mathbf{x}^N \rangle \langle \epsilon_{J'}, \mathbf{x}^N \rangle \bar{a}_{i_{|I|} j_{|J|}}(\pi_1(\mathbf{x}^N) + X_0) 1_{|I|,|J|\geqslant 1}, \\
\mathcal{K}(\mathbf{x}^N, A) &:= \int_{\mathbb{R}^d} 1_{A\setminus\{0\}}(\delta(\mathbf{x}^N, \xi)) \bar{K}(\pi_1(\mathbf{x}^N), d\xi),
\end{aligned}$$

for $\mathbf{x}^N \in T^N(\mathbb{R}^d)$ and $\delta : T^N(\mathbb{R}^d) \times \mathbb{R}^d \to T^N(\mathbb{R}^d)$ defined via

$$\delta(\mathbf{x}^N, \xi)_I := \sum_{\epsilon_{I_1}\otimes\epsilon_{I_2}=\epsilon_I} \frac{1}{|I_2|!} \langle \epsilon_{I_1}, \mathbf{x}^N \rangle \langle \epsilon_{I_2}, \xi^{\otimes |I_2|} \rangle 1_{\{|I_2|>0\}},$$

for each $\xi \in \mathbb{R}^d$ and $|I| \leqslant N$. By Lemma 2.2 in Filipović and Larsson (2020) it suffices to check that $\zeta_{IJ} \in \mathrm{Pol}_2(T^N(\mathbb{R}^d))$ and

$$\int_{\mathbb{R}^d} \delta(\mathbf{x}^N, \xi)_{J_1} \cdots \delta(\mathbf{x}^N, \xi)_{J_k} \bar{K}(\pi_1(\mathbf{x}^N), d\xi) \in \mathrm{Pol}_k(T^N(\mathbb{R}^d)).$$

For fixed ij we have that $\bar{a}_{ij}(\pi_1(\mathbf{x}^N)+X_0) = \sum_{|H_1|,|H_2|\leqslant 1} \alpha_{H_1H_2} \langle \epsilon_{H_1}, \mathbf{x}^N \rangle \langle \epsilon_{H_2}, \mathbf{x}^N \rangle$. Since $|I'|, |J'| \leqslant N-1$ we get that

$$\zeta_{IJ}(\mathbf{x}^N) := \sum_{|H_1|,|H_2|\leqslant 1} \alpha_{H_1H_2} \langle \epsilon_{I'} \sqcup e_{H_1}, \mathbf{x}^N \rangle \langle \epsilon_{J'} \sqcup e_{H_2}, \mathbf{x}^N \rangle 1_{|I|,|J|\geqslant 1}$$

and the first claim follows. The second claim follows similarly by the polynomial property of $\bar{m}_{i_1\cdots i_{|I|}}$ and we can conclude that the truncated signature $\mathbb{X}^N$ of X is a $T^N(\mathbb{R}^d)$-valued polynomial process. (3.3.10) follows then by an application of Theorem 2.5 in Filipović and Larsson (2020) for the basis $\{\langle \epsilon_I, \cdot\rangle \colon |I| \leqslant N\}$.

□

Remark 3.3.11. (i) In the proof of Proposition 3.3.10 we have shown that the truncated signature of any polynomial process having characteristics as in (3.3.9) is a polynomial process on the truncated tensor algebra. Furthermore, we have used the results in Filipović and Larsson (2020) to derive the representation in (3.3.10). Similarly, one could use the properties of polynomial processes to show that the conditions of Theorem 3.3.8 are satisfied and then derive the representation (3.3.9) from there.

(ii) Notice that Proposition 3.3.10 is an extension of Proposition 2.3.

(iii) Let $X = (X_t)_{t\in[0,T]}$ be a market primary process as in Definition 2.3.1 and let $S = (S_t)_{t\in[0,T]}$ be a Lévy-type signature model as in Definition 2.3.6. Then, the joint process $Z := (X, S)$ is a Sig-jump-diffusion.

4 Holomorphic jump-diffusions

The present chapter is based on Cuchiero et al. (2024e), a joint work with Christa Cuchiero and Sara-Svaluto Ferro.

4.1 Introduction

The goal of this chapter is to introduce a class of jump-diffusion processes, called *holomorphic*, for which the calculation of the expected values of the restrictions of holomorphic functions to $\mathbb{R}^d$, of the process' marginals, reduces to solving a sequence-valued linear ordinary differential equation (ODE). This thus constitutes an important extension of *polynomial processes,* introduced in Cuchiero et al. (2012); Filipović and Larsson (2016), for which merely moments can be computed by solving a (finite dimensional) linear ODE.

At the core of our analysis lie duality considerations which are a key concept in many areas of mathematics and have also played an important role in the analysis of stochastic processes. Indeed, duality theory for Markov processes with respect to a duality function goes back to several contributions in the early fifties, e.g., Karlin and McGregor (1957), where classifications of birth and death processes are considered. Since then it has been extended in several directions (see e.g., Holley and Stroock (1979); Ethier and Kurtz (1986)) and applied in the context of interacting particle systems, queuing theory and population genetics.

The concept of dual processes can be formalized as follows: let $T > 0$ be some finite time-horizon and consider two time-homogeneous Markov processes $(X_t)_{t\in[0,T]}$ and $(U_t)_{t\in[0,T]}$ with respective state spaces S and U. Then X and U are *dual* with respect to some measurable function $H : S \times U \to \mathbb{R}$ if for all $x \in S$, $u \in U$ and $t \in [0, T]$

$$\mathbb{E}_x[H(X_t, u)] = \mathbb{E}_u[H(x, U_t)], \tag{4.1.1}$$

holds, and both the left and right-hand sides are well-defined. Here, $\mathbb{E}_x$ denotes the expected value for the Markov process X starting at $X_0 = x$ and similarly for U. Modulo technical conditions, the dual relation (4.1.1) holds if and only if the generators denoted by $\mathcal{A}$ and $\mathcal{B}$ satisfy

$$\mathcal{A}H(\cdot, u)(x) = \mathcal{B}H(x, \cdot)(u), \quad \text{for all } x \in S, u \in U, \tag{4.1.2}$$

(see e.g., Jansen and Kurt (2014)). One prominent example is the Wright-Fisher diffusion, also called the Jacobi process and denoted now by X, whose dual process with respect to

$H(x,u) = x^u$ with $u \in \mathbb{N}$ is the Kingman coalescent U, i.e., we have

$$\mathbb{E}_x[X_t^u] = \mathbb{E}_u[x^{U_t}].$$

The Wright-Fisher diffusion is also an example of a polynomial process and for all functions $H(x,u) : [0,1] \times \mathbb{R}^n \to \mathbb{R}$ given by $H(x,u) = \sum_{i=0}^{n-1} u_i x^i$ it holds by the so-called *moment formula* (see e.g., Filipović and Larsson (2016)) that

$$\mathbb{E}_x[H(X_t,u)] = \mathbb{E}_x\left[\sum_{i=0}^{n-1} u_i X_t^i\right] = \sum_{i=0}^{n-1} c(t)_i x^i = H(x,c(t)) = \mathbb{E}_u[H(x,c(t))],$$

where $c(t)$ is the solution of a linear ordinary differential equation (ODE). This means that in this case the dual process is given by the coefficients $U_t = c(t)$ (with $U_0 = c(0) = u$) of the polynomial $x \mapsto H(x,c(t))$, which are as solution of a linear ODE deterministic. This property actually defines polynomial processes in the sense that the expected value of a polynomial of X_t for $t \in [0,T]$ is a polynomial in the initial value $X_0 = x$. Another class of stochastic processes which admit a deterministic dual process with respect to certain functions H are *affine processes* (see Duffie et al. (2003); Cuchiero and Teichmann (2013)). In this case the duality functions $H : S \times U \to \mathbb{R}$ are given by $H(x,u) = \exp(\langle u,x\rangle)$ where $S \subseteq \mathbb{R}^d$, $U \subseteq \mathbb{C}^d$ and $\langle\cdot,\cdot\rangle$ denotes the scalar product on $\mathbb{R}^d$ (extended to $\mathbb{C}^d$). The corresponding dual process is a solution to a generalized Riccati ODE.

These deterministic dual processes for H ranging in the important and law determining[1] function classes of polynomials and exponential functions (when viewed as functions in x) are the crucial property for the popularity of affine and polynomial processes for all kind of applications including finance, population genetics and physics. Indeed, the computation of expected values for these functions H reduces to solving simple deterministic ODEs.

It is therefore natural to search of other stochastic processes and functions which admit a deterministic dual process. The most natural extension of both polynomials and exponentials are *entire functions*, or the subclass of *holomorphic functions*, which are given as power series on $\mathbb{C}^d$, whose radius of convergence is not necessarily the whole space. The goal of the present paper is thus to define and specify a class of processes X such that

$$\mathbb{E}_x[H(X_t,\mathbf{u})] = H(x,\mathbf{c}(t)), \tag{4.1.3}$$

where H is (the restriction to $\mathbb{R}^d$ of) a holomorphic function on $\mathbb{C}^d$, given by the power series $H(x,\mathbf{u}) = \sum_{\alpha\in\mathbb{N}_0^d} \mathbf{u}_\alpha x^\alpha$ where $\alpha \in \mathbb{N}_0^d$ is a multi-index and $\mathbf{u}_\alpha \in \mathbb{C}$ are the coefficients. This rather wide extension of polynomial processes is twofold: first, it allows go beyond them and second, we can establish an analog of the moment formula, which we call *holomorphic formula* given by (4.1.3), also for affine and certain polynomial processes (see Section 4.3.5), including for instance the Wright-Fisher diffusion (see Section 6.2 of Cuchiero et al. (2023c)).

[1] In the polynomial case the law is determined by the moments under certain exponential moment conditions, while the characteristic function in the affine case always determines the law.

In contrast to the polynomial case, $(\mathbf{c}(t))_{t\in[0,T]}$ in (4.1.3) is a sequence-valued infinite dimensional deterministic process, given as a solution of an infinite dimensional linear ODE. This already indicates that the analysis becomes much more delicate, as we have to deal with existence of such solutions, convergence of the corresponding power series and many other subtleties from complex analysis.

The fact that we consider *jump-diffusions* instead of just continuous processes makes things even further involved. Indeed, already establishing the necessary condition 4.1.2, which in the current setup means that the generator of X maps holomorphic functions to holomorphic ones, is due to the appearance of jumps not at all straightforward (see Theorem 4.3.6). Indeed, one has to define specific subsets of holomorphic functions, denoted by $\mathcal{V}$, where this holds true, see e.g., the set defined in (4.3.7). This is the very reason why we actually consider *$\mathcal{V}$-holomorphic jump-diffusionses* which we define as solutions to the martingale problem specified by the (extended) generator $\mathcal{A}$ which maps $\mathcal{V}$ to general holomorphic functions. Analyzing then different sets $\mathcal{V}$ in dependence of the specificities of the compensator of the jumps requires to apply a number of results of complex analysis and is subject of Section 4.3.1 (see in particular Theorem 4.3.12). For such $\mathcal{V}$-holomorphic jump-diffusionses we then establish the holomorphic formula (4.1.3), under technical conditions involving in particular the existence of solutions to the sequence-valued linear ODEs and certain moment conditions. This can be viewed as a verification result (see Theorem 4.3.20). To apply this result we establish further sufficient conditions which are then verified for Lévy processes, affine processes and certain non-polynomial jump diffusions on compact sets (see Section 4.3.5), which is one of the main contributions of this paper.

In analogy to affine processes, we can, of course, also consider processes where the expected value of the exponential of a holomorphic function is given as the exponential of a holomorphic function whose coefficients solve a sequence-valued Riccati ODE. This is subject of Section 4.4 where we introduce the class of *affine-holomorphic jump-diffusions*, explain their relation to holomorphic ones and elaborate on their characteristics as well as appropriate subsets of holomorphic functions that are mapped to (general) holomorphic functions by the Riccati operator. Here, the jumps are again the essential part which makes the analysis intricate (see Theorem 4.4.3). Under technical conditions, involving in particular the existence of solutions of the infinite dimensional Riccati ODEs, we then prove the so-called affine-holomorphic formula (see Theorem 4.4.10).

In terms of applications our results can be used for pricing and hedging real analytic claims in finance and new duality relations in population genetics as e.g., in Casanova and Spanò (2018); Jochen Blath and Adrián González Casanova and Noemi Kurt and Maite Wilke-Berenguer (2016). Reading the duality relation backwards, the (affine)-holomorphic formula also allows to develop numerical schemes for solving infinite dimensional linear and Riccati ODEs based on the stochastic representations.

The remainder of the paper is organized as follows. In the subsections below we clarify the relations to the literature and fix the terminology related to holomorphic functions. Section 4.2 introduces all necessary notation, general jump-diffusion processes and convergent power series for given state spaces. Section 4.3 and Section 4.4 are then dedicated to the analysis of holomorphic and affine-holomorphic jump-diffusions respectively. The paper concludes with several appendices, containing in particular the proofs and important results from complex analysis (see Appendix 4.D).

4.1.1 Relation to the literature

Our expressions for the expected valued of holomorphic functions either in terms of the holomorphic formula or the affine-holomorphic formula are related to several recent works in the literature. Indeed, in the one-dimensional case similar results have been obtained in Cuchiero et al. (2023c), however only for continuous processes and real analytic functions, which is a significantly simpler setting. Note that the main focus of Cuchiero et al. (2023c) are actually continuous signature SDEs which reduce in the one dimensional case to continuous holomorphic jump-diffusions. We also refer to Remark 4.3.2 for another relation with signature SDEs. Developing signature jump-diffusions and in turn the theory of holomporphic jump-diffusions on the extended tensor algebra is subject of future work and already partly realized in Chapter 3. Here, again the tools of complex analysis play a major role.

Our expressions for the expected values of power series in terms of power series are also related to similar expansions, obtained e.g., in Friz et al. (2022b,a); Friz and Gatheral (2022); Fukasawa and Matsushita (2021); Alos et al. (2020). For applications to Fourier pricing in finance we refer to Abi Jaber and Gérard (2024); Jaber and Lin (2024).

In terms of polynomial processes which – as already mentioned – constitue a special case of holomorphic jump-diffusions, let us in particular mention Benth et al. (2021) and the recent paper Benth et al. (2020) where an abstract far reaching approach to polynomial processes is proposed. This covers also infinite dimensional polynomial processes considered in Benth et al. (2018); Cuchiero and Svaluto-Ferro (2021); Cuchiero et al. (2024a) as well. The latter can also be related to lifts of polynomial Volterra processes which were recently analyzed in Abi Jaber et al. (2024).

4.1.2 Terminology

We aim to recall several notions related to holomorphic functions and fix some terminology that we shall use throughout.

A holomorphic function with values in $\mathbb{C}$ is a map defined on (some subset of) the complex plane that is complex differentiable at every point within its domain (see Definition 4.D.1, where more generally $\mathbb{C}^m$-valued holomorphic functions are introduced). Despite the resemblance of this class with the class of just differentiable functions of one real variable, the former satisfy much stronger properties. A holomorphic function is actually infinitely many times complex differentiable, that is, the existence of the first derivative guarantees the existence of the derivatives of any order. In fact, more is true: every holomorphic function is *complex analytic*, in the sense that it has a power series expansion near every point. This is also the reason why the term (complex) analytic is frequently used as a synonym for holomorphic. Moreover, any real analytic function on some open set on the real line can be extended to a complex analytic, and thus holomorphic, function on some open set of the complex plane. (However, not every real analytic function defined on the whole real line can be extended to a complex function defined on the whole complex plane). In particular, any real convergent power series can be extended to a complex holomorphic function on some open disc of the complex plane.

4.2 Preliminaries

This section is primarily dedicated to introduce necessary notation and the notion of jump-diffusion processes with which we shall work. It is important to note that this chapter employs notation independent of that used in the other chapters.

4.2.1 Notation

4.2.1.1 Multi-index notation

Fix $d \in \mathbb{N}$. For $z \in \mathbb{C}^d$ we write $z_1, \ldots, z_d$ for the components of z and set

$$|z| := (|z_1|, \ldots, |z_d|) \qquad \text{and} \qquad \|z\| := \sqrt{|z_1|^2 + \ldots + |z_d|^2},$$

where $|z_j|$ denotes denote the modulus of z_j. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ we use the following notation: $|\alpha| := \alpha_1 + \cdots + \alpha_d$, $\alpha! := \alpha_1! \ldots \alpha_d!$, $z^\alpha := z_1^{\alpha_1} \ldots z_d^{\alpha_d}$ for $z \in \mathbb{C}^d$, and for $d = 1$, we disregard the brackets and simply write $\alpha \in \mathbb{N}_0$. We set

$$h^{(\alpha)}(z) := D^\alpha h(z) = \frac{\partial D^{|\alpha|} h}{\partial^{\alpha_1} z_1 \ldots \partial^{\alpha_d} z_d}(z),$$

for sufficiently regular maps $h : U \to \mathbb{C}$ and some open set $U \subseteq \mathbb{C}^d$ (or $U \subseteq \mathbb{R}^d$), with $z \in U$, and $\|h\|_\infty := \sup_{z \in U} \|h(z)\|$. Moreover, we write $\nabla h(z)$ and $\nabla^2 h(z)$ to denote the gradient and Hessian matrix of h at z, respectively. For $d = 1$ we also write h' and h'', respectively.

4.2.1.2 The space of sequences

We introduce the notation for representing power series in d variables. Throughout, we use bold letters to denote sequences $\mathbf{u} := (\mathbf{u}_\alpha)_{\alpha \in \mathbb{N}_0^d}$ with $\mathbf{u}_\alpha \in \mathbb{C}$ indexed by multi-indices and set $|\mathbf{u}| := (|\mathbf{u}_\alpha|)_{\alpha \in \mathbb{N}_0^d}$. To denote vectors and matrices of sequences we then make use of underlining. Specifically, we write

$$\underline{\mathbf{u}} := (\underline{\mathbf{u}}^1, \ldots, \underline{\mathbf{u}}^d) \qquad \text{and} \qquad \underline{\underline{\mathbf{u}}} := \begin{pmatrix} \underline{\underline{\mathbf{u}}}^{11} & \cdots & \underline{\underline{\mathbf{u}}}^{1d} \\ \vdots & \ddots & \vdots \\ \underline{\underline{\mathbf{u}}}^{d1} & \cdots & \underline{\underline{\mathbf{u}}}^{dd} \end{pmatrix},$$

where $\underline{\mathbf{u}}^i = (\underline{\mathbf{u}}^i_\alpha)_{\alpha\in\mathbb{N}_0^d}$ and $\underline{\underline{\mathbf{u}}}^{ij} = (\underline{\underline{\mathbf{u}}}^{ij}_\alpha)_{\alpha\in\mathbb{N}_0^d}$ for $\underline{\mathbf{u}}^i_\alpha, \underline{\underline{\mathbf{u}}}^{ij}_\alpha \in \mathbb{C}$. We let $\mathbf{1} := (\mathbf{1}_\alpha)_{\alpha\in\mathbb{N}_0^d}$ denote the sequence such that $\mathbf{1}_\alpha = 1$ if $\alpha = (0, \dots, 0)$ and $\mathbf{1}_\alpha = 0$ otherwise. To simplify the notation, we denote by $(\epsilon_i)_{i\in\{1,\dots,d\}}$ the canonical basis of $\mathbb{R}^d$ and write

$$\underline{\mathbf{u}}^{\epsilon_i} := \underline{\mathbf{u}}^i, \qquad \text{and} \quad \underline{\underline{\mathbf{u}}}^{2\epsilon_i} := \underline{\underline{\mathbf{u}}}^{ii}, \qquad \underline{\underline{\mathbf{u}}}^{\epsilon_i+\epsilon_j} := \underline{\underline{\mathbf{u}}}^{ij}.$$

This is useful to access also the components of $\underline{\mathbf{u}}$ and $\underline{\underline{\mathbf{u}}}$ via the multi-index-notation, namely $\underline{\mathbf{u}}^\beta$ and $\underline{\underline{\mathbf{u}}}^\beta$ for $|\beta| = 1, 2$, respectively.

Finally, in this paper integrals of sequence-valued maps are always computed componentwise. Precisely, given a measure F on a measurable space E and a map $((\mathbf{f}(\cdot))_\alpha)_{\alpha\in\mathbb{N}_0^d}$ on E such that $(\mathbf{f}(y))_\alpha \in \mathbb{C}$, we define

$$\Big(\int_E \mathbf{f}(y)\ F(dy)\Big)_\alpha := \int_E (\mathbf{f}(y))_\alpha\ F(dy),$$

whenever the involved quantities are well defined. The same extends to vector and matrices of sequence valued maps.

4.2.1.3 The set of holomorphic functions on polydiscs

Let $d \in \mathbb{N}$, $R = (R_1, \dots, R_d) \in (0, \infty]^d$, $a \in \mathbb{C}^d$ and denote by $P^d_R(a)$ the complex *polydisc* centered at a with *polyradius* R:

$$P^d_R(a) := \{z \in \mathbb{C}^d : |z_i - a_i| < R_i \text{ for all } i \in \{1, \dots, d\}\},$$

where $|\cdot|$ denotes the complex modulus. This in particular implies that

$$P^d_R(a) = P^1_{R_1}(a_1) \times \cdots \times P^1_{R_d}(a_d) \subseteq \mathbb{C}^d,$$

where $P^1_{R_j}(a_j)$ denotes the complex disc centered at a_j and radius R_j. Observe that for $R_j = \infty$ we get $P^1_{R_j}(a_j) = \mathbb{C}$. If $R \in (0, \infty)^d$ we let $T^d_R(a)$ denote the boundary of the polydisc $P^d_R(a)$, also called *polytorus*, defined via

$$T^d_R(a) := \{z \in \mathbb{C}^d : |z_i - a_i| = R_i \text{ for all } i \in \{1, \dots, d\}\}. \tag{4.2.1}$$

The closure $\overline{P^d_R(a)}$ of the polydisc is given by $\overline{P^d_R(a)} = \{z \in \mathbb{C}^d : |z_i - a_i| \leqslant R_i \text{ for all } i\}$. Moreover, for any $R \in (0, \infty]$ we denote the polyradius $(R, \dots, R)$ again by R. Given two polyradii $M, N \in (0, \infty]^d$, we write $M > N$ if for all $j = 1, \dots, d$ $M_j > N_j$.

We denote by $H(P^d_R(0), \mathbb{C}^m)$, $m \in \mathbb{N}$, the set of holomorphic functions on $P^d_R(0)$ with values in $\mathbb{C}^m$ (see Definition 4.D.1). When $m = 1$, we simply write $H(P^d_R(0))$. When a holomorphic function on a polydisc has a power series representation, it can be identified through the corresponding coefficients (see Proposition 4.D.2(ii)). Setting

$$|\mathbf{u}|_z := \sum_{\alpha\in\mathbb{N}_0^d} \Big|\mathbf{u}_\alpha \frac{z^\alpha}{\alpha!}\Big|, \tag{4.2.2}$$

we define

$$\mathcal{P}_R^d(0)^* := \{\mathbf{u} = (\mathbf{u}_\alpha)_{\alpha\in\mathbb{N}_0^d} \ : \ \mathbf{u}_\alpha \in \mathbb{C} \text{ and } |\mathbf{u}|_z < \infty \text{ for all } z \in P_R^d(0)\}.$$

For $\mathbf{u} \in \mathcal{P}_R^d(0)^*$, we let $h_{\mathbf{u}} \in H(P_R^d(0))$ denote the holomorphic function on $P_R^d(0)$ determined by the corresponding convergent power series

$$h_{\mathbf{u}}(z) := \sum_{\alpha\in\mathbb{N}_0^d} \mathbf{u}_\alpha \frac{z^\alpha}{\alpha!}, \qquad z \in P_R^d(0).$$

Similarly, we write $\underline{\mathbf{u}} \in (\mathcal{P}_R^d(0)^*)^d$, $\underline{\underline{\mathbf{u}}} \in (\mathcal{P}_R^d(0)^*)^{d\times d}$ if each $\underline{\underline{\mathbf{u}}}^{ij}, \underline{\mathbf{u}}^i \in \mathcal{P}_R^d(0)^*$ and denote by $h_{\underline{\mathbf{u}}} \in H(P_R^d(0), \mathbb{C}^d)$ and $h_{\underline{\underline{\mathbf{u}}}} \in H(P_R^d(0), \mathbb{C}^{d\times d})$ the corresponding holomorphic functions.

4.2.1.4 Operations on holomorphic functions and sequences

Now we introduce the operations on holomorphic functions that will be used throughout the paper and describe how these can be translated into operations on the corresponding sequence of coefficients, and vice-versa. The relevant operations in our setting are algebraic operations, differentiation and exponentiation.

- **Linear operations**: Through componentwise linear operations we obtain the relation
$$\lambda h_{\mathbf{u}}(z) + \mu h_{\mathbf{v}}(z) = h_{\lambda\mathbf{u}+\mu\mathbf{v}}(z)$$
for each $z \in P_R^d(0)$, $\mathbf{u}, \mathbf{v} \in \mathcal{P}_R^d(0)^*$ and $\lambda, \mu \in \mathbb{C}$.
- **Product**: We consider the symmetric bilinear map on $\mathcal{P}_R^d(0)^*$ given by
$$\begin{aligned} \mathcal{P}_R^d(0)^* \times \mathcal{P}_R^d(0)^* &\longrightarrow \mathcal{P}_R^d(0)^* \\ (\mathbf{u}, \mathbf{v}) &\longmapsto \mathbf{u} * \mathbf{v} \end{aligned}$$
where $\mathbf{u} * \mathbf{v} := ((\mathbf{u} * \mathbf{v})_\alpha)_{\alpha\in\mathbb{N}_0^d}$ for
$$(\mathbf{u} * \mathbf{v})_\alpha = \sum_{\gamma+\beta=\alpha} \frac{\alpha!}{\beta!\gamma!} \mathbf{u}_\beta \mathbf{v}_\gamma, \quad \alpha \in \mathbb{N}_0^d.$$
Notice that for all $z \in P_R^d(0)$ and $k \in \mathbb{N}$
$$h_{\mathbf{u}}(z) h_{\mathbf{v}}(z) = h_{\mathbf{u}*\mathbf{v}}(z) \qquad \text{and} \qquad (h_{\mathbf{u}}(z))^k = h_{\mathbf{u}^{*k}}(z),$$
where $\mathbf{u}^{*1} = \mathbf{u}$ and $\mathbf{u}^{*k} := \mathbf{u}^{*(k-1)} * \mathbf{u}$ for each $k > 1$. We also set $\mathbf{u}^{*0} := (1, 0, 0, \ldots)$.
Through the multi-index notation, these representations extend to vector-valued functions. More precisely, setting $\underline{\mathbf{u}} := (\underline{\mathbf{u}}^1, \ldots, \underline{\mathbf{u}}^d) \in (\mathcal{P}_R^d(0)^*)^d$ we define $\underline{\mathbf{u}}^{*\beta} := (\underline{\mathbf{u}}^1)^{*\beta_1} * \cdots * (\underline{\mathbf{u}}^d)^{*\beta_d}$ and then get
$$(h_{\underline{\mathbf{u}}}(z))^\beta = h_{\underline{\mathbf{u}}^{*\beta}}(z).$$

- **Differentiation**: For $\mathbf{u} \in \mathcal{P}_R^d(0)^*$ and $\beta \in \mathbb{N}_0^d$ we set $\mathbf{u}^{(\beta)} := (\mathbf{u}_{\alpha+\beta})_{\alpha \in \mathbb{N}_0^d}$. Observe that $\mathbf{u}^{(\beta)} \in \mathcal{P}_R^d(0)^*$ and

$$h_{\mathbf{u}}^{(\beta)}(z) = \sum_{\alpha \in \mathbb{N}_0^d} \mathbf{u}_{\alpha+\beta} \frac{z^\alpha}{\alpha!} = h_{\mathbf{u}^{(\beta)}}(z),$$

for all $z \in P_R^d(0)$.

- **Exponentiation**: For $\mathbf{u} \in \mathcal{P}_R^d(0)^*$, we denote by $\mathbf{exp^*}(\mathbf{u}) \in \mathcal{P}_R^d(0)^*$ the coefficients determining the power series representation on $P_R^d(0)$ of $\exp(h_{\mathbf{u}})$, namely

$$\exp(h_{\mathbf{u}}(z)) = h_{\mathbf{exp^*}(\mathbf{u})}(z),$$

for all $z \in P_R^d(0)$.

- **Composition**: Fix $\underline{\mathbf{v}} \in (\mathcal{P}_R^d(0)^*)^d$ and $\mathbf{u} \in \mathcal{P}_N^d(0)^*$ for N such that

$$N_j > \sup_{z \in \mathcal{P}_R^d(0)} \big(|z_j| + |h_{\underline{\mathbf{v}}}(z)_j|\big).$$

Then $h_{\mathbf{u}}(\cdot + h_{\underline{\mathbf{v}}}(\cdot)) \in H(P_R^d(0))$. If we additionally assume that

$$N_j > \sup_{z \in \mathcal{P}_R^d(0)} \big(|z_j| + |\underline{\mathbf{v}}^j|_z\big),$$

setting

$$(\mathbf{u} \circ^s \underline{\mathbf{v}})_\alpha := \sum_{\beta \in \mathbb{N}_0^d} \frac{1}{\beta!} (\mathbf{u}^{(\beta)} * \underline{\mathbf{v}}^{*\beta})_\alpha$$

it holds $\mathbf{u} \circ^s \underline{\mathbf{v}} \in \mathcal{P}_R^d(0)^*$ and

$$h_{\mathbf{u} \circ^s \underline{\mathbf{v}}}(z) = h_{\mathbf{u}}(z + h_{\underline{\mathbf{v}}}(z))$$

for all $z \in P_R^d(0)$. The superscript ^{s} is mnemonic for "shift".

Proof. The first claim follows since the composition of holomorphic functions is holomorphic (see Proposition 1.2.2. in Scheidemann (2005)). Next, observe that since for all $z \in P_R^d(0)$ we have $|z_j| + h_{|\underline{\mathbf{v}}^j|}(|z|) < N_j$, we get $h_{|\mathbf{u}|}(|z| + h_{|\underline{\mathbf{v}}|}(|z|)) < \infty$. Moreover, by Proposition 4.D.2(ii) for all $z \in P_R^d(0)$ we get

$$\begin{aligned}
h_{|\mathbf{u}|}(|z| + h_{|\underline{\mathbf{v}}|}(|z|)) &= \sum_{\beta \in \mathbb{N}_0^d} \frac{1}{\beta!} h_{|\mathbf{u}|}^{(\beta)}(|z|) h_{|\underline{\mathbf{v}}|}(|z|)^\beta \\
&= \sum_{\beta \in \mathbb{N}_0^d} \frac{1}{\beta!} \sum_{\alpha \in \mathbb{N}_0^d} (|\mathbf{u}|^{(\beta)} * |\underline{\mathbf{v}}|^{*\beta})_\alpha \frac{|z|^\alpha}{\alpha!} \\
&= \sum_{\alpha \in \mathbb{N}_0^d} \Big(\sum_{\beta \in \mathbb{N}_0^d} \frac{1}{\beta!} (|\mathbf{u}|^{(\beta)} * |\underline{\mathbf{v}}|^{*\beta})_\alpha \Big) \frac{|z|^\alpha}{\alpha!}.
\end{aligned}$$

Thus, $|\mathbf{u}| \circ^s |\underline{\mathbf{v}}| \in \mathcal{P}_R^d(0)^*$. Since $|(\mathbf{u} \circ^s \underline{\mathbf{v}})_\alpha| \leqslant (|\mathbf{u}| \circ^s |\underline{\mathbf{v}}|)_\alpha$ the claim follows. $\square$

4.2.2 Jump-diffusion processes

Let $S \subseteq \mathbb{R}^d$, $b : S \to \mathbb{R}^d$, $a : S \to \mathbb{S}^d_+$ measurable functions and $K(\cdot, d\xi)$ a transition kernel from S to $\mathbb{R}^d$ which satisfies $K(x, \{0\}) = 0$, and $\int_{\mathbb{R}^d} \|\xi\| \wedge \|\xi\|^2 K(x, d\xi) < \infty$, for all $x \in S$. Let $M(S; \mathbb{C})$ denote the set of measurable maps on S with values in $\mathbb{C}$, and consider the operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \to M(S; \mathbb{C})$ such that

$$\mathcal{A}f(x) = \nabla f(x)^\top b(x) + \frac{1}{2}\mathrm{Tr}(a(x)\nabla^2 f(x)) + \int_{\mathbb{R}^d} f(x + \xi) - f(x) - \nabla f(x)^\top \xi \; K(x, d\xi),$$

for each $x \in S$ and $f \in \mathcal{D}(\mathcal{A})$, where

$$\mathcal{D}(\mathcal{A}) := \{f \in C^2(\mathbb{R}^d; \mathbb{C}) \;\: : \forall x \in S \int_{\mathbb{R}^d} |f(x + \xi) - f(x) - \nabla f(x)^\top \xi| \; K(x, d\xi) < \infty\}.$$

In particular, observe that $\mathcal{D}(\mathcal{A})$ includes all bounded $f \in C^2(\mathbb{R}^d; \mathbb{C})$.

Fix then $T > 0$. We say that $X = (X_t)_{t \in [0,T]}$ is an S*-valued jump-diffusion* with characteristics

$$\left(\int_0^\cdot b(X_{s-})ds, \int_0^\cdot a(X_{s-})ds, K(X_{s-}, d\xi)ds \right)$$

if X is a special semimartingale on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ such that for all bounded $f \in C^2(\mathbb{R}^d; \mathbb{C})$ the process $N^f := (N^f_t)_{t\in[0,T]}$ given by

$$N^f_t := f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s)ds, \quad t \in [0, T] \tag{4.2.3}$$

defines a local martingale[2]. We refer to $\mathcal{A}$ as the *extended generator* of X. In this chapter, we stick to the truncation function $\chi(\xi) = \xi$, and for simplicity, we refer to the coefficients (b, a, K) as characteristics of the jump-diffusion X.

Next, we introduce a class of functions that will form a subset of the domain of the extended generator of the process under consideration and discuss some key properties of the transition kernels.

4.2.2.1 Convergent power series on a given set

Here we introduce the space of holomorphic functions defined on polydiscs containing a given subset $S \subseteq \mathbb{R}^d$, which will serve as the state space of a jump-diffusion process. For $i \in \{1, \ldots, d\}$ set

$$R_i(S) := \sup\{|x_i| \;\: : x \in S\} \qquad \text{and} \qquad R(S) := (R_1(S), \ldots, R_d(S)).$$

[2]See Theorem II.2.42 in Jacod and Shiryaev (1987).

Note that $P^d_{R(S)+\varepsilon}(0)$ includes S for each $\varepsilon > 0$, and define

$$\mathcal{O}(S) := \bigcup_{\varepsilon>0} \mathcal{O}_\varepsilon(S)$$

for

$$\mathcal{O}_\varepsilon(S) := \Big\{ f : S \to \mathbb{C} \colon f = h|_S \text{ for some } h \in H(P^d_{R(S)+\varepsilon}(0)) \Big\}.$$

For a function $f : A \to \mathbb{C}$ for some $S \subseteq A \subseteq \mathbb{C}^d$, we write $f \in \mathcal{O}(S)$ if $f|_S \in \mathcal{O}(S)$. Observe that given $f_1, f_2 \in \mathcal{O}(S)$ it holds $\lambda f_1 + \mu f_2 \in \mathcal{O}(S)$ for each $\lambda, \mu \in \mathbb{R}$, showing that $\mathcal{O}(S)$ is a linear space. Maps in $\mathcal{O}(S)$ have two important properties:

(i) they admit a power series representation: setting

$$\mathcal{S}^* := \{ \mathbf{u} \in \mathcal{P}^d_{R(S)}(0)^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)} \text{ for some } h \in H(P^d_{R(S)+\varepsilon}(0)) \text{ and } \varepsilon > 0 \}, \tag{4.2.4}$$

it holds that for all $f \in \mathcal{O}(S)$, $f = h_{\mathbf{u}}|_S$, for some $\mathbf{u} \in \mathcal{S}^*$ (see Proposition 4.D.2(ii)). Notice that the power series representation of each $f \in \mathcal{O}(S)$ might not be unique. For the purposes of this chapter, this is however not relevant. We only need such a representation to exist (see Section 4.3.3 and in particular Remark 4.3.21(ii)).

(ii) on S they coincide with the restriction of a complex-valued smooth function on $\mathbb{R}^d$, making them eligible to be elements of $\mathcal{D}(\mathcal{A})$. It is important to observe that if $\mathcal{A}$ is the extended generator of X, then $\mathcal{A}f|_S$ just depends on $f|_S$ for all $f \in \mathcal{D}(\mathcal{A})$. This can be proved using that for each $f \in \mathcal{D}(\mathcal{A})$ such that $f|_S \equiv 0$ the process $(f(X_t))_{t\geqslant 0}$ needs to have vanishing drift. For similar results in different contexts see for instance Theorem 2.8 in Cuchiero et al. (2018) or the discussion after Definition 2.3 in Larsson and Svaluto-Ferro (2020). This observation in particular implies that $K(x, (S-x)^c) = 0$, where $(S-x)^c$ denotes the complement of the set $S - x$.

4.2.2.2 Kernels extension

Similarly, as in the previous section, where we were interested in the restriction to S of holomorphic functions defined on polydiscs containing S, we are now interested in kernels with the same type of property.

For a transition kernel $K(x, d\xi)$ from S to $\mathbb{R}^d$ we write $K \in \mathcal{O}(S) := \bigcup_{\varepsilon>0} \mathcal{O}_\varepsilon(S)$ if for all $x \in S$, $K(x, d\xi) = K_\varepsilon(x, d\xi)$, for some transition kernel $K_\varepsilon(z, d\xi)$ from S_ε to $\mathbb{C}^d$ where S_ε denotes an open set in $\mathbb{C}^d$ such that $S \subseteq S_\varepsilon \subseteq P^d_{R(S)+\varepsilon}(0)$, for some $\varepsilon > 0$. Moreover, we require that for all $|\beta| \geqslant 2$

$$\int_{\mathbb{C}^d} \xi^\beta K_\varepsilon(\cdot, d\xi) \in \Big\{ f : S_\varepsilon \to \mathbb{C} \colon f = h|_{S_\varepsilon} \text{ for some } h \in H(P^d_{R(S)+\varepsilon}(0)) \Big\}. \tag{4.2.5}$$

Observe in particular that for each $K \in \mathcal{O}_\varepsilon(S)$ it holds

$$\int_{\mathbb{R}^d} \xi^\beta K(\cdot, d\xi) \in \mathcal{O}_\varepsilon(S),$$

for each $|\beta| \geqslant 2$. Also in this context we write $\mathcal{O}(S) := \bigcup_{\varepsilon>0} \mathcal{O}_\varepsilon(S)$.

4.3 Holomorphic jump-diffusions

Fix now $S \subseteq \mathbb{R}^d$ and $T > 0$.

Definition 4.3.1. Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with extended generator $\mathcal{A}$ and fix a linear subset $\mathcal{V} \subseteq \mathcal{O}(S)$. We say that X is an S-valued $\mathcal{V}$-*holomorphic jump-diffusion* if for each $f \in \mathcal{V}$ it holds $f \in \mathcal{D}(\mathcal{A})$, $\mathcal{A}f \in \mathcal{O}(S)$, and the process N^f introduced in equation (4.2.3) defines a local martingale.

Let $X = (X_t)_{t\in[0,T]}$ be an S-valued $\mathcal{V}$-holomorphic jump-diffusion, for some linear subset $\mathcal{V} \subseteq \mathcal{O}(S)$ and set

$$\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)},\ h \in \mathcal{V}\}, \tag{4.3.1}$$

for $\mathcal{S}^*$ introduced in equation (4.2.4). Following the discussion on the properties of the functions in $\mathcal{O}(S)$ in Section 4.2.2.1, one can see that there always exists a linear map $L : \mathcal{V}^* \to \mathcal{O}(S)$, such that for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S. \tag{4.3.2}$$

Observe that since by the identity theorem for holomorphic functions (see Proposition 4.D.2(iii)) the power series representation of the maps in $\mathcal{O}(S)$ is not unique, there might be more than one map L (corresponding to different representations of the functions $\mathcal{A}h$ on S, $h \in \mathcal{V}$) for which (4.3.2) holds true. However, we only need such a map to exist (see Section 4.3.3 and, in particular, Remark 4.3.21(ii)).

Remark 4.3.2. Recall the notion of polynomial jump-diffusions given in Definition 1 in Filipović and Larsson (2020) and note that the class of holomorphic jump-diffusions is a considerable extension of the class of polynomial jump-diffusions. It is also interesting to observe that for an S-valued $\mathcal{V}$-holomorphic jump-diffusion $X = (X_t)_{t\in[0,T}$, the infinite-dimensional process

$$\mathbb{X} := ((1, \mathbb{X}^1_t, \mathbb{X}^2_t, \ldots, \mathbb{X}^i_t, \ldots))_{t\in[0,T]}, \tag{4.3.3}$$

where, for $i \in \mathbb{N}$ and all $t \in [0, T]$,

$$\mathbb{X}^i_t := \Big(\frac{1}{\alpha^1!} X_t^{\alpha^1}, \ldots, \frac{1}{\alpha^{k_i}!} X_t^{\alpha^{k_i}}\Big) \in \mathbb{R}^{k_i}, \qquad \alpha^j \in \mathbb{N}^d_0 \text{ with } |\alpha^j| = i$$

where k_i is the number of multi-indices $\alpha \in \mathbb{N}_0^d$ such that $|\alpha| = i$, can be interpreted as a $\mathcal{V}^*$-polynomial jump-diffusion on the extended tensor algebra of $\mathbb{R}^d$, in the sense of Definition 3.17 in Cuchiero et al. (2023c), for $\mathcal{V}^*$ defined in (4.3.1).

Notice in particular that each $\mathbf{u} \in \mathcal{S}^*$ can be uniquely associated with some symmetric element in the set (3.2.1) introduced in Chapter 3 (see Section 1.1.3 for the precise definition of a symmetric element in the extended tensor algebra). Denoting such an element by $T(\mathbf{u})$, by Lemma 4.4(iv) in Cuchiero et al. (2023c), we get that for each $t \in [0,1]$ $h_{\mathbf{u}}(X_t)$ is equal to the pairing (in the sense of equation (3.2.2)) of $T(\mathbf{u})$ with the signature (see Section 1.4) of X at time t.

4.3.1 Characteristics of holomorphic jump-diffusions

In this section, we provide a discussion on some sufficient and necessary conditions for a jump-diffusion to be a holomorphic jump-diffusion. We do not elaborate on existence results here. We assume instead the existence of a jump-diffusion process on a certain state space S and study the conditions on its characteristics such that the corresponding process is a holomorphic one. We start with a simple result establishing necessary conditions.

Lemma 4.3.3. Consider a subset $\mathcal{V}$ such that

$$\{p|_S \text{ for some polynomial } p : \mathbb{R}^d \to \mathbb{R}\} \subseteq \mathcal{V} \subseteq \mathcal{O}(S) \tag{4.3.4}$$

and let $X = (X_t)_{t\in[0,T]}$ be an S-valued $\mathcal{V}$-holomorphic jump-diffusion. Then $\int_{\mathbb{R}^d} \|\xi\|^k K(x, d\xi) < \infty$ for all for all $k \geqslant 2$ and $x \in S$, and the characteristics (b, a, K) of X satisfy

$$\begin{aligned}
&b_i(\cdot) \in \mathcal{O}(S), \quad \text{for all } i \in \{1,\dots,d\}, &(4.3.5)\\
&a_{ij}(\cdot) + \int_{\mathbb{R}^d} \xi_i\xi_j K(\cdot, d\xi) \in \mathcal{O}(S), \quad \text{for all } i,j \in \{1,\dots,d\}^2,\\
&\int_{\mathbb{R}^d} \xi^\beta K(\cdot, d\xi) \in \mathcal{O}(S), \quad \text{for all } \beta \in \mathbb{N}_0^d,\ |\beta| \geqslant 3.
\end{aligned}$$

Proof. The proof follows the proof of Lemma 1 in Filipović and Larsson (2020) by applying $\mathcal{A}$ to all monomials. □

Remark 4.3.4. Each S-valued polynomial jump-diffusion is an S-valued $\mathcal{V}$-holomorphic jump-diffusion for $\mathcal{V} = \{p|_S \text{ for some polynomial } p : \mathbb{R}^d \to \mathbb{R}\}$. However, holomorphic jump-diffusions allow for more general drift and diffusion coefficients (e.g., higher order polynomials and holomorphic functions) as well as more general kernels. For example, excluding higher order polynomials from $\mathcal{V}$ would permit to include in the class of holomorphic jump-diffusions jump-diffusions with kernels that do not admit all moments (see Corollary 4.3.15 and the discussion thereafter). An extreme example in this sense is given by $\mathcal{V}$ consisting entirely of bounded holomorphic functions. This shows how holomorphic jump-diffusions substantially extend the class of polynomial jump-diffusions.

In the classical polynomial case, if the characteristics satisfiy the type of conditions expressed by equation (4.3.5), the corresponding jump-diffusion process is automatically polynomial (see Lemma 1 in Filipović and Larsson (2020)). For holomorphic jump-diffusions, this is not the case. Indeed, in addition to the holomorphic dependence of the characteristics on the state variables further assumptions have to be made. Before illustrating this in the following theorem, we introduce some notation that will be used throughout the chapter.

Notation 4.3.5. Throughout, given an S-valued jump-diffusion X with drift and diffusion (b, a) such that $b_i \in \mathcal{O}(S)$ and $a_{ij} \in \mathcal{O}(S)$, we let

$$\underline{\mathbf{b}} \in (\mathcal{S}^*)^d, \qquad \underline{\underline{\mathbf{a}}} \in (\mathcal{S}^*)^{d\times d}$$

denote some coefficients determining the power series representation of b and a, respectively. This in particular, implies that

$$b_i = h_{\underline{\mathbf{b}}^i}|_S \qquad \text{and} \qquad a_{ij} = h_{\underline{\underline{\mathbf{a}}}^{ij}}|_S.$$

For each $|\beta| \geqslant 2$ and transition kernel $K \in \mathcal{O}_\varepsilon(S)$ we then denote by $\mathbf{m}^\beta \in \mathcal{P}^d_{R(S)+\varepsilon}(0)^*$ the coefficients determining the power series representation of $\int_{\mathbb{C}^d} \xi^\beta K_\varepsilon(\cdot, d\xi)$ and thus satisfying

$$\int_{\mathbb{C}^d} \xi^\beta K_\varepsilon(\cdot, d\xi) = h_{\mathbf{m}^\beta}|_{S_\varepsilon}.$$

Similarly, for a transition kernel K with $\int_{\mathbb{R}^d} \xi^\beta K(\cdot, d\xi) \in \mathcal{O}(S)$, we let $\mathbf{m}^\beta \in \mathcal{S}^*$ be some coefficients such that $\int_{\mathbb{R}^d} \xi^\beta K(\cdot, d\xi) = h_{\mathbf{m}^\beta}|_S$.

Notice that by the assumptions in Section 4.2.2, for all $\alpha \in \mathbb{N}_0^d$ it holds $\underline{\underline{\mathbf{a}}}^{ij}_\alpha, \underline{\mathbf{b}}^i_\alpha, \mathbf{m}^\beta_\alpha \in \mathbb{R}$. The proof of the following theorem can be found in Appendix 4.B.1.

Theorem 4.3.6. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Fix $\varepsilon > 0$ and assume that $b_j, a_{ij}, K \in \mathcal{O}_\varepsilon(S)$. Suppose that the following conditions hold true.*

(i) for all $z \in S_\varepsilon$, $\int_{\|\xi\|>1} \exp(|\xi_1| + \cdots + |\xi_d|) K_\varepsilon(z, d\xi) < \infty$;

(ii) the real-valued map $\sum_{|\beta|\geqslant 2} \frac{1}{\beta!} |h_{\mathbf{m}^\beta}(\cdot)|$ is locally bounded on $P^d_{R(S)+\varepsilon}(0)$.

Fix $G = (G_1, \ldots, G_d)$ such that

$$G_j > \sup_{z\in S_\varepsilon} \sup_{\xi\in \mathrm{supp}(K_\varepsilon(z,\cdot))} |z_j| + |\xi_j|, \tag{4.3.6}$$

and set

$$\begin{aligned} &\mathcal{V} \subseteq \{h \in H(P^d_G(0)) \,:\, (h^{(\beta)})_{|\beta|\geqslant 2} \text{ is locally uniformly bounded on } P^d_G(0)\}, \\ &\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)},\ h \in \mathcal{V}\}. \end{aligned} \tag{4.3.7}$$

Then X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$

$$\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S,$$

where $L : \mathcal{V}^* \to \mathcal{S}^*$ is given by

$$L(\mathbf{u})_\alpha = \sum_{|\beta|=1} (\mathbf{u}^{(\beta)} * \underline{\mathbf{b}}^\beta)_\alpha + \sum_{|\beta|=2} \frac{1}{\beta!}(\mathbf{u}^{(\beta)} * (\underline{\underline{\mathbf{a}}}^\beta + \mathbf{m}^\beta))_\alpha + \sum_{|\beta|\geqslant 3} \frac{1}{\beta!}(\mathbf{u}^{(\beta)} * \mathbf{m}^\beta)_\alpha, \quad (4.3.8)$$

for all $\mathbf{u} \in \mathcal{V}^*$, $\alpha \in \mathbb{N}_0^d$.

Remark 4.3.7. (i) Observe that by Proposition 4.D.2(ii), for each $h \in H(P_G^d(0))$ with $(h^{(\beta)})_{|\beta|\geqslant 2}$ locally uniformly bounded on $P_G^d(0)$ it holds

$$|h(z)| = \Big| \sum_{|\beta|\geqslant 0} \frac{1}{\beta!} h^{(\beta)}(0) z^\beta \Big| \leqslant C \sum_{|\beta|\geqslant 0} \frac{1}{\beta!} |z|^\beta = C \exp(|z_1| + \cdots + |z_d|),$$

for some $C > 0$ and all $z \in P_{G-\delta}^d(0)$, with $\delta > 0$. This shows how condition (i) in Theorem 4.3.6 is strictly related to the defining property of $\mathcal{V}$. As one can imagine, this implies that one can relax the integrability condition on K_ε by choosing a more restrictive growth condition on the derivatives of the elements of $\mathcal{V}$. Viceversa, one can obtain the result for a larger set of functions $\mathcal{V}$ by imposing a stronger integrability condition in (i). This trade-off appears in many results of the chapter and is exploited for instance in Corollary 4.3.15 (see also Remark 4.3.16).

(ii) Suppose that $S_\varepsilon = P_{R(S)+\varepsilon}^d(0)$. By the monotone convergence theorem, it holds

$$\begin{aligned}\sum_{|\beta|\geqslant 2} \frac{1}{\beta!} |h_{\mathbf{m}^\beta}(z)| &\leqslant \sum_{|\beta|\geqslant 2} \frac{1}{\beta!} \int_{\mathbb{C}^d} |\xi|^\beta K_\varepsilon(z, d\xi) \\ &= \int_{\mathbb{C}^d} \exp(|\xi_1| + \cdots + |\xi_d|) - 1 - (|\xi_1| + \cdots + |\xi_d|) K_\varepsilon(z, d\xi).\end{aligned}$$

Using that $0 \leqslant \exp(y) - 1 - y \leqslant y^2 1_{\{y\leqslant 1\}} + \exp(y) 1_{\{y>1\}}$ for each $y \in \mathbb{R}_+$, by continuity of $\int_{\mathbb{C}^d} (|\xi_1|^2 + \cdots + |\xi_d|^2) K_\varepsilon(\cdot, d\xi)$ we can see that condition (ii) of Theorem 4.3.6 is automatically satisfied if the map $\int_{\|\xi\|>1} \exp(|\xi_1| + \cdots + |\xi_d|) K_\varepsilon(\cdot, d\xi)$ is locally bounded on S_ε.

(iii) Finally, we refer to Remark 4.D.3 for a discussion about the assumption in Equation (4.3.6).

(iv) Notice that any S-values jump-diffusion verifying the conditions of Theorem 4.B.1 is in particular a Sig-jump-diffusion in the sense of Definition 3.2.3. This follows by a simple adaptation of the proof of Lemma 3.3.9 to the present setting.

It is important to highlight that when $d = 1$, the properties of holomorphic functions in one variable allow us to considerably relax the conditions on the kernel K needed to deduce the result of Theorem 4.3.6. Indeed, as a consequence of the simplification of the Identity theorem for holomorphic functions (see Proposition 4.D.2(iii)), instead of checking the pointwise convergence of (4.B.2) for each $z \in S_\varepsilon$, here it suffices to check it only for each $x \in S$. We illustrate the precise statement in the following corollary, whose proof is analogous to the one of Theorem 4.3.6. For more details on the differences in this one-dimensional setting, we refer to Remark 4.B.1.

Corollary 4.3.8. *Let $S \subseteq \mathbb{R}$ be a subset with an accumulation point in $\mathbb{R}$ and $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Fix $\varepsilon > 0$ and assume that $b_j, a_{ij} \in \mathcal{O}_\varepsilon(S)$ and*

$$\int_{\mathbb{R}} \xi^\beta K(\cdot, d\xi) \in \mathcal{O}_\varepsilon(S), \quad \text{for all } \beta \geqslant 2. \tag{4.3.9}$$

Suppose furthermore that the following conditions hold true.

(i) for all $x \in S$, $\int_{|\xi|>1} \exp(|\xi|)K(x, d\xi) < \infty$;

(ii) the real-valued map $\sum_{\beta\geqslant 2} \frac{1}{\beta!}|h_{\mathbf{m}^\beta}(\cdot)|$ is locally bounded on $P^1_{R(S)+\varepsilon}(0)$.

Fix

$$G > \sup_{x\in S} \sup_{\xi\in \operatorname{supp}(K(x,\cdot))} |x| + |\xi| \tag{4.3.10}$$

and set $\mathcal{V}$, $\mathcal{V}^*$ as in equation (4.3.7). Then X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$

$$\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S,$$

where $L : \mathcal{V}^* \to \mathcal{S}^*$ is the operator defined in equation (4.3.8), whose explicit form for $d = 1$ reads

$$L(\mathbf{u})_\alpha = (\mathbf{u}^{(1)} * \mathbf{b})_\alpha + \frac{1}{2}(\mathbf{u}^{(2)} * (\mathbf{a} + \mathbf{m}^2))_\alpha + \sum_{\beta=3}^{\infty} \frac{1}{\beta!}(\mathbf{u}^{(\beta)} * \mathbf{m}^\beta)_\alpha. \tag{4.3.11}$$

for all $\mathbf{u} \in \mathcal{V}^*$, $\alpha \in \mathbb{N}_0$.

Next, we shall make stronger assumptions on the parametrization of the jump kernel K of an S-valued jump diffusion X. Indeed, we only consider kernels with holomorphic jump size, defined as follows.

Definition 4.3.9. The jump kernel K is said to have holomorphic jump size if it is of the form

$$K(\cdot, A) = \lambda(\cdot) \int_E 1_{A\setminus\{0\}}(j(\cdot, y))F(dy),$$

where,

(i) F is a non-negative measure on a measurable space E;

(ii) $\lambda \in \mathcal{O}(S)$ and its restriction to S is positive real-valued;

(iii) $j : S \times E \to \mathbb{R}^d$ such that $j = j_\varepsilon|_{S\times E}$ for some $\varepsilon > 0$ and some measurable function

$$j_\varepsilon : P^d_{R(S)+\varepsilon}(0) \times E \to \mathbb{C}^d$$

with $j_\varepsilon(\cdot, y) \in H(P^d_{R(S)+\varepsilon}(0), \mathbb{C}^d)$ for each $y \in E$.

Before developing the analysis of these kernels further, we introduce another notation that will be employed throughout the chapter.

Notation 4.3.10. Consider a jump kernel K with holomorphic jump size j and let j_ε be the corresponding extension as in Definition 4.3.9. For each $y \in E$ we let

$$\underline{\boldsymbol{j}}(y) \in (\mathcal{S}^*)^d,$$

denote the coefficients determining the power series representation of $j_\varepsilon(\cdot, y)$. This implies that

$$h_{\underline{\boldsymbol{j}}(y)}|_S = j(\cdot, y).$$

Similarly, we choose $\boldsymbol{\lambda} \in \mathcal{S}^*$ such that $h_{\boldsymbol{\lambda}}|_S = \lambda$.

Finally, for each $|\beta| \geqslant 2$ and $|\alpha| \geqslant 0$ recall that

$$\left(\int_E \underline{\boldsymbol{j}}(y)^{*\beta}\ F(dy)\right)_\alpha = \int_E (\underline{\boldsymbol{j}}(y)^{*\beta})_\alpha\ F(dy),$$

whenever the involved quantities are well defined.

Notice that for all $\alpha \in \mathbb{N}_0^d$, it holds $\underline{\boldsymbol{j}}(y)^i_\alpha$, $\boldsymbol{\lambda}_\alpha \in \mathbb{R}$. Moreover, the map $E \ni y \mapsto \underline{\boldsymbol{j}}(y)_\alpha$ is a measurable function as for all $y \in E$, $\underline{\boldsymbol{j}}(y)_\alpha = D^\alpha j_\varepsilon(z, y)|_{z=0}$ with j_ε measurable (see Proposition 4.D.2(iv)).

Under the assumption of holomorphic jump sizes, the hypothesis of Theorem 4.3.6 become more explicit. The obtained result is stated in the next corollary. Recall the notation introduced in (4.2.2).

Corollary 4.3.11. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Assume that $b_j, a_{ij} \in \mathcal{O}(S)$ and K is a kernel with holomorphic jump size. Suppose that for some $\varepsilon > 0$ the following conditions hold true.*

(i) for all $|\beta| \geqslant 2$ and $z \in P^d_{R(S)+\varepsilon}(0)$, $\int_E |\underline{\boldsymbol{j}}(y)^{*\beta}|_z\ F(dy) < \infty$;

(ii) the map $\int_{\|j(\cdot,y)\|>1} \exp(|j_{\varepsilon,1}(\cdot, y)| + \cdots + |j_{\varepsilon,d}(\cdot, y)|)\ F(dy)$ is locally bounded on $P^d_{R(S)+\varepsilon}(0)$.

Let $\mathcal{V}$ and $\mathcal{V}^*$ be defined as in equation (4.3.7). Then, $K \in \mathcal{O}(S)$ and X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion. Moreover, for all $\mathbf{u} \in \mathcal{V}^*$, $\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S$, where $L : \mathcal{V}^* \to \mathcal{S}^*$ is the operator introduced in equation (4.3.8), and for all $|\beta| \geqslant 2$,

$$\mathbf{m}^\beta = \boldsymbol{\lambda} * \int_E \underline{\boldsymbol{j}}(y)^{*\beta}\ F(dy). \tag{4.3.12}$$

Proof. Fix $\varepsilon > 0$ such that $\lambda, b_j, a_{ij} \in \mathcal{O}_\varepsilon(S)$, j_ε satisfies the conditions of Definition 4.3.9(iii), and the conditions (i) and (ii) of the corollary are satisfied. Let K_ε denote the transition kernel from $P^d_{R(S)+\varepsilon}(0)$ to $\mathbb{C}^d$ given by

$$K_\varepsilon(\cdot, A) = h_{\boldsymbol{\lambda}}(\cdot)\int_E 1_{A\setminus\{0\}}(j_\varepsilon(\cdot, y))F(dy).$$

Notice that by the dominated convergence theorem and condition (i), it holds $\int_E \underline{\boldsymbol{j}}(y)^{*\beta}\, F(dy) \in \mathcal{P}^d_{R(S)+\varepsilon}(0)^*$ and for all $z \in P^d_{R(S)+\varepsilon}(0)$

$$\int_{\mathbb{C}^d} \xi^\beta\, K(z, d\xi) = \lambda(z)\int_E j^\beta_\varepsilon(z, y)\, F(dy) = \sum_{\alpha\in\mathbb{N}^d_0}\Big(\boldsymbol{\lambda} * \int_E \underline{\boldsymbol{j}}(y)^{*\beta}\, F(dy)\Big)_\alpha \frac{z^\alpha}{\alpha!}.$$

As a consequence, $\big(\boldsymbol{\lambda} * \int_E \underline{\boldsymbol{j}}(y)^{*\beta}\, F(dy)\big) \in \mathcal{P}^d_{R(S)+\varepsilon}(0)^*$, implying that $K \in \mathcal{O}_\varepsilon(S)$, with $S_\varepsilon = P^d_{R(S)+\varepsilon}(0)$. By Remark 4.3.7(ii) condition (ii) of Theorem 4.3.6 follows by condition (ii) in the statement. Thus, all the hypotheses of Theorem 4.3.6 are verified, and the claim follows. □

Even though Corollary 4.3.11 gives explicit conditions for Theorem 4.3.6, direct verification of such conditions may prove to be rather complicated. Relying on powerful results from complex analysis it is however possible to obtain simpler sufficient conditions. Particularly useful for our purpose is Morera's theorem (see Proposition 4.D.2(vi) and Lemma 4.A.1 for the adaptation to our needs) providing sufficient conditions for the integral of an holomorphic map to be holomorphic. This property is very useful to show that the extended generator of an S-valued jump-diffusion maps a subset of $\mathcal{O}(S)$ into $\mathcal{O}(S)$. The corresponding result is provided in Theorem 4.3.12.

Recall that $\circ^s$ and integration of sequence-valued maps have been introduced in Section 4.2.1.4 and Section 4.2.1.2, respectively. In the following, we let $L^1(E, F)$ denote the space of L^1 maps on the measurable space E with non-negative measure F. The proof of the following theorem can be found in Section 4.B.2.

Theorem 4.3.12. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Assume that $b_j, a_{ij} \in \mathcal{O}(S)$, K is a kernel with holomorphic jump size, and let F be the corresponding non-negative measure on the measurable space E. Fix $G = (G_1, \ldots, G_d)$ for G_j such that for some $\varepsilon > 0$*

$$G_j > \sup_{z\in P^d_{R(S)+\varepsilon}(0)}\ \sup_{y\in\mathrm{supp}(F)} |z_j| + |\underline{\boldsymbol{j}}(y)^j|_z. \tag{4.3.13}$$

For each $h \in H(P^d_G(0))$, $z \in P^d_{R(S)+\varepsilon}(0)$, and $y \in E$ set

$$Jh(z, y) := \lambda(z)\big(h(z + j_\varepsilon(z, y)) - h(z) - \nabla h(z)^\top j_\varepsilon(z, y)\big)$$

and

$$\mathcal{D}(J) := \{h \in H(P^d_G(0)) \colon Jh(z,\cdot) \in L^1(E,F) \text{ for each } z \in P^d_{R(S)+\varepsilon}(0)\}. \tag{4.3.14}$$

Set

$$\begin{aligned}&\mathcal{V} \subseteq \{h \in \mathcal{D}(J) \colon \text{the map } P^d_{R(S)+\varepsilon} \ni z \to Jh(z,\cdot) \in L^1(E,F) \text{ is continuous}\},\\ &\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)},\ h \in \mathcal{V}\}.\end{aligned}$$

Then X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S,$$

where $L : \mathcal{V}^* \to \mathcal{S}^*$ is given by

$$\begin{aligned}L(\mathbf{u}) = &\sum_{|\beta|=1} \mathbf{u}^{(\beta)} * \underline{\mathbf{b}}^\beta + \sum_{|\beta|=2} \frac{1}{\beta!}\mathbf{u}^{(\beta)} * \underline{\underline{\mathbf{a}}}^\beta \\ &+ \boldsymbol{\lambda} * \int_E \mathbf{u} \circ^s \underline{\boldsymbol{j}}(y) - \mathbf{u} - \sum_{|\beta|=1} \mathbf{u}^{(\beta)} * \underline{\boldsymbol{j}}(y)^{*\beta}\ F(dy).\end{aligned} \tag{4.3.15}$$

Remark 4.3.13. Observe that the operator $L : \mathcal{V}^* \to \mathcal{S}^*$ defined in Equation (4.3.15) can be expressed in the following more explicit form when $d = 1$

$$L(\mathbf{u}) = \mathbf{u}^{(1)} * \mathbf{b} + \frac{1}{2}\mathbf{u}^{(2)} * \mathbf{a} + \boldsymbol{\lambda} * \int_E (\mathbf{u} \circ^s \mathbf{j}(y)) - \mathbf{u} - \mathbf{u}^{(1)} * \mathbf{j}(y)\ F(dy), \tag{4.3.16}$$

for all $\mathbf{u} \in \mathcal{V}^*$.

The next step consists in investigating the sets $\mathcal{V}$ satisfying the condition of Theorem 4.3.12. We start the analysis by considering the case of general jump diffusions with possible unbounded jump sizes. To specify the integrability properties of the kernels of such processes, we exploit the notion of a weight function.

Definition 4.3.14. Given a polydisc $P^d_G(0)$, for some $G > 0$, and a nondecreasing map

$$v : \mathbb{R}_+ \to (0,\infty),$$

which we call *weight function*, we denote by

$$H_v(P^d_G(0)) := \{h \in H(P^d_G(0)) \colon \|h\|_v < \infty\},$$

the corresponding weighted spaces of holomorphic functions on $P^d_G(0)$, where

$$\|h\|_v := \sup_{z \in P^d_G(0)} |h(z)| v(\|z\|)^{-1}.$$

The proof of the following corollary can be found in Section 4.B.3.

Corollary 4.3.15. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Assume that $b_j, a_{ij} \in \mathcal{O}(S)$, K is a kernel with holomorphic jump size, and let F be the corresponding non-negative measure on the measurable space E. Fix a weight function v, $\varepsilon > 0$, and suppose that for each $z \in P^d_{R(S)+\varepsilon}(0)$, there exists $\delta_z > 0$ such that*

$$\begin{aligned} &\int_E \sup_{w\in P^d_{\delta_z}(z)} \|j_\varepsilon(w,y)\| \wedge \|j_\varepsilon(w,y)\|^2 \; F(dy) < \infty; \\ &\int_E \sup_{w\in P^d_{\delta_z}(z)} 1_{\{\|j_\varepsilon(w,y)\|>1\}} v(\|w + j_\varepsilon(w,y)\|) \; F(dy) < \infty. \end{aligned} \tag{4.3.17}$$

Let $G \in (0,\infty]^d$ satisfy (4.3.13) and set

$$\mathcal{V} := H_v(P^d_G(0)) \qquad \text{and} \qquad \mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* : h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)},\ h \in \mathcal{V}\}$$

Then X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$, $\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S$, where $L : \mathcal{V}^* \to \mathcal{S}^*$ is the operator defined in equation (4.3.15).

Remark 4.3.16. Hereafter, we list some interesting examples of weight functions. Observe furthermore that condition (4.3.17) concerns integrability conditions of the extension of the jump kernel.

(i) **Weight functions with sublinear growth**: we say that a weight functions v has sublinear growth
$$v(t) \leqslant Ct$$
for some $C > 0$ and each $t > 1$. In such case, the second condition in (4.3.17) is implied by the first one.

(ii) **Rapidly increasing weight functions**: we say that a weight functions v is rapidly increasing if satisfies $\lim_{t\to\infty} t^k v(t)^{-1} = 0$ for each $k \geqslant 0$. In this case, one has that
$$\begin{aligned} &\{p|_S \text{ for some polynomial } p : \mathbb{R}^d \to \mathbb{R}\} \\ &\qquad \subseteq \{h \in H(P^d_G(0)) : \sup_{z\in P^d_G(0)} |h(z)| v(\|z\|)^{-1} < \infty\}, \end{aligned}$$
implying that the conditions of Lemma 4.3.3 are satisfied. If v is sub-multiplicative (i.e., $v(t_1 + t_2) \leqslant C' v(t_1) v(t_2)$), the dominated convergence theorem with the bound
$$\begin{aligned} &\sup_{w\in P^d_{\delta_z}(z)} \|j_\varepsilon(w,y)\|^{|\beta|} \\ &\quad \leqslant C \sup_{w\in P^d_{\delta_z}(z)} \Big(\|j_\varepsilon(w,y)\|^2 1_{\{\|j_\varepsilon(w,y)\|\leqslant 1\}} + 1_{\{\|j_\varepsilon(w,y)\|>1\}} v(\|w + j_\varepsilon(w,y)\|) \Big), \end{aligned}$$

and an application of Lemma 4.A.1 yields that the coefficients $\mathbf{m}^\beta$ determining the power series representation of $\int_{\mathbb{R}^d} \xi^\beta\, K_\varepsilon(\cdot, d\xi)$ are given by Equation (4.3.12), for all $|\beta| \geqslant 2$.

(iii) **Entire functions of finite order and type**: if S is given as an unbounded subset of $\mathbb{R}$, an interesting choice for the weight function v is

$$v(t) := \exp(\tau t^\eta)$$

for some $\tau, \eta \in (0, \infty)$. This allows to work with the concept of order and type of an entire function.

- An entire function $h \in H(\mathbb{C})$ is said to be of (finite) *order* $\eta \in \mathbb{R}$ if

$$\rho = \inf\{c > 0\colon |h(z)| < \exp(|z|^c) \text{ for sufficiently large } |z|\}. \tag{4.3.18}$$

- An entire function $h \in H(\mathbb{C})$ of finite order $\eta \in \mathbb{R}$ is then said to be of (finite) *type* $\tau \in \mathbb{R}$ if

$$\tau = \inf\{a > 0\colon |h(z)| < \exp(a|z|^\eta) \text{ for sufficiently large } |z|\}. \tag{4.3.19}$$

- An entire function is said of *exponential type* τ if its order $\eta < 1$, or $\eta = 1$ and its type τ is finite.

More generally, we say that an entire function is of *exponential type* if its order $\eta < 1$, or $\eta = 1$, and its type is finite.

In particular, setting $\mathcal{V} = H_v(\mathbb{C})$, we obtain that $\mathcal{V}$ contains every entire function of order strictly smaller than η and every entire function of order η and type strictly smaller than τ. The advantage of working with the class of entire functions of finite order is that conditions for characterizing the sequence determining their power series representation have been studied extensively (see e.g., Theorem 2 and Theorem 3 in Levin et al. (1996) and Proposition 4.D.4).

(iv) **Locally uniformly bounded jump size**: assume that for every $z \in P^d_{R(S)+\varepsilon}(0)$ there exists $\delta_z > 0$ such that

$$\sup_{w \in P^d_{\delta_z}(z),\ y \in \mathrm{supp}(F)} \|j(w, y)\| < \infty. \tag{4.3.20}$$

In this case the second condition in (4.3.17) is implied by the first one for each weight function v and the result of Corollary 4.3.15 holds for $\mathcal{V} = H(P^d_G(0))$. This in particular implies that X is an S-valued $H(P^d_G(0))$-holomorphic jump-diffusion. This is of particular interest when S is bounded and condition (4.3.20) is automatically satisfied. Also in this case by the dominated convergence theorem and Lemma 4.A.1 we can conclude that $\mathbf{m}^\beta$ satisfies (4.3.12), for all $|\beta| \geqslant 2$.

It is also interesting to notice that including extra conditions on the kernel K as in Corollary 4.3.15 permits to obtain the holomorphic property for a set of functions $\mathcal{V}$ considerably larger than the one defined in Theorem 4.3.6 (see in particular Equation (4.3.7)).

Consider for instance the case of a locally uniformly bounded jump size (iv), and notice that here the set $\mathcal{V}$ is strictly larger than the corresponding set defined in Equation (4.3.7). The same holds true in the above case (iii). Indeed, assume that $\eta, \tau > 1$ and consider the entire function

$$f(z) := \sum_{n=0}^{\infty} \frac{e^n n!}{n^n} \frac{z^n}{n!}, \quad z \in \mathbb{C}.$$

By Theorem 2 and Theorem 3 in Levin et al. (1996) f is of order and type 1, implying that $f \in \mathcal{V} = H_v(\mathbb{C})$. However, its derivatives in 0 are not uniformly bounded.

4.3.2 Examples of holomorphic jump-diffusions

This section is dedicated to provide an overview of different instances of holomorphic jump-diffusions. For simplicity, we consider the case $d = 1$. We already saw that classical polynomial processes are $\mathcal{V}$-holomorphic if $\mathcal{V}$ consists of polynomials. We show now that this is the case also for larger sets $\mathcal{V}$.

Example 4.3.17. Let $B = (B_t)_{t\in[0,T]}$ be a one-dimensional Brownian motion and $N(dy, dt)$ a Poisson random measure with compensator $F(dy) \times dt$ on $E \times [0, T]$, for some measurable space E. Consider the following stochastic differential equation:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t + \int_E \delta(X_{t-}, y)(N(dy, dt) - F(dy)dt), \qquad X_0 = x_0 \in \mathbb{R}. \tag{4.3.21}$$

Suppose that

$$b(x) := \mathbf{b}_0 + \mathbf{b}_1 x, \qquad \sigma(x) := \boldsymbol{\sigma}_0 + x\boldsymbol{\sigma}_1, \quad \text{and} \quad j(x, y) := \mathbf{j}_0(y) + x\mathbf{j}_1(y),$$

for some functions $\mathbf{j}_i : E \to \mathbb{R}$, such that $\int_E |\mathbf{j}_i(y)|^k F(dy) < \infty$ for all $k \geqslant 2$, $i = 0, 1$. Then, there exists a unique strong $\mathbb{R}$-valued solution $X = (X_t)_{t\in[0,T]}$ of the equation (4.3.21), which is in fact a polynomial jump-diffusion (see Example 2.6 in Filipović and Larsson (2020)). Furthermore, notice that setting

$$K(x, A) := \int_E 1_{A\setminus\{0\}}(j(x, y))F(dy),$$

it holds that K is a transition kernel with holomorphic jump size in the sense of Definition 4.3.9. Moreover, in some cases, X is also a holomorphic jump-diffusion, for other $\mathcal{V}$ different from polynomials.

(i) If $\sup_{y\in E} |\mathbf{j}_i(y)| < \infty$, for $i = 0, 1$, the conditions of Remark 4.3.16(iv) are satisfied. Let $\mathcal{V}$ and $\mathcal{V}^*$ be the sets defined in the same remark. Then X is a $\mathcal{V}$-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$, $\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_{\mathbb{R}}$, where L is given in equation (4.3.16) with

$$\mathbf{b} := (\mathbf{b}_0, \mathbf{b}_1, 0, \dots), \quad \mathbf{a} := (\boldsymbol{\sigma}_0^2, 2\boldsymbol{\sigma}_0\boldsymbol{\sigma}_1, 2\boldsymbol{\sigma}_1^2, 0, \dots) \quad \text{and} \quad \boldsymbol{\lambda} := (1, 0, 0, \dots).$$

(ii) Alternatively, if for every $z \in \mathbb{C}$ it holds

$$\int_E \sup_{w \in P^1_{\delta_z}(z)} 1_{\{|\mathbf{j}_0(y)+x\mathbf{j}_1(y)|>1\}} v(|\mathbf{j}_0(y) + w(1 + \mathbf{j}_1(y))|)F(dy) < \infty,$$

for some $\delta_z > 0$ and some weight function v, then the hypothesis of Corollary 4.3.15 are satisfied. Letting $\mathcal{V}$ and $\mathcal{V}^*$ be the sets defined in the same corollary, we get that the process X is $\mathcal{V}$-holomorphic.

Relying on existence results given in Theorem III.2.32 in Jacod and Shiryaev (1987), one could also go beyond linear coefficients and consider solutions of stochastic differential equations with entire coefficients.

Example 4.3.18. Let $B = (B_t)_{t\in[0,T]}$ be a one-dimensional Brownian motion. Consider the following stochastic differential equation:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \qquad X_0 = x_0 \in \mathbb{R}. \tag{4.3.22}$$

Assume that b and σ are bounded entire functions with bounded first derivatives. Then, equation (4.3.22) admits a unique solution $X = (X)_{t\in[0,T]}$ which is furthermore a holomorphic $\mathcal{V}$-holomorphic jump-diffusion, for $\mathcal{V}$ being the class of entire functions.

Alternatively, examples of holomorphic jump-diffusions can be provided by studying martingale problems. Indeed, recall that we defined holomorphic jump-diffusions as a solution of a martingale problem, i.e. as $X = (X_t)_{t\in[0,T]}$ such that the process N^f introduced in (4.2.3) is a local martingale for a set of test functions f. Given some coefficients a, b, λ and j, sufficient (and necessary) conditions for the existence of a solution to the corresponding martingale problem can be obtained by verifying the hypothesis of Theorem 4.5.4 in Ethier and Kurtz (1986) (modulo explosion). If the state space is $\mathbb{R}$, this reduces to check that the diffusion coefficient a is nonnegative. To guarantee that X does not explode, one can then resource to Theorem 4.3.8 in Ethier and Kurtz (1986), which translates in checking that $\mathcal{A}1 = 0$, when the state space is compact.

We here present some coefficients a, b, λ and j, for which the solution of the corresponding martingale problem is a holomorphic problem.

Example 4.3.19. Fix $S \subseteq \mathbb{R}$, $G > 0$ as in (4.3.10), set $b, a, \lambda \in \mathcal{O}(S)$, and let F be a non-negative measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We now analyze the form of the jump size j in the compensator of the jumps $K(x, A) := \int_{\mathbb{R}} 1_{A\setminus\{0\}}(j(x, y))F(dy)$ such that a jump-diffusion $X = (X_t)_{t\in[0,T]}$, given as solution to the martingale problem for the triplet (b, a, K), is a holomorphic jump-diffusion. Consider the following three specifications:

(i) Let $j : P^1_G(0) \times \mathbb{R} \to \mathbb{R}$ be such that for all $(z, y) \in P^1_G(0) \times \mathbb{R}$, $j(z, y) := \sum_{\alpha\in\mathbb{N}_0} \boldsymbol{j}_\alpha(y)\frac{z^\alpha}{\alpha!}$, for $\mathbf{j}_\alpha : \mathbb{R} \to \mathbb{R}$ measurable and bounded functions, for all $\alpha \in \mathbb{N}_0$. Set

$$s(y) := \sup_{\alpha\in\mathbb{N}_0} |\mathbf{j}_\alpha(y)|, \quad y \in \mathbb{R}.$$

Assume that $\sup_{y\in\mathbb{R}} s(y) < \infty$ and $\int_{\mathbb{R}} s(y)^2\ F(dy) < \infty$. Fix $z \in P^1_G(0)$, $\delta_z > 0$ and note that

$$\sup_{y\in\mathbb{R}} \sup_{w\in P^1_{\delta_z}(z)} |j(w,y)| \leqslant \sup_{\alpha\in\mathbb{N}_0} \sup_{y\in\mathbb{R}} |\mathbf{j}_\alpha(y)| \exp(r+|z|) < \infty,$$
$$\int_{\mathbb{R}} \sup_{w\in P^1_{\delta_z}(z)} |j(w,y)|^2\ F(dy) \leqslant \exp(2(r+|z|)) \int_{\mathbb{R}} s(y)^2\ F(dy) < \infty.$$

In particular, a polynomial jump size might be considered: $j(z,y) := \sum_{\alpha=0}^{N} \boldsymbol{j}_\alpha(y) \frac{z^\alpha}{\alpha!}$, where for all $\alpha \leqslant N$, $\mathbf{j}_\alpha : \mathbb{R} \to \mathbb{R}$ with $\sup_{y\in\mathbb{R}} |\mathbf{j}(y)_\alpha| < \infty$ and $\int_{\mathbb{R}} |\mathbf{j}_\alpha(y)|^2\ F(dy) < \infty$.

(ii) Let $f \in \mathcal{O}(S)$ and assume that F is a measure of bounded support $E := \operatorname{supp}(F)$ such that $\int_E |y|^2 F(dy) < \infty$. Set for all $(z,y) \in P^1_G(0) \times E$,

$$j(z,y) := f(z+y) = \sum_{\alpha\in\mathbb{N}_0} f^{(\alpha)}(y) \frac{z^\alpha}{\alpha!}.$$

Then for all $z \in P^1_G(0)$ and some $\delta_z > 0$,

$$\sup_{y\in E} \sup_{w\in P^1_{\delta_z}(z)} |j(w,y)| = \sup_{y\in E} \sup_{w\in P^1_{\delta_z}(z)} |f(w+y)| < \infty,$$
$$\int_E \sup_{w\in P^1_{\delta_z}(z)} |j(w,y)|^2\ F(dy) \leqslant \int_E \sup_{w\in P^1_{\delta_z}(z)} |f(w+y)|^2 F(dy) \leqslant \int_E C|y|^2 F(dy) < \infty,$$

for some $C > 0$.

(iii) Let $j : P^1_G(0) \times \mathbb{R} \to \mathbb{R}$ be such that for all $(z,y) \in P^1_G(0) \times \mathbb{R}$, $j(z,y) := \sum_{\alpha\in\mathbb{N}_0} \boldsymbol{j}_\alpha(y) \frac{z^\alpha}{\alpha!}$, where for all α, $\mathbf{j}_\alpha : \mathbb{R} \to \mathbb{R}$ with $|\mathbf{j}(y)_\alpha| < \ell(y)$ for some $\ell : \mathbb{R} \to \mathbb{R}_+$ satisfying $\ell(y) \geqslant |y|$ and $\int_{\mathbb{R}} \exp(C\ell(y))F(dy) < \infty$ for each $C > 0$. Notice that for all $z \in \mathbb{C}$ and some $\delta_z > 0$,

$$\int_{\mathbb{R}} \sup_{w\in P^1_{\delta_z}(z)} 1_{\{|j(w,y)|\leqslant 1\}} |j(w,y)|^2\ F(dy) \leqslant C \int_{\mathbb{R}} \ell(y)^2 F(dy) < \infty,$$
$$\int_{\mathbb{R}} \sup_{w\in P^1_{\delta_z}(z)} 1_{\{|j(w,y)|> 1\}} \exp(|j(w,y)|)\ F(dy) \leqslant \int_{\mathbb{R}} \exp(C\ell(y))F(dy) < \infty.$$

Note that the integrability condition on ℓ is guaranteed if $\int_{\mathbb{R}} \exp(\ell(y)^2)F(dy) < \infty$.

If the jump size is specified as in (i) or (ii), then the conditions of Remark 4.3.16(iv) are satisfied, and X is a $\mathcal{V}$-holomorphic jump-diffusion, for $\mathcal{V}$ being the set defined in the same remark. If instead j is given as in (iii), then the conditions of Corollary 4.3.15 hold true and X is $\mathcal{V}$-holomorphic, for $\mathcal{V}$ denoting the set of entire functions of exponential type (see Remark 4.3.16(iii) for the definition of entire function of exponential type).

4.3.3 The holomorphic formula

Fix $\mathcal{V} \subseteq \mathcal{O}(S)$ and let $X = (X_t)_{t\in[0,T]}$ be an S-valued $\mathcal{V}$-holomorphic jump-diffusion. Denote by $\mathcal{A}$ its extended generator and recall that by the properties of the functions in $\mathcal{O}(S)$, there exists a linear map $L : \mathcal{V}^* \to \mathcal{O}(S)$, such that for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S, \tag{4.3.23}$$

where $\mathcal{V}^*$ denotes the set of coefficients determining some power series representation on $P^d_{R(S)}(0)$ of the functions in $\mathcal{V}$. In this section, we rely on the *duality theory*, as presented in Chapter 4 of Ethier and Kurtz (1986) (see also Section 2.1 in Cuchiero et al. (2023c)), to compute expected values of $h_{\mathbf{u}}(X_t)$ for $t \in [0,T]$ and $h_{\mathbf{u}} \in \mathcal{V}$. In particular, we show that condition (4.3.23) is in fact the key property that allows recognizing a sequence valued-solution of the linear ODE

$$\partial_t \mathbf{c}(t) = L(\mathbf{c}(t)), \quad \mathbf{c}(0) = \mathbf{u}$$

as (one possible choice of) a dual process[3] of X. This then implies that computing expected values of holomorphic functions of holomorphic jump-diffusions reduces to solving an (infinite-dimensional) system of linear ODEs.

Recall from Section 4.2.1.2 that integrals of sequences are defined componentwise.

Theorem 4.3.20. *Set $X_0 = x_0 \in S$, fix $\mathbf{u} \in \mathcal{V}^*$, and suppose that the following conditions hold true.*

(i) The sequence-valued linear ODE

$$\mathbf{c}(t) = \mathbf{u} + \int_0^t L(\mathbf{c}(s))ds, \qquad t \in [0,T], \tag{4.3.24}$$

admits a $\mathcal{V}^*$-valued solution $(\mathbf{c}(t))_{t\in[0,T]}$ such that

$$h_{\mathbf{c}(t)}(x) = h_{\mathbf{u}}(x) + \int_0^t h_{L(\mathbf{c}(s))}(x)ds, \tag{4.3.25}$$

for all $x \in S$, $t \in [0,T]$.

(ii) The process $(N_t^{h_{\mathbf{c}(s)}})_{t\in[0,T]}$ given in equation (4.2.3) defines a true martingale for each $s \in [0,T]$.

(iii) $\int_0^T \int_0^T \mathbb{E}[|\mathcal{A}h_{\mathbf{c}(s)}(X_t)|]dsdt < \infty$.

Then for each $\mathbf{u} \in \mathcal{V}*$ it holds that

$$\mathbb{E}[h_{\mathbf{u}}(X_T)] = h_{\mathbf{c}(T)}(x_0). \tag{4.3.26}$$

[3]We refer to page 188 in Ethier and Kurtz (1986) for the precise notion of a dual process.

Proof. We verify the conditions of Lemma A.1 in Cuchiero et al. (2023c). Set

$$Y^1(s,t) := h_{\mathbf{c}(s)}(X_t), \qquad Y^2(s,t) := h_{L(\mathbf{c}(s))}(X_t).$$

By continuity of $Y^1(\cdot,t)(\omega)$ and measurability (on $[0,T]\times\Omega$) of $Y^1(s,\cdot)$ the two maps $Y^1, Y^2 : [0,T]\times[0,T]\times\Omega \longrightarrow \mathbb{C}$ are measurable functions. Next, observe that by the assumption (ii) it holds that for each $s,t \in [0,T]$, the process $(N_t^{h_{\mathbf{c}(s)}})_{t\in[0,T]}$, whose explicit form reads

$$h_{\mathbf{c}(s)}(X_t) - h_{\mathbf{c}(s)}(x_0) - \int_0^t \mathcal{A}h_{\mathbf{c}(s)}(X_u)du, \tag{4.3.27}$$

is a true martingale. Moreover, by assumption (i) it holds that for every $s,t \in [0,T]$,

$$h_{\mathbf{c}(s)}(X_t) - h_{\mathbf{c}(0)}(X_t) - \int_0^s \mathcal{A}h_{\mathbf{c}(u)}(X_t)du = 0. \tag{4.3.28}$$

Finally, taking expectation in both equations (4.3.27), (4.3.28) and by assumption (iii), all the hypotheses of the above-mentioned lemma are satisfied, and thus the claim follows. □

Remark 4.3.21. (i) An inspection of the proof shows that (4.3.24) is not necessary. Specifically, condition (i) can be replaced by the assumption of the existence of a $\mathcal{V}^*$-valued map $(\mathbf{c}(t))_{t\in[0,T]}$ such that

$$h_{\mathbf{c}(t)}(x) = h_{\mathbf{u}}(x) + \int_0^t h_{L(\mathbf{c}(s))}$$

for all $x \in S$, $t \in [0,T]$.

On the other hand if $\int_0^T |L(\mathbf{c}(s))|_x ds < \infty$ for all $x \in S$, $t \in [0,T]$ then condition (4.3.24) implies (4.3.25).

(ii) Recall that there might be more than one map L (corresponding to different representations of the functions $\mathcal{A}h$ on S for $h \in \mathcal{V}$) for which (4.3.23) holds true. However, this variability is not an issue for the objectives of this chapter. The duality approach requires only the existence of a dual process, which in this context then refers to the existence of a solution to some sequence-valued ODE that satisfies (4.3.23).

(iii) Notice that the claim of Theorem 4.3.20 for the $\mathcal{V}$-holomorphic jump-diffusion X coincides with the assertion of Theorem 3.21 in Cuchiero et al. (2023c) for the (infinite-dimensional) $\mathcal{V}^*$-polynomial process $\mathbb{X}$ introduced in Equation (4.3.3).

(iv) Let X be a S-valued polynomial jump-diffusion and recall from Remark 4.3.4 that X is a S-valued $\mathcal{V}$-holomorphic jump-diffusion, for $\mathcal{V} = \{p|_S$ for some polynomial $p : \mathbb{R}^d \to \mathbb{R}\}$. In this case, the holomorphic formula (4.3.26) coincides with the so-called moment formula in Theorem 1 in Filipović and Larsson (2020) and Theorem 2.7

in Cuchiero et al. (2012). Notice in particular that the sequence-valued linear ODE in (4.3.24) reduces to a finite-dimensional system of linear ODEs, and (i),(ii),(iii) are always satisfied (see also Section 4.2 in Cuchiero and Svaluto-Ferro (2021)).

(v) Observe that in principle there could exists two S-valued $\mathcal{V}$-holomorphic jump-diffusions X and Y sharing the same generator $\mathcal{A}$. If condition (i) of Theorem 4.3.20 is satisfied for some $(\mathbf{c}(t))_{t\in[0,T]}$ and conditions (ii) and (iii) are satisfied for both X and Y we can conclude that

$$\mathbb{E}[h_{\mathbf{u}}(X_T)] = \mathbb{E}[h_{\mathbf{u}}(Y_T)],$$

for each $\mathbf{u} \in \mathcal{V}^*$. If instead $\mathcal{V}$ is rich enough, e.g., it contains some exponential functions, uniqueness hold.

4.3.4 Sufficient conditions for the application of the holomorphic formula

This section is dedicated to the study of sufficient conditions for the application of Theorem 4.3.20. To start, we provide sufficient conditions for the existence of a $\mathcal{V}^*$-valued map $(\mathbf{c}(t))_{t\in[0,T]}$ such that condition (4.3.25) holds. Recall that by Remark 4.3.21 (i) this can substitute assumption (i) in Theorem 4.3.20. Suppose that X is a time-homogeneous Markov process with semigroup $(P_t)_{t\in[0,T]}$, that is

$$P_t f(x) = \mathbb{E}[f(X_t)|X_0 = x], \qquad x \in S,$$

for each measurable function $f : S \to \mathbb{R}$ such that $P_t|f|(x) < \infty$ for each $x \in S$. The next lemma is an adaptation of Lemma 2.6 in Cuchiero et al. (2012). Observe that the action of the corresponding extended generator (in the sense of Definition 2.3 in Cuchiero et al. (2012)) on functions $f \in \mathcal{V}$ corresponds here to $\mathcal{A}f$.

Lemma 4.3.22. Let $f \in \mathcal{V}$ and $\mathbf{u} \in \mathcal{V}^*$ be such that $f = h_{\mathbf{u}}|_S$. Suppose that:

(i) the process N^f introduced in (4.2.3) is a true martingale;

(ii) $\int_0^T P_s|\mathcal{A}f|(x)ds < \infty$ for all $x \in S$;

(iii) $P_t f \in \mathcal{V}$ for all $t \in [0,T]$;

(iv) $S \ni x \mapsto P_t\mathcal{A}f(x)$ is continuous.

Then for all $x \in S$, $P_t\mathcal{A}f(x) = \mathcal{A}P_t f(x)$ and each $\mathcal{V}^*$-valued map $(\mathbf{c}(t))_{t\in[0,T]}$ such that

$$P_t f(x) = h_{\mathbf{c}(t)}(x),$$

satisfies (4.3.25) with initial condition $\mathbf{u}$.

Proof. Since by assumption N^f is a true martingale and for each $x \in S$, $\int_0^T P_s|\mathcal{A}f|(x)ds < \infty$, by Fubini theorem, for each $t \in [0,T]$,

$$P_t f(x) - P_0 f(x) - \int_0^t P_s\mathcal{A}f(x)ds = \mathbb{E}[N_t^f - N_0^f|X_0 = x] = 0. \tag{4.3.29}$$

Fix $s \in [0,T]$ and notice that by definition of $\mathcal{V}$-holomorphic jump-diffusion, since $P_s f \in \mathcal{V}$, the process $N^{P_s f}$ given by

$$N_t^{P_s f} = P_s f(X_t) - P_s f(X_0) - \int_0^t \mathcal{A} P_s f(X_r) dr, \qquad t \in [0,T],$$

is a local martingale. Next, we show that the process

$$P_s f(X_t) - P_s f(X_0) - \int_0^t P_s \mathcal{A} f(X_r) dr, \qquad t \in [0,T], \tag{4.3.30}$$

is a true martingale. Since for each $t \in [0,T]$ and each $x \in S$ $\int_0^t P_s|\mathcal{A}f|(x)ds < \infty$, and N^f is a true martingale, $P_t|f|(x) < \infty$ and thus $\mathbb{E}[|P_s f(X_t)|] \leqslant P_{s+t}|f|(x) < \infty$ and

$$\int_0^t \mathbb{E}[|P_s \mathcal{A} f(X_r)|] dr \leqslant \int_0^t P_{s+r}|\mathcal{A}f|(x) dr < \infty,$$

showing that (4.3.30) is integrable. By taking conditional expectation, we thus obtain

$$\begin{aligned}
&\mathbb{E}[P_s f(X_t) - P_s f(X_u) - \int_u^t P_s \mathcal{A} f(X_r) dr | \mathcal{F}_u] \\
&\quad = \mathbb{E}[P_s f(X_{t-u}) - P_s f(X_0) - \int_u^t P_s \mathcal{A} f(X_{r-u}) dr | X_0 = x]|_{x=X_u} \\
&\quad = P_{s+t-u} f(X_u) - P_s f(X_u) - \int_u^t P_{s+r-u} \mathcal{A} f(X_u) dr \\
&\quad = P_{s+t-u} f(X_u) - P_s f(X_u) - \int_s^{s+t-u} P_r \mathcal{A} f(X_u) dr = 0,
\end{aligned}$$

and the claim follows. By uniqueness of the decomposition of a special semimartingale, it follows that $\int_0^t \mathcal{A} P_s f(X_r) dr = \int_0^t P_s \mathcal{A} f(X_r) dr$ and that $N^{P_s f}$ is given by (4.3.30) and is in fact a true martingale. This implies condition (ii) of Theorem 4.3.20. Finally, since $P_s f \in \mathcal{V}$ implies $\mathcal{A} P_s f \in \mathcal{O}(S)$, and by assumption $P_s \mathcal{A} f$ is continuous, we can conclude that for each $x \in S$,

$$\mathcal{A} P_s f(x) = P_s \mathcal{A} f(x).$$

Then, (4.3.25) follows by (4.3.29). □

Remark 4.3.23. An inspection of the proof of Lemma 4.3.22 shows that its assumptions imply condition (ii) of Theorem 4.3.20.

Furthermore, notice that since $\int_0^T \int_0^T P_t|\mathcal{A} P_s f|(x) ds dt \leqslant \int_0^T \int_0^T P_{t+s}|\mathcal{A} f|(x) ds dt$, if we additionally assume that

$$\int_0^T \int_0^T P_{t+s}|\mathcal{A} h_{\mathbf{u}}|(x) ds dt < \infty$$

then condition (iii) of the same theorem holds. This can be interesting in view of Remark 4.3.21.

Next, we specify some conditions for the assumptions (ii) and (iii) in Theorem 4.3.20 to be satisfied. Additionally to the assumption made at the beginning of the section, we suppose that for each $\mathbf{u} \in \mathcal{V}^*$, there exists a $\mathcal{V}^*$-valued solution of the linear ODE (4.3.24) which satisfies (4.3.25), with initial value $\mathbf{u}$, that we denote by $(\mathbf{c}(t))_{t\in[0,T]}$.

The first result pertains to holomorphic jump-diffusions whose extended generator $\mathcal{A}$ acts between (the restriction on S of) weighted spaces of holomorphic functions (see Definition 4.3.14).

Lemma 4.3.24. Assume that $\mathcal{V} \subseteq H_v(\mathbb{C}^d)$ and that $\mathcal{A}(\mathcal{V}) \subseteq H_w(\mathbb{C}^d)$, for some weight functions v and w. Suppose furthermore that one of the following conditions holds true for each $s \in [0,T]$.

(i) $\int_0^T \|h_{L(\mathbf{c}(s))}\|_w ds < \infty$ and $\mathbb{E}[\sup_{t\leqslant T} v(\|X_t\|)]$, $\mathbb{E}[\sup_{t\leqslant T} w(\|X_t\|)] < \infty$.

(ii) $\|h^p_{\mathbf{c}(s)}\|_v$, $\|h^p_{L(\mathbf{c}(s))}\|_w$, $\int_0^T \|h_{L(\mathbf{c}(s))}\|_w ds < \infty$ for $p > 1$, and $\mathbb{E}[v(\|X_T\|)]$, $\int_0^T \mathbb{E}[w(\|X_t\|)]dt < \infty$.

Then[4] conditions (ii) and (iii) of Theorem 4.3.20 are satisfied.

Proof. Since for each $s \in [0,T]$, $\mathbf{c}(s) \in \mathcal{V}^*$, and X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion, it holds that the process $(N_t^{h_{\mathbf{c}(s)}})_{t\in[0,T]}$ given in equation (4.2.3) defines a local martingale for every $s \in [0,T]$. Consider the first set of assumptions and note that since

$$\begin{aligned}\mathbb{E}[\sup_{t\leqslant T} |N_t^{h_{\mathbf{c}(s)}}|] &\leqslant 2\mathbb{E}[\sup_{t\leqslant T} |h_{\mathbf{c}(s)}(X_t)|] + T\mathbb{E}[\sup_{t\leqslant T} |\mathcal{A}h_{\mathbf{c}(s)}(X_t)|] \\ &\leqslant \|h_{\mathbf{c}(s)}\|_v \mathbb{E}[\sup_{t\leqslant T} v(\|X_t\|)] + T\|h_{L(\mathbf{c}(s))}\|_w \mathbb{E}[\sup_{t\leqslant T} w(\|X_t\|)] < \infty,\end{aligned}$$

we can conclude that $(N_t^{h_{\mathbf{c}(s)}})_{t\in[0,T]}$ is a true martingale and thus condition (ii) of Theorem 4.3.20. Similarly, since

$$\int_0^T \int_0^T \mathbb{E}[|\mathcal{A}h_{\mathbf{c}(s)}(X_t)|]dsdt \leqslant T\int_0^T \|h_{L(\mathbf{c}(s))}\|_w ds\ \mathbb{E}[\sup_{t\leqslant T} w(\|X_t\|)] < \infty,$$

we can conclude that condition (iii) of the same theorem holds too.

With the second set of assumptions by Doob's inequality, we have that

$$\begin{aligned}\mathbb{E}[\sup_{t\leqslant T} |N_t^{h_{\mathbf{c}(s)}}|^p] &\leqslant C\mathbb{E}[|N_T^{h_{\mathbf{c}(s)}}|^p] \leqslant 2C'\mathbb{E}[|h_{\mathbf{c}(s)}(X_T)|^p] + C'\int_0^T \mathbb{E}[|\mathcal{A}h_{\mathbf{c}(s)}(X_t)|^p]dt \\ &\leqslant 2C'\|h^p_{\mathbf{c}(s)}\|_v \mathbb{E}[v(\|X_T\|)] + C'\|h^p_{L(\mathbf{c}(s))}\|_w \int_0^T \mathbb{E}[w(\|X_t\|)]dt < \infty,\end{aligned}$$

[4]An inspection on the proof shows that the supremum over $\mathbb{C}^d$ defining $\|\cdot\|_v$ can be replaced by a supremum over $\mathbb{R}^d$.

proving condition (ii) and

$$\int_0^T \int_0^T \mathbb{E}[|\mathcal{A}h_{\mathbf{c}(s)}(X_t)|]dsdt \leqslant \int_0^T \int_0^T \|h_{L(\mathbf{c}(s))}\|_w ds \; \mathbb{E}[w(\|X_t\|)]dt < \infty,$$

proving condition (iii) and concluding the proof. □

In the next result, we use a Gronwall-type argument to deduce some integrability conditions needed for proving (ii) and (iii) of Theorem 4.3.20. A similar argument is used in the classical case (see Theorem 2.10 in Cuchiero et al. (2012)) to prove finiteness of moments of polynomial jump-diffusions and in the infinite dimensional setting (see Definition 3.18 and Lemma 3.19 in Cuchiero and Svaluto-Ferro (2021)) for similar purposes.

Definition 4.3.25. Let $\mathcal{A}$ be the extended generator of an S-valued jump-diffusion, and fix $g : \mathbb{R}^d \to \mathbb{R}_+$, with $g \in \mathcal{D}(\mathcal{A})$. We say that $\mathcal{A}$ is *g-cyclical* if $|\mathcal{A}g(x)|g(x)^{-1} < \infty$ for all $x \in S$.

Remark 4.3.26. (i) Notice that from the polynomial property of the extended generator $\mathcal{A}$ of a polynomial jump-diffusion, one can deduce that $\mathcal{A}$ is g-cyclical for $g(x) := 1 + \|x\|^{2k}$, $x \in \mathbb{R}$, for every $k \in \mathbb{N}$. Observe however that, since in the setting of holomorphic jump-diffusions we deal with the larger class of convergent power series (and not only with polynomials of finite degree), the same cyclical argument does not generally follow directly.

(ii) Following the same reasoning in the proof of Lemma 3.19 in Cuchiero and Svaluto-Ferro (2021) (see also Theorem 2.10 in Cuchiero et al. (2012)), we get that if the extended generator of an S-valued jump-diffusion $X = (X_t)_{t\in[0,T]}$ is g-cyclical, then

$$\mathbb{E}[g(X_t)] \leqslant g(x_0)\exp(Ct), \tag{4.3.31}$$

for all $t \in [0,T]$. Thus, the map g can play the role of a weight function to deduce sufficient conditions for (ii) and (iii) in Theorem 4.3.20. To simplify the notation, for a map f we set

$$\|f\|_g := \sup_{x\in S} |f(x)|g(x)^{-1}. \tag{4.3.32}$$

Note that contrary to the setting in Definition 4.3.14, in (4.3.32), we consider the supremum only over the set S.

The proof of the next lemma follows the proof of Lemma 4.3.24 combined with (4.3.31).

Lemma 4.3.27. Assume that the operator $\mathcal{A}$ is g-cyclical, for some function $g : \mathbb{R}^d \to \mathbb{R}_+$. Fix $\mathbf{u} \in \mathcal{V}^*$ and let $(\mathbf{c}(t))_{t\in[0,T]}$ satisfy condition (i) of Theorem 4.3.20 with $\mathbf{c}(0) = \mathbf{u}$. If for each $s \in [0,T]$ it holds $\|h^p_{\mathbf{c}(s)}\|_g, \|h^p_{L(\mathbf{c}(s))}\|_g$, and $\int_0^T \|h_{L(\mathbf{c}(s))}\|_g ds < \infty$ for some $p > 1$, then conditions (ii) and (iii) of Theorem 4.3.20 are satisfied.

To conclude, we discuss the case of holomorphic jump-diffusions with values in a bounded state space S. Recall that $R(S)$ denotes the polyradius of the smallest closed polydisc, which includes S (see Section 4.2.2.1) and that for $\mathbf{u} \in \mathcal{V}^*$, $z \in \mathbb{C}^d$, the notation $|\mathbf{u}|_z$ has been introduced in Equation (4.2.2).

Lemma 4.3.28. Assume that S is a bounded set and that

$$\int_0^T |L(\mathbf{c}(s))|_{R(S)} ds < \infty. \tag{4.3.33}$$

Then condition (4.3.24) implies (4.3.25) and conditions (ii) and (iii) of Theorem 4.3.20 are satisfied.

Proof. By the definition of a holomorphic jump-diffusion, if $\mathbf{c}(s) \in \mathcal{V}^*$ then the process $(N_t^{h_{\mathbf{c}(s)}})_{t\in[0,T]}$ is a local martingale, for each $s \in [0,T]$. Since bounded local martingales are martingales, condition (ii) is always satisfied. Next, notice that

$$\int_0^T \int_0^T \mathbb{E}[|h_{L(\mathbf{c}(s))}(X_t)|] ds dt \leqslant T \int_0^T |L(\mathbf{c}(s))|_{R(S)} ds < \infty,$$

proving that condition (iii) is satisfied, too. □

4.3.5 Applications

In this section, we discuss some applications of the preceding theory. As a first example, we consider continuous-time Markov chains with a finite-state space. Next, we examine the set of Lévy processes, affine processes, and finally, we present some examples of jump diffusions that are not polynomial to which Theorem 4.3.20 applies.

4.3.5.1 Continuous time Markov chains with a finite state space

Let $S := \{x_1, \dots, x_N\} \subseteq \mathbb{R}^d$ and note that every map $f : S \to \mathbb{R}$ can be seen as the restriction to S of an entire map bounded on $\mathbb{R}$.

Let $X = (X_t)_{t\in[0,T]}$ be a continuous-time Markov chains with a finite-state space S with $X_0 = x_0 \in \mathbb{R}^d$ and generator

$$\mathcal{A}f(x_i) = \sum_{j=1}^N \lambda_{ij}(f(x_j) - f(x_i)).$$

Set $\mathcal{V} := \{h : S \to \mathbb{R}\}$ and $h_{\mathbf{u}_i}(x_j) = 1_{\{i=j\}}$ for some $\mathbf{u}_i \in \mathcal{S}^*$ and each i, j. Then, the operator L given by (4.3.11) reads

$$L(\mathbf{u}_k) = \sum_{i=1}^N \mathbf{u}_i \sum_{j=1}^N \lambda_{ij}(1_{\{k=j\}} - 1_{\{k=i\}}),$$

and for each $h \in \mathcal{V}$,

$$\mathbb{E}[h(X_t)] = ((h_{\mathbf{u}_1}(X_0), \dots, h_{\mathbf{u}_N}(X_0)) \exp(t\widetilde{L})(h(x_1), \dots, h(x_N))^\top,$$

where $\widetilde{L} \in \mathbb{R}^{N\times N}$ is given by $\widetilde{L}_{ik} = \sum_{j=1}^N \lambda_{ij}(1_{\{k=j\}} - 1_{\{k=i\}})$.

We illustrate now how the conditions of Theorem 4.3.20 can be verified. Observe first that for each k and i

$$\mathcal{A}h_{\mathbf{u}_k}(x_i) = \sum_{\ell=1}^{\infty} \frac{1}{\ell!} h_{\mathbf{u}_k}^{(\ell)}(x_i) \sum_{j=1}^{N} \lambda_{ij}(x_j - x_i)^k = h_{L(\mathbf{u}_k)}(x_i),$$

showing that X is $\mathcal{V}$-holomporphic. Since for each $h \in \mathcal{V}$ it holds that $h \in \operatorname{span}\{h_{\mathbf{u}_1}, \dots, h_{\mathbf{u}_N}\}$, the sequence-valued ODE given by (4.3.24) can be interpreted as an N-dimensional system of linear ODEs, and for $\mathbf{c}(0) = \mathbf{u}_k$ we can explicitly write

$$\mathbf{c}(t) = \sum_{j=1}^{N} \mathbf{u}_j \exp(t\widetilde{L})_{jk}.$$

Since conditions (ii) and (iii) of Theorem 4.3.20 follow by boundedness of the state space the claim follows.

Note that the same approach can be used if the extended generator is mapping the span of a finite number of holomorphic functions to itself, also for non-finite state spaces. This is the case for polynomial processes, where the bases is given e.g., by monomials. It is also possible to consider an infinite number of basis elements, but more conditions need to be verified. A possible approach to guarantee the existence of the solution of the ODE is given in Lemma 5.11 of Cuchiero et al. (2023c).

4.3.5.2 Lévy processes

It is well known that all Lévy processes whose extended generator is well-defined on the space of polynomials are polynomial jump-diffusions (see e.g., Lemma 1 in Filipović and Larsson (2020)). In the next proposition, we prove that they are also $\mathcal{V}$-holomorphic jump-diffusions for some suitable $\mathcal{V}$ (containing but not being limited to the polynomials) and demonstrate the validity of the holomorphic formula (4.3.26) for a large class of holomorphic functions. For simplicity, we deal with processes with values on (a subset of) $\mathbb{R}$. Notice however that most of the analysis could be extended to the more general multidimensional case.

The proof of the next theorem is given in Appendix 4.B.4. Recall the notion of the entire function of exponential type from Remark 4.3.16(iii) and that in this chapter, we stick to the "truncation function" $\chi(\xi) = \xi$

Theorem 4.3.29. *Fix $S \subseteq \mathbb{R}$ with and let $X = (X_t)_{t\in[0,T]}$ be an S-valued Lévy process, with characteristics (b, a, F). Assume that $\int_{|y|>1} \exp(\alpha|y|)F(dy) < \infty$ for some $\alpha > 0$.*

Set

$$\mathcal{V} \subseteq \{h \in H(\mathbb{C})\colon h \text{ is of exponential type } \tau < \alpha\},$$
$$\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^*\colon h_{\mathbf{u}} = h,\ h \in \mathcal{V}\}.$$

Then,

(i) X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion.

(ii) For all $h \in \mathcal{V}$, the process N^h given in equation (4.2.3) is a true martingale.

(iii) For all $\mathbf{u} \in \mathcal{V}^*$, there exists a $\mathcal{V}^*$-valued solution $(\mathbf{c}(t))_{t\in[0,T]}$ of (4.3.25) with initial condition $\mathbf{u}$ which satisfies condition (iii) of Theorem 4.3.20.

In particular, for each $\mathbf{u} \in \mathcal{V}^*$, the holomorphic formula holds true:

$$\mathbb{E}[h_{\mathbf{u}}(X_T)] = h_{\mathbf{c}(T)}(x_0).$$

Remark 4.3.30. (i) Set $v_\tau(t) := \exp(\tau t)$ and note that $\mathcal{V} = \bigcup_{\tau<\alpha} H_{v_\tau}(\mathbb{C})$. Observe moreover that imposing enough integrability on F, the result of Proposition 4.3.29 can be extended to all entire functions of finite (but arbitrarily large) exponential type. In particular, if $\int_{|y|>1} \exp(\alpha|y|)F(dy) < \infty$ for all $\alpha > 0$ then the results of the cited proposition hold for

$$\mathcal{V} \subseteq \{h \in H(\mathbb{C})\colon h \text{ is of finite exponential type }\} = \bigcup_{\tau>0} H_{v_\tau}(\mathbb{C}).$$

This is in particular the case if F has bounded support.

(ii) If $\int_{|y|>1} \exp(|y|)F(dy) < \infty$, with a slight adaptation, the proof of Proposition 4.3.29 can be adapted to other sets $\mathcal{V}$. An example is given by the set

$$\mathcal{V} := \{h \in H(\mathbb{C})\colon h = h_{\mathbf{u}} \text{ for some } \mathbf{u} \text{ such that } \sup_{n\in\mathbb{N}} |\mathbf{u}_n| < \infty\}.$$

Note that $\mathcal{V} = H_v(\mathbb{C})$ for $v(t) = \exp(t)$. In this case, the only missing piece is given by checking that $h', h'' \in \mathcal{V}$ for each $h \in \mathcal{V}$, which can be deduced from the representation of the derivatives introduced in Section 4.2.1.4.

Next we show that when the Lévy measure F has bounded support, the results of Proposition 4.3.15 hold for a larger set $\mathcal{V}$. In particular, instead of defining $\mathcal{V}$ through the growth rate of the complex extension of its elements, one can impose conditions on the growth rate of functions and their derivatives only on $\mathbb{R}$. The proof of the Corollary 4.3.31 can be found in Appendix 4.B.5. In the following, given a holomorphic function h, we write $h|_{\mathbb{R}}$ for its restriction on $\mathbb{R}$ and denote by $|h|_{\mathbb{R}}|$ its absolute value.

Corollary 4.3.31. *Fix $S \subseteq \mathbb{R}$ and let $X = (X_t)_{t\in[0,T]}$ be an S-valued Lévy process with characteristics (b, a, F). Assume that F has bounded support and set*

$$\mathcal{V} := \{h \in H(\mathbb{C})\colon |h|_{\mathbb{R}}|, |h'|_{\mathbb{R}}|, |h''|_{\mathbb{R}}| \leqslant C\exp(a|\cdot|) \text{ for some } a \in \mathbb{R}, C > 0\}$$
$$\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^*\colon h_{\mathbf{u}} = h,\ h \in \mathcal{V}\}.$$

The following conditions hold true.

(i) X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion.

(ii) The process N^h given in equation (4.2.3) is a true martingale for each $h \in \mathcal{V}$.

(iii) Fix $\mathbf{u} \in \mathcal{V}^*$. Assume that $\int_{\mathbb{R}} |h_{\mathbf{u}}(x)|dx < \infty$ or $\int_{\mathbb{R}} |h_{\mathbf{u}}(x)|^2 dx < \infty$ and the map $\hat{h}_{\mathbf{u}} : \mathbb{R} \to \mathbb{C}$ given by

$$\hat{h}_{\mathbf{u}}(u) := \int_{\mathbb{R}} h_{\mathbf{u}}(x) e^{-iux} dx \tag{4.3.34}$$

satisfies

$$\int_{\mathbb{R}} \exp(|ux|)|\hat{h}_{\mathbf{u}}(u)|du < \infty, \tag{4.3.35}$$

for all $x \in \mathbb{R}$. Then there exists a $\mathcal{V}^*$-valued solution $(\mathbf{c}(t))_{t\in[0,T]}$ of (4.3.25) with initial condition $\mathbf{u}$ which satisfies condition (ii) and (iii) of Theorem 4.3.20. In particular, the holomorphic formula holds true:

$$\mathbb{E}[h_{\mathbf{u}}(X_T)] = h_{\mathbf{c}(T)}(x_0).$$

Remark 4.3.32. (i) Observe that, modulo a normalization constant, the map defined in (4.3.34) represents the Fourier transform of $h_{\mathbf{u}}$.

(ii) Notice that contrary to the setting of Proposition 4.3.29, functions on of the form $h(z) = \exp(-z^{2n})$, $n \in \mathbb{N}$, $z \in \mathbb{C}$ are included in the set $\mathcal{V}$ introduced in Corollary 4.3.31.

(iii) If furthermore the characteristics of the Lévy process X satisfies $\int_{|\xi|\leqslant 1} |\xi|F(d\xi) < \infty$ and $a = 0$, Corollary 4.3.31 holds true by considering the larger set

$$\mathcal{V} := \{h \in H(\mathbb{C})\colon |h|_{\mathbb{R}}|, |h'|_{\mathbb{R}}|| \leqslant C \exp(a|\cdot|) \text{ for some } a \in \mathbb{R}, C > 0\}.$$

4.3.5.3 Affine processes

It is well known that under some integrability conditions, affine processes are polynomial jump-diffusions (see Corollary 3.3 in Filipović and Larsson (2020)). Here, we go further and show that affine processes are holomorphic jump-diffusions for a larger class of holomorphic functions than polynomials and show the validity of the holomorphic formula (4.3.26). Notably, the literature encompasses various formulations of affine processes, reflecting subtle differences in their definitions. In this context, we will always consider affine processes as specified in Definition 3.1 in Filipović and Larsson (2020). This is a relaxed definition compared to the definition of an affine process in Duffie et al. (2003), because it is directly given in terms of the point-wise action of the extended generator on exponential-affine functions. The proof of Proposition 4.3.33 can be found in Appendix 4.B.6.

Proposition 4.3.33. Let $X = (X_t)_{t\in[0,T]}$ be an affine process on $S \subseteq \mathbb{R}$. Suppose that for each $x \in S$, $K(x, d\xi) = \nu_0(d\xi) + x\nu_1(d\xi)$ for some signed measure ν_0, ν_1 such that both $|\nu_0|$ and $|\nu_1|$ have bounded support. Set

$$\begin{aligned}\mathcal{V} &:= \{h \in H(\mathbb{C})\colon h, h', h'' \text{ are bounded on } \mathbb{R}\},\\ \mathcal{V}^* &:= \{\mathbf{u} \in \mathcal{S}^*\colon h_{\mathbf{u}} = h,\ h \in \mathcal{V}\}.\end{aligned}$$

The following conditions hold true.

(i) X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion.

(ii) The process N^h given in equation (4.2.3) is a true martingale for each $h \in \mathcal{V}$.

(iii) Fix $\mathbf{u} \in \mathcal{V}^*$. Assume that $h_{\mathbf{u}}(x) = \int_{-\tau}^{\tau} \exp((\varepsilon+iu)x)g(u)du$ for some $g : [-\tau, \tau] \to \mathbb{R}$ with $\int_{-\tau}^{\tau} |g(u)|du < \infty$, $\varepsilon, \tau \in [0, \infty)$ and

$$\mathbb{E}[\exp((\varepsilon + iu)X_t)|X_0 = x] = \exp(\phi(t, \varepsilon + iu) + \psi(t, \varepsilon + iu)x),$$

for all $u \in [-\tau, \tau]$ and $t \in [0, T]$. Then there exists a $\mathcal{V}^*$-valued solution $(\mathbf{c}(t))_{t\in[0,T]}$ of (4.3.25) with initial condition $\mathbf{u}$ which satisfies condition (ii) and (iii) of Theorem 4.3.20.

In particular, the holomorphic formula holds true:

$$\mathbb{E}[h_{\mathbf{u}}(X_T)] = h_{\mathbf{c}(T)}(x_0).$$

Remark 4.3.34. (i) Observe that by the Paley–Wiener theorem (see Proposition 4.D.4(iv)) if h is of exponential type and

$$\int_{\mathbb{R}} |h(x)|^2 \exp(-2\varepsilon x)dx < \infty,$$

then

$$h(z) = \exp(\varepsilon z) \int_{-\tau}^{\tau} \exp(iuz)g(u)du,$$

for some $g \in L^2(-\tau, \tau)$, $\tau > 0$ and $z \in \mathbb{C}$.

Other classes of functions for which the proof of Theorem 4.3.20 works is given by polynomials and the Fourier basis. In this case we indeed already know from the classical theory that the semigroup maps such functions to entire functions.

(ii) As mentioned at the beginning of the section, for simplicity, here we are dealing with processes with values on (a subset of) $\mathbb{R}$. Notice, however, that the above analysis could be extended to affine processes on more general state spaces. A first example in this direction concerns affine processes on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$, for some $m, n \in \mathbb{N}_0$, as introduced in Duffie et al. (2003).

4.3.5.4 Beyond polynomial processes

Consider the jump-diffusion on $[0,1]$ whose coefficients (b,a,ν) with respect to the truncation function $\chi(\xi)=\xi$ are given by

$$b(x)=0, \qquad a(x)=x(1-x)(1-x/2), \qquad \nu(x,\cdot)=\frac{(1-x)(1-x/2)}{x}1_{\{x\neq 0\}}\delta_{-x}(\cdot),$$

where δ_{-x} denotes the Dirac measure in $(-x)$. Since its generator $\mathcal{A}$ is given by

$$\mathcal{A}f(x)=\frac{1}{2}a(x)f''(x)+\frac{(1-x)(1-x/2)}{x}\big(f(0)-f(x)+xf'(x)\big),$$

we can see that it is also a $[0,1]$-valued $\mathcal{V}$-holomorphic jump-diffusion for $\mathcal{V}=\mathcal{O}(S)$. Observe also that this is a continuous martingale that can perform a jump to the state 0, where it will then get absorbed. Setting $f_k(x):=(x/2)^k$ we can see that the generator of this process satisfies

$$\mathcal{A}f_k(x)=\frac{(k+2)(k-1)}{4}(f_{k-1}+2f_{k+1}-3f_k)1_{\{k\geqslant 2\}},$$

which implies that $\mathcal{A}$ is of the form described in Lemma 5.11 in Cuchiero et al. (2023c) for $\lambda_k:=(1/2)^k$, $\beta_{0j}=\beta_{1j}=0$,

$$\beta_{kj}=\frac{(k+2)(k-1)}{4}1_{\{j=k-1\}}+\Big(\frac{(k+2)(k-1)}{2}\Big)1_{\{j=k+1\}}, \qquad k\geqslant 2$$

and $\beta_k=0$. Following the examples in Section 6.2 of Cuchiero et al. (2023c) we can then apply Lemma 5.11 of the same chapter to conclude that there exists a solution $(\mathbf{c}(t))_{t\in[0,T]}$ of (4.3.24) satisfying (4.3.25) for each initial condition $\mathbf{u}\in\mathcal{V}^*$ where, setting $S:=[0,1]$,

$$\mathcal{V}^*:=\{\mathbf{u}\in\mathcal{S}^*:\mathbf{u}_i=\mu_i\frac{i!}{2^i}\text{ for some }\mu_i\geqslant 0\text{ such that }\sum_{i=0}^{\infty}\mu_i<\infty\}.$$

Since condition (ii) of Theorem 4.3.20 is always satisfied for bounded state spaces and condition (iii) of the same theorem can be verified using that $\mathcal{A}f_k(x)\geqslant 0$ for each $k\in\mathbb{N}_0$ and $x\in[0,1]$ we can conclude that Theorem 4.3.20 can be applied yielding

$$\mathbb{E}[h_{\mathbf{u}}(X_T)]=h_{\mathbf{c}(T)}(x_0). \tag{4.3.36}$$

for each $\mathbf{u}\in\mathcal{W}^*$. This includes in particular $\mathbf{u}=(1,u,u^2,\dots)$ corresponding to $h_{\mathbf{u}}(x)=\exp(ux)$ for each $u\in\mathbb{R}_+$.

An inspection of the proof of Lemma 5.11 also gives us a constructive description of the process $(\mathbf{c}(t))_{t\in[0,T]}$ appearing in (4.3.36). Assuming for simplicity that $\sum_{i=0}^{\infty}\mathbf{u}_i 2^i/i!=1$ we indeed get that

$$\mathbf{c}(T)_i:=\mathbb{P}(Z_T=i)i!/2^i,$$

where $(Z_t)_{t\in[0,T]}$ is the $\mathbb{N}_0$-valued process satisfying

$$\mathbb{P}(Z_0=i)=\mathbf{u}_i\frac{2^i}{i!}$$

and jumping from state k to state j after an $\exp(\beta_{kj})$-distributed random time.

4.4 Affine-holomorphic jump-diffusions

Fix $\mathcal{S} \subseteq \mathbb{R}^d$ and $T > 0$.

Definition 4.4.1. Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with extended generator $\mathcal{A}$ and fix a subset $\mathcal{V} \subseteq \mathcal{O}(S)$. We say that X is an S-valued $\mathcal{V}$-*affine-holomorphic jump-diffusion* if there exists an operator $\mathcal{R} : \mathcal{V} \to \mathcal{O}(S)$ such that for each $f \in \mathcal{V}$ it holds that $\exp(f) \in \mathcal{D}(\mathcal{A})$,

$$\mathcal{A}\exp(f) = \exp(f)\mathcal{R}(f),$$

and the process $N^{\exp(f)}$ introduced in equation (4.2.3) defines a local martingale.

Let $X = (X_t)_{t\in[0,T]}$ be a S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion, for some subset $\mathcal{V} \subseteq \mathcal{O}(S)$, and let $\mathcal{V}^*$ as in (4.3.1). Notice that by the properties of the functions in $\mathcal{O}(S)$ there always exists a map $R : \mathcal{V}^* \to \mathcal{O}(S)$ such that for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A}\exp(h_{\mathbf{u}}) = \exp(h_{\mathbf{u}})R(h_{\mathbf{u}})|_S.$$

Observe furthermore that also here, such a map might not be unique.

Remark 4.4.2. Note that every S-valued affine jump diffusion in the sense of Definition 2 in Filipović and Larsson (2020) is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion, for $\mathcal{V} = \{h : \mathbb{R}^d \to \mathbb{C} : \ h(x) := iu^\top x,\ u \in \mathbb{R}^d\}$. Moreover for an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion, the infinite-dimensional process defined in Equation (4.3.3) can be viewed as an $\mathcal{V}^*$-affine process in the extended tensor algebra of $\mathbb{R}^d$ in the sense of Definition 3.6 in Cuchiero et al. (2023c).

4.4.1 Characteristics of affine-holomorphic jump-diffusions

In this section, we establish some sufficient conditions for a jump-diffusion process to be affine-holomorphic. In alignment with the section on holomorphic jump-diffusions, we do not elaborate on the existence results here. Specifically, considering kernels with holomorphic jump sizes we investigate the criteria on its characteristics guaranteeing the affine-holomorphic property.

Recall that for $\mathbf{u}, \mathbf{v} \in \mathcal{S}^*$, we denote by $\mathbf{exp^*(u)} \in \mathcal{S}^*$ some coefficients determining the power series representation of $\exp(h_{\mathbf{u}})$ on S, namely $h_{\mathbf{exp^*(u)}}|_S := \exp(h_{\mathbf{u}})|_S$, by $\mathbf{u} \circ^s \mathbf{v}$ some coefficients determining the power series representation of $h_{\mathbf{u}}(\cdot + h_{\underline{\mathbf{v}}}(\cdot))$ on S, whenever the latter is well defined (see Section 4.2.1.4), and that $\mathbf{1} := (\mathbf{1}_\alpha)_{\alpha\in\mathbb{N}_0^d}$ denotes the sequence such that $\mathbf{1}_\alpha = 1$ if $\alpha = (0, \ldots, 0)$ and $\mathbf{1}_\alpha = 0$ otherwise. Recall also the following notation introduced in Theorem 4.3.12. We have

$$\begin{aligned} &Jh(z, y) := \lambda(z)(h(z + j_\varepsilon(z, y)) - h(z) - \nabla h(z)^\top j_\varepsilon(z, y)), \\ &\mathcal{D}(J) := \{h \in H(P^d_G(0)) \colon Jh(z, \cdot) \in L^1(E, F) \text{ for each } z \in P^d_{R(S)+\varepsilon}(0)\}, \end{aligned} \tag{4.4.1}$$

for each $h \in H(P^d_G(0))$, $z \in P^d_{R(S)+\varepsilon}(0)$, and $y \in E$. The proof of the next theorem follows the proof of Theorem 4.3.12.

Theorem 4.4.3. *Let $X = (X_t)_{t\in[0,T}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Assume that $b_j, a_{ij} \in \mathcal{O}(S)$, K is a kernel with holomorphic jump size, and let F be the corresponding non-negative measure on the measurable space E. Let G satisfy (4.3.13) and for each $h \in H(P^d_G(0))$, $z \in P^d_{R(S)+\varepsilon}(0)$, and $y \in E$ set $\mathcal{D}(J)$ and Jh as in (4.4.1). Set*

$$\begin{aligned}
&\mathcal{V} \subseteq \{h \in H(P^d_G(0)) \colon e^h \in \mathcal{D}(J) \text{ and } P^d_{R(S)+\varepsilon} \ni z \mapsto Je^h(z, \cdot) \in L^1(E, F) \text{ is continuous}\},\\
&\mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)},\ h \in \mathcal{V}\}.
\end{aligned}$$

Then X is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A} \exp(h_{\mathbf{u}}) = \exp(h_{\mathbf{u}}) R(h_{\mathbf{u}})|_S,$$

where $R : \mathcal{V}^* \to \mathcal{S}^*$ is given by

$$\begin{aligned}
R(\mathbf{u}) := &\sum_{|\beta|=1} \mathbf{u}^{(\beta)} * \underline{\mathbf{b}}^{\beta} + \sum_{|\beta_1|,|\beta_2|=1} \frac{1}{(\beta_1+\beta_2)!} \underline{\underline{\mathbf{a}}}^{\beta_1+\beta_2} * (\mathbf{u}^{(\beta_1)} * \mathbf{u}^{(\beta_2)} + \mathbf{u}^{(\beta_1+\beta_2)}) \qquad (4.4.2)\\
&+ \boldsymbol{\lambda} * \int_E \mathbf{exp}^* \left(\mathbf{u} \circ^s \underline{\mathbf{j}}(y) - \mathbf{u}\right) - \mathbf{1} - \sum_{|\beta|=1} \mathbf{u}^{(\beta)} * \underline{\boldsymbol{j}}(y)^{*\beta} \ F(dy).
\end{aligned}$$

In the next lemma, we show that with a slightly stronger assumption on $h \in \mathcal{V}$ we obtain a nicer representation of the operator R in (4.4.2). The proof of the next result is given in Appendix 4.C.1.

Lemma 4.4.4. If $h_{\mathbf{u}} \in \mathcal{V}$ satisfies $h_{\mathbf{u}} \in \mathcal{D}(J)$ and the map $P^d_{R(S)+\varepsilon} \ni z \mapsto Jh(z, \cdot) \in L^1(E, F)$ is continuous, we also get the representation

$$\begin{aligned}
R(\mathbf{u}) := L(\mathbf{u}) + &\sum_{|\beta_1|,|\beta_2|=1} \frac{1}{(\beta_1+\beta_2)!} \underline{\underline{\mathbf{a}}}^{\beta_1+\beta_2} * (\mathbf{u}^{(\beta_1)} * \mathbf{u}^{(\beta_2)}) \qquad (4.4.3)\\
&+ \boldsymbol{\lambda} * \int_E \mathbf{exp}^* \left(\mathbf{u} \circ^s \underline{\mathbf{j}}(y) - \mathbf{u}\right) - \mathbf{1} - \left(\mathbf{u} \circ^s \underline{\mathbf{j}}(y) - \mathbf{u}\right) \ F(dy),
\end{aligned}$$

for L as in equation (4.3.15).

As for the holomorphic case, several corollaries can then be deduced by Theorem 4.4.3. The proof of the following result is given in Appendix 4.C.2.

Corollary 4.4.5. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Fix a weight function v in the sense of Definition 4.3.14. Assume that $b_j, a_{ij} \in \mathcal{O}(S)$, K is a kernel with holomorphic jump size, and* *let F be the corresponding non-negative measure on the measurable space E. Fix $\varepsilon > 0$ and suppose that for each $z \in P^d_{R(S)+\varepsilon}(0)$, there exists $\delta_z > 0$ such that*

$$\int_E \sup_{w \in P^d_{\delta_z}(z)} \|j_\varepsilon(w,y)\| \wedge \|j_\varepsilon(w,y)\|^2 \; F(dy) < \infty, \tag{4.4.4}$$

$$\int_E \sup_{w \in P^d_{\delta_z}(z)} 1_{\{\|j_\varepsilon(w,y)\|>1\}} \exp(m \; v(\|w + j_\varepsilon(w,y)\|)) \; F(dy) < \infty, \tag{4.4.5}$$

for some $m \in \mathbb{R}_+$. Let $G \in (0,\infty]^d$ satisfy (4.3.13) and set

$$\mathcal{V} := H_v(P^d_G(0)) \qquad \text{and} \qquad \mathcal{V}^* := \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)}, \; h \in \mathcal{V}\}.$$

Then X is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A} \exp(h_{\mathbf{u}}) = \exp(h_{\mathbf{u}}) R(h_{\mathbf{u}})|_S,$$

where $R \colon \mathcal{V}^* \to \mathcal{S}^*$ is given by (4.4.3).

Also in this case similar considerations as in Remark 4.3.16 apply.

Remark 4.4.6. Suppose that jump sizes are locally uniformly bounded in the sense of (4.3.20). In this case condition (4.4.5) is implied by (4.4.4) for each weight function v and the result of Corollary 4.4.5 holds for $\mathcal{V} = H(P^d_G(0))$. This is of particular interest when S is bounded and condition (4.3.20) is automatically satisfied.

Finally, we analyze the class of jump-diffusion processes with jump sizes that do not depend on the current value of the process. The proof of the following proposition can be found in Appendix 4.C.3. Here $\Im(z)$ denotes the imaginary part of z.

Corollary 4.4.7. *Let $X = (X_t)_{t\in[0,T]}$ be an S-valued jump-diffusion with characteristics (b, a, K) and extended generator $\mathcal{A}$. Assume that $b, a, \lambda \in \mathcal{O}(S)$ and $K(x, d\xi) = \lambda(x)F(d\xi)$. Let $G \in (0,\infty]^d$ satisfy (4.3.13) and set*

$$\begin{aligned}
\mathcal{V} &\subseteq \{h \in H(P^d_G(0)) \colon |h(z)| \leqslant g_h(\Im(z)), \; \text{for } g_h : \mathbb{R} \to \mathbb{R}_+ \text{ continuous}, \; z \in \mathbb{C}\},\\
\mathcal{V}^* &:= \{\mathbf{u} \in \mathcal{S}^* \colon h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)}, \; h \in \mathcal{V}\}.
\end{aligned}$$

Then X is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion and for all $\mathbf{u} \in \mathcal{V}^*$,

$$\mathcal{A} \exp(h_{\mathbf{u}}) = \exp(h_{\mathbf{u}}) R(h_{\mathbf{u}})|_S,$$

where $R \colon \mathcal{V}^* \to \mathcal{S}^*$ is given by (4.4.3) for $\mathbf{j}(\xi) = \xi\mathbf{1}$.

Remark 4.4.8. (i) Observe that for $d = 1$ the operator R given by (4.4.3) reads

$$\begin{aligned}
R(\mathbf{u}) = L(\mathbf{u}) &+ \frac{1}{2}(\mathbf{a} * \mathbf{u}^{(1)} * \mathbf{u}^{(1)})_\alpha \\
&+ \boldsymbol{\lambda} * \int_{\mathbb{R}} \mathbf{exp}^* \left(\mathbf{u} \circ^s \mathbf{j}(\xi) - \mathbf{u}\right) - \mathbf{1} - \left(\mathbf{u} \circ^s \mathbf{j}(\xi) - \mathbf{u}\right) \; F(d\xi),
\end{aligned} \tag{4.4.6}$$

where L denotes the operator given in Equation (4.3.16).

(ii) Fix again $d = 1$. By Proposition 4.D.4, the set B_τ of all entire functions of order not exceeding 1 and of type not exceeding $\tau \in (0, \infty)$, which are bounded on $\mathbb{R}$ is included in the set $\mathcal{V}$ specified in Corollary 4.4.7.

More generally, we can extend the above result beyond B_τ and consider for instance functions of the form $h(z) := \exp(-z^2)$ for every $z = (x + iy) \in \mathbb{C}$, for which a direct computation shows that $|\exp(-z^2)| \leqslant \exp(y^2)$ for every z.

(iii) Observe that a slight adaptation of Corollary 4.4.7 applies in particular if for all $x \in S$

$$K(x, d\xi) = F_0(d\xi) + xF_1(d\xi),$$

for some signed measures $F_0(d\xi)$, $F_1(d\xi)$ on $\mathbb{R}^d$ for which $\int_{\mathbb{R}^d} \|\xi\| \wedge \|\xi\|^2 |F_i|(d\xi) < \infty$, for $i = 0, 1$. That is, K is the compensator of the jump measure of a S-valued affine jump-diffusion (see e.g., Filipović and Larsson (2020)).

Additionally, kernels of the form $K(x, d\xi) = \sum_{j=0}^N \lambda_j(x) F_j(d\xi)$, for some positive entire functions $\lambda_j(x)$ and positive measures $F_j(d\xi)$ and limits thereof could also be considered (see for instance Section 2.3 in Cuchiero et al. (2018)).

Remark 4.4.9. Note that often a $\mathcal{V}$-holomorphic jump-diffusion is also $\mathcal{V}$-affine-holomorphic. This is, however, not true in general, as one can see considering $\mathcal{V}$ consisting of linear maps only and $\mathcal{A}$ being the generator of a diffusion process.

Viceversa, observe that each $\mathcal{V}$-affine-holomorphic jump-diffusion is an $\mathcal{W}$-holomorphic jump-diffusion for $\mathcal{W} := \{\exp(h) \colon h \in \mathcal{V}\}$.

4.4.2 The affine-holomorphic transform formula

Fix $\mathcal{V} \subseteq \mathcal{O}(S)$ and let $X = (X_t)_{t\in[0,T]}$ be an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion. Denote by $\mathcal{A}$ its extended generator, and by $R : \mathcal{V}^* \to \mathcal{S}^*$ a linear operator such that $\mathcal{A} \exp(h_{\mathbf{u}}) = \exp(h_{\mathbf{u}}) h_{R(\mathbf{u})}|_S$, for all $\mathbf{u} \in \mathcal{V}^*$, where $\mathcal{V}^*$ denotes the set of coefficients determining some power series representation of the functions in $\mathcal{V}$.

Parallel to the discussion pertaining to holomorphic jump-diffusions in Section 4.3.3, we here exploit duality methods to compute the expected value of $\exp(h_{\mathbf{u}}(X_t))$, for $t \in [0, T]$ and $h_{\mathbf{u}} \in \mathcal{V}$.

Recall from Section 4.2.1.2 that integrals of sequences are defined componentwise.

Theorem 4.4.10. *Set $X_0 = x_0 \in S$, fix $\mathbf{u} \in \mathcal{V}^*$, and suppose that the following conditions hold true.*

(i) The sequence-valued ODE

$$\boldsymbol{\psi}(t) = \mathbf{u} + \int_0^t R(\boldsymbol{\psi}(s))ds, \qquad t \in [0, T], \tag{4.4.7}$$

admits a $\mathcal{V}^*$-valued solution $(\boldsymbol{\psi}(t))_{t\in[0,T]}$ such that

$$h_{\boldsymbol{\psi}(t)}(x) = h_{\mathbf{u}}(x) + \int_0^t h_{R(\boldsymbol{\psi}(s))}(x)ds, \tag{4.4.8}$$

for all $x \in S$, $t \in [0,T]$.

(ii) The process $N^{\exp(h_{\boldsymbol{\psi}(s)})}$ given in equation (4.2.3) defines a true martingale for each $s \in [0,T]$.

(iii) $\int_0^T \int_0^T \mathbb{E}[|\mathcal{A}\exp(h_{\boldsymbol{\psi}(s)})(X_t)|]dsdt < \infty$.

Then it holds that

$$\mathbb{E}[\exp(h_{\mathbf{u}}(X_T))] = \exp(h_{\boldsymbol{\psi}(T)})(x_0). \tag{4.4.9}$$

The proof of Theorem 4.4.10 follows the proof of Theorem 4.3.20. However, for the sake of completeness, we include it below.

Proof. We verify the conditions of Lemma A.1 in Cuchiero et al. (2023c). Set

$$Y^1(s,t) := \exp(h_{\boldsymbol{\psi}(s)})(X_t), \qquad Y^2(s,t) := \exp(h_{\boldsymbol{\psi}(s)})h_{R(\boldsymbol{\psi}(s))}(X_t),$$

By continuity of $Y^1(\cdot,t)(\omega)$ and measurability (on $[0,T]\times\Omega$) of $Y^1(s,\cdot)$ the two maps $Y^1, Y^2 : [0,T]\times[0,T]\times\Omega \longrightarrow \mathbb{C}$ are measurable functions. Next, observe that by the assumption (ii) it holds that for each $s,t \in [0,T]$, the process $(N_t^{\exp(h_{\mathbf{c}(s)})})_{t\in[0,T]}$, whose explicit form reads

$$\exp(h_{\boldsymbol{\psi}(s)})(X_t) - \exp(h_{\boldsymbol{\psi}(s)})(x_0) - \int_0^t \exp(h_{\boldsymbol{\psi}(s)})h_{R(\boldsymbol{\psi}(s))}(X_u)du, \tag{4.4.10}$$

is a true martingale. Moreover, by assumption (i) it holds that for every $s,t \in [0,T]$,

$$h_{\boldsymbol{\psi}(s)}(X_t) - h_{\boldsymbol{\psi}(0)}(X_t) - \int_0^s h_{R(\boldsymbol{\psi}(u))}(X_t)du = 0. \tag{4.4.11}$$

Finally, taking expectation in both equations (4.4.10), (4.4.11) and by assumption (iii), all the hypothesis of the above-mentioned lemma are satisfied, and thus the claim follows. □

Remark 4.4.11. (i) Also in this case, an inspection of the proof shows that (4.4.7) is not necessary. Specifically, condition (i) can be replaced by the assumption of the existence of a $\mathcal{V}^*$-valued map $(\boldsymbol{\psi}(t))_{t\in[0,T]}$ such that

$$h_{\boldsymbol{\psi}(t)}(x) = h_{\mathbf{u}}(x) + \int_0^t h_{L(\boldsymbol{\psi}(s))}(x)ds$$

for all $x \in S$, $t \in [0,T]$.

On the other hand if $\int_0^T |R(\boldsymbol{\psi}(s))|_x ds < \infty$ for all $x \in S$, $t \in [0,T]$ then condition (4.4.7) implies (4.4.8).

(ii) Recall that the operator R might not be unique. Again, this is not an issue, as explained in Remark 4.3.21(ii).

(iii) Notice that the claim of Theorem 4.4.10 for the $\mathcal{V}$-affine-holomorphic jump-diffusion X coincides with the assertion of Theorem 3.9 in Cuchiero et al. (2023c) for the (infinite-dimensional) $\mathcal{V}^*$-affine process $\mathbb{X}$ introduced in Equation (4.3.3) and discussed in Remark 4.4.2.

(iv) Let X be an S-valued affine jump-diffusion and recall from Remark 4.4.2 that X is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion, for $\mathcal{V} = \{h : \mathbb{R}^d \to \mathbb{C} : \ h(x) := iu^\top x, \ u \in \mathbb{R}^d\}$. In this case, formula (4.3.26) coincides with the so-called affine transform formula in Theorem 2 in Filipović and Larsson (2020).

The remaining part of the section is dedicated to the study of sufficient conditions for the application of Theorem 4.4.10. In particular, we specify some conditions for the assumptions (ii) and (iii) in Theorem 4.4.10 to be satisfied. Additionally to the assumption made at the beginning of the section, we suppose that for each $\mathbf{u} \in \mathcal{V}^*$, there exists a $\mathcal{V}^*$-valued solution of the linear ODE (4.4.7), with initial value $\mathbf{u}$, that we denote by $(\boldsymbol{\psi}(t))_{t\in[0,T]}$.

The next lemma pertains to affine-holomorphic jump-diffusions whose extended generator $\mathcal{A}$ acts between (the restriction on S of) weighted spaces of entire functions (see Definition 4.3.14). The proof of the following result follows the proof of Lemma 4.3.24.

Lemma 4.4.12. Assume that $\{\exp(h) \colon h \in \mathcal{V}\} \subseteq H_v(\mathbb{C}^d)$ and $\{\mathcal{A}(\exp(h)) \colon h \in \mathcal{V}\} \subseteq H_w(\mathbb{C}^d)$, for some weight functions v and w. Suppose moreover that $\mathbb{E}[\sup_{t\leqslant T} v(\|X_t\|)]$, $\mathbb{E}[\sup_{t\leqslant T} w(\|X_t\|)]$, and $\int_0^T \|h_{\mathbf{exp}^*(\boldsymbol{\psi}(s))*\mathcal{R}(\boldsymbol{\psi}(s))}\|_w ds$ are finite. Then conditions (ii) and (iii) of Theorem 4.4.10 are satisfied.

Finally, we discuss the case of affine-holomorphic jump-diffusions with values in a bounded state space S. An example of such processes is given by an affine process with compact state space (see Krühner and Larsson (2018)) for which Corollary 4.4.7 directly applies (see Remark 4.4.8(iii)). This result is the equivalent of Lemma 4.3.28 and since the proof is analogous we will omit it.

Lemma 4.4.13. Assume that S is a bounded set and that

$$\int_0^T |\,\mathbf{exp}^*(\boldsymbol{\psi}(s)) * R(\boldsymbol{\psi}(s))|_{R(S)} ds < \infty. \tag{4.4.12}$$

Then condition (4.4.7) implies (4.4.8) and conditions (ii) and (iii) of Theorem 4.4.10 are satisfied.

4.A A sufficient condition for interchanging summation and integration

The next lemma states that the continuity of the map (4.A.1) is a sufficient condition for interchanging summation and integration. Recall that integrals of (vectors of) sequence-valued maps are computed componentwise (see Section 4.2.1.2).

Lemma 4.A.1. Let F be a non-negative measure on a measurable space E, $R \in (0, \infty]^d$ and for some $\varepsilon > 0$, let $f : P^d_{R+\varepsilon}(0) \times E \to \mathbb{C}^d$ be a map such that

(i) for all $z \in P^d_{R+\varepsilon}(0)$, $f^j(z, \cdot) \in L^1(E, F)$ and the map

$$P^d_{R+\varepsilon}(0) \ni z \longmapsto f^j(z, \cdot) \in L^1(E, F) \tag{4.A.1}$$

is continuous for each j;

(ii) for all fixed $y \in E$ it holds $f(\cdot, y) \in H(P^d_{R+\varepsilon}(0), \mathbb{C}^d)$.

Let $\underline{\mathbf{f}}(y) \in (\mathcal{P}^d_R(0)^*)^d$ denote the coefficients determining the power series representation of $f(\cdot, y)$ on $P^d_R(0)$. Then

$$\int_E f(\cdot, y)\ F(dy) \in H(P^d_{R+\varepsilon}(0), \mathbb{C}^d) \tag{4.A.2}$$

and $\int_E \underline{\mathbf{f}}(y)\ F(dy)$ are the coefficients determining its power series representation.

Proof. Fix $i, j \in \{1, \dots, d\}$, $a \in P^d_{R+\varepsilon}(0)$, and $\delta > 0$ such that $P^d_\delta(a) \subset P^d_{R+\varepsilon}(0)$. Consider the functions $f^j_i : P^1_\delta(a_i) \to \mathbb{C}$ and $g_i : P^1_\delta(a_i) \to \mathbb{C}$ given by

$$\begin{aligned} f^j_i(z_i, y) &:= f^j(a_1, \dots, a_{i-1}, z_i, a_{i+1}, \dots, a_d, y), \\ g^j_i(z_i) &:= \int_E f^j_i(z_i, y) F(dy). \end{aligned}$$

By Hartogs' theorem (see Proposition 4.D.2(vii)) in order to show (4.A.2) it suffices to show that $g^j_i \in H(P^1_{R_i+\varepsilon}(0), \mathbb{C})$. By Morera's theorem (see Proposition 4.D.2(vi)) this follows by showing that g^j_i is continuous on $P^1_\delta(a_i)$ and

$$\int_\triangle g^j_i(z_i) dz_i = 0,$$

for every triangle $\triangle \subset P^1_\delta(a_i)$. Observe that condition (i) implies continuity of the map $\int_E f^j(\cdot, y) F(dy)$ on $P^d_{R+\varepsilon}(0)$ and thus of g^j_i on $P^1_\delta(a_i)$. Next, fix a triangle $\triangle \subset P^1_\delta(a_i)$ and note that by (i) we get

$$\int_\triangle \int_E |f^j_i(z_i, y)| F(dy) dz_i = \int_\triangle \|f^j_i(z_i, \cdot)\|_{L_1} dz_i < \infty.$$

Therefore, by Fubini's theorem $\int_\triangle \int_E f_i^j(z_i, y)F(dy)dz_i = \int_E \int_\triangle f_i^j(z_i, y)dz_i F(dy)$. Since for each $y \in E$ the map $f_i^j(\cdot, y)$ is holomorphic on $P_\delta^1(a_i)$ by (ii), by Goursat's theorem (see Proposition 4.D.2(v)) we can conclude that

$$\int_\triangle g_i^j(z_i)dz_i = \int_E \int_\triangle f_i^j(z_i, y)dz_i F(dy) = 0.$$

Next, by the Taylor expansion (see Proposition 4.D.2(ii)), for all $z \in P_R^d(0)$ it holds

$$\int_E f^j(z, y)\ F(dy) = \sum_{\alpha \in \mathbb{N}_0^d} D^\alpha \left(\int_E f^j(\cdot, y)\ F(dy) \right)(0)\ \frac{z^\alpha}{\alpha!}.$$

The claimed representation of the coefficients of $\int_E f(\cdot, y)F(dy)$ can thus be proven verifying that

$$D^\alpha \left(\int_E f^j(\cdot, y)\ F(dy) \right)(0) = \int_E \underline{\mathbf{f}}(y)_\alpha^j F(dy). \tag{4.A.3}$$

By the Cauchy integral formula (see Proposition 4.D.2(i)), fixing $0 < M < R$

$$D^\alpha \left(\int_E f^j(\cdot, y)\ F(dy) \right)(0) = \frac{\alpha!}{(2\pi i)^d} \int_{T_M^d(0)} \int_E f^j(z, y)\ F(dy) \frac{1}{z^{\alpha+1}}\ dz,$$

where $T_M^d(0)$ denotes the polytorus centered at $0 \in \mathbb{C}^d$ and with polyradius M in the sense of (4.2.1). Since by condition (i)

$$\int_{T_M^d(0)} \int_E |f^j(z, y)|\ F(dy) \frac{1}{|z|^{\alpha+1}}\ dz < \infty,$$

by Fubini's theorem we get

$$D^\alpha \left(\int_E f^j(\cdot, y)\ F(dy) \right)(0) = \int_E \frac{\alpha!}{(2\pi i)^d} \int_{T_M^d(0)} f^j(z, y)\ \frac{1}{z^{\alpha+1}}\ dz\ F(dy).$$

By condition (ii), a further application of the Cauchy integral formula yields

$$\frac{\alpha!}{(2\pi i)^d} \int_{T_M^d(0)} f^j(z, y)\ \frac{1}{z^{\alpha+1}} dz = D^\alpha f^j(\cdot, y)(0) = \underline{\mathbf{f}}(y)_\alpha^j$$

and (4.A.3) follows. □

Remark 4.A.2. An inspection of the proof shows that the continuity assumption in (i) could be replaced by requiring the map $\int_E f^j(\cdot, y)F(dy)$ to be continuous on $P_{R+\varepsilon}^d(0)$ for each j, and the map $\int_E \|f(\cdot, y)\| F(dy)$ to be in $L_{loc}^1(P_R^d(0))$.

4.B Proofs of Section 4.3

Before providing the proofs of Section 4.3.1, we recall that by Definition 4.3.1, to prove that X is an S-valued $\mathcal{V}$-holomorphic jump-diffusion we need to show that for each $f \in \mathcal{V}$ it holds $f \in \mathcal{D}(\mathcal{A})$, $\mathcal{A}f \in \mathcal{O}(S)$, and the process N^f introduced in Equation (4.2.3) is a local martingale.

4.B.1 Proof of Theorem 4.3.6

Fix $h \in \mathcal{V}$, $z \in S_\varepsilon$, and $\xi \in \operatorname{supp}(K_\varepsilon(z,\cdot))$. Observe that since $G_j > |z_j| + |\xi_j|$, there exists $R > 0$ such that $z + \xi \in P^d_R(z)$ and $\overline{P^d_R(z)} \subset P^d_G(0)$. Thus, by Proposition 4.D.2(ii) (see also Remark 4.D.3),

$$h(z+\xi) - h(z) - \nabla h(z)^\top \xi = \sum_{|\beta|\geqslant 2} \frac{1}{\beta!} h^{(\beta)}(z)\xi^\beta.$$

Furthermore, by definition of $\mathcal{V}$, we have that $C(z) := \sup_{|\beta|\geqslant 2} |h^{(\beta)}(z)| < \infty$ and hence

$$\frac{1}{\beta!}|h^{(\beta)}(z)\xi^\beta| \leqslant C(z)\frac{1}{\beta!}|\xi|^\beta.$$

Since $\sum_{|\beta|\geqslant 0} \frac{1}{\beta!}|\xi|^\beta = \exp(|\xi_1| + \ldots + |\xi_d|)$ we get the bound

$$\begin{aligned}
&\int_{\xi\in\mathbb{C}^d} |h(z+\xi) - h(z) - \nabla h(z)^\top \xi| K_\varepsilon(z,d\xi) \\
&\quad\leqslant \int_{\mathbb{C}^d} \sum_{|\beta|\geqslant 2} C(z)\frac{1}{\beta!}|\xi|^\beta K_\varepsilon(z,d\xi) \\
&\quad\leqslant C(z)\int_{\|\xi\|\leqslant 1} \|\xi\|^2 K_\varepsilon(z,d\xi) + C(z)\int_{\|\xi\|>1} \exp(|\xi_1| + \ldots + |\xi_d|) K_\varepsilon(z,d\xi), \qquad (4.B.1)
\end{aligned}$$

which is finite by condition (4.2.5) and condition (i). This in particular implies that $h \in \mathcal{D}(\mathcal{A})$. Next, by (4.B.1) and the dominated convergence theorem, we get that

$$\int_{\mathbb{C}^d} \sum_{|\beta|\geqslant 2} \frac{1}{\beta!} h^{(\beta)}(z)\xi^\beta K_\varepsilon(z,d\xi) = \sum_{|\beta|\geqslant 2} \frac{1}{\beta!} h^{(\beta)}(z) h_{\mathbf{m}^\beta}(z).$$

The map $\mathcal{A}h$ can thus be written as the following pointwise limit of functions in $\mathcal{O}(S)$:

$$\mathcal{A}h(x) = \lim_{n\to\infty}\Bigg(\sum_{|\beta|=1} h^{(\beta)}(x) h_{\underline{\mathbf{b}}^\beta}(x) + \sum_{|\beta|=2} \frac{1}{\beta!} h^{(\beta)}(x) h_{\underline{\underline{\mathbf{a}}}^\beta}(x) + \sum_{|\beta|=2}^{n} \frac{1}{\beta!} h^{(\beta)}(x) h_{\mathbf{m}^\beta}(x)\Bigg), \qquad (4.B.2)$$

for each $x \in S$. Moreover, the limit on the right-hand side converges also for each $z \in S_\varepsilon$. Since the sequence $(h^{(\beta)})_{|\beta|\geqslant 2}$ is locally uniformly bounded on $P^d_G(0)$ by assumption, the sequence in (4.B.2) is locally uniformly bounded on $P^d_{(R(S)+\varepsilon)\wedge G}(0)$ by condition (ii). Since $S_\varepsilon \subseteq P^d_{(R(S)+\varepsilon)\wedge G}(0)$, the Vitali-Porter theorem (see Proposition 4.D.2(iv)) yields then that the limit in (4.B.2) is well defined for each $z \in P^d_{(R(S)+\varepsilon)\wedge G}(0)$, belongs to $H(P^d_{(R(S)+\varepsilon)\wedge G}(0))$, and the convergence holds locally uniformly. This in particular implies that $\mathcal{A}h \in \mathcal{O}(S)$ and for all $z \in P^d_{R(S)}(0)$ by Proposition 4.D.2(ii) it holds

$$\mathcal{A}h(z) = \sum_{|\alpha|\geqslant 0} D^\alpha\left(\mathcal{A}h\right)(0)\frac{z^\alpha}{\alpha!}.$$

Let $\mathbf{u} \in \mathcal{V}^*$ be such that $h_{\mathbf{u}} = h|_{P^d_{R(S)}(0)}$. By Weierstrass' theorem (see Proposition 4.D.2(viii)) we get

$$\begin{aligned} D^\alpha(\mathcal{A}h_{\mathbf{u}})(0) &= \sum_{|\beta|=1} D^\alpha\big(h^{(\beta)}_{\mathbf{u}} h_{\underline{\mathbf{b}}^\beta}\big)(0) + \sum_{|\beta|=2}\frac{1}{\beta!}D^\alpha\big(h^{(\beta)}_{\mathbf{u}} h_{\underline{\underline{\mathbf{a}}}^\beta}\big)(0) + \sum_{|\beta|\geqslant 2}\frac{1}{\beta!}D^\alpha\big(h^{(\beta)}_{\mathbf{u}} h_{\mathbf{m}^\beta}\big)(0) \\ &= \sum_{|\beta|=1}(\mathbf{u}^{(\beta)} * \underline{\mathbf{b}}^\beta)_\alpha + \sum_{|\beta|=2}\frac{1}{\beta!}(\mathbf{u}^{(\beta)} * (\underline{\underline{\mathbf{a}}}^\beta + \mathbf{m}^\beta))_\alpha + \sum_{|\beta|\geqslant 3}\frac{1}{\beta!}(\mathbf{u}^{(\beta)} * \mathbf{m}^\beta)_\alpha, \end{aligned}$$

for all $|\alpha| \geqslant 0$. Since this expression corresponds to $L(\mathbf{u})_\alpha$ for L as in (4.3.8), this implies that $\mathcal{A}h_{\mathbf{u}} = h_{L(\mathbf{u})}|_S$. Finally, we need to show that the process N^h introduced in Equation (4.2.3) is a local martingale. By Theorem II.1.8 in Jacod and Shiryaev (1987) it then suffices to show that for all $h \in \mathcal{V}$

$$\int_0^T \int_{\mathbb{R}^d} |h(X_{s^-}+\xi) - h(X_{s^-}) - \nabla h(X_{s^-})^\top \xi|\ K(X_{s^-}, d\xi)\ ds < \infty\ a.s. \tag{4.B.3}$$

Notice first that the process $(X_s)_{s\in[0,T]}$ is a.s. càdlàg . Thus, it suffices to show that the right-hand side of (4.B.1) is bounded on compact subsets of S. Observe that since the sequence $(h^{(\beta)})_{|\beta|\geqslant 2}$ is locally uniformly bounded on $P^d_G(0)$ we know that $C(z)$ is bounded on compact subsets of S. By condition (4.2.5), the same holds for the first term of (4.B.1). For the second term instead, note that setting $F(\xi) := \prod_{i=1}^d(\exp(-\xi_i) + \exp(\xi_i))$, we get

$$\exp(|\xi_1| + \cdots + |\xi_d|)1_{\{\|\xi\|>1\}} \leqslant F(\xi)1_{\{\|\xi\|>1\}} \leqslant F(\xi) - 1 - \nabla F(0)^\top\xi + C'\|\xi\|^2,$$

for some $C' > 0$. Since $F \in \mathcal{V}$, applying the Vitali Porter theorem as before, we get that

$$\int_{\mathbb{R}^d} F(\xi) - 1 - \nabla F(0)^\top \xi\ K(x, d\xi)$$

lies in $\mathcal{O}(S)$. Using again that $\int_{\mathbb{R}^d}\|\xi\|^2 K(\cdot, d\xi) \in \mathcal{O}(S)$ the claim follows.

Remark 4.B.1. The proof of Corollary 4.3.8 is analogous to the one of Theorem 4.3.6. We stress that when $d = 1$, the Vitali-Porter theorem guarantees that a locally bounded sequence of holomorphic functions converges uniformly on compact subsets of some domain if pointwise convergence holds on a set containing an accumulation point in the considered domain (see Proposition 4.D.2(iv) for more details). For this reason, instead of checking the pointwise convergence of (4.B.2) for each $z \in S_\varepsilon$ it thus suffices to check it for each $x \in S$.

This explains why in this particular case, instead of assuming $K \in \mathcal{O}(S)$, we simply require that condition (4.3.9) is satisfied. Similarly, condition (i) in Theorem 4.3.6 is then replaced by the weaker one which concerns only the set S which is by assumption a set of accumulation points of $P^1_{R(S)+\varepsilon}(0)$.

4.B.2 Proof of Theorem 4.3.12

First observe that $\mathcal{V} \subseteq \mathcal{D}(J)$ directly implies that $\mathcal{V} \subseteq \mathcal{D}(\mathcal{A})$. Next, fix $h_{\mathbf{u}} \in \mathcal{V}$. Recall from Section 4.2.1.4 that the assumptions on G guarantee that $h_{\mathbf{u}}(\cdot + j_\varepsilon(\cdot, y)) \in H(P^d_{R(S)+\varepsilon}(0))$ and the corresponding coefficients are given by $\mathbf{u} \circ^s \underline{\mathbf{j}}(y)$. By Lemma 4.A.1 we thus get that the map $z \mapsto \int_E Jh_{\mathbf{u}}(z,y)F(dy)$ lies in $H(P^d_{R(S)+\varepsilon}(0))$ and the corresponding coefficients are given by

$$\boldsymbol{\lambda} * \int_E \mathbf{u} \circ^s \underline{\boldsymbol{j}}(y) - \mathbf{u} - \sum_{|\beta|=1} \mathbf{u}^{(\beta)} * \underline{\boldsymbol{j}}(y)^{*\beta}\ F(dy).$$

This implies that $\mathcal{A}(\mathcal{V}) \subseteq \mathcal{O}(S)$. As the continuity of the map $z \to Jh_{\mathbf{u}}(z,\cdot)$ implies the continuity of the map

$$z \mapsto \int_E |Jh_{\mathbf{u}}(z,y)|F(dy),$$

the claim follows as in the proof of Theorem 4.3.6.

4.B.3 Proof of Corollary 4.3.15

We verify the conditions of Theorem 4.3.12. Fix $h \in \mathcal{V}$ and note that by definition of G there is an $\eta > 0$ such that setting $R_z := \eta + \sup_{y \in \mathrm{supp}(F)} |j_\varepsilon(z,y)| \wedge 1$ we get

$$\overline{P^d_{R_z}(z)} \subseteq P^d_G(0)$$

for each $z \in P^d_{R(S)+\varepsilon}(0)$. Since $|j_\varepsilon(z,y)|1_{\{\|j_\varepsilon(z,y)\| \leqslant 1\}}$ is bounded away from R_z, we get by Proposition 4.D.2(ii)

$$|h(z+j_\varepsilon(z,y)) - h(z) - \nabla h(z)^\top j_\varepsilon(z,y)|1_{\{\|j_\varepsilon(z,y)\| \leqslant 1\}} \leqslant C\|f|_{T^d_{R_z}(z)}\|_\infty \|j_\varepsilon(z,y)\|^2 1_{\{\|j_\varepsilon(z,y)\| \leqslant 1\}},$$

for some C not depending on z and y. Observe also that by definition of $\mathcal{V}$ it holds $|h(w+\xi)| \leqslant C' v(\|w+\xi\|)$ and for $\|\xi\| > 1$ we get

$$|h(z+\xi) - h(z) - \nabla h(z)^\top \xi| \leqslant C'|v(\|z+\xi\|)| + (|h(z)| + \|\nabla h(z)\|)\|\xi\|,$$

for some $C' > 0$. By continuity of h and its derivatives we thus obtain

$$\sup_{w \in P^d_{\delta_z}(z)} |Jh(w,y)| \leqslant C'' \sup_{w \in P^d_{\delta_z}(z)} \Big(v(\|w + j_\varepsilon(w,y)\|) 1_{\{\|j_\varepsilon(w,y)\|>1\}} + \|j_\varepsilon(w,y)\| \wedge \|j_\varepsilon(w,y)\|^2 \Big),$$

for some $C'' > 0$. The claim now follows by the dominated convergence theorem. □

Remark 4.B.2. An inspection of the proof of Corollary 4.3.15 shows that the claim also follows by considering the slightly larger set of holomorphic functions given by

$$\{h \in H(P^d_G(0)) \colon \sup_{z \in P^d_{R(S)+\varepsilon}(0)} |h(z)| v(\|z\|)^{-1} < m\}.$$

The proof of Theorem 4.B.4 and Corollary 4.3.31 strongly relate to the Lévy-Kinchine formula for the moment generating function (Theorem 25.17 in Sato (2013)), which states that, under some integrability conditions, for each $|\tau| \leqslant \alpha$ it holds

$$P_t \exp(\tau(\cdot))(x) = \exp(\tau x + t\psi(\tau)), \tag{4.B.4}$$

for $\psi(\tau) = b\tau + \frac{1}{2}a\tau^2 + \int_E \exp(\tau y) - 1 - \tau y F(dy)$, and P_t denoting the semigroup $(P_t)_{t \in [0,T]}$ given by $P_t h(x) = \mathbb{E}[h(X_t + x)]$, for each $t \in [0,T]$, $x \in \mathbb{R}$ and measurable map h for which $\mathbb{E}[|h(X_t + x)|] < \infty$.

4.B.4 Proof of Theorem 4.3.29

Note that

$$\mathcal{V} = \{h \in H(\mathbb{C}) \colon |h(z)| \leqslant C \exp(\tau |z|) \text{ for } \tau < \alpha,\ C > 0\}. \tag{4.B.5}$$

(i): Since the jump size j_ε is constant in its first argument, condition (4.3.17) for $v(t) := \exp(\alpha t)$ is verified. The statement then follows by Corollary 4.3.15.

(ii): Observe that $\mathcal{A}$ is g-cyclical for $g(x) := \exp(\alpha x) + \exp(-\alpha x)$ in the sense of Definition 4.3.25. Moreover, by definition of $\mathcal{V}$ for each $h \in \mathcal{V}$ there is a $p > 1$ such that $\|h^p\|_g < \infty$. Next, by Proposition 4.D.4 we know that $h', h'' \in \mathcal{V}$. Noting that

$$|h(x+y) - h(x) - h'(x)y| \leqslant \sup_{|y| \leqslant 1} |h''(x+y)||y|^2 + (|h(x+y)| + |h(x)| + |h'(x)||y|) 1_{|y|>1},$$

we also get that $\mathcal{A}h \in \mathcal{V}$ and hence $\|(\mathcal{A}h)^p\|_g < \infty$ for some (possibly different) $p > 1$. Proceeding as in the proof of Lemma 4.3.24(ii) the claim follows.

(iii): Consider the semigroup $(P_t)_{t \in [0,T]}$ given by $P_t h(x) = \mathbb{E}[h(X_t + x)]$, for $h \in \mathcal{V}$, $t \in [0,T]$, $x \in \mathbb{R}$. Notice that it is well-defined since $P_t|h|(x) = \mathbb{E}[|h(X_t + x)|] < \infty$, by Theorem 25.17 in Sato (2013) and definition of $\mathcal{V}$. Fix $h_{\mathbf{u}} \in \mathcal{V}$ and recall that we already showed that $\mathcal{A}h_{\mathbf{u}} \in \mathcal{V}$ and thus $|\mathcal{A}h_{\mathbf{u}}| \leqslant C \exp(\alpha |\cdot|)$. To show that $\int_0^t P_s |\mathcal{A}h_{\mathbf{u}}|(x) ds < \infty$ it suffices to note that

$$\exp(\alpha |x|) \leqslant \exp(\alpha x) + \exp(-\alpha x) \tag{4.B.6}$$

and (4.B.4) can be applied. Next, we prove that for all $t \in [0,T]$, $P_t h_{\mathbf{u}} \in \mathcal{V}$. Fix $t \in [0,T]$ and for each $z \in \mathbb{C}$ set

$$P_t h_{\mathbf{u}}(z) = \mathbb{E}[h_{\mathbf{u}}(X_t + z)].$$

Observe that by (4.B.5) there is a $\tau < \alpha$ such that

$$\mathbb{E}[|h_{\mathbf{u}}(X_t + z)|] \leqslant C\mathbb{E}[\exp(\tau|X_t + z|)] \leqslant C\exp(\tau|z|)\mathbb{E}[\exp(\tau|X_t|)], \tag{4.B.7}$$

for each $z \in \mathbb{C}$. Therefore by the dominated convergence theorem the map $z \mapsto \mathbb{E}[h_{\mathbf{u}}(X_t + z)]$ is continuous and $\int_\triangle \mathbb{E}[|h_{\mathbf{u}}(X_t + z)|]dz$ is finite for every triangle $\triangle \subset \mathbb{C}$. By Fubini's and Goursat's theorem (see Proposition 4.D.2(v)), we then get $\int_\triangle \mathbb{E}[h_{\mathbf{u}}(X_t + z)]dz = \mathbb{E}[\int_\triangle h_{\mathbf{u}}(X_t + z)dz] = 0$. Finally, by Morera's theorem (see Proposition 4.D.2(vi)) $P_t h_{\mathbf{u}} \in H(\mathbb{C})$ and by (4.B.7) and (4.B.5) $P_t h_{\mathbf{u}} \in \mathcal{V}$. Finally, notice that since $\mathcal{A}h_{\mathbf{u}} \in \mathcal{V}$, by the previous reasoning $P_t\mathcal{A}h_{\mathbf{u}} \in \mathcal{V}$, implying in particular that $S \ni x \mapsto P_t\mathcal{A}h_{\mathbf{u}}(x)$ is continuous, concluding the proof.

Finally, let us denote by $\mathbf{c}(t)_{t\in[0,T]}$ such solution and recall from Lemma 4.3.22 that for all $x \in S$, $h_{\mathbf{c}(t)}(x) = P_t h_{\mathbf{u}}(x)$ and $\mathcal{A}P_t h_{\mathbf{u}}(x) = P_t\mathcal{A}h_{\mathbf{u}}(x)$. Since $\mathbf{c}(t) \in \mathcal{V}^*$, by (ii) condition (ii) of Theorem 4.3.20 is verified. Next, notice that again by Lemma 4.3.22 and (4.B.4)

$$\int_0^T \int_0^T \mathbb{E}[|\mathcal{A}h_{\mathbf{c}(s)}(X_t)|]dsdt = \int_0^T \int_0^T \mathbb{E}[|P_s\mathcal{A}h_{\mathbf{u}}(X_t)|]dsdt < \infty.$$

Since all the conditions of Theorem 4.3.20 are verified, the claim follows.

4.B.5 Proof of Corollary 4.3.31

Set

$$\mathcal{B} = \{h \in H(\mathbb{C})\colon |h|_{\mathbb{R}}| \leqslant C\exp(a|\cdot|) \text{ for some } a \in \mathbb{R}, C > 0\}.$$

and notice that $\mathcal{B} \subseteq \mathcal{V}$.

Condition (i) follows by Remark 4.3.16(iv) and (ii) follows as in Proposition 4.3.29. We highlight in particular that $\mathcal{A}(\mathcal{V}) \subseteq \mathcal{B}$.

(iii): We proceed as in the proof of Theorem 4.3.29 showing that the conditions of Lemma 4.3.22 are satisfied. Consider the semigroup $(P_t)_{t\in[0,T]}$ given by $P_t h(x) = \mathbb{E}[h(X_t + x)]$, for $h \in \mathcal{B}$, $t \in [0,T]$, $x \in \mathbb{R}$. Since for each $h \in \mathcal{V}$ it holds $\mathcal{A}h \in \mathcal{B}$, following the proof of Theorem 4.3.29(iii) we get that $\int_0^T P_s|\mathcal{A}h|(x)ds < \infty$ for all $x \in S$ and N^h is a true martingale.

Fix $h_{\mathbf{u}} \in \mathcal{V}$, for $\mathbf{u}$ as in (iii). Observe that by Proposition 4.D.5 we get

$$h_{\mathbf{u}}(x) = \int_{\mathbb{R}} \hat{h}_{\mathbf{u}}(u)\exp(iux)du, \tag{4.B.8}$$

for a.e. $x \in \mathbb{R}$. Since $h_{\mathbf{u}}$ is continuous on $\mathbb{R}$ by assumption, and the right-hand side of (4.B.8) is continuous by (4.3.35) and dominated converge theorem, the equality in (4.B.8) holds for every $x \in \mathbb{R}$. Finally, for $t \in [0,T]$ and $x \in \mathbb{R}$, by Fubini's theorem we have that

$$P_t h_{\mathbf{u}}(x) = \int_{\mathbb{R}} \mathbb{E}[\exp(iuX_t)] \exp(iux)\hat{h}_{\mathbf{u}}(u)du.$$

Since the expectation on the right hand side is bounded by 1, by (4.3.35) we can extend this map to $\mathbb{C}$ and use Morera's theorem to deduce that $P_t h_{\mathbf{u}}$ is an entire map bounded on $\mathbb{R}$. Moreover, by (4.B.4) and Leibniz integral rule we can conclude that $P_t h_{\mathbf{u}} \in \mathcal{V}$. Since $\mathcal{A}h_{\mathbf{u}} \in \mathcal{B}$ by (4.B.4) $P_t\mathcal{A}h_{\mathbf{u}}$ is continuous on S. The last part of the proof follows as in Theorem 4.3.29. □

4.B.6 Proof of Proposition 4.3.33

(i): This follows by Remark 4.3.16(iv).

(ii): Observe that for $h \in \mathcal{V}$ by the characterization of the characteristic of an affine process (see Lemma 2 in Filipović and Larsson (2020)) we have that $|\mathcal{A}h(x)| \leqslant C(1+|x|)$. Since X is also a polynomial process we know that

$$\mathbb{E}[\sup_{t\in[0,T]}(1+|X_t|)] \leqslant C\mathbb{E}[(1+X_T^2)] < \infty,$$

showing that N^h is a uniformly integrable local martingale and thus a true martingale.

Next fix $\mathbf{u}$ as in (iii) and consider the semigroup given by $P_t h(x) := \mathbb{E}[h(X_t)|X_0 = x]$ for each measurable h such that $\mathbb{E}[|h(X_t)||X_0 = x] < \infty$. We verify that the remaining conditions of Lemma 4.3.22 are satisfied. Observe that by Fubini, we have that

$$P_t h_{\mathbf{u}}(x) = \mathbb{E}\Big[\int_{-\tau}^{\tau} \exp((\varepsilon+iu)X_t)g(u)du\Big|X_0 = x\Big] = \int_{-\tau}^{\tau} \exp(\phi(t,\varepsilon+iu)+\psi(t,\varepsilon+iu)x)g(u)du.$$

Since by the dominated convergence theorem we know that $u \mapsto \phi(t,\varepsilon+iu)$ and $u \mapsto \psi(t,\varepsilon+iu)$ are continuous for each t and ϵ, the map

$$x \mapsto \exp(\phi(t,\varepsilon+iu)+\psi(t,\varepsilon+iu)x)g(u)$$

can be extended to $\mathbb{C}$ and since its module is bounded on compacts an application of Fubini, Goursat and Morera (see Proposition 4.D.2(v) and (vi)) yield that $P_t h_{\mathbf{u}} \in H(\mathbb{C})$.

Next, note that since affine processes are Feller (see Theorem 2.7 in Duffie et al. (2003)) we know that P_t is mapping the set of bounded continuous functions into itself. This in particular implies that

$$P_t\mathcal{A}h_{\mathbf{u}}(x) = \lim_{N\to\infty} \mathbb{E}[\mathcal{A}h_{\mathbf{u}}(X_t) \wedge N|X_0 = x],$$

which is a sequence of continuous functions in x. Observe that

$$|P_t\mathcal{A}h_{\mathbf{u}}(x) - \mathbb{E}[\mathcal{A}h_{\mathbf{u}}(X_t) \wedge N|X_0 = x]| \leqslant \mathbb{E}[|\mathcal{A}h_{\mathbf{u}}(X_t)|1_{\{\mathcal{A}h_{\mathbf{u}}(X_t)>N\}}|X_0 = x],$$

which by Cauchy-Schwartz and Markov's inequality can be bounded by

$$\begin{aligned}&\mathbb{E}[|\mathcal{A}h_{\mathbf{u}}(X_t)|^2|X_0=x]^{1/2}\mathbb{P}(\{\mathcal{A}h_{\mathbf{u}}(X_t)>N\}|X_0=x)^{1/2}\\&\leqslant \mathbb{E}[|\mathcal{A}h_{\mathbf{u}}(X_t)|^2|X_0=x]^{1/2}\frac{\mathbb{E}[|\mathcal{A}h_{\mathbf{u}}(X_t)||X_0=x]^{1/2}}{N^{1/2}}.\end{aligned}$$

By the polynomial property of X, proceeding as in (ii) we get that the convergence is uniform on compacts and thus $P_t\mathcal{A}h_{\mathbf{u}}$ is continuous on S. Similarly, again by the polynomial property of X we get that for each $x\in S$, $P_t|\mathcal{A}h_{\mathbf{u}}|(x)$ is bounded by a continuous function in t and thus $\int_0^T P_t|\mathcal{A}h_{\mathbf{u}}|(x)dt<\infty$ for each $x\in S$. Observe furthermore that same argument implies also that $\int_0^T\int_0^T P_{t+s}|\mathcal{A}h_{\mathbf{u}}|(x)dsdt<\infty$. The claim follows by Remark 4.3.23. □

4.C Proofs of Section 4.4

Before presenting the proofs of Section 4.4.1, recall that by Definition 4.4.1, in order to prove that X is an S-valued $\mathcal{V}$-affine-holomorphic jump-diffusion we need to show that for each $f\in\mathcal{V}$ it holds $\exp(f)\in\mathcal{D}(\mathcal{A})$, there exists a map $\mathcal{R}:\mathcal{V}\to\mathcal{O}(S)$ such that for all $f\in\mathcal{V}$, $\mathcal{A}\exp(f)=\exp(f)\mathcal{R}(f)$ and the process $N^{\exp(f)}$ introduced in equation (4.2.3) defines a local martingale.

4.C.1 Proof of Lemma 4.4.4

First observe that from $\{e^h: h\in\mathcal{V}\}\cup\mathcal{V}\subseteq\mathcal{D}(J)$ it follows $\{e^h: h\in\mathcal{V}\}\cup\mathcal{V}\in\mathcal{D}(\mathcal{A})$. This in particular implies that the map $P^d_{R(S)+\varepsilon}(0)\ni z\mapsto Je^h(z,\cdot)-e^{h(z)}Jh(z,\cdot)\in L^1(E,F)$ is well defined and continuous for each $h\in\mathcal{V}$. This integrability then yields that

$$\begin{aligned}\mathcal{A}\exp(h)(z)&=\exp(h(z))\left(\nabla h(z)^\top b(z)+\frac{1}{2}\mathrm{Tr}\big(a(z)\left(\nabla^2h(z)+\nabla h(z)^\top\nabla h(z)\right)\big)\right)\\&\quad+\int_E Je^h(z,y)F(dy)\\&=\exp(h(z))\left(\mathcal{A}h(z)+\frac{1}{2}\mathrm{Tr}\big(a(z)\left(\nabla^2h(z)+\nabla h(z)^\top\nabla h(z)\right)\big)\right.\\&\quad\left.+\lambda(z)\int_E\exp\left(h(z+j_\varepsilon(z,\xi))-h(z)\right)-1-\left(h(z+j(z,\xi))-h(z)\right)\,F(d\xi)\right),\end{aligned}$$

for each $h\in\mathcal{V}$, $z\in P^d_{R(S)+\varepsilon}(0)$. The claim now follows as in the proof of Theorem 4.3.12. □

4.C.2 Proof of Corollary 4.4.5

We prove that the conditions of Theorem 4.4.3 and Lemma 4.4.4 are satisfied. Proceeding as in the proof of Corollary 4.3.15, for each $z\in P^d_{R(S)+\varepsilon}(0)$ and $y\in E$ both quantities

$\sup_{w\in P^d_{\delta_z}(z)} |Jh(w,y)|$ and $\sup_{w\in P^d_{\delta_z}(z)} |Je^h(w,y)|$ can be bounded by

$$C \sup_{w\in P^d_{\delta_z}(z)} \Big(\exp(m\ v(\|w + j_\varepsilon(w,y)\|))1_{\{\|j_\varepsilon(w,y)\|>1\}} + \|j_\varepsilon(w,y)\| \wedge \|j_\varepsilon(w,y)\|^2\Big).$$

Observe that the constant m enters in the bound due to the second quantity, of which we know how to bound the exponent and not the whole function.

By conditions (4.4.4), (4.4.5) this implies that $h, e^h \in \mathcal{D}(J)$. Finally, observe that the maps

$$z \longmapsto F^1_z(\cdot) := Je^h(z,\cdot) \qquad \text{and} \qquad z \longmapsto F^2_z(\cdot) := Jh(z,\cdot),$$

are both continuous by the dominated convergence theorem (see the proof of Corollary 4.3.15 for the detailed argument). □

4.C.3 Proof of Corollary 4.4.7

First observe that since $j(\cdot,\xi) = \xi$ is a real valued constant on S, the same holds for its extension $j_\varepsilon(\cdot,\xi) = \xi$ on $P^d_{R(S)+\varepsilon}(0)$. This in particular implies that

$$\int_{\mathbb{R}^d} \sup_{w\in P^1_{R(S)+\varepsilon}(0)} \|j_\varepsilon(w,\xi)\| \wedge \|j_\varepsilon(w,\xi)\|^2 F(d\xi) = \int_{\mathbb{R}^d} \|\xi\| \wedge \|\xi\|^2 F(d\xi) < \infty.$$

Moreover, using that $\xi \in \mathbb{R}^d$, by definition of $\mathcal{V}$ we get that $|h(z+\xi)|1_{\{\|\xi\|>1\}} \leqslant g_h(\Im(z))1_{\{\|\xi\|>1\}}$. Since for each $z \in P^d_{R(S)+\varepsilon}(0)$ and $\delta_z \in (0,\infty)$ such that $P^d_{\delta_z}(z) \subseteq P^d_{R(S)+\varepsilon}(0)$ by continuity of g_h it holds

$$\int_{\mathbb{R}^d} \sup_{z\in P^1_{\delta_z}(z)} \exp(g_h(\Im(z)))1_{\{\|\xi\|>1\}} F(d\xi) < \infty$$

the claim follows as for Corollary 4.4.5. □

4.D A primer on holomorphic functions

The goal of this section is to provide accurate statements and precise references for the needed results from complex analysis. We will use the notation relative to poldiscs and polytorus introduced in Section 4.2.1.3, as well as the multi-index notation introduced in Section 4.2.1.1.

Definition 4.D.1. (*Definition 1.2.1 in* Scheidemann (2005)) Let $U \subseteq \mathbb{C}^d$ be an open set and $m \in \mathbb{N}$.

(i) **Complex differentiable functions**

A function $f : U \to \mathbb{C}^m$ is called *complex differentiable* at $z_0 \in U$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ and a $\mathbb{C}$-linear map

$$Lf(z_0) : \mathbb{C}^d \to \mathbb{C}^m$$

such that for all $z \in U$ with $\|z - z_0\| \leqslant \delta$, the inequality

$$\|f(z) - f(z_0) - Lf(z_0)(z - z_0)\| \leqslant \varepsilon \|z - z_0\|$$

holds.

(ii) **Holomorphic functions**

A function $f : U \to \mathbb{C}^m$ is called *holomorphic* on U if it is complex differentiable at all $z_0 \in U$. A holomorphic function on the whole $\mathbb{C}^d$ is called *entire*.

Proposition 4.D.2. Let $U \subseteq \mathbb{C}^d$ be an open set.

(i) **Cauchy integral formula** (Theorem 1.3.3 in Scheidemann (2005))

Let $f : U \to \mathbb{C}$ be a holomorphic function and fix $z_0 \in U$ and $R \in (0, \infty)^d$ such that $\overline{P_R^d(z_0)} \subseteq U$. Then

$$D^\alpha f(z_0) = \frac{\alpha!}{(2\pi i)^d} \int_{T_R^d(z_0)} \frac{f(w)}{(z_0 - w)^{\alpha+1}} dw,$$

where $\alpha + 1 := (\alpha_1 + 1, \ldots, \alpha_d + 1)$, and

$$|D^\alpha f(z_0)| \leqslant \frac{\alpha!}{R^\alpha} \|f|_{T_R^d(z_0)}\|_\infty.$$

(ii) **Taylor expansion** (Corollary 1.5.9 in Scheidemann (2005))

Let $f : U \to \mathbb{R}$ be a holomorphic functions and fix $z_0 \in U$ and $R \in (0, \infty]^d$ such that $\overline{P_R^d(z_0)} \subseteq U$. Then

$$f(z) = \sum_{\alpha \in \mathbb{N}_0^d} \frac{1}{\alpha!} D^\alpha f(z_0)(z - z_0)^\alpha, \qquad \text{for all } z \in P_R^d(z_0). \tag{4.D.1}$$

By (i) and Example 1.5.7 in Scheidemann (2005), for $R \in (0, \infty)^d$ the reminder's term can be bounded as follows

$$\Big| \sum_{|\alpha| \geqslant k+1} \frac{1}{\alpha!} D^\alpha f(z_0)(z - z_0)^\alpha \Big| \leqslant \|f|_{T_R^d(z_0)}\|_\infty \prod_{i=1}^d \frac{1}{1 - \frac{|z_i - z_{0i}|}{R_i}} \sum_{|\alpha| = k+1} \frac{|z - z_0|^\alpha}{R^\alpha}.$$

(iii) **Identity theorem** (Conclusion 1.2.12.2 in Scheidemann (2005)) Let $f : U \to \mathbb{C}$ be a holomorphic function. If U is connected and $f = 0$ on some open set $E \subseteq U$, then $f(z) = 0$ for all $z \in U$. For $d = 1$ the set E is just required to have an accumulation point in U.

(iv) **Vitali-Porter theorem** (Exercise 1.4.37 in Scheidemann (2005))

Fix an open subset $E \subseteq U$ and consider a sequence $\{f_n\}_n$ of holomorphic functions $f_n : U \to \mathbb{C}$. Suppose that $\{f_n\}_n$ is a locally bounded sequence (uniformly bounded on every compact set of the domain of definition) and

$$f(z) := \lim_{n\to\infty} f_n(z) \tag{4.D.2}$$

exists for all $z \in E$. Then the limit in (4.D.2) is well defined for each $z \in U$, the resulting map $f : U \to \mathbb{C}$ is holomorphic, and the convergence holds locally uniformly. For $d = 1$ the set E is just required to have an accumulation point in U (see p.44 Schiff (1993)).

(v) **Goursat's theorem** (Theorem 1.1 in Stein and Shakarchi (2010))

Fix $d = 1$ and let $\triangle \subseteq U$ be a triangle whose interior is contained in U. If $f : U \to \mathbb{C}$ is holomorphic then

$$\int_\triangle f(z)dz = 0.$$

(vi) **Morera's theorem** (Theorem 5.1 in Stein and Shakarchi (2010))

Fix $d = 1, R \in (0, \infty)$, $a \in \mathbb{C}$, and let $f : P^1_R(a) \to \mathbb{C}$ be a continuous map. If for any triangle $\triangle \subset P^1_R(a)$ it holds

$$\int_\triangle f(z)dz = 0,$$

then f is holomorphic.

(vii) **Hartogs' theorem** (see p.28 in Shabat (1992))

Let $f : U \to \mathbb{C}$ be a partially holomorphic function, meaning that it is holomorphic in each variable while the other variables are held constant. Then f is holomorphic.

(viii) **Weierstrass' theorem** (Theorem 1.4.20 in Scheidemann (2005)).

$H(U,\mathbb{C})$ is a closed subspace of $C(U,\mathbb{C})$ with respect to the locally uniform convergence. Moreover, for every $\alpha \in \mathbb{N}_0^d$ the linear operator $D^\alpha : H(U,\mathbb{C}) \to H(U,\mathbb{C})$ is continuous.

Remark 4.D.3. If $U = P^d_\infty(0)$, and f thus entire, the representation in (4.D.1) holds for all $z, z_0 \in \mathbb{C}$. If this is not the case, then requiring that $z, z_0 \in U$ is not sufficient to guarantee that $\overline{P^d_{|z-z_0|}(z_0)} \subseteq U$ and thus to derive a power series representation of $f(z)$ centered in z_0 as in (4.D.1). For this specific reason in Theorem 4.3.6 we need to impose condition (4.3.6) instead of

$$G_j > \sup_{z\in S_\varepsilon} \max\{|z_j|, \sup_{\xi\in\mathrm{supp}(K_\varepsilon(z,\cdot))} |z_j + \xi_j|\}.$$

We also consider a few results concerning holomorphic functions of finite order and type (see (4.3.18) and (4.3.19)).

Proposition 4.D.4. (Theorem 2.4.1, Theorem 6.2.14, Theorem 6.2.4 and Theorem 11.1.2 in Boas (2011) and Theorem 19.3 in Rudin (1986))

(i) Order and type of an entire function do not change under differentiation.

(ii) Non-constant entire functions of order strictly smaller than 1 are unbounded on lines.

(iii) Let $\tau \in (0, \infty)$ and let B_τ be the set of all entire functions of exponential type τ, which are bounded on the real axis. Then $h \in B_\tau$ implies

$$|h(z)| \leqslant \sup_{x \in \mathbb{R}} |h(x)| \exp(\tau |\Im(z)|).$$

Moreover, in this case we also have that $h' \in B_\tau$, $\sup_{x\in\mathbb{R}} |h'(x)| \leqslant \sup_{x\in\mathbb{R}} |h(x)|\tau$ for all $z \in \mathbb{C}$, and thus in particular

$$|h'(z)| \leqslant \sup_{x \in \mathbb{R}} |h(x)|\tau \exp(\tau |\Im(z)|).$$

(iv) **Paley–Wiener theorem** Let $h \in H(\mathbb{C})$ be an entire function of exponential type τ such that

$$\int_{\mathbb{R}} |h(x)|^2 dx < \infty.$$

Then there exists $g \in L^2(-\tau, \tau)$ such that $h(z) = \int_{-\tau}^{\tau} g(t)e^{itz}dt$ for each $z \in \mathbb{C}$.

In the spirit of the Paley-Wiener theorem, but for general holomorphic functions, we have the following result (see Theorem 9.11 and Theorem 9.14 in Rudin (1986)).

Proposition 4.D.5 (Inversion theorems)**.** Fix $h \in H(\mathbb{C})$ such that

$$\int_{\mathbb{R}} |h(x)|dx < \infty \quad \text{or} \quad \int_{\mathbb{R}} |h(x)|^2 dx < \infty$$

and suppose that the map $\hat{h} : \mathbb{R} \to \mathbb{C}$ given by (4.3.34) satisfies $\int_{\mathbb{R}} |\hat{h}(u)|du < \infty$. Then $h(x) = \int_{\mathbb{R}} \hat{h}(t) \exp(itx)dt$, for a.e. $x \in \mathbb{R}$.

5 Functional Itô-formula and Taylor expansion of non-anticipative maps of rough paths

The present chapter is based on Cuchiero et al. (2024c), a joint work with Christa Cuchiero and Xin Guo.

5.1 Introduction

The Itô-formula is one of the main tools of stochastic calculus, extending the chain rule from classical calculus to stochastic analysis and asserting that any sufficiently smooth function of a semimartingale is itself a semimartingale. Recognizing that uncertainty often affects the present situation through both the current state and the entire process' history, the seminal paper Dupire (2019) extended the Itô-formula to path functionals which may depend on the entire trajectory of a continuous semimartingale (with values in a $\mathbb{R}^d$) up to the present time t, i.e., so-called non-anticipative path functionals. This then opened the door to many subsequent studies as e.g., Cont and Fournie (2010a,b, 2013); Ananova and Cont (2017) Litterer and Oberhauser (2011), Cosso and Russo (2016), Jazaerli and F. Saporito (2017), Keller and Zhang (2016), Cont and Perkowski (2018), Viens and Zhang (2019), Riga (2016) Houdré and Víquez (2024) and the references therein.

The original formulation of the functional Itô-formula incorporates vertical and horizontal derivatives (detailed in Section 2 in Dupire (2019)), reflecting the functional's dependence on the trajectory and current time, respectively. This formula is derived using a Taylor expansion of the functionals along an approximation of the path. This approach allows for a second-order approximation of the functional, which is quadratic in the path increments, making it well-suited for studying functionals of all those paths for which such approximations converge. It is not surprising, therefore, that this technique applies not only to functionals of semimartingales but also to more general paths of finite quadratic variation. This is, in fact, the finding of the subsequent paper Cont and Fournie (2013), which establishes a pathwise functional Itô-formula for non-anticipative functionals of càdlàg paths of finite quadratic variation in the sense of Föllmer, providing furthermore an extension of the results in Föllmer (1981).

Despite the significant impact of these results, many path functionals of interest are still excluded. Indeed, the framework is limited to functionals that depend continuously on the trajectories of the path with respect to (some variants of) the uniform topology (see Section 1.2 in Cont and Fournie (2013) for details). This excludes several important examples, such as standard Itô/Stratonovich integrals and the Itô-map, which describes the correspondence between the solution of a stochastic differential equation and the driving signal. To clarify this statement, consider the sequence of real-valued paths $X^{1,n}, X^{2,n} : [0, 2\pi] \to \mathbb{R}$ defined via

$$X_t^{1,n} := -n^{-\frac{1}{3}} \cos(nt), \qquad X_t^{2,n} := n^{-\frac{1}{3}} \sin(nt), \qquad \text{for all } t \in [0, 2\pi],$$

for $n \in \mathbb{N}$ (see Allan (2021)). Then,

$$\int_0^{2\pi} X_s^{2,n} dX_s^{1,n} = n^{\frac{1}{3}} \pi, \tag{5.1.1}$$

implying that $X^{1,n}, X^{2,n} \to 0$ in the uniform topology, but $\int_0^{2\pi} X_s^{2,n} dX_s^{1,n} \to \infty$.

To address these limitations Cont and Fournie (2010a) expanded the framework to include functionals that depend on both the semimartingale and its quadratic variation, and developed a weak functional calculus that works well in a probabilistic setup, whereas does not retain the pathwise perspective. Other efforts to broaden this applicability can also be found in Oberhauser (2016), which uses instead the old work of Bichteler.

The non-negligible limitations, excluding, e.g., iterated integrals, raise the question of which topology to consider on the space of paths to determine a (pathwise) functional Itô-formula that can encompass these examples. A direct answer to this problem is provided by rough path theory, pioneered by Terry Lyons (see Lyons (1998); Lyons et al. (2007)), for which, in fact, one of the fundamental results consists in identifying a natural family of topologies on paths space such that iterated Itô/Stratonovich integrals, or more generally solutions to controlled differential equations, are continuous maps with respect to the driving signal.

Relying on this powerful theory, we can thus prove a pathwise *(rough) functional Itô-formula* that covers these examples. To this end, we equip the path spaces with the stronger variation topology (see Definition 1.2.2) and consider functionals of weakly geometric càdlàg p-rough paths for $p \in [1, 3)$ (see Definition 1.2.1). This means that when dealing with paths of finite p-variation for $p \geqslant 2$ (which is the case for sample paths of a semimartingale), we consider functionals that depend on a lifted path, i.e., the path itself (with values in $\mathbb{R}^d$) and (some of) its rough paths lift.

Our proof technique is based on a *density approach building on linear functions of the signature* (see Section 5.2.4), which are, in fact, specific non-anticipative functionals of weakly geometric rough paths that exhibit powerful approximating properties (see e.g., Section 2.2 as well as Kalsi et al. (2020), Bayer et al. (2023), Cuchiero et al. (2023b)). Our approach, therefore, consists of establishing an Itô-formula for these specific functionals and then extending it to more general maps via a density argument. This is indeed one possible proof technique also used in classical stochastic calculus (see, e.g., Theorem 5.7

in Miller and Silvestri (2017)), as well as for measure-valued processes (see, e.g., Guo et al. (2023), Jacka and Tribe (2003)).

To elaborate further, the functional Itô-formula is derived following the subsequent steps.

Firstly, inspired by the paper Chevyrev and Friz (2019) (see also Marcus (1978, 1981) and Williams (2001)) and in particular, by the signature construction (see Section 1.3.3), we introduce the class of so-called *non-anticipative Marcus canonical* path functionals. Roughly speaking, these are all those maps

$$F : [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$$

that depend on weakly geometric rough paths in a non-anticipative way (see Definition 5.2.2) and which are invariant under Marcus transformation, i.e.,

$$F(t, \mathbf{X}) = F(\psi_R^{-1}(\tau_{\mathbf{X},R}(t), \widetilde{\mathbf{X}}),$$

for $\widetilde{\mathbf{X}}$ denoting some Marcus transformed path of $\mathbf{X}$ with respect to some pair (R, ψ_R). We refer to Section 1.3.2 and Definition 5.2.4 for the precise definition of the involved quantities, and also to Fujiwara and Kunita (1985), Applebaum and Tang (2001), Kurtz et al. (1995), Chevyrev et al. (2020) for other instances where a similar approach is used. Then, adopting the Lie group valued paths point of view on the space of weakly geometric rough paths, we introduce a notion of vertical derivatives for Marcus canonical paths functionals, inspired by the (Euclidean) one introduced in Dupire (2019). Indeed, we call the quantity

$$U^i F(t, \mathbf{X}) := \frac{d}{dh} F(\psi_R^{-1}(\tau_{\mathbf{X},R}(t), \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant \psi_R^{-1}(\tau_{\mathbf{X},R}(t)\}})|_{h=0}, \tag{5.1.2}$$

vertical derivative of F in the direction $i = 1, \ldots, d$ at $(t, \mathbf{X})$, when the derivative exists (see Definition 5.2.14).

Notice that (5.1.2) might be interpreted as a directional derivative (or rather Lie vertical derivative) of the functional F in the specific direction determined by the vertical perturbation. This approach goes in the direction of Qian and Tudor (2011) where a first attempt for studying a differential structure of the (non-linear) space of rough paths has been made (see also Schmeding (2022) for a discussion on this topic).

We remark that when inserting a càdlàg path in (5.1.2) the Marcus transformation of the path $\mathbf{X}$ is computed before the vertical perturbation. This is, in fact, a key aspect of this definition. On the one hand, it allows the preservation of the Marcus property of the functionals also at the level of the (functional) derivative (see Proposition 5.2.16). On the other hand, if the original path $\mathbf{X}$ admits a jump a time t, the Marcus property of the functional F allows to interpret the derivative in (5.1.2) as a derivative involving a delayed perturbation of the original path

$$\mathbf{X}^t \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant t+\varepsilon\}},$$

for some $\varepsilon > 0$ independent on the specific pair (R, ψ_R) (see Remark 5.2.15(ii)). This specification is particularly relevant when computing the *higher-order vertical derivatives*, which is done by an iterative application of the procedure in (5.1.2) (see Definition 5.2.19). In this case, by definition, the perturbations are always computed at a jump time, resulting therefore in a notion of vertical derivative of order $k \in \mathbb{N}$, $k \geqslant 2$, which involves delayed perturbations of the form

$$\mathbf{X}^t \otimes \exp^{([p])}(h_k \epsilon_{i_k}) 1_{\{\cdot \geqslant t+\varepsilon_k\}} \otimes \cdots \otimes \exp^{([p])}(h_1 \epsilon_{i_1}) 1_{\{\cdot \geqslant t+\varepsilon_k+\cdots+\varepsilon_1\}},$$

for some $h_k, \ldots, h_1 \in \mathbb{R}$, $\varepsilon_k, \ldots, \varepsilon_1 > 0$, and $i_k, \ldots, i_1 = 1, \ldots, d$.

Via the Marcus property of the paths we are thus able to establish a differential calculus for path functionals that allows for non-commutative derivation orders. This is one remarkable key feature of our framework, as it enables both the derivation of a functional Itô-formula for linear functions of the signature (see Lemma 5.2.28) and a Taylor expansion in terms of the signature of sufficiently regular non-anticipative Marcus canonical path functional (see Section 5.5). These results would not follow from a direct application of the differential calculus introduced in Dupire (2019) and adopted in Cont and Fournie (2010a), Cont and Fournie (2013). Indeed, in their framework, the higher-order derivatives involve vertical perturbations of the path at the current time: which then results in higher order (functional) derivatives that take values in fact in the space of symmetric tensors over $\mathbb{R}^d$ (see e.g., Cont and Perkowski (2018) as well as Remarks 5.2.23 and 5.5.8(ii)).

With these notions of derivatives, a functional Itô-formula for linear functions of the signature follows then just by the the definition of the signature of weakly geometric càdlàg rough paths.

Relying on that, we then prove a universal approximation theorem (UAT) for vertically differentiable path functionals, which states that any C^K-non-anticipative Marcus canonical path functionals F (in the sense of Definition 5.3.12) evaluated at some *tracking jumps* (tj)-extended path $\widehat{\mathbf{X}}$ (see Definiton 5.3.3) can be approximated uniformly in time together with its derivatives by linear functionals on the signature and their derivatives (see Theorem 5.3.15 for the precise formulation). Given the non-linear structure of the vertical (Lie) derivatives, deriving this result has proven to be a very delicate task. Similarly to Theorem 2.2.6, in order to recover a point-separation property and to express the functional's dependency on the current time, we need first to specify the set of tj-extended paths, which are all those that admit a time-extended Marcus transformation. Then, the claim follows via an intricate application of Nachbin-type theorems (see Nachbin (1949)), as well as some key ideas from Lie group theory.

Finally, relying on this UAT and some interpolation arguments, we prove the (rough) functional Itô-formula: any $C^{[p]+1}$-non-anticipative Marcus canonical, for which the functional itself, together with its derivatives and the remainder functional (see Definition 5.2.12) is $[[p], [p]+1)$-var continuous (see Definition 5.2.24), satisfies a (rough) functional Itô-formula (see Theorems 5.4.1, 5.4.4 for the precise statements), which furthermore matches some standard as well as functional Itô-formula in the literature (see Section 5.4.3).

The second main result of the chapter is the *functional Taylor expansion* in terms of the *signature*, whose proof relies on the established non-commutative calculus. More specifically, as a byproduct of the functional Itô-formula, we prove a functional Taylor expansion in terms of signature for tj-extended multidimensional weakly geometric rough paths (see Theorems 5.5.1, 5.5.5 and the following discussion). To the best of our knowledge, this is the first purely deterministic (rough) Taylor expansion of functionals of weakly geometric càdlàg rough paths. It nevertheless shares similarities with the work by Buckdahn et al. (2015), where a rough Taylor expansion is derived by identifying the vertical derivatives with the abstract notion of Gubinelli derivatives, expansions coming from control theory e.g., Fliess (1981, 1983, 1986), Beauchard et al. (2023), as well as stochastic Taylor expansions (see e.g., Litterer and Oberhauser (2011), Klöden and Platen (1992), Arous (1989)) and the recently established results in Dupire and Tissot-Daguette (2023). There the authors consider functionals of continuous one-dimensional time-extended paths with finite quadratic variation in the sense of Föllmer (see Remark 5.5.8(ii) for a discussion on this specific expansion).

The chapter is structured as follows. In Section 5.2, we introduce the space of non-anticipative Marcus-canonical path functionals, the differential calculus for this type of functionals as well as the considered p-variation topologies. In Section 5.3, we introduce the set of so-called tj-extended paths and present the UAT for vertically differentiable path functionals. Finally, in Section 5.4 and Section 5.5, we prove the functional Itô-formula and the Taylor expansion, respectively. In the Appendix, we collect the technical proofs of Section 5.3 and recall different pathwise integration approaches to which we compare the rough setting.

Throughout the chapter, we let $p \in [1,3)$ and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Moreover, we refer to Chapter 1 for the notation and main results concerning weakly geometric càdlàg rough paths and their signature.

5.2 Functionals of weakly geometric càdlàg rough paths

5.2.1 Non-anticipative Marcus canonical path functionals

Inspired by the notion of Marcus-type-RDEs (see Chevyrev and Friz (2019)), we here introduce the class of so-called *non-anticipative Marcus-canonical path functionals*. To this end, we start with the notion of non-anticipative path functionals, which in turn relies on the definition of *stopped-paths*, which we recall below. In the following, we let $p \in [1,3)$.

Definition 5.2.1. Given $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and $t \in [0,1]$, we define the *stopped weakly geometric càdlàg rough path of* $\mathbf{X}$ stopped at time t as the càdlàg path

$$\mathbf{X}^t : [0,1] \to G^{[p]}(\mathbb{R}^d)$$

given by

$$\mathbf{X}_u^t := \mathbf{X}_u 1_{\{u<t\}} + \mathbf{X}_t 1_{\{u \geqslant t\}},$$

for all $u \in [0,1]$.

Definition 5.2.2. Let $F : [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$. We say that F is a *non-anticipative path functional* if for all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$, it holds that

$$F(t, \mathbf{X}) = F(t, \mathbf{X}^t).$$

Remark 5.2.3. (i) The variable t has the role of a parameter and not of a component of the path $\mathbf{X}$. In fact, it is the parameter needed to specify that the path functional is non-anticipative.

(ii) Every non-anticipative path functional F admits a representation of the form

$$F(t, \mathbf{X}) = \widetilde{F}(\mathbf{X}^t),$$

for some path functional $\widetilde{F} : D([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$.

Finally, the space of non-anticipative Marcus canonical path functional is defined as the set of all non-anticipative path functionals that are invariant under a Marcus transformation (defined in Section 1.3.2), in the precise sense specified below.

For $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$, recall the notion of a Marcus-transformed path of $\mathbf{X}$ associated to some pair (R, ψ_R) given in Definition 1.3.8.

Definition 5.2.4. Let $F : [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$ be a non-anticipative path functional. We say that F is a *non-anticipative Marcus-canonical path functional* if

(i) for all $\mathbf{X} \in C([0,1], G^{[p]}(\mathbb{R}^d))$, $t \in [0,1]$, ϕ time reparametrization, it holds that

$$F(t, \mathbf{X}_\phi) = F(\phi(t), \mathbf{X});$$

(ii) for all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$

$$F(t, \mathbf{X}) = F(\psi_R^{-1}(\tau_{\mathbf{X},R}(t)), \widetilde{\mathbf{X}}),$$

where $\widetilde{\mathbf{X}}$ denotes the Marcus-transformed path of $\mathbf{X}$ with respect to some pair (R, ψ_R).

We denote the space of such functionals by $\mathcal{M}^0_{[p]}$.

Notation 5.2.5. For $m \in \mathbb{N}_0$, we set $((\mathcal{M}^0_{[p]})^d)^{\otimes 0} := \mathcal{M}^0_{[p]}$ and write $F \in ((\mathcal{M}^0_{[p]})^d)^{\otimes m}$ if

$$F : [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes m}$$

is a path functional whose components are non-anticipative Marcus canonical path functionals in the sense of Definition 5.2.4.

Remark 5.2.6. (i) The symbol $\mathcal{M}^0_{[p]}$ does not include the dimension d, as it will always be clear from the context.

(ii) The value of F at $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$ is independent of the particular Marcus transformation. Indeed, let $\mathbf{Z}$, $\widetilde{\mathbf{Z}}$ be two Marcus-transformations of $\mathbf{X}$ associated (R, ψ_R) and $(\widetilde{R}, \psi_{\widetilde{R}})$, respectively. By Remark 1.3.10 (iii), there exists a time reparametrization ϕ such that $\mathbf{Z}_\phi = \widetilde{\mathbf{Z}}$ and for all $t \in [0,1]$,

$$\phi(\psi^{-1}_{\widetilde{R}}(\tau_{\mathbf{X},\widetilde{R}}(t))) = \psi^{-1}_R(\tau_{\mathbf{X},R}(t)).$$

By (i) in Definition 5.2.4, we then get

$$F(\psi^{-1}_{\widetilde{R}}(\tau_{\mathbf{X},\widetilde{R}}(t)), \widetilde{\mathbf{Z}}) = F(\phi(\psi^{-1}_{\widetilde{R}}(\tau_{\mathbf{X},\widetilde{R}}(t))), \mathbf{Z}) = F(\psi^{-1}_R(\tau_{\mathbf{X},R}(t)), \mathbf{Z}). \tag{5.2.1}$$

(iii) Observe that a non-anticipative path functional defined only on the set of continuous paths that satisfies condition (i) of Definition 5.2.4 can always be extended to a path functional on the entire set of càdlàg path via condition (ii) of Definition 5.2.4. The resulting functional is a well-defined Marcus canonical path functional (see (ii) of this remark). Moreover, this extension is also unique. This is in fact the approach proposed in the inspiring paper Chevyrev and Friz (2019).

Notation 5.2.7. Given the independence on the specific Marcus transformed path discussed in Remark 5.2.6 and to ease notation, we set $\mu_t := \psi^{-1}_R(\tau_{\mathbf{X},R}(t))$, for all $t \in [0,1]$, $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$ and a pair (R, ψ_R), whenever there is no ambiguity.

Example 5.2.8. We now present some examples of Marcus-canonical non-anticipative path functionals.

(i) Let $F(t, \mathbf{X}) := Y_t$, where Y denotes the solution of a Marcus-type RDE in the sense of Chevyrev and Friz (2019) (see also equation (1.3.9) and the discussion thereafter) driven by some $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$, at time t. Then, by definition of such solutions, F is a non-anticipative Marcus canonical path functional.

(ii) Consider the functional $F(t, \mathbf{X})$ defined via

$$F(t, \mathbf{X}) := \sup_{s \in [0,t]} \|\mathbf{X}_s\|,$$

for all $(t, \mathbf{X}) \in [0,T] \times D^p([0,1], G^{[p]}(\mathbb{R}^d))$. Then F is a Marcus canonical non-anticipative path-functional.

(iii) For every $\mathbf{X} \in C^1([0,1], G^1(\mathbb{R}^2))$, set $\mathbf{X}^1_\cdot := \langle \epsilon_1, \mathbf{X}_\cdot \rangle =$ and $\mathbf{X}^2_\cdot := \langle \epsilon_2, \mathbf{X}_\cdot \rangle$, and consider the non-anticipative path functional (defined only on the set of continuous paths) via $F(t, \mathbf{X}) := \int_0^t \mathbf{X}^1_s d\mathbf{X}^2_s$. Since F verifies (i) of Definition 5.2.4, following the discussion in Remark 5.2.6 (iii), we can extend it to the set $D^1([0,1], G^1(\mathbb{R}^2))$ via condition (ii). A direct computation shows that the resulting functional (that we still denote by F) explicitly reads as

$$F(t, \mathbf{X}) := \int_0^t \mathbf{X}^1_{s-} d\mathbf{X}^2_s + \frac{1}{2} \sum_{s \leqslant t} \Delta \mathbf{X}^1_s \Delta \mathbf{X}^2_s, \tag{5.2.2}$$

for all $(t, \mathbf{X}) \in D^1([0,1], G^1(\mathbb{R}^2))$.

Remark 5.2.9. It is important to note that not every non-anticipative path functional that is well defined on the space of càdlàg paths is a Marcus canonical path functional. Consider for instance the functional defined via $F(t, \mathbf{X}) := \int_0^t \mathbf{X}^1_{s-} d\mathbf{X}^2_s$, for all $(t, \mathbf{X}) \in D^1([0,1], G^1(\mathbb{R}^2))$. Then, F is not Marcus canonical. Indeed, for $\mathbf{X} \in D^1([0,1], G^1(\mathbb{R}^2)) \backslash C^1([0,1], G^1(\mathbb{R}^2))$, let $\widetilde{\mathbf{X}}$ be its Marcus transformed path with respect to some pair (R, ψ_R). Then,

$$\begin{aligned} F(\psi_R^{-1}(\tau_{\mathbf{X},R}(t)), \widetilde{\mathbf{X}}) &= \int_0^{\psi_R^{-1}(\tau_{\mathbf{X},R}(t)} \widetilde{\mathbf{X}}^1_s d\widetilde{\mathbf{X}}^2_s \\ &= \int_0^t \mathbf{X}^1_{s-} d\mathbf{X}^2_s + \frac{1}{2} \sum_{s \leqslant t} \Delta \mathbf{X}^1_s \Delta \mathbf{X}^2_s \neq F(t, \mathbf{X}). \end{aligned}$$

Similarly, the path functional given by $F(t, \mathbf{X}) := \sum_{s \leqslant t} \Delta \mathbf{X}^1_s \Delta \mathbf{X}^2_s$, for all $(t, \mathbf{X}) \in D^1([0,1], G^1(\mathbb{R}^2))$ is not Marcus canonical as $0 = F(\psi_R^{-1}(\tau_{\mathbf{X},R}(t)), \widetilde{\mathbf{X}}) \neq F(t, \mathbf{X})$ for pure jump paths.

Next, we introduce the notion of path functionals that are invariant under reparametrization, which is the same as condition (i) in Definition 5.2.4, however on the whole space $D([0,1], G^{[p]}(\mathbb{R}^d))$ and not only on $C([0,1], G^{[p]}(\mathbb{R}^d))$.

Definition 5.2.10. Let $F : [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$ be a non-anticipative path functional. We say that F is *invariant under reparametrization* if for all $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$, $t \in [0,1]$ and ϕ time reparametrization,

$$F(t, \mathbf{X}_\phi) = F(\phi(t), \mathbf{X}).$$

We now show that every $F \in \mathcal{M}^0_{[p]}$ satisfies this property.

Proposition 5.2.11. Let $F \in \mathcal{M}^0_{[p]}$. Then F is invariant under reparametrization.

Proof. Let $\mathbf{X} \in D([0,1], G^{[p]}(\mathbb{R}^d))$ and ϕ a time reparametrization, one can show that there exists another time reparametrization η such that $\mathbf{Z}_\eta = \widetilde{\mathbf{Z}}$, where $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ denote the Marcus transformations of the càdlàg paths $\mathbf{X}_\phi$ and $\mathbf{X}$ with respect to some pairs (R, ψ_R) and $(\widetilde{R}, \psi_{\widetilde{R}})$, respectively. Since η can be chosen to satisfy

$$\eta(\psi_{\widetilde{R}}^{-1}(\tau_{\mathbf{X},\widetilde{R}}(\phi(t)))) = \psi_R^{-1}(\tau_{\mathbf{X}_\phi,R}(t)),$$

for all $t \in [0,1]$, the claim follows as in equation (5.2.1). □

Finally, we introduce the notion of remainder path functional related to some functionals of $G^2(\mathbb{R}^d)$-valued paths.

Definition 5.2.12. Let $F \in ((\mathcal{M}^0_{[p]})^d)^{\otimes m}$ and $F' \in ((\mathcal{M}^0_{[p]})^d)^{\otimes m+1}$, for $m \in \mathbb{N}_0$. We define the *remainder path functional* (related to F and F') as follows:

$$R^{F,F'} : \Delta_1 \times D([0,1], G^2(\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes m},$$

given by

$$R^{F,F'}((s,t),\mathbf{X}) := F(t,\mathbf{X}) - F(s,\mathbf{X}) - F'(s,\mathbf{X})\pi_1(\mathbf{X}_{s,t}),$$

for all $((s,t),\mathbf{X}) \in \Delta_1 \times D([0,1], G^2(\mathbb{R}^d))$.

Remark 5.2.13. For all $(s,t) \in \Delta_1$, $F'(s,\mathbf{X})\pi_1(\mathbf{X})_{s,t}$ denotes the contraction of the tensors $F'(s,\mathbf{X}) \in (\mathbb{R}^d)^{\otimes m+1}$ and $\pi_1(\mathbf{X}_{s,t}) \in \mathbb{R}^d$. Thus, $F'(s,\mathbf{X})\pi_1(\mathbf{X})_{s,t} \in (\mathbb{R}^d)^{\otimes m}$.

5.2.2 Vertically differentiable path functionals

Vertical derivatives Recall that $\mathcal{M}^0_{[p]}$ denotes the space of non-anticipative Marcus canonical path functionals.

Definition 5.2.14. Let $F \in \mathcal{M}^0_{[p]}$. We say that F is vertically differentiable at $(t,\mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$ in the direction $i = 1,\dots,d$ if the map

$$\mathbb{R} \ni h \mapsto F(\mu_t, \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i)1_{\{\cdot \geqslant \mu_t\}})$$

is differentiable at 0, for some Marcus transformed path $\widetilde{\mathbf{X}}$ and $\mu_t \in [0,1]$ given in Notation 5.2.7. In this case we call

$$\frac{d}{dh} F(\mu_t, \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i)1_{\{\cdot \geqslant \mu_t\}})|_{h=0}. \tag{5.2.3}$$

the vertical derivative of F at $(t,\mathbf{X})$ in the direction i.

If F is vertically differentiable in all directions $i = 1,\dots,d$ at all $(t,\mathbf{X})$, we say that F is vertically differentiable. We denote the space of such functionals by $\mathcal{M}^1_{[p]}$.

Remark 5.2.15. (i) We highlight that the path $\mathbf{Z} := \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i)1_{\{\cdot \geqslant \mu_t\}})$ denotes the path belonging to $D([0,1], G^{[p]}(\mathbb{R}^d))$ such that for all $s \in [0,1]$,

$$\mathbf{Z}_s := \begin{cases} \exp^{(1)}(\widetilde{X}_s + h\epsilon_i 1_{\{s \geqslant \mu_t\}}), & \text{if } p \in [1,2), \\ \exp^{(2)}(\widetilde{X}_s + h\epsilon_i 1_{\{s \geqslant \mu_t\}}, \widetilde{A}_s + \frac{1}{2}[X_s, h\epsilon_i 1_{\{s \geqslant \mu_t\}}]) & \text{if } p \in [2,3), \end{cases}$$

where $\exp^{(1)}$, $\exp^{(2)}$ are given in equation (1.1.2) and

$$\widetilde{X} : [0,1] \to \mathbb{R}^d, \qquad \widetilde{A} : [0,1] \to [\mathbb{R}^d, \mathbb{R}^d]$$

are the continuous paths determining $\widetilde{\mathbf{X}}$ (see Remark 1.2.2). Notice that the above expression for $p \in [2,3)$ follows by an application of the Campbell–Baker–Hausdorff formula (see Theorem 7.24 in Friz and Victoir (2010)).

(ii) In (5.2.3) the Marcus transformation of the path $\mathbf{X}$ is computed before the vertical perturbation. This is, in fact, a key aspect of this definition. On the one hand, it allows the preservation of the Marcus property of the functionals also at the level of the (functional) derivative (see Proposition 5.2.16). On the other hand, if the original path $\mathbf{X}$ admits a jump a time t, the Marcus property of the functional F allows to interpret the derivative in (5.2.3) as a derivative involving a delayed perturbation of the original path:

$$\mathbf{X}^t + h\epsilon_i 1_{\{\cdot \geqslant t+\varepsilon\}}, \tag{5.2.4}$$

for some $\varepsilon > 0$ independent on the specific pair employed for computing $\widetilde{\mathbf{X}}$. Here we assume for notational convenience that $p \in [1,2)$ and identity $G^1(\mathbb{R}^d)$ with $\mathbb{R}^d$. Indeed, assume $\mathbf{X}$ admits a jump only at time t. Then, a direct computation (which is nothing else than two consecutive applications of a Marcus transformation, first to the path $\mathbf{X}$ and then to the perturbed path $\widetilde{\mathbf{X}} + h\epsilon_i 1_{\{\cdot \geqslant \mu_t\}}$) shows that (5.2.3) is calculated by computing

$$\frac{d}{dh} F(t + \delta_2 + \delta_1, \mathbf{Y}^{[i]}(h))|_{h=0},$$

for $\mathbf{Y}^{[i]}(h)$ defined via (given the non-anticipative nature of the functional F, we specify only its values up to time $t + \delta_2 + \delta_1$)

$$\mathbf{Y}^{[i]}(h)_s := \begin{cases} \mathbf{X}_s & \text{if } s \in [0,t], \\ \mathbf{X}_{t-} + \frac{s-t}{\delta_2} \Delta \mathbf{X}_t & \text{if } s \in [t, t+\delta_2], \\ \mathbf{X}_{t-} + \Delta \mathbf{X}_t + \frac{s-t-\delta_1}{\delta_1} h\epsilon_i & \text{if } s \in [t+\delta_2, t+\delta_2+\delta_1], \end{cases}$$

for some $\delta_2, \delta_1 > 0$ such that $t + \delta_2 + \delta_1 < 1$, $\mu_t = t + \delta_2$. Since the path $\mathbf{Y}^{[i]}(h)$ stopped at time $t + \delta_2 + \delta_1$ corresponds to the Marcus transformation of the path in (5.2.4), we can interpret the notion of the vertical derivative in (5.2.3) as a derivative involving a delayed perturbation of the original path. This aspect affects in particular the higher order vertical derivatives where a jump at the current time always occurs (see Remark 5.2.23 Example 5.2.22).

(iii) The value of the vertical derivatives of $F \in \mathcal{M}^1_{[p]}$ at some $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$ is independent of the particular Marcus-transformed path. Indeed, reasoning as in Remark 5.2.6, we see that for every fixed $\xi \in \mathbb{R}^d$,

$$\begin{aligned} &F\big(\psi_{\tilde{R}}^{-1}(\tau_{\mathbf{X},\tilde{R}}(t)), \widetilde{\mathbf{Z}} \otimes \exp^{([p])}(\xi 1_{\{\cdot \geqslant \psi_{\tilde{R}}^{-1}(\tau_{\mathbf{X},\tilde{R}}(t))\}})\big) \\ &\qquad = F\big(\phi(\psi_{\tilde{R}}^{-1}(\tau_{\mathbf{X},\tilde{R}}(t))), \mathbf{Z} \otimes \exp^{([p])}(\xi 1_{\{\cdot \geqslant \psi_{R}^{-1}(\tau_{\mathbf{X},R}(t))\}})\big) \\ &\qquad = F\big(\psi_{R}^{-1}(\tau_{\mathbf{X},R}(t)), \mathbf{Z} \otimes \exp^{([p])}(\xi 1_{\{\cdot \geqslant \psi_{R}^{-1}(\tau_{\mathbf{X},R}(t))\}})\big), \end{aligned}$$

where $\mathbf{Z}$, $\widetilde{\mathbf{Z}}$ denote the Marcus-transformations of $\mathbf{X}$ associated with (R, ψ_R) and $(\widetilde{R}, \psi_{\widetilde{R}})$, respectively, and the last equality follows from Proposition 5.2.11.

(iv) If $p \in [1, 2)$ the notion of vertical derivatives given in Definition 5.2.14 corresponds to the notion of vertical derivatives introduced in Definition 8 in Cont and Fournie (2010a) (see also Dupire (2019)) when evaluated at a continuous path.

Similarly, notice that if for $p \in [2, 3)$ a non-anticipative path functional depends only on $\pi_1(\mathbf{X})$ for all $(t, \mathbf{X})$, then (5.2.3) matches again the notion of vertical derivative introduced in Definition 8 in Cont and Fournie (2010a) evaluated at continuous paths.

For $F \in \mathcal{M}^1_{[p]}$, it is easy to verify that the functional given by

$$\begin{aligned} [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) &\to \mathbb{R} \\ (t, \mathbf{X}) &\longmapsto \frac{d}{dh} F(\mu_t, \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant \mu_t\}})|_{h=0}, \end{aligned} \tag{5.2.5}$$

inherits the property of being a non-anticipative Marcus canonical path functional, for all $i = 1, \ldots, d$. For clarity of exposition, we explicitly restate this property in the following proposition.

Proposition 5.2.16. Let $F \in \mathcal{M}^1_{[p]}$. Then the path functional given in equation (5.2.5) is a non-anticipative Marcus canonical path functional.

Proof. Fix $(t, \mathbf{X}) \in [0, 1] \times C([0, 1], G^{[p]}(\mathbb{R}^d))$, ϕ a time-reparametrization and $\xi \in \mathbb{R}^d$. Since $F \in \mathcal{M}^1_{[p]} \subset \mathcal{M}^0_{[p]}$, by Proposition 5.2.11 we get that

$$\begin{aligned} F(\phi(t), \mathbf{X} \otimes \exp^{([p])}(\xi) 1_{\{\cdot \geqslant \phi(t)\}}) =& F(\phi(t), \mathbf{X}_{\phi \circ \phi^{-1}} \otimes \exp^{([p])}(\xi) 1_{\{\phi^{-1}(\cdot) \geqslant t\}}) \\ =& F(t, \mathbf{X}_\phi \otimes \exp^{([p])}(\xi) 1_{\{\cdot \geqslant t\}}), \end{aligned}$$

which implies condition (i) in Definition 5.2.4. Condition (ii) follows by definition of the functional given in equation (5.2.5) □

Remark 5.2.17. Let $F \in \mathcal{M}^0_{[p]}$ and assume that for all $(t, \mathbf{X}) \in [0, 1] \times D^p([0, 1], G^{[p]}(\mathbb{R}^d))$ and some $i = 1, \ldots, d$, the map

$$\mathbb{R} \ni h \mapsto F(t, \mathbf{X} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant t\}})$$

is differentiable at 0. Then, the non-anticipative functional defined via

$$\begin{aligned} [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d)) &\to \mathbb{R} \\ (t, \mathbf{X}) &\longmapsto \frac{d}{dh} F(t, \mathbf{X} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant t\}})|_{h=0}, \end{aligned} \tag{5.2.6}$$

is not Marcus canonical in general. Indeed, consider the path functional introduced in Example 5.2.8 (iii) and denote it by F. Then, a direct computation shows that for all $(t, \mathbf{X}) \in [0,1] \times D^1([0,1], G^1(\mathbb{R}^2))$,

$$\frac{d}{dh} F(t, \mathbf{X} \otimes \exp^{(1)}(h\epsilon_1) 1_{\{\cdot \geqslant t\}})|_{h=0} = \frac{1}{2} \Delta \mathbf{X}_t^2,$$

implying that the functional in (5.2.6) is not Marcus canonical (see the discussion in Remark 5.2.9). Since the analysis of the main results of the chapter (see Sections 5.4 and 5.5) rely on this property, we have chosen the notion vertical differentiability as in Definition 5.2.14 with the Marcus transformed path. Notice that adopting this point of view, yields

$$\frac{d}{dh} F(t, \widetilde{\mathbf{X}} \otimes \exp^{(1)}(h\epsilon_1) 1_{\{\cdot \geqslant t\}})|_{h=0} = 0,$$

for all $(t, \mathbf{X}) \in [0,1] \times D^1([0,1], G^1(\mathbb{R}^2))$ as the Marcus transformed path $\widetilde{\mathbf{X}}$ never has a jump component.

Given the results of Proposition 5.2.16 the next definition follows naturally.

Definition 5.2.18. For all $i = 1, \ldots, d$, we define the differential operators

$$\begin{aligned} U^i : \mathcal{M}^1_{[p]} &\to \mathcal{M}^0_{[p]} \\ F &\mapsto U^i(F), \end{aligned}$$

where for all $F \in \mathcal{M}^1_{[p]}$, $U^i(F)$ denotes the non-anticipative Marcus canonical path functional such that for all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$,

$$U^i(F)(t, \mathbf{X}) := \frac{d}{dh} F(\mu_t, \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant \mu_t\}})|_{h=0},$$

for some Marcus transformed path $\widetilde{\mathbf{X}}$ and $\mu_t \in [0,1]$ given in Notation 5.2.7.

Vertical derivatives of higher order We introduce the notion of higher-order vertical derivatives.

Definition 5.2.19. Let $F \in \mathcal{M}^0_{[p]}$ and $K \in \mathbb{N}$. We say that F is K times vertically differentiable if for all $l = 1, \ldots, K$ and all $(i_1, \ldots, i_{l-1}) \in \{1, \ldots, d\}^l$, the path functionals $U^{i_{l-1}} \ldots U^{i_1} F$, defined recursively by

$$\begin{aligned} &U^{i_0} F := F, && \text{for } l = 1, \\ &U^{i_{l-1}} \ldots U^{i_1} F := U^{i_{l-1}}(U^{i_{l-2}} \ldots U^{i_1} F), && \text{for } l = 2, \ldots, K, \end{aligned}$$

are vertically differentiable at all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$.

In this case, we call

$$U^{i_l}\dots U^{i_1}F(t,\mathbf{X}) := U^{i_l}(U^{i_{l-1}}\dots U^{i_1}F)(t,\mathbf{X}). \tag{5.2.7}$$

vertical derivative of order l of F at $(t,\mathbf{X})$ in the directions $(i_1,\dots,i_l)$. We denote the space of such functionals by $\mathcal{M}^K_{[p]}$.

Notation 5.2.20. In the following, for $K \in \mathbb{N}$, $F \in \mathcal{M}^K_{[p]}$, $l = 1,\dots,K$, we let $\nabla^l F$ denote the $(\mathbb{R}^d)^{\otimes l}$-valued the path functional such that for all $|I| = l$,

$$\nabla^l F(t,\mathbf{X})_I := U^{i_l}\dots U^{i_1}(F)(t,\mathbf{X}),$$

for all $(t,\mathbf{X}) \in [0,1] \times D([0,1],G^{[p]}(\mathbb{R}^d))$. Notice that $\nabla^l F$ is a $(\mathbb{R}^d)^{\otimes l}$-valued path functional whose components are in $\mathcal{M}^{K-l}_{[p]}$. For notational convenience, we also write $\nabla^0 F := F$ and $\nabla F := \nabla^1 F$.

Remark 5.2.21. (i) Notice that the computation of the higher-order vertical derivatives requires the application of iterative Marcus transformations. Since by definition the higher order derivatives are computed at a jump time, by Remark 5.2.15(ii) these calculations involve in fact delayed perturbations. This results in a differential calculus for path functionals that allows for non-commutative derivation orders (see Example 5.2.22 and Remark 5.2.23).

(ii) By definition, the computation of the higher-order vertical derivatives can be done by considering the following iterative procedure. For simplicity, we here consider càdlàg paths with values in $G^1(\mathbb{R}^d)$ and identify $G^1(\mathbb{R}^d)$ with $\mathbb{R}^d$. Let $F : [0,1] \times D([0,1],G^1(\mathbb{R}^d)) \to \mathbb{R}$ and assume $F \in \mathcal{M}^2_1$.

Set

$$\mathbf{X}^{[0]} := \mathbf{X}, \qquad \mathbf{Y}^{[0]} := \widetilde{\mathbf{X}^{[0]}}, \quad \text{and} \quad \mu^{[0]}_t := \psi^{-1}_{R_0}(\tau_{\mathbf{X}^{[0]},R_0}(t)),$$

for some pair (R_0,ψ_{R_0}),

$$\begin{aligned}
\mathbf{X}^{[i_2]}(h_2) &:= \mathbf{Y}^{[0]} + h_2\epsilon_{i_2}1_{\{\cdot \geqslant \mu^{[0]}_t\}},\\
\mathbf{Y}^{[i_2]}(h_2) &:= \widetilde{\mathbf{X}^{[i_2]}(h_2)},\\
\mu^{[i_2]}_t(h_2) &:= \psi^{-1}_{R_2}(\tau_{\mathbf{X}^{[i_2]}(h_2),R_2}(\mu^{[0]}_t)),
\end{aligned}$$

for some $h_2 \in \mathbb{R}$ and pair (R_2,ψ_{R_2}) that might depend on h_2, and finally

$$\begin{aligned}
\mathbf{X}^{[i_2,i_1]}(h_2,h_1) &:= \mathbf{Y}^{[i_2]}(h_2) + h_1\epsilon_{i_1}1_{\{\cdot \geqslant \mu^{[i_2]}_t(h_2)\}},\\
\mathbf{Y}^{[i_2,i_1]}(h_2,h_1) &:= \widetilde{\mathbf{X}^{[i_2,i_1]}(h_2,h_1)},\\
\mu^{[i_2,i_1]}_t(h_2,h_1) &:= \psi^{-1}_{R_1}(\tau_{\mathbf{X}^{[i_2,i_1]}(h_2,h_1),R_1}(\mu^{[i_2]}_t(h_2))),
\end{aligned} \tag{5.2.8}$$

for some $h_1 \in \mathbb{R}$, and pair (R_1, ψ_{R_1}) that might depend on h_1. For notational convenience, we set $\mu_t^{[i_2]} := \mu_t^{[i_2]}(h_2)$ and $\mu_t^{[i_2,i_1]} := \mu_t^{[i_2,i_1]}(h_2, h_1)$.

Then, for $i_1 = 1, \ldots, d$,

$$U^{i_1} F(\mu_t^{[i_2]}, \mathbf{Y}^{[i_2]}(h_2)) = \frac{d}{dh_1} F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1))|_{h_1=0}, \tag{5.2.9}$$

and in particular, $U^{i_1} F(t, \mathbf{X}) = U^{i_1} F(\mu_t^{[i_2]}, \mathbf{Y}^{[i_2]}(0))$.

For $i_1, i_2 = 1, \ldots, d$, it holds that

$$\begin{aligned} U^{i_2} U^{i_1} F(t, \mathbf{X}) =& \frac{d^2}{dh_1 dh_2} F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1))|_{h_1=h_2=0} \\ =& \frac{d}{dh_2} \Big(\frac{d}{dh_1} F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1))|_{h_1=0} \Big)|_{h_2=0} \\ =& \frac{d}{dh_2} \Big(U^{i_1} F(\mu_t^{[i_2]}, \mathbf{Y}^{[i_2]}(h_2)) \Big)|_{h_2=0}. \end{aligned} \tag{5.2.10}$$

Example 5.2.22. Let us consider again the non-anticipative Marcus canonical path functional introduced in Example 5.2.8(iii). In the following, we make explicit the iterative computations described in Remark 5.2.21(ii) and show that for all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^1(\mathbb{R}^2))$, one gets that

$$U^2 U^2 F(t, \mathbf{X}) = U^1 U^1 F(t, \mathbf{X}) = U^2 U^1 F(t, \mathbf{X}) = U^1 F(t, \mathbf{X}) = 0,$$

whereas,

$$U^2 F(t, \mathbf{X}) = \mathbf{X}_t^1 \quad \text{and} \quad U^1 U^2 F(t, \mathbf{X}) = 1.$$

We identify $G^1(\mathbb{R}^2)$ with $\mathbb{R}^2$. Fix $(t, \mathbf{X}) \in C^1([0,1], G^1(\mathbb{R}^2))$, for simplicity.

Following the previous notation, we set

(i) $\mathbf{X}^{[0]} := \mathbf{X}, \qquad \mathbf{Y}^{[0]} := \mathbf{X}^{[0]}, \quad \text{and} \quad \mu_t^{[0]} := t.$

(ii) For $h_2 \in \mathbb{R}$, $h_2 \neq 0$, and $i_2 = 1, 2$, set

$$\mathbf{X}^{[i_2]}(h_2) := \mathbf{X} + h_2 \epsilon_{i_2} 1_{\{\cdot \geqslant t\}},$$

and consider the pair (R_2, ψ_{R_2}) given by $R_2 := \delta_2$, for some $\delta_2 > 0$ such that (without loss of generality) $t + \delta_2 < 1$, and $\psi_{R_2} : [0,1] \to [0, 1+\delta_2]$ an increasing continuous bijection such that

$$\psi_{R_2}([0, t+\delta_2]) = [0, t+\delta_2] \quad \text{and} \quad \psi_{R_2}([t+\delta_2, 1]) = [t+\delta_2, 1+\delta_2].$$

Define

$$\mathbf{Z}_s^{[i_2]} := \begin{cases} \mathbf{X}_s, & \text{if } s \in [0,t], \\ \mathbf{X}_t + \frac{s-t}{\delta_2} h_2 \epsilon_{i_2} & \text{if } s \in [t, t+\delta_2], \\ \mathbf{X}_{s-\delta_2} + h_2 \epsilon_{i_2} & \text{if } s \in [t+\delta_2, 1+\delta_2]. \end{cases}$$

and $\mathbf{Y}^{[i_2]}(h_2) := \mathbf{Z}^{[i_2]}(h_2)_{\psi_{R_2}}$. Notice that

$$\mathbf{Y}^{[i_2]}(h_2) = \widetilde{\mathbf{X}^{[i_2]}(h_2)}$$

with respect to the pair (R_2, ψ_{R_2}) and

$$\mu_t^{[i_2]} := \psi_{R_2}^{-1}(\tau_{\mathbf{X}^{[i_2]}(h_2),R_2}(t)) = t + \delta_2.$$

For $h_2 = 0$, set simply $\psi_{R_2}(t) := t$ and $\mathbf{Y}^{[i_2)]}(h_2) := \mathbf{X}$.

(iii) Similarly, let $h_1 \in \mathbb{R}$, $h_1 \neq 0$, $i_1 = 1, 2$, and set

$$\mathbf{X}^{[i_2,i_1]}(h_2, h_1) := \mathbf{Y}^{[i_2]}(h_2) + h_1\epsilon_{i_1} 1_{\{\cdot \geqslant \mu_t^{[i_2]}(h_2)\}}.$$

Consider the pair (R_1, ψ_{R_1}) given by $R_1 := \delta_1$, for some $\delta_1 > 0$ such that (without loss of generality) $t + \delta_2 + \delta_1 < 1$, and $\psi_{R_1} : [0, 1] \to [0, 1 + \delta_1]$ such that

$$\psi_{R_1}([0, t+\delta_2+\delta_1]) = [0, t+\delta_2+\delta_1] \quad \text{and} \quad \psi_{R_1}([t+\delta_2+\delta_1, 1]) = [t+\delta_2+\delta_1, 1+\delta_1].$$

Define

$$\mathbf{Z}^{[i_2,i_1]}(h_2, h_1)_s := \begin{cases} \mathbf{Y}^{[i_2]}(h_2)_s & \text{if } s \in [0, t + \delta_2], \\ \mathbf{Y}^{[i_2]}(h_2)_{t+\delta_2} + \frac{s-t-\delta_2}{\delta_1} h_1\epsilon_{i_1} & \text{if } s \in [t + \delta_2, t + \delta_2 + \delta_1], \\ \mathbf{Y}^{[i_2]}(h_2)_{s-\delta_1} + h_1\epsilon_{i_1} & \text{if } s \in [t + \delta_2 + \delta_1, 1 + \delta_2 + \delta_1], \end{cases}$$

and $\mathbf{Y}^{[i_2,i_1]}(h_2, h_1) := \mathbf{Z}^{[i_2,i_1]}(h_2, h_1)_{\psi_{R_1}}$.

Notice that

$$\mathbf{Y}^{[i_2,i_1]}(h_2, h_1) = \widetilde{\mathbf{X}^{[i_2,i_1]}(h_2, h_1)}$$

with respect to the pair (R_1, ψ_{R_1}) and

$$\mu_t^{[i_2,i_1]} := \psi_{R_1}^{-1}(\tau_{\mathbf{X}^{[i_2,i_1]}(h_2,h_1),R_1}(t + \delta_2)) = t + \delta_2 + \delta_1.$$

For $h_1 = 0$, set $\psi_{R_1}(t) := t$ and $\mathbf{Y}^{[i_2,i_1)]}(h_2, h_1) := \mathbf{Y}^{[i_2]}(h_2)$.

Let us now compute $F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1))$. By definition of F

$$F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1)) = \int_0^{\mu_t^{[i_2,i_1]}} \mathbf{Y}_s^{[i_2,i_1],1}(h_2, h_1) d\mathbf{Y}_s^{[i_2,i_1],2}(h_2, h_1). \tag{5.2.11}$$

Since $\mu_t^{[i_2,i_1]} = t + \delta_2 + \delta_1$, splitting the integral in (5.2.11) over the intervals $[0, t]$, $[t, t + \delta_2]$, $[t + \delta_2, t + \delta_2 + \delta_1]$, by definition of $\mathbf{Y}^{[i_2,i_1]}(h_2, h_1)$, we then get

$$F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2, h_1)) = \int_0^t \mathbf{X}_s^1 d\mathbf{X}_s^1 + h_2\mathbf{X}_t^1 1_{\{i_1=2\}} + \big(h_1\mathbf{X}_t^1 + h_1h_2 1_{\{i_2=1\}}\big) 1_{\{i_1=2\}}.$$

Since by (5.2.9) and (5.2.10),

$$U^{i_1}F(t,\mathbf{X}) = \Big(\frac{d}{dh_1}F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(0,h_1))|_{h_1=0}\Big)$$

and

$$U^{i_2}U^{i_1}F(t,\mathbf{X}) = \frac{d}{dh_2}\Big(\frac{d}{dh_1}F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2,h_1))|_{h_1=0}\Big)|_{h_2=0},$$

the claim follows. Notice that if $\mathbf{X}$ is càdlàg , one must simply apply the same reasoning to a Marcus transformed path $\widetilde{\mathbf{X}}$ of $\mathbf{X}$.

Remark 5.2.23. As already mentioned above, this example shows that the notion of vertical derivatives of higher order introduced in Definition 5.2.19 establishes a differential calculus for path functionals that allows for non-commutative derivation orders. This is in contrast to the setup in Dupire (2019), Cont and Fournie (2010a), Cont et al. (2016) (see e.g., page 133 in Cont et al. (2016)).

This different behavior can be explained by the following considerations. For a non-anticipative path functional F, the computation of the second-order vertical derivatives at some $(t,\mathbf{X})$ in the sense of Dupire (2019), Cont and Fournie (2010a) is obtained by computing the differential quotient at the path $\mathbf{X}^t + h_1\epsilon_{i_1}1_{\cdot\geqslant t} + h_2\epsilon_{i_2}1_{\cdot\geqslant t}$ (see Definition 9 in Cont and Fournie (2013)). In the present setting, the Marcus property of the functionals (see (ii) in Definition 5.2.4) and the definition of the vertical derivatives (see Definition 5.2.19) yield that the second order derivatives are calculated by computing the differential quotient at the path $\mathbf{Y}^{[i_2,i_1)]}(h_2,h_1)$, introduced in (5.2.8).

More precisely, the computation of the second-order vertical derivatives considered by Dupire (2019) Cont and Fournie (2010a) explicitly reads as

$$\partial_{i_1}\partial_{i_2}F(t,\mathbf{X}) = \frac{d^2}{dh_1dh_2}F(t,\mathbf{X} + (h_2\epsilon_{i_2} + h_1\epsilon_{i_1})1_{\{\cdot\geqslant t\}})|_{h_1=h_2=0}, \tag{5.2.12}$$

for $i_1, i_2 = 1,\ldots,d$, while in the present framework,

$$U^{i_1}U^{i_2}F(t,\mathbf{X}) = \frac{d^2}{dh_1dh_2}F(\mu_t^{[i_2,i_1]}, \mathbf{Y}^{[i_2,i_1]}(h_2,h_1))|_{h_1=h_2=0}.$$

Note that $\partial_{i_1}\partial_{i_2}F(t,\mathbf{X}) = \partial^2_{\xi_{i_1}\xi_{i_2}}G(\xi)|_{\xi=0}$, for $G(\xi) := F(t,\mathbf{X}+\xi 1_{\{\cdot\geqslant t\}})$, $\xi\in\mathbb{R}^d$.

To better understand the parametrization of the path $\mathbf{Y}^{[i_2,i_1]}(h_2,h_1)$, consider again Example 5.2.22 and observe that by construction for all $t\in[0,1]$, $s\in[0,\mu_t^{[i_2,i_1]}]$,

$$\mathbf{Y}^{[i_2,i_1]}(h_2,h_1)_s := \begin{cases} \mathbf{X}_s & \text{if } s\in[0,t], \\ \mathbf{X}_t + \frac{s-t}{\delta_2}h_2\epsilon_{i_2} & \text{if } s\in[t,t+\delta_2], \\ \mathbf{X}_t + h_2\epsilon_{i_2} + \frac{s-t-\delta_1}{\delta_1}h_1\epsilon_{i_1} & \text{if } s\in[t+\delta_2, t+\delta_2+\delta_1], \end{cases} \tag{5.2.13}$$

implying that the stopped path $\mathbf{Y}^{[i_2,i_1]}(h_2,h_1)^{\mu_t^{[i_2,i_1]}}$ is equal to $\mathbf{X}$ until time t and then from t onwards is piecewise linear. This however is nothing else than the Marcus tranformation of a càdlàg path (with delayed perturbations) given by

$$\mathbf{X}^t + h_2\epsilon_{i_2}1_{\{\cdot\geqslant t\}} + h_1\epsilon_{i_1}1_{\{\cdot\geqslant t+\varepsilon\}},$$

for some $\varepsilon > 0$. Hence, our notion of vertical derivative involving the Marcus transformation can be interpreted as a derivative involving a delayed vertical perturbation of the original path. Furthermore, the delay $\varepsilon > 0$ is independent of the specific parametrization of the path (5.2.13).

This, in turn, induces the non-commutative behavior of the second-order derivatives, which is in contrast to the second-order vertical derivatives considered by Dupire (2019) Cont and Fournie (2010a), where the perturbation occurs at the same time.

To conclude, we discuss again the functional in Example 5.2.22 which is defined via $F(t,\mathbf{X}) := \int_0^t \mathbf{X}^1_{s-} d\mathbf{X}^2_s + \frac{1}{2}\Delta\mathbf{X}^1_s\Delta\mathbf{X}^2_s$, for all $(t,\mathbf{X}) \in D^1([0,1],G^1(\mathbb{R}^2))$. A direct computation shows that for $i_2, i_2 = 1,2$, $i_1 \neq i_2$,

$$\partial_{i_1}\partial_{i_2}F(t,\mathbf{X}) = \frac{1}{2},$$

which is in line with (5.2.12) and Schwarz's theorem, while we have already proved that $U^1U^2F(t,\mathbf{X}) = 1$ and $U^2U^1F(t,\mathbf{X}) = 0$.

This non-commutative framework is particularly relevant when dealing with weakly geometric p-rough paths, for $p \in [2,3)$, where the second order derivatives appear in the Itô-formula (see Sections 5.4.2, 5.5.2). But, of course, it also plays a significant role for finite variation paths in view of the Taylor expansions (see Theorem 5.5.1).

5.2.3 Var-continuous path functionals

In this section, we introduce the required regularity conditions for the path functionals to derive the results presented in the subsequent sections. Unless otherwise specified, we let $p \in [1,3)$ and for $q \geqslant p$, we denote by d_q the q-variation pseudometric on the space of càdlàg paths as defined in Section 1.2.1.

Definition 5.2.24. Fix $m \in \mathbb{N}_0$. We say that a path functional $F \in ((\mathcal{M}^0_{[p]})^d)^{\otimes m}$ is a $[p,[p]+1)$-*var-continuous* if for every $(\mathbf{X}^n,\mathbf{X})_{n\in\mathbb{N}} \subset C^p([0,1],G^{[p]}(\mathbb{R}^d))$ such that for some $q \in [p,[p]+1)$

$$\lim_{n\to\infty} d_q(\mathbf{X}^n,\mathbf{X}) = 0,$$

it holds that

$$\lim_{n\to\infty} d_q(F(\cdot,\mathbf{X}^n),F(\cdot,\mathbf{X})) = 0.$$

Recall the notion of remainder functional given in Definition 5.2.12.

Definition 5.2.25. Fix $m \in \mathbb{N}_0$. For $p \in [2,3)$ and $F \in ((\mathcal{M}_2^0)^d)^{\otimes m}$, $F' \in ((\mathcal{M}_2^0)^d)^{\otimes m+1}$, we say that a remainder path functionals $R^{F,F'}$ is a $[p,3)$*-var-continuous* if for every $(\mathbf{X}^n, \mathbf{X})_{n\in\mathbb{N}} \subset C^p([0,1], G^2(\mathbb{R}^d))$ such that for some $q \in [p,3)$

$$\lim_{n\to\infty} d_q(\mathbf{X}^n, \mathbf{X}) = 0,$$

it holds that

$$\lim_{n\to\infty} d_{q/2}(R((\cdot,\cdot), \mathbf{X}^n), R((\cdot,\cdot), \mathbf{X})) = 0.$$

5.2.4 Linear functionals of the signature as non-anticipative Marcus canonical path functionals

This section is devoted to the study of linear functionals of the signature and their properties. To be consistent with the notation used in the rest of the chapter, for some $\mathbf{u} \in T(\mathbb{R}^d)$, we denote by $F^{\mathbf{u}}$ the linear functional of the signature given by

$$[0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^d)) \ni (t, \mathbf{X}) \mapsto F^{\mathbf{u}}(t, \mathbf{X}) := \langle \mathbf{u}, \mathbb{X}_t \rangle. \tag{5.2.14}$$

The next proposition lists the most important properties of these functionals.

Proposition 5.2.26. Let $F^{\mathbf{u}} : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$ be a linear functionals of the signature, for some $\mathbf{u} \in T(\mathbb{R}^d)$. Then,

(i) $F^{\mathbf{u}}$ is a non-anticipative Marcus canonical path functional.

(ii) $F^{\mathbf{u}}$ is infinitely many times vertically differentiable and for all $k \in \mathbb{N}$,

$$\nabla^k F^{\mathbf{u}}(t, \mathbf{X}) = \langle \mathbf{u}^{(k)}, \mathbb{X}_t \rangle, \tag{5.2.15}$$

for all $(t, \mathbf{X}) \in [0,1] \times D([0,1], G^{[p]}(\mathbb{R}^d))$.

(iii) $F^{\mathbf{u}}$ is $[[p],[p]+1)$-var continuous.

Proof. (i): It follows by the definition of the signature of a weakly geometric càdlàg rough path (see Section 1.3.3).

(ii): Fix $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and let $\widetilde{\mathbf{X}}$ denote its Marcus transformed path with respect to some pair (R, ψ_R), with signature $\widetilde{\mathbb{X}}$, and for $t \in [0,1]$, let $\mu_t \in [0,1]$ be defined as in Notation 5.2.7. Fix $i = 1, \ldots, d$, $h \in \mathbb{R}$, and consider the vertically perturbed path

$$\mathbf{Y} := \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i) 1_{\{\cdot \geqslant \mu_t\}} \in D^p([0,1], G^{[p]}(\mathbb{R}^d)). \tag{5.2.16}$$

By Chen's relation (see Corollary 1.3.19) and the minimal jump extension property (see equation (1.3.7)), it holds that the signature of $\mathbf{Y}$, that we denote by $\mathbb{Y}$, at time μ_t reads as follows:

$$\mathbb{Y}_{\mu_t} = \mathbb{Y}_{\mu_t^-} \otimes \Delta\mathbb{Y}_{\mu_t} = \widetilde{\mathbb{X}}_{\mu_t} \otimes \exp(h\epsilon_i).$$

Notice that we use that the path $s \mapsto \widetilde{\mathbb{X}}_s$ has continuous components. We then get

$$\begin{aligned} U^i F^{\mathbf{u}}(t,\mathbf{X}) :=& \frac{d}{dh} F^{\mathbf{u}}(\mu_t, \widetilde{\mathbf{X}} \otimes \exp^{([p])}(h\epsilon_i)1_{\{\cdot \geqslant \mu_t\}})|_{h=0} \\ =& \frac{d}{dh} \langle \mathbf{u}, \widetilde{\mathbb{X}}_{\mu_t} \otimes \exp(h\epsilon_i)\rangle|_{h=0}. \end{aligned}$$

An explicit computation yields $U^i F^{\mathbf{u}}(t,\mathbf{X}) = \langle \mathbf{u}^{(1)}_{(i)}, \widetilde{\mathbb{X}}_{\mu_t}\rangle$ for $\mathbf{u}^{(1)}_{(i)} \in T(\mathbb{R}^d)$ introduced in equation (1.1.7). Moreover, by definition of the signature of weakly geometric càdlàg rough paths, $\langle \mathbf{u}^{(1)}_{(i)}, \widetilde{\mathbb{X}}_{\mu_t}\rangle = \langle \mathbf{u}^{(1)}_{(i)}, \mathbb{X}_t\rangle$. Thus, computing the vertical derivative in all the directions $j = 1, \dots, d$, we get $\nabla F^{\mathbf{u}}(t,\mathbf{X}) = \langle \mathbf{u}^{(1)}, \mathbb{X}_t\rangle$, and more generally, by iterating the same reasoning, for all $k \in \mathbb{N}$, $\nabla^k F^{\mathbf{u}}(t,\mathbf{X}) = \langle \mathbf{u}^{(k)}, \mathbb{X}_t\rangle$.

(iii): It follows by Corollary 10.28 in Friz and Victoir (2010). □

Remark 5.2.27. (i) Notice that (ii) in Proposition 5.2.26 could be deduced also directly from equations (1.3.12), (1.3.13).

(ii) Proposition 5.2.26 can be easily extended to multivalued linear path functionals. In particular recall the shifts introduced in (1.1.7) a fix $\mathbf{u} \in T(\mathbb{R}^d)$. Then, the path functional $F^{\mathbf{u}^{(j)}}$ defined via

$$F^{\mathbf{u}^{(j)}}(t,\mathbf{X}) := \langle \mathbf{u}^{(j)}, \mathbb{X}_t\rangle, \tag{5.2.17}$$

for all $(t,\mathbf{X}) \in [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^d))$, $j \in \mathbb{N}$, is $[[p],[p]+1)$-var continuous and its components belong to $\mathcal{M}^K_{[p]}$ for all $K \in \mathbb{N}$.

Now, as a direct combination of Proposition 5.2.26 and equations (1.3.12), (1.3.13) the functional Itô-formula for linear functionals of the signature holds true.

Lemma 5.2.28. Let $F^{\mathbf{u}} : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^d)) \to \mathbb{R}$ be a linear functional of the signature, for some $\mathbf{u} \in T(\mathbb{R}^d)$. Then, the following functional Itô-formulas hold true.

(i) if $p \in [1,2)$,

$$\begin{aligned} F^{\mathbf{u}}(t,\mathbf{X}) =& F^{\mathbf{u}}(0,\mathbf{X}) + \int_0^t \nabla F^{\mathbf{u}}(s^-,\mathbf{X}) d\mathbf{X}_s \\ &+ \sum_{0<s\leqslant \cdot} F^{\mathbf{u}}(s,\mathbf{X}) - F^{\mathbf{u}}(s^-,\mathbf{X}) - \Delta F^{\mathbf{u}}(s^-,\mathbf{X})\Delta \mathbf{X}_s, \end{aligned}$$

where the integral term is a Young integral of $\nabla F^{\mathbf{u}}(\cdot,\mathbf{X}) \in D^p([0,1], \mathbb{R}^{d+1})$ with respect to $\mathbf{X}$;

(ii) if $p \in [2,3)$,

$$\begin{aligned} F^{\mathbf{u}}(t,\mathbf{X}) =& F^{\mathbf{u}}(0,\mathbf{X}) + \int_0^t F^{\mathbf{u}}(s^-,\mathbf{X}) d\mathbf{X}_s \\ &+ \sum_{0<s\leqslant \cdot} F^{\mathbf{u}}(s,\mathbf{X}) - F^{\mathbf{u}}(s^-,\mathbf{X}) - \nabla F^{\mathbf{u}}(s^-,\mathbf{X})\Delta X_s - \nabla^2 F^{\mathbf{u}}(s^-,\mathbf{X})\Delta \mathbb{X}^{(2)}_s, \end{aligned}$$

where the integral term is a rough integral of the controlled rough path

$$\left(\nabla F^{\mathbf{u}}(\cdot, \mathbf{X}), \nabla^2 F^{\mathbf{u}}(\cdot, \mathbf{X})\right) \in \mathcal{V}_{\mathbf{X}}^{p,\frac{p}{2}}$$

with respect to $\mathbf{X}$.

Remark 5.2.29. Observe that in general for each $\mathbf{u} \in T(\mathbb{R}^d)$, and t, $\langle \mathbf{u}^{(2)}, \mathbb{X}_t \rangle \in (\mathbb{R}^d)^{\otimes 2}$ is not symmetric. Therefore, a differential calculus that admits only commutative higher-order vertical derivatives would not allow the derivation of Lemma 5.2.28.

5.3 Universal approximation theorem for vertically differentiable path functionals

In this section, we exploit the approximation properties of the signature of weakly geometric càdlàg rough paths to derive a universal approximation theorem (UAT) for vertically differentiable non-anticipative Marcus-canonical path functionals. We prove that every sufficiently regular vertically differentiable path functional can be approximated uniformly in time together with its derivatives by linear functions of the signature and their derivatives. This result for path functionals is parallel to the classic Nachbin theorem, which provides an analog to the Stone–Weierstrass theorem for algebras of C^K functions on $\mathbb{R}^d$, for $K \in \mathbb{N}$ (see Nachbin (1949)).

Firstly, we introduce the set of so-called *tracking-jumps-extended paths* and then derive the main result of the section, stated in Theorem 5.3.15.

5.3.1 The subspace of the tj-extended paths

In this section, we introduce the subset of paths considered for deriving the UAT for vertically differentiable functionals, the so-called *tracking jumps-extended paths*. The distinctive feature of these paths is that they always admit a Marcus-transformed path that is time-extended (see Remark 5.3.4(ii)). This is a key aspect both for the development of the proof of Theorem 5.3.15 and the main results in Sections 5.4 and 5.5.

Definition 5.3.1. Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$. We say that a càdlàg strictly increasing piecewise linear function $\upsilon : [0,1] \to [0,1]$ such that $\upsilon_0 = 0$ and $\upsilon_1 = 1$ is a *tracking jumps-path* (tj)-path of $\mathbf{X}$ if

$$\{s \in (0,1] \ : \ \Delta \upsilon_s \neq 0\} = \{s \in (0,1] \ : \ \Delta \mathbf{X}_s \neq 0\}.$$

Remark 5.3.2. (i) Given $\mathbf{X}^p \in D([0,1], G^{[p]}(\mathbb{R}^d))$ it is always possible to build (non-uniquely) a tracking-jumps path associated with it. If instead $\mathbf{X}$ is continuous, the tracking-jumps path is unique and given by the identity path $\mathrm{Id}_t := t$ for all $t \in [0,1]$.

(ii) One can always build a tracking-jumps path associated with any càdlàg path with values in $\mathbb{R}^d$, regardless of its regularity. Indeed, the existence of such a path is related only to the càdlàg property of the original path.

Recall the notion of extended weakly geometric càdlàg rough path given in Definition 1.2.16. Now, we introduce the subset of the space of weakly geometric càdlàg rough paths consisting of tj-extended paths.

Definition 5.3.3. We say that $\widehat{\mathbf{X}} \in D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})))$ is a *tj-extended weakly geometric càdlàg p-rough path* if it is the v-extension of some $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$, for v being some of the tracking-jumps path of $\mathbf{X}$. We denote the set of such paths by $\widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ and the subset of tj-extended weakly geometric continuous p-rough path by $\widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$.

Remark 5.3.4. (i) Let $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^d))$ and observe that by Remark 5.3.2 there exists more than one tj-extension $\widehat{\mathbf{X}}$ of it. In the continuous setting instead, uniqueness holds and every tj-extended path is a time-extended path (see Definition 2.2.3).

(ii) Notice that for all $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ there always exists a Marcus transformed path $\mathbf{Z}$ of $\widehat{\mathbf{X}}$ with respect to some pair (R, ψ_R), for which $\mathbf{Z}$ is a time-extended weakly geometric continuous rough path. Indeed, since a tj path is simply a strictly increasing piecewise linear path v such that $v(0) = 0$ and $v(1) = 1$, and jumps whenever the other components of the path do, a careful choice of a pair (R, ψ_R) always allows to build a Marcus transformed path for which the transformed component v is exactly the time.

(iii) We remark that while in Chapter 2 $\widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ denotes the set of all time-extended paths, in the present chapter instead, it denotes the different set of tj-extended paths.

Notation 5.3.5. We use the index 0 to denote the tj-component of any tj-extended path $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$. This notation will be used throughout the chapter.

5.3.2 UAT for vertically differentiable path functionals

Here we present the main result of the section: Theorem 5.3.15. To this end, we shall introduce some notation and provide new definitions. Since we will be mostly concerned with paths with values in $G^{[p]}(\mathbb{R}^{d+1})$, we introduce the relevant notions by considering directly the space $\mathbb{R}^{d+1}$.

Notation 5.3.6. Let $K \in \mathbb{N}$ and $\mathfrak{g}^K(\mathbb{R}^{d+1})$ be the free step-K nilpotent Lie algebra over $\mathbb{R}^{d+1}$ (see Section 1.1.2). Recall that for every $\boldsymbol{\xi} := (0, \boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(K)}) \in \mathfrak{g}^K(\mathbb{R}^{d+1})$,

$$\boldsymbol{\xi}^{(j)} := \pi_j(\boldsymbol{\xi}) \in \underbrace{[\mathbb{R}^{d+1}, [\mathbb{R}^{d+1}, \dots [\mathbb{R}^{d+1}, \mathbb{R}^{d+1}]]]}_{(j-1) \text{ brackets}},$$

for $j = 1, \ldots, K$. Set $M := \dim(\mathfrak{g}^K(\mathbb{R}^{d+1}))$ and

$$M_j := \dim(\underbrace{[\mathbb{R}^{d+1}, [\mathbb{R}^{d+1}, \ldots [\mathbb{R}^{d+1}, \mathbb{R}^{d+1}]]]}_{(j-1) \text{ brackets}}).$$

Notice that $M = \sum_{j=1}^K M_j$. For $\beta \in \mathbb{N}_0^M$, we write $\beta = (\beta_1, \ldots, \beta_K)$, with $\beta_j \in \mathbb{N}_0^{M_j}$, and set

$$|\beta| := \sum_{j=1}^K \sum_{i=1}^{M_j} \beta_j^i, \qquad \text{and} \quad |\beta|_{\mathfrak{g}^K(\mathbb{R}^{d+1})} := \sum_{j=1}^K \sum_{i=1}^{M_j} j \beta_j^i. \tag{5.3.1}$$

Furthermore, we let $\partial_{\boldsymbol{\xi}^{(j)}}^{\beta_j}$ denote the differential operator such that

$$\partial_{\boldsymbol{\xi}^{(j)}}^{\beta_j} f(\boldsymbol{\xi}) := \frac{\partial^{|\beta_j|} f}{\partial^{\beta_j^1} \boldsymbol{\xi}^{(j),1} \ldots \partial^{\beta_j^{M_j}} \boldsymbol{\xi}^{(j),M_j}}(\boldsymbol{\xi}),$$

for $\boldsymbol{\xi}^{(j),l}$ denoting the l-th component of $\boldsymbol{\xi}^{(j)}$, $|\beta_j| := \sum_{i=1}^{M_j} \beta_j^i$, and $f : \mathcal{U} \to \mathbb{R}$ some sufficiently regular map defined on some open set $\mathcal{U} \subseteq \mathfrak{g}^K(\mathbb{R}^{d+1})$, with $\boldsymbol{\xi} \in \mathcal{U}$.

Finally, we denote by $D_{\boldsymbol{\xi}}^{\beta} := \partial_{\boldsymbol{\xi}^{(1)}}^{\beta_1} \ldots \partial_{\boldsymbol{\xi}^{(K)}}^{\beta_K}$ the operator given by a consecutive application of $\partial_{\boldsymbol{\xi}^{(1)}}^{\beta_1}, \ldots, \partial_{\boldsymbol{\xi}^{(K)}}^{\beta_K}$.

Recall that $G((\mathbb{R}^{d+1}))$ denotes the set of group-like elements, introduced in equation (1.1.6).

Assumption 5.3.7. *$F : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$ is a Marcus canonical non-anticipative path functional such that for all $(t, \mathbf{X}) \in [0,1] \times C^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$,*

$$F(t, \mathbf{X}) = g(\mathbb{X}_t), \tag{5.3.2}$$

for some $g : G((\mathbb{R}^{d+1})) \to \mathbb{R}$, and $\mathbb{X}$ denoting the signature of $\mathbf{X}$.

Remark 5.3.8. (i) By the definition of Marcus canonical path functionals and the signature of càdlàg rough paths, the equality (5.3.2) is valid in fact on the whole space $[0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$.

(ii) Let S be the map that associates to every continuous weakly geometric p-rough paths its signature at time 1. For $t \in [0,1]$, let $\widehat{C}_t^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ denote the space of time-extended paths stopped at time t. One can show that on the set

$$\bigcup_{t \in [0,1]} \widehat{C}_t^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$$

S is an injective map (see e.g., the proof of Theorem 2.2.6). Therefore the restriction of the map S to its image, which we denote by $\mathcal{S}$, is a bijection. Furthermore, every non-anticipative path functional F is such that

$$F(t, \mathbf{X}) = \widetilde{F}(\mathbf{X}^t),$$

for some $\widetilde{F} : D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ and all $(t, \mathbf{X}) \in [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$. Therefore, F always admits a representation of the form

$$F(t, \widehat{\mathbf{X}}) = \bar{g}(\widehat{\mathbb{X}}_t), \tag{5.3.3}$$

for $\bar{g} : \mathcal{S} \to \mathbb{R}$ given by $\bar{g} := \widetilde{F} \circ S^{-1}$ and every time extended continuous path $\widehat{\mathbf{X}}$. However, in general, the set $\mathcal{S} \subsetneq G((\mathbb{R}^{d+1}))$ might not be big enough for the application of the reasonings in the proofs of Lemma 5.3.9 and Theorem 5.3.15.

The proof of the following lemma is given in Appendix 5.A.1. Recall that

$$\exp(\mathbf{x}) := 1 + \sum_{k=1}^{\infty} \frac{\mathbf{x}^{\otimes k}}{k!} \in T((\mathbb{R}^{d+1}))$$

for each $\mathbf{x} \in T((\mathbb{R}^{d+1}))$ such that $\pi_0(\mathbf{x}) = 0$.

Lemma 5.3.9. Let $F \in \mathcal{M}^K_{[p]}$, for some $K \in \mathbb{N}$. Fix $\mathbf{X} \in [0,1] \times C^p([0,1], G^{[p]}(\mathbb{R}^d))$ and denote by $\mathbb{X}$ its signature. Under Assumption 5.3.7, the map

$$g^{\mathbf{X},K}(t, \boldsymbol{\xi}) : [0,1] \times \mathfrak{g}^K(\mathbb{R}^{d+1}) \to \mathbb{R} \tag{5.3.4}$$

given by

$$g^{\mathbf{X}.K}(t, \boldsymbol{\xi}) := g(\mathbb{X}_t \otimes \exp(\boldsymbol{\xi}, 0 \ldots, 0))$$

for all $(t, \boldsymbol{\xi})$ is well-defined and its derivatives at zero $D^\beta_{\boldsymbol{\xi}} g^{\mathbf{X},K}(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0}$ exist for all $\beta \in \mathbb{N}_0^M$, with $|\beta|_{\mathfrak{g}^K(\mathbb{R}^{d+1})} = K$.

For the development of the following results, we need the map $g^{\mathbf{X},K}$ to satisfy some stronger continuity conditions expressed by the next assumption.

Assumption 5.3.10. *For all $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^d))$, the map*

$$[0,1] \times \mathcal{U}(0) \ni (t, \boldsymbol{\xi}) \mapsto D^\beta_{\boldsymbol{\xi}} g^{\mathbf{X},K}(t, \boldsymbol{\xi})$$

is jointly continuos for all $\beta \in \mathbb{N}_0^M$ with $|\beta|_{\mathfrak{g}^K(\mathbb{R}^{d+1})} = K$, and $\mathcal{U}(0)$ being some open neighborhood of $0 \in \mathfrak{g}^K(\mathbb{R}^{d+1})$.

Remark 5.3.11. If Assumption 5.3.10 is verified, the map $\mathcal{U}(0) \ni \boldsymbol{\xi} \mapsto D^\beta g^{\mathbf{X},K}(t, \boldsymbol{\xi})$ does not depend on the order of the partial derivatives (see Step 2 in the proof of Theorem 5.3.15).

Definition 5.3.12. Let $F : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$. We write $F \in C^K$ if $F \in \mathcal{M}^K_{[p]}$ and Assumptions 5.3.7 and 5.3.10 are verified.

An example of path functional satisfying Definition 5.3.12 is given by the linear functionals of the signature $F^{\mathbf{u}}$, for some $\mathbf{u} \in T(\mathbb{R}^{d+1})$ (see Section 5.2.4). Then, by Proposition 5.2.26 $F^{\mathbf{u}} \in C^K$ for all $K \in \mathbb{N}$, and the map (5.3.4) explicitly reads as

$$g^{\mathbf{X},K}(t, \boldsymbol{\xi}) := \langle \mathbf{u}, \mathbb{X}_t \otimes \exp(\boldsymbol{\xi}, 0 \ldots, 0) \rangle,$$

for all $(t, \boldsymbol{\xi})$ and $\mathbf{X}$.

Remark 5.3.13. For $m \in \mathbb{N}_0$, $K \in \mathbb{N}$ and F a $(\mathbb{R}^d)^{\otimes m}$-valued path functional, we write $F \in C^K$ if its components are C^K path functionals.

Notice that if $F \in C^K$, then for all $l = 1, \ldots, K$, $\nabla^l F \in C^{K-l}$.

Notation 5.3.14. Throught the chapter, write $\|\cdot\|$ to denote the Euclidean norm on $\mathbb{R}^m$, for all $m \in \mathbb{N}$.

The proof of the following theorem is given in Appendix 5.A.2. Recall that in this chapter $\widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ indicates the set of tj-extended paths introduced in Definition 5.3.3.

Theorem 5.3.15. *Let $F : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$. Assume $F \in C^K$. Then, for all $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ there exists $(\mathbf{u}_n)_{n \in \mathbb{N}} \in T(\mathbb{R}^{d+1})$ such that*

$$\lim_{n \to \infty} \sup_{t \in [0,1]} \sum_{j=0}^{K} \|\nabla^j F(t, \widehat{\mathbf{X}}) - \nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}})\| = 0. \tag{5.3.5}$$

Remark 5.3.16. Recall from Section 5.2.4 that for $\mathbf{u} \in T(\mathbb{R}^{d+1})$, $j = 0, \ldots, K$, $(t, \widehat{\mathbf{X}})$, $\nabla^j F^{\mathbf{u}}(t, \widehat{\mathbf{X}}) = \langle \mathbf{u}^{(j)}, \widehat{\mathbb{X}}_t \rangle$. Therefore, equation (5.3.5) explicitly reads as

$$\lim_{n \to \infty} \sup_{t \in [0,1]} \sum_{j=0}^{K} \|\nabla^j F(t, \widehat{\mathbf{X}}) - \langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_t \rangle\| = 0.$$

5.4 Functional Itô-formula

5.4.1 The case $p \in [1,2)$

In this section, we present the functional Itô-formula for maps of càdlàg rough paths of finite p-variation, for $p \in [1,2)$. Recall the notion of C^2-non-anticipative Marcus canonical path functional given in Definition 5.3.12 and the one of var-continuous path functional in Definition 5.2.24.

Theorem 5.4.1. *Let $p \in [1,2)$ and $F : [0,1] \times D^p([0,1], G^1(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^2-non-anticipative Marcus canonical path functional such that $F, \nabla F$ are $[p,2)$-var-continuous. Then, for every $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^1(\mathbb{R}^{d+1}))$ the path $\nabla F(\cdot, \widehat{\mathbf{X}})$ is a càdlàg path of finite p'-variation, for all $p' > p$ such that $\frac{1}{p'} + \frac{1}{p} > 1$, and for all $t \in [0,1]$,*

$$F(t,\widehat{\mathbf{X}})-F(0,\widehat{\mathbf{X}}) = \int_0^t \nabla F(s^-,\widehat{\mathbf{X}})d\mathbf{X}_s \qquad (5.4.1)$$
$$+ \sum_{s\leqslant t} F(s,\widehat{\mathbf{X}}) - F(s^-,\widehat{\mathbf{X}}) - \nabla F(s^-,\widehat{\mathbf{X}})\Delta\widehat{\mathbf{X}}_s.$$

The integral in (5.4.1) is understood as a Young integral and the summation term is well defined as an absolutely summable series.

The proof of Theorem 5.4.1 will make use of the following lemma. Recall the notation introduced in Section 5.2.4 for linear functionals of the signature and that d_q denotes the q-variation pseudometric on the space of càdlàg paths (see Section 1.2.1).

Lemma 5.4.2. . Let $p \in [1,2)$, $F : [0,1] \times D^p([0,1], G^1(\mathbb{R}^{d+1})) \to \mathbb{R}$ be such that $F \in \mathcal{M}_1^2$ and $\widehat{\mathbf{X}} \in \widehat{C}^1([0,1], G^1(\mathbb{R}^{d+1}))$ a time-extended piecewise linear path. Assume that there exists $(\mathbf{u}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+1})$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sum_{j=0}^{2} \|\nabla^j F(t,\widehat{\mathbf{X}}) - \nabla^j F^{\mathbf{u}_n}(t,\widehat{\mathbf{X}})\| = 0. \qquad (5.4.2)$$

Then,

(i) $\sup_{n\in\mathbb{N}} \|\nabla F^{\mathbf{u}_n}(t,\widehat{\mathbf{X}})\|_{1\text{-}var} < \infty$;

(ii) for all $p' > 1$, $\lim_{n\to\infty} d_{p'}\big(\nabla F^{\mathbf{u}_n}(t,\widehat{\mathbf{X}}), \nabla F(\cdot,\widehat{\mathbf{X}})\big) = 0$.

Proof. Recall that for all $n \in \mathbb{N}$, $j = 0,\dots,2$, $\nabla^j F^{\mathbf{u}_n}(t,\widehat{\mathbf{X}}) = \langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_t\rangle$, where $\widehat{\mathbb{X}}$ denotes the signature of $\widehat{\mathbf{X}}$.

(i): Let $(t_i)_i \subset [0,1]$ be the grid such that $\widehat{\mathbf{X}}$ is a linear path on $[t_i, t_{i+1}]$. Notice that by Chen's relation and the invariance property of the signature under reparametrization (see Proposition 1.3.6) for every $u \in [t_i, t_{i+1}]$,

$$\widehat{\mathbb{X}}_u = \widehat{\mathbb{X}}_{t_i} \otimes \exp(\theta \widehat{\mathbf{X}}_{t_i,t_{i+1}}),$$

for some $\theta \in [0,1]$. Fix $s_1, s_2 \in [t_i, t_{i+1}]$, $s_1 \leqslant s_2$, for some i. Then, it holds that

$$\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_2}\rangle - \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_1}\rangle = g(1) - g(0),$$

for $g : [0,1] \to \mathbb{R}^d$ such that $g(\theta) := \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta\widehat{\mathbf{X}}_{s_1,s_2})\rangle$, for all $\theta \in [0,1]$. Thus, a first-order Taylor expansion of each of the components of g at 0 yields that

$$\|\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_2}\rangle - \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_1}\rangle\| \leqslant \sup_{\theta\in[0,1]} \|\langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta\widehat{\mathbf{X}}_{s_1,s_2})\rangle\| \|\widehat{\mathbf{X}}_{s_1,s_2}\|. \qquad (5.4.3)$$

Since by (5.4.2),

$$\sup_{n\in\mathbb{N}} \sup_{s_1,s_2\in[t_i,t_{i+1}]} \sup_{\theta\in[0,1]} \|\langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta \widehat{\mathbf{X}}_{s_1,s_2})\rangle\| < \infty,$$

equation (5.4.3) and $\|\widehat{\mathbf{X}}\|_{1\text{-}var} < \infty$ yield that $\sup_{n\in\mathbb{N}} \|\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_\cdot\rangle\|_{1\text{-}var[t_i,t_{i+1}]} < \infty$. Repeating the same reasoning on every $[t_i, t_{i+1}]$, the claim follows.

(ii): It follows from (i) (5.4.2) and interpolation (see Lemma 5.12 and Lemma 5.27 in Friz and Victoir (2010)). □

We are now ready to present the proof of Theorem 5.4.1.

Proof of Theorem 5.4.1. We split the proof into three main steps. In Step 1 we prove the assertion for functionals evaluated at a time-extended piecewise linear path. Then, in Step 2 we extend it to the whole space of time-extended continuous paths by a density argument. Finally, in Step 3, exploiting that the functionals considered are of Marcus type, an adaptation of the proof of Theorem 38 in Friz and Shekhar (2017) leads to the general result.

Step 1: Let $\widehat{\mathbf{X}} \in \widehat{C}^1([0,1], G^1(\mathbb{R}^{d+1}))$ be a time-extended peicewise linear path. Since F is a C^2-non-anticipative path functional, by Theorem 5.3.15 there exists a sequence $(\mathbf{u}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+1})$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sum_{j=0}^{2} \|\nabla^j F(t, \widehat{\mathbf{X}}) - \nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}})\| = 0.$$

Recall that for all $n \in \mathbb{N}$, $j = 0, \dots, 2$, $\nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}) = \langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_t\rangle$, where $\widehat{\mathbb{X}}$ denotes the signature of $\widehat{\mathbf{X}}$. By Lemma 5.4.2 (i) and Lemma 5.12 in Friz and Victoir (2010) $\|\nabla F(\cdot, \widehat{\mathbf{X}})\|_{1\text{-}var} < \infty$ and thus in particular

$$\|\nabla F(\cdot, \widehat{\mathbf{X}})\|_{p'\text{-}var} < \infty, \tag{5.4.4}$$

for all $p' > 1$. Fix $p' > 1$. By Lemma 5.4.2 (ii) $\lim_{n\to\infty} d_{p'}\big(\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_\cdot\rangle, \nabla F(\cdot, \widehat{\mathbf{X}})\big) = 0$. Moreover, by Proposition 6.11 in Friz and Victoir (2010), for all $n \in \mathbb{N}$,

$$\begin{aligned} d_1&\left(\int_0^{\cdot} \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_s\rangle d\widehat{\mathbf{X}}_s, \int_0^{\cdot} \nabla F(s, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s\right) \\ &\leqslant C\left(d_{CC}\big(\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_0\rangle, \nabla F(0, \widehat{\mathbf{X}})\big) + d_{p'}\big(\langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_\cdot\rangle, \nabla F(\cdot, \widehat{\mathbf{X}})\big)\right), \end{aligned}$$

for some $C > 0$. Since by equation (1.3.12) for every $n \in \mathbb{N}$ and $t \in [0,1]$, it holds that $\langle \mathbf{u}_n, \widehat{\mathbb{X}}_t\rangle = \langle \mathbf{u}_n, \widehat{\mathbb{X}}_0\rangle + \int_0^t \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_s\rangle d\widehat{\mathbf{X}}_s$, the claim follows.

Step 2: Fix $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^1(\mathbb{R}^{d+1}))$. By Theorem 5.23 in Friz and Victoir (2010), there exists a sequence of piecewise linear time-extended paths $(\widehat{\mathbf{X}}^M)_{M\in\mathbb{N}}$ such that

$$\lim_{M\to\infty} \sup_{t\in[0,1]} d_{CC}(\widehat{\mathbf{X}}_t^M, \widehat{\mathbf{X}}_t) = 0, \qquad \text{and} \qquad \sup_{M\in\mathbb{N}} \|\widehat{\mathbf{X}}^M\|_{p\text{-}var} \leqslant C\|\widehat{\mathbf{X}}\|_{p\text{-}var}, \tag{5.4.5}$$

for some $C > 0$. By Step 1, for every fixed M and $t \in [0,1]$, it holds that

$$F(t, \widehat{\mathbf{X}}^M) - F(0, \widehat{\mathbf{X}}^M) = \int_0^t \nabla F(s, \widehat{\mathbf{X}}^M) d\widehat{\mathbf{X}}_s^M.$$

By conditions (5.4.5) and interpolation (see Lemma 5.12 and Lemma 5.27 in Friz and Victoir (2010)), for all $p' > p$, $\lim_{M\to\infty} d_{p'}(\widehat{\mathbf{X}}^M, \widehat{\mathbf{X}}) = 0$. Since both F and ∇F are by assumption $[p,2)$-var continuous,

$$\lim_{M\to\infty} d_{p'}(F(\cdot, \widehat{\mathbf{X}}^M), F(\cdot, \widehat{\mathbf{X}})) = 0 \qquad \text{and} \qquad \lim_{M\to\infty} d_{p'}(\nabla F(\cdot, \widehat{\mathbf{X}}^M), \nabla F(\cdot, \widehat{\mathbf{X}})) = 0.$$

This in particular implies that for all $p' > p$ such that $\frac{1}{p'} + \frac{1}{p} > 1$, $\|\nabla F(\cdot, \widehat{\mathbf{X}})\|_{p'\text{-}var} < \infty$, as by Step 1 (see equation (5.4.4)) for all M, $\|\nabla F(\cdot, \widehat{\mathbf{X}}^M)\|_{p'\text{-}var} < \infty$.

Next, fix q such that $2 > q > p$. Observe that $\widehat{\mathbf{X}}, \widehat{\mathbf{X}}^M \in \widehat{C}^q([0,1], G^1(\mathbb{R}^{d+1}))$, for all $M \in \mathbb{N}$, and since $\nabla F(\cdot, \widehat{\mathbf{X}}), \nabla F(\cdot, \widehat{\mathbf{X}}^M) \in C^q([0,1], \mathbb{R}^{d+1})$ as well, the Young integral of the integrands $\nabla F(\cdot, \widehat{\mathbf{X}}), \nabla F(\cdot, \widehat{\mathbf{X}}^M)$ with respect to $\widehat{\mathbf{X}}, \widehat{\mathbf{X}}^M$, respectively, is well-defined. Finally, by Proposition 6.11 in Friz and Victoir (2010), for all $M \in \mathbb{N}$,

$$\begin{aligned} d_q&\left(\int_0^\cdot \nabla F(s, \widehat{\mathbf{X}}^M) d\widehat{\mathbf{X}}_s^M, \int_0^\cdot \nabla F(s, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s\right) \\ &\leqslant C\Big(d_q(\widehat{\mathbf{X}}^M, \widehat{\mathbf{X}}) + d_{CC}\big(\nabla F(0, \widehat{\mathbf{X}}^M), \nabla F(0, \widehat{\mathbf{X}})\big) + d_q\big(\nabla F(\cdot, \widehat{\mathbf{X}}^M), \nabla F(\cdot, \widehat{\mathbf{X}})\big)\Big), \end{aligned}$$

for some $C > 0$. The claim follows as in Step 1.

Step 3: Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^1(\mathbb{R}^{d+1}))$. For notational convenience, let us denote by $\mathbf{Z} \in \widehat{C}^p([0,1], G^1(\mathbb{R}^{d+1})$ the time-extended Marcus transformed path of $\widehat{\mathbf{X}}$ and let $\mu_t \in [0,1]$ be such that $\mathbf{Z}_{\mu_t} = \widehat{\mathbf{X}}_t$ for all $t \in [0,1]$ (see Notation 5.2.7). By Step 1 and Step 2, $\nabla F(\cdot, \mathbf{Z})$ is a continuous path of finite p'-variation, for all $p' > p$ such that $\frac{1}{p'} + \frac{1}{p} > 1$. Since ∇F is a Marcus canonical path functional, for all $t \in [0,1]$,

$$\nabla F(t, \widehat{\mathbf{X}}) = \nabla F(\mu_t, \mathbf{Z}),$$

implying by definition of μ_t that $\nabla F(\cdot, \widehat{\mathbf{X}})$ is a càdlàg path of finite p'-variation.

Next, assume first that $\widehat{\mathbf{X}}$ admits only one jump at time $a \in (0,1]$. Suppose that $t < a$. By the arguments in Step 2 applied to $\mathbf{Z}$,

$$F(\mu_t, \mathbf{Z}) = F(0, \mathbf{Z}) + \int_0^{\mu_t} \nabla F(s, \mathbf{Z}) d\mathbf{Z}_s.$$

Observe that $\nabla F \in \mathcal{M}_1^1 \subset \mathcal{M}_1^0$ implies that it is a path functional invariant under reparametrization (see Proposition 5.2.11). This, combined with the property of the Young integral, yields that the functional

$$\begin{aligned} G : [0,1] \times C^p([0,1], G^1(\mathbb{R}^{d+1})) &\to \mathbb{R} \\ (t, \mathbf{Y}) \mapsto G(t, \mathbf{Y}) &:= \int_0^t \nabla F(s, \mathbf{Y}) d\mathbf{Y}_s \end{aligned}$$

is invariant under reparametrization too. Thus, for $t < a$, we get that $\int_0^{\mu_t} \nabla F(s, \mathbf{Z}) d\mathbf{Z}_s = \int_0^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s$, as the stopped path $\mathbf{Z}^{\mu_t}$ is nothing else than a time reparametrization of the continuous path $\widehat{\mathbf{X}}^t$. Since it also holds that $F(\mu_t, \mathbf{Z}) = F(t, \widehat{\mathbf{X}})$, $F(0, \mathbf{Z}) = F(0, \mathbf{X})$, the claim follows. Next, suppose that $t \geqslant a$. Then,

$$\begin{aligned} F(t, \mathbf{X}) - F(0, \mathbf{X}) =& F(\mu_t, \mathbf{Z}) - F(\mu_a, \mathbf{Z}) \\ &+ F(\mu_a, \mathbf{Z}) - F(\mu_{a^-}, \mathbf{Z}) \\ &+ F(\mu_{a^-}, \mathbf{Z}) - F(0, \mathbf{Z}). \end{aligned}$$

By a similar argument as before, we get that

$$\begin{aligned} &F(\mu_t, \mathbf{Z}) - F(\mu_a, \mathbf{Z}) = \int_{\mu_a}^{\mu_t} \nabla F(s, \mathbf{Z}) d\mathbf{Z}_s = \int_a^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s, \\ &F(\mu_a, \mathbf{Z}) - F(\mu_{a^-}, \mathbf{Z}) = F(a, \widehat{\mathbf{X}}) - F(a^-, \widehat{\mathbf{X}}), \\ &F(\mu_{a^-}, \mathbf{Z}) - F(0, \mathbf{Z}) = \int_0^{\mu_{a^-}} \nabla F(s, \mathbf{Z}) d\mathbf{Z}_s = \int_0^a \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s - \nabla F(a^-, \widehat{\mathbf{X}}) \Delta \widehat{\mathbf{X}}_a. \end{aligned}$$

Combining all three terms, the claim follows. The argument extends trivially to a càdlàg paths with finitely many jumps. Finally, let $\widehat{\mathbf{X}}$ be a càdlàg path with countable many jumps. By the arguments in Step 2 applied to $\mathbf{Z}$, for $\varepsilon > 0$, there exists $\eta > 0$ and a partition $\pi_{[0,\mu_t]}$ with $|\pi_{[0,\mu_t]}| < \eta$ such that

$$\Big\| F(\mu_t, \mathbf{Z}) - F(0, \mathbf{Z}) - \sum_{s_i \in \pi_{[0,\mu_t]}} \nabla F(s_i, \mathbf{Z}) \mathbf{Z}_{s_i, s_{i+1}} \Big\| < \frac{\varepsilon}{2}.$$

Since $\widehat{\mathbf{X}}$ has finite p-variation, for $p \in [1, 2)$, we can find a finite set $B(\varepsilon, t) \subset [0, t]$ of jump times of $\widehat{\mathbf{X}}$ such that $\sum_{s \leqslant t,\ s \notin B(\varepsilon,t)} \|\Delta \widehat{\mathbf{X}}_s\|^2 < \frac{\varepsilon}{2}$. Without loss of generality, we may assume that if $t_j \in B(\varepsilon, t)$, then $\mu_{t_j^-}, \mu_{t_j} \in \pi_{[0,\mu_t]}$. Thus, repeating the arguments applied in the first part of the proof, we get that there exists a partition $\pi_{[0,t]}$ such that

$$\begin{aligned} \Big\| F(t, \widehat{\mathbf{X}}) - F(0, \widehat{\mathbf{X}}) - &\sum_{s_i \in \pi_{[0,t]}} \nabla F(s_i, \widehat{\mathbf{X}}) \widehat{\mathbf{X}}_{s_i, s_{i+1}} \\ - &\sum_{s \in B(\varepsilon,t)} F(s, \widehat{\mathbf{X}}) - F(s^-, \widehat{\mathbf{X}}) - \nabla F(s^-, \widehat{\mathbf{X}}) \Delta \widehat{\mathbf{X}}_s \Big\| < \frac{\varepsilon}{2}. \end{aligned} \tag{5.4.6}$$

Next, let $(\mathbf{u}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+1})$ be such that $\lim_{n\to\infty}\sup_{t\in[0,1]}\sum_{j=0}^{2}\|\nabla^j F(t,\widehat{\mathbf{X}})-\langle \mathbf{u}_n^{(j)},\widehat{\mathbb{X}}_t\rangle\| = 0$. Recall that the existence of this sequence is guaranteed by Theorem 5.3.15. Since by Remark 5.A.1 (i), for all $n \in \mathbb{N}$,

$$\sum_{s\notin B(\varepsilon,t)} \|\langle \mathbf{u}_n, \widehat{\mathbb{X}}_s\rangle - \langle \mathbf{u}_n, \widehat{\mathbb{X}}_{s-}\rangle - \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s-}\rangle \Delta\widehat{\mathbf{X}}_s\|$$

$$\leqslant \sup_{n\in\mathbb{N}}\sup_{s\notin B(\varepsilon,t)}\sup_{\theta\in[0,1]} \|\langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s-}\otimes \exp(\theta \log^{(1)}(\Delta\widehat{\mathbf{X}}_s))\rangle\| \sum_{s\notin B(\varepsilon,t)} \|\Delta\widehat{\mathbf{X}}_s\|^2 < \infty,$$

an application of the dominated convergence theorem yields that

$$\sum_{s\notin B(\varepsilon,t)} \|F(s,\widehat{\mathbf{X}}) - F(s^-,\widehat{\mathbf{X}}) - \nabla F(s^-,\widehat{\mathbf{X}})\Delta\widehat{\mathbf{X}}_s\| < \infty.$$

Thus, repeating the above argument, we can take $\varepsilon > 0$ such that

$$\begin{aligned} F(t,\widehat{\mathbf{X}}) - F(0,\mathbf{X}) = &\lim_{(RRS)|\pi_{[0,t]}|\to 0} \sum_{s_i\in\bar{\pi}_{[0,t]}} \nabla F(s_i,\widehat{\mathbf{X}})\widehat{\mathbf{X}}_{s_i,s_{i+1}} \\ &- \sum_{s\leqslant t} F(s,\widehat{\mathbf{X}}) - F(s^-,\widehat{\mathbf{X}}) - \nabla F(s^-,\widehat{\mathbf{X}})\Delta\widehat{\mathbf{X}}_s. \end{aligned} \tag{5.4.7}$$

Since $\widehat{\mathbf{X}}$ is a càdlàg path, by Proposition 2.4 in Friz and Zhang (2017), the convergence in equation (5.4.7) holds in Mesh Riemann-Stieltjes sense (see Definition 1.2.4) and the claim follows.

□

Remark 5.4.3. An inspection of the above proof shows that we can replace the hypothesis of F being $[p,2)$-var-continuous with the assumption that

$$\lim_{n\to\infty}\sup_{t\in[0,1]} d_{CC}(\mathbf{X}_t^n, \mathbf{X}_t) = 0$$

implies

$$\lim_{n\to\infty}\sup_{t\in[0,1]} d_{CC}(F(t,\mathbf{X}^n), F(t,\mathbf{X})) = 0,$$

for every $(\mathbf{X}^n,\mathbf{X})_{n\in\mathbb{N}} \subset C^p([0,1], G^{[p]}(\mathbb{R}^d))$.

5.4.2 The case $p \in [2,3)$

In this section, we present the functional Itô-formula for maps of càdlàg p-rough paths for $p \in [2,3)$. Recall the notions of controlled rough path introduced in Definition 1.2.9, remainder functional in Definition 5.2.12, C^3-non-anticipative Marcus canonical path functional given in Definition 5.3.12, and finally of Marcus-like weakly geometric càdlàg rough path in Definition 1.2.1.

Theorem 5.4.4. *Let $p \in [2,3)$ and $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^3-non-anticipative Marcus canonical path functional such that $F, \nabla F, \nabla^2 F$ and $R^{\nabla F, \nabla^2 F}$ are $[p,3)$-var-continuous. Then, for every $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, with $\widehat{\mathbf{X}}$ being Marcus-like, it holds that the pair $\big(\nabla F(\cdot, \widehat{\mathbf{X}}), \nabla^2 F(\cdot, \widehat{\mathbf{X}})\big) \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$, for all $p' > p$ such that $\frac{1}{p'} + \frac{2}{p} > 1$, and for all $t \in [0,1]$,*

$$\begin{aligned} F(t, \widehat{\mathbf{X}}) - F(0, \widehat{\mathbf{X}}) = &\int_0^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_s \\ &+ \sum_{s \leqslant t} F(s, \widehat{\mathbf{X}}) - F(s^-, \widehat{\mathbf{X}}) - \nabla F(s^-, \widehat{\mathbf{X}}) \Delta \widehat{X}_s - \nabla^2 F(s^-, \widehat{\mathbf{X}}) \Delta \widehat{\mathbb{X}}^{(2)}_s. \end{aligned} \tag{5.4.8}$$

The integral in (5.4.8) is understood as a rough integral and the summation term is well defined as an absolutely summable series.

Similarly to the results in the previous section, the proof of Theorem 5.4.4 will make use of the following lemma, where algebraic properties of the signature of piecewise linear paths are exploited to infer some key analytical properties.

Recall from Section 1.3.3 that the truncated signature at level 2 of a piecewise linear path (defined as in equation (1.2.8)) belongs to $\widehat{C}^1([0,1], G^2(\mathbb{R}^{d+1})) \subseteq \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$. Recall furthermore that for $\mathbf{u} \in T(\mathbb{R}^{d+1})$, the functional $F^{\mathbf{u}}$ has been introduced in Section 5.2.4 and that $R^{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}}((\cdot,\cdot),\cdot)$ denotes the remainder (see Definition 5.2.12) related to the functionals $F^{\mathbf{u}^{(1)}}$, $F^{\mathbf{u}^{(2)}}$ introduced in equation (5.2.17).

Lemma 5.4.5. Let $p \in [2,3)$, $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be such that $F \in \mathcal{M}^3_2$ and $\widehat{\mathbf{X}} \in \widehat{C}^1([0,1], G^2(\mathbb{R}^{d+1}))$ be the truncated signature at level 2 of a time-extended piecewise linear path. Assume that there exists $(\mathbf{u}_n)_{n \in \mathbb{N}} \subset T(\mathbb{R}^{d+1})$ such that

$$\lim_{n \to \infty} \sup_{t \in [0,1]} \sum_{j=0}^{3} \|\nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}) - \nabla^j F(t, \widehat{\mathbf{X}})\| = 0. \tag{5.4.9}$$

Then,

(i) $\sup_{n \in \mathbb{N}} \left(\|\nabla^2 F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}})\|_{1\text{-}var} + \|R^{\mathbf{u}_n^{(1)}, \mathbf{u}_n^{(2)}}((\cdot,\cdot), \widehat{\mathbf{X}})\|_{\frac{1}{2}\text{-}var} \right) < \infty$;

(ii) for all $p' > 1$, $r > \frac{1}{2}$,

$$\begin{aligned} &\lim_{n \to \infty} d_{p'}\big(\nabla^2 F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}), \nabla^2 F(\cdot, \widehat{\mathbf{X}})\big) = 0, \\ &\lim_{n \to \infty} d_r\big(R^{\mathbf{u}_n^{(1)}, \mathbf{u}_n^{(2)}}((\cdot,\cdot), \widehat{\mathbf{X}}), R^{\nabla F, \nabla^2 F}((\cdot,\cdot), \widehat{\mathbf{X}})\big) = 0. \end{aligned}$$

Proof. Recall that for all $n \in \mathbb{N}$, $j = 0, \ldots, 3$, $\nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}) = \langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_t \rangle$, where $\widehat{\mathbb{X}}$ denotes the signature of $\widehat{\mathbf{X}}$.

(i): Let $(t_i)_i \subset [0,1]$ be the grid such that $\widehat{X} := \pi_1(\widehat{\mathbf{X}})$ is a linear path on $[t_i, t_{i+1}]$, and notice that for every $u \in [t_i, t_{i+1}]$, $\widehat{\mathbb{X}}_u = \widehat{\mathbb{X}}_{t_i} \otimes \exp(\theta \widehat{X}_{t_i,t_{i+1}})$, for some $\theta \in [0,1]$. Fix $s_1, s_2 \in [t_i, t_{i+1}]$, $s_1 \leqslant s_2$, for some i. Then,

$$\langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_2} \rangle - \langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \rangle = g(1) - g(0),$$

for $g : [0,1] \to (\mathbb{R}^{d+1})^{\otimes 2}$ such that $g(\theta) := \langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta \widehat{X}_{s_1,s_2}) \rangle$ for all $\theta \in [0,1]$. Thus, a first-order Taylor expansion of the components of g at 0 yields that

$$\| \langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_2} \rangle - \langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \rangle \| \leqslant \sup_{\theta \in [0,1]} \| \langle \mathbf{u}_n^{(3)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta \widehat{X}_{s_1,s_2}) \rangle \| \| \widehat{X}_{s_1,s_2} \|.$$

Similarly, a second-order Taylor expansion of the map $g' : [0,1] \to \mathbb{R}^d$ defined via

$$g'(\theta) := \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta \widehat{X}_{s_1,s_2}) \rangle,$$

for all $\theta \in [0,1]$, yields that

$$\begin{aligned} \| R^{\mathbf{u}_n^{(1)}, \mathbf{u}_n^{(2)}}((s_1,s_2), \widehat{\mathbf{X}}) \| =& \| \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_2} \rangle - \langle \mathbf{u}_n^{(1)}, \widehat{\mathbb{X}}_{s_1} \rangle - \langle \mathbf{u}_n^{(2)}, \widehat{\mathbb{X}}_{s_1} \rangle \widehat{X}_{s_1,s_2} \| \\ \leqslant& \frac{1}{2} \sup_{\theta \in [0,1]} \| \langle \mathbf{u}_n^{(3)}, \widehat{\mathbb{X}}_{s_1} \otimes \exp(\theta \widehat{X}_{s_1,s_2}) \rangle \| \| \widehat{X}_{s_1,s_2} \|^2. \end{aligned} \tag{5.4.10}$$

The claim follows as in the proof of (i) of Lemma 5.4.5.

(ii): The assertion follows from (i) and interpolation (see Lemma 5.12 and Lemma 5.27 in Friz and Victoir (2010)). □

Proof of Theorem 5.4.4. The structure of the proof is very similar to the one of Theorem 5.4.1. We split it into three main steps. In Step 1, the assertion is proved in the simpler setting of functionals evaluated at the truncated signature at level 2 of a time-extended piecewise linear path. In Step 2, we extend the result to the whole space of continuous paths by a density argument. Finally, in Step 3, we conclude by addressing the case of functionals evaluated at a general càdlàg rough path.

Step 1: Let $\widehat{\mathbf{X}} \in \widehat{C}^1([0,1], G^2(\mathbb{R}^{d+1}))$ be the truncated signature at level 2 of a time-extended piecewise linear path. Since F is a C^3-non-anticipative path functional, by Theorem 5.3.15, there exists a sequence $(\mathbf{u}_n)_{n \in \mathbb{N}} \subset T(\mathbb{R}^{d+1})$ such that

$$\lim_{n \to \infty} \sup_{t \in [0,1]} \sum_{j=0}^{3} \| \nabla^j F(t, \widehat{\mathbf{X}}) - \nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}) \| = 0.$$

Recall that for all $n \in \mathbb{N}$, $j = 0, \ldots, 3$, $\nabla^j F^{\mathbf{u}_n}(t, \widehat{\mathbf{X}}) = \langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_t \rangle$, where $\widehat{\mathbb{X}}$ denotes the signature of $\widehat{\mathbf{X}}$. Thus, by Lemma 5.4.5 (i) and Lemma 5.12 in Friz and Victoir (2010),

$$\| \nabla^2 F(\cdot, \widehat{\mathbf{X}}) \|_{1\text{-}var} + \| R^{\nabla F, \nabla^2 F}(\cdot, \cdot, \widehat{\mathbf{X}}) \|_{\frac{1}{2}\text{-}var} < \infty, \tag{5.4.11}$$

implying in particular that $\big(\nabla F(\cdot,\widehat{\mathbf{X}}),\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big)\in\mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$, for all $p'>p$ such that $\frac{1}{p'}+\frac{2}{p}>1$. Fix such a p' and notice that by equation (1.2.6), $r>\frac{p}{2}>\frac{1}{2}$. By Lemma 5.4.5 (ii),

$$\lim_{n\to\infty} d_{p'}\big(\nabla^2 F(\cdot,\widehat{\mathbf{X}}),\langle \mathbf{u}_n^{(2)},\widehat{\mathbb{X}}_\cdot\rangle\big)=0,$$
$$\lim_{n\to\infty} d_r\big(R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}}),R^{\mathbf{u}_n^{(1)},\mathbf{u}_n^{(2)}}(\cdot,\cdot,\widehat{\mathbf{X}})\big)=0.$$

Moreover, an application of Theorem 8.10 in Friz and Victoir (2010) and Proposition 2.7 in Allan et al. (2023b) yields that for all $n\in\mathbb{N}$,

$$\begin{aligned} d_p\bigg(\int_0^\cdot\langle \mathbf{u}_n^{(1)},\widehat{\mathbb{X}}_s\rangle d\widehat{\mathbf{X}}_s,\int_0^\cdot \nabla F(s,\widehat{\mathbf{X}})d\widehat{\mathbf{X}}_s\bigg)&\\ \leqslant C\bigg(d_{CC}\big(\langle \mathbf{u}_n^{(1)},\widehat{\mathbb{X}}_0\rangle,\nabla F(0,\widehat{\mathbf{X}})\big)&+d_{CC}\big(\langle \mathbf{u}_n^{(2)},\widehat{\mathbb{X}}_0\rangle,\nabla^2 F(0,\widehat{\mathbf{X}})\big)\\ +d_{p'}\big(\langle \mathbf{u}_n^{(2)},\widehat{\mathbb{X}}\rangle,\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big)&+d_r\big(R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}}),R^{\mathbf{u}_n^{(1)},\mathbf{u}_n^{(2)}}(\cdot,\cdot,\widehat{\mathbf{X}})\big)\bigg)\end{aligned}$$

for some $C>0$. Since by equation (1.3.13) for every $n\in\mathbb{N}$ and $t\in[0,1]$, it holds that $\langle \mathbf{u}_n,\widehat{\mathbb{X}}_t\rangle=\langle \mathbf{u}_n,\widehat{\mathbb{X}}_0\rangle+\int_0^t\langle \mathbf{u}_n^{(1)},\widehat{\mathbb{X}}_s\rangle d\widehat{\mathbf{X}}_s$, the claim follows.

Step 2: Fix $\widehat{\mathbf{X}}\in\widehat{C}^p([0,1],G^2(\mathbb{R}^{d+1}))$. By Theorem 8.12 in Friz and Victoir (2010), there exists a sequence of piecewise linear time-extended paths such that their truncated signature at level 2, denoted by $(\widehat{\mathbf{X}}^M)_M$, satisfy

$$\lim_{M\to\infty}\sup_{t\in[0,1]} d_{CC}(\widehat{\mathbf{X}}^M_t,\widehat{\mathbf{X}}_t)=0,\qquad\text{and}\qquad \sup_{M\in\mathbb{N}}\|\widehat{\mathbf{X}}^M\|_{p\text{-}var}\leqslant C\|\widehat{\mathbf{X}}\|_{p\text{-}var},\tag{5.4.12}$$

for some $C>0$. By Step 1, for every fixed M and $t\in[0,1]$,

$$F(t,\widehat{\mathbf{X}}^M)-F(0,\widehat{\mathbf{X}}^M)=\int_0^t\nabla F(s,\widehat{\mathbf{X}}^M)d\widehat{\mathbf{X}}^M_s.$$

By conditions (5.4.12) and interpolation (see Lemma 5.12 and Lemma 8.16 in Friz and Victoir (2010)), for all $p'>p$, $\lim_{M\to\infty}d_{p'}(\widehat{\mathbf{X}}^M,\widehat{\mathbf{X}})=0$. Thus, since $F,\nabla F,\nabla^2 F$ and $R^{\nabla F,\nabla^2 F}$ are by assumption $[p,3)$-var continuous, it holds that

$$\begin{aligned}&\lim_{M\to\infty} d_{p'}(F(\cdot,\widehat{\mathbf{X}}^M),F(\cdot,\widehat{\mathbf{X}}))=0,\\ &\lim_{M\to\infty} d_{p'}(\nabla F(\cdot,\widehat{\mathbf{X}}^M),\nabla F(\cdot,\widehat{\mathbf{X}}))=0,\\ &\lim_{M\to\infty} d_{p'}(\nabla^2 F(\cdot,\widehat{\mathbf{X}}^M),\nabla^2 F(\cdot,\widehat{\mathbf{X}}))=0,\\ &\lim_{M\to\infty} d_{\frac{p'}{2}}\big(R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}}^M),R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}})\big)=0.\end{aligned}\tag{5.4.13}$$

Fix p' and r as in Step 1 and notice that since for all M, $\big(\nabla F(\cdot,\widehat{\mathbf{X}}^M),\nabla^2 F(\cdot,\widehat{\mathbf{X}}^M)\big) \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$, by the convergences in (5.4.13) $\big(\nabla F(\cdot,\widehat{\mathbf{X}}),\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big) \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$ too.

Next, fix q such that $3 > q > p$ and set $r = \frac{q}{2}$. Observe that for all $M \in \mathbb{N}$, $\widehat{\mathbf{X}}^M, \widehat{\mathbf{X}} \in \widehat{D}^q([0,1],G^2(\mathbb{R}^{d+1}))$ and by (5.4.11), $\big(\nabla F(\cdot,\widehat{\mathbf{X}}^M),\nabla^2 F(\cdot,\widehat{\mathbf{X}}^M)\big) \in \mathcal{V}^{q,q/2}_{\widehat{\mathbf{X}}^M}$. Therefore, by (5.4.13), it also holds that $\big(\nabla F(\cdot,\widehat{\mathbf{X}}),\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big) \in \mathcal{V}^{q,q/2}_{\widehat{\mathbf{X}}}$. Finally, an application of Theorem 8.10 in Friz and Victoir (2010) and Proposition 2.7 in Allan et al. (2023b) yields that there exists a constant $C > 0$, such that for all $M \in \mathbb{N}$,

$$\begin{aligned}
&d_q\bigg(\int_0^\cdot \nabla F(s,\widehat{\mathbf{X}}^M)d\widehat{\mathbf{X}}^M_s, \int_0^\cdot \nabla F(s,\widehat{\mathbf{X}})d\widehat{\mathbf{X}}_s\bigg) \qquad (5.4.14)\\
&\quad\leqslant C\bigg(d_q(\widehat{\mathbf{X}}^M,\widehat{\mathbf{X}}) + d_{CC}\big(\nabla F(0,\widehat{\mathbf{X}}^M),\nabla F(0,\widehat{\mathbf{X}})\big) + d_{CC}\big(\nabla^2 F(0,\widehat{\mathbf{X}}^M),\nabla^2 F(0,\widehat{\mathbf{X}})\big)\\
&\qquad + d_q\big(\nabla^2 F(\cdot,\widehat{\mathbf{X}}^M),\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big) + d_{q/2}\big(R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}}^M), R^{\nabla F,\nabla^2 F}(\cdot,\cdot,\widehat{\mathbf{X}})\big)\bigg).
\end{aligned}$$

The claim follows.

Step 3: The proof follows step by step the reasonings in Step 3 of the proof of Theorem 5.4.1. We report here just the main differences.

Fix $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1],G^1(\mathbb{R}^{d+1}))$, with $\widehat{\mathbf{X}}$ Marcus like and set $\widehat{X} := \pi_1(\widehat{\mathbf{X}})$. For notational convenience, let us denote by $\mathbf{Z} \in \widehat{C}^p([0,1],G^2(\mathbb{R}^{d+1}))$ the time-extended Marcus transformed path of $\widehat{\mathbf{X}}$ and let $\mu_t \in [0,1]$ be such that $\mathbf{Z}_{\mu_t} = \widehat{\mathbf{X}}_t$ for all $t \in [0,1]$ (see Notation 5.2.7). It holds that $\big(\nabla F(\cdot,\widehat{\mathbf{X}}),\nabla^2 F(\cdot,\widehat{\mathbf{X}})\big) \in \mathcal{V}^{p',r}_{\widehat{\mathbf{X}}}$. Next, assume first that $\widehat{\mathbf{X}}$ admits only one jump at time $a \in (0,1]$. If $t < a$, the claim follows as in the Young case. If $t \geqslant a$, one just needs to notice that since we deal here with level 2 rough integration,

$$\begin{aligned}
&F(\mu_t,\mathbf{Z}) - F(\mu_a,\mathbf{Z}) = \int_a^t \nabla F(s^-,\widehat{\mathbf{X}})d\widehat{\mathbf{X}}_s,\\
&F(\mu_a,\mathbf{Z}) - F(\mu_{a^-},\mathbf{Z}) = F(a,\widehat{\mathbf{X}}) - F(a^-,\widehat{\mathbf{X}}),\\
&F(\mu_{a^-},\mathbf{Z}) - F(0,\mathbf{Z}) = \int_0^a \nabla F(s^-,\widehat{\mathbf{X}})d\widehat{\mathbf{X}}_s - \nabla F(a^-,\widehat{\mathbf{X}})\Delta\widehat{X}_a - \nabla^2 F(s^-,\widehat{\mathbf{X}})\Delta\widehat{\mathbb{X}}^{(2)}_a.
\end{aligned}$$

The argument extends trivially to a càdlàg paths with finitely many jumps. Finally, for $\widehat{\mathbf{X}}$ a càdlàg path with countable many jumps, observe that since $\widehat{\mathbf{X}}$ has finite p-variation, for $p \in [2,3)$, there exists a finite set $B(\varepsilon,t) \subset [0,t]$ of jump times of $\widehat{\mathbf{X}}$ such that $\sum_{s\leqslant t,\ s\notin B(\varepsilon,t)} \|\Delta\widehat{X}_s\|^3 < \frac{\varepsilon}{2}$, for $\varepsilon > 0$. Note furthermore that the very last step of the proof follows from Proposition 2.6 in Friz and Zhang (2017) instead of Proposition 2.4 of the same paper.

□

Remark 5.4.6. (i) An inspection of the above proof shows that we can replace the hypothesis of F, ∇F being $[p,3)$-var-continuous with the assumption that

$$\lim_{n\to\infty}\sup_{t\in[0,1]} d_{CC}(\mathbf{X}^n_t,\mathbf{X}_t)=0$$

implies

$$\lim_{n\to\infty}\sup_{t\in[0,1]} d_{CC}(F(t,\mathbf{X}^n),F(t,\mathbf{X}))=0, \qquad \lim_{n\to\infty}\sup_{t\in[0,1]} d_{CC}(\nabla F(t,\mathbf{X}^n),\nabla F(t,\mathbf{X}))=0,$$

for every $(\mathbf{X}^n,\mathbf{X})_{n\in\mathbb{N}}\subset C^p([0,1],G^{[p]}(\mathbb{R}^d))$.

(ii) Instead of assuming $R^{\nabla^1 F,\nabla^2 F}$ to be $[p,3)$-var continuous, one can also assume that $F,\nabla F,\nabla^2 F,\nabla^3 F$ and are $[p,3)$-var-continuous. Indeed, by Therem 5.3.15 and equation (5.4.10) for every (signature at level 2 of) a piecewise linear time-extended path $\widehat{\mathbf{X}}^M$ as in **Step 2**, one get

$$\begin{aligned}\|R^{\nabla^1 F,\nabla^2 F}((s,t),\widehat{\mathbf{X}}^M)\| &= \|\nabla^1 F(t,\widehat{\mathbf{X}}^M)-\nabla^1 F(s,\widehat{\mathbf{X}}^M)-\nabla^2 F(s,\widehat{\mathbf{X}}^M)\widehat{X}^M_{s,t}\|\\ &\leqslant C\sup_{s\in[0,1]}\|\nabla^3 F(u,\widehat{\mathbf{X}}^M)\|\|\widehat{X}^M_{s,t}\|^2,\end{aligned}$$

for $\widehat{X}_{s,t}:=\pi_1(\widehat{\mathbf{X}}_{s,t})$ and $(s,t)\in\Delta_1$. Therefore, if $\nabla^3 F$ is $[p,3)$-var continuous, the sequence $\|R^{\nabla^1 F,\nabla^2 F}((\cdot,\cdot),\widehat{\mathbf{X}}^M)\|$ is uniformly bounded in $p/2$-variation, and by interpolation for all $p'>p$,

$$\lim_{M\to\infty} d_{\frac{p'}{2}}\big(R^{\nabla F,\nabla^2 F}((\cdot,\cdot),\widehat{\mathbf{X}}^M),R^{\nabla F,\nabla^2 F}((\cdot,\cdot),\widehat{\mathbf{X}})\big)=0.$$

(iii) Observe that in view of the consistency between the Young and the rough integral (see Remark 1.2.18 (iii)) Itô-formula (5.4.8) explicitly reads as

$$\begin{aligned}F(t,\widehat{\mathbf{X}})-F(0,\widehat{\mathbf{X}}) = &\int_0^t U^0F(s^-,\widehat{\mathbf{X}})dv_s+\int_0^t \nabla F(s^-,\widehat{\mathbf{X}})d\mathbf{X}_s\\ &+\sum_{s\leqslant t}\Big(F(s,\widehat{\mathbf{X}})-F(s^-,\widehat{\mathbf{X}})-\nabla F(s^-,\widehat{\mathbf{X}})\Delta\widehat{X}_s\\ &\qquad -\frac{1}{2}\sum_{i,j=0}^d U^iU^jF(s^-,\widehat{\mathbf{X}})\Delta\langle\epsilon_i,\widehat{\mathbf{X}}_s\rangle\Delta\langle\epsilon_j,\widehat{\mathbf{X}}_s\rangle 1_{\{j\neq 0\}}\Big),\end{aligned}$$

for all $(t,\widehat{\mathbf{X}})\in[0,1]\times\widehat{D}^p([0,1],G^2(\mathbb{R}^{d+1}))$ given as v-extended path of some $\mathbf{X}\in D^p([0,1],G^2(\mathbb{R}^d))$, for v some of its tj-path. Here, the index 0 denotes the tj-path component, $U^0F(t,\widehat{\mathbf{X}})$ the vertical derivatives in the direction 0, and the first integral on the right-hand rise is a Young integral of the path $s\mapsto U^0F(s,\widehat{\mathbf{X}})$ with respect to v. Notice that if in particular $\widehat{\mathbf{X}}$ is continuous, $v(s)=s$ for all $s\in[0,1]$ and

$$\int_0^t U^0F(s^-,\widehat{\mathbf{X}})dv_s=\int_0^t U^0F(s^-,\widehat{\mathbf{X}})ds.$$

We highlight that the consideration of tj-extended paths is necessary not only for deriving the approximation result in Theorem 5.3.15, but also to express the dependence of a non-anticipative path functional on the parameter t. Consider, for instance, a path functional F such that for all $(t, \mathbf{X}) \in [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1}))$,

$$F(t, \mathbf{X}) = \sin(\langle \epsilon_0, \mathbf{X}_t \rangle).$$

Then for all $\widehat{\mathbf{X}}$ tj-extended path, $F(t, \widehat{\mathbf{X}}) = \sin(v(t))$, $U^0 F(t, \widehat{\mathbf{X}}) = \cos(v(t))$, $U^0 U^0 F(t, \widehat{\mathbf{X}}) = -\sin(v(t))$ and all the other (first and second order) vertical derivatives are equal to 0. Since the hypothesis of Theorem 5.4.4 are satisfied, we get that

$$\sin(v_t) - \sin(v_0) = \int_0^t \cos(v_{s-}) dv_s + \sum_{s \leqslant t} \sin(v_s) - \sin(v_{s-}) - \cos(v_{s-}) \Delta v_s,$$

which coincides with the change of variable formula for paths of finite variation. In particular, if $\widehat{\mathbf{X}}$ is continuous, the path functional explicitly reads as $F(t, \mathbf{X}) = \sin(t)$, and the consideration of time-extended paths is necessary in order to embed the functional F in the present framework.

(iv) Theorems 5.4.1 and 5.4.4 provide an Itô-formula for C^2 and C^3 non-anticipative Marcus canonical path functional, respectively (see Definition 5.3.12). This means that our framework includes all the path functionals for which for all $i = 0, \ldots, 2$ (or $i = 0, \ldots, 3$) the maps $[0,1] \ni t \mapsto \nabla^i F(t, \widehat{\mathbf{X}})$ is continuous if $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$, and càdlàg if $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ (see Definition 5.3.12). This means that functionals of the form $F(t, \mathbf{X}) := \mathbf{X}_{s \wedge t}$, for some fixed $s \in (0,1]$ are not included in our setup. A similar argument applies to the functional $F(t, \mathbf{X}) := \mathbf{X}_{t-}$, as the map $[0,1] \ni t \mapsto F(t, \mathbf{X})$ is càglàd for $\mathbf{X}$ càdlàg .

Finally, notice that also the delayed functional defined via $F(t, \mathbf{X}) := \mathbf{X}_{t-\delta}$, for some $\delta > 0$ is not included in this setup as F is not invariant under reparametrization. The same holds true in the setting of Cont and Fournie (2013) (see Example 6 therein), however for different reasons.

5.4.3 Connections to the literature

The Itô-formula for rough paths In this section, we show that the functional Itô-formula formulated in Theorem 5.4.4 matches the existing Itô-formulas for rough paths that have already been formulated in the literature. Throughout, we fix $p \in [2, 3)$.

(i) Let $f \in C^3(\mathbb{R}^{d+1})$ and consider the path functional given by

$$[0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1})) \ni (t, \widehat{\mathbf{X}}) \mapsto F(t, \widehat{\mathbf{X}}) := f(\widehat{X}_t),$$

where we set $\widehat{X} := \pi_1(\widehat{\mathbf{X}})$. Observe that F is a C^3-non-anticipative Marcus canonical path functional such that for all $i = 1, 2, 3$, $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$,

$$\nabla^i F(t, \widehat{\mathbf{X}}) = \nabla^i f(\widehat{X}_t),$$

where on the right-hand side, $\nabla^i f$ denotes the $(\mathbb{R}^{d+1})^{\otimes i}$-valued maps given by the i-th (standard) partial derivatives of the map f. Moreover, from standard properties of regular functions on $\mathbb{R}^{d+1}$ we get that $F, \nabla F, \nabla^2 F$ and $R^{\nabla F, \nabla^2 F}$ are $[p,3)$-var-continuous. Therefore, an application of Theorem 5.4.4 yields that for all $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, with $\widehat{\mathbf{X}}$-Marcus like, it holds that

$$\begin{aligned} f(\widehat{X}_t) - f(\widehat{X}_0) &= \int_0^t \nabla f(\widehat{X}_{s-}) d\widehat{\mathbf{X}}_s \\ &+ \sum_{s \leqslant t} f(\widehat{X}_s) - f(\widehat{X}_{s-}) - \nabla f(\widehat{X}_{s-}) \Delta \widehat{X}_s - \nabla^2 f(\widehat{X}_{s-}) \frac{1}{2} \Delta \widehat{X}_s^{\otimes 2}. \end{aligned} \tag{5.4.15}$$

The formula in equation (5.4.15) has been derived in Theorem 2.12 in Friz and Zhang (2017) by adapting the notion of *reduced rough paths* (see Definition 5.7 in Friz and Hairer (2014)) to the jumps setting. Notice that the reasoning employed in Friz and Zhang (2017) exploit the Taylor expansion of the regular function f on $\mathbb{R}^{d+1}$. It is therefore completely different from the proof techniques used in Theorem 5.4.4.

(ii) Let $f \in C^3(\mathbb{R})$ and consider the path functional given by

$$[0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1})) \ni (t, \widehat{\mathbf{X}}) \mapsto F(t, \widehat{\mathbf{X}}) := f(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle), \tag{5.4.16}$$

for some $\mathbf{u} \in T(\mathbb{R}^{d+1})$. An application of the rules of derivation for compound functions yields that F is a C^3-non-anticipative Marcus canonical path functional such that

$$\begin{aligned} \nabla F(t, \widehat{\mathbf{X}}) &= f'(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle, \\ \nabla^2 F(t, \widehat{\mathbf{X}}) &= f''(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle^\top + f'(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(2)}, \widehat{\mathbb{X}}_t \rangle, \\ \nabla^3 F(t, \widehat{\mathbf{X}}) &= f'''(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle^\top \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle^\top \\ &\quad + 3 f''(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(2)}, \widehat{\mathbb{X}}_t \rangle \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle^\top + f'(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) \langle \mathbf{u}^{(3)}, \widehat{\mathbb{X}}_t \rangle, \end{aligned}$$

for all $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, with f', f'', f''' denoting the maps corresponding to the first, second, and third derivatives of f, respectively. A further application of the standard properties of regular functions on $\mathbb{R}$ yields that $F, \nabla F, \nabla^2 F$ and $R^{\nabla F, \nabla^2 F}$ are $[p,3)$-var-continuous. Therefore the assertion of Theorem 5.4.4 holds true for all $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, with $\widehat{\mathbf{X}}$-Marcus like. Furthermore, if in particular $\widehat{\mathbf{X}} \in (t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$, the functional Itô-formula in equation (5.4.8) coincides with the Itô-formula for controlled rough paths stated in Theorem 7.7 in Friz and Hairer (2014). Following their notation, it holds that

$$Y = \langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle, \quad Y' = \langle \mathbf{u}^{(1)}, \widehat{\mathbb{X}}_t \rangle, \quad Y'' = \langle \mathbf{u}^{(2)}, \widehat{\mathbb{X}}_t \rangle.$$

Observe that the above reasoning can be easily generalized to path functionals of the form

$$[0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1})) \ni (t, \widehat{\mathbf{X}}) \mapsto F(t, \widehat{\mathbf{X}}) := g(\langle \mathbf{u}_1, \widehat{\mathbb{X}}_t \rangle, \ldots, \langle \mathbf{u}_m, \widehat{\mathbb{X}}_t \rangle),$$

for some $g \in C^3(\mathbb{R}^m)$, $\mathbf{u}_1, \dots \mathbf{u}_m \in T(\mathbb{R}^{d+1})$.

Föllmer, RIE and stochastic integration theories Here, we elaborate on the connections of the results derived in Section 5.4 with Föllmer (see Föllmer (1981)), RIE (see Perkowski and Prömel (2016), Allan et al. (2023b)), and stochastic integration (see Jacod and Shiryaev (1987)) theories, whose main properties have been summarized in Appendix 5.B. Recall the notion of the path of finite quadratic variation along a sequence of partitions given in Definition 5.B.1. In the following, we let $\widehat{D}([0,1], \mathbb{R}^{d+1})$ denote the space of càdlàg paths obtained by extending the paths in $D([0,1], \mathbb{R}^d)$ throught the addition of a tj-path component (see Definition 5.3.1 and Remark 5.3.2 (ii)).

Corollary 5.4.7 (Föllmer)**.** *Let $p \in [2,3)$ and $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^3-non-anticipative Marcus canonical path functional such that $F, \nabla F, \nabla^2 F$ and $R^{\nabla F, \nabla^2 F}$ are $[p,3)$-var-continuous, and $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ a sequence of partitions with vanishing mesh size. Let $\widehat{X} \in \widehat{D}([0,1], \mathbb{R}^{d+1})$ be a path with finite quadratic variation in the sense of Föllmer along $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$, and $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$ a Marcus-like weakly geometric rough path such that $\pi_1(\widehat{\mathbf{X}}) = \widehat{X}$. Assume that for all $t \in [0,1]$, Anti$(\nabla^2 F(t, \widehat{\mathbf{X}})) = 0$. Then, the following limit exists (along the sequence of partitions $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$),*

$$\int_0^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{X}_s := \lim_{n\to\infty} \sum_{s_i^n \in \pi^n_{[0,1]}} \nabla F(s_i^n, \widehat{\mathbf{X}}) \widehat{X}_{s_i^n, s_{i+1}^n}, \tag{5.4.17}$$

and

$$\begin{aligned} F(t, \widehat{\mathbf{X}}) - F(0, \widehat{\mathbf{X}}) = &\int_0^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{X}_s + \frac{1}{2}\int_0^t \nabla^2 F(s^-, \widehat{\mathbf{X}}) d[\widehat{X}, \widehat{X}]^c_s \\ &+ \sum_{s\leqslant t} F(s, \widehat{\mathbf{X}}) - F(s^-, \widehat{\mathbf{X}}) - \nabla F(s^-, \widehat{\mathbf{X}}) \Delta \widehat{X}_s. \end{aligned} \tag{5.4.18}$$

The integral in (5.4.18) with respect to the continuous quadratic variation path $[\widehat{X}, \widehat{X}]^c$ is understood as a Young integral, and the summation term is well defined as an absolutely summable series.

Remark 5.4.8. (i) Since the continuous quadratic variation path $[\widehat{X}, \widehat{X}]^c$ is of finite variation, the integral in (5.4.18) is equal to a Lebesgue-Stieltjes integral (see Remarks 5.B.2 and 1.2.7 (ii)).

Proof. Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$ be a Marcus-like weakly geometric rough path such that $\pi_1(\widehat{\mathbf{X}}) = \widehat{X}$. Since by Theorem 5.3.15,

$$\{s \in (0,1] \, : \, \Delta\nabla^2 F(s, \widehat{\mathbf{X}}) \neq 0\} = \{s \in (0,1] \, : \, \Delta\widehat{\mathbf{X}}_s \neq 0\},$$

and $\widehat{\mathbf{X}}$ is Marcus-like, by Remark 5.B.2,

$$\lim_{n\to\infty}\sum_{s_i^n\in\pi_{[0,1]}^n}\nabla^2F(s_i^n,\widehat{\mathbf{X}})\widehat{X}_{s_i^n,s_{i+1}^n}^{\otimes 2}=\int_0^t\nabla^2F(s,\widehat{\mathbf{X}})d[\widehat{X},\widehat{X}]_s^c+\sum_{s\leqslant t}\nabla^2F(s,\widehat{\mathbf{X}})\Delta\widehat{X}_s^{\otimes 2}.$$

Moreover, by Theorem 5.4.4, $\big(\nabla F(\cdot,\widehat{\mathbf{X}}),\nabla^2F(\cdot,\widehat{\mathbf{X}})\big)\in\mathcal{V}_{\widehat{\mathbf{X}}}^{p',r}$, for all $p'>p$ such that $\frac{1}{p'}+\frac{2}{p}>1$. Since for all $t\in[0,1]$, $Anti(\nabla^2F(t,\widehat{\mathbf{X}}))=0$, by definition of the rough integral, we then deduce the existence of the limit in equation (5.4.17). Finally, formula (5.4.18) is derived from equation (5.4.8). □

Next, we proceed our discussion by considering the special case of càdlàg rough paths over some $\widehat{X}\in\widehat{D}([0,1],\mathbb{R}^{d+1})$ which satisfies (RIE) (see Property 5.B.6) with respect to some $p\in(2,3)$ and some sequence of nested partitions $(\pi_{[0,1]}^n)_{n\in\mathbb{N}}$. Recall from Proposition 5.B.9 that any path that satisfies (RIE) along a sequence of partition, also has quadratic variation in the sense of Föllmer along the same partition.

Corollary 5.4.9 ((RIE) property)**.** *Let $p\in(2,3)$ and $F:[0,1]\times D^p([0,1],G^2(\mathbb{R}^{d+1}))\to\mathbb{R}$ be a C^3-non-anticipative Marcus canonical path functional such that $F,\nabla F,\nabla^2F$ and $R^{\nabla F,\nabla^2F}$ are $[p,3)$-var-continuous, and $(\pi_{[0,1]}^n)_{n\in\mathbb{N}}$ a sequence of nested partitions with vanishing mesh size. Let $\widehat{X}\in\widehat{D}([0,1],\mathbb{R}^{d+1})$ be a tj-extended path which satisfies (RIE) with respect to p and $(\pi_{[0,1]}^n)_{n\in\mathbb{N}}$, and $\widehat{\mathbf{X}}\in\widehat{D}^p([0,1],G^2(\mathbb{R}^{d+1}))$ the rough path defined in equation (5.B.6) such that $\pi_1(\widehat{\mathbf{X}})=\widehat{X}$. Then, for all $t\in[0,1]$, it holds that*

$$\begin{aligned}F(t,\widehat{\mathbf{X}})-F(0,\widehat{\mathbf{X}})&=\int_0^t\nabla F(s^-,\widehat{\mathbf{X}})d\widehat{\mathbf{X}}_s+\frac{1}{2}\int_0^t\nabla^2F(s^-,\widehat{\mathbf{X}})d[\widehat{X},\widehat{X}]_s^c\\&\quad+\sum_{s\leqslant t}F(s,\widehat{\mathbf{X}})-F(s^-,\widehat{\mathbf{X}})-\nabla F(s^-,\widehat{\mathbf{X}})\Delta\widehat{X}_s.\end{aligned}\tag{5.4.19}$$

The first term on the left-hand side of (5.4.19) is interpreted as in equation (5.4.17), the second term is a Young integral with respect to $[X,X]^c$, and the summation term is well defined as an absolutely summable series.

Remark 5.4.10. Notice that Remark 5.4.8 applies also here. Therefore, the Young integral with respect to $[X,X]^c$ in (5.4.17) is equal to a Lebesgue-Stieltjes integral.

Proof. Observe that by definition, $\widehat{\mathbf{X}}$ is a Marcus-like rough path. Therefore it holds that

$$\{s\in(0,1]\,:\,\Delta\widehat{\mathbf{X}}_s\neq 0\}=\{s\in(0,1]\,:\,\Delta\widehat{X}_s\neq 0\}.$$

Moreover, by Proposition 2.14 in Allan et al. (2023b), we also get that

$$\{s\in(0,1]\,:\,\Delta\widehat{X}_s\neq 0\}\subseteq\bigcup_{n\in\mathbb{N}}\pi_{[0,1]}^n.$$

Since by Theorem 5.3.15, for $j = 1, 2$,

$$\{s \in (0,1] \;:\; \Delta\nabla^j F(s, \widehat{\mathbf{X}}) \neq 0\} = \{s \in (0,1] \;:\; \Delta\widehat{\mathbf{X}}_s \neq 0\},$$

the claim follows by Theorem 5.4.4 and Proposition 5.B.9. □

Remark 5.4.11. (i) Notice that for $\widehat{X} \in \widehat{D}([0,1], \mathbb{R}^{d+1})$ a tj-extended path as in Corollary 5.4.9, $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$ as given in (5.B.6) is a tj-extended weakly geometric rough path (see Definition 5.3.3). Therefore, Theorem 5.3.15 holds true for F evaluated at such a path.

(ii) The Itô-formula (5.4.19) coincides with the formula derived in Corollary 5.4.7. Observe that the former has been deduced under weaker assumptions on $\nabla^2 F$, but stronger ones on the path $\widehat{X}$. Specifically, in Corollary 5.4.7, we have supposed that for all $t \in [0,1]$, $Anti(\nabla^2 F(t, \widehat{\mathbf{X}})) = 0$ to ensure the convergence of the limit in (5.4.17). In contrast, the latter convergence was achieved without any assumptions on the antisymmetric part of $\nabla^2 F$ when $\widehat{X}$ satisfies property (RIE). Property (RIE) is indeed a stronger requirement for a path than having finite quadratic variation in the sense of Föllmer (see Proposition 5.B.9).

Finally, we conclude this section by analyzing path functionals evaluated at some random rough paths given as Marcus lift of some càdlàg semimartingale (see Proposition 1.4.1).

Corollary 5.4.12 (Semimartingales (stochastic))**.** *For $p \in (2,3)$, let $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^3-non-anticipative Marcus canonical path functional such that $F, \nabla F, \nabla^2 F$ and $R^{\nabla F, \nabla^2 F}$ are $[p,3)$-var-continuous. For $\widehat{X}$ a tj-extended càdlàg semimartingale, denote by $\widehat{\mathbf{X}}$ its Marcus lift as defined in equation (1.4.1). Assume that the processes $\nabla F(\cdot, \widehat{\mathbf{X}}) := (\nabla F(t, \widehat{\mathbf{X}}))_{t\in[0,1]}, \nabla^2 F(s, \widehat{\mathbf{X}}) := (\nabla^2 F(\cdot, \widehat{\mathbf{X}}))_{t\in[0,1]}$ are locally bounded and predictable. Then, for all $t \in [0,1]$, a.s.*

$$\begin{aligned} F(t, \widehat{\mathbf{X}}) - F(0, \widehat{\mathbf{X}}) = &\int_0^t \nabla F(s^-, \widehat{\mathbf{X}}) d\widehat{X}_s + \frac{1}{2}\int_0^t \nabla^2 F(s^-, \widehat{\mathbf{X}}) d[\widehat{X}, \widehat{X}]^c_s \\ &+ \sum_{s \leqslant t} F(s, \widehat{\mathbf{X}}) - F(s^-, \widehat{\mathbf{X}}) - \nabla F(s^-, \widehat{\mathbf{X}}) \Delta\widehat{X}_s. \end{aligned} \tag{5.4.20}$$

The first term on the left-hand side of (5.4.20) is an Itô-integral, the second term is a Young integral with respect to $[\widehat{X}, \widehat{X}]^c$, and the summation term is well defined as an a.s. absolutely summable series.

Remark 5.4.13. Since the continuous quadratic variation path $[\widehat{X}, \widehat{X}]^c$ a.s. of finite variation, the integral in (5.4.20) is a.s. equal to a Lebesgue-Stieltjes integral.

Proof. The assertion follows by Theorem 5.4.4, and Lemma 4.35 in Chevyrev and Friz (2019) combined with a similar argument as in Proposition 4.16 in Chevyrev and Friz (2019). □

Remark 5.4.14. (i) The functional Itô-formula in (5.4.20) evaluated at continuous paths coincides with the functional Itô-formula derived in Theorem 1 in Dupire (2019) and Theorem 4.1 in Cont and Fournie (2013) whenever the relative regularity conditions and functional representations coincide. Furthermore, we remark that in this semimartingale context, the antisymmetric part of the second-order vertical derivatives does not matter as its contribution vanishes $a.s.$ (see Lemma 4.35 in Chevyrev and Friz (2019)).

Notice in particular that since we consider here time-extended paths, the integral of the horizontal derivative of the functional (see Definition 3.1 in Cont and Fournie (2013)) with respect to the time considered in Theorem 4.1 in Cont and Fournie (2013) coincides with the integral of the vertical derivative of the functional with respect to the time component.

(ii) Recall that the framework in Dupire (2019), Cont and Fournie (2013), Cont and Fournie (2010a) for deriving a functional Itô-formula requires continuity of the functionals and their derivatives with respect to the supremum norm, as well as the verification of some boundness preserving properties (see Definition 2.5 in Cont and Fournie (2013)). This excludes several important examples of non-anticipative functionals, such as iterated Stratonovich (or Itô) stochastic integrals, which are instead covered in the present framework. Indeed, recall that linear functionals of the signature are the prototype of functionals for which Theorem 5.5.5 applies, and by the discussion in Section 1.4, linear functionals of the signature of a continuous semimartingale with Marcus lift, which reduces to the Stratonovich lift, are simply iterated Stratonovich integrals.

More generally, the regularity conditions on the functionals and their derivatives expressed with respect to a stronger topology (p-variation versus uniform topology) allow here for the consideration of a larger set of regular functionals.

(iii) Corollary 5.4.12 easily extends to the case of tj-extended semimartingales perturbed by tj-extended paths of finite q-variation and their canonical Marcus-like rough path lift, as introduced in Section 5.1 in Chevyrev and Friz (2019).

5.5 Functional Taylor expansion

The Taylor expansion is a fundamental concept in classical calculus, offering explicit polynomial approximations of smooth functions. The aim of this section is to derive a *functional Taylor expansion* of sufficiently regular Marcus-canonical path functionals in terms of the signature.

5.5.1 The case $p \in [1, 2)$

Here, as a byproduct of the functional Itô-formula proved in Theorem 5.4.1, we derive a functional Taylor expansion for non-anticipative maps of càdlàg rough paths. We start by proving the result for functionals evaluated at some continuous paths. Then, relying on the Marcus property of the functionals considered, a Taylor expansion for functionals of càdlàg paths will be proved in a subsequent corollary.

Theorem 5.5.1. *Let $p \in [1,2)$, $K \geqslant 2$ and $F : [0,1] \times D^p([0,1], G^1(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^K-non-anticipative Marcus canonical path functional such that $F, \nabla F, \dots, \nabla^{K-1}F$ are $[p,2)$-var-continuous. Let $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^1(\mathbb{R}^{d+1}))$ and denote by $\widehat{\mathbb{X}}$ its signature. Then, for every $t \in [0,1]$, it holds that*

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-2} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-1}(t, \widehat{\mathbf{X}}), \tag{5.5.1}$$

where $\mathcal{R}^F_{K-1}(t, \widehat{\mathbf{X}}) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{K-2}} \nabla^{K-1} F(t_{K-1}, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_{K-1}} \dots d\widehat{\mathbf{X}}_{t_1}$ is defined via iterated Young integrals.

Remark 5.5.2. Recall from Corollary 1.2.8 that for any $Y \in C^{p'}([0,1], G^1(\mathbb{R}^{d+1}))$ and $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^1(\mathbb{R}^{d+1}))$, with $p' > p$, $\frac{1}{p'} + \frac{1}{p} > 1$, the path defined via the Young integral as $\int_0^{\cdot} Y_s d\widehat{\mathbf{X}}_s$ has finite p-variation. Therefore, the iterated integrals defining the remainder term $\mathcal{R}^F_{K-1}(\cdot, \widehat{\mathbf{X}})$ are well-defined.

Proof. Let $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^1(\mathbb{R}^{d+1}))$. We prove the assertion by induction. If $K = 2$, then the expansion (5.5.1) is simply the statement of Theorem 5.4.1. Assume $K > 2$ and that the assertion holds true for all $l = 2, \dots, K-1$. Since by assumption F is C^K, it is in particular C^{K-1}, thus by the induction hypothesis, it holds that

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-3} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}),$$

for $\mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{K-3}} \nabla^{K-2} F(t_{K-2}, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_{K-2}} \dots d\widehat{\mathbf{X}}_{t_1}$. Next, since $\nabla^{K-2}F$ is a C^2-non-anticipative Marcus canonical path functional (see Remark 5.3.13) and $\nabla^{K-2}F, \nabla^{K-1}F$ are $[p,2)$-var-continuous, a component-wise application of Theorem 5.4.1 to $\nabla^{K-2}F$ yields that for all $s \in [0,1]$,

$$\nabla^{K-2} F(s, \widehat{\mathbf{X}}) = \nabla^{K-2} F(0, \widehat{\mathbf{X}}) + \int_0^s \nabla^{K-1} F(r, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_r. \tag{5.5.2}$$

The claim follows by replacing the expression in equation (5.5.2) into $\mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}})$. □

Corollary 5.5.3. *Let $p \in [1,2)$, $K \geqslant 2$ and $F : [0,1] \times D^p([0,1], G^1(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^K-non-anticipative Marcus canonical path functional such that $F, \nabla F, \dots, \nabla^{K-1}F$ are $[p,2)$-var-continuous. Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^1(\mathbb{R}^{d+1}))$, denote by $\widehat{\mathbb{X}}$ its signature and by* $\mathbf{Z}$

its time-extended Marcus-transformed path with respect to some pair (R, ψ_R)*. Then, for all* $t \in [0,1]$*, it holds that*

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-2} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-1}(t, \widehat{\mathbf{X}}),$$

where $\mathcal{R}^F_{K-1}(t, \widehat{\mathbf{X}}) := \int_0^{\psi_R^{-1}(\tau_{\widehat{\mathbf{X}},R}(t))} \int_0^{t_1} \dots \int_0^{t_{k-2}} \nabla^{K-1} F(t_{k-1}, \mathbf{Z}) d\mathbf{Z}_{t_{k-1}} \dots d\mathbf{Z}_{t_1}$ *is defined via iterated Young integrals.*

Proof. Let $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^1(\mathbb{R}^{d+1}))$ and $\mathbf{Z}$ the time-extended Marcus-transformed path of $\widehat{\mathbf{X}}$. By Theorem 5.5.1 the expansion (5.5.1) holds true when evaluating the functional F at $(\psi_R^{-1}(\tau_{\widehat{\mathbf{X}},R}(t)), \mathbf{Z})$. Since for all $j = 0, \dots, K-2$, $\nabla^j F$, $\widehat{\mathbb{X}}^{(j)}$ are Marcus canonical path functional, the claim follows. □

Remark 5.5.4. Notice that since the proofs of Theorem 5.5.1 and Corollary 5.5.3 rely on the functional Itô-formula stated in Theorem 5.4.1, the considerations regarding a possible weakening of the hypothesis on the functional F discussed in Remark 5.4.3 also apply here.

5.5.2 The case $p \in [2,3)$

In this section, we derive a Taylor expansion for a sufficiently regular path functional of càdlàg p-rough paths, for $p \in [2,3)$, via an iterative application of the functional Itô-formula (5.4.8).

Theorem 5.5.5. *Let* $p \in [2,3)$*,* $K \geqslant 3$ *and* $F : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$ *be a* C^K*- non-anticipative path functional such that* $F, \nabla F, \dots, \nabla^{K-1} F$ *and* $R^{\nabla^{K-2}F, \nabla^{K-1}F}$ *are* $[p,3)$*-var-continuous. Let* $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$ *and denote by* $\widehat{\mathbb{X}}$ *its signature. Then, for every* $t \in [0,1]$ *it holds that*

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-3} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}), \tag{5.5.3}$$

where $\mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{K-3}} \nabla^{K-2} F(t_{K-2}, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_{K-2}} \dots d\widehat{\mathbf{X}}_{t_1}$ *is defined via iterated rough integrals.*

Remark 5.5.6. The iterated rough integrals determining the remainder term $\mathcal{R}^F_{K-2}(\cdot, \widehat{\mathbf{X}})$ are defined as follows. Set $G^{-1}(\cdot, \widehat{\mathbf{X}}) := \nabla^{K-1} F(\cdot, \widehat{\mathbf{X}})$, $G^0(\cdot, \widehat{\mathbf{X}}) := \nabla^{K-2} F(\cdot, \widehat{\mathbf{X}})$ and

$$G^j(\cdot, \widehat{\mathbf{X}}) := \int_0^{\cdot} \int_0^{t_1} \dots \int_0^{t_{j-1}} \nabla^{K-2} F(t_j, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_j} \dots d\widehat{\mathbf{X}}_{t_1}, \tag{5.5.4}$$

for $j = 1, \dots, K-2$. Then, $G^j(\cdot, \widehat{\mathbf{X}})$ is the rough integral of $(G^{j-1}(\cdot, \widehat{\mathbf{X}}), G^{j-2}(\cdot, \widehat{\mathbf{X}}))$ with respect to $\widehat{\mathbf{X}}$, which is well-defined by Corollary 1.2.13.

Proof. Similar to the proof of Theorem 5.4.4, here we first derive the Taylor expansion for functionals evaluated at the truncated signature at level 2 of a time-extended piecewise linear path. Then, we extend it to general continuous paths by a density argument. This approach simplifies the proof development as it enables working with Young integrals (and thus Riemann sums) instead of truly rough integrals (and thus compensated Riemann sums, see Proposition 1.2.11) in the derivation of the remainder term $\mathcal{R}^F_{K-2}$. This is possible because any rough integral with respect to the level-2 signature of a $\mathbb{R}^d$-valued path of finite variation is simply a Young integral with respect to the path itself (see e.g., Lemma 1.2.15).

Step 1: Let $\widehat{\mathbf{X}} \in \widehat{C}^1([0,1], G^2(\mathbb{R}^{d+1}))$ be the truncated signature at level 2 of a time-extended piecewise linear path. We prove the assertion by induction. If $K = 3$, then the expansion in equation (5.5.3) is simply the statement of Theorem 5.4.4. Assume $K > 3$ and that the assertion holds true for all $j = 3, \dots, K-1$. Since F is in C^K, it is in particular in C^{K-1}. By the induction hypothesis,

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-4} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-3}(t, \widehat{\mathbf{X}}),$$

for $\mathcal{R}^F_{K-3}(t, \widehat{\mathbf{X}}) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{K-4}} \nabla^{K-3} F(t_{K-3}, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_{K-3}} \dots d\widehat{\mathbf{X}}_{t_1}$. Thus, if we prove that

$$\mathcal{R}^F_{K-3}(t, \widehat{\mathbf{X}}) = \nabla^{K-3} F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(K-3)}_{0,t} + \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}), \tag{5.5.5}$$

the claim follows. Since $\nabla^{K-3} F$ is a C^3 non-anticipative Marcus canonical path functional (see Remark 5.3.13), and $\nabla^{K-3} F, \nabla^{K-2} F, \nabla^{K-1} F$, and $R^{\nabla^{K-2} F, \nabla^{K-1} F}$ are $[p, 3)$-var-continuous, a component-wise application of Theorem 5.4.4 to $\nabla^{K-3} F$ yields that for all $s \in [0,1]$,

$$\nabla^{K-3} F(s, \widehat{\mathbf{X}}) = \nabla^{K-3} F(0, \widehat{\mathbf{X}}) + \int_0^s \nabla^{K-2} F(r, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_r. \tag{5.5.6}$$

Finally, since any rough integral with respect to $\widehat{\mathbf{X}}$ is nothing else than a Young integral with respect to $\pi_1(\widehat{\mathbf{X}})$, (see e.g., Lemma 1.2.15), the claim follows by replacing the expression in equation (5.5.6) into $\mathcal{R}^F_{K-3}(t, \widehat{\mathbf{X}})$.

Step 2: Fix $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$. By Theorem 5.23 in Friz and Victoir (2010), there exists a sequence of piecewise linear time-extended paths such that their truncated signature at level 2, denoted by $(\widehat{\mathbf{X}}^M)_{M \in \mathbb{N}}$, satisfies

$$\lim_{M \to \infty} d_{CC} \sup_{t \in [0,1]} (\widehat{\mathbf{X}}^M_t, \widehat{\mathbf{X}}_t) = 0, \qquad \text{and} \qquad \sup_{M \in \mathbb{N}} \|\widehat{\mathbf{X}}^M\|_{p\text{-}var} \leqslant C \|\widehat{\mathbf{X}}\|_{p\text{-}var},$$

for some $C > 0$. By Step 1, for every fixed M and $t \in [0,1]$,

$$F(t, \widehat{\mathbf{X}}^M) = \sum_{j=0}^{K-3} \nabla^j F(0, \widehat{\mathbf{X}}^M) \widehat{\mathbb{X}}_{0,t}^{M,(j)} + \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}^M).$$

Thus, since for all $j = 0, \ldots, K-3$, $\nabla^j F$ and $\widehat{\mathbb{X}}^{(j)}$ are $[p,3)$-var-continuous (see Proposition 5.2.26), the claim follows if we prove that for every $t \in [0,1]$,

$$\lim_{M \to \infty} \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}^M) = \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}).$$

To this end, set $G^{-1} := \nabla^{K-1} F$, $G^0 := \nabla^{K-2} F$ and define for $j = 1, \ldots, K-2$,

$$G^j(t, \widehat{\mathbf{X}}) := \int_0^t \int_0^{t_1} \ldots \int_0^{t_{j-1}} \nabla^{K-2} F(t_j, \widehat{\mathbf{X}}) d\widehat{\mathbf{X}}_{t_j} \ldots d\widehat{\mathbf{X}}_{t_1}.$$

We show that there exists $p := p^{(0)} < p^{(1)} < \ldots p^{(K-2)} < 3$ such that for all $j = 1, \ldots, K-2$,

$$\begin{aligned} &\lim_{M \to \infty} d_{p^{(j)}}(G^{j-2}(\cdot, \widehat{\mathbf{X}}^M), G^{j-2}(\cdot, \widehat{\mathbf{X}})) = 0, \\ &\lim_{M \to \infty} d_{p^{(j)}/2}\big(R^{G^{j-1}, G^{j-2}}((\cdot,\cdot), \widehat{\mathbf{X}}^M), R^{G^{j-1}, G^{j-2}}((\cdot,\cdot), \widehat{\mathbf{X}})\big) = 0, \end{aligned} \tag{5.5.7}$$

and $\lim_{M \to \infty} d_{CC}(G^{j-1}(0, \widehat{\mathbf{X}}^M), G^{j-1}(0, \widehat{\mathbf{X}}^M)) = 0$. Then, a similar argument as in the Step 2 of the proof of Theorem 5.4.4 (see equation (5.4.14)) yields that for all $j = 1, \ldots, K-2$,

$$\lim_{M \to \infty} d_{p^{(j)}\text{-}var}\big(G^j(\cdot, \widehat{\mathbf{X}}^M), G^j(\cdot, \widehat{\mathbf{X}})\big) = 0. \tag{5.5.8}$$

Since $G^{K-2} = \mathcal{R}^F_{K-2}$, the claim follows.

We reason by induction. Let $j = 1$ and fix $p^{(1)} > p$. By interpolation, we get $\lim_{M \to \infty} d_{p^{(1)}\text{-}var}\big(\widehat{\mathbf{X}}^M, \widehat{\mathbf{X}}\big) = 0$. Since by assumption G^{-1}, G^0, and $R^{G^0, G^{-1}}$ are $[p,3)$-var continuous, the claim follows.

Assume the assertion holds true for all $l = 1, \ldots, j$, with $K - 3 \geqslant j > 1$, and fix $p^{(j+1)} > p^{(j)} > p^{(j-1)}$. By the induction hypothesis,

$$\lim_{M \to \infty} d_{p^{(j)}\text{-}var}\big(G^{j-1}(\cdot, \widehat{\mathbf{X}}^M), G^{j-1}(\cdot, \widehat{\mathbf{X}})\big) = 0. \tag{5.5.9}$$

Moreover, an application of equation (2.3.3) in Theorem 31 in Friz and Shekhar (2017), yields that

$$\sup_{M \in \mathbb{N}} \|R^{G^j, G^{j-1}}((\cdot,\cdot), \widehat{\mathbf{X}}^M)\|_{p^{(j)}/2\text{-}var} < \infty.$$

Therefore, by interpolation

$$\lim_{M \to \infty} d_{p^{(j+1)}/2}\big(R^{G^j, G^{j-1}}((\cdot,\cdot), \widehat{\mathbf{X}}^M), R^{G^j, G^{j-1}}((\cdot,\cdot), \widehat{\mathbf{X}})\big) = 0.$$

Since the convergence in (5.5.9) holds true also with respect $d_{p^{(j+1)}}$, the claim follows. $\square$

Finally, we derive a Taylor expansion for functionals of càdlàg paths exploiting the Marcus-canonical properties of the functionals considered. The proof of the following corollary follows step by step the proof of Corollary 5.5.3. Therefore, we omit it.

Corollary 5.5.7. *Let $p \in [2,3)$ and $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^K, $K \geqslant 3$, non-anticipative Marcus canonical path functional such that $F, \nabla F, \ldots, \nabla^{K-1}F$ and $R^{\nabla^{K-2}F, \nabla^{K-1}F}$ are $[p,3)$-var-continuous. Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, denote by $\widehat{\mathbb{X}}$ its signature, and by $\mathbf{Z}$ its time-extended Marcus-transformed path with respect to some pair (R, ψ_R). Then, for all $t \in [0,1]$, it holds that*

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{K-3} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t} + \mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}),$$

where $\mathcal{R}^F_{K-2}(t, \widehat{\mathbf{X}}) := \int_0^{\psi_R(\tau_{\widehat{\mathbf{X}},R}(t))} \int_0^{t_1} \ldots \int_0^{t_{K-3}} \nabla^{K-1} F(t_{k-1}, \mathbf{Z}) d\mathbf{Z}_{t_{k-1}} \ldots d\mathbf{Z}_{t_1}$ is defined via iterated rough integral.

Remark 5.5.8. (i) Since the proofs of Theorem 5.5.5 and Corollary 5.5.7 rely on Theorem 5.4.4, the considerations regarding a possible modifications of the conditions on the functional F discussed in Remark 5.4.6(i)(ii) also apply here.

(ii) A Taylor expansion of path functionals in terms of signature has been recently proved in Dupire and Tissot-Daguette (2023). In particular, adopting the pathwise framework pioneered by Föllmer (1981) and combining it with the (functional) differential calculus introduced in Dupire (2019), they prove a pathwise functional Taylor expansion for continuous paths of finite quadratic variations in the sense of Föllmer (see Theorem 3.10 in Dupire and Tissot-Daguette (2023)). However, recalling the discussion in Remark 5.2.23 on the commutativity of the derivation order, it is not surprising that their expansion is presented only for functionals of *continuous one-dimensional time-extended paths* of finite quadratic variation in the sense of Föllmer, for which the dependency with respect to the time and path component is captured via the horizontal and vertical derivatives, respectively.

Indeed, in this framework, only the horizontal and vertical derivatives do not commute (see the discussion at page. 33 in Cont et al. (2016)). Therefore, considering time-extended one-dimensional paths, they can indeed recover all terms in the signature that appear in the expansion. In a multidimensional setting, however, since the higher-order (mixed) vertical derivatives commute, by Lemma 1.1.9, any expansion in terms of the signature can not be obtained.

This remark further emphasizes the importance of establishing a differential calculus on path functionals that allows a non-commutative order of differentiation, without which it would not be possible to achieve Taylor expansions in terms of the signature.

Analytic non-anticipative path functionals Building on the Taylor expansions derived in the preceding section, we mimic here the classical analysis approach to the study of smooth functions and introduce a notion of analytical path functional. For simplicity, we address these concepts specifically in the case of càdlàg p-rough path, for $p \in [2,3)$. However, note that they can also be formulated within the simpler setting of càdlàg p-rough path, for $p \in [1,2)$. We start by providing the definition of the concatenation path.

Definition 5.5.9. Let $\mathbf{X} = \exp^{(2)}(X, A^X)$ and $\mathbf{Y} = \exp^{(2)}(Y, A^Y) \in D([0,1], G^2(\mathbb{R}^d))$. We define the *concatenation path* of $\mathbf{X}$ and $\mathbf{Y}$ at time $t \in [0,1]$ as the path

$$\mathbf{X} \oplus_t \mathbf{Y} \in D([0,1], G^2(\mathbb{R}^d)),$$

defined as follows: for all $u \in [0,1]$

$$(\mathbf{X} \oplus_t \mathbf{Y})_u := \exp^{([p])}((X_u, A^X_u)1_{\{u \leqslant t\}}) \otimes \exp^{([p])}((Y_u - Y_t + X_t, A^Y_u - A^Y_t - A^X_t)1_{\{u > t\}}).$$

Remark 5.5.10. Notice that $\mathbf{X} \oplus_t \mathbf{Y} \in D([0,1], G^{[p]}(\mathbb{R}^d))$. Moreover $(\mathbf{X} \oplus_t \mathbf{Y})^t = \mathbf{X}^t$ and if both $\mathbf{X}$, $\mathbf{Y}$ have finite p-variation, for some $p > 0$, the same holds for the concatenation path (for all $t \in [0,1]$).

Next, we use the concept of concatenation path to derive the Taylor expansion of a path functional within a neighborhood of a given path.

Corollary 5.5.11. *Let $p \in [2,3)$, $K \geqslant 3$ and $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^K-non-anticipative path functional such that $F, \nabla F, \ldots, \nabla^{K-1}F$ and $R^{\nabla^{K-2}F, \nabla^{K-1}F}$ are $[p,3)$-var-continuous. Then, for all $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$, $\widehat{\mathbf{Y}} \in \widehat{C}^p([0,1], G^2(\mathbb{R}^{d+1}))$, $t, s \in [0,1]$, $t \leqslant s$ it holds that*

$$F(s, \widehat{\mathbf{X}} \oplus_t \widehat{\mathbf{Y}}^s) = \sum_{j=0}^{K-3} \nabla^j F(t, \widehat{\mathbf{X}}) \widehat{\mathbb{Y}}^{(j)}_{t,s} + \mathcal{R}^F_{K-2}(s, \widehat{\mathbf{X}} \oplus_t \widehat{\mathbf{Y}}), \tag{5.5.10}$$

where, $\mathcal{R}^F_{K-2}(s, \widehat{\mathbf{X}} \oplus_t \widehat{\mathbf{Y}}) := \int_t^s \int_t^{s_1} \ldots \int_t^{s_{K-4}} \nabla^{K-2} F(s_{K-2}, \widehat{\mathbf{Y}} \oplus_t \widehat{\mathbf{Y}}) d\widehat{\mathbf{Y}}_{s_{K-2}} \ldots d\mathbf{X}_{s_1}$, and $\widehat{\mathbb{Y}}$ denotes the signature of $\widehat{\mathbf{Y}}$.

Proof. The proof is exactly the same as the proof of Theorem 5.5.5, except that one needs to derive the expansion between s and t instead of 0 and t. □

Remark 5.5.12. The formula (5.5.10) for F evaluated at a càdlàg concatenation path can be derived as in Corollary 5.5.7, and results in an expansion whose remainder term $\mathcal{R}^F_{K-2}$ is expressed in terms of the time-extended Marcus transformed path of the concatenation path.

Recall from Remark 5.5.8(i) that the formula (5.5.10) can be derived also under the assumption of F being a C^K-non-anticipative path functional such that

$$F, \nabla F, \ldots, \nabla^{K-1}F, \nabla^K F$$

are $[p,3)$-var-continuous. Now, we use this slightly stronger assumption in order to introduce the notion of analytical and entire path functional. In the following, we say that a Marcus-canonical path functional is C^∞ if it is C^K, for all $K \in \mathbb{N}$.

Definition 5.5.13. Let $p \in [2,3)$ and $F : [0,1] \times D^p([0,1], G^2(\mathbb{R}^{d+1})) \to \mathbb{R}$ be a C^∞-Marcus-canonical non-anticipative path functional such that all its vertical derivatives $\nabla^j F$ are $[p,3)$-var-continuous. Fix $(t, \widehat{\mathbf{X}}) \in [0,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$.

(i) We say that F is *real analytic* at $(t, \widehat{\mathbf{X}})$ if there exists $\delta > 0$ such that for all $(s, \widehat{\mathbf{Y}}) \in [t,1] \times \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$ with $|s-t| + \|\widehat{\mathbf{Y}}\|_{p\text{-}var[t,s]} < \delta$ it holds that

$$F(s, \widehat{\mathbf{X}} \oplus_t \widehat{\mathbf{Y}}) = \sum_{j=0}^{\infty} \nabla^j F(t, \widehat{\mathbf{X}}) \widehat{\mathbb{Y}}^{(j)}_{t,s}.$$

(ii) We say that F is *entire* if for all $(t, \widehat{\mathbf{X}}) \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, it holds that

$$F(t, \widehat{\mathbf{X}}) = \sum_{j=0}^{\infty} \nabla^j F(0, \widehat{\mathbf{X}}) \widehat{\mathbb{X}}^{(j)}_{0,t}.$$

Example 5.5.14. Let $\mathbf{u} \in T^N(\mathbb{R}^{d+1})$, for some $N \in \mathbb{N}$ and consider the linear functional of the signature $F^{\mathbf{u}}$ (see Section 5.2.4). Then, by Proposition 5.2.26 $F^{\mathbf{u}}$ is an entire functional and for all $(t, \widehat{\mathbf{X}}) \in \widehat{D}^p([0,1], G^2(\mathbb{R}^{d+1}))$, setting $F^{\mathbf{u}}(t, \widehat{\mathbf{X}}) := \langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle$, it holds that

$$\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle = \sum_{j=0}^{N} \langle \mathbf{u}^{(j)}, \widehat{\mathbb{X}}_0 \rangle \widehat{\mathbb{X}}^{(j)}_{0,t}. \tag{5.5.11}$$

Notice that equation (5.5.11) coincides with Chen's relation (see Proposition 1.3.6). More generally, one can also consider path functional of the form $F(t, \widehat{\mathbf{X}}) := f(\langle \mathbf{u}, \mathbb{X}_t \rangle)$ for an entire function $f : \mathbb{R} \to \mathbb{R}$. Similar computations as in (5.4.16) show that F is an entire path functional.

5.A Proof of Section 5.3

Before presenting the proofs of the main results of the section, we list some key notions and introduce a lighter notation. Recall that

$$G((\mathbb{R}^{d+1})) := \{\mathbf{x} \in T((\mathbb{R}^{d+1})) \mid \pi_{\leqslant N}(\mathbf{x}) \in G^N(\mathbb{R}^{d+1}) \text{ for all } N \in \mathbb{N}\},$$

for $G^N(\mathbb{R}^{d+1}) := \exp^{(N)}(\mathfrak{g}^N(\mathbb{R}^{d+1}))$, $\exp^{(N)}$ given in 2.3.3, and

$$\mathfrak{g}^N(\mathbb{R}^{d+1}) := \{0\} \oplus \mathbb{R}^{d+1} \oplus [\mathbb{R}^{d+1}, \mathbb{R}^{d+1}] \oplus \cdots \oplus [\mathbb{R}^{d+1}, [\mathbb{R}^d, \ldots [\mathbb{R}^{d+1}, \mathbb{R}^{d+1}]]], \tag{5.A.1}$$

with the last term in (5.A.1) composed of $(N-1)$-brackets. Notice that $G^N(\mathbb{R}^{d+1})$ is a Carnot group (see Definition 2.2.1 in Bonfiglioli et al. (2007)). Moreover, set

$$\mathfrak{g}((\mathbb{R}^{d+1})) := \{\mathbf{x} \in T((\mathbb{R}^{d+1})) \mid \pi_{\leqslant N}(\mathbf{x}) \in \mathfrak{g}^N(\mathbb{R}^{d+1}) \text{ for all } N \in \mathbb{N}\}.$$

Then $G((\mathbb{R}^{d+1})) = \exp(\mathfrak{g}((\mathbb{R}^{d+1})))$ (see e.g., Bank et al. (2024), Schmeding (2022)) and is closed with respect to the tensor multiplication $\otimes$.

Recall furthermore that by Definition 5.2.4 if $\mathbf{X} \in D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$, $\widetilde{\mathbf{X}}$ denotes its Marcus transformed path with respect to some pair (R, ψ_R) and that by construction for all $t \in [0,1]$,

$$\mathbf{X}_t = \widetilde{\mathbf{X}}_{\psi_R^{-1}(\tau_{\mathbf{X},R}(t))}.$$

Fix $F \in C^K$ and let g be the map for which Assumption 5.3.7 holds. For $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$, let $g^{\mathbf{X},K}$ be the map introduced in equation (5.3.4). To simplify the notation, we write $g^K := g^{\mathbf{X},K}$ whenever no confusion arises, and write $\mathfrak{g}^K$ and $|\beta|_{\mathfrak{g}^K}$ in place of $\mathfrak{g}^K(\mathbb{R}^{d+1})$ and $|\beta|_{\mathfrak{g}^K(\mathbb{R}^{d+1})}$ (see Notation 5.3.6), respectively. Finally, we denote with the index 0 the first component of $\mathbf{X}$.

5.A.1 Proof of Lemma 5.3.9

Fix $\mathbf{X} \in C^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ and let $(t, \boldsymbol{\xi}) \in [0,1] \times \mathfrak{g}^K$. By definition of signature and $G((\mathbb{R}^{d+1}))$,

$$\mathbb{X}_t \otimes \exp(\boldsymbol{\xi}, 0, \ldots, 0) \in G((\mathbb{R}^{d+1})).$$

Therefore, by Assumption 5.3.7, the map g^K is well-defined. Next, fix $t \in [0,1]$. We show that for all $j = 1, \ldots, K$ and $|\beta|_{\mathfrak{g}^K} = j$, the derivatives $D^\beta_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0}$ exist.

The claim follows if we prove that for all $j = 1, \ldots, K$, $i_1, \ldots, i_j = 0, 1, \ldots, d$,

$$U^{i_j} \ldots U^{i_1} F(t, \mathbf{X}) = \frac{d^j}{dh_j \ldots dh_1} g(\mathbb{X}_t \otimes \exp(h_j \epsilon_{i_j}) \otimes \cdots \otimes \exp(h_1 \epsilon_{i_1}))|_{h_j = \cdots = h_1 = 0}. \tag{5.A.2}$$

Indeed, since by assumption $F \in \mathcal{M}^K_{[p]}$, the quantities $U^{i_j} \ldots U^{i_1} F(t, \mathbf{X})$ are well-defined. Furthermore, by Propositions 20.1.7 (see in particular equation (20.20)) and 20.1.9 (see also Proposition 20.1.4) in Bonfiglioli et al. (2007), setting

$$L(t, h_1\epsilon_{i_1}, \ldots, h_j\epsilon_{i_j}) := g(\mathbb{X}_t \otimes \exp(h_j\epsilon_{i_j}) \otimes \cdots \otimes \exp(h_1\epsilon_{i_1})),$$

it holds that for all $|\beta|_{\mathfrak{g}^K} = j$,

$$D^\beta_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} \in \operatorname{span}\Big\{ \frac{d^j}{dh_j \ldots dh_1} L(t, h_1\epsilon_{i_1}, \ldots, h_j\epsilon_{i_j})|_{h_1 = \cdots = h_j = 0},\ i_1, \ldots, i_j = 0, \ldots, d \Big\},$$

from which we deduce that $D^{\beta}_{\boldsymbol{\xi}} g^K(t,\boldsymbol{\xi})|_{\boldsymbol{\xi}=0}$ exists. In order to show (5.A.2) we apply the iterative procedure described in Remark 5.2.21 (ii) (see also Example 5.2.22) and compute the higher order vertical derivatives of F.

Fix $\delta > 0$ and for $i_K = 0, 1, \ldots, d$, $h_K \in (-\delta, \delta)$ set

$$\begin{aligned}
\mathbf{X}^{[i_K]}(h_K) &:= \mathbf{X} + h_{i_K}\epsilon_{i_K}1_{\{\cdot \geqslant \mu_t^{[0]}\}},\\
\mathbf{Y}^{[i_K]}(h_K) &:= \widetilde{\mathbf{X}^{[i_K]}(h_K)},\\
\mu_t^{[i_K]}(h_K) &:= \tau_{\mathbf{X}^{[i_K]}(h_K),R_K}(t).
\end{aligned}$$

For $j = K-1, \ldots, 1$, $i_j = 0, 1, \ldots, d$, $h_j \in (-\delta, \delta)$, set

$$\begin{aligned}
\mathbf{X}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j) &:= \mathbf{Y}^{[i_K,\ldots,i_{j+1}]}(h_K,\ldots,h_{j+1}) + h_j\epsilon_{i_j}1_{\{\cdot \geqslant \mu_t^{[i_K,\ldots,i_{j+1}]}\}},\\
\mathbf{Y}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j) &:= \widetilde{\mathbf{X}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j)},\\
\mu_t^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j) &:= \psi_{R_j}^{-1}(\tau_{\mathbf{X}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j),R_j}(\mu_t^{[i_K,\ldots,I_{j+1}]}(h_K,\ldots,h_{j+1}))),
\end{aligned}$$

for $\widetilde{\mathbf{X}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_j)}$ denoting the Marcus transformed path of $\mathbf{X}^{i_K,\ldots,i_j]}(h_K,\ldots,h_j)$ with respect to some pair (R_j, ψ_{R_j}) that might depend on $(h_K,\ldots,h_j)$.

Then, by definition of higher-order vertical derivatives (see Definition 5.2.19), it holds that for all $j = 1, \ldots, K$

$$\begin{aligned}
&U^{i_j}\ldots U^{i_1}F(t,\mathbf{X})\\
&= \frac{d^j}{dh_j\ldots dh_1}F(\mu_t^{[i_K,\ldots,i_1]}(0,\ldots,0,h_j,\ldots,h_1),\mathbf{Y}^{[i_K,\ldots,i_j]}(0,\ldots,0,h_j,\ldots,h_1))|_{h_j=\cdots=h_1=0}.
\end{aligned}$$

Finally, Assumption 5.3.7 and a direct computation of the signature at time $\mu_t^{[i_K,\ldots,i_1]}(h_K,\ldots,h_1)$ of the continuous path $\mathbf{Y}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_1)$ yield that

$$F(\mu_t^{[i_K,\ldots,i_1]}(h_K,\ldots,h_1),\mathbf{Y}^{[i_K,\ldots,i_j]}(h_K,\ldots,h_1)) = g(\mathbb{X}_t \otimes \exp(h_j\epsilon_{i_j})\otimes\cdots\otimes\exp(h_1\epsilon_{i_1})).$$

The claim follows. □

5.A.2 Proof of Theorem 5.3.15

Before delving into the proof of Theorem 5.3.15 we outline its main steps. Recall the simplified notation introduced at the beginning of the present section.

Sketch of the proof Fix $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$.

Step 1: We show that for $t \in [0,1]$, $i_1, \ldots, i_j \in \{0, \ldots, d\}^j$, $j = 1, \ldots, K$,

$$\begin{aligned}
&F(t, \widehat{\mathbf{X}}) = g^K(t, 0), \\
&U^{i_j} \ldots U^{i_1} F(t, \widehat{\mathbf{X}}) \in \operatorname{span}\{D^{\beta}_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} \ : \ \beta \in \mathbb{N}_0^M, \ |\beta|_{\mathfrak{g}^K} = j\}.
\end{aligned}$$

Step 2: Since $F \in C^K$, Assumption 5.3.10 an adaptation of the proof of the Weierstrass approximation theorem (see Theorem 1.6.2 in Narasimhan (1985)) to the present setting yields that there exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ on $[0,1] \times \mathcal{U}(0)$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{\boldsymbol{\xi}\in H} \sum_{\beta\in\mathbb{N}_0^M, \ |\beta|_{\mathfrak{g}^K} \leqslant K} \| D^{\beta}_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi}) - D^{\beta}_{\boldsymbol{\xi}} p_n(t, \boldsymbol{\xi}) \| = 0,$$

for some compact set $H \subset \mathcal{U}(0)$ such that $0 \in H$.

Step 3: We define $\mathbf{Y} \in \widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+2}))$ as the time-extended path of $\widehat{\mathbf{X}}$ (Section 1.2.4 and in particular Remark 1.2.18 (iv)), and denote this auxiliary component by the index -1 and its signature by $\mathbb{Y}$. Let p be a polynomial on $[0,1] \times \mathcal{U}(0)$. Notice that the map

$$[0,1] \ni t \mapsto \tilde{p}(t)(\cdot) := \Big(\mathcal{U}(0) \ni \boldsymbol{\xi} \mapsto p(t, \boldsymbol{\xi})\Big)$$

belongs to $C\big([0,1], C^K(\mathcal{U}(0))\big)$, where $C^K(\mathcal{U}(0))$ denotes the set of K-times differentiable functions on $\mathcal{U}(0)$, which we endow with the topology of uniform convergence on compacts of the function and all its derivatives up to order K. We exploit the Stone-Weierstrass theorem for vector-valued maps (see e.g., Theorem 3.3 in Cuchiero et al. (2023b)) to show that there exists a sequence $(\mathbf{v}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+2})$ whose indices are in $\{-1, 0, \ldots, d\}$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sup_{\boldsymbol{\xi}\in H} \sum_{\beta\in\mathbb{N}_0^M, \ |\beta| \leqslant K} \| D^{\beta}_{\boldsymbol{\xi}} p(t, \boldsymbol{\xi}) - D^{\beta}_{\boldsymbol{\xi}} \langle \mathbf{v}_n, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \ldots, 0) \rangle \| = 0, \quad (5.A.3)$$

where $i : g^K(\mathbb{R}^{d+1}) \to g^K(\mathbb{R}^{d+2})$ denotes the embedding of $g^K(\mathbb{R}^{d+1})$ into $g^K(\mathbb{R}^{d+2})$ obtained by setting equal to 0 all the additional components.

Step 4: A combination of Step 2 and Step 3 yields the existence of a sequence $(\mathbf{v}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+2})$ whose indices are in $\{-1, 0, \ldots, d\}$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sum_{\beta\in\mathbb{N}_0^M, \ |\beta|_{\mathfrak{g}^K} \leqslant K} \| D^{\beta}_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} - D^{\beta}_{\boldsymbol{\xi}} \langle \mathbf{v}_n, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \ldots, 0) \rangle|_{\boldsymbol{\xi}=0} \| = 0. \tag{5.A.4}$$

Step 5: We show that there exists a sequence $(\mathbf{u}_n)_{n\in\mathbb{N}} \subset T(\mathbb{R}^{d+1})$ whose indices are in $\{0,\dots,d\}$ such that for all $n \in \mathbb{N}$, $t \in [0,1]$, $\beta \in \mathbb{N}_0^M$, $|\beta|_{\mathfrak{g}^K} \leqslant K$,

$$D^{\beta}_{\boldsymbol{\xi}}\langle \mathbf{v}_n, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \dots, 0)\rangle|_{\boldsymbol{\xi}=0} = D^{\beta}_{\boldsymbol{\xi}}\langle \mathbf{u}_n, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0)\rangle|_{\boldsymbol{\xi}=0}.$$

Step 6: For $\mathbf{u} \in T(\mathbb{R}^{d+1})$, consider the path functional given by

$$[0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^d)) \ni (t, \mathbf{X}) \mapsto F^{\mathbf{u}}(t, \mathbf{X}) := \langle \mathbf{u}, \mathbb{X}_t\rangle.$$

Recall that for such functional, the map g^K in (5.3.4) reads as $g^K(t,\boldsymbol{\xi}) := \langle \mathbf{u}, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0)\rangle$. An application of Step 1 to the functional $F^{\mathbf{u}}$ yields that for all $t \in [0,1]$, $i_1, \dots, i_j \in \{0, \dots, d\}^j$, $j = 1, \dots, K$,

$$\begin{aligned}
&F^{\mathbf{u}}(t, \widehat{\mathbf{X}}) = \langle \mathbf{u}, \widehat{\mathbb{X}}_t\rangle,\\
&U^{i_j} \dots U^{i_1} F^{\mathbf{u}}(t, \widehat{\mathbf{X}}) \in \operatorname{span}\{D^{\beta}_{\boldsymbol{\xi}}\langle \mathbf{u}, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0)\rangle|_{\boldsymbol{\xi}=0} \ : \ \beta \in \mathbb{N}_0^M,\ |\beta|_{\mathfrak{g}^K} = j\},
\end{aligned}$$

which concludes the first part of the proof.

Step 7: For $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$, we deduce the claim by exploiting the Marcus property of the functionals.

Proof. Fix $\widehat{\mathbf{X}} \in \widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$.

Step 1: First of all notice that $F(t, \mathbf{X}) = g(\mathbb{X}_t) = g^K(t, 0)$ by Assumption 5.3.7 and definition of g^K in (5.3.4). Then, an inspection of the proof of Lemma 5.3.9 shows that for all $j = 1, \dots, K$, $i_1, \dots, i_j = 0, \dots, d$

$$U^{i_j} \dots U^{i_1} F(t, \widehat{\mathbf{X}}) = \frac{d^j}{dh_j \dots dh_1} g(\widehat{\mathbb{X}}_t \otimes \exp(h_j \epsilon_{i_j}) \otimes \cdots \otimes \exp(h_1 \epsilon_{i_1}))|_{h_j = \cdots = h_1 = 0}.$$

Since by Propositions 20.1.4 and 20.1.5 (see in particular equation 20.19), in Bonfiglioli et al. (2007), for all $j = 1, \dots, K$,

$$\begin{aligned}
&\frac{d^j}{dh_j \dots dh_1} g(\widehat{\mathbb{X}}_t \otimes \exp(h_j \epsilon_{i_j}) \otimes \dots \otimes \exp(h_1 \epsilon_{i_1}))|_{h_j = \cdots = h_1 = 0} \qquad (5.A.5)\\
&\qquad \in \operatorname{span}\{D^{\beta}_{\boldsymbol{\xi}} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} \ : \beta \in \mathbb{N}_0^M,\ |\beta|_{\mathfrak{g}^K = j}\}.
\end{aligned}$$

The claim follows.

Step 2: It follows directly by the assumption on g^K and a simple adaptation of the proof of Theorem 1.6.2 in Narasimhan (1985) to the present setting.

Step 3: Let p be a polynomial on $[0,1]\times\mathcal{U}(0)$, $\mathbf{Y}\in\widehat{C}^p([0,1],G^{[p]}(\mathbb{R}^{d+2}))$ be the time-extended path of $\widehat{\mathbf{X}}$, and denote the auxiliary time component by the index -1 and its signature by $\mathbb{Y}$. Consider the following set of maps

$$\mathcal{W}:=span\left\{t\mapsto(\mathcal{U}(0)\ni\boldsymbol{\xi}\mapsto w(t)(\boldsymbol{\xi}):=\langle e_I,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}),0,\ldots,0)\rangle),|I|\geqslant 0\right\}.$$

We verify the hypothesis of the Stone -Weiestrass theorem for vector-valued maps to show the convergence in equation (5.A.3). We must prove that $\mathcal{W}$ satisfies the following conditions:

(i) it is a $\mathcal{A}$-submodule of $C([0,1],C^K(\mathcal{U}(0)))$, for some point separating subalgebra $\mathcal{A}\subseteq C([0,1])$ that vanishes nowhere;

(ii) $\mathcal{W}(t):=\{w(t)\ :\ w\in\mathcal{W}\}$ is dense in $C^K(\mathcal{U}(0))$.

(i): Notice that $\mathcal{W}\subseteq C([0,1],C^K(\mathcal{U}(0))$, as the functions in $\mathcal{W}$ are simply linear combinations of polynomials in $\boldsymbol{\xi}$ whose coefficients are continuous in t. Let $\mathcal{A}$ denote the space of polynomials on $[0,1]$. Then, $\mathcal{W}$ is an $\mathcal{A}$-submodule since the map $[0,1]\ni t\mapsto p(t)w(t)\in\mathcal{W}$ for every $w\in\mathcal{W}$ and p polynomial on $[0,1]$.

(ii): We show that the set $\mathcal{W}(t):=\{w(t)\ :\ w\in\mathcal{W}\}$, which explicitly reads as

$$\mathcal{W}(t)=\operatorname{span}\{\mathcal{U}(0)\ni\boldsymbol{\xi}\mapsto\langle\epsilon_I,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}),0,\ldots,0)\rangle\colon |I|\geqslant 0\},\tag{5.A.6}$$

is dense in $C^K(\mathcal{U}(0))$. To this end, we apply the Nachbin theorem (see Nachbin (1949)). We need to verify that for all $t\in[0,1]$, the set $\mathcal{W}(t)$ satisfies the following properties:

(i) it is a linear subspace of $C^K(\mathcal{U}(0))$;

(ii) it is a sub-algebra that contains a non-zero constant function and separates points;

(iii) for all $\boldsymbol{\xi}\in\mathcal{U}(0)$ and $\mathbf{y}\in\mathfrak{g}^K(\mathbb{R}^{d+1})$ with $\mathbf{y}\neq 0$, there exists $f\in\mathcal{W}(t)$ such that $\nabla_{\boldsymbol{\xi}}f(\boldsymbol{\xi})\mathbf{y}\neq 0$.

(i): It is clear.

(ii): The shuffle properties of group-like elements yields that $\mathcal{W}(t)$ is a sub-algebra (see Section 1.1.3). Moreover, it contains a map constantly equal to 1:,

$$\mathcal{U}(0)\ni\boldsymbol{\xi}\mapsto\langle\epsilon_\varnothing,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}),0,\ldots,0)\rangle,$$

and separates point. Indeed, let $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\in\mathcal{U}(0)$ with $\boldsymbol{\xi}_1\neq\boldsymbol{\xi}_2$. If for some $i=0,\ldots,d$, $\langle\epsilon_i,\boldsymbol{\xi}_1\rangle\neq\langle\epsilon_i,\boldsymbol{\xi}_2\rangle$, then

$$\langle\epsilon_i,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}_1),0,\ldots,0)\rangle\neq\langle\epsilon_i,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}_2),0,\ldots,0)\rangle.$$

Otherwise, let $J:=(i_1,\ldots,i_j)\in\{0,\ldots,d\}^j$ $j=2,\ldots,K$ be such that $\langle\epsilon_J,\boldsymbol{\xi}_1\rangle\neq\langle\epsilon_J,\boldsymbol{\xi}_2\rangle$ and $\langle\epsilon_I,\boldsymbol{\xi}_1\rangle=\langle\epsilon_I,\boldsymbol{\xi}_1\rangle$ for all $I\in\{0,\ldots,d\}^i$ with $i<j$. Then,

$$\langle\epsilon_J,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}_1),0,\ldots,0)\rangle\neq\langle\epsilon_J,\mathbb{Y}_t\otimes\exp(i(\boldsymbol{\xi}_2),0,\ldots,0)\rangle.$$

(iii): Let $\boldsymbol{\xi} \in \mathcal{U}(0)$ and $\mathbf{y} \in \mathfrak{g}^K(\mathbb{R}^{d+1})$ with $\mathbf{y} \neq 0$. If for some $i = 0, \dots, d$, $\langle \epsilon_i, \mathbf{y} \rangle \neq 0$, then

$$\nabla_{\boldsymbol{\xi}} \langle \epsilon_i, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \dots, 0) \rangle \mathbf{y} = \langle \epsilon_i, \mathbf{y} \rangle \neq 0.$$

If instead for all $i = 0, \dots, d$, $\langle \epsilon_i, \mathbf{y} \rangle \neq 0$, let $J := (i_1, \dots, i_j) \in \{0, \dots, d\}^j$ $j = 2, \dots, K$ be such that $\langle \epsilon_J, \mathbf{y} \rangle \neq 0$ and $\langle \epsilon_I, \mathbf{y} \rangle = 0$ for all $I \in \{0, \dots, d\}^i$ with $i < j$. Then,

$$\nabla_{\boldsymbol{\xi}} \langle \epsilon_J, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \dots, 0) \rangle \mathbf{y} = \langle \epsilon_J, \mathbf{y} \rangle \neq 0.$$

The claim follows.

Step 4: A direct consequence of Step 2 and Step 3.

Step 5: Let $(\mathbf{v}_n)_{n \in \mathbb{N}} \subset T(\mathbb{R}^{d+2})$ be the sequence whose indices are in $\{-1, 0, \dots, d\}$ and for which the convergence in (5.A.4) holds. The assertion follows if we prove that

(i) we can extract a subsequence, that without loss of generality, we still denote by $(\mathbf{v}_n)_{n \in \mathbb{N}}$, whose last K indices of each of its terms differ from -1.

Recall that here -1 denotes the index corresponding to the auxiliary time component of $\mathbf{Y}$. Indeed, in such a case, one can find a sequence $(\mathbf{u}_n)_{n \in \mathbb{N}} \subset T(\mathbb{R}^{d+1})$ whose indices are in $\{0, \dots, d\}$ such that for all $n \in \mathbb{N}$, $t \in [0,1]$, $\beta \in \mathbb{N}_0^M$, $|\beta|_{\mathfrak{g}^K} \leqslant K$,

$$\begin{aligned}
&\langle \mathbf{v}_n, \mathbb{Y}_t \rangle = \langle \mathbf{u}_n, \widehat{\mathbb{X}}_t \rangle, \\
&D_{\boldsymbol{\xi}}^{\beta} \langle \mathbf{v}_n, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \dots, 0) \rangle_{|\boldsymbol{\xi}=0} = D_{\boldsymbol{\xi}}^{\beta} \langle \mathbf{u}_n, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0) \rangle_{|\boldsymbol{\xi}=0},
\end{aligned}$$

concluding the proof.

In order to prove (i), assume first that for all $j = 1, \dots, K$, there exist $i_j \in \{0, \dots, d\}$ and $t_j \in [0,1]$ such that

$$\langle \epsilon_{i_j}^{\otimes j}, \nabla^j F(t_j, \widehat{\mathbf{X}}) \rangle \neq 0. \tag{5.A.7}$$

Under this assumption, condition (i) is verified. Indeed, otherwise, for some $j = 1, \dots, K$, and some $N \in \mathbb{N}$, the last j-th component of each $\mathbf{v}_n$ with $n \geqslant N$ would be equal to -1. This would imply that for all $n \geqslant N$ and $t \in [0,1]$,

$$D_{\boldsymbol{\xi}}^{\beta} \langle \mathbf{v}_n, \mathbb{Y}_t \otimes \exp(i(\boldsymbol{\xi}), 0, \dots, 0) \rangle_{|\boldsymbol{\xi}=0} = 0,$$

for all $\beta \in \mathbb{N}_0^M$ with $|\beta|_{\mathfrak{g}^K} = j$. In particular, by Step 4, for all $t \in [0,1]$

$$D_{\boldsymbol{\xi}}^{\beta} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} = 0. \tag{5.A.8}$$

Finally, by Step 1 (5.A.8) contradicts (5.A.7) and the claim follows.

Assume now that condition (5.A.7) does not hold. This means that there exist $1 \leqslant m_1 < \dots < m_l \leqslant K$ such that for all $i \in \{0, \dots, d\}$ and for all $t \in [0,1]$, $\langle \epsilon_i^{\otimes m_k}, \nabla^{m_k} F(t, \widehat{\mathbf{X}}) \rangle = 0$, for all $k = 1, \dots, l$. Let M be the biggest of such m_k and consider the functional $F_0 : [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1})) \to \mathbb{R}$ defined via

$$F_0(t, \mathbf{X}) := F(t, \mathbf{X}) + \langle \epsilon_0^{\otimes M}, \mathbb{X}_t \rangle,$$

for all $(t, \mathbf{X}) \in [0,1] \times D^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$, with $\mathbb{X}_t$ denoting the signature of $\mathbf{X}$ at time t. Notice that for all $t \in (0,1]$, and for all m_k, $\langle \epsilon_0^{\otimes m_k}, \nabla^{m_k} F_0(t, \widehat{\mathbf{X}}) \rangle = \langle \epsilon_0^{\otimes M - m_k}, \widehat{\mathbb{X}}_t \rangle \neq 0$. Therefore,

a) if $M = 1$, condition (5.A.7) is verified for the functional F_0. An application of the previous steps to the functional F_0 yields that there exists a sequence $(\mathbf{u}_n^0)_{n \in \mathbb{N}} \subset T(\mathbb{R}^{d+1})$ whose indices are in $\{0, \dots, d\}$ such that

$$\lim_{n \to \infty} \sup_{t \in [0,1]} \sum_{\beta \in \mathbb{N}_0^M,\ |\beta|_{\mathfrak{g}^K} \leqslant K} \| D_{\boldsymbol{\xi}}^{\beta} g^K(t, \boldsymbol{\xi})|_{\boldsymbol{\xi}=0} + D_{\boldsymbol{\xi}}^{\beta} \langle \epsilon_0^{\otimes M}, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0) \rangle|_{\boldsymbol{\xi}=0} - D_{\boldsymbol{\xi}}^{\beta} \langle \mathbf{u}_n^0, \widehat{\mathbb{X}}_t \otimes \exp(\boldsymbol{\xi}, 0, \dots, 0) \rangle|_{\boldsymbol{\xi}=0} \| = 0.$$

The claim follows by considering the sequence $\mathbf{u}_n := \mathbf{u}_n^0 - \epsilon_0^{\otimes M}$, $n \in \mathbb{N}$;

b) if $M > 1$ and for all $1 \leqslant j < M$, i_j for which F satisfies condition (5.A.7) at $\widehat{\mathbf{X}}$ is different from 0, then condition (5.A.7) is verified for the functional F_0 and the claim follows as in a).

c) if $M > 1$ and for some $1 \leqslant j < M$, i_j for which F satisfies condition (5.A.7) is equal to 0, denote such js by $1 \leqslant j_1 < \dots < j_l < M$, $1 \leqslant l < M$. If for all $k = 1, \dots, l$

$$\langle \epsilon_0^{\otimes j_k}, \nabla^{j_k} F(t, \widehat{\mathbf{X}}) \rangle + \langle \epsilon_0^{\otimes M - j_k}, \widehat{\mathbb{X}}_t \rangle \neq 0,$$

for some $t \in [0,1]$, the claim follows as in a). Otherwise, let J be the smallest of the j_k for which for all $t \in [0,1]$,

$$\langle \epsilon_0^{\otimes j_k}, \nabla^{j_k} F(t, \widehat{\mathbf{X}}) \rangle + \langle \epsilon_0^{\otimes M - j_k}, \widehat{\mathbb{X}}_t \rangle = 0.$$

Consider the path functional $F_1(t, \mathbf{X}) := F(t, \mathbf{X}) + \alpha_1 \langle \epsilon_0^{\otimes M}, \mathbb{X}_t \rangle$ for $\alpha_1 \in \mathbb{R}$, $\alpha_1 \notin \{1, 0\}$. Observe that for all $j_k \geqslant J$ and for all $t \in (0,1]$, $\langle \epsilon_0^{\otimes j_k}, \nabla^{j_k} F_1(t, \widehat{\mathbf{X}}) \rangle \neq 0$.

Finally, if for some j_k, with $j_k < J$

$$\langle \epsilon_0^{\otimes j_k}, \nabla^{j_k} F(t, \widehat{\mathbf{X}}) \rangle + \alpha_1 \langle \epsilon_0^{\otimes M - j_k}, \widehat{\mathbb{X}}_t \rangle = 0,$$

for all $t \in [0,1]$, let $J_1 < J$ be the smallest of such j_k. Consider the path functional $F_2(t, \mathbf{X}) := F(t, \mathbf{X}) + \alpha_2 \alpha_1 \langle \epsilon_0^{\otimes M}, \mathbb{X}_t \rangle$ for $\alpha_2 \in \mathbb{R}$, $\alpha_2 \notin \{1, 0, \frac{1}{\alpha_1}\}$ and observe that for all $j_k \geqslant J_1$ and for all $t \in (0,1]$, $\langle \epsilon_0^{\otimes j_k}, \nabla^{j_k} F_2(t, \widehat{\mathbf{X}}) \rangle \neq 0$. An iterative application of this reasoning yields that condition (5.A.7) is verified for a functional of the form $\widetilde{F}(t, \mathbf{X}) := F(t, \mathbf{X}) + \alpha \langle \epsilon_0^{\otimes M}, \mathbb{X}_t \rangle$ for some properly choosen $\alpha \in \mathbb{R}$, $\alpha \neq 0$. The claim follows as in a).

Step 6: It is a simple application of Step 1 to the functional $F^{\mathbf{u}}$.

Step 7: Fix $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ and denote by $\mathbf{Z}$ the Marcus transformed path such that $\mathbf{Z} \in \widehat{C}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$. Recall that by definition of tj-extended paths, such Marcus-transformed path always exists (see Remark 5.3.4). An application of the previous steps of the proof to $\mathbf{Z}$ yields that there exists $(\mathbf{u}_n)_{n\in\mathbb{N}} \in T(\mathbb{R}^{d+1})$ such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sum_{j=0}^{K} \|\nabla^j F(t,\mathbf{Z}) - \nabla^j F^{\mathbf{u}_n}(t,\mathbf{Z})\| = 0. \tag{5.A.9}$$

Thus, in particular,

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \sum_{j=0}^{K} \|\nabla^j F(\mu_t,\mathbf{Z}) - \nabla^j F^{\mathbf{u}_n}(\mu_t,\mathbf{Z})\| = 0,$$

for μ_t introduced in Notation 5.2.7. Since we deal here with Marcus canonical path functionals, the claim follows. □

Remark 5.A.1. (i) Let $\widehat{\mathbf{X}} \in \widehat{D}^p([0,1], G^{[p]}(\mathbb{R}^{d+1}))$ and $(\mathbf{u}_n)_{n\in\mathbb{N}} \in T(\mathbb{R}^{d+1})$ be the sequence such that the convergence in equation (5.A.9) holds. Observe that by the construction of the Marcus transformed path, for all $j = 0, \ldots, K$,

$$\sup_{s\in[0,1]} \sup_{n\in\mathbb{N}} \sup_{\theta\in[0,1]} \|\langle \mathbf{u}_n^{(j)}, \widehat{\mathbb{X}}_{s-} \otimes \exp(\theta \log^{([p])}(\Delta\widehat{\mathbf{X}}_s))\rangle\| < \infty.$$

5.B Föllmer and RIE Integration Theories

5.B.1 Föllmer integration

In this section, we briefly revisit the concept of quadratic variation and the associated pathwise integration theory as introduced by Föllmer (1981).

In the following, we let $\mathcal{B}([0,1])$ denote the Borel σ-algebra on $[0,1]$.

Definition 5.B.1. Let $X \in D([0,1],\mathbb{R})$ and $\pi^n_{[0,1]}$, $n \in \mathbb{N}$, a sequence of partitions with vanishing mesh size. We say that X has finite *quadratic variation along* $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ *in the sense of Föllmer* if the sequence of measures $(\nu_n)_{n\in\mathbb{N}}$ on $([0,1],\mathcal{B}([0,1])$ defined by

$$\nu_n := \sum_{s_i^n \in \pi^n_{[0,1]}} \|X_{s_i^n, s_{i+1}^n}\| \delta_{s_i^n}$$

converges weakly to a measure ν such that the map

$$[0,1] \ni t \mapsto [X]^c_t := \nu([0,t]) - \sum_{0<s\leqslant t} \|\Delta X_s\|^2 \tag{5.B.1}$$

is continuous and increasing. We then call the function $[X]$ given by $[X]_t := \nu([0,t])$ the *quadratic variation* of X along $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$. We say that a path $X \in D([0,1],\mathbb{R}^d)$, $d \geqslant 1$, has finite quadratic variation along $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ if for all $i,j \in \{1,\ldots,d\}$, the condition above holds for X^i and $X^i + X^j$. In this case, we set

$$[X^i, X^j] := \frac{1}{2}([X^i + X^j] - [X^i] - [X^j]).$$

Remark 5.B.2. Notice that for all $i,j \in \{1,\ldots,d\}$, $[X^i,X^j] = [X^j,X^i]$ and the map $[0,1] \ni t \mapsto [X,X]_t \in (\mathbb{R}^d)^{\otimes 2}$ is a càdlàg path of finite 1-variation. This implies that for any $Z \in D([0,1],\mathcal{L}((\mathbb{R}^d)^{\otimes 2},\mathbb{R}^m))$, $m \in \mathbb{N}$, the Young integral of Z with respect to $[X,X]$, that we denote by $\int_0^{\cdot} Z_u d[X,X]_u$, is well defined (see Section 1.2.3 and in particular Remark 1.2.7 (ii)). Moreover, if

$$\{s \in (0,1] \ : \ \Delta Z_s \neq 0\} \subseteq \{s \in (0,1] \ : \ \Delta X_s \neq 0\},$$

the same argument as in the proof of Théorème in Föllmer (1981) yields that for all $t \in (0,1]$,

$$\lim_{n\to\infty} \sum_{s_i^n \in \pi^n_{[0,1]}} Z_{s_i^n} X^{\otimes 2}_{s_i^n \wedge t, s_{i+1}^n \wedge t} = \int_0^t Z_{u-} d[X,X]_u, \tag{5.B.2}$$

where on the right-hand rise the integral is interpreted as a Young integral. Notice that in view of (5.B.1),

$$\int_0^t Z_{u-} d[X,X]_u = \int_0^t Z_{u-} d[X,X]^c_u + \sum_{0<u\leqslant t} Z_{u-} \Delta X_u^{\otimes 2}.$$

We refer also to Lemma 5.11 in Friz and Hairer (2014) for a proof of the equality (5.B.2) when X is continuous, and Lemma 2.6 in Hirai (2017) when X is càdlàg and $Z = f(X)$, for some suitably chosen $\mathbb{R}^m$-valued function f.

Theorem 5.B.3 (see Théorème in Föllmer (1981))**.** *Let $X \in D([0,1],\mathbb{R}^d)$ be a path with finite quadratic variation along a sequence of partition $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ with vanishing mesh size and $f \in C^2(\mathbb{R}^d)$. Then, for all $t \in [0,1]$, the limit*

$$\int_0^t \nabla f(X_{u-}) dX_u := \lim_{n\to\infty} \sum_{si \in \pi^n_{[0,1]}} \nabla f(X_{s_i^n}) X_{s_i^n \wedge t, s_{i+1}^n \wedge t} \tag{5.B.3}$$

exists and the resulting integral satisfies a pathwise Itô-formula.

Remark 5.B.4. Let us highlight that the Föllmer integral (5.B.3) is well-defined only for gradient-type functions and not for general functions.

5.B.2 RIE property

Another line of research that lies in the middle between rough and Föllmer integration theory is that which exploits the so-called RIE property of a path together with a given sequence of partition. This theory was first introduced by Perkowski and Prömel (2016) for continuous paths and then generalized to the càdlàg setting by Allan et al. (2023b). Hereafter, we recall this property along with the most important findings regarding the paths satisfying it.

Definition 5.B.5. A sequence $(\pi^n_{[0,t]})_{n\in\mathbb{N}}$ of partitions is called nested if $\pi^n_{[0,t]} \subseteq \pi^{n+1}_{[0,t]}$ for all $n \in \mathbb{N}$.

Property 5.B.6 (RIE)**.** *Let $p \in (2,3)$ and $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ a sequence of nested partitions with vanishing mesh size. For $X \in D([0,1],\mathbb{R}^d)$, define $X^n : [0,1] \to \mathbb{R}^d$ by*

$$X^n_u := X_1 1_{\{u=1\}} + \sum_{s^n_i \in \pi^n_{[0,1]}} X_{s^n_i} 1_{\{s^n_i \leqslant u < s^n_{i+1}\}},$$

for all $u \in [0,1]$, $n \in \mathbb{N}$. Assume that

(i) The sequence $(X^n)_{n\in\mathbb{N}}$ converges uniformly to X as $n \to \infty$.

(ii) The Riemann sums

$$\int_0^t X^n_{u^-} dX_u := \sum_{s^n_i \in \pi^n_{[0,1]}} X_{s^n_i} X_{s^n_i \wedge t, s^n_{i+1} \wedge t} \tag{5.B.4}$$

converges uniformly as $n \to \infty$ to a limit which we denote by $\int_0^t X_{u^-} dX_u$, for all $t \in [0,1]$.

(iii) There exists a control function w such that[1]

$$\sup_{(s,t)\in\Delta_1} \frac{\|X_{s,t}\|^p}{w(s,t)} + \sup_{n\in\mathbb{N}} \sup_{t^n_i < t^n_j \in \pi^n_{[0,1]}} \frac{\| \int_{t^n_i}^{t^n_j} X^n_{u^-} dX_u - X_{t^n_i} X_{t^n_i, t^n_j} \|^p}{w(t^n_i, t^n_j)} \leqslant 1. \tag{5.B.5}$$

Definition 5.B.7. A path $X \in D([0,1],\mathbb{R}^d)$ is said to *satisfy* (RIE) with respect to p and $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ if p, $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ and X together satisfy property (RIE).

Remark 5.B.8. Notice that we denote the Riemann sums in (5.B.4) and its limit by $\int_0^t X^n_{u^-} dX_u$ and $\int_0^t X_{u^-} dX_u$, respectively. This notation is different from the original paper Allan et al. (2023b), where the authors write instead $\int_0^t X^n_u dX_u$ and $\int_0^t X_u dX_u$. We adopt a different notation to be consistent with the other chapters (see Section 1.2).

[1] In (5.B.5) we adopt the convention that $\frac{0}{0} := 0$.

Next, we collect in the form of a proposition the main properties of the paths satisfying (RIE) that we use throughout the paper. Recall the notion of quadratic variation along a sequence of partitions in the sense of Föllmer given in Definition 5.B.1. and the one of the controlled rough path given in Definition 1.2.9.

Proposition 5.B.9. Let $p \in (2,3)$, $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ a sequence of nested partition with vanishing mesh size and $X \in D([0,1],\mathbb{R}^d)$ with $X_0 = 0$ satisfying (RIE) with respect to p and $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$. Then, the following conditions are satisfied.

(i) The path X has quadratic variation $[X,X]$ along $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$ in the sense of Föllmer.

(ii) For $t \in [0,1]$, let $\int_0^t X_{u^-} dX_s$ denote the limit of the Riemann sums $(\int_0^t X^n_{s^-} dX_s)_{n\in\mathbb{N}}$ defined as in Property (RIE) 5.B.6 (ii). For $(s,t) \in \Delta_1$, set $\int_s^t X_{u^-} dX_u := \int_0^t X_{u^-} dX_u - \int_0^s X_{u^-} dX_u$, and define

$$\mathbb{X}^{(2)}_{s,t} := \int_s^t X_{u^-} dX_u - X_s X_{s,t} + \frac{1}{2}[X,X]_{s,t},$$

and

$$\mathbf{X}_t := (1, X_t, \mathbb{X}^{(2)}_{0,t}) \tag{5.B.6}$$

Then, $\mathbf{X} \in D^p([0,1], G^2(\mathbb{R}^d))$.

(iii) For $\mathbf{X}$ as in (5.B.6), let (Y, Y') be a pair of a controlled rough path, i.e., $(Y,Y') \in \mathcal{V}^{q,r}_{\mathbf{X}}$, for some $q \geqslant p$. Assume that

$$\{s \in (0,1] \,:\, \Delta Y_s \neq 0\} \subseteq \bigcup_{n\in\mathbb{N}} \pi^n_{[0,1]},$$
$$\{s \in (0,1] \,:\, \Delta Y'_s \neq 0\} \subseteq \{s \in (0,1] \,:\, \Delta X_s \neq 0\}.$$

Then, for all $t \in [0,1]$, the following limit (along the sequence of partition $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$) exists

$$\int_0^t Y_{s^-} dX_s := \lim_{n\to\infty} \sum_{s_i^n \in \pi^n_{[0,1]}} Y_{s_i^n} X_{s_i^n \wedge t, s_{i+1}^n \wedge t}, \tag{5.B.7}$$

and the rough integral of (Y,Y') with respect to $\mathbf{X}$ as defined in equation (1.2.7) is given by

$$\int_0^t Y_{s^-} d\mathbf{X}_s = \int_0^t Y_{s^-} dX_s + \frac{1}{2}\int_0^t Y'_{s^-} d[X,X]_s, \tag{5.B.8}$$

for all $t \in [0,1]$.

Remark 5.B.10. Notice that in view of equation (5.B.7), the right-hand side of (5.B.8) has to be interpreted as a limit of Riemann sums along the sequence of partitions $(\pi^n_{[0,1]})_{n\in\mathbb{N}}$.

Proof. (i): See Proposition 2.18 in Allan et al. (2023b).

(ii): By Proposition 2.18 in Allan et al. (2023b) (see in particular equation (2.14)) $\mathbf{X} \in D([0,1], G^2(\mathbb{R}^d))$ and by Lemma 2.13 in Allan et al. (2023b) and Remark 5.B.2 it has finite p-variation.

(iii): The existence of the limit in equation (5.B.7) is proved in Theorem 2.15 in Allan et al. (2023b). The equality in equation (5.B.8) follows by the definition of the rough integral (1.2.7) with respect to $\mathbf{X}$, Theorem 2.15 in Allan et al. (2023b) and Remark 5.B.2.

□

Bibliography

E. Abi Jaber and L.-A. Gérard. Signature volatility models: pricing and hedging with fourier. *Available at SSRN*, 2024.

E. Abi Jaber, C. Cuchiero, L. Pelizzari, S. Pulido, and S. Svaluto-Ferro. Polynomial volterra processes. *Preprint, arXiv:2403.14251*, 2024.

E. Akyildirim, M. Gambara, J. Teichmann, and S. Zhou. Applications of Signature Methods to Market Anomaly Detection. *Preprint, arXiv:2201.02441*, 2022.

A. L. Allan. Rough path theory lecture notes. *Rough Path Theory class, ETH Zürich*, Spring 2021.

A. L. Allan, C. Cuchiero, C. Liu, and D. J. Prömel. Model-free portfolio theory: A rough path approach. *Mathematical Finance*, 33(3):709–765, 2023a.

A. L. Allan, C. Liu, and D. J. Prömel. A càdlàg rough path foundation for robust finance. *Finance and Stochastics*, 28:215–257, 2023b.

E. Alos, J. Gatheral, and R. Radoicić. Exponentiation of conditional expectations under stochastic volatility. *Quantitative Finance*, 20(1):13–27, 2020.

A. Ananova and R. Cont. Pathwise integration with respect to paths of finite quadratic variation. *Journal de Mathématiques Pures et Appliquées*, 107(6):737–757, 2017.

D. Applebaum and F. Tang. Stochastic flows of diffeomorphisms on manifolds driven by infinite-dimensional semimartingales with jumps. *Stochastic Processes and their Applications*, 92:219–236, 04 2001.

G. B. Arous. Flots et séries de Taylor stochastiques. *Probability Theory and Related Fields*, 81:29–77, 1989.

I. P. Arribas, C. Salvi, and L. Szpruch. Sig-sdes model for quantitative finance. In *Proceedings of the First ACM International Conference on AI in Finance*, ICAIF '20, New York, NY, USA, 2021. Association for Computing Machinery.

P. Bank, C. Bayer, P. P. Hager, S. Riedel, and T. Nauen. Stochastic control with signatures. *Preprint, arXiv:2406.01585*, 2024.

C. Bayer, P. P. Hager, S. Riedel, and J. Schoenmakers. Optimal stopping with signatures. *The Annals of Applied Probability*, 33(1):238–273, 2023.

K. Beauchard, J. L. Borgne, and F. Marbach. On expansions for nonlinear systems Error estimates and convergence issues. *Comptes Rendus. Mathématique*, 361:97–189, 2023.

F. E. Benth, N. Detering, and P. Krühner. Multilinear processes in Banach space. *Preprint, arXiv*, 2018.

F. E. Benth, N. Detering, and P. Kruhner. Abstract polynomial processes. *Preprint, arXiv:2010.02483*, 2020.

F. E. Benth, N. Detering, and P. Krühner. Independent increment processes: a multilinearity preserving property. *Stochastics*, 93(6):803–832, 2021.

P. Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., Second edition, 1999.

R. P. Boas. *Entire functions: Entire Functions.* Academic press, 2011.

A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie groups and potential theory for their sub-Laplacians.* Springer Monographs in Mathematics, 2007.

H. Brezis. *Function Analysis, Sobolev Spaces and Partial Differential Equations.* Springer New York, NY, 2011.

R. Buckdahn, C. I. G. Keller, J. Ma, and J. Zhang. Fully nonlinear stochastic and rough pdes: Classical and viscosity solutions. *Probability, Uncertainty and Quantitative Risk*, 5:1–59, 2015.

H. Bühler, B. Horvath, T. Lyons, I. P. Arribas, and B. Wood. A data-driven market simulator for small data environments. *ERN: Neural Networks & Related Topics (Topic)*, 2020.

A. Casanova and D. Spanò. Duality and fixation in Ξ-Wright–Fisher processes with frequency-dependent selection. *The Annals of Applied Probability*, 28(1):250–284, 2018.

T. Cass and T. Lyons. Evolving communities with individual preferences. *Proceedings of the London Mathematical Society*, 110(1):83–107, 2015.

K. T. Chen. Integration of paths, geometric invariants and a generalized baker- hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

K. T. Chen. Iterated path integrals. *Bulletin of the American Mathematical Society*, 83 (5):831–879, 1977.

I. Chevyrev. Random walks and Lévy processes as rough paths. *Probability Theory and Related Fields*, 170(3–4):891–932, May 2017.

I. Chevyrev. Rough path theory. *Preprint, arXiv:2402.10331*, 2024.

I. Chevyrev and P. K. Friz. Canonical RDEs and general semimartingales as rough paths. *The Annals of Probability*, 47(1):420–463, 2019.

I. Chevyrev, P. K. Friz, A. Korepanov, and I. Melbourne. Superdiffusive limits for deterministic fast–slow dynamical systems. *Probability theory and related fields*, 178 (3):735–770, 2020.

R. Cont and D. A. Fournie. Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis*, 259(4):1043–1072, 2010a.

R. Cont and D. A. Fournie. A functional extension of the Itô formula. *Comptes Rendus Mathematique*, 348(1):57–61, 2010b.

R. Cont and D. A. Fournie. Functional Itô calculus and stochastic integral representation of martingales. *The Annals of Probability*, 41(1):109 – 133, 2013.

R. Cont and N. Perkowski. Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity. *Transactions of the American Mathematical Society, Series B*, 2018.

R. Cont, V. Bally, and L. Caramellino. *Stochastic integration by parts and Functional Itô calculus (Lectures Notes of the Barcelona Summer School on Stochastic Analysis, Centro de Recerca de Matematica, July 2012).* Springer (Basel), Jan. 2016.

A. Cosso and F. Russo. Functional itô versus banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 19(04):1650024, 2016.

C. Cuchiero and S. Svaluto-Ferro. Infinite-dimensional polynomial processes. *Finance and Stochastics*, 25:383–426, 2021.

C. Cuchiero and J. Teichmann. Path properties and regularity of affine processes on general state spaces. *Séminaire de probabilités XLV*, pages 201–244, 2013.

C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740, 2012.

C. Cuchiero, M. Larsson, and S. Svaluto-Ferro. Polynomial jump-diffusions on the unit simplex. *The Annals of Applied Probability*, 28(4):2451–2500, 2018.

C. Cuchiero, L. Gonon, L. Grigoryeva, J.-P. Ortega, and J. Teichmann. Discrete-time signatures and randomness in reservoir computing. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11):6321–6330, 2022.

C. Cuchiero, G. Gazzani, and S. Svaluto-Ferro. Signature-based models: Theory and calibration. *SIAM Journal on Financial Mathematics*, 14(3):910–957, 2023a.

C. Cuchiero, P. Schmocker, and J. Teichmann. Global universal approximation of functional input maps on weighted spaces. *Preprint, arXiv:2306.03303*, 2023b.

C. Cuchiero, S. Svaluto-Ferro, and J. Teichmann. Signature SDEs from an affine and polynomial perspective. *Preprint, arXiv:2302.01362*, 2023c.

C. Cuchiero, L. Di Persio, F. Guida, and S. Svaluto-Ferro. Measure-valued affine and polynomial diffusions. *Stochastic Processes and their Applications*, page 104392, 2024a.

C. Cuchiero, G. Gazzani, J. Möller, and S. Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based models. *Mathematical Finance*, pages 1–53, 2024b.

C. Cuchiero, X. Guo, and F. Primavera. Funtional Itô formula and Taylor expansion of non-anticipative maps of rough paths. *In preparation*, 2024c.

C. Cuchiero, F. Primavera, and S. Svaluto-Ferro. Universal approximation theorems for continuous functions of càdlàg paths and Lévy-type signature models. *Forthcoming in Finance and Stochastics*, 2024d.

C. Cuchiero, F. Primavera, and S. Svaluto-Ferro. Holomorphic jump-diffusions. *In preparation*, 2024e.

C. Cuchiero, F. Primavera, and S. Svaluto-Ferro. Sig-jump-diffusions from a polynomial perspective. *In preparation*, 2024f.

R. Dudley, R. Norvaivsa, A. universitet. Mathematical Physics, and Stochastics. *An Introduction to P-variation and Young Integrals: With Emphasis on Sample Functions of Stochastic Processes.* Lecture notes. MaPhySto, Department of Mathematical Sciences, University of Aarhus, 1998.

D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

B. Dupire. Functional Itô calculus. *Quantitative Finance*, 19(5):721–729, 2019.

B. Dupire and V. Tissot-Daguette. Functional expansions. *Preprint, arXiv:2212.13628*, 2023.

S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., 1986.

D. Filipović and M. Larsson. Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972, 2016.

D. Filipović and M. Larsson. Polynomial jump-diffusion models. *Stochastic Systems*, 10 (1):71–97, 2020.

M. Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bulletin de la société mathématique de France*, 109:3–40, 1981.

M. Fliess. On the concept of derivatives and taylor expansions for nonlinear input-output systems. In *The 22nd IEEE Conference on Decision and Control*, pages 643–646, 1983.

M. Fliess. Vers une notion de dérivation fonctionnelle causale. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 3(1):67–76, 1986.

P. Friz and J. Gatheral. Diamonds and forward variance models. *Preprint, arXiv:2205.03741*, 2022.

P. Friz and M. Hairer. *A Course on Rough Paths.* Springer Cham, 2014.

P. Friz and A. Shekhar. General rough integration, Lévy rough paths and a Lévy–Kintchine-type formula. *The Annals of Probability*, 45(4):2707–2765, 2017.

P. Friz and N. B. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications.* Cambridge University Press, 2010.

P. Friz and H. Zhang. Differential equations driven by rough paths with jumps. *Journal of Differential Equations*, 264(10):6226–6301, 2017.

P. K. Friz, J. Gatheral, and R. Radoicić. Forests, cumulants, martingales. *The Annals of Probability*, 50(4):1418–1445, 2022a.

P. K. Friz, P. P. Hager, and N. Tapia. Unified signature cumulants and generalized Magnus expansions. *Forum of Mathematics, Sigma*, 10:e42, 2022b.

T. Fujiwara and H. Kunita. Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. *Journal of Mathematics of Kyoto University*, 25: 71–106, 1985.

M. Fukasawa and K. Matsushita. Realized cumulants for martingales. *Electronic Communications in Probability*, 26(none):1–10, 2021.

H. Föllmer. Calcul d'Itô sans probabilites. *Lecture Notes in Mathematics -Springer-Verlag*, 850:143–150, 01 1981.

H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 205–223. North-Holland, Amsterdam, 1986.

X. Guo, H. Pham, and X. Wei. Itô's formula for flows of measures on semimartingales. *Stochastic Processes and their Applications*, 159:350–390, 2023.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the econometric society*, pages 1029–1054, 1982.

Y. Hirai. Remarks on Föllmer's pathwise Itô calculus. *Osaka Journal of Mathematics*, 56:631–660, 2017.

R. Holley and D. Stroock. Dual processes and their application to infinite interacting systems. *Advances in mathematics*, 32(2):149–174, 1979.

C. Houdré and J. Víquez. An optimal functional itô's formula for lévy processes. *Preprint, arXiv:2406.00601*, 2024.

S. L. Jaber, Eduardo Abi and X. Lin. Fourier-laplace transforms in polynomial ornstein-uhlenbeck volatility models. *Preprint arXiv:2405.02170*, 2024.

S. Jacka and R. Tribe. Comparisons for measure valued processes with interactions. *The Annals of Probability*, 31(3):1679–1712, 2003.

J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer Berlin, Heidelberg, Second edition, 1987.

S. Jansen and N. Kurt. On the notion(s) of duality for Markov processes. *Probability Surveys*, 11:59 – 120, 2014.

S. Jazaerli and Y. F. Saporito. Functional itô calculus, path-dependence and the computation of greeks. *Stochastic Processes and their Applications*, 127(12):3997–4028, 2017.

Jochen Blath and Adrián González Casanova and Noemi Kurt and Maite Wilke-Berenguer. A new coalescent for seed-bank models. *The Annals of Applied Probability*, 26(2):857 – 891, 2016.

J. Kallsen and J. Muhle-Karbe. Method of moment estimation in time-changed Lévy models. *Statistics & Decisions*, 28(2):169–194, 2011.

J. Kallsen and A. Pauwels. Variance-optimal hedging in general affine stochastic volatility models. *Advances in Applied Probability*, 42(1):83–105, 2010.

J. Kalsi, T. Lyons, and I. Perez-Arribas. Optimal execution with rough path signatures. *SIAM Journal on Financial Mathematics*, 11(2):470–493, 2020.

S. Karlin and J. McGregor. The classification of birth and death processes. *Transactions of the American Mathematical Society*, 86(2):366–400, 1957.

C. Keller and J. Zhang. Pathwise itô calculus for rough paths and rough pdes with path dependent coefficients. *Stochastic Processes and their Applications*, 126(3):735–766, 2016. ISSN 0304-4149.

F. J. Kiraly and H. Oberhauser. Kernels for sequentially ordered data. *Journal of Machine Learning Research*, 20(31):1–45, 2019.

P. E. Klöden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer, 1992.

P. Krühner and M. Larsson. Affine processes with compact state space. *Electronic Journal of Probability*, pages 1–23, 2018.

H. Kunita. Representation of martingales with jumps and applications to mathematical finance. In H. Kunita, S. Watanabe, and Y. Takahashi, editors, *Stochastic analysis and related topics in Kyoto. In honour of Kiyoshi Itô.*, pages 209–232. Tokyo: Mathematical Society of Japan, 2004.

T. G. Kurtz, E. Pardoux, and P. Protter. Stratonovich stochastic differential equations driven by general semimartingales. *Annales de l'I.H.P. Probabilités et statistiques*, 31 (2):351–377, 1995.

M. Larsson and S. Svaluto-Ferro. Existence of probability measure valued jump-diffusions in generalized Wasserstein spaces. *Electronic Journal of Probability*, 25(none):1 – 25, 2020.

B. Levin, Y. Lyubarskii, M. Sodin, and V. Tkachenko. *Lectures on entire functions*, volume 150 of *Translations of Mathematical Monographs*. American Mathematical Society, 1996.

D. Levin, T. Lyons, and H. Ni. Learning from the past, predicting the statistics for the future, learning an evolving system. *Preprint, arXiv:1309.0260*, 2013.

S. Liao, H. Ni, M. Sabaté-Vidales, L. Szpruch, M. Wiese, and B. Xiao. Sig-wasserstein gans for conditional time series generation. *Mathematical Finance*, 34(2):622–670, 2024.

C. Litterer and H. Oberhauser. On a Chen–Fliess approximation for diffusion functionals. *Monatshefte für Mathematik*, 175:577–593, 2011.

T. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

T. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*. Springer, 2007.

T. Lyons, S. Nejad, and I. Perez Arribas. Non-parametric Pricing and Hedging of Exotic Derivatives. *Applied Mathematical Finance*, 27(6):457–494, 2020.

W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on pure and applied mathematics*, 7(4):649–673, 1954.

S. Marcus. Modeling and analysis of stochastic differential equations driven by point processes. *IEEE Transactions on Information Theory*, 24(2):164–172, 1978.

S. Marcus. Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics*, 4(3):223–245, 1981.

J. Miller and V. Silvestri. Stochastic calculus lecture notes. *University of Cambridge Part III course Stochastic Calculus, given Lent 2017*, 2017.

M. Min and R. Hu. Signatured deep fictitious play for mean field games with common noise. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7736–7747. PMLR, 18–24 Jul 2021.

L. Nachbin. Sur les algèbres denses de fonctions diffèrentiables sur une variètè. *Comptes Rendus hebdomadaires des séances de l'Académie des sciences Paris*, 228:1549–1551, 1949.

R. Narasimhan. *Analysis on Real and Complex Manifolds*. North-Holland mathematical library. North-Holland, 1985.

H. Oberhauser. The functional itō formula under the family of continuous semimartingale measures. *Stochastics and Dynamics*, 16(04):1650010, 2016.

N. Perkowski and D. Prömel. Pathwise stochastic integrals for model free finance. *Bernoulli*, 22:2486–2520, 11 2016.

H. Pham. On quadratic hedging in continuous time. *Mathematical Methods of Operations Research*, 51:315–339, 2000.

P. Protter and K. Shimbo. No arbitrage and general semimartingales. *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, 4:267–283, 2006.

Z. Qian and J. Tudor. Differential structure and flow equations on rough path space. *Bulletin des Sciences Mathématiques*, 135(6):695–732, 2011. Special issue in memory of Paul Malliavin.

C. Riga. Pathwise functional calculus and applications to continuous-time finance (phd thesis). *Preprint, arXiv:1602.04523*, 2016.

W. Rudin. *Real and Complex Analysis*. McGraw-Hill Science/Engineering/Math, 1986.

K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2013.

V. Scheidemann. *Introduction to Complex Analysis in Several Variables*. Birkhäuser Basel, 2005.

J. L. Schiff. *Normal Families*. Springer, 1993.

A. Schmeding. *An Introduction to Infinite-Dimensional Differential Geometry*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2022.

M. Schweizer. Risk-minimality and orthogonality of martingales. *Stochastics and Stochastic Reports*, 30(2):123–131, 1990.

M. Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2):339–363, 1991.

M. Schweizer. On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, 13(5):573–599, 1995.

M. Schweizer. A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanic, and M. Musiela, editors, *Handbooks in Mathematical Finance: Option Pricing, Interest Rates and Risk Management*, pages 538–574. Cambridge University Press, 2001.

B. V. Shabat. *Introduction to complex analysis: functions of several variables*, volume 110. American Mathematical Soc., 1992.

E. M. Stein and R. Shakarchi. *Complex analysis*, volume 2. Princeton University Press, 2010.

H. Sussmann. Product expansions of exponential lie series and the discretization of stochastic differential equations. In *Stochastic Differential Systems, Stochastic Control Theory and Applications*, pages 563–582. Springer, 1988.

F. Viens and J. Zhang. A martingale approach for fractional brownian motions and related path dependent pdes. *The Annals of Applied Probability*, 29(6):pp. 3489–3540, 2019.

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

H. Zhou. Itô conditional moment generator and the estimation of short-rate processes. *Journal of Financial Econometrics*, 1(2):250–271, 2003.