On pathwise functional Itô calculus and its applications to mathematical finance

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von
Iryna Voloshchenko
aus Kiev, Ukraine

Mannheim, 2016

Dekan:	Professor Dr. Heinz Jürgen Müller, Universität Mannheim
Referent:	Professor Dr. Alexander Schied, Universität Mannheim
Korreferent:	Professor Dr. Rainer Dahlhaus, Universität Heidelberg

Tag der mündlichen Prüfung: 6. September 2016

Abstract

In recent years, pathwise Itô calculus has been particularly popular in mathematical finance and economics. This is due to the fact that the results derived with the help of the pathwise Itô calculus are robust with respect to model risk that might stem from a misspecification of probabilistic dynamics. In this sense, there is also a close link to robust statistics. The only assumption on the underlying paths is that they admit the quadratic variation in the sense of Föllmer.

In this thesis, we will be particularly interested in the *functional* extension of Föllmer's pathwise calculus, since it is natural to assume that randomness impacts the current situation not simply by influencing the current state of the process but through its entire past evolution. In a first part, we derive the associativity property of the pathwise Itô integral in a functional setting for continuous integrators. This allows us to establish existence and uniqueness results for a class of linear functional Itô differential equations. With this at hand, we turn to financial applications.

First, we use functional pathwise Itô calculus to prove a strictly pathwise version of the master formula in Fernholz' stochastic portfolio theory. This adds a new case study in which continuous-time trading strategies can be constructed in a probability-free manner by means of pathwise Itô calculus. Moreover, the portfolio-generating function may depend on the entire history of the asset trajectories and on an additional continuous trajectory of bounded variation. Our results are illustrated by several examples and shown to work on empirical market data.

Second, we consider a strictly pathwise setting for Delta hedging exotic options, based on Föllmer's pathwise Itô calculus. Price trajectories are d-dimensional continuous functions whose pathwise quadratic variations and covariations are determined by a given local volatility matrix. The existence of Delta hedging strategies in this pathwise setting is established via existence results for recursive schemes of parabolic Cauchy problems and via the existence of functional Cauchy problems on path space. Our main results establish the nonexistence of pathwise arbitrage opportunities in classes of strategies containing these Delta hedging strategies and under relatively mild conditions on the local volatility matrix.

Zusammenfassung

In den letzten Jahren hat das pfadweise Itô Kalkül sowohl in der Finanzmathematik als auch in der Wirtschaft zunehmend mehr an Bedeutung gewonnen. Dies ist darauf zurückzuführen, dass es “robuste” Resultate in dem Sinne liefert, dass sie das Modellrisiko, welches durch eine fehlerhafte probabilistische Modellierung der zugrundeliegenden Wertentwicklung entstehen könnte, eingrenzen.

In dieser Arbeit liegt der Schwerpunkt auf der *funktionalen* Erweiterung von Föllmer's pfadweisem Kalkül, welche dadurch motiviert wird, dass in vielen praktisch relevanten Situationen das Ergebnis von der ganzen vergangenen Entwicklung und nicht nur von dem Wert im betrachteten Zeitpunkt abhängt. Im ersten Schritt leiten wir die Assoziativität des pfadweisen funktionalen Itô Integrals für stetige Integratoren her, was uns dann erlaubt, Existenz- und Eindeutigkeitsresultate für eine Klasse von linearen funktionalen Itô Differentialgleichungen zu zeigen. Damit wenden wir uns den folgenden zwei finanzmathematischen Fragestellungen zu.

Erstens benutzen wir die Assoziativitätseigenschaft, um eine pfadweise Version der Masterformel aus Fernholz' stochastischer Portfoliotheorie zu zeigen. Dies kann als eine zusätzliche Fallstudie betrachtet werden, in welcher zeitstetige Handelsstrategien ganz ohne probabilistische Annahmen konstruiert werden. Zudem dürfen die portfoliogenerierenden Funktionen von der ganzen Vergangenheit der Preisentwicklung der Vermögenswerte und von einer zusätzlichen Komponente von beschränkter Variation abhängen. Unsere Resultate werden anhand von empirischen Marktdaten durch Beispiele untermauert.

Zweitens betrachten wir einen pfadweisen Modellrahmen für das Delta–Hedging exotischer Optionen. Die Preistrajektorien werden dabei durch d-dimensionale stetige Funktionen modelliert, deren Kovariationsstruktur durch eine vorgegebene lokale Volatilitätsmatrix bestimmt wird. Die Existenz der Delta Hedging–Strategien wird durch Existenzresultate für rekursiv gegebene parabolische Cauchy–Probleme und funktionale Cauchy–Probleme auf dem Pfadraum gewährleistet. Wir zeigen die Nichtexistenz von pfadweisen Arbitragegelegenheiten in einer Klasse von Strategien, die solche Delta Hedging-Strategien beinhalten, unter relativ milden Annahmen an die lokale Volatilitätsmatrix.

Contents

Chapter 1

Introduction

1.1 Summary of results

In many situations uncertainty influences the outcome by affecting not only the current state of the world but the entire past. For instance, the quality of a harvest may depend on the current temperature *and* on the pattern of past temperatures. In finance, the price or hedging strategy of a path-dependent option may in general depend on the entire past evolution of the underlying price trajectory; see, e.g., [34] for more examples. Motivated by these arguments, Dupire [34] and Cont and Fournié [20, 21] have recently introduced a new type of stochastic calculus, known as *functional Itô calculus*. It essentially relies on an extension of the classical Itô formula to functionals depending on all of the past values of the underlying path, and not only on the current value. The approach taken in [20] is a direct extension of the non-probabilistic Itô formula of Föllmer [46] to non-anticipative functionals on Skorohod space. These functionals must possess certain directional derivatives that may be computed pathwise, but without requiring Fréchet differentiability. An alternative approach, which to some extent still relies on probabilistic arguments, was introduced by Cosso and Russo [22]; it is based on the theory of stochastic calculus via regularization [26, 27, 28, 29, 30, 78].

In Chapter 2, we first describe an elementary setting for continuous-time modeling of asset prices and discuss what assumptions the price trajectories have to fulfill so as to allow for a reasonable integration theory for self-financing strategies. Proposition 2.1.11, extending and elaborating an argument by Föllmer [47], shows that what we must assume is that the price paths admit a continuous quadratic variation structure in the sense of [46] so as to allow for pathwise (functional) Itô calculus. In a second step, we compare the approach to functional Itô calculus from [20] with the alternative approach from [22]. However, functionals of interest often depend on additional arguments such as quadratic variation, moving average, or running maximum of the underlying path, which are not sufficiently regular for the framework of [20] (see also the discussions in [21] and [34]), respectively [22].

In Chapter 3, which is based on [82], we will therefore extend the functional change of variables formula from [20] to functionals F depending on an additional variable A that corresponds to a general path of bounded variation. This requires us to extend the notions of the horizontal and vertical derivatives to functionals of this type. This extension then allows us to derive the following associativity rule: Assume that ξ is suitably integrable with respect to a continuous path X that satisfies the assumptions of Föllmer's pathwise Itô calculus and let η be suitably integrable for $Y(t) := \int_0^t \xi(s)\,\mathrm{d}X(s)$. Then, the main result in Chapter 3, corresponding to [82, Theorem 3.1], states that $\eta\xi$ is again suitably integrable with respect to X, and we have the intuitive cancellation property

$$\int_0^T \eta(s)\,\mathrm{d}Y(s) = \int_0^T \eta(s)\xi(s)\,\mathrm{d}X(s).$$

Although in standard stochastic calculus the associativity of the stochastic integral follows immediately from an application of the Kunita–Watanabe characterization, in our present strictly pathwise setting this characterization is not available. Thus, the fact that we can only use analytical techniques renders the proof of this associativity property considerably more involved.

Nevertheless, just as in standard stochastic calculus, associativity is a fundamental property of the Itô integral and of crucial importance for a number of applications. For instance, in [80], a basic version of the associativity rule was derived, which allows for a pathwise treatment of constant-proportion portfolio insurance strategies (CPPI), and shows that it is possible to translate the Doss–Sussmann method to the pathwise Itô calculus (see [67, Section 2.3]). In Section 3.4, we will use the associativity rule in pathwise functional Itô calculus in order to prove existence and uniqueness results for pathwise linear Itô differential equations whose coefficients are non-anticipative functionals.

Originally, our desire to derive an associativity rule within functional pathwise Itô calculus was motivated by the fact that it is needed for analyzing functionally dependent strategies in a pathwise version of Stochastic portfolio theory (SPT). In our Chapter 4, which is based on [84], we will see how this application can be carried out.

Stochastic portfolio theory was introduced by Fernholz [37, 38, 40]; see also Karatzas and Fernholz [45] for an overview. On one hand, this framework provides theoretical methods that can be used to analyze portfolio behavior and the structure of equity markets. On the other hand, SPT has been successful in many practical applications concerning portfolio analysis and optimization. SPT aims at constructing investment strategies that outperform a certain reference portfolio such as the market portfolio, $\mu(t)$, as shown in, e.g., [44]. The focus in standard SPT is mainly on functionally generated portfolios that are constructed from functions $G(t, \mu(t))$ that depend on the current state of the market portfolio, $\mu(t)$. The performance of such functionally generated portfolios, with respect to the market portfolio, is described in a very convenient way by the so-called *master formula* of SPT, which, under certain conditions, may allow for (relative) arbitrage opportunities. For instance, in [90] a variant of the strategy from [44] is studied, namely, a diversity-weighted portfolio with negative parameter $p < 0$; in [59], the generating functions are interpreted as Lyapunov functions, i.e., via the property that $G(t, \mu(t))$ is a supermartingale under an appropriate change of measure. In [87], the classical master formula is extended to the case where the generating function, G, may additionally depend on the current state of a continuous trajectory A with components of bounded variation.

In practice, the construction of portfolios often involves knowing not just the current market prices or capitalizations but also past data such as econometric estimates, moving or rolling averages, running maxima, realized covariances, Bollinger bands, etc. Therefore, it is natural to ask whether we can establish a master formula for portfolios that are based on functionals depending on the entire past evolution of the market portfolio, $\mu^t := (\mu(s))_{0\le s\le t}$, and possibly also on other factors. Our main result in Chapter 4, Theorem 4.2.5, which corresponds to [84, Theorem 2.9], gives an affirmative answer to this question: It establishes a master formula for portfolios that arise from sufficiently smooth functionals of the form $G(t, \mu^t, A^t_\mu)$, where $A^t_\mu = (A_\mu(s))_{0\le s\le t}$ is an additional m-dimensional continuous trajectory, which may depend on μ in an adaptive manner, and whose components are of bounded variation. With this at hand, we turn to discussing concrete examples of portfolios that are generated by functions of mixtures of current asset prices and their (adjusted) moving averages. Our analysis is carried out both on a mathematical level and on empirical market data from Reuters Datastream.

Moreover, following [84], Chapter 4 deals with the basis for the modeling framework of SPT. Usually, price processes for SPT are modeled as Itô processes, but at the same time it has often been remarked that the (standard) master formula yields a path-by-path representation of the associated relative wealth. Thus the following questions might arise:

- To what extent does the derivation of the results of SPT rely on a stochastic model?
- Must price processes really be modeled as Itô processes driven by Brownian motion or can we relax this condition and consider more general processes, perhaps even beyond the class of semimartingales?
- Is it possible to get rid of the nullsets inherent in stochastic models and prove the master formula in a strictly probability-free way?

Our approach gives affirmative answers to the questions raised above. To this end, we show that the results of SPT can be derived within the strictly pathwise Föllmer's Itô calculus [46] and its functional extension of Dupire [34] and Cont and Fournié [20, 21]. In particular, we will use in Chapter 4 the slightly extended formalism of [82] (see Chapter 3) and heavily rely on the pathwise functional associativity rule, Theorem 3.3.1. Thus, the only assumption on the trajectories of the price evolution is that they are continuous and admit continuous quadratic variations and covariations in the sense of [46]. This assumption is satisfied by all typical sample paths of a continuous semimartingale but also by non-semimartingales, such as fractional Brownian motion with Hurst index $H \ge 1/2$ and many deterministic fractal curves (see [67, 81]).

In the context of model uncertainty, avoiding the choice of a probabilistic model as done in the pathwise framework becomes very useful for practical applications. It is known from previous discussions that, for example, hedging strategies for variance swaps and related derivatives can be constructed in a purely pathwise manner (see [25, 49]), while [80] adds the analogous result on CPPI strategies; see also [11, 18, 23, 79] for similar analyses on other financial problems. In this sense, our results also contribute to *robust finance*, which aims at reducing the reliance on a probabilistic model and, thus, to model uncertainty. Robustness results for discrete-time SPT were previously obtained also by Pal and Wong [68], where the relative performance of portfolios with respect to a certain benchmark is analyzed using the discrete-time energy-entropy framework [69, 99, 100].

In the final Chapter 5, which follows [83], we discuss strictly pathwise hedging of exotic derivatives. A theory of hedging European options of the form $H = h(S(T))$ for one-dimensional asset price trajectories $S = (S(t))_{0\le t\le T}$ was developed by Bick and Willinger [11], using Föllmer's strictly pathwise approach [46] to Itô calculus. In particular, [11] established that if S is strictly positive and admits a pathwise quadratic variation of the form $[S,S](t) = \int_0^t a(s,S(s))\,\mathrm{d}s$, where $a(s,x) > 0$, then a solution v to the terminal-value problem

$$\begin{cases} v \in C^{1,2}([0,T)\times\mathbb{R}_+)\cap C([0,T]\times\mathbb{R}_+), \\ \frac{\partial v}{\partial t} + a\frac{\partial^2 v}{\partial x^2} = 0 \text{ in } [0,T)\times\mathbb{R}_+, \\ v(T,x) = h(x),\ x\in\mathbb{R}_+, \end{cases} \tag{1.1.1}$$

gives rise to a self-financing trading strategy with portfolio value $v(t,S(t))$ that perfectly replicates the payoff $H = h(S(T))$ in a strictly pathwise sense. Thus, the initial value, $v(0,S(0))$, represents the cost that is required to replicate the payoff H, which, in standard continuous-time finance, is usually interpreted as an arbitrage-free price for H. In our strictly pathwise situation, however, we first need to exclude the existence of strictly pathwise arbitrage in order for this latter interpretation to make sense.

In a first step, we take as starting point the approach from [11] and extend their results to the situation with a d-dimensional price trajectory, $S(t) = (S_1(t),\dots,S_d(t))^\top$, and an exotic option whose payoff is given by $H = h(S(t_0),\dots,S(t_N))$, where $t_0 < t_1 < \cdots < t_N$ are the fixing times of daily closing prices and h is a certain function. In practice, most European-style exotic derivatives (i.e., such derivatives that pay off at maturity T) are given in this form. We show that such options can be hedged in a strictly pathwise sense if a certain recursive scheme of terminal-value problems (1.1.1) can be solved, by using ideas from [79]. We also provide sufficient conditions for the existence and uniqueness of the corresponding solutions.

In a second step, we turn to discussing the absence of strictly pathwise arbitrage in a class of strategies that are based on solutions of recursive schemes of terminal-value problems and comprise, in particular, the Delta hedging strategies of exotic derivatives of the form $H = h(S(t_0),\dots,S(t_N))$. Our main result in this chapter, Theorem 5.2.4, which corresponds to [83, Theorem 3.3], establishes the non-existence of admissible arbitrage opportunities in a strictly pathwise sense under the condition that the covariation of the price trajectory is of the following form

$$\mathrm{d}[S_i,S_j] = \begin{cases} a_{ij}(t,S(t))\,\mathrm{d}t & \text{if } S \text{ takes values in all of } \mathbb{R}^d, \\ a_{ij}(t,S(t))S_i(t)S_j(t)\,\mathrm{d}t & \text{if } S \text{ takes values in } \mathbb{R}^d_+, \end{cases} \tag{1.1.2}$$

with a continuous, bounded, and positive definite matrix $a(t,x) = (a_{ij}(t,x))$. Here, admissibility is understood in the usual sense, namely, we require that the portfolio value of a strategy must be bounded from below for all considered price trajectories.

Our result on the absence of arbitrage is related to [2, Theorem 4], where the absence of pathwise arbitrage is established in the one-dimensional case for constant $a > 0$ and a certain class of smooth strategies. However, there are several differences between this and our result. First, we consider a more general class of price trajectories that are based on local instead of constant volatility and allow for an arbitrary number d of traded assets, whose prices may either be strictly positive or of Bachelier type. Second, in our class of trading strategies the natural Delta hedging strategies for path-dependent exotic options are included and, third, our proof uses completely different techniques; while Alvarez et al. [2] transfer the absence of arbitrage from the probabilistic Black–Scholes model to a pathwise context, by applying a continuity argument, our proof does not involve any probabilistic asset pricing model. Instead, our proof relies on Stroock's and Varadhan's idea for a probabilistic proof [89] of Nirenberg's strong parabolic maximum principle.

In Section 5.3, we turn to the case where the option's payoff may depend on the entire past evolution of the asset price trajectory. Hence, Föllmer's pathwise Itô formula needs to be replaced by the corresponding functional extension; we will use the formulation in Theorem 3.2.1. Furthermore, the previously considered Cauchy problem, as given in (1.1.1), and the associated iterated scheme need to be replaced by a functional version of the Cauchy problem on path space; this was, for instance, studied in Peng and Wang [74] and Ji and Yang [58]. Our results on hedging strategies and the absence of pathwise arbitrage can be naturally extended to this functional setting.

There are many other approaches to hedging and arbitrage in the face of model risk. For continuous-time results, we refer, for instance, to Lyons [63], Hobson [54, 55], Vovk [92, 93, 94], Bender et al. [9], Davis et al. [25], Biagini et al. [10], Beiglböck et al. [8], Schied et al. [84], and the references therein. For discrete-time settings, we refer the reader, for instance, to Acciaio et al. [1], Bouchard and Nutz [14], Föllmer and Schied [48, Section 7.4], Riedel [77], and again the references therein.

1.2 Acknowledgements

I am profoundly grateful to my supervisor Prof. Alexander Schied. Not only many ideas, which form the basis for this thesis, were initially proposed by him, but he also always found time to discuss any issues and mathematical difficulties I may have had concerning this thesis. In particular, I am deeply grateful for his precious advice and his extensive support when writing the papers on which this thesis relies; from our collaboration I could learn a lot. I could also learn a lot from various lectures, seminars, workshops, and other activities organized by the Graduate School 1953 Mannheim/Heidelberg *Statistical Modeling of Complex Systems and Processes—Advanced Non-parametric Approaches.* Furthermore, I would like to thank the organizers of the 8th European Summer School in Financial Mathematics for inviting me to Le Mans to give a talk, which was an extremely valuable experience.

A special thank goes also to all my (former) colleagues: Alexander Kalinin, Mourad Lazgham, Elias Strehle, Tao Zhang, as well as my fellow students from GRK 1953 for many interesting discussions and our common activities at the University of Mannheim and at the University of Heidelberg. I would also like to thank Susanne Spether for all those pleasant and very useful conversations we had together and for her support during my stay at the "Lehrstuhl für Wirtschaftsmathematik I". I am likewise indebted to Prof. Li Chen and to Prof. Martin U. Schmidt for helpful comments concerning Chapter 5 of this thesis. I would also like to thank my coauthor Leo Speiser without whose prior work Chapter 4 would not have been possible. I gratefully acknowledge support by Deutsche Forschungsgemeinschaft through the Research Training Group RTG 1953. Last but not least, I would like to thank my family for their understanding and constant support during my dissertation project.

Chapter 2

Continuous time modeling of asset prices

In mainstream finance, the price evolution of a risky asset is usually modeled as a stochastic process defined on some probability space and hence is subject to model uncertainty. In a number of situations, however, it is possible to construct continuous-time strategies on a path-by-path basis and without making any probabilistic assumptions on the asset price evolution. Let us start with considering two assets, in which trading is allowed. Typically we assume that we have a riskless bond and a risky asset. The prices of the riskless bond, $B(t)$, will be described via a continuously compounded interest with short rate $r(t)$. Since usually prices of risky assets, denoted by $S(t)$, can rise or fall in rather an unpredictable manner, as is illustrated by any stock or index chart, we will not impose any particular assumptions on it for the moment.

A key concept in mathematical finance is the notion of *self-financing* strategies. First recall the situation, where trading can take place only at finitely many time points $0 = t_0 < t_1 < \cdots < t_N < T$ and the corresponding strategies ξ and η are constant on each interval $[t_i, t_{i+1})$ and on $[t_N, T]$. Then, $\big(\xi(t_i), \eta(t_i)\big)_{i=0,\ldots,N-1}$ will be self-financing if and only if the terminal value of the portfolio $(\xi(t_i), \eta(t_i))$ of the i-th trading period coincides with the initial value of the portfolio $\big(\xi(t_{i+1}), \eta(t_{i+1})\big)$ of the next trading period. That is,

$$\xi(t_i)S(t_{i+1}) + \eta(t_i)B(t_{i+1}) = \xi(t_{i+1})S(t_{i+1}) + \eta(t_{i+1})B(t_{i+1}).$$

Intuitively this means that the portfolio should always be rearranged so as to preserve its present value. It follows easily that the accumulated gains and losses resulting from asset price fluctuations should then represent the only source of variations in the portfolio value, i.e., the trading strategy $\big(\xi(t_i), \eta(t_i)\big)_{i=0,\ldots,N-1}$ is self-financing if and only if

$$V(t_i) = V(0) + \sum_{k=1}^{i} \xi(t_{k-1})\big(S(t_k) - S(t_{k-1})\big) + \sum_{k=1}^{i} \eta(t_{k-1})\big(B(t_k) - B(t_{k-1})\big). \tag{2.0.1}$$

Let us now consider the case, where prices $S(t)$ and $B(t)$ are available at each time $0 \leq t \leq T$ of the trading period. Then, the portfolio value of a continuous-time trading strategy

$\big(\xi(t), \eta(t)\big)_{0 \le t \le T} = \big(\xi, \eta\big)$ is given by

$$V(t) = \xi(t)S(t) + \eta(t)B(t), \quad 0 \le t \le T. \tag{2.0.2}$$

If we let the partition $\{t_0, \ldots, t_N\}$ of the considered time interval $[0, T]$ become finer and finer, the right-hand side of equation (2.0.1) should naturally converge toward the sum of an *integral of* ξ *with respect to* S and an *integral of* η *with respect to* B, which should be given as the limit of the Riemann sums in (2.0.1). Denoting these integrals by $\int_0^t \xi(s)\,\mathrm{d}S(s)$ and $\int_0^t \eta(s)\,\mathrm{d}B(s)$, respectively, we would obtain the following condition in order for the continuous-time trading strategy $\big(\xi, \eta\big)$ to be self-financing:

$$V(t) = V(0) + \int_0^t \xi(s)\,\mathrm{d}S(s) + \int_0^t \eta(s)\,\mathrm{d}B(s), \quad 0 \le t \le T. \tag{2.0.3}$$

The integral with respect to the bond price, $\int_0^t \eta(s)\,\mathrm{d}B(s)$, is indeed a Riemann-Stieltjes integral, and criteria for its existence are well known. However, the question arises what assumptions have to be imposed on the path $t \to S(t)$ in order to have a "reasonable" integration theory for self-financing strategies.

2.1 Functions of bounded variation

The first solution one would probably think of is to define $\int_0^t \xi(s)\,\mathrm{d}S(s)$ via the limit of the Riemann sums in (2.0.1), which involves the concept of functions of bounded variation. In the following, we recall some results from classical integration theory for functions of bounded variation; see, e.g., [75], or [95].

Definition 2.1.1. Let $[0, T]$ be a subinterval of $[0, \infty)$.

(i) A *partition* of the time interval $[0, T]$ is a finite set $\mathbb{T} = \{t_0, t_1, \ldots, t_n\} \subset [0, T]$ such that $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots t_n = T$.

(ii) We will denote the successor of $t \in \mathbb{T}_n$ by t', i.e.,

$$t' = \begin{cases} \min\{u \in \mathbb{T}_n \,|\, u > t\} & \text{if } t < T, \\ T & \text{if } t = T. \end{cases}$$

(iii) The *mesh* of a partition $\mathbb{T}$ is defined as $|\mathbb{T}| := \sup_{t \in \mathbb{T}} |t' - t|$.

(iv) A sequence of partitions $(\mathbb{T}_n)_{n \in \mathbb{N}}$ is called a *refining sequence of partitions* if each $\mathbb{T}_n$ is a finite partition of the interval $[0, T]$ and satisfies $\mathbb{T}_1 \subset \mathbb{T}_2 \subset \cdots$ as well as $\mathrm{mesh}(\mathbb{T}_n) \to 0$ as $n \uparrow \infty$.

Definition 2.1.2. For a right-continuous function $t \mapsto A(t)$ on the time interval $[0, T] \subset [0, \infty)$, we define the *total variation* of A over $[0, T]$ as

$$V_{[0,T]}(A) := \sup_{\mathbb{T}} \sum_{t \in \mathbb{T}} |A(t') - A(t)|, \tag{2.1.1}$$

where the supremum is taken over the class of all partitions $\mathbb{T}$ of the finite interval $[0, T]$. We say that $t \to A(t)$ is *of bounded variation* if $V_{[0,T]}(A) < \infty$. The class of all such functions on $[0, T]$ will be denoted in the following by $BV([0, T])$. Analogously, the class of all such functions that are moreover continuous will be denoted by $CBV([0, T]) = BV([0, T]) \cap C([0, T])$.

It is a well known fact that if $(\mathbb{T}_n)$ is a refining sequence of partitions, then

$$V_{[0,T]}(A) = \lim_{n\to\infty} \sum_{t' \in \mathbb{T}_n} |A(t') - A(t)| \in [0, \infty]. \tag{2.1.2}$$

Clearly, continuously differentiable functions on $[0, T]$ are of bounded variation; more generally, monotone finite functions on $[0, T]$ are of bounded variation. Conversely, we can give the following characterization of the class $BV([0, T])$.

Proposition 2.1.3 ([75, Proposition 4.2])**.** *A function $t \to A(t)$ on $[0, T]$ belongs to the class $BV([0, T])$ if and only if it can be represented as the difference of two non-decreasing right-continuous functions $t \to A^+(t)$ and $t \to A^-(t)$ of class $BV([0, T])$, i.e.,*

$$A = A^+ - A^-.$$

If in addition A is continuous, then A^+ and A^- can be chosen as continuous functions.

Remark 2.1.4 (Quantile transform)**.** Let $t \to A(t)$ be a non-decreasing right-continuous function on $[0, T]$, then we can define its *generalized right-continuous inverse function*,

$$q(s) := \inf\{t \in [0, T] \mid A(t) > s\},$$

which is itself a non-decreasing right-continuous function (see [75, Lemma 4.8] for a detailed discussion). We set by convention $q(0-) = 0$ and $q(s-) = \lim_{u\downarrow s} q(u) = \inf\{t \in [0, T] \mid A(t) \geq s\}$.

(a) As pointed out in [75], the functions A and q do not play symmetric roles in the sense that if A is continuous, q is still only right-continuous; in this case, we have $A(q(s)) = s$, and $q(A(s)) > s$ if s lies in an interval of constancy of A. The right-continuity of the function q does not stem from the right-continuity of A, but is given inherently in its definition via a strict inequality. Moreover, if A is strictly non-decreasing, then q is a continuous function: $q = A^{-1}$.

(b) The function A is the generalized right-continuous inverse function of q, that is,

$$A(t) = \inf\{s \in (A(0), A(T)] \mid q(s) > t\}.$$

(c) Let λ denote the Lebesgue measure on the interval $(A(0), A(T)]$, and define the measure ν on $[0,T]$ as the image $\lambda \circ q^{-1}$ of λ under $q : (A(0), A(T)] \mapsto [0,T]$, then

$$\nu([0,t]) = A(t).$$

If in addition A is continuous, we have $\lambda = \nu \circ A^{-1}$, respectively $\nu = \lambda \circ A$.

Theorem 2.1.5 ([75, Theorem 4.3])**.** *There is a one-to-one correspondence between finite (signed) measures ν on $[0,T]$ and right-continuous functions with bounded variation, which is given via*

$$A(t) = \nu([0,t]) = \nu^+([0,t]) - \nu^-([0,t]), \tag{2.1.3}$$

where

$$\nu^+([0,t]) = A^+(t), \quad \textit{respectively,} \quad \nu^-([0,t]) = A^-(t).$$

Definition 2.1.6. Let $f, A : [0,T] \to \mathbb{R}$ be such that f is bounded, measurable and A is of class $BV([0,T])$. Then, the *Lebesgue-Stieltjes integral* of f with respect to A is defined as the Lebesgue integral of f with respect to the measure ν, i.e.,

$$\int_{(0,t]} f(s)\,\mathrm{d}A(s) := \int_{(0,t]} f(s)\nu(\,\mathrm{d}s) = \int_{(0,t]} f(s)\nu^+(\,\mathrm{d}s) - \int_{(0,t]} f(s)\nu^-(\,\mathrm{d}s).$$

Remark 2.1.7. If A is of class $CBV([0,T])$, we can simply write $\int_0^t f(s)\,\mathrm{d}A(s)$ instead of $\int_{[0,t]} f(s)\,\mathrm{d}A(s)$, or $\int_{(0,t]} f(s)\,\mathrm{d}A(s)$, since in this case $\nu(\{t\}) = 0$. Furthermore, it can be shown that for $A \in CBV([0,T])$ and right-continuous f, the Lebesgue-Stieltjes integral $\int_0^t f(s)\,\mathrm{d}A(s)$ coincides with the *Riemann-Stieltjes integral*: It is the limit of both the upper and the lower Riemann sums, and in particular,

$$\int_0^t f(s)\,\mathrm{d}A(s) = \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n, s\le t} f(s)\big(A(s') - A(s)\big)$$

for $(\mathbb{T}_n)$ a refining sequence of partitions of $[0,T]$.

Proposition 2.1.8 ([95, Theorem I.5 c])**.** *For $A \in BV([0,T])$, respectively $CBV([0,T])$, and a measurable bounded function $f$on $[0,T]$, the function $t \to \int_{[0,t]} f(s)\,\mathrm{d}A(s)$ is again of class $BV([0,T])$, respectively $CBV([0,T])$.*

The tools from Lebesgue integration theory allow us to infer that Stieltjes integrals satisfy the following computational identities.

Proposition 2.1.9. *(i)* ***Associativity of the integral:*** *Let A be of class $BV([0,T])$ and f, g measurable, bounded functions on $[0,T]$. For the integral function $B(t) := \int_{[0,t]} g(s)\,\mathrm{d}A(s)$ we have*

$$\int_{[0,t]} f(s)\,\mathrm{d}B(s) = \int_{[0,t]} f(s)g(s)\,\mathrm{d}A(s). \tag{2.1.4}$$

(iii) ***Integration-by-parts formula or product rule:*** *Let A and B be of class $BV([0,T])$, then*

$$A(t)B(t) - A(0)B(0) = \int_{(0,t]} B(s-)\,\mathrm{d}A(s) + \int_{(0,t]} A(s)\,\mathrm{d}B(s). \tag{2.1.5}$$

(iv) ***Fundamental Theorem of Calculus (FTC):*** *Let A be of class $CBV([0,T])$ and f continuously differentiable, then*

$$f\big(A(t)\big) - f\big(A(0)\big) = \int_0^t f'\big(A(s)\big)\,\mathrm{d}A(s). \tag{2.1.6}$$

The fundamental theorem of calculus in the form of (2.1.6) can be extended to functionals F depending on the entire past evolution of A, and not only on its current value. The subsequent theorem is a corollary of Theorem 3.2.1.

Theorem 2.1.10. *Let $A \in CBV^m([0,T], S)$, where $S \subset \mathbb{R}^m$ Borel, i.e., A is a continuous function with values in S whose components are of bounded variation. Suppose moreover that $F = F(t, A^t)$ is a left-continuous non-anticipative functional (see Definition 3.1.1 and Definition 3.1.3) of class $\mathbb{C}^{1,1}([0,T])$ (see Definition 3.1.8) such that its first-order vertical derivative with respect to A in the sense of Definition 3.1.5, denoted by $\nabla_A F$, and its horizontal derivative, defined by*

$$\mathcal{D}F(t, A^t) := \lim_{h \to 0^+} \frac{F(t, A^{t-h}) - F((t-h), A^{t-h})}{h}, \tag{2.1.7}$$

if this limit is finite, are boundedness-preserving in the sense of (3.1.5)*. Then,*

$$F(T, A^T) - F(0, A^0) = \int_0^T \mathcal{D}F(s, A^s)\,\mathrm{d}s + \int_0^T \nabla_A F(s, A^s)\,\mathrm{d}A(s). \tag{2.1.8}$$

Let us now return to the question, which was the starting point for the above discussion. Will we obtain a "reasonable" integration theory for self-financing trading strategies if we suppose that S has bounded variation? Assume, for the moment, that $B \equiv 1$, then a trading strategy (ξ, η) will be self-financing if and only if its portfolio value satisfies

$$V(t) = V(0) + \int_0^t \xi(s)\,\mathrm{d}S(s), \quad 0 \le t \le T.$$

Moreover, for any risky component ξ for which the integral $\int_0^t \xi(s)\,\mathrm{d}S(s)$ is well-defined, there exists a riskless component η such that the pair (ξ, η) is a self-financing trading strategy; choose

$$\eta(t) = V(0) + \int_0^t \xi(s)\,\mathrm{d}S(s) - \xi(t)S(t). \tag{2.1.9}$$

Then, the strategy defined by

$$V(0) := 0, \quad \xi(t) := 2\big(S(t) - S(0)\big),$$

and η as in (2.1.9) will be self-financing, and ξ is of the form $\xi(t) = f'\big(S(t)\big)$ for the function $f(x) = \big(x - S(0)\big)^2$. Thus, in conjunction with the Fundamental Theorem of Calculus in the form of (2.1.6), we have

$$V(t) = V(0) + \int_0^t \xi(s)\,\mathrm{d}S(s) = \int_0^t f'\big(S(s)\big)\,\mathrm{d}S(s) = f\big(S(t)\big) - f\big(S(0)\big) = \big(S(t) - S(0)\big)^2.$$

Hence, applying this simple strategy we would stay clear of any possible losses. Moreover, as soon as the price process $S(t)$ moves away from the starting value $S(0)$, we would obtain a strictly positive profit. That is, without any knowledge of the stock price evolution other than that it satisfies the Fundamental Theorem of Calculus (2.1.6), the above strategy yields an arbitrage opportunity!

This follows from the following proposition, giving reasonable necessary and sufficient conditions for the existence of $\int_0^t \xi(s)\,\mathrm{d}S(s)$. Note that this proposition extends and elaborates an argument by Föllmer [47]. For the sake of completeness, we present here the multi-dimensional version (see [83, Proposition 2.1]).

Proposition 2.1.11. *Let $t \mapsto S(t) \in \mathbb{R}^d$ be a continuous function on $[0,T]$. For $i,j \in \{1,\dots,d\}$ and $K_{ij} \in \mathbb{R}$ with $K_{ij} = K_{ji}$, we consider the trading strategy $\xi^{ij} = (\xi_1^{ij},\dots,\xi_d^{ij})^\top$ defined by*

$$\xi_k^{ij}(t) = \begin{cases} 2\big(S_i(t) + S_j(t) - K_{ij}\big) & \text{if } i \neq j \text{ and } k = i \text{ or } k = j, \\ 2\big(S_i(t) - K_{ii}\big) & \text{if } i = j \text{ and } k = i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.10}$$

Then $\int_0^t \xi^{ij}(t)\,\mathrm{d}S(t)$ exists for all t and all i,j as the finite limit of the corresponding Riemann sums, i.e.,

$$\int_0^t \xi^{ij}(s)\,\mathrm{d}S(s) = \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n,\, s\le t} \xi^{ij}(s)(S(s') - S(s)), \tag{2.1.11}$$

if and only if the covariations *(along a given refining sequence of partitions $(\mathbb{T}_n)$ of $[0,T]$),*

$$[S_i,S_j](t) := \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n,\, s\le t} (S_i(s') - S_i(s))(S_j(s') - S_j(s)), \tag{2.1.12}$$

exist in $\mathbb{R}$ for all t and all i,j. In this case, we have

$$\int_0^t \xi^{ii}(s)\,\mathrm{d}S(s) = \big(S_i(t) - K_{ii}\big)^2 - \big(S_i(0) - K_{ii}\big)^2 - [S_i,S_i](t), \tag{2.1.13}$$

and, for $i \neq j$,

$$\int_0^t \xi^{ij}(s)\,\mathrm{d}S(s) = \big(S_i(t)+S_j(t) - K_{ij}\big)^2 - \big(S_i(0) + S_j(0) - K_{ij}\big)^2 - \sum_{k,\ell\in\{i,j\}} [S_k,S_\ell](t). \tag{2.1.14}$$

Proof. First, consider the case $i = j$. Then,

$$\begin{aligned}\xi^{ii}(s) \cdot (S(s') - S(s)) &= 2(S_i(s) - K_{ii})(S_i(s') - S_i(s)) \\ &= (S_i(s') - K_{ii})^2 - (S_i(s) - K_{ii})^2 - (S_i(s') - S_i(s))^2.\end{aligned}$$

Summing over $s \in \mathbb{T}_n$ yields

$$\sum_{s \in \mathbb{T}_n,\, s \le t} \xi^{ii}(s) \cdot (S(s') - S(s)) = (S_i(t_n) - K_{ii})^2 - (S_i(0) - K_{ii})^2 - \sum_{s \in \mathbb{T}_n,\, s \le t} (S_i(s') - S_i(s))^2, \quad (2.1.15)$$

with $t_n = \max\{s' \mid s \in \mathbb{T}_n,\ s \le t\} \searrow t$ as $n \uparrow \infty$. Clearly, the limit of the left-hand side exists if and only if the limit of the right-hand side exists, which gives the result for the case $i = j$. For $i \neq j$, we conclude analogously as above by using the already established existence of $[S_k, S_k](t)$ for all k and t, in conjunction with the fact that $\sum_{k,\ell \in \{i,j\}} [S_k, S_\ell] = [S_i + S_j, S_i + S_j]$. □

Thus, if we wish to work with the very simple strategies of the form (2.1.10), we must necessarily assume that the asset price trajectory S admits all pathwise quadratic variations and covariations of the form (2.1.12). The other way around, if we suppose that the quadratic variation of S_i exists and vanishes identically (which is, for instance, the case if S_i is Hölder continuous for some exponent $\alpha > 1/2$), then for ξ^{ii} as in (2.1.10) and $K_{ii} = S_i(0)$, the integral $\int_0^t \xi(s)\, \mathrm{d}S(s)$ exists for all t. Letting $\eta(t) := \int_0^t \xi(s)\, \mathrm{d}S(s) - \xi(t) \cdot S(t)$, we arrive at a self-financing trading strategy with portfolio value $V(t) = (S_i(t) - S_i(0))^2$, which gives us an arbitrage opportunity as soon as S_i is not constant.

These two aspects imply that it is crucial to require that price trajectories S of a risky asset possess all covariations $[S_i, S_j]$ in the sense of (2.1.12). Föllmer [46] showed that, if in addition the covariations are continuous functions of t, Itô's formula holds in a strictly pathwise sense (see also [86] for additional background and an English translation of [46]). This pathwise Itô formula has recently been extended by Dupire, Cont and Fournié [19, 20, 21, 34] to the functional setting, where the outcome may depend on all of the past values of the underlying trajectory, and not only on its current value, which is the subject of the next section.

2.2 Functional Itô calculus

In [20], the strictly pathwise Itô formula [46] is extended to non-anticipative functionals on the space $D([0,T], \mathbb{R}^d)$ of $\mathbb{R}^d$-valued càdlàg paths. It is required that functionals should admit certain directional (pathwise) derivatives, but, importantly, no Fréchet differentiability is imposed. Alternatively, Cosso and Russo [22] have introduced the functional Itô calculus via regularization, which is closely related to Banach space valued calculus via regularization for window processes; see [26, 27, 28, 29, 30, 78]. Although this alternative approach is close to a purely pathwise approach, there is still a probability space in the background.

In the following, we will summarize and compare the approach taken in [20] with the alternative approach from [22].

2.2.1 Functionals on spaces of paths

In [22], they fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $T \in (0, \infty)$, take a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual conditions, and consider a real-valued continuous (respectively $\mathbb{P}$-a.s. integrable) process $X = (X(t))_{t\in[0,T]}$ (respectively $Y = (Y(t))_{t\in[0,T]}$). Since every real continuous process X is naturally extended to the real line by setting $X(t) = X(0),\ t \leq 0$, and $X(t) = X(T),\ t \geq T$, this allows them to define a $C([-T, 0])$-valued process $\mathbb{X} = (\mathbb{X}_t)_{t\in\mathbb{R}}$, called the *window process* associated to X, which is given by

$$\mathbb{X}_t := \{X(t + x),\ x \in [-T, 0]\}, \quad t \in \mathbb{R}. \tag{2.2.1}$$

Definition 2.2.1 ([22, Definition 2.1])**.** If for every $t \in [0, T]$, the limit

$$\int_0^t Y(s)\,\mathrm{d}^- X(s) := \lim_{\epsilon\to 0+} \int_0^t Y(s) \frac{X(s+\epsilon) - X(s)}{\epsilon}\,\mathrm{d}s \tag{2.2.2}$$

exists in probability and if the resulting random function admits a continuous modification, that process is denoted by $\int_0^\cdot Y\,\mathrm{d}^- X$ and called the *forward integral of* Y *with respect to* X.

Recall that a family $\big(H(t)^{(\epsilon)}\big)_{t\in[0,T]}$ is said to converge to $(H(t))_{t\in[0,T]}$ in the *ucp sense* if and only if $\sup_{0\leq t\leq T} |H(t)^{(\epsilon)} - H(t)|$ goes to 0 in probability, as $\epsilon \to 0^+$. Then the following can be derived (see [22, Proposition 2.1]): If the limit in (2.2.2) exists in the ucp sense, then the forward integral $\int_0^\cdot Y\,\mathrm{d}^- X$ of Y with respect to X exists.

Definition 2.2.2 ([22, Definition 2.3])**.** The *covariation or bracket* of two continuous processes X and Y is defined as

$$[X, Y](t) = [Y, X](t) = \lim_{\epsilon\to 0+} \frac{1}{\epsilon} \int_0^t (X(s+\epsilon) - X(s))(Y(s+\epsilon) - Y(s))\,\mathrm{d}s, \quad t \in [0, T], \tag{2.2.3}$$

if the limit exists in probability, provided that the limiting function admits a continuous version (this is the case, for instance, if the limit holds in the ucp sense). If $X = Y$, X is said to be a *finite quadratic variation process* and we set $[X] := [X, X]$.

Note that the forward integral and the covariation generalize the classical Itô integral and covariation for semimartingales. In particular, the following holds (for a proof, see [78]).

Proposition 2.2.3. *(i) Let S^1, S^2 be continuous $\mathcal{F}$-semimartingales. Then, $[S^1, S^2] = [M^1, M^2]$, where M^1 (respectively M^2) is the local martingale part of S^1 (respectively S^2).*

(ii) Let A be a continuous bounded variation process and Y be a càdlàg process (or viceversa), then $[A] = [Y, A] = 0$. Moreover, $\int_0^\cdot Y\,\mathrm{d}^- A = \int_0^\cdot Y\,\mathrm{d}A$ is the Lebesgue Stieltjes integral.

(iii) If W is a Brownian motion and Y is an $\mathcal{F}$-progressively measurable process such that $\int_0^T Y(s)^2\,\mathrm{d}s < \infty$, $\mathbb{P}$-a.s., then $\int_0^\cdot Y\,\mathrm{d}^- W = \int_0^\cdot Y\,\mathrm{d}W$ is the Itô integral of Y with respect to W.

Theorem 2.2.4 ([22, Theorem 2.1]). *Let $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ be of class $C^{1,2}([0,T] \times \mathbb{R})$ and $X = (X(t))_{t\in[0,T]}$ be a real continuous finite quadratic variation process. Then, the following Itô formula holds,* $\mathbb{P}$*-a.s.,*

$$F(t, X(t)) = F(0, X(0)) + \int_0^t \partial_t F(s, X(s))\,\mathrm{d}s + \int_0^t \partial_x F(s, X(s))\,\mathrm{d}^- X(s) + \frac{1}{2}\int_0^t \partial_x^2 F(s, X(s))\,\mathrm{d}[X](s). \tag{2.2.4}$$

Note that although for simplicity we have assumed here that X is real-valued all these considerations can be directly transferred to the multi-dimesional case.

On the other hand, Cont and Fournié [20] work with arbitrary but fixed paths, instead of working with processes. In particular, they use the notion of quadratic variation in the sense of Föllmer [46], which we recall in the following.

Definition 2.2.5 (Quadratic variation). Let $T > 0$, $(\mathbb{T}_n) = (\mathbb{T}_n)_{n\in\mathbb{N}}$ be a refining sequence of partitions of $[0,T]$, and $X, Y \in C([0,T],\mathbb{R})$. We say that X and Y admit the *continuous covariation* $[X,Y]$ along $(\mathbb{T}_n)_{n\in\mathbb{N}}$ if and only if for all $t \in [0,T]$ the sequence

$$\sum_{\substack{s\in\mathbb{T}_n\\ s\le t}} (X(s') - X(s))(Y(s') - Y(s)) \tag{2.2.5}$$

converges to a finite limit, denoted $[X,Y](t)$, and if $t \mapsto [X,Y](t)$ is continuous. If $X = Y$, we say that X admits the *continuous quadratic variation* $[X]$ along $(\mathbb{T}_n)_{n\in\mathbb{N}}$ (notation: $X \in QV$), and we set $[X] := [X,X]$. We say that $X \in C([0,T],\mathbb{R}^d)$ admits the *continuous quadratic variation* along $(\mathbb{T}_n)$ (notation: $X \in QV^d$) if and only if the functions X_i, $i = 1,\dots,d$, and $X_i + X_j$, $i,j = 1,\dots,d$, $i \ne j$, do. Writing $\mathcal{S}_+^d$ for the class of symmetric nonnegative definite $d \times d$ matrices, the quadratic variation of $X \in C([0,T],\mathbb{R}^d)$ is given by the $\mathcal{S}_+^d$-valued function $[X]$, defined by

$$[X]_{ii} = [X_i], \quad [X]_{ij} = \frac{1}{2}\Big([X_i + X_j] - [X_i] - [X_j]\Big) = [X_i, X_j], \quad i \ne j. \tag{2.2.6}$$

Note that the quadratic variation depends strongly on the particular choice of the refining sequence of partitions. For example, it is shown in [50, p.47] that for any continuous function $X : [0,1] \mapsto \mathbb{R}$ there exists a refining sequence of partitions along which the quadratic variation of X is identically zero. An example where $[X,Y]$ does not exist even though both $[X]$ and $[Y]$ exist is also given in [81, Proposition 2.7]. In addition, QV^d is not a vector space [81]. Also note that for $X, A \in QV$ with $[A] \equiv 0$, the sum $X + A$ also belongs to QV. Moreover,

$$[X,A](t) = 0 \quad \text{and} \quad [X + A](t) = [X](t) \quad \text{for all } t.$$

In case of $A \in CBV([0,T])$ this is equivalent to [80, Remark 8].

For $X \in QV$, the pathwise Föllmer integral is defined as the limit of the corresponding non-anticipative Riemann sums along $(\mathbb{T}_n)$:

$$\int_0^t Y(s)\,\mathrm{d}^{(\mathbb{T}_n)} X(s) := \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n} Y(s)(X(s') - X(s)). \tag{2.2.7}$$

Here, Y must be suitably integrable with respect to X, and in order to define (2.2.7) we need to a priori fix the refining sequence of partitions, along which the quadratic variation is computed. In [20], they therefore use the notation $\int_0^t \partial_x F(s, X(s))\, \mathrm{d}^{(\mathbb{T}_n)} X(s)$ so as to explicitly account for this dependence.

Theorem 2.2.6 (Pathwise Itô formula [46])**.** *Let $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ be of class $C^{1,2}([0,T] \times \mathbb{R})$ and $X \in QV$. Then, the following pathwise Itô formula holds*

$$\begin{aligned} F(t, X(t)) = F(0, X(0)) &+ \int_0^t \partial_t F(s, X(s))\, \mathrm{d}s \\ &+ \int_0^t \partial_x F(s, X(s))\, \mathrm{d}^{(\mathbb{T}_n)} X(s) + \frac{1}{2} \int_0^t \partial_x^2 F(s, X(s))\, \mathrm{d}[X](s). \end{aligned} \tag{2.2.8}$$

One of the main issues in *functional* Itô calculus is the definition of functional pathwise derivatives, whose introduction necessitates of discontinuities/jumps. In [20], Cont and Fournié deal with this issue by considering functionals defined on spaces of càdlàg paths. Let $T > 0$ and $D \subset \mathbb{R}^n$ be an arbitrary subset of $\mathbb{R}^n$. By convention, a "D-valued càdlàg function" is a right-continuous function $f : [0,T] \mapsto D$ with left limits such that for each $t \in (0,T]$, $f(t-) \in D$, $\Delta f(t) := f(t) - f(t-)$ denotes the jump of f at time t, and f_t denotes the restriction of f to the interval $[0,t]$. They fix an open subset $U \subset \mathbb{R}^d$ and a Borel subset $S \subset \mathbb{R}^m$, $m \in \mathbb{N}$, and denote by $\mathcal{U}_t = D([0,t], U)$ the space of U-valued càdlàg functions, respectively by $C([0,t], U)$ the space of continuous functions with values in U (analogously for S). Cont and Fournié [20] work with *non-anticipative* functionals, i.e., families $Y : [0,T] \times \mathcal{U}_T \times \mathcal{S}_T \mapsto \mathbb{R}$ such that for all $(t, X, V) \in [0,T] \times \mathcal{U}_T \times \mathcal{S}_T$, $Y(t, X, V) = Y(t, X_t, V_t)$. A non-anticipative functional can be represented as $Y(t, X, V) = F_t(X_t, V_t)$, where $(F_t)_{t \in [0,T]}$ is a family of maps $F_t : \mathcal{U}_t \times \mathcal{S}_t \mapsto \mathbb{R}$. For a path $\omega \in \mathcal{U}_T \times \mathcal{S}_T$, they denote by ω_{t-} the path defined by

$$\omega_{t-}(u) = \omega(u), \quad u \in [0,t), \quad \omega_{t-}(t) = \omega(t-). \tag{2.2.9}$$

Note that ω_{t-} is càdlàg and should *not* be confounded with the left-continuous path $u \mapsto \omega(u-)$. This definition induces the following notion of predictability: A functional $Y : [0,T] \times \mathcal{U}_T \times \mathcal{S}_T \mapsto \mathbb{R}$ is *predictable* if and only if for all $(t, \omega) \in [0,T] \times \mathcal{U}_T \times \mathcal{S}_T$, $Y(t, \omega) = Y(t, \omega_{t-})$. Cont and Fournié [20] focus on functionals

$$F = (F_t)_{t \in [0,T]}, \quad F_t : \mathcal{U}_t \times \mathcal{S}_t \mapsto \mathbb{R},$$

where F has a "predictable" dependence with respect to the second argument, i.e.,

$$\forall t \in [0,T],\ \forall (X, V) \in \mathcal{U}_t \times \mathcal{S}_t, \quad F_t(X_t, V_t) = F_t(X_t, V_{t-}). \tag{2.2.10}$$

Such functionals F can then be viewed as maps on the vector bundle space $\mathcal{Y} = \bigcup_{t \in [0,T]} \mathcal{U}_t \times \mathcal{S}_t$.

On the other hand, in [22], Russo et al. consider a map

$$\begin{aligned} \mathcal{W} : [0,T] \times C([-T,0]) &\mapsto \mathbb{R} \\ (t, \eta) &\to \mathcal{W}(t, \eta). \end{aligned}$$

It is readily observed that these two formulations are equivalent to one another: If we have a family $F = (F_t)_{t\in[0,T]}$ of maps $F_t : C([0,t]) \to \mathbb{R}$, the map $\mathcal{W} : [0,T] \times C([-T,0]) \to \mathbb{R}$ can be defined as

$$\mathcal{W}(t,\eta) := F_t(\eta(\cdot + T)\,\big|_{[0,t]}), \quad (t,\eta) \in [0,T] \times C([-T,0]).$$

The other way around, if we have given a map $\mathcal{W} : [0,T] \times C([-T,0]) \to \mathbb{R}$, then the family of maps $F = (F_t)_{t\in[0,T]}$ can be defined as

$$F_t(\widetilde{\eta}) := \mathcal{W}(t,\eta), \quad (t,\widetilde{\eta}) \in [0,T] \times C([0,t]), \tag{2.2.11}$$

where $\eta(x) := \widetilde{\eta}(x+t)\mathbb{1}_{[-t,0]}(x) + \widetilde{\eta}(-t)\mathbb{1}_{[-T,-t)}(x)$, $x \in [-T,0]$. Note that the map $\mathcal{W}$ contains more information than F, since in (2.2.11) the values of $\mathcal{W}$ at $(t,\eta) \in [0,T] \times C([-T,0])$ where η is non-constant on the interval $[-T,-t]$ are not taken into account. However, the equivalence between these two formulations is established by the fact that when considering the decomposition of $\mathcal{W}$ in terms of self-financing strategies, this additional information is not crucial, as can be seen from the derivation of the Itô formula. Thus, the representation in [22] allows one to work with a single map instead of working with a family of maps. Moreover, the time variable and the underlying path itself have two distinct roles, as is the case for the time variable and the space variable in the classical Itô calculus. The point is that this allows them to focus only on the definition of horizontal and vertical derivatives (as emphasized in [22, Remark 2.1]), and to define the horizontal derivative independently of the time derivative. Since one can *not*, in general, extend in a unique way a functional $\mathcal{W}$ defined on $C([-T,0])$ to $D([-T,0])$, in [22] they introduce an intermediate space between $C([-T,0])$ and $D([-T,0])$, denoted by $\mathfrak{C}([-T,0])$, the set of bounded trajectories on $[-T,0]$, which are continuous on $[-T,0)$ and possibly have a jump at 0.

Definition 2.2.7 ([22, Definition 2.6])**.** The set $\mathfrak{C}([-T,0])$ comprises the bounded functions $\eta : [-T,0] \to \mathbb{R}$ such that η is continuous on $[-T,0)$, equipped with the following topology.
Convergence. The set $\mathfrak{C}([-T,0])$ is endowed with a topology inducing the following convergence: $(\eta_n)_n$ converges to η in $\mathfrak{C}([-T,0])$ as n tends to infinity if the following conditions apply:

(i) $\|\eta_n\|_\infty := \sup_{x\in[-T,0]} |\eta_n(x)| \le C$ for any $n \in \mathbb{N}$ for some positive constant C independent of n;

(ii) $\sup_{x\in K} |\eta_n(x) - \eta(x)| \to 0$ as n tends to infinity for any compact set $K \subset [-T,0)$;

(iii) $\eta_n(0) \to \eta(0)$ as n tends to infinity.

Topology. For each compact $K \subset [-T,0)$, define the seminorm p_K on $\mathfrak{C}([-T,0])$ by

$$p_K(\eta) = \sup_{x\in K} |\eta(x)| + |\eta(0)|, \quad \forall \eta \in \mathfrak{C}([-T,0]).$$

Let $M > 0$ and $\mathfrak{C}_M([-T,0])$ be the set of functions in $\mathfrak{C}([-T,0])$, which are bounded by M. Still denote by p_K the restriction of p_K to $\mathfrak{C}_M([-T,0])$ and consider the topology on $\mathfrak{C}_M([-T,0])$ induced by the collection of seminorms $(p_K)_K$. Then, $\mathfrak{C}([-T,0])$ is endowed with the smallest topology (inductive topology) turning all the inclusions $i_M : \mathfrak{C}_M([-T,0]) \to \mathfrak{C}([-T,0])$ into continuous maps.

2.2.2 Functional derivatives

In order to define the functional derivatives, in [22] the "past" is separated from the "present" of the path $\eta \in \mathfrak{C}([-T,0])$. Informally, the horizontal derivative will call in the past values of η, namely $\{\eta(x) : x \in [-T,0)\}$, while the vertical derivative will call in only the present value of η, namely $\eta(0)$.

Definition 2.2.8 ([22, Definition 2.7])**.** The set $\mathfrak{C}([-T,0))$ comprises the bounded continuous functions $\gamma : [-T,0) \to \mathbb{R}$, equipped with the following topology.
Convergence. The set $\mathfrak{C}([-T,0))$ is endowed with a topology inducing the following convergence: $(\gamma_n)_n$ converges to γ in $\mathfrak{C}([-T,0))$ as n tends to infinity if the following conditions apply:

(i) $\sup_{x\in[-T,0)} |\gamma_n(x)| \leq C$ for any $n \in \mathbb{N}$ for some positive constant C independent of n;

(ii) $\sup_{x\in K} |\gamma_n(x) - \gamma(x)| \to 0$ as n tends to infinity for any compact set $K \subset [-T,0)$.

Topology. For each compact $K \subset [-T,0)$, define the seminorm q_K on $\mathfrak{C}([-T,0))$ by

$$q_K(\gamma) = \sup_{x\in K} |\gamma(x)|, \quad \forall \gamma \in \mathfrak{C}([-T,0)).$$

Let $M > 0$ and $\mathfrak{C}_M([-T,0))$ be the set of functions in $\mathfrak{C}([-T,0))$, which are bounded by M. Still denote q_K the restriction of q_K to $\mathfrak{C}_M([-T,0))$ and consider the topology on $\mathfrak{C}_M([-T,0))$ induced by the collection of seminorms $(q_K)_K$. Then, $\mathfrak{C}([-T,0))$ is endowed with the smallest topology (inductive topology) turning all the inclusions $i_M : \mathfrak{C}_M([-T,0)) \to \mathfrak{C}([-T,0))$ into continuous maps.

For every functional $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$, the space $\mathfrak{C}([-T,0))$ can be exploited to introduce a map $\widetilde{w} : \mathfrak{C}([-T,0)) \times \mathbb{R} \to \mathbb{R}$, which separates the "past" and the"present".

Definition 2.2.9 ([22, Definition 2.8])**.** Let $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$ and define $\widetilde{w} : \mathfrak{C}([-T,0)) \times \mathbb{R} \to \mathbb{R}$ as

$$\widetilde{w}(\gamma, a) = w(\gamma \mathbb{1}_{[-T,0)} + a\mathbb{1}_{\{0\}}), \quad \forall (\gamma, a) \in \mathfrak{C}([-T,0)) \times \mathbb{R}. \tag{2.2.12}$$

In particular,

$$w(\eta) = \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0)), \quad \forall \eta \in \mathfrak{C}([-T,0]).$$

Definition 2.2.10 ([22, Definition 2.9])**.** Consider a functional $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$ and a path $\eta \in \mathfrak{C}([-T,0])$.

(i) The functional w admits a *horizontal derivative* at η if the following limit exists and is finite

$$D^H w(\eta) := \lim_{\epsilon \to 0^+} \frac{w(\eta(\cdot)\mathbb{1}_{[-T,0)} + \eta(0)\mathbb{1}_{\{0\}}) - w(\eta(\cdot - \epsilon)\mathbb{1}_{[-T,0)} + \eta(0)\mathbb{1}_{\{0\}})}{\epsilon} \tag{2.2.13}$$

with $\widetilde{w}$ from Definition 2.2.9. Alternatively, if $\widetilde{w}$ is as in (2.2.12), then $\widetilde{w}$ admits a *horizontal derivative* at $(\gamma, a) \in \mathfrak{C}([-T,0)) \times \mathbb{R}$ if the following limit exists and is finite

$$D^H \widetilde{w}(\gamma, a) := \lim_{\epsilon \to 0^+} \frac{\widetilde{w}(\gamma(\cdot), a) - \widetilde{w}(\gamma(\cdot - \epsilon), a)}{\epsilon}. \tag{2.2.14}$$

Clearly, if $D^H w(\eta)$ exists, then $D^H \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0))$ exists and they are equal; the other way around, if $D^H \widetilde{w}(\gamma, a)$ exists, then $D^H w(\gamma \mathbb{1}_{[-T,0)} + a\mathbb{1}_{\{0\}})$ exists and they are equal.

(ii) The functional w admits a *first-order vertical derivative* at η if the first-order partial derivative at $(\eta\big|_{[-T,0)}, \eta(0))$ of $\widetilde{w}$ with respect to its second argument, denoted by $\partial_a \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0))$, exists and we set

$$D^V w(\eta) := \partial_a \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0)). \tag{2.2.15}$$

(iii) The functional w admits a *second-order vertical derivative* at η if the second-order partial derivative at $(\eta\big|_{[-T,0)}, \eta(0))$ of $\widetilde{w}$ with respect to its second argument, denoted by $\partial^2_{aa} \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0))$, exists and we set

$$D^{VV} w(\eta) := \partial^2_{aa} \widetilde{w}(\eta\big|_{[-T,0)}, \eta(0)). \tag{2.2.16}$$

Definition 2.2.11 ([22, Definition 2.10])**.** Let $\mathcal{W} : C([-T,0]) \to \mathbb{R}$ and $\eta \in C([-T,0])$. Suppose that there exists a unique extension $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$ of $\mathcal{W}$ (e.g., if $\mathcal{W}$ is continuous with respect to the topology of $\mathfrak{C}([-T,0])$). Then,

(i) the *horizontal derivative* of $\mathcal{W}$ at η is defined as

$$D^H \mathcal{W}(\eta) := D^H w(\eta);$$

(ii) the *first-order vertical derivative* of $\mathcal{W}$ at η is defined as

$$D^V \mathcal{W}(\eta) := D^V w(\eta);$$

(iii) the *second-order vertical derivative* of $\mathcal{W}$ at η is defined as

$$D^{VV} \mathcal{W}(\eta) := D^{VV} w(\eta).$$

In [20], Cont and Fournié introduce the following notions of horizontal extension and the vertical perturbation of a càdlàg path in order to define the functional derivatives.

Definition 2.2.12. Let $X \in D([0,T], U)$ and X_t be its restriction to $[0,t]$ with $t < T$.

(i) For $h \geq 0$, the *horizontal extension* $X_{t,h} \in D([0,t+h], \mathbb{R}^d)$ of X_t to $[0,t+h]$ is defined as

$$X_{t,h}(u) = X(u), \quad u \in [0,t]; \qquad X_{t,h}(u) = X(t), \quad u \in (t, t+h]. \tag{2.2.17}$$

(ii) For $h \in \mathbb{R}^d$ sufficiently small, the *vertical perturbation* X^h_t of X_t is defined as the càdlàg path obtained by shifting the endpoint by the quantity h:

$$X^h_t(u) = X_t(u), \quad u \in [0,t), \qquad X^h_t(t) = X(t) + h, \tag{2.2.18}$$

or, equivalently, $X^h_t(u) = X_t(u) + h\mathbb{1}_{t=u}$. By convention, the vertical perturbation precedes the horizontal extension, so $X^u_{t,h}$ is still a càdlàg path.

Definition 2.2.13 ([20, Definition 6]). The *horizontal derivative* at $(X, V) \in \mathcal{U}_t \times \mathcal{S}_t$ of a non-anticipative functional $F = (F_t)_{t\in[0,T)}$ is defined as

$$\mathcal{D}_t F(X, V) := \lim_{h\to 0^+} \frac{F_{t+h}(X_{t,h}, V_{t,h}) - F_t(X, V)}{h} \tag{2.2.19}$$

if this limit exists and is finite. If (2.2.19) is well-defined for all $(X, V) \in \mathcal{Y}$, the map

$$\begin{aligned} \mathcal{D}_t F : \mathcal{U}_t \times \mathcal{S}_t &\mapsto \mathbb{R} \\ (X, V) &\to \mathcal{D}_t F(X, V) \end{aligned} \tag{2.2.20}$$

defines a non-anticipative functional $\mathcal{D}F = (\mathcal{D}_t F)_{t\in[0,T)}$, the *horizontal derivative* of F.

Definition 2.2.14 ([20, Definition 8]). A non-anticipative functional F is said to be *vertically differentiable with respect to* X at $(X, V) \in D([0,t], \mathbb{R}^d) \times \mathcal{S}_t$ if the following map

$$\begin{aligned} \mathbb{R}^d &\mapsto \mathbb{R} \\ e &\to F_t(X^e, V) \end{aligned}$$

is differentiable at 0. Its gradient at 0

$$\nabla_x F_t(X, V) = (\partial_i F_t(X, V)\ i = 1, \dots, d)\,, \quad \partial_i F_t(X, V) = \lim_{h\to 0} \frac{F_t(X^{he_i}, V) - F_t(X, V)}{h}, \tag{2.2.21}$$

is called the *vertical derivative with respect to* X of F_t at (X, V). Here, e_i denotes the i-th unit vector in $\mathbb{R}^d$. If (2.2.21) is well-defined for all $(X, V) \in \mathcal{Y}$, the map

$$\begin{aligned} \nabla_x F : \mathcal{U}_t \times \mathcal{S}_t &\mapsto \mathbb{R}^d \\ (X, V) &\to \nabla_x F_t(X, V) \end{aligned} \tag{2.2.22}$$

defines a non-anticipative functional $\nabla_x F = (\nabla_x F_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$, which is called the *vertical derivative with respect to* X of F.

Thus, since Russo et al. work [22] with a single map instead of considering a family of maps, and time and the underlying path itself do not interplay, it follows that the horizontal derivative $\mathcal{D}_t F$ from (2.2.19) can be written as the sum of $D^H w(\eta)$ and the time derivative. Also note that the definition of the horizontal derivative $D^H w(\eta)$ is based on a limit on the left, while in [20] the definition of the horizontal derivative is based on a limit on the right. In particular (see also [22, Remark 2.6]), applying the approach from [20] would lead to the following alternative formulation of $D^H w(\eta)$:

$$D^{H,+} w(\eta) := \lim_{\epsilon\to 0^+} \frac{w(\eta(\cdot + \epsilon)\mathbb{1}_{[-T,0)} + \eta(0)\mathbb{1}_{\{0\}}) - w(\eta(\cdot)\mathbb{1}_{[-T,0)} + \eta(0)\mathbb{1}_{\{0\}})}{\epsilon}. \tag{2.2.23}$$

To better see the difference between (2.2.13) and (2.2.23), consider a real-valued continuous finite quadratic variation process X with associated window process $\mathbb{X}$. Then, the definition (2.2.23) of $D^{H,+} w(\mathbb{X}_t)$ takes into account the increment $\widetilde{w}(\mathbb{X}_t(\cdot + \epsilon)\big|_{[-T,0)}, X(t))) -$

$\widetilde{w}(\mathbb{X}_t(\cdot)\big|_{[-T,0)}, X(t))$, comparing the present value of $w(\mathbb{X}_t) = \widetilde{w}(\mathbb{X}_t\big|_{[-T,0)}, X(t))$ with an hypothetical future value $\widetilde{w}(\mathbb{X}_t\big|_{[-T,0)}(\cdot+\epsilon), X(t))$, obtained assuming a constant time evolution for X. On the other hand, the definition (2.2.13) of $D^H w(\eta)$ considers the increment $\widetilde{w}(\mathbb{X}_t\big|_{[-T,0)}, X(t)) - \widetilde{w}(\mathbb{X}_{t-\epsilon}\big|_{[-T,0)}, X(t))$, where only the present and past values of X are taken into account (and where the trajectory of X is extended in a constant way before time 0). In particular, since (2.2.13) does not call in the future evolution of the path, no future time behavior for X needs to be specified in [22], but only a past evolution before time 0.

2.2.3 Functional Itô formula

In [20], Cont and Fournié introduce the following notion of distance: For $T \ge t' = t+h \ge t \ge 0$, $(X,V) \in \mathcal{U}_t \times \mathcal{S}_t$, and $(X',V') \in D([0,t+h],\mathbb{R}^d) \times \mathcal{S}_{t+h}$, define

$$d_\infty\left((X,V),(X',V')\right) = \sup_{u\in[0,t+h]} |X_{t,h}(u) - X'(u)| + \sup_{u\in[0,t+h]} |V_{t,h}(u) - V'(u)| + h. \qquad (2.2.24)$$

Then, the pair $(\mathcal{Y}, d_\infty)$ represents a metric space, and if the paths $(X,V),(X',V')$ are defined on the same time interval, $d_\infty\left((X,V),(X',V')\right)$ simply coincides with the distance in supremum norm. The following regularity properties are used in [20]:

(i) A non-anticipative functional $F = (F_t)_{t\in[0,T]}$ is said to be *left-continuous* if

$$\forall t \in [0,T],\ \forall \epsilon > 0,\ \forall (X,V) \in \mathcal{U}_t \times \mathcal{S}_t,\ \exists \lambda > 0,\ \forall h \in [0,t],\ \forall (X',V') \in \mathcal{U}_{t-h} \times \mathcal{S}_{t-h},$$

$$d_\infty\left((X,V),(X',V')\right) < \lambda \quad\Rightarrow\quad |F_t(X,V) - F_{t-h}(X',V')| < \epsilon. \qquad (2.2.25)$$

The set of all left-continuous non-anticipative functionals is denoted by $\mathbb{F}_l^\infty$ (see [20, Definition 3]).

(ii) A non-anticipative functional $F = (F_t)_{t\in[0,T]}$ is *boundedness-preserving* (see [20, Definition 5]) if it is bounded on each bounded set of paths, i.e., $F \in \mathbb{B}$ if and only if for any compact subset $K \subset U$ and any $R > 0$, there exists a constant $C_{K,R}$ such that

$$\forall t \in [0,T],\ \forall (X,V) \in D([0,t],K)\times\mathcal{S}_t,\ \sup_{s\in[0,t]} |v(s)| < R \ \Rightarrow\ |F_t(X,V)| < C_{K,R}. \quad (2.2.26)$$

In particular, if F is boundedness-preserving, then it is "locally bounded" in the neighborhood of any given path. That is,

$$\forall (X,V) \in \mathcal{U}_T \times \mathcal{S}_T,\ \exists C > 0,\ \lambda > 0,\ \forall t \in [0,T],\ \forall (X',V') \in \mathcal{U}_t \times \mathcal{S}_t,$$

$$d_\infty\left((X,V),(X',V')\right) < \lambda \quad\Rightarrow\quad \forall t \in [0,T], \quad |F_t(X',V')| \le C. \qquad (2.2.27)$$

Definition 2.2.15. If the functional $F = (F_t)_{t\in[0,T]}$ admits a horizontal, respectively vertical, derivative $\mathcal{D}F$, respectively $\nabla_x F$, one may iterate these operations in order to define higher order horizontal and vertical derivatives.

(i) Let $I \subset [0,T]$ be a subinterval of $[0,T]$. The set $\mathbb{C}^{j,k}(I)$ is comprised of all non-anticipative functionals $F = (F_t)_{t\in I}$ satisfying the following conditions:

(a) F is continuous at fixed times: $F_t : \mathcal{U}_t \times \mathcal{S}_t \mapsto \mathbb{R}$ is continuous with respect to the supremum norm;

(b) F admits j horizontal derivatives and k vertical derivatives with respect to X at all $(X,V) \in \mathcal{U}_t \times \mathcal{S}_t,\ t \in I$;

(c) $\mathcal{D}^i F,\ i \le j,\ \nabla_x^m F,\ m \le k$, are continuous at fixed times $t \in I$.

(ii) A non-anticipative functional $F \in \mathbb{C}^{1,2}([0,T))$ is called *regular* if

(a) $F,\ \nabla_x F,\ \nabla_x^2 F \in \mathbb{F}_l^\infty$,

(b) $\nabla_x^2 F,\ \mathcal{D}F$ satisfy the local boundedness property (2.2.27).

On the other hand, in [22], Russo et al. require the following regularity assumptions.

Definition 2.2.16 ([22, Definition 2.10/Definition 2.12])**.** (a) The map $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$ is of class $\mathfrak{C}^{1,2}$(past $\times$ present) if the following conditions are satisfied:

(i) w is continuous;

(ii) $D^H w$ exists everywhere on $\mathfrak{C}([-T,0])$ and for every $\gamma \in \mathfrak{C}([-T,0])$ the map

$$(\epsilon, a) \mapsto D^H \widetilde{w}(\gamma(\cdot - \epsilon), a), \qquad (\epsilon, a) \in [0,\infty) \times \mathbb{R},$$

is continuous on $[0,\infty) \times \mathbb{R}$;

(iii) $D^V w$ and $D^{VV} w$ exist everywhere on $\mathfrak{C}([-T,0])$ and are continuous.

(b) The map $\mathcal{W} : C([-T,0]) \to \mathbb{R}$ is of class $\mathcal{C}^{1,2}$(past $\times$ present) if $\mathcal{W}$ admits a (necessarily unique) extension $w : \mathfrak{C}([-T,0]) \to \mathbb{R}$ that is of class $\mathfrak{C}^{1,2}$(past $\times$ present).

Remark 2.2.17. As pointed out in [22, Remark 2.4], in the previous definition the same class $\mathfrak{C}^{1,2}$(past $\times$ present) of functionals is still obtained if point (ii) in (a) is replaced by

(ii)' $D^H w$ exists everywhere on $\mathfrak{C}([-T,0])$ and for every $\gamma \in \mathfrak{C}([-T,0])$ there exists $\delta(\gamma) > 0$ such that the map

$$(\epsilon, a) \mapsto D^H \widetilde{w}(\gamma(\cdot - \epsilon), a), \qquad (\epsilon, a) \in [0,\infty) \times \mathbb{R}, \tag{2.2.28}$$

is continuous on $[0,\delta(\gamma)) \times \mathbb{R}$.

For $\mathcal{W} : [0,T] \times C([-T,0]) \to \mathbb{R}$, the functional derivatives $D^H\mathcal{W}, D^V\mathcal{W}, D^{VV}\mathcal{W}$ (respectively, $D^H w, D^V w, D^{VV} w$) as defined in Definition 2.2.11 (respectively, Definition 2.2.10) can be directly transferred to this time-dependent case. Moreover, given the family of functionals $w : [0,T] \times \mathfrak{C}([-T,0]) \to \mathbb{R}$, one can define a map $\widetilde{w} : [0,T] \times \mathfrak{C}([-T,0]) \times \mathbb{R} \to \mathbb{R}$, allowing for the following definition.

Definition 2.2.18 ([22, Definition 2.13 and Definition 2.14])**.** Let I be $[0,T]$ or $[0,T)$.

(a) The map $w : I \times \mathfrak{C}([-T,0]) \to \mathbb{R}$ is of class $\mathfrak{C}^{1,2}((I \times \text{past}) \times \text{present})$ if the following conditions hold:

   (i) w is continuous;

   (ii) $\partial_t w$ exists everywhere on $I \times \mathfrak{C}([-T,0])$ and is continuous;

   (iii) $D^H w$ exists everywhere on $I \times \mathfrak{C}([-T,0])$ and for every $\gamma \in \mathfrak{C}([-T,0])$ the map

$$(t,\epsilon,a) \mapsto D^H \widetilde{w}(t,\gamma(\cdot - \epsilon),a), \qquad (t,\epsilon,a) \in I \times [0,\infty) \times \mathbb{R},$$

   is continuous on $I \times [0,\infty) \times \mathbb{R}$;

   (iv) $D^V w$ and $D^{VV} w$ exist everywhere on $I \times \mathfrak{C}([-T,0])$ and are continuous.

(b) The map $\mathcal{W} : I \times C([-T,0]) \to \mathbb{R}$ is of class $\mathcal{C}^{1,2}((I \times \text{past}) \times \text{present})$ if $\mathcal{W}$ admits a (necessarily unique) extension $w : I \times \mathfrak{C}([-T,0]) \to \mathbb{R}$ that is of class $\mathfrak{C}^{1,2}((I \times \text{past}) \times \text{present})$.

We will now review and compare the main results from [20] and [22].

Theorem 2.2.19 ([22, Theorem 2.2])**.** *Let* $\mathcal{W} : C([-T,0]) \to \mathbb{R}$ *be of class* $\mathcal{C}^{1,2}$(*past* $\times$ *present*) *and* $X = (X(t))_{t\in[0,T]}$ *a real-valued continuous finite quadratic variation process. Then, the following* functional Itô's formula *holds,* $\mathbb{P}$*-a.s.,*

$$\mathcal{W}(\mathbb{X}_t) = \mathcal{W}(\mathbb{X}_0) + \int_0^t D^H \mathcal{W}(\mathbb{X}_s)\,\mathrm{d}s + \int_0^t D^V \mathcal{W}(\mathbb{X}_s)\,\mathrm{d}^- X(s) + \frac{1}{2}\int_0^t D^{VV}\mathcal{W}(\mathbb{X}_s)\,\mathrm{d}[X](s), \quad 0 \le t \le T, \tag{2.2.29}$$

where $\mathbb{X} = (\mathbb{X}_t)_{t\in\mathbb{R}}$ *is the window process associated to* X*, and* $\int_0^t D^V \mathcal{W}(\mathbb{X}_s)\,\mathrm{d}^- X(s)$ *denotes the forward integral.*

Theorem 2.2.20 ([22, Theorem 2.3])**.** *Let* $\mathcal{W} : [0,T] \times C([-T,0]) \to \mathbb{R}$ *be of class* $\mathcal{C}^{1,2}(([0,T] \times$ *past*$) \times$ *present*) *and* $X = (X(t))_{t\in[0,T]}$ *a real-valued continuous finite quadratic variation process. Then, the following* functional Itô's formula *holds,* $\mathbb{P}$*-a.s.,*

$$\mathcal{W}(t,\mathbb{X}_t) = \mathcal{W}(0,\mathbb{X}_0) + \int_0^t \big(\partial_t \mathcal{W}(s,\mathbb{X}_s) + D^H \mathcal{W}(s,\mathbb{X}_s)\big)\,\mathrm{d}s + \int_0^t D^V \mathcal{W}(s,\mathbb{X}_s)\,\mathrm{d}^- X(s) + \frac{1}{2}\int_0^t D^{VV}\mathcal{W}(s,\mathbb{X}_s)\,\mathrm{d}[X](s) \tag{2.2.30}$$

for all $0 \le t \le T$.

Remark 2.2.21. Note that the classical time-dependent Itô's formula for finite quadratic variation processes (2.2.4) follows from (2.2.30) by setting $\mathcal{W}(t,\eta) = F(t,\eta(0))$ for any $(t,\eta) \in [0,T] \times C([-T,0])$, where $F \in C^{1,2}([0,T] \times \mathbb{R})$. In this case the unique continuous extension of $\mathcal{W}$ is the map $w : [0,T] \times \mathfrak{C}([-T,0]) \to \mathbb{R}$ defined by $w(t,\eta) = F(t,\eta(0))$, for all $(t,\eta) \in [0,T] \times \mathfrak{C}([-T,0])$. In addition, $D^H\mathcal{W} \equiv 0$, while $D^V\mathcal{W} = \partial_x F$ and $D^{VV}\mathcal{W} = \partial^2_{xx}F$, with $\partial_x F$ (respectively $\partial^2_{xx}F$) denoting the first-order (respectively second-order) partial derivative of F with respect to its second argument.

Proof of Theorem 2.2.19. For a detailed version of the following proof we refer to [22]. For any fixed $t \in [0,T]$, consider

$$I_0(\epsilon,t) = \int_0^t \frac{\mathcal{W}(\mathbb{X}_{s+\epsilon}) - \mathcal{W}(\mathbb{X}_s)}{\epsilon}\,\mathrm{d}s = \frac{1}{\epsilon}\int_t^{t+\epsilon} \mathcal{W}(\mathbb{X}_s)\,\mathrm{d}s - \frac{1}{\epsilon}\int_0^{\epsilon} \mathcal{W}(\mathbb{X}_s)\,\mathrm{d}s, \quad \epsilon > 0.$$

Since $(\mathcal{W}(\mathbb{X}_s))_{s\geq 0}$ is continuous, $I_0(\epsilon,t)$ converges ucp to $\mathcal{W}(\mathbb{X}_t) - \mathcal{W}(\mathbb{X}_0)$, i.e., $\sup_{0\leq t\leq T}|I_0(\epsilon,t) - (\mathcal{W}(\mathbb{X}_t) - \mathcal{W}(\mathbb{X}_0))|$ converges to 0 in probability for $\epsilon \to 0^+$.

On the other hand, $I_0(\epsilon,t)$ can be written as follows in terms of $\widetilde{w}$ from Definition 2.2.11:

$$I_0(\epsilon,t) = \int_0^t \frac{\widetilde{w}(\mathbb{X}_{s+\epsilon}\big|_{[-T,0)}, X(s+\epsilon)) - \widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s))}{\epsilon}\,\mathrm{d}s.$$

We can further split $I_0(\epsilon,t)$ into two terms, as follows

$$I_1(\epsilon,t) = \int_0^t \frac{\widetilde{w}(\mathbb{X}_{s+\epsilon}\big|_{[-T,0)}, X(s+\epsilon)) - \widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s+\epsilon))}{\epsilon}\,\mathrm{d}s, \tag{2.2.31}$$

$$I_2(\epsilon,t) = \int_0^t \frac{\widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s+\epsilon)) - \widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s))}{\epsilon}\,\mathrm{d}s. \tag{2.2.32}$$

Then, $I_1(\epsilon,t)$ converges ucp to $\int_0^t D^H w(\mathbb{X}_s)\,\mathrm{d}s$, as $\epsilon \to 0^+$, by applying the regularity assumption (ii) from Definition 2.2.16, the fundamental theorem of calculus, and an appropriate shift in time. On the other hand, $I_2(\epsilon,t)$ can be written by means of a standard Taylor expansion as the sum of the following three terms:

$$I_{21}(\epsilon,t) = \int_0^t \partial_a\widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s))\frac{X(s+\epsilon) - X(s)}{\epsilon}\,\mathrm{d}s, \tag{2.2.33}$$

$$I_{22}(\epsilon,t) = \int_0^t \partial^2_{aa}\widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s))\frac{(X(s+\epsilon) - X(s))^2}{\epsilon}\,\mathrm{d}s, \tag{2.2.34}$$

$$\begin{aligned} I_{23}(\epsilon,t) = \int_0^t \bigg(\int_0^1 &(1-\alpha)\big((\partial^2_{aa}\widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s) + \alpha(X(s+\epsilon) - X(s))) \\ &- \partial^2_{aa}\widetilde{w}(\mathbb{X}_{s|[-T,0)}, X(s))\big)\frac{(X(s+\epsilon) - X(s))^2}{\epsilon}\,\mathrm{d}\alpha\bigg)\,\mathrm{d}s. \end{aligned} \tag{2.2.35}$$

It can be shown using the given regularity assumptions that

$$I_{22}(\epsilon,t) \to \frac{1}{2}\int_0^t \partial^2_{aa}\widetilde{w}(\mathbb{X}_s\big|_{[-T,0)}, X(s))\,\mathrm{d}[X]s = \frac{1}{2}\int_0^t D^{VV}w(\mathbb{X}_s)\,\mathrm{d}[X](s),$$

in the ucp sense as $\epsilon \to 0^+$. Since X has finite quadratic variation, it can be deduced that $I_{23}(\epsilon, t) \to 0$ ucp as $\epsilon \to 0^+$. Hence, since $I_0(\epsilon, t),\ I_1(\epsilon, t),\ I_{22}(\epsilon, t)$ and $I_{23}(\epsilon, t)$ all converge ucp, it follows that the forward integral exists, i.e.,

$$I_{21}(\epsilon, t) \to \int_0^t \partial_a \widetilde{w}(\mathbb{X}_s \big|_{[-T,0)}, X(s))\, \mathrm{d}^- X(s) = \int_0^t D^V w(\mathbb{X}_s)\, \mathrm{d}^- X(s)$$

in the ucp sense, as $\epsilon \to 0^+$. □

On the other hand, in [20] the following change of variables formula is shown.

Theorem 2.2.22 ([20, Theorem 3])**.** *Let $(X, V) \in C([0,T], U) \times \mathcal{S}_T$ such that $X \in QV^d$ along a given refining sequence of partitions $(\mathbb{T}_n)$ of $[0,T]$ and $\sup_{t \in [0,T] \setminus \mathbb{T}_n} |V(t) - V(t-)| \to 0$. Denote*

$$X^n(t) := \sum_{s \in \mathbb{T}_n} X(s') I_{[s,s')}(t) + X(T) I_{\{T\}}(t), \quad 0 \le t \le T, \tag{2.2.36}$$

$$V^n(t) := \sum_{s \in \mathbb{T}_n} V(s) I_{[s,s')}(t) + V(T) I_{\{T\}}(t), \quad 0 \le t \le T, \tag{2.2.37}$$

and set

$$h^n_s := s' - s = s'_n - s_n. \tag{2.2.38}$$

Then for any non-anticipative functional $F \in \mathbb{C}^{1,2}([0,T))$ that is regular in the sense of Definition 2.2.15 and predictable with respect to the second argument in the sense of (2.2.10), the following limit

$$\lim_{n \to \infty} \sum_{s \in \mathbb{T}_n} \nabla_x F_s(X^n_{s-}, V^n_{s-})\, (X(s') - X(s)) \tag{2.2.39}$$

is well-defined. Denoting this limit by

$$\int_0^T \nabla_x F_s(X_s, V_s)\, \mathrm{d}^{(\mathbb{T}_n)} X(s),$$

we have

$$\begin{aligned} F_T(X_T, V_T) - F_0(X_0, V_0) = &\int_0^T \mathcal{D}_s F(X_s, V_s)\, \mathrm{d}s + \frac{1}{2} \sum_{i,j=1}^d \int_0^T \nabla^2_{x,ij} F_s(X_s, V_s)\, \mathrm{d}[X]_{ij}(s) \\ &+ \int_0^T \nabla_x F_s(X_s, V_s)\, \mathrm{d}^{(\mathbb{T}_n)} X(s). \end{aligned} \tag{2.2.40}$$

Proof. For a detailed version of the following proof see [20]. For $s \in \mathbb{T}_n$, consider the decomposition

$$\begin{aligned} F_{s'}(X^n_{s'-}, V^n_{s'-}) - F_s(X^n_{s-}, V^n_{s-}) = &F_{s'}(X^n_{s'-}, V^n_{s,h^n_s}) - F_s(X^n_s, V^n_s) \\ &+ F_s(X^n_s, V^n_{s-}) - F_s(X^n_{s-}, V^n_{s-}). \end{aligned} \tag{2.2.41}$$

Summing over $s \in \mathbb{T}_n$, the left-hand side of (2.2.41) gives $F_T(X^n_{T-}, V^n_{T-}) - F_0(X_0, V_0)$, which converges to $F_T(X_T, V_T) - F_0(X_0, V_0)$, by left-continuity of F, continuity of X, and the predictability property of F with respect to the second argument. The first term on the right-hand side of (2.2.41) sums up to

$$\sum_{s \in \mathbb{T}_n} F_{s'}(X^n_{s'-}, V^n_{s'-}) - F_s(X^n_s, V^n_s) = \int_0^T \mathcal{D}_u F\left(X^n_{s_n(u),u-s_n(u)}, V^n_{s_n(u),u-s_n(u)}\right) \mathrm{d}u, \qquad (2.2.42)$$

by the definition of the horizontal derivative. Here, $s_n(u) := s_n$ such that $u \in [s, s') = [s_n, s'_n)$. By the dominated convergence theorem, (2.2.42) converges to

$$\int_0^T \mathcal{D}_u F\left(X_u, V_{u-}\right) \mathrm{d}u = \int_0^T \mathcal{D}_s F\left(X_s, V_s\right) \mathrm{d}s, \qquad (2.2.43)$$

since $V_s = V_{s-}$ $\mathrm{d}s$-almost everywhere. The second term on the right-hand side of (2.2.41) sums up to

$$\begin{aligned} &\sum_{s \in \mathbb{T}_n} \nabla_x F_s\left(X^n_{s-}, V^n_{s-}\right) \delta X^n_s \\ &+ \frac{1}{2} \sum_{i,j=1}^d \sum_{s \in \mathbb{T}_n} \nabla^2_{x,ij} F_s\left(X^n_{s-}, V^n_{s-}\right) \delta X^n_{i,s} \delta X^n_{j,s} + \sum_{s \in \mathbb{T}_n} r^n_s, \end{aligned} \qquad (2.2.44)$$

by a standard second-order Taylor expansion. Here, $\delta X^n_s := X(s') - X(s)$ and the error terms r^n_s are bounded in absolute value by

$$K|\delta X^n_s|^2 \sup_{x \in \mathcal{B}(X(s))} \left\| \nabla^2_x F_s\left(X^{n,x-X(s)}_{s-}, V^n_{s-}\right) - \nabla^2_x F_s\left(X^n_{s-}, V^n_{s-}\right) \right\|. \qquad (2.2.45)$$

It can then be shown (see [20] for details) that the second term in (2.2.45) converges to

$$\frac{1}{2} \sum_{i,j=1}^d \int_0^T \nabla^2_{x,ij} F_s(X_s, V_{s-}) \,\mathrm{d}[X]_{ij}(s) = \frac{1}{2} \sum_{i,j=1}^d \int_0^T \nabla^2_{x,ij} F_s(X_s, V_s) \,\mathrm{d}[X]_{ij}(s),$$

since $\nabla^2 F$ is predictable in the second argument. Similarly, the remainder term $\sum_{s \in \mathbb{T}_n} r^n_s$ converges to 0. Since all terms considered converge, the limit

$$\lim_{n \to \infty} \sum_{s \in \mathbb{T}_n} \nabla_x F_s(X^n_{s-}, V^n_{s-}) \left(X(s') - X(s)\right)$$

exists, and the theorem follows. □

Thus, both approaches to proving the functional Itô formula appear to use the same basic idea. They both consider a decomposition of a (small) change in the functional value into two terms, where the first one converges to the horizontal derivative term, by the fundamental theorem of calculus, and the second one to the vertical derivatives terms, by a second-order Taylor expansion. Nevertheless, there is a certain difference in the sense that the framework used in [20], and thus their proof of the change of variables formula, is completely probability-free and only relies on real analysis techniques. The approach in [20] also appears to be somewhat more intuitive, in the sense that we do not have to deal with window processes, ucp convergence, etc. Moreover, in [20] the functional F is allowed to depend on an additional argument, V, which may allow for more flexibility compared to the setting in [22]. Note, however, that the role of V is rather limited, due to the restrictive nature of the regularity requirements on the horizontal derivative (2.2.19). There is also another technical difference. Since in [20] the horizontal derivative is defined via a limit on the *right*, instead of on the left, this difference leads to an additional regularity condition in order to prove the functional Itô formula (2.2.29); see [22, Remark 2.6] for more details.

In view of many applications in mathematical finance, one would wish to be able to deal with non-anticipative functionals F that depend on additional arguments such as the running maximum $\max_{u\leq t} X(u)$ or the quadratic variation $[X]$ of a trajectory X. These quantities, however, may not be absolutely continuous with respect to t. Thus, functionals depending on such quantities are not covered by the Itô formula from [20] (nor from [22], of course). In the next step we will therefore extend their change of variables formulae to functionals F that are allowed to depend on an additional variable A that corresponds to a general path of bounded variation. We will also extend the notions of the horizontal and vertical derivatives to functionals of this type. This extension will be very natural and convenient for the proof of a pathwise functional associativity rule.

Chapter 3

The associativity rule in pathwise functional Itô calculus

The present chapter is based on [82]. Let ξ be a locally admissible integrand for a continuous path X that admits the continuous quadratic variation in the sense of Föllmer [46], and let η be a locally admissible integrand for the path $Y(t) := \int_0^t \xi(s)\,\mathrm{d}X(s)$. The main result in this chapter states that then $\eta\xi$ is again a locally admissible integrand for X, and

$$\int_0^T \eta(s)\,\mathrm{d}Y(s) = \int_0^T \eta(s)\xi(s)\,\mathrm{d}X(s).$$

This intuitive cancellation rule is often called the *associativity* of the integral. As pointed out in the introduction, in our present strictly pathwise setting we have no analogue of the Kunita–Watanabe characterization of the stochastic integral. In particular, the proof of the associativity property becomes surprisingly involved, as we are only allowed to use analytical tools.

The chapter is organized as follows. First, we provide a slight extension of the functional change of variables formula from [20], namely a change of variables formula for non-anticipative functionals depending on an additional bounded variation component. With this at hand, we can state and show in Section 3.3 the associativity of the pathwise Itô integral $\int \xi(s)\,\mathrm{d}X(s)$. Our applications to pathwise Itô differential equations are given in Section 3.4.

3.1 Preliminaries

In the following we will first describe our framework, slightly extending the definitions and notations introduced in [20] and [34]. In the second step, we will derive our slightly extended functional change of variables formula for a continuous path X. However, as the definition of functional derivatives requires us to apply discontinuous shocks even in case X is continuous, we still have to consider functionals defined on the space of càdlàg paths.

For the sake of simplicity, we keep our notation as close as possible to the one in [20]. More precisely, let $T > 0$ and $D \subset \mathbb{R}^n$ be an arbitrary subset of $\mathbb{R}^n$. As above, a "D-valued càdlàg function" is a right-continuous function $f : [0,T] \mapsto D$ with left limits such that for each $t \in (0,T]$, $f(t-) \in D$. By $\Delta f(t) := f(t) - f(t-)$ we denote the jump of f at time t. By $f^t = (f(u \wedge t))_{0 \le u \le T}$ we denote the stopped path at t. As in [20], we use curly letters to denote the class of càdlàg functions with values in a certain set. We write $\mathcal{D}^T = D([0,T],D)$ for the space of D-valued càdlàg functions on $[0,T]$, and $\mathcal{D}^t \subset \mathcal{D}^T$ for the space of D-valued càdlàg paths stopped at time t. Analogously, $\mathcal{D}_I = D(I,D)$ denotes the space of D-valued càdlàg functions on a subinterval $I \subset [0,T]$, and $\mathcal{D}_I^t \subset \mathcal{D}_I$ is the set of D-valued càdlàg paths on I stopped at time t. By $C(I,D)$ we denote the set of continuous functions on I with values in D. Note that we work with *stopped* paths instead of restrictions as in [20]. This is simpler and clearly equivalent to working with restricted paths (see also [21]). In the sequel, we let $U \subset \mathbb{R}^d$ be an open subset of $\mathbb{R}^d$ and $W \subset \mathbb{R}^m$ a Borel subset of $\mathbb{R}^m$.

Definition 3.1.1. A family $F : [0,T] \times \mathcal{U}^T \times \mathcal{W}^T \mapsto \mathbb{R}$ of functionals is said to be *non-anticipative* if

$$\forall (t,X,A) \in [0,T] \times \mathcal{U}^T \times \mathcal{W}^T, \quad F(t,X,A) = F(t,X^t,A^t). \tag{3.1.1}$$

Here, $\mathcal{U}^T$ and $\mathcal{W}^T$ are special cases of the previously introduced generic notation $\mathcal{D}^T$. In contrast to [20], we will require that the function A in (3.1.1) has components of bounded variation. The reader is referred to the closing paragraph in Chapter 2 for further details on the motivation for the introduction of A. If I is a time interval, the class of right-continuous functions $A : I \mapsto \mathbb{R}^m$ whose components are of bounded variation will be denoted by $BV^m(I)$. The subset of continuos functions in $BV^m(I)$ will be denoted by $CBV^m(I)$. We will also use the notation $\mathcal{D}_{I,BV}^t := \mathcal{D}_I^t \cap BV^m(I)$ and $\mathcal{D}_{I,CBV}^t := \mathcal{D}_I^t \cap CBV^m(I)$, and in particular the generic notation for the specific choices of U and W.

For the definition of functional derivatives, we need to introduce the following notion, which is in analogy to Definition 2.2.12.

Definition 3.1.2 (Perturbation on path spaces)**.** Let $X \in D([0,T],U)$ and X^t be the stopped path at t. For $h \in \mathbb{R}^d$ sufficiently small, the *vertical perturbation* $X^{t,h}$ of the stopped path X^t is defined as the càdlàg path obtained by shifting the value at t by the quantity h:

$$X^{t,h}(u) = X^t(u), \quad u \in [0,t), \quad X^{t,h}(u) = X(t) + h, \quad u \in [t,T], \tag{3.1.2}$$

or, equivalently, $X^{t,h}(u) = X^t(u) + h\mathbb{1}_{[t,T]}(u)$.

Since we work with stopped instead of restricted paths, we can use the standard supremum norm on path space:

$$d_\infty((X,A),(X',A')) = \sup_{u \in [0,T]} |X(u) - X'(u)| + \sup_{u \in [0,T]} |A(u) - A'(u)| \tag{3.1.3}$$

for $(X,A),\ (X',A') \in \mathcal{U}^T \times \mathcal{W}^T$.

Definition 3.1.3 (Regularity properties)**.** (i) A non-anticipative functional F is said to be *left-continuous* (notation: $F \in \mathbb{F}_l^\infty$) if

$$\forall t \in [0,T],\ \forall \epsilon > 0,\ \forall (X,A) \in \mathcal{U}^t \times \mathcal{W}_{BV}^t,$$

$$\exists \lambda > 0 \text{ such that } \forall h \in [0,t],\ \forall (X',A') \in \mathcal{U}^{t-h} \times \mathcal{W}_{BV}^{t-h},$$

$$d_\infty\left((X,A),(X',A')\right) + h < \lambda \Rightarrow |F(t,X,A) - F((t-h),X',A')| < \epsilon. \tag{3.1.4}$$

(ii) A non-anticipative functional F is said to be *boundedness-preserving* (notation: $F \in \mathbb{B}$) if for any compact subset $K \subset U$ there exists a constant C_K such that

$$\forall t \in [0,T],\ \forall (X,A) \in \mathcal{K}^t \times \mathcal{W}_{BV}^t, \quad |F(t,X,A)| < C_K. \tag{3.1.5}$$

Note that if $F \in \mathbb{B}$, then it is *locally bounded* in the neighborhood of any given path. That is,

$$\forall (X,A) \in \mathcal{U}^T \times \mathcal{W}_{BV}^T,\ \exists C > 0,\ \lambda > 0 \text{ such that}$$

$$\forall t \in [0,T],\ \forall (X',A') \in \mathcal{U}^t \times \mathcal{W}_{BV}^t,$$

$$d_\infty\big((X^t,A^t),(X',A')\big) < \lambda \quad \Rightarrow \quad \forall t \in [0,T], \quad |F(t,X',A')| \le C. \tag{3.1.6}$$

We next introduce our notion of *horizontal derivative* (with respect to some measure), which is motivated by the desire to lessen smoothness assumptions on those functionals, for which a change of variables formula can be derived. This generalizes the functional fundamental theorem of calculus for paths with bounded variation from Theorem 2.1.10 and extends the horizontal derivative from [20] and [22].

Definition 3.1.4 (Horizontal derivative)**.** Let F be a non-anticipative functional and $(X,A) \in \mathcal{U}^T \times \mathcal{W}_{BV}^T$. Since the components of A are functions of bounded variation and, hence, correspond to finite measures ν_k, $k = 1,\dots,m$, on $[0,T]$, we can introduce the vector-valued measure ν on $[0,T]$ via $\nu(\mathrm{d}s) := (\mathrm{d}s, A_1(\mathrm{d}s),\dots,A_m(\mathrm{d}s))^\top$. The *horizontal derivative* of F at (t,X^t,A^t) (with respect to ν) is defined as the vector

$$\mathcal{D}F(t,X^t,A^t) = \big(\mathcal{D}_0F(t,X^t,A^t), \mathcal{D}_1F(t,X^t,A^t),\dots,\mathcal{D}_mF(t,X^t,A^t)\big)^\top,$$

where

$$\mathcal{D}_0F(t,X^t,A^t) := \lim_{h\to 0^+} \frac{F(t,X^{t-h},A^{t-h}) - F(t-h,X^{t-h},A^{t-h})}{h} \tag{3.1.7}$$

$$\mathcal{D}_kF(t,X^t,A^t) := \lim_{h\to 0^+} \frac{F(t,X^{t-h},A_1^{t-h},\dots,A_k^t,\dots,A_m^{t-h}) - F(t,X^{t-h},A^{t-h})}{\nu_k((t-h,t])} \tag{3.1.8}$$

for $k = 1,\ \dots,m$, if the corresponding limits exist. In addition, it will be convenient to set $\mathcal{D}F(t,X^t,A^t) = 0$ for $t = 0$.

If (3.1.7) and (3.1.8) are well-defined for all (X, A), then the map

$$\begin{aligned} \mathcal{D}F : [0,T] \times \mathcal{U}^T \times \mathcal{W}_{BV}^T &\mapsto \mathbb{R}^{m+1} \\ (t, X, A) &\to \mathcal{D}F(t, X^t, A^t) \end{aligned} \tag{3.1.9}$$

defines a non-anticipative vector-valued functional $\mathcal{D}F : [0,T] \times \mathcal{U}^T \times \mathcal{W}_{BV}^T \mapsto \mathbb{R}^{m+1}$, the extended *horizontal derivative* of F.

Note that in contrast to the definition of the horizontal derivative in [20] as time derivative, the definition (3.1.7) is based on a limit on the left, and thus only the past evolution of the underlying path is taken into account while no assumptions on the possible future values are made whatsoever. This modification is inspired by [22, Definition 2.9] (see also Definition 2.2.10); for more details on this subject the reader is referred to the closing paragraph in Subsection 2.2.2.

Denote $(e_i,\ i = 1, \ldots, d)$ the canonical basis in $\mathbb{R}^d$.

Definition 3.1.5 (Vertical derivative)**.** A non-anticipative functional F is said to be *vertically differentiable with respect to* X at $(X^t, A^t) \in \mathcal{U}^t \times \mathcal{W}_{BV}^t$ if the map $\mathbb{R}^d \ni e \to F(t, X^{t,e}, A^t)$ is differentiable at 0. Its gradient at 0,

$$\begin{aligned} \nabla_X F(t, X^t, A^t) &= \left(\partial_i F(t, X^t, A^t)\right)_{i=1,\ldots,d}, \text{ where} \\ \partial_i F(t, X^t, A^t) &= \lim_{h \to 0} \frac{F(t, X^{t,he_i}, A^t) - F(t, X^t, A^t)}{h}, \end{aligned} \tag{3.1.10}$$

is called the *vertical derivative* of F at (t, X^t, A^t), with respect to X. If (3.1.10) is well-defined for all (X, A), the map

$$\begin{aligned} \nabla_X F : [0,T] \times \mathcal{U}^T \times \mathcal{W}_{BV}^T &\mapsto \mathbb{R}^d \\ (t, X, A) &\to \nabla_X F(t, X^t, A^t) \end{aligned} \tag{3.1.11}$$

defines a non-anticipative functional $\nabla_X F$ with values in $\mathbb{R}^d$, which we call the *vertical derivative of F with respect to X.*

Example 3.1.6. Let $X \in \mathcal{U}^T,\ A \in \mathcal{W}_{BV}^T$ in the following. If $F(t, X^t, A^t) = f(X(t), A(t))$ for a smooth function $f(x, a)$, then $\mathcal{D}F(t, X^t, A^t) = \nabla_a f(X(t), A(t))$ and $\nabla_X F(t, X^t, A^t) = \nabla_x f(X(t), A(t))$, where $\nabla_a f(X(t), A(t))$ is the gradient of the function $f(x, a)$ with respect to the first argument, evaluated at $(X(t), A(t))$, and, analogously, $\nabla_x f(X(t), A(t))$ is the gradient of the function $f(x, a)$ with respect to the second argument, evaluated at $(X(t), A(t))$. For $X(t) \neq 0$, consider the special case of the geometric mean $F(t, X^t, A^t) = \prod_{i=1}^d \left(\widetilde{X}_i(t)\right)^{\frac{1}{d}}$ of the convex combinations

$$\widetilde{X}_i(t) := \alpha X_i(t) + (1 - \alpha) A_i(t), \quad \alpha \in (0, 1), \tag{3.1.12}$$

where $A_i(t) := \int\limits_0^t X_i(s)\,\mathrm{d}s,\ i = 1, \dots d$. We have,

$$\mathcal{D}F(t, X^t, A^t) = \left(\mathcal{D}_1 F(t, X^t, A^t), \dots, \mathcal{D}_d F(t, X^t, A^t)\right)^\top,$$

with

$$\begin{aligned}\mathcal{D}_i F(t, X^t, A^t) &= \frac{1-\alpha}{d} \left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i}}^{d} \left(\widetilde{X}_k(t)\right)^{\frac{1}{d}} \\ &= \frac{1-\alpha}{d} \cdot \frac{F(t, X^t, A^t)}{\widetilde{X}_i(t)},\end{aligned}$$

and

$$\partial_i F(t, X^t, A^t) = \frac{\alpha}{d} \left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i}}^{d} \left(\widetilde{X}_k(t)\right)^{\frac{1}{d}} = \frac{\alpha}{d} \cdot \frac{F(t, X^t, A^t)}{\widetilde{X}_i(t)},$$

for $i = 1, \dots, d$. This will be useful in Example 4.3.1 below.

Example 3.1.7. Sometimes, a quantity of interest can either be considered as a path-dependent functional of $X \in \mathcal{U}^T$ only or as an additional trajectory of bounded variation. The latter possibility allows us to include functionals that may not be regular enough for the setting of [20] or [34]. This illustrates one advantage of our extended approach.

(i) Consider the time average of the first component, $F(t, X^t) = \int_0^t X_1(s)\,\mathrm{d}s$. Alternatively, it can be represented as $A_X(t)$, where $A_X(t) := \int_0^t X_1(s)\,\mathrm{d}s$, because the time average is a function of bounded variation. In the first approach, we have $\mathcal{D}F(t, X^t) = X_1(t-)$ and $\partial_X F(t, X^t) = 0$. In the second approach, we have $\mathcal{D}A_X(t) = f'(A_X(t)) = 1$ with $f(a) = a$. Clearly, $\partial_X A_X(t) = 0$.

(ii) Consider the functional $F(t, X^t) = [X_1](t)$, where $[X_1]$ is the pathwise quadratic variation of the first component in the sense of Föllmer [46] (see Definition 2.2.5). Alternatively, it can be represented as $A_X(t) := [X_1](t)$. In the first approach, we have $\mathcal{D}F(t, X^t) = 0$ and $\partial_X F(t, X^t) = 2(X_1(t) - X_1(t-))$. In the second approach, we have $\mathcal{D}A_X(t) = 1$ and $\partial_X A_X(t) = 0$.

(iii) Consider the running maximum of the first component $F(t, X^t) = \max\limits_{0\le s\le t} X_1(s)$. Alternatively, it can be represented as $A_X(t) := \max\limits_{0\le s\le t} X_1(s)$, since the running maximum is a function of bounded variation. Then, F is not (vertically) differentiable in the first approach, and we would have to resort to smoothing techniques [34]. In the second approach, however, the horizontal derivative with respect to the measure corresponding to A_X does exist, and we have $\mathcal{D}A_X(t) = 1$.

If the functional F admits horizontal and vertical derivatives $\mathcal{D}F$ and $\nabla_X F$, we may iterate the above operations in order to define higher order horizontal and vertical derivatives.

Definition 3.1.8. Let $I \subset [0,T]$ be a subinterval of $[0,T]$ with nonempty interior, $\mathring{I}$. We denote by $\mathbb{C}^{j,k}(I)$ the set of all non-anticipative functionals F on $\bigcup_{t\in I} \mathcal{U}_I^t \times \mathcal{W}_{I,BV}^t$ such that:

- F is continuous at fixed times t, locally uniformly in t. That is,

$$\begin{gathered}\forall t \in I \subset [0,T],\ \forall \epsilon > 0,\ \forall (X,A) \in \mathcal{U}_I^t \times \mathcal{W}_{I,BV}^t,\\ \exists \lambda > 0 \text{ such that } \forall (X',A') \in \mathcal{U}_I^{t'} \times \mathcal{W}_{I,BV}^{t'},\\ d_\infty\left((X,A),(X',A')\right) + |t-t'| < \lambda \Rightarrow |F(t',X,A) - F(t',X',A')| < \epsilon. \end{gathered} \tag{3.1.13}$$

- F admits j horizontal derivatives and k vertical derivatives with respect to X at all $(X,A) \in \mathcal{U}_I^t \times \mathcal{W}_{I,BV}^t$, $t \in I$.

- $\mathcal{D}^l F$, $l \le j$, $\nabla_X^n F$, $n \le k$, are left-continuous on I.

3.2 Functional change of variables formula

To prove our associativity result we will need the following change of variables formula, which is a slight extension of the corresponding formulae from [20] and [34]. From now on we fix a refining sequence of partitions $(\mathbb{T}_n)$ of $[0,T]$, along which X admits the continuous quadratic variation in the sense of Definition 2.2.5.

Theorem 3.2.1. *Let $(X,A) \in QV^d \cap \mathcal{U}^T \times \mathcal{W}_{CBV}^T$ and denote*

$$X^n(t) := \sum_{s\in\mathbb{T}_n} X(s')\mathbb{1}_{[s,s')}(t) + X(T)\mathbb{1}_{\{T\}}(t), \quad 0 \le t \le T, \tag{3.2.1}$$

$$A^n(t) := \sum_{s\in\mathbb{T}_n} A(s)\mathbb{1}_{[s,s')}(t) + A(T)\mathbb{1}_{\{T\}}(t), \quad 0 \le t \le T. \tag{3.2.2}$$

$$h_s^n := s' - s,\ s, s' \in \mathbb{T}_n. \tag{3.2.3}$$

Suppose moreover that F is a left-continuous non-anticipative functional of class $\mathbb{C}^{1,2}([0,T])$ such that $\mathcal{D}F$, $\nabla_X F$, $\nabla_X^2 F \in \mathbb{B}$. Denote $X^{n,s-}$ the n-th approximation of X stopped at time $s-$. Then the pathwise Itô integral along $(\mathbb{T}_n)$, defined as

$$\int_0^T \nabla_X F(s,X^s,A^s)\,\mathrm{d}X(s) := \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n} \nabla_X F(s,X^{n,s-},A^{n,s}) \cdot (X(s') - X(s)), \tag{3.2.4}$$

exists and

$$\begin{aligned}F(T,X^T,A^T) - F(0,X^0,A^0) = \int_0^T \mathcal{D}F(s,X^s,A^s)\nu(\,\mathrm{d}s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^T \nabla_{X,ij}^2 F(s,X^s,A^s)\,\mathrm{d}[X]_{ij}(s)\\ + \int_0^T \nabla_X F(s,X^s,A^s)\,\mathrm{d}X(s).\end{aligned} \tag{3.2.5}$$

Proof. The proof uses the same Föllmer-type discretization techniques as are used in the proof of [20, Theorem 3]. Denote $\delta X^n_s = X(s') - X(s)$ for $s, s' \in \mathbb{T}_n$ and $|\mathbb{T}_n| := \sup_{s\in\mathbb{T}_n} |s' - s|$ the mesh of $\mathbb{T}_n$. Since X and A are continuous, and hence uniformly continuous on $[0, T]$, it follows that

$$\lambda_n := \sup_{\alpha,\beta\in[0,T],|\alpha-\beta|\leq|\mathbb{T}_n|} \Big(|A(\alpha) - A(\beta)| + |X(\alpha) - X(\beta)|\Big) \to 0 \text{ as } n \to \infty. \tag{3.2.6}$$

Since $\nabla^2_X F,\ \mathcal{D}F \in \mathbb{B}$, for n large enough there exists a constant $C > 0$ such that

$$\forall t < T,\ \forall (X', A') \in \mathcal{U}^t \times \mathcal{W}^t_{BV},$$

$$d_\infty \big((X^t, A^t), (X', A')\big) < \lambda_n \quad \Rightarrow \quad |\mathcal{D}F(t, X', A')| \leq C,\ \|\nabla^2_X F(t, X', A')\| \leq C.$$

For $s \in \mathbb{T}_n$, consider the following decomposition of increments into "horizontal" and "vertical" terms:

$$\begin{aligned} F(s', X^{n,s'-}, A^{n,s'}) - F(s, X^{n,s-}, A^{n,s}) &= F(s', X^{n,s'-}, A^{n,s'}) - F(s, X^{n,s}, A^{n,s}) \\ &\quad + F(s, X^{n,s}, A^{n,s}) - F(s, X^{n,s-}, A^{n,s}). \end{aligned} \tag{3.2.7}$$

The first term on the right-hand side of (3.2.7) is equal to $\psi(h^n_s) - \psi(0)$, where $\psi(u) := F(s + u, X^{n,s}, A^{n,s+u})$. Since F admits a horizontal derivative in the sense of (3.1.7), (3.1.8), we can write

$$\begin{aligned} F(s', X^{n,s'-}, A^{n,s'}) - F(s, X^{n,s}, A^{n,s}) &= \int_{(0,s'-s]} \mathcal{D}F(s + u, X^{n,s}, A^{n,s+u-})\nu_n(\,\mathrm{d}u) \\ &= \int_{(s,s']} \mathcal{D}F\left(u, X^{n,s}, A^{n,u-}\right) \nu_n(\,\mathrm{d}u), \end{aligned} \tag{3.2.8}$$

where ν_n is the measure corresponding to the n-th approximation A^n of A, i.e.,

$$\nu_n(\,\mathrm{d}s) = \left(\,\mathrm{d}s, A_{n,1}(\,\mathrm{d}s), \dots, A_{n,m}(\,\mathrm{d}s)\right)^\top.$$

For the second term on the right-hand side of (3.2.7), we have

$$F(s, X^{n,s}, A^{n,s}) - F(s, X^{n,s-}, A^{n,s}) = \phi(\delta X^n_s) - \phi(0), \tag{3.2.9}$$

where $\phi(u) = F(s, X^{n,s-,u}, A^{n,s})$. Since $F \in \mathbb{C}^{1,2}([0, T])$, the function ϕ is well-defined and twice continuously differentiable in the neighborhood $\mathcal{B}(X(s), \lambda_n) \subset U$, with

$$\nabla\phi(u) = \nabla_X F(s, X^{n,s-,u}, A^{n,s}), \tag{3.2.10}$$

$$\nabla^2\phi(u) = \nabla^2_X F(s, X^{n,s-,u}, A^{n,s}). \tag{3.2.11}$$

Hence, a second-order Taylor expansion of ϕ at $u = 0$ yields

$$\begin{aligned} F(s, X^{n,s}, A^{n,s}) - F(s, X^{n,s-}, A^{n,s}) &= \nabla_X F(s, X^{n,s-}, A^{n,s})\delta X^n_s \\ &\quad + \frac{1}{2}\sum_{i,j=1}^{d} \nabla^2_{X,ij} F(s, X^{n,s-}, A^{n,s})\delta X^n_{i,s}\delta X^n_{j,s} + r^n_s, \end{aligned} \tag{3.2.12}$$

where for some $\theta \in [0,1]$,

$$r_s^n := \frac{1}{2}\sum_{i,j=1}^{d}\left(\nabla^2_{X,ij}F(s, X^{n,s-,\theta\delta X_s^n}, A^{n,s}) - \nabla^2_{X,ij}F(s, X^{n,s-}, A^{n,s})\right)\delta X^n_{i,s}\delta X^n_{j,s}.$$

We now sum over $s \in \mathbb{T}_n$.

- The left-hand side of (3.2.7) sums up to $F(T, X^{n,T-}, A^{n,T}) - F(0, X^0, A^0)$, which converges to $F(T, X^{T-}, A^T) - F(0, X^0, A^0)$, by left-continuity of F. Since X and A are continuous, this is equal to $F(T, X^T, A^T) - F(0, X^0, A^0)$.

- For the first term on the right-hand side of (3.2.7), we have

$$\sum_{s\in\mathbb{T}_n} F(s', X^{n,s'-}A^{n,s'}) - F(s, X^{n,s}, A^{n,s}) = \int_{(0,T]} \mathcal{D}F\left(u, X^{n,s(u)}, A^{n,u-}\right)\nu_n(\,\mathrm{d}u), \quad (3.2.13)$$

  in conjunction with (3.2.8). Here, we set $s(u) := s$ such that $u \in [s, s'),\ s, s' \in \mathbb{T}_n$. The integrand converges to $\mathcal{D}F(u, X^u, A^u)$, is bounded by C, and both are left-continuous in u by [20, Proposition 1]. Moreover, the sequence of finite measures ν_n, corresponding to the approximations A_n of A, converges vaguely to the atomless measure ν, corresponding to A, so we can use a "diagonal lemma" for vague convergence of measures in the form of [20, Lemma 12] to obtain that (3.2.13) converges to

$$\int_{(0,T]} \mathcal{D}F(u, X^u, A^u)\nu(\,\mathrm{d}u) = \int_0^T \mathcal{D}F(s, X^s, A^s)\nu(\,\mathrm{d}s). \quad (3.2.14)$$

- For the second term on the right-hand side of (3.2.7), we have

$$\sum_{s\in\mathbb{T}_n} F(s, X^{n,s}, A^{n,s}) - F(s, X^{n,s-}, A^{n,s}) = \sum_{s\in\mathbb{T}_n} \nabla_X F(s, X^{n,s-}, A^{n,s})\cdot\delta X_s^n$$
$$+ \frac{1}{2}\sum_{i,j=1}^{d}\sum_{s\in\mathbb{T}_n} \nabla^2_{X,ij}F(s, X^{n,s-}, A^{n,s})\delta X^n_{i,s}\delta X^n_{j,s} + \sum_{s\in\mathbb{T}_n} r_s^n. \quad (3.2.15)$$

  The quantity $\nabla^2_X F(s, X^{n,s-}, A^{n,s})\mathbb{1}_{u\in(s,s']}$ is bounded by C, converges to $\nabla^2_X F(u, X^u, A^u)$, by left-continuity of $\nabla^2_X F$, and both are left-continuous in u, by [20, Proposition 1]. Since $X \in QV^d$, we can again apply [20, Lemma 12] to infer that

$$\frac{1}{2}\sum_{i,j=1}^{d}\sum_{s\in\mathbb{T}_n} \nabla^2_{X,ij}F(s, X^{n,s-}, A^{n,s})\delta X^n_{i,s}\delta X^n_{j,s}$$
$$\to \frac{1}{2}\sum_{i,j=1}^{d}\int_0^T \nabla^2_{X,ij}F(s, X^s, A^s)\,\mathrm{d}[X]_{ij}(s). \quad (3.2.16)$$

  Using the same argument, since $|r_s^n|$ is bounded by $\epsilon_s^n|\delta X_s^n|^2$, where ϵ_s^n converges to 0, the remainder term, $\sum_{s\in\mathbb{T}_n} r_s^n$, converges to 0.

Since all terms considered converge, the limit

$$\lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\nabla_X F(s,X^{n,s-},A^{n,s})\cdot(X(s')-X(s))$$

exists, and the theorem follows. □

Lemma 3.2.2. *Suppose that F satisfies the regularity assumptions from Theorem 3.2.1. Then for any $(X,A)\in C([0,T],U)\times\mathcal{W}^T_{CBV}$ the map $t\mapsto F(t,X^t,A^t)$ is continuous.*

Proof. We first show the left-continuity of $t\mapsto F(t,X^t,A^t)$. Since X and A are continuous, and hence uniformly continuous on $[0,T]$, we have that for $h>0$ sufficiently small,

$$d_\infty\big((X^{t-h},A^{t-h}),(X^t,A^t)\big)=\sup_{u\in[t-h,t]}|X(u)-X(t-h)|+\sup_{u\in[t-h,t]}|A(u)-A(t-h)|\to 0 \text{ as } h\to 0^+.$$

Thus, the left-continuity of F implies that $F(t-h,X^{t-h},A^{t-h})-F(t,X^t,A^t)\to 0$ as $h\to 0^+$. Analogously, since X and A are continuous, and hence uniformly continuous on $[0,T]$, we have that for $h>0$ sufficiently small,

$$d_\infty\big((X^{t+h},A^{t+h}),(X^t,A^t)\big)=\sup_{u\in[t,t+h]}|X(u)-X(t)|+\sup_{u\in[t,t+h]}|A(u)-A(t)|\to 0 \text{ as } h\to 0^+.$$

This implies that

$$F(t+h,X^{t+h},A^{t+h})-F(t+h,X^t,A^t)\to 0 \text{ as } h\to 0^+, \tag{3.2.17}$$

due to the continuity at fixed times, which holds locally uniformly. Moreover, since X and A are continuous and the horizontal derivative is boundedness-preserving, we infer that

$$F(t+h,X^t,A^t)-F(t,X^t,A^t)=\int_t^{t+h}\mathcal{D}_0F(u,X^u,A^u)\,\mathrm{d}u\to 0 \text{ as } h\to 0^+. \tag{3.2.18}$$

Using (3.2.17) and (3.2.18) yields

$$\begin{aligned}F(t+h,X^{t+h},A^{t+h})-F(t,X^t,A^t)&=F(t+h,X^{t+h},A^{t+h})-F(t+h,X^t,A^t)\\&\quad+\ F(t+h,X^t,A^t)-F(t,X^t,A^t)\to 0 \text{ as } h\to 0^+,\end{aligned}$$

which shows the right-continuity. □

Since the first and second integrals in (3.2.5) are Riemann-Stieltjes integrals, and hence continuous as functions in t, using Lemma 3.2.2 gives the following result.

Corollary 3.2.3. *The pathwise Itô integral defined in* (3.2.4)*, as a function in t, is continuous.*

The limit in (3.2.4) depends on the choice of the refining sequence of partitions $(\mathbb{T}_n)$ (see [20, Remark 5]), but, as opposed to [20], we do not make this dependence explicit in the notation to keep things simple. Also note that it may be tempting to write

$$\int_0^T\nabla_x F(s,X^s,A^s)\,\mathrm{d}X(s)=\sum_{i=1}^d\int_0^T\partial_i F(s,X^s,A^s)\,\mathrm{d}X_i(s),$$

but it is not obvious whether the integrals on the right-hand side of the last equation exist individually as pathwise Itô integrals, i.e., as the limits

$$\int_0^T \partial_i F(s, X^s, A^s) \, \mathrm{d}X^i(s) = \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n} \partial_i F(s, X^{n,s-}, A^{n,s}) \left(X_i(s') - X_i(s)\right).$$

Theorem 3.2.1 implies in particular that the pathwise Itô integral is well-defined for integrands that are vertical derivatives of non-anticipative $\mathbb{C}^{1,2}$-functionals satisfying the regularity conditions from Theorem 3.2.1. This allows us to formalize the notion of a locally admissible integrand, as follows.

Definition 3.2.4. Let $X \in QV^d \cap \mathcal{U}^T$. A function $t \mapsto \xi(t) \in \mathbb{R}^d$ is called a *locally admissible integrand for* X if there exists a partition $\mathbb{T} = \{t_0, \ldots, t_N\}$ of $[0, T]$ such that for all $k = 1, \ldots, N$ there are $m_k \in \mathbb{N}$, $A_k \in \mathcal{W}_{[t_{k-1},t_k],CBV}$, where $W \subset \mathbb{R}^{m_k}$, and F_k as in Theorem 3.2.1 so that

$$\xi(t) = \sum_{k=1}^N \nabla_X F_k(t, X_k^t, A_k^t) \mathbb{1}_{t\in[t_{k-1},t_k)}.$$

Here, $X_k := X\big|_{[t_{k-1},t_k]}$ is the restriction of X to $[t_{k-1}, t_k]$, and we require

$$F_k(t_k, X_k^{t_k}, A_k^{t_k}) = F_{k+1}(t_k, X_k^{t_k}, A_k^{t_k}), \quad k = 1, \ldots, N-1. \tag{3.2.19}$$

Remark 3.2.5. Let us point out that the above definition allows for a large class of integrands, which include generalized Delta hedging strategies for many exotic and plain-vanilla options in complete market models such as local volatility models that are relevant for practical applications. It thus extends [80, Remark 4] to the functional setting. Moreover, A in the above definition can for instance be a continuous function of the running average of X, $t \mapsto \int_{(t-\delta)+}^t X(s)\,\mathrm{d}s$, or its running maximum, $t \mapsto \max_{(t-\delta)+\le s\le t} X(s)$, since these functions are of bounded variation on $[0, T]$.

Remark 3.2.6. Also note that the above definition of a locally admissible integrand can be equivalently written as follows: A function $t \to \xi(t) \in \mathbb{R}^d$ is called a *locally admissible integrand for* $X \in QV^d \cap \mathcal{U}^T$ if for every $t \in [0, T]$ there exists $\epsilon > 0$ such that there are $m_\epsilon \in \mathbb{N}$, a continuous function A_ϵ of bounded variation on $[t-\epsilon, t+\epsilon]$, and F_ϵ as in Theorem 3.2.1 such that

$$\xi(t) = \nabla_X F_\epsilon(t, X_\epsilon^t, A_\epsilon^t) \mathbb{1}_{t\in[t-\epsilon,t+\epsilon)}.$$

Here, $X_\epsilon := X\big|_{[t-\epsilon,t+\epsilon]}$ is the restriction of X to $[t-\epsilon, t+\epsilon]$.

The following result extends the covariation formula for the pathwise (classical) Itô integral from [86, 80]. Note that in the preprint [3] a pathwise isometry formula is also derived, but our proof is shorter and works under slightly less restrictive assumptions.

Proposition 3.2.7. *Suppose that $\xi = (\xi_1, \dots, \xi_d)$ and $\eta = (\eta_1, \dots, \eta_d)$ are two locally admissible integrands for $X \in QV^d$. Then the pathwise Itô integrals $\int_0^t \xi(s)\,\mathrm{d}X(s)$ and $\int_0^t \eta(s)\,\mathrm{d}X(s)$ admit the continuous covariation*

$$\Big[\int_0^\cdot \xi(s)\,\mathrm{d}X(s), \int_0^\cdot \eta(s)\,\mathrm{d}X(s)\Big](t) = \sum_{i,j=1}^d \int_0^t \xi_i(s)\eta_j(s)\,\mathrm{d}[X_i, X_j](s).$$

Proof. First step: Let F be a left-continuous non-anticipative functional of class $\mathbb{C}^{1,1}([0,T])$ such that $\mathcal{D}F,\ \nabla_X F \in \mathbb{B}$. We will show that $Y(t) := F(t, X^t, A^t)$ has the continuous quadratic variation

$$[Y](t) = \sum_{i,j=1}^d \int_0^t \partial_i F(s, X^s, A^s)\partial_j F(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s).$$

First, observe that we can approximate the increments of F for the path (X, A) by the respective increments along the piecewise constant approximation (X^n, A^n) from (3.2.1), (3.2.2), due to the left-continuity of F. As above, denote $\delta X_s^n := X(s') - X(s),\ s, s' \in \mathbb{T}_n$, and $s(u) := s$ such that $u \in [s, s')$. Then, a first-order Taylor expansion analogous to the one from the proof of Theorem 3.2.1 yields for the increments along the piecewise constant approximation (X^n, A^n):

$$\begin{aligned} F(s', X^{n,s'-}, A^{n,s'}) - F(s, X^{n,s-}, A^{n,s}) &= \int_{(s,s']} \mathcal{D}F(u, X^{n,s}, A^{n,u-})\,\mathrm{d}\nu_n(u) \\ &+ \nabla_X F(s, X^{n,s-}, A^{n,s})\delta X_s^n + r_s^n, \end{aligned}$$

where, for some $\theta \in [0,1]$,

$$r_s^n := \big(\nabla_X F(s, X^{n,s-,\theta\delta X_s^n}, A^{n,s}) - \nabla_X F(s, X^{n,s-}, A^{n,s})\big)\,\delta X_s^n,$$

and hence $|r_s^n| \le \epsilon_s^n |\delta X_s^n|$ with $\epsilon_s^n \to 0$. Thus,

$$\begin{aligned} &\sum_{s\in\mathbb{T}_n, s\le t} \Big(F(s', X^{n,s'-}, A^{n,s'}) - F(s, X^{n,s-}, A^{n,s})\Big)^2 \\ &= \sum_{s\in\mathbb{T}_n, s\le t}\sum_{i,j=1}^d \partial_i F(s, X^{n,s-}, A^{n,s})\partial_j F(s, X^{n,s-}, A^{n,s})(X_i(s') - X_i(s))(X_j(s') - X_j(s)) \\ &+ \sum_{s\in\mathbb{T}_n, s\le t}\sum_{l,j=0}^m \int_{(s,s']} \mathcal{D}_l F(u, X^{n,s}, A^{n,u-})\,\mathrm{d}\nu_{n,l}(u) \int_{(s,s']} \mathcal{D}_j F(u, X^{n,s}, A^{n,u-})\,\mathrm{d}\nu_{n,j}(u) \\ &+ 2\sum_{s\in\mathbb{T}_n, s\le t} \int_{(s,s']} \mathcal{D}F(u, X^{n,s}, A^{n,u-})\nu_n(\,\mathrm{d}u)\nabla_X F(s, X^{n,s-}, A^{n,s})\delta X_s^n + \sum_{s\in\mathbb{T}_n, s\le t} R_s^n. \end{aligned}$$

Since all appearing approximations have a d_∞-distance from (X^s, A^s) less than λ_n from (3.2.6), and $\mathcal{D}F,\ \nabla_X F \in \mathbb{B}$, we infer that

$$|R_s^n| \le C\epsilon_s^n \left(|(X(s') - X(s)|^2 + \sum_{i=1}^d\sum_{l=0}^m |(X_i(s') - X_i(s)||\nu_{n,l}(\,\mathrm{d}s)|\right).$$

Moreover, $\mathcal{D}F\left(u, X^{n,s(u)}, A^{n,u-}\right)$ and $\nabla_X F(s, X^{n,s-}, A^{n,s})\mathbb{1}_{u\in(s,s']}$ are bounded, converge to $\mathcal{D}F\left(u, X^{u}, A^{u}\right)$ and $\nabla_X F(u, X^u, A^u)$, respectively, and all paths are left-continuous in u (see proof of Theorem 3.2.1). Thus we can use a "diagonal lemma" for vague convergence of measures in the form of [20, Lemma 12]. This gives us that the first term on the right-hand side of the above equation converges to

$$\sum_{i,j=1}^{d}\int_0^t \partial_i F(s, X^s, A^s)\partial_j F(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s).$$

The second and third terms on the right-hand side of the above equation converge to 0, since ν corresponds to a continuous function whose components are of bounded variation, and hence all appearing covariations vanish (see [80, Remark 8]). The same argument gives that the error term converges to 0, since ϵ_s^n converges to 0.

Second step: Note that we can assume without loss of generality that $\xi = \eta$, by means of polarization. This is because if we define $\Sigma(t) := Y(t) + Z(t)$, where $Z(t) := G(t, X^t, A^t)$ with G as in the first step, then Σ also satisfies the requirements from the first step. Hence, Σ has the following continuous quadratic variation

$$[\Sigma](t) = \sum_{i,j=1}^{d}\int_0^t (\partial_i F(s, X^s, A^s) + \partial_i G(s, X^s, A^s))(\partial_j F(s, X^s, A^s) + \partial_j G(s, X^s, A^s))\,\mathrm{d}[X_i, X_j](s).$$

Thus, by the polarization identity,

$$\begin{aligned}
&[Y, Z](t) = \frac{1}{2}\left([Y+Z](t) - [Y](t) - [Z](t)\right)\\
&= \frac{1}{2}\Bigg(\sum_{i,j=1}^{d}\int_0^t (\partial_i F(s, X^s, A^s) + \partial_i G(s, X^s, A^s))(\partial_j F(s, X^s, A^s) + \partial_j G(s, X^s, A^s))\,\mathrm{d}[X_i, X_j](s)\\
&\quad - \sum_{i,j=1}^{d}\int_0^t \partial_i F(s, X^s, A^s)\partial_j F(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s)\\
&\quad - \sum_{i,j=1}^{d}\int_0^t \partial_i G(s, X^s, A^s)\partial_j G(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s)\Bigg)\\
&= \sum_{i,j=1}^{d}\int_0^t \partial_i F(s, X^s, A^s)\partial_j G(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s).
\end{aligned}$$

Moreover, by concentrating on small time intervals, we can assume without loss of generality that ξ is of the form $\xi(t) = \nabla_X F(t, X^t, A^t)$, $t \in [0, T]$, for A, F as in Definition 3.2.4. The change of variables formula, Theorem 3.2.1, thus implies

$$\begin{aligned}
&F(T, X^T, A^T) - F(0, X^0, A^0) - \int_0^T \mathcal{D}F(s, X^s, A^s)\nu(\,\mathrm{d}s)\\
&\quad - \frac{1}{2}\sum_{i,j=1}^{d}\int_0^T \nabla^2_{X,\,ij} F(s, X^s, A^s)\,\mathrm{d}[X]_{ij}(s) = \int_0^T \nabla_X F(s, X^s, A^s)\,\mathrm{d}X(s).
\end{aligned}$$

We introduce

$$B(t) := -\int_0^t \mathcal{D}F(s, X^s, A^s)\nu(\,\mathrm{d}s) - \frac{1}{2}\sum_{i,j=1}^d \int_0^t \nabla^2_{X,ij}F(s, X^s, A^s)\,\mathrm{d}[X]_{ij}(s),$$

which belongs to the class $CBV([0,T])$, by standard properties of Stieltjes integrals. In particular, we have $[B] \equiv 0$. Moreover, setting $Y(t) := F(t, X^t, A^t) - F(0, X^0, A^0)$, we obtain

$$\Big[\int_0^{\cdot} \nabla_X F(s, X^s, A^s)\,\mathrm{d}X(s)\Big](t) = [Y + B](t) = [Y](t).$$

Applying the first step yields

$$\begin{aligned}\Big[\int_0^{\cdot} \xi(s)\,\mathrm{d}X(s)\Big](t) &= \Big[\int_0^{\cdot} \nabla_X F(s, X^s, A^s)\,\mathrm{d}X(s)\Big](t)\\ &= [Y](t) = \sum_{i,j=1}^d \int_0^t \partial_i F(s, X^s, A^s)\partial_j F(s, X^s, A^s)\,\mathrm{d}[X_i, X_j](s),\end{aligned}$$

which concludes the proof. □

3.3 The associativity rule in pathwise functional Itô calculus

We are now ready to state and show the associativity property of the pathwise functional Itô integral (3.2.4). This extends [80, Theorem 13] to the functional setting. It turns out that associativity is crucial when discussing Itô differential equations in the pathwise setting (see [67, 80]). This fact will be illustrated for functional Itô differential equations in the subsequent section.

Let $X \in QV^d \cap \mathcal{U}^T$ and $\xi_{(1)}, \ldots, \xi_{(\kappa)}$ be locally admissible integrands for X. We define

$$Y_{(\ell)}(t) := \int_0^t \xi_{(\ell)}(s)\,\mathrm{d}X(s), \quad \ell = 1, \ldots, \kappa. \tag{3.3.1}$$

Then, Proposition 3.2.7 implies that the continuous trajectory $Y = \big(Y_{(1)}, \ldots, Y_{(\kappa)}\big)$ admits the continuous covariations

$$[Y_{(k)}, Y_{(\ell)}](t) = \sum_{i,j=1}^d \int_0^t \xi_{(k),i}(s)\xi_{(\ell),j}(s)\,\mathrm{d}[X_i, X_j](s). \tag{3.3.2}$$

Theorem 3.3.1. *Let X, $\xi_{(1)}, \ldots, \xi_{(\kappa)}$ be as above and Y as in (3.3.1). Moreover, let $\eta = (\eta_1, \ldots \eta_\kappa)$ be a locally admissible integrand for Y. Then $\sum_{\ell=1}^\kappa \eta_\ell \xi_{(\ell)}$ is a locally admissible integrand for X and*

$$\int_0^T \eta(s)\,\mathrm{d}Y(s) = \int_0^T \sum_{\ell=1}^\kappa \eta_\ell(s)\xi_{(\ell)}(s)\,\mathrm{d}X(s). \tag{3.3.3}$$

For the proof we need the following auxiliary lemmas. The first one is a product rule for vertical derivatives, the second one a chain rule for both vertical and horizontal derivatives. Both extend statements from [34].

Lemma 3.3.2. *Let F, G be two non-anticipative vertically differentiable functionals such that $F,\ G,\ \nabla_X F,\ \nabla_X G \in \mathbb{F}_l^\infty$ and $F,\ G,\ \nabla_X F,\ \nabla_X G \in \mathbb{B}$. Then the product FG is again a non-anticipative vertically differentiable functional such that $FG,\ \nabla_X(FG) \in \mathbb{F}_l^\infty$ and $FG,\ \nabla_X(FG) \in \mathbb{B}$. Moreover,*

$$\partial_i(FG) = \partial_i F G + F \partial_i G \quad \text{for all } \ i = 1, \dots, d. \tag{3.3.4}$$

Proof. Let $X \in \mathcal{U}^T,\ A \in \mathcal{W}_{BV}^T$ in the following. For $i \in \{1, \dots, d\}$, we have

$$\begin{aligned}
\partial_i(FG)(t, X^t, A^t) &= \lim_{h\to 0} \frac{(FG)(t, X^{t,he_i}, A^t) - (FG)(t, X^t, A^t)}{h} \\
&= \lim_{h\to 0} \frac{G(t, X^{t,he_i}, A^t)(F(t, X^{t,he_i}, A^t) - F(t, X^t, A^t))}{h} \\
&\quad + \lim_{h\to 0} \frac{F(t, X^t, A^t)(G(t, X^{t,he_i}, A^t) - G(t, X^t, A^t))}{h} \\
&= G(t, X^t, A^t)\partial_i F(t, X^t, A^t) + F(t, X^t, A^t)\partial_i F(t, X^t, A^t).
\end{aligned}$$

Hence, FG is vertically differentiable with respect to X, since F and G are, and (3.3.4) holds.

Now we want to show that FG satisfies the required regularity assumptions, i.e., we have to show that $FG, \nabla_X(FG) \in \mathbb{F}_l^\infty$ and that $FG,\ \nabla_X(FG) \in \mathbb{B}$. The product FG and its gradient are boundedness-preserving, since all functionals appearing on the right-hand side of (3.3.4) are, and since the product of two boundedness-preserving functionals is again boundedness-preserving. Indeed, if for any compact subset $K \subset U$ there exist constants $C_{K,1},\ C_{K,2}$ such that

$$\forall t \in [0,T],\ \forall (X,A) \in \mathcal{K}^t \times \mathcal{W}_{BV}^t, \quad |\Lambda(t,X,A)| < C_{K,1},\ |\Psi(t,X,A)| < C_{K,2},$$

it follows that there exists a constant C_K such that

$$\forall t \in [0,T],\ \forall (X,A) \in \mathcal{K}^t \times \mathcal{W}_{BV}^t, \quad |(\Lambda\Psi)(t,X,A)| = |\Lambda(t,X,A)\Psi(t,X,A)| < C_K.$$

Moreover, all functionals appearing on the right-hand side of (3.3.4) are left-continuous, by our assumptions. Thus, both FG and its gradient, $\nabla_X(FG)$, are left-continuous, since it is easily checked that the product of two (locally bounded) left-continuous functionals is again left-continuous. More precisely, if

$$\forall t \in [0,T],\ \forall \epsilon_1,\ \epsilon_2 > 0,\ \forall (X,A) \in \mathcal{U}^t \times \mathcal{W}_{BV}^t, \exists \lambda_1,\ \lambda_2 > 0 \text{ such that}$$

$$\forall h \in [0,t],\ \forall (X',A') \in \mathcal{U}^{t-h} \times \mathcal{W}_{BV}^{t-h},\ d_\infty\left((X,A),(X',A')\right) < \lambda_1 \wedge \lambda_2 \Rightarrow$$

$$\begin{aligned}
&|\Lambda(t,X,A) - \Lambda((t-h),X',A')| < \epsilon_1, \\
&|\Psi(t,X,A) - \Psi((t-h),X',A')| < \epsilon_2,
\end{aligned}$$

it follows that

$$\forall t \in [0,T],\ \forall \epsilon > 0,\ \forall (X,A) \in \mathcal{U}^t \times \mathcal{W}^t_{BV}, \exists \lambda > 0 \text{ such that}$$

$$\forall h \in [0,t],\ \forall (X',A') \in \mathcal{U}^{t-h} \times \mathcal{W}^{t-h}_{BV},\ d_\infty\left((X,A),(X',A')\right) < \lambda \Rightarrow$$

$$\begin{aligned}|(\Lambda\Psi)(t,X,A) - (\Lambda\Psi)((t-h),X',A')| &\leq |(\Psi(t,X,A)(\Lambda(t,X,A) - \Lambda((t-h),X',A'))| \\ &+ |\Lambda((t-h),X',A')(\Psi(t,X,A) - \Psi((t-h),X',A'))| < \epsilon.\end{aligned}$$

This concludes the proof. □

Lemma 3.3.3. *Let* $H : [0,T] \times \mathcal{V}^T \times \mathcal{W}^T_{BV} \mapsto \mathbb{R}$*, where* $V \subset \mathbb{R}^\kappa$ *open,* $W \subset \mathbb{R}^n$ *Borel, and* $\widetilde{F} : [0,T] \times \mathcal{U}^T \times \widetilde{\mathcal{W}}^T_{BV} \mapsto \mathbb{R}^\kappa$*, where* $\widetilde{W} \subset \mathbb{R}^{\widetilde{m}}$ *Borel, be two left-continuous non-anticipative functionals such that* $H,\ \widetilde{F} \in \mathbb{C}^{1,2}([0,T])$ *and satisfy the regularity conditions from Theorem 3.2.1. Then, the composition* $\widetilde{H}(t,X,(D,\widetilde{A})) := H(t,\widetilde{F}^t(\cdot,X,\widetilde{A}),D)$ *defines a non-anticipative real-valued functional on* $[0,T] \times \mathcal{U}^T \times \widetilde{\mathcal{M}}^T_{BV}$*, where* $\widetilde{M} \subset \mathbb{R}^{n+\widetilde{m}}$ *Borel, that is of class* $\mathbb{C}^{1,2}([0,T])$ *and satisfies the regularity conditions from Theorem 3.2.1. Moreover,*

$$\partial_i \widetilde{H} = \sum_{\ell=1}^{\kappa} \partial_\ell H \partial_i \widetilde{F}_{(\ell)}, \quad i = 1, \ldots, d; \tag{3.3.5}$$

$$\nabla^2_{X,ij}\widetilde{H} = \sum_{\ell=1}^{\kappa}\left(\sum_{m=1}^{\kappa} \nabla^2_{Y,\ell m} H \partial_j \widetilde{F}_{(m)} \partial_i \widetilde{F}_{(\ell)} + \partial_\ell H \nabla^2_{X,ij}\widetilde{F}_{(\ell)}\right), \quad i,j = 1,\ldots,d; \tag{3.3.6}$$

$$\mathcal{D}\widetilde{H} = \left(\mathcal{D}_0 H + \sum_{\ell=1}^{\kappa}\partial_\ell H \mathcal{D}_0 \widetilde{F}_{(\ell)}, \sum_{\ell=1}^{\kappa}\partial_\ell H \mathcal{D}_1 \widetilde{F}_{(\ell)}, \ldots, \sum_{\ell=1}^{\kappa}\partial_\ell H \mathcal{D}_{\widetilde{m}} \widetilde{F}_{(\ell)}, \mathcal{D}_1 H, \ldots, \mathcal{D}_n H\right)^\top. \tag{3.3.7}$$

Proof. We can assume without loss of generality that $d = \kappa = \widetilde{m} = 1$. Since H is vertically differentiable with respect to Y, it follows that

$$H(t, Y^{t,h'}, D^t) = H(t, Y^t, D^t) + P(t, Y^t, D^t)h' + o(|h'|);$$

analogously, since $\widetilde{F}$ is vertically differentiable with respect to X, it follows that

$$\widetilde{F}(t, X^{t,h}, \widetilde{A}^t) = \widetilde{F}(t, X^t, \widetilde{A}^t) + Q(t, X^t, \widetilde{A}^t)h + o(|h|).$$

This implies, with $h' = \widetilde{F}(t, X^{t,h}, \widetilde{A}^t) - \widetilde{F}(t, X^t, \widetilde{A}^t)$,

$$\begin{aligned}&\widetilde{H}(t, X^{t,h}, D^t, \widetilde{A}^t) = H(t, \widetilde{F}^{t,h'}(\cdot, X, \widetilde{A}), D^t) = H(t, Y^{t,h'}, D^t) \\ &= H(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t) + P(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)(\widetilde{F}(t, X^{t,h}, \widetilde{A}^t) - \widetilde{F}(t, X^t, \widetilde{A}^t)) + o(|h'|) \\ &= H(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t) + P(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)Q(t, X^t, \widetilde{A}^t)h + o(|h|).\end{aligned}$$

Thus, $\widetilde{H}$ is vertically differentiable with respect to X, and its vertical derivative is given by

$$\partial_X \widetilde{H}(t, X^t, D^t, \widetilde{A}^t) = P(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)Q(t, X^t, \widetilde{A}^t) = \partial_Y H(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)\partial_X \widetilde{F}(t, X^t, \widetilde{A}^t),$$

which shows (3.3.5). Since H and $\widetilde{F}$ are twice vertically differentiable with respect to Y and X, respectively, an application of the chain rule in the form of (3.3.5) and the product rule from the preceding lemma directly yields (3.3.6).

Now we turn to the proof of (3.3.7). Clearly, there is nothing to show for the last n components of the vector $\mathcal{D}\widetilde{H}$. For the horizontal derivative with respect to $\widetilde{A}$, we proceed as follows. Since $\widetilde{F}$ is horizontally differentiable with respect to $\widetilde{A}$, it follows that

$$\widetilde{F}(t, X^{t-h}, \widetilde{A}^t) = \widetilde{F}(t, X^{t-h}, \widetilde{A}^{t-h}) + \Phi(t, X^t, \widetilde{A}^t)\widetilde{\nu}((t-h, t]) + o(|\widetilde{\nu}((t-h, t])|).$$

This implies, with $h' = \widetilde{F}(t, X^{t-h}, \widetilde{A}^t) - \widetilde{F}(t, X^{t-h}, \widetilde{A}^{t-h})$,

$$\begin{aligned}
&\widetilde{H}(t, X^{t-h}, D^{t-h}, \widetilde{A}^t) = H(t, \widetilde{F}^{t-h,h'}(\cdot, X, \widetilde{A}), D^{t-h}) = H(t, \widetilde{F}^{t-h}(\cdot, X, \widetilde{A}), D^{t-h}) \\
&\quad + P(t, \widetilde{F}^{t-h}(\cdot, X, \widetilde{A}), D^{t-h})(\widetilde{F}(t, X^{t-h}, \widetilde{A}^t) - \widetilde{F}(t, X^{t-h}, \widetilde{A}^{t-h})) + o(|h'|) \\
&= H(t, \widetilde{F}^{t-h}(\cdot, X, \widetilde{A}), D^{t-h}) + P(t, \widetilde{F}^{t-h}(\cdot, X, \widetilde{A}), D^{t-h})\Phi(t, X^t, \widetilde{A}^t)\widetilde{\nu}((t-h, t]) + o(|\widetilde{\nu}((t-h, t])|).
\end{aligned}$$

Thus, $\widetilde{H}$ is horizontally differentiable with respect to $\widetilde{A}$ with

$$\mathcal{D}\widetilde{H}(t, X^t, D^t, \widetilde{A}^t) = P(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)\Phi(t, X^t, \widetilde{A}^t) = \partial_Y H(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D^t)\mathcal{D}\widetilde{F}(t, X^t, \widetilde{A}^t),$$

which shows (3.3.7). Analogously, the statement follows for the first component of the vector $\mathcal{D}\widetilde{H}$. As shown in Lemma 3.3.2, products of boundedness-preserving (respectively, left-continuous) functionals are again boundedness-preserving (respectively, left-continuous). Hence, $\widetilde{H} \in \mathbb{C}^{1,2}([0, T])$ and satisfies the regularity conditions from Theorem 3.2.1. □

Proof of Theorem 3.3.1. As in the proof of Proposition 3.2.7, by concentrating on small intervals we can assume without loss of generality that $\xi_{(\ell)}$ is of the form $\xi_{(\ell)}(t) = \nabla_X F_{(\ell)}(t, X^t, A^t_{(\ell)})$, $t \in [0, T]$, for $A_{(\ell)} \in \mathcal{W}^{\ell,T}_{CBV}$, where $W^\ell \subset \mathbb{R}^{m_\ell}$, and $F_{(\ell)}$ is as in Definition 3.2.4. Then, Theorem 3.2.1 yields

$$\begin{aligned}
F_{(\ell)}(T, X^T, A^T_{(\ell)}) - F_{(\ell)}(0, X^0, A^0_{(\ell)}) &= \int_0^T \mathcal{D}F_{(\ell)}(s, X^s, A^s_{(\ell)})\nu_{(\ell)}(\mathrm{d}s) \\
+ \frac{1}{2}\sum_{i,j=1}^d \int_0^T \nabla^2_{X,ij} F_{(\ell)}(s, X^s, A^s_{(\ell)})\,\mathrm{d}[X]_{ij}(s) &+ \int_0^T \nabla_X F_{(\ell)}(s, X^s, A^s_{(\ell)})\,\mathrm{d}X(s),
\end{aligned}$$

where $\nu_{(\ell)}(\mathrm{d}s) := \big(\mathrm{d}s, A_{(\ell),1}(\mathrm{d}s), \ldots, A_{(\ell),m_\ell}(\mathrm{d}s)\big)^\top$. Introducing

$$A_{(\ell),m_\ell+1}(t) := \int_0^t \mathcal{D}F_{(\ell)}(s, X^s, A^s_{(\ell)})\nu_{(\ell)}(\mathrm{d}s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \nabla^2_{X,ij} F_{(\ell)}(s, X^s, A^s_{(\ell)})\,\mathrm{d}[X]_{ij}(s) \quad (3.3.8)$$

and setting $\widetilde{A}_{(\ell)} := \big(A_{(\ell),1}, \dots, A_{(\ell),m_\ell}, A_{(\ell),m_\ell+1}\big)^\top$, we can write

$$Y_{(\ell)}(t) = F_{(\ell)}(t, X^t, A^t_{(\ell)}) - F_{(\ell)}(0, X^0, A^0_{(\ell)}) - A_{(\ell),m_\ell+1}(t) =: \widetilde{F}_{(\ell)}(t, X^t, \widetilde{A}^t_{(\ell)}). \tag{3.3.9}$$

Here, $\widetilde{A}_{(\ell)}$ is a continuous function whose components are of bounded variation, by standard properties of Stieltjes integrals (see, e.g., [95, Theorem I.5 c]). Moreover, $\widetilde{F}_{(\ell)}$ is a non-anticipative functional of class $\mathbb{C}^{1,2}([0,T])$ with $\nabla_X \widetilde{F}_{(\ell)}(t, X, \widetilde{A}_{(\ell)}) = \nabla_X F_{(\ell)}(t, X, A_{(\ell)})$, and the regularity conditions from Theorem 3.2.1 are satisfied for $F_{(\ell)}$ being replaced by $\widetilde{F}_{(\ell)}$. Denoting

$$\widetilde{F}(t, X, \widetilde{A}) := \Big(\widetilde{F}_{(1)}(t, X, \widetilde{A}_{(1)}), \dots, \widetilde{F}_{(\kappa)}(t, X, \widetilde{A}_{(\kappa)})\Big)^\top,$$

the identity (3.3.9) reads

$$Y(t) = \widetilde{F}(t, X^t, \widetilde{A}^t). \tag{3.3.10}$$

Again by concentrating on small intervals, η can be written without loss of generality as $\eta(t) = \nabla_Y H(t, Y^t, D^t)$, $t \in [0,T]$, for $D \in \mathcal{W}^T_{CBV}$, where $W \subset \mathbb{R}^m$, and H is as in Definition 3.2.4. Using (3.3.10) as well as the notation $\nabla_X \widetilde{F}(t, X, \widetilde{A})$ for the matrix of vertical derivatives of $\widetilde{F}$, Lemma 3.3.3 yields

$$\begin{aligned}
\sum_{\ell=1}^{\kappa} \eta_\ell(t)\xi_{(\ell)}(t) &= \sum_{\ell=1}^{\kappa} \partial_\ell H(t, Y^t, D^t) \nabla_X \widetilde{F}_{(\ell)}(t, X^t, \widetilde{A}^t_{(\ell)}) \\
&= \nabla_Y H\left(t, \widetilde{F}^t(\cdot, X^\cdot, \widetilde{A}^\cdot), D^t\right) \cdot \nabla_X \widetilde{F}(t, X^t, \widetilde{A}^t) \\
&= \nabla_X \widetilde{H}(t, X^t, \widetilde{D}^t),
\end{aligned}$$

where $\widetilde{D} = (D, \widetilde{A}) \in \mathcal{W}^T_{CBV} \times \widetilde{\mathcal{W}}^T_{CBV}$, with $W \subset \mathbb{R}^m$, $\widetilde{W} \subset \mathbb{R}^{\widetilde{m}}$ Borel, and $\widetilde{m} = \sum_{\ell=1}^{\kappa} m_\ell + \kappa$. Moreover,

$$\widetilde{H}(t, X, (D, \widetilde{A})) := H\left(t, \widetilde{F}^t(\cdot, X, \widetilde{A}), D\right)$$

defines a non-anticipative left-continuous functional that satisfies the requirements of Definition 3.2.4, also by Lemma 3.3.3. We hence infer that $\sum_{\ell=1}^{\kappa} \eta_\ell \xi_{(\ell)}$ is admissible for X.

Using (3.3.10) we get

$$\begin{aligned}
\int_0^T \eta(s)\,\mathrm{d}Y(s) &= \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n} \sum_{\ell=1}^{\kappa} \partial_\ell H(s, Y^{n,s-}, D^{n,s})(F_{(\ell)}(s', X^{s'}, A^{s'}_{(\ell)}) - F_{(\ell)}(s, X^s, A^s_{(\ell)})) \\
&\quad - \lim_{n\to\infty} \sum_{s\in\mathbb{T}_n} \sum_{\ell=1}^{\kappa} \partial_\ell H(s, Y^{n,s-}, D^{n,s})(A_{(\ell),m_\ell+1}(s') - A_{(\ell),m_\ell+1}(s)),
\end{aligned} \tag{3.3.11}$$

where $Y^{n,s-}$ denotes the n-th approximation of Y, stopped at time $s-$. Since the quantity $\partial_\ell H(s, Y^{n,s-}, D^{n,s})\mathbb{1}_{u\in(s,s']}$ is bounded and converges to $\eta_\ell(u)$, by left-continuity of $\nabla_Y H$, and both are left-continuous in u, by [20, Proposition 1], and since $A_{(\ell),m_\ell+1}$ is a continuous function of bounded variation, and hence corresponds to an atomless measure on $[0,T]$, we can use a "diagonal lemma" for vague convergence of measures in the form of [20, Lemma 12]. Thus, with the associativity of the Stieltjes integral (see, e.g., [95, Theorem I.6 b]) and the definition of $A_{(\ell),m_\ell+1}$, we infer for the second term in (3.3.11):

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s,Y^{n,s-},D^{n,s})(A_{(\ell),m_\ell+1}(s')-A_{(\ell),m_\ell+1}(s))\\
&=\sum_{\ell=1}^{\kappa}\int_0^T\eta_\ell(s)\mathcal{D}F_{(\ell)}(s,X^s,A^s_{(\ell)})\nu_{(\ell)}(\mathrm{d}s)\\
&\quad+\frac{1}{2}\sum_{\ell=1}^{\kappa}\sum_{i,j=1}^{d}\int_0^T\eta_\ell(s)\nabla^2_{X,ij}F_{(\ell)}(s,X^s,A^s_{(\ell)})\,\mathrm{d}[X]_{ij}(s).
\end{aligned}\tag{3.3.12}$$

For the first term in (3.3.11), as in the proof of Proposition 3.2.7, we approximate the increments of $F_{(\ell)}$ for the path $(X,A_{(\ell)})$ by the respective increments along the piecewise constant approximation $(X^n,A^n_{(\ell)})$. As in the proof of Theorem 3.2.1, for $s\in\mathbb{T}_n$ we consider the following decomposition of the increments of $F_{(\ell)}$ along the piecewise constant approximation $(X^n,A^n_{(\ell)})$ into "horizontal" and "vertical" terms:

$$\begin{aligned}
F_{(\ell)}(s',X^{n,s'-},A^{n,s'}_{(\ell)})-F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)})&=F_{(\ell)}(s',X^{n,s'-},A^{n,s'}_{(\ell)})-F_{(\ell)}(s,X^{n,s},A^{n,s}_{(\ell)})\\
&\quad+F_{(\ell)}(s,X^{n,s},A^{n,s}_{(\ell)})-F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)}).
\end{aligned}\tag{3.3.13}$$

The first term on the right-hand side of (3.3.13) is equal to $\psi_{(\ell)}(h^n_s)-\psi_{(\ell)}(0)$, where $\psi_{(\ell)}(u):=F_{(\ell)}(s+u,X^{n,s},A^{n,s+u}_{(\ell)})$. Since $F_{(\ell)}$ admits a horizontal derivative in the sense of (3.1.7), (3.1.8), we can write

$$F_{(\ell)}(s',X^{n,s'-},A^{n,s'}_{(\ell)})-F_{(\ell)}(s,X^{n,s},A^{n,s}_{(\ell)})=\int_{(0,s'-s]}\mathcal{D}F_{(\ell)}(s+u,X^{n,s},A^{n,s+u-}_{(\ell)})\nu_{(\ell),n}(\mathrm{d}u),\tag{3.3.14}$$

where $\nu_{(\ell),n}$ is the measure corresponding to the n-th approximation $A^n_{(\ell)}$ of $A_{(\ell)}$. For the second term on the right-hand side of (3.3.13), we have

$$F_{(\ell)}(s,X^{n,s},A^{n,s}_{(\ell)})-F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)})=\phi_{(\ell)}(\delta X^n_s)-\phi_{(\ell)}(0),\tag{3.3.15}$$

where $\phi_{(\ell)}(u)=F_{(\ell)}(s,X^{n,s-,u},A^{n,s}_{(\ell)})$. Since $F_{(\ell)}\in\mathbb{C}^{1,2}([0,T])$, the function ϕ is well-defined and twice continuously differentiable in an open neighborhood $\mathcal{B}$ of $X(s)$ in U, with

$$\nabla\phi_{(\ell)}(u)=\nabla_X F_{(\ell)}(s,X^{n,s-,u},A^{n,s}_{(\ell)}),\tag{3.3.16}$$

$$\nabla^2\phi_{(\ell)}(u)=\nabla^2_X F_{(\ell)}(s,X^{n,s-,u},A^{n,s}_{(\ell)}).\tag{3.3.17}$$

Hence, a second-order Taylor expansion of $\phi_{(\ell)}$ at $u=0$ yields

$$\begin{aligned}
F_{(\ell)}(s,X^{n,s},A^{n,s}_{(\ell)})-F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)})&=\nabla_X F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)})\delta X^n_s\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}\nabla^2_{X,ij}F_{(\ell)}(s,X^{n,s-},A^{n,s}_{(\ell)})\delta X^n_{i,s}\delta X^n_{j,s}+r^n_{s,\ell},
\end{aligned}\tag{3.3.18}$$

where for some $\theta \in [0,1]$,

$$r^n_{s,\ell} := \frac{1}{2}\sum_{i,j=1}^{d}\left(\nabla^2_{X,ij}F_{(\ell)}(s, X^{n,s-,\theta\delta X^n_s}, A^{n,s}_{(\ell)}) - \nabla^2_{X,ij}F_{(\ell)}(s, X^{n,s-}, A^{n,s}_{(\ell)})\right)\delta X^n_{i,s}\delta X^n_{j,s}.$$

Since $(X, A_{(\ell)})$, (Y, D) are continuous and hence, uniformly continuous on $[0,T]$, it follows that

$$\lambda_{n,\ell} := \sup_{\alpha,\beta\in[0,T],|\alpha-\beta|\le|\mathbb{T}_n|}\Big(|A_{(\ell)}(\alpha) - A_{(\ell)}(\beta)| + |X(\alpha) - X(\beta)| + |D(\alpha) - D(\beta)| + |Y(\alpha) - Y(\beta)|\Big)$$

converges to 0 as $n \to \infty$. Thus, since $\nabla^2_X F_{(\ell)}$, $\mathcal{D}F_{(\ell)}$, and $\nabla_Y H$ are boundedness-preserving, we infer that for n large enough there exists a constant $C > 0$ such that

$$\begin{aligned}
&\forall t < T,\ \forall (X', A'_{(\ell)}) \in \mathcal{U}^t \times \mathcal{W}^{\ell,t}_{BV},\ (Y', D') \in \mathcal{V}^t \times \mathcal{W}^t_{BV},\\
&d_\infty\left((X^t, A^t_{(\ell)}), (X', A'_{(\ell)})\right) + d_\infty\left((Y^t, D^t), (Y', D')\right) < \lambda_{n,\ell}\\
&\Rightarrow |\partial_\ell H(t, Y', D')\mathcal{D}F_{(\ell)}(t, X', A'_{(\ell)})| \le C, \quad \|\partial_\ell H(t, Y', D')\nabla^2_X F_{(\ell)}(t, X', A'_{(\ell)})\| \le C.
\end{aligned}$$

We now sum over $s \in \mathbb{T}_n$ and $\ell = 1, \dots, \kappa$. As in the proof of Theorem 3.2.1, we have that

$$\sum_{s\in\mathbb{T}_n, s\le t}\left(F_{(\ell)}(s', X^{n,s'-}, A^{n,s'}_{(\ell)}) - F_{(\ell)}(s, X^{n,s}, A^{n,s}_{(\ell)})\right) = \int_{(0,t]}\mathcal{D}F_{(\ell)}\left(u, X^{n,s(u)}, A^{n,u-}_{(\ell)}\right)\nu_{\ell,n}(\,\mathrm{d}u)$$

converges for all $t \in [0,T]$ to $B_{(\ell)}(t) := \int_0^t \mathcal{D}F_{(\ell)}(s, X^s, A^s_{(\ell)})\nu_{(\ell)}(\,\mathrm{d}s)$, using a "diagonal lemma" for vague convergence of measures in the form of [20, Lemma 12]. As above, the quantity $\partial_\ell H(s, Y^{n,s-}, D^{n,s})\mathbb{1}_{u\in(s,s']}$ is bounded, converges to $\eta_\ell(u)$, by left-continuity of $\nabla_Y H$, and both are left-continuous in u, by [20, Proposition 1]. Thus, since $B_{(\ell)}$ is continuous and of bounded variation, and hence corresponds to an atomless measure on $[0,T]$, another application of [20, Lemma 12], in conjunction with the associativity of the Stieltjes integral (see [95, Theorem I.6 b]), yields that

$$\begin{aligned}
&\sum_{\ell=1}^{\kappa}\sum_{s\in\mathbb{T}_n}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\left(F_{(\ell)}(s', X^{n,s'-}, A^{n,s'}_{(\ell)}) - F_{(\ell)}(s, X^{n,s}, A^{n,s}_{(\ell)})\right)\\
&\quad\to \sum_{\ell=1}^{\kappa}\int_0^T \eta_\ell(s)\mathcal{D}F_{(\ell)}(s, X^s, A^s_{(\ell)})\nu_{(\ell)}(\,\mathrm{d}s).
\end{aligned} \tag{3.3.19}$$

Moreover, using (3.3.18) we obtain that

$$\begin{aligned}
&\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\left(F_{(\ell)}(s, X^{n,s}, A^{n,s}_{(\ell)}) - F_{(\ell)}(s, X^{n,s-}, A^{n,s}_{(\ell)})\right)\\
&\quad= \sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla_X F_{(\ell)}(s, X^{n,s-}, A^{n,s}_{(\ell)})\cdot\delta X^n_s\\
&\qquad+ \frac{1}{2}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\sum_{i,j=1}^{d}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla^2_X F_{(\ell),\,ij}(s, X^{n,s-}, A^{n,s}_{(\ell)})\delta X^n_{i,s}\delta X^n_{j,s}\\
&\qquad+ \sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})r^n_{s,\ell}.
\end{aligned} \tag{3.3.20}$$

The quantity $\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla_X^2 F_{(\ell),\,ij}(s, X^{n,s-}, A_{(\ell)}^{n,s})\mathbb{1}_{u\in(s,s']}$ is bounded by C, converges to $\eta_\ell(u)\nabla_X^2 F_{(\ell),\,ij}(u, X^u, A_{(\ell)}^u)$, by left-continuity of $\nabla_Y H$, $\nabla_X^2 F_{(\ell)}$, and both are left-continuous in u (by [20, Proposition 1]). Since $X \in QV^d$, an application of [20, Lemma 12] gives that

$$\begin{aligned}&\frac{1}{2}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\sum_{i,j=1}^{d}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla_X^2 F_{(\ell),\,ij}(s, X^{n,s-}, A_{(\ell)}^{n,s})\delta X_{i,s}^n\delta X_{j,s}^n\\ &\quad\to \frac{1}{2}\sum_{\ell=1}^{\kappa}\sum_{i,j=1}^{d}\int_0^T \eta_\ell(u)\nabla_X^2 F_{(\ell),\,ij}(u, X^u, A_{(\ell)}^u)\,\mathrm{d}[X]_{ij}(u)\end{aligned}\tag{3.3.21}$$

as $n\to\infty$. Using the same argument, since $|\partial_\ell H(s, Y^{n,s-}, D^{n,s})r_{s,\ell}^n|$ is bounded by $\epsilon_{s,\ell}^n|\delta X_s^n|^2$, where $\epsilon_{s,\ell}^n$ converges to 0, the remainder term, $\sum_{\ell=1}^{\kappa}\sum_{s\in\mathbb{T}_n}\partial_\ell H(s, Y^{n,s-}, D^{n,s})r_{s,\ell}^n$, converges to 0.

Applying (3.3.19) and (3.3.21) we thus obtain for the first term in (3.3.11):

$$\begin{aligned}&\lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\left(F_{(\ell)}(s', X^{s'}, A_{(\ell)}^{s'}) - F_{(\ell)}(s, X^s, A_{(\ell)}^s)\right)\\ &\quad= \lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\left(F_{(\ell)}(s', X^{n,s'-}, A_{(\ell)}^{n,s'}) - F_{(\ell)}(s, X^{n,s-}, A_{(\ell)}^{n,s})\right)\\ &\quad= \sum_{\ell=1}^{\kappa}\int_0^T \eta_\ell(s)\mathcal{D}F_{(\ell)}(s, X^s, A_{(\ell)}^s)\nu_{(\ell)}(\,\mathrm{d}s)\\ &\qquad+ \lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla_X F_{(\ell)}(s, X^{n,s-}, A_{(\ell)}^{n,s})\delta X_s^n\\ &\qquad+ \frac{1}{2}\sum_{\ell=1}^{\kappa}\sum_{i,j=1}^{d}\int_0^T \eta_\ell(s)\nabla_X^2 F_{(\ell),\,ij}(u, X^s, A_{(\ell)}^s)\,\mathrm{d}[X]_{ij}(s).\end{aligned}\tag{3.3.22}$$

Thus, (3.3.11) becomes, in conjunction with (3.3.12),

$$\int_0^T \eta(s)\,\mathrm{d}Y(s) = \lim_{n\to\infty}\sum_{s\in\mathbb{T}_n}\sum_{\ell=1}^{\kappa}\partial_\ell H(s, Y^{n,s-}, D^{n,s})\nabla_X F_{(\ell)}(s, X^{n,s-}, A_{(\ell)}^{n,s})\cdot\delta X_s^n.$$

Since we have already established in the first step that $\sum_{\ell=1}^{\kappa}\eta_\ell(s)\xi_{(\ell)}(s)$ is admissible for X, this concludes the proof. □

3.4 Applications to Itô Differential Equations

The associativity rule derived in the preceding section guarantees that informal computations with Itô differentials typically lead to correct statements. It is therefore of fundamental importance and crucial for many applications. For instance, the elementary associativity rule from [80] was derived in order to show that CPPI strategies can be constructed in a purely pathwise manner. Moreover, it enables us to translate the Doss–Sussmann method to the pathwise Itô calculus (see [67, Section 2.3]). Our motivation for establishing the associativity rule within functional pathwise Itô calculus originated from the wish to give a strictly pathwise treatment of stochastic portfolio theory; see [84] and the discussion in Chapter 4. To illustrate already at this point why the associativity rule is so important, we will now use it to prove existence and uniqueness results for pathwise linear Itô differential equations whose coefficients are non-anticipative functionals.

Let $X \in QV^d$, $Y \in QV$ be such that all covariations $[Y, X_i]$, $i = 1, \dots, d$, exist and are continuous, and let $\sigma = (\sigma_1, \dots, \sigma_d)$ be a locally admissible integrand for X. Then, $Z \in C([0,T], \mathbb{R})$ is called a solution of the *linear Itô differential equation*

$$\mathrm{d}Z(t) = \mathrm{d}Y(t) + Z(t)\sigma(t)\,\mathrm{d}X(t), \tag{3.4.1}$$

with initial condition $Z(0) = z$, if the map $t \to Z(t)\sigma(t)$ is a locally admissible integrand for X and if Z satisfies the integral form of (3.4.1): $Z(t) = z + \int_0^t Z(s)\sigma(s)\,\mathrm{d}X(s)$, $0 \le t \le T$.

Theorem 3.4.1 (Existence and uniqueness of the homogeneous linear IDE)**.** *Suppose that σ is a locally admissible integrand for $X \in QV^d$. Then, for any $z \in \mathbb{R}$, the homogeneous linear Itô differential equation*

$$\mathrm{d}Z(t) = Z(t)\sigma(t)\,\mathrm{d}X(t), \tag{3.4.2}$$

with initial condition $Z(0) = z$, has the unique solution

$$Z(t) = z \exp\left(\int_0^t \sigma(s)\,\mathrm{d}X(s) - \frac{1}{2} \sum_{i,j=1}^d \int_0^t \sigma_i(s)\sigma_j(s)\,\mathrm{d}[X]_{ij}(s) \right). \tag{3.4.3}$$

The solution Z of the above equation is the *Doléans-Dade* exponential of

$$I(t) := \int_0^t \sigma(s)\,\mathrm{d}X(s), \tag{3.4.4}$$

and will be denoted by $\mathcal{E}(I)(t)$ in the following.

Proof of Theorem 3.4.1. Let us first show that Z satisfies the required equation. Since I from (3.4.4) has the continuous quadratic variation $[I](t) = \sum_{i,j=1}^d \int_0^t \sigma_i(s)\sigma_j(s)\,\mathrm{d}[X]_{ij}(s)$, by Proposition 3.2.7, setting

$$K(t) := I(t) - \frac{1}{2}[I](t) = I(t) - \sum_{i,j=1}^d \int_0^t \sigma_i(s)\sigma_j(s)\,\mathrm{d}[X]_{ij}(s),$$

and applying Föllmer's pathwise Itô formula [46] to K and $f(k) = e^k$ we obtain:

$$Z(t) = f(K(t)) = Z(0) + \int_0^t Z(s)\,\mathrm{d}K(s) + \frac{1}{2}\int_0^t Z(s)\,\mathrm{d}[K](s) = z + \int_0^t Z(s)\sigma(s)\,\mathrm{d}X(s).$$

Here we have used Theorem 3.3.1, with $\eta = Z$ and $\xi = \sigma$, the associativity of the Stieltjes integral from [95, Theorem I.6 b], and the fact that $[K](t) = [I](t)$, due to [80, Remark 8].

Let us now turn to the proof of the uniqueness of solutions. Throughout the proof we let Z be as defined in (3.4.3) and suppose that $\widetilde{Z}$ is another solution of (3.4.2) with initial condition z. We distinguish three cases.

<u>I. $z > 0$:</u> Since $Z > 0$ in this case, applying the pathwise Itô formula [46] to the function $f(z, \widetilde{z}) = \widetilde{z}/z$ and the paths Z, $\widetilde{Z}$ yields that

$$\frac{\widetilde{Z}(t)}{Z(t)} = 1 + \int_0^t \begin{pmatrix} 1/Z(s) \\ -\widetilde{Z}(s)/Z^2(s) \end{pmatrix} \mathrm{d}\begin{pmatrix} \widetilde{Z}(s) \\ Z(s) \end{pmatrix} + \int_0^t \frac{\widetilde{Z}(s)}{Z^3(s)}\,\mathrm{d}[Z](s) - \int_0^t \frac{1}{Z^2(s)}\,\mathrm{d}[Z, \widetilde{Z}](s). \tag{3.4.5}$$

Since both Z and $\widetilde{Z}$ satisfy the integral form of (3.4.2), the associativity result from Theorem 3.3.1 gives for $\kappa = 2$, $\eta = \begin{pmatrix} 1/Z \\ -\widetilde{Z}/Z^2 \end{pmatrix}$, $\xi_{(1)}(s) = \widetilde{Z}(s)\sigma(s)$, and $\xi_{(2)}(s) = Z(s)\sigma(s)$ that

$$\int_0^t \begin{pmatrix} 1/Z(s) \\ -\widetilde{Z}(s)/Z^2(s) \end{pmatrix} \mathrm{d}\begin{pmatrix} \widetilde{Z}(s) \\ Z(s) \end{pmatrix} = \int_0^t \left(\frac{1}{Z(s)}\widetilde{Z}(s) - \frac{\widetilde{Z}(s)}{Z^2(s)} Z(s) \right) \sigma(s)\,\mathrm{d}X(s) = 0,$$

and so the pathwise Itô integral vanishes. Moreover, using Proposition 3.2.7 for the quadratic variation $[Z]$ and covariation $[Z, \widetilde{Z}]$ we obtain, in conjunction with the associativity of the Stieltjes integral from [95, Theorem I.6 b],

$$\int_0^t \frac{\widetilde{Z}(s)}{Z^3(s)}\,\mathrm{d}[Z](s) = \sum_{i,j=1}^d \int_0^t \sigma_i(s)\sigma_j(s)\frac{\widetilde{Z}(s)}{Z(s)}\,\mathrm{d}[X]_{ij}(s) = \int_0^t \frac{1}{Z^2(s)}\,\mathrm{d}[Z, \widetilde{Z}](s).$$

Plugging these results back into (3.4.5) we arrive at $\frac{\widetilde{Z}(t)}{Z(t)} \equiv 1$, which is the desired uniqueness in case $z > 0$.

<u>II. $z < 0$:</u> If $\widetilde{Z}$ is any solution of (3.4.2) with $\widetilde{Z}(0) = z$, then $\widehat{Z} := -\widetilde{Z}$ also is a solution of (3.4.2) with initial condition $\widehat{Z}(0) = -z > 0$. Hence, uniqueness holds by the previous argument.

<u>III. $z = 0$:</u> Suppose by way of contradiction that there exists a non-zero solution $\widetilde{Z}$ of (3.4.2) with $\widetilde{Z}(0) = 0$. For instance, this is the case if there exists $t > 0$ such that $\widetilde{Z}(t) > 0$, the case $\widetilde{Z}(t) < 0$ can be treated in the same manner. Then, the stopping time $\tau_n := \inf\{t > 0 \,\big|\, \widetilde{Z}(t) = \frac{1}{n}\}$ will be finite for sufficiently large $n \in \mathbb{N}$. Introducing the time-shifted paths $\widehat{Z}(t) := \widetilde{Z}(t + \tau_n)$, $\widehat{\sigma}(t) := \sigma(t + \tau_n)$, and $\widehat{X}(t) := X(t + \tau_n)$, we have for sufficiently large n,

$$\widehat{Z}(t) = \frac{1}{n} + \int_0^t \widehat{Z}(s)\widehat{\sigma}(s)\,\mathrm{d}\widehat{X}(s), \quad t \geq 0,$$

which is equivalent to $\widehat{Z}(t) = \frac{1}{n}\mathcal{E}\left(\int_0^\cdot \widehat{\sigma}(s)\,\mathrm{d}\widehat{X}(s)\right)(t)$, by previous arguments. It follows that

$$\widetilde{Z}(t+\tau_n) = \frac{1}{n}\exp\left(\int_{\tau_n}^{t+\tau_n} \sigma(s)\,\mathrm{d}X(s) - \frac{1}{2}\sum_{i,j=1}^{d}\int_{\tau_n}^{t+\tau_n} \sigma_i(s)\sigma_j(s)\,\mathrm{d}[X_i,X_j](s)\right) \tag{3.4.6}$$

for $t \geq 0$. Clearly, the integrals inside the exponential function are uniformly bounded in n. Thus, letting n go to infinity in (3.4.6) yields $\widetilde{Z}(t) = 0$ for all $t \geq \lim_n \tau_n$, which is the desired contradiction. Hence, $Z(t) = 0$ is the only solution with initial value $Z(0) = 0$. □

For the general *inhomogeneous* linear Itô differential equation (3.4.1) the following result extends the standard “variation of constants” technique.

Theorem 3.4.2 (Existence and uniqueness of the inhomogeneous linear IDE)**.** *Suppose that σ is a locally admissible integrand for $X \in QV^d$ and $Y \in QV$ is such that all covariations $[Y, X_i]$, $i = 1, \ldots, d$, exist and are continuous. Then, for*

$$Z^0(t) := \mathcal{E}\Big(\int_0^\cdot \sigma(s)\,\mathrm{d}X(s)\Big)(t),$$

and any $z \in \mathbb{R}$, the inhomogeneous linear Itô differential equation

$$\mathrm{d}Z(t) = \mathrm{d}Y(t) + Z(t)\sigma(t)\,\mathrm{d}X(t), \tag{3.4.7}$$

with initial condition $Z(0) = z$, has the unique solution

$$Z(t) = Z^0(t)\left(z + \int_0^t \frac{1}{Z^0(s)}\,\mathrm{d}Y(s) - \int_0^t \frac{1}{Z^0(s)}\,\mathrm{d}\Big[Y, \int_0^\cdot \sigma(s)\,\mathrm{d}X(s)\Big](s)\right). \tag{3.4.8}$$

Proof. If Z and $\widetilde{Z}$ are solutions of (3.4.7) with the same initial condition, then their difference $\widehat{Z} := Z - \widetilde{Z}$ satisfies $\mathrm{d}\widehat{Z}(t) = \widehat{Z}(t)\sigma(t)\,\mathrm{d}X(t)$ with initial condition $\widehat{Z}(0) = 0$. Hence, Theorem 3.4.1 implies $\widehat{Z} \equiv 0$, which shows the uniqueness of solutions. In the second step, an application of Föllmer's pathwise product rule and of the associativity result from Theorem 3.3.1 shows that Z indeed solves (3.4.7).

More precisely, we introduce

$$A(t) := z + \int_0^t \frac{1}{Z^0(s)}\,\mathrm{d}Y(s) - \int_0^t \frac{1}{Z^0(s)}\,\mathrm{d}\Big[Y, \int_0^\cdot \sigma(s)\,\mathrm{d}X(s)\Big](s). \tag{3.4.9}$$

Thus, applying Föllmer's pathwise Itô formula [46] to the paths Z_0, A with the function $f(z_0, a) = z_0 a$ yields

$$Z(t) = z + \int_0^t A(s)\,\mathrm{d}Z^0(s) + \int_0^t Z^0(s)\,\mathrm{d}A(s) + [Z^0, A](t).$$

Since Z^0 satisfies $Z^0(t) = 1 + \int_0^t Z^0(s)\sigma(s)\,\mathrm{d}X(s)$, as shown previously in Theorem 3.4.1, we infer in conjunction with the integral form of A,

$$\begin{aligned} Z(t) = z &+ \int_0^t Z^0(s)A(s)\sigma(s)\,\mathrm{d}X(s) + \int_0^t Z^0(s)\frac{1}{Z^0(s)}\,\mathrm{d}Y(s) \\ &- \int_0^t Z^0(s)\frac{1}{Z^0(s)}\,\mathrm{d}\Big[Y, \int_0^{\cdot} \sigma(s)\,\mathrm{d}X(s)\Big](s) + [Z^0, A](t). \end{aligned} \tag{3.4.10}$$

Furthermore, by Proposition 3.2.7,

$$\begin{aligned} [Z^0, A](t) &= \Big[\int_0^{\cdot} Z^0(s)\sigma(s)\,\mathrm{d}X(s), \int_0^{\cdot} \frac{1}{Z^0(s)}\,\mathrm{d}Y(s)\Big](t) = \sum_{i=1}^{d} \int_0^t \sigma_i(s)\,\mathrm{d}[X_i, Y](s) \\ &= \Big[Y, \int_0^{\cdot} \sigma(s)\,\mathrm{d}X(s)\Big](t). \end{aligned}$$

Hence, we obtain

$$Z(t) = z + \int_0^t Z(s)\sigma(s)\,\mathrm{d}X(s) + Y(t).$$

which concludes the proof. □

Chapter 4

Model-free portfolio theory and its functional master formula

The purpose of this chapter, which follows [84], is twofold. On the one hand, we will deal with an extension of the master formula in Stochastic portfolio theory (SPT) [37, 38, 40, 45] to path-dependent portfolio generating functions. On the other hand, since our framework is based on the pathwise Itô calculus developed by Föllmer [46], Dupire [34], and Cont, Fournié [20], in the slightly extended form of [82] (as described in the previous chapter), we will be able to overcome uncertainty issues in specifying a probabilistic model. Thus, we will obtain a new case study in which continuous-time trading strategies can be constructed in a probability-free manner by means of pathwise Itô calculus.

The chapter is organized as follows. In Section 4.1, we introduce our market model; in particular, we introduce the notion of the market portfolio and of the excess growth rate of a portfolio in a strictly pathwise setting. Our main result on the pathwise functional extension of the master equation is stated in Theorem 4.2.5, whose proof essentially relies on the associativity of the pathwise functional Itô integral, Theorem 3.3.1. In Section 4.3, we discuss simulation results, obtained using the stock data base available from Reuters Datastream.

4.1 Markets and portfolios

Pathwise Itô calculus can be used to model financial markets without probabilistic assumptions on the underlying asset price dynamics; see, e.g., [9, 11, 25, 47, 49, 79, 84, 86] for corresponding case studies. In the following, we consider a financial market model consisting of d risky assets and a locally riskless bond. The price of the bond is given by

$$\mathrm{d}B(t) = B(t)r(t)\,\mathrm{d}t, \quad B(0) = 1, \tag{4.1.1}$$

where $r : [0, \infty) \to \mathbb{R}$ is a measurable short rate function satisfying $\int_0^T |r(s)|\,\mathrm{d}s < \infty$ for all $T > 0$. The prices of the risky assets are described by a d-dimensional continuous path S with values in an open subset $U \subset \mathbb{R}_+^d$, where $\mathbb{R}_+ := (0, \infty)$. Our only requirement is that these asset prices allow for the functional pathwise Itô calculus [46, 20, 34] in the slightly extended form of [82], and that trading strategies are such that the corresponding pathwise Itô integrals exist. For the remainder of this chapter we fix a refining sequence of partitions $(\mathbb{T}_n)$ of $[0,\infty)$, i.e., $\mathbb{T}_n = \{t_0, t_1, \dots\}$ is such that $0 = t_0 < t_1 < \dots$ and $t_k \to \infty$ as $k \to \infty$, we have $\mathbb{T}_1 \subset \mathbb{T}_2 \subset \cdots$, and the mesh of $\mathbb{T}_n$ tends to zero on each compact interval as $n \uparrow \infty$. We assume that $S \in QV^d$ in the sense of Definition 2.2.5, adjusted for $[0,\infty)$, i.e., for any $1 \le i, j \le d$ and all $t \ge 0$ the sequence

$$\sum_{\substack{s\in\mathbb{T}_n \\ s\le t}} (S_i(s') - S_i(s))\,(S_j(s') - S_j(s)) \tag{4.1.2}$$

converges to a finite limit, denoted $[S_i, S_j](t)$, such that $t \to [S_i, S_j](t)$ is continuous. The functional Itô formula for continuous paths, Theorem 3.2.1, allows us to define admissible integrands ξ for S as follows, so as to guarantee the existence of the pathwise Itô integral $\int_0^T \xi(s)\,\mathrm{d}S(s)$, $T > 0$, as the limit of non-anticipative Riemann sums in the sense of (3.2.4).

Definition 4.1.1. A function $\xi : [0,\infty) \mapsto \mathbb{R}^d$ is called a *locally admissible integrand* for S if for every $T > 0$ there exists a partition $\mathbb{T} = \{t_0, \dots, t_N\}$ of $[0,T]$ such that $\forall k = 1,\dots,N$, there are constants $m_k \in \mathbb{N}$, functions $A_k \in \mathcal{W}_{[t_{k-1},t_k],CBV}$, where $W \subset \mathbb{R}^{m_k}$, and non-anticipative functionals F_k as in Theorem 3.2.1 such that

$$\xi(t) = \sum_{k=1}^{N} \nabla_X F_k(t, S_k^t, A_k^t)\mathbb{1}_{t\in[t_{k-1},t_k)}.$$

Here, $\nabla_X F_k$ is the vertical derivative of F_k with respect to X in the sense of Definition 3.1.5, $S_k := S\big|_{[t_{k-1},t_k]}$ is the restriction of S to $[t_{k-1}, t_k]$, and we require

$$F_k(t_k, S_k^{t_k}, A_k^{t_k}) = F_{k+1}(t_k, S_k^{t_k}, A_k^{t_k}) \quad \text{for} \quad k = 1,\dots,N-1.$$

Definition 4.1.2. Suppose that ξ is a locally admissible integrand for S and η is a real-valued measurable function on $[0,\infty)$ such that $\int_0^T |\eta(s)|\,\mathrm{d}|B|(s) < \infty$ for all $T > 0$. Then the pair (ξ, η) is called a *trading strategy*, where $\xi_i(t)$ corresponds to the number of shares held in the i-th stock at time t, and $\eta(t)$ is the number of shares held in the riskless bond at time t.

In the following, we focus on the class of self-financing trading strategies that do not require the infusion or withdrawal of capital after time $t = 0$, as motivated at the beginning of Chapter 2.

Definition 4.1.3. Let ξ be a locally admissible integrand for S and η a real-valued measurable function on $[0,\infty)$ such that $\int_0^T |\eta(s)|\,\mathrm{d}|B|(s) < \infty$ for all $T > 0$. The trading strategy (ξ, η) is said to be *self-financing* if the associated wealth $V(t) := \xi(t)\cdot S(t) + \eta(t)B(t)$ satisfies the identity

$$V(t) = V(0) + \int_0^t \xi(s)\,\mathrm{d}S(s) + \int_0^t \eta(s)\,\mathrm{d}B(s), \quad t \ge 0. \tag{4.1.3}$$

Note that the prices $S(t) = (S_1(t), \dots, S_d(t))^\top$, $B(t)$ are given in currency units that are payable at time t. However, euros that are payable today and euros that are payable in, for instance, one year from now do not have the equal value (this can be seen in the change of the bond price). In order to compare currency units that are payable at different times, we introduce the *discounted* prices, as follows

$$\widetilde{S}(t) := \left(\frac{S_1(t)}{B(t)}, \dots, \frac{S_d(t)}{B(t)}\right)^\top, \ \widetilde{V}(t) := \frac{V(t)}{B(t)} = \xi(t) \cdot \widetilde{S}(t) + \eta(t), \ \widetilde{B}(t) \equiv 1. \tag{4.1.4}$$

Lemma 4.1.4. *It holds that $S \in QV^d$ if and only if $\widetilde{S} \in QV^d$.*

Proof. Without loss of generality assume $d = 1$. First, let $S \in QV$. Föllmer's pathwise Itô formula [46] applied to the function $f(s, a) = s \cdot a$ and to the paths S and $1/B$ yields that

$$\widetilde{S}(t) = \widetilde{S}(0) + \int_0^t \frac{1}{B(s)} \,\mathrm{d}S(s) - \int_0^t \frac{1}{B(s)} \widetilde{S}(s) \,\mathrm{d}B(s), \tag{4.1.5}$$

in conjunction with the fact that $[S, 1/B]$ vanishes, due to [80, Remark 8]. Since $\int_0^t \frac{1}{B(s)} \widetilde{S}(s) \,\mathrm{d}B(s)$ is a Riemann-Stieltjes integral, its quadratic variation along $(\mathbb{T}_n)$ vanishes. Since $\int_0^t \frac{1}{B(s)} \,\mathrm{d}S(s)$ is a pathwise Itô integral, it belongs to QV, by Proposition 3.2.7. Moreover, both admit (vanishing) covariation along $(\mathbb{T}_n)$, by polarization. It thus follows that $\widetilde{S} \in QV$. On the other hand, if $\widetilde{S} \in QV$, an application of Föllmer's pathwise Itô formula [46] to the function $f(s, a) = s \cdot a$ with the paths $\widetilde{S}$ and B and a completely analogous argument yield that $S \in QV$. $\square$

The following proposition shows that when focusing on self-financing strategies, we can work without loss of generality with the discounted prices. It is standard in stochastic calculus, however, in our strictly pathwise setting its proof needs the associativity rule from Theorem 3.3.1.

Proposition 4.1.5. *If ξ is a locally admissible integrand for S, then ξ is also a locally admissible integrand for $\widetilde{S}$, and a trading strategy (ξ, η) is self-financing if and only if its associated discounted wealth satisfies*

$$\widetilde{V}(t) = \widetilde{V}(0) + \int_0^t \xi(s) \,\mathrm{d}\widetilde{S}(s), \quad t \geq 0. \tag{4.1.6}$$

In this case, the riskless component η is given by

$$\eta(t) = V(0) + \int_0^t \xi(s) \,\mathrm{d}\widetilde{S}(s) - \xi(t) \cdot \widetilde{S}(t), \tag{4.1.7}$$

and thus, the riskless component of a self-financing trading strategy is uniquely determined by the initial investment $V(0)$ and the risky component ξ. In particular, if ξ is a locally admissible integrand for S and $w \in \mathbb{R}$, then there exist a real-valued function η such that the pair (ξ, η) is a self-financing trading strategy with $V(0) = w$.

Proof. Again, without loss of generality assume $d = 1$. Since ξ is locally admissible for S, for every $T > 0$ there exists a partition $\mathbb{T} = \{t_0, \ldots, t_N\}$ of $[0, T]$ such that $\xi(t) = \sum_{k=1}^N \partial_X F_k(t, S_k^t, A_k^t)\mathbb{1}_{t\in[t_{k-1},t_k)}$ with A_k, F_k as in Definition 4.1.1. Let B_k denote the restriction of B to $[t_{k-1}, t_k]$, then ξ can be written as $\xi(t) = \sum_{k=1}^N \partial_X \widehat{F}_k(t, \widetilde{S}_k^t, \widehat{A}_k^t)\mathbb{1}_{t\in[t_{k-1},t_k)}$, with $\widehat{A}_k = (A_k, B_k)$ and $\widehat{F}_k(t, X_k, \widehat{A}_k) = F_k(t, B_k X_k, A_k)/B_k(t)$. Here, $B_k X_k$ means pointwise multiplication. Hence, ξ is locally admissible for $\widetilde{S}$.

To see the equivalence of (4.1.3) and (4.1.6), we first assume that (ξ, η) is self-financing in the sense of (4.1.3). Then,

$$\begin{aligned}
\widetilde{V}(t) &= \widetilde{V}(0) + \int_0^t \frac{1}{B(s)}\,\mathrm{d}V(s) - \int_0^t \frac{1}{B(s)}\widetilde{V}(s)\,\mathrm{d}B(s)\\
&= \int_0^t \frac{1}{B(s)}\xi(s)\,\mathrm{d}S(s) + \int_0^t \frac{1}{B(s)}\eta(s)\,\mathrm{d}B(s)\\
&\quad - \int_0^t \frac{1}{B^2(s)}(\xi(s)S(s) + \eta(s)B(s))\,\mathrm{d}B(s)\\
&= \int_0^t \frac{1}{B(s)}\xi(s)\,\mathrm{d}S(s) - \int_0^t \frac{1}{B(s)}\xi(s)\widetilde{S}(s)\,\mathrm{d}B(s). \qquad (4.1.8)
\end{aligned}$$

Here, we have used Föllmer's pathwise Itô formula [46] applied to the function $f(v, a) = v \cdot a$ and to the paths V and $1/B$ in the first step. In the second step, we have used (4.1.3) and the associativity rule in functional pathwise Itô calculus, Theorem 3.3.1, and we have inserted the definition of $\widetilde{V}$. Using (4.1.5) and once again Theorem 3.3.1, we infer that

$$\int_0^t \frac{1}{B(s)}\xi(s)\,\mathrm{d}S(s) - \int_0^t \frac{1}{B(s)}\xi(s)\widetilde{S}(s)\,\mathrm{d}B(s) = \int_0^t \xi(s)\,\mathrm{d}\widetilde{S}(s), \qquad (4.1.9)$$

which is (4.1.6). To show the converse direction, we assume that (4.1.6) holds, apply (4.1.9), and reverse the steps in (4.1.8) so as to obtain (4.1.3).

To show the last statement, define

$$\widetilde{V}(t) := w + \int_0^t \xi(s)\,\mathrm{d}\widetilde{S}(s), \quad \eta(t) := \widetilde{V}(t) - \xi(t)\cdot\widetilde{S}(t).$$

The pathwise Itô integral on the right-hand side is well-defined, since for every $T > 0$ there exists a partition $\mathbb{T} = \{t_0, \ldots, t_N\}$ of $[0, T]$ such that

$$\xi(t) = \sum_{k=1}^N \nabla_X \widetilde{F}_k(t, \widetilde{S}_k^t, (A_k^t, B_k^t))\mathbb{1}_{t\in[t_{k-1},t_k)},$$

where $\widetilde{F}_k(t, X_k, (A_k, B_k)) = F_k(t, B_k X_k, A_k)/B_k(t)$ with A_k, F_k as in Definition 4.1.1, and (A_k, B_k) has components of bounded variation on $[0, T]$. It follows that $V(t) = \widetilde{V}(t)B(t)$ is the portfolio value of (ξ, η) and the above argument implies that (ξ, η) is self-financing. □

Hence, when dealing with self-financing trading strategies, it is sufficient to focus on the initial wealth w and the risky component ξ. Often it is convenient to describe such self-financing trading strategies by the vector $\pi(t) = (\pi_1(t), \dots, \pi_d(t))^\top$, where $\pi_i(t)$ denotes the proportion of the current wealth $V^\pi(t)$ that is invested into the i-th asset at time t, i.e.,

$$\xi_i(t) = \frac{\pi_i(t)V^\pi(t)}{S_i(t)}, \quad i = 1, \dots, d, \quad \text{and} \quad \eta(t) = \frac{\left(1 - \sum_{i=1}^d \pi_i(t)\right) V^\pi(t)}{B(t)}. \tag{4.1.10}$$

Taking this point of view, the trading strategy (4.1.10) will be self-financing if and only if the associated wealth, V^π, satisfies the Itô differential equation

$$\mathrm{d}V^\pi(t) = V^\pi(t)\, \frac{\pi(t)}{S(t)}\, \mathrm{d}S(t) + V^\pi(t)\, \frac{\left(1 - \sum_{i=1}^d \pi_i(t)\right)}{B(t)}\, \mathrm{d}B(t), \tag{4.1.11}$$

where

$$\frac{\pi(t)}{S(t)} := \left(\frac{\pi_1(t)}{S_1(t)}, \dots, \frac{\pi_d(t)}{S_d(t)}\right)^\top.$$

The corresponding code is given in Table A.1.

By Theorem 3.4.1, it follows that

$$\begin{aligned} V^\pi(t) &= V(0) \cdot \mathcal{E}\left(\int_0^\cdot \frac{\pi(s)}{S(s)}\, \mathrm{d}S(s)\right) \exp\left(\int_0^t (1 - \sum_{i=1}^d \pi_i(s)) r(s)\, \mathrm{d}s\right) \\ &= V(0) \cdot \exp\left(\int_0^t \frac{\pi(s)}{S(s)}\, \mathrm{d}S(s) - \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\pi_i(s)\pi_j(s)}{S_i(s)S_j(s)}\, \mathrm{d}[S_i, S_j](s) + \int_0^t (1 - \sum_{i=1}^d \pi_i(s)) r(s)\, \mathrm{d}s\right) \end{aligned} \tag{4.1.12}$$

for any π such that π/S is a locally admissible integrand for S and the second and third integrals in (4.1.12) exist as Riemann-Stieltjes integrals. If we disallow borrowing from or investing in the money market, we obtain the following definition.

Definition 4.1.6. An $\mathbb{R}^d$-valued measurable function π is called a *portfolio* if π/S is a locally admissible integrand for S and the second and third integrals in (4.1.12) exist as Riemann-Stieltjes integrals, and if

$$\pi_1(t) + \cdots + \pi_d(t) = 1, \quad t \geq 0. \tag{4.1.13}$$

A portfolio is called *long-only* if $\pi_i(t) \geq 0$ for all i and all t.

A long-only portfolio thus takes values in the simplex in $\mathbb{R}^d$,

$$\Delta^d = \{(\pi_1, \dots, \pi_d)^\top \in \mathbb{R}^d \mid \pi_1 \geq 0 \dots \pi_d \geq 0 \text{ and } \pi_1 + \cdots + \pi_d = 1\}.$$

In the following, we will also need the following notation

$$\Delta^d_+ = \{(\pi_1, \dots, \pi_d)^\top \in \Delta^d \mid \pi_1 > 0 \dots \pi_d > 0\}.$$

Moreover, we will denote by $V^{w,\pi}$ the wealth associated to a portfolio π with initial investment w.

As in [45, Section 2], we normalize the market, i.e., we suppose that at any time t each stock has only one share outstanding. Then, the stock prices $S_i(t)$ represent the capitalizations of the individual companies, and the quantities

$$S^{\mathrm{Total}}(t) := S_1(t) + \cdots + S_d(t) \quad \text{and} \quad \mu_i(t) := \frac{S_i(t)}{S^{\mathrm{Total}}(t)}, \quad i = 1, \ldots, d, \tag{4.1.14}$$

correspond to the total capitalization of the market and the relative capitalizations of the individual firms, which are called the respective *market weights*. Clearly, we have

$$0 < \mu_i(t) < 1 \quad \text{for all} \quad i = 1, \ldots, d \quad \text{and} \quad \sum_{i=1}^{d} \mu_i(t) = 1. \tag{4.1.15}$$

By (4.1.11), the associated portfolio value satisfies

$$\mathrm{d}V^{w,\mu}(t) = V^{w,\mu}(t) \frac{\mu(t)}{S(t)} \,\mathrm{d}S(t),$$

or, equivalently,

$$\frac{\mathrm{d}V^{w,\mu}(t)}{V^{w,\mu}(t)} = \frac{\mu(t)}{S(t)} \,\mathrm{d}S(t) = \frac{\mathrm{d}S^{\mathrm{Total}}(t)}{S^{\mathrm{Total}}(t)},$$

since

$$\frac{\mu_i(t)}{S_i(t)} = \frac{1}{S^{\mathrm{Total}}(t)} \quad \text{and} \quad \sum_{i=1}^{d} \mathrm{d}S_i(t) = \mathrm{d}\left(\sum_{i=1}^{d} S_i(t) \right) = \mathrm{d}S^{\mathrm{Total}}(t).$$

With $V^{w,\mu}(0) = w$ this gives us

$$V^{w,\mu}(t) = \frac{w}{S^{\mathrm{Total}}(0)} S^{\mathrm{Total}}(t), \tag{4.1.16}$$

whence we infer that

$$\frac{\mu_i(t) V^{w,\mu}(t)}{S_i(t)} = \frac{V^{w,\mu}(t)}{S^{\mathrm{Total}}(t)} = \frac{w}{S^{\mathrm{Total}}(0)}$$

is the number of shares in the i-th asset μ holds at time t (for $i = 1, \ldots, d$). The corresponding code is given in Table A.2.

In our model-free version of portfolio theory, we do not wish to make assumptions on the structure of the covariations $[S_i, S_j]$ apart from the fact that the sequence in (4.1.2) converges for all t and $[S_i, S_j]$(t) is continuous. In particular, we do not assume that $[S_i, S_j](t)$ is *absolutely* continuous in t. Growth rates and covariances, which are functions in [45], therefore need to be modeled as measures in our strictly pathwise setting.

Definition 4.1.7. The *covariance* of the stocks in the market is described by the positive semidefinite matrix-valued Radon measure $a = (a_{ij})_{1 \leq i,j \leq d}$ defined as

$$a_{ij}(\mathrm{d}t) := \mathrm{d}[\log S_i, \log S_j](t) = \frac{1}{S_i(t) S_j(t)} \,\mathrm{d}[S_i, S_j](t), \quad i, j = 1, \ldots, d. \tag{4.1.17}$$

The *excess growth rate* of the portfolio π is defined as the signed Radon measure

$$\gamma_\pi^*(\mathrm{d}t) := \frac{1}{2}\left(\sum_{i=1}^d \pi_i(t)a_{ii}(\mathrm{d}t) - \sum_{i,j=1}^d \pi_i(t)\pi_j(t)a_{ij}(\mathrm{d}t)\right). \tag{4.1.18}$$

For any portfolio π, we define the *covariances* of the individual stocks *relative to the portfolio* π as follows for $i, j = 1, \ldots, d$,

$$\tau_{ij}^\pi(\mathrm{d}t) : = (\pi(t) - e_i)^\top a(\mathrm{d}t)(\pi(t) - e_j). \tag{4.1.19}$$

It follows from Lemma 4.2.1 below that $\tau^\pi = \left(\tau_{ij}^\pi\right)$ is always a positive semidefinite matrix-valued Radon measure. Moreover, if π is long-only, Lemma 4.2.2 gives that the excess growth rate is a positive Radon measure. Note that

$$\tau_{ij}^\pi(\mathrm{d}t) = a_{ij}(\mathrm{d}t) - \sum_{j=1}^d \pi_j(t)a_{ij}(\mathrm{d}t) - \sum_{i=1}^d \pi_i(t)a_{ij}(\mathrm{d}t) + \sum_{i,j=1}^d \pi_i(t)\pi_j(t)a_{ij}(\mathrm{d}t). \tag{4.1.20}$$

Thus, the relative covariances of the individual stocks from (4.1.19) satisfy the elementary property

$$\sum_{j=1}^d \pi_j(t)\tau_{ij}^\pi(\mathrm{d}t) = 0, \quad i = 1, \ldots, d. \tag{4.1.21}$$

More precisely, (4.1.20) yields for any $i = 1, \ldots, d$,

$$\begin{aligned}\sum_{j=1}^d \pi_j(t)\tau_{ij}^\pi(\mathrm{d}t) &= \sum_{j=1}^d \pi_j(t)a_{ij}(\mathrm{d}t) - \sum_{j=1}^d\sum_{k=1}^d \pi_j(t)\pi_k(t)a_{ik}(\mathrm{d}t)\\ &\quad - \sum_{j=1}^d\sum_{k=1}^d \pi_j(t)\pi_k(t)a_{jk}(\mathrm{d}t) + \sum_{i=1}^d\sum_{j=1}^d \pi_i(t)\pi_j(t)a_{ij}(\mathrm{d}t).\end{aligned}$$

Since

$$\sum_{j=1}^d\sum_{k=1}^d \pi_j(t)\pi_k(t)a_{ik}(\mathrm{d}t) = \sum_{k=1}^d \pi_k(t)a_{ik}(\mathrm{d}t),$$

thanks to the fact that $\pi_1(t) + \cdots + \pi_d(t) = 1$, we get

$$\begin{aligned}\sum_{j=1}^d \pi_j(t)\tau_{ij}^\pi(\mathrm{d}t) &= \sum_{j=1}^d \pi_j(t)a_{ij}(\mathrm{d}t) - \sum_{j=1}^d\sum_{k=1}^d \pi_j(t)\pi_k(t)a_{ik}(\mathrm{d}t)\\ &= \sum_{j=1}^d \pi_j(t)a_{ij}(\mathrm{d}t) - \sum_{k=1}^d \pi_k(t)a_{ik}(\mathrm{d}t)\\ &= 0.\end{aligned}$$

The next lemma gives the logarithmic representation of the wealth $V^\pi(t) := V^{1,\pi}(t)$ corresponding to the portfolio π with the initial investment $w = \$1$. It also provides a motivation for calling γ_π^* the "excess growth rate" of the portfolio π.

Lemma 4.1.8. *The return from a one-unit investment according to the portfolio* π *is given by*

$$\mathrm{d}\log V^{\pi}(t) = \pi(t)\,\mathrm{d}\log S(t) + \gamma^*_{\pi}(\mathrm{d}t), \tag{4.1.22}$$

where $\log S(t)$ *denotes the vector of the log-prices.*

Proof. On the one hand, Föllmer's pathwise Itô formula [46], applied to $f(x) = \log x$ and the paths S_i individually, yields

$$\mathrm{d}\log S_i(t) = \frac{1}{S_i(t)}\,\mathrm{d}S_i(t) - \frac{1}{2(S_i(t))^2}\,\mathrm{d}[S_i](t).$$

On the other hand, (4.1.12) implies, in conjunction with the portfolio condition (4.1.13),

$$\mathrm{d}\log V^{\pi}(t) = \frac{\pi(t)}{S(t)}\,\mathrm{d}S(t) - \frac{1}{2}\sum_{i,j=1}^{d}\frac{\pi_i(t)\pi_j(t)}{S_i(t)S_j(t)}\,\mathrm{d}[S_i, S_j](t).$$

Combining the above two equations we infer, using the associativity of the Stieltjes integral from [96, Theorem I.6 b] and the associativity of the pathwise functional Itô integral, Theorem 3.3.1,

$$\mathrm{d}\log V^{\pi}(t) = \pi(t)\,\mathrm{d}\log S(t) + \frac{1}{2}\sum_{i=1}^{d}\pi_i(t)\,\mathrm{d}[\log S_i](t) - \frac{1}{2}\sum_{i,j=1}^{d}\pi_i(t)\pi_j(t)\,\mathrm{d}[\log S_i, \log S_j](t),$$

which implies the assertion via the definition of γ^*_{π}. □

In particular, equation (4.1.22) yields the following dynamics for the market weights

$$\mathrm{d}\log \mu_i(t) = \big(e_i - \mu(t)\big)\,\mathrm{d}\log S(t) - \gamma^*_{\mu}(\mathrm{d}t). \tag{4.1.23}$$

Lemma 4.1.9. *Equivalently,* (4.1.23) *can be written as*

$$\frac{\mathrm{d}\mu_i(t)}{\mu_i(t)} = \big(e_i - \mu(t)\big)\,\mathrm{d}\log S(t) - \gamma^*_{\mu}(\,\mathrm{d}t) + \frac{1}{2}\tau^{\mu}_{ii}(\,\mathrm{d}t), \quad i = 1, \ldots, d. \tag{4.1.24}$$

Proof. Let $i \in \{1, \ldots, d\}$ be given. In the first step, observe that Proposition 3.2.7 implies, in conjunction with (4.1.23) and [80, Remark 8],

$$\begin{aligned}
\mathrm{d}[\log \mu_i](t) &= \mathrm{d}\Big[\int_0^{\cdot} \big(e_i - \mu(s)\big)\,\mathrm{d}\log S(s)\Big](t) \\
&= \sum_{k,l=1}^{d} \big((e_i)_k - \mu_k(t)\big)\big((e_i)_l - \mu_l(t)\big)\,\mathrm{d}[\log S_k, \log S_l](t) \\
&= \big(\mu(t) - e_i\big)^{\top} a(\,\mathrm{d}t)\big(\mu(t) - e_i\big) = \tau^{\mu}_{ii}(\,\mathrm{d}t).
\end{aligned} \tag{4.1.25}$$

Analogously (see also Lemma 4.2.1), it can be inferred that

$$\tau^{\mu}_{ij}(\,\mathrm{d}t) = \mathrm{d}[\log \mu_i, \log \mu_j](t) = \frac{1}{\mu_i(t)\mu_j(t)}\,\mathrm{d}[\mu_i, \mu_j](t), \quad i, j = 1, \ldots, d. \tag{4.1.26}$$

Thus, (4.1.23) becomes

$$\mu_i(t) = \exp\left(\int_0^t \big(e_i - \mu(s)\big)\,\mathrm{d}\log S(s) - \frac{1}{2}\int_0^t \tau_{ii}^{\mu}(\,\mathrm{d}s)\right)\cdot \exp\left(\frac{1}{2}\int_0^t \tau_{ii}^{\mu}(\,\mathrm{d}s) - \int_0^t \gamma_{\mu}^{*}(\mathrm{d}s)\right).$$

Denoting $I(t) = \int_0^t \big(e_i - \mu(s)\big)\,\mathrm{d}\log S(s)$, we infer with (4.1.25) that $[I](t) = \int_0^t \tau_{ii}^{\mu}(\,\mathrm{d}s)$. Thus, applying Föllmer's pathwise Itô formula [46] to the function $f(k,l) = e^{k+l}$ and to the paths

$$K(t) := I(t) - \frac{1}{2}[I](t), \quad L(t) := \frac{1}{2}\int_0^t \tau_{ii}^{\mu}(\,\mathrm{d}s) - \int_0^t \gamma_{\mu}^{*}(\mathrm{d}s)$$

yields that

$$\begin{aligned}\mu_i(t) &= \mu_i(0) + \int_0^t \mu_i(s)\,\mathrm{d}K(s) + \frac{1}{2}\int_0^t \mu_i(s)\,\mathrm{d}[K](s) + \int_0^t \mu_i(s)\,\mathrm{d}L(s)\\ &= \mu_i(0) + \int_0^t \mu_i(s)\big(e_i - \mu(s)\big)\,\mathrm{d}\log S(s) - \int_0^t \mu_i(s)\gamma_{\mu}^{*}(\mathrm{d}s) + \frac{1}{2}\int_0^t \mu_i(s)\tau_{ii}^{\mu}(\,\mathrm{d}s),\end{aligned}$$

where we have used the associativity rule for pathwise functional Itô calculus in the form of Theorem 3.3.1, with $\eta = \mu_i$ and $\xi = e_i - \mu$, the associativity of the Stieltjes integral from [96, Theorem I.6 b], and the fact that $[K](t) = [I](t)$, as can be seen from [80, Remark 8]. But this is what we had to show. □

4.2 Derivation of the pathwise functional master formula

Let us first collect some useful properties of the relative covariances τ_{ij}^{π} of the individual stocks from (4.1.19). Note that the proofs of the following lemmas are slight adaptations of the respective proofs from [45, Section 3] to our functional pathwise setting. For every stock i and every portfolio π, we denote by

$$R_i^{\pi}(t) := \log\left(\frac{S_i(t)}{V^{w,\pi}(t)}\right)\Bigg|_{w=S_i(0)}, \quad 0 \le t < \infty,$$

the relative return of the i-th stock with respect to the portfolio π.

Lemma 4.2.1. *For every portfolio π, for all $1 \le i,j \le d$, we have*

$$\tau_{ij}^{\pi}(\mathrm{d}t) = \mathrm{d}[R_i^{\pi}, R_j^{\pi}](t); \quad \textit{in particular,} \quad \tau_{ii}^{\pi}(\mathrm{d}t) = \mathrm{d}[R_i^{\pi}](t) \ge 0,$$

and $\tau^{\pi} = (\tau_{ij}^{\pi})_{1\le i,j\le d}$ is a positive semidefinite matrix-valued Radon measure.

Proof. As above, we get with the dynamics (4.1.22),

$$\begin{aligned}\mathrm{d}R_i^{\pi}(t) = {}& \mathrm{d}\log\left(\frac{S_i(t)}{V^{w,\pi}(t)}\right)\Bigg|_{w=S_i(0)} = \mathrm{d}\log S_i(t) - \mathrm{d}\log V^{w,\pi}(t)\\ &= \mathrm{d}\log S_i(t) - \pi(t)\,\mathrm{d}\log S(t) - \gamma_{\pi}^{*}(\,\mathrm{d}t)\\ &= \big(e_i - \pi(t)\big)\,\mathrm{d}\log S(t) - \gamma_{\pi}^{*}(\,\mathrm{d}t).\end{aligned}$$

Thus, using [80, Remark 8], we infer that the vector of the relative returns $R^\pi = (R_1^\pi \dots, R_d^\pi)^\top$ admits the continuous covariations

$$[R_i^\pi, R_j^\pi](t) = \left[\int_0^\cdot \big(e_i - \pi(s)\big) \,\mathrm{d} \log S(s), \int_0^\cdot \big(e_j - \pi(s)\big) \,\mathrm{d} \log S(s) \right](t). \tag{4.2.1}$$

As in the proof of Lemma 4.1.9, we thus get for $i, j = 1, \dots, d$,

$$\begin{aligned} \mathrm{d}[R_i^\pi, R_j^\pi](t) &= \sum_{k,l=1}^d \big((e_i)_k - \pi_k(t)\big)\big((e_j)_l - \pi_l(t)\big) \,\mathrm{d}[\log S_k, \log S_l](t) \\ &= \sum_{k,l=1}^d \left(\pi(t) - e_i\right)_k \left(\pi(t) - e_j\right)_l a_{kl}(\,\mathrm{d}t) \\ &= \big(\pi(t) - e_i\big)^\top a(\,\mathrm{d}t)\big(\pi(t) - e_j\big) = \tau_{ij}^\pi(\,\mathrm{d}t), \end{aligned} \tag{4.2.2}$$

which concludes the proof. □

Lemma 4.2.2. *For any pair of portfolios π and ρ we have the following* numéraire invariance property

$$\gamma_\pi^*(\mathrm{d}t) = \frac{1}{2} \left(\sum_{i=1}^d \pi_i(t) \tau_{ii}^\rho(\mathrm{d}t) - \sum_{i=1}^d \sum_{j=1}^d \pi_i(t)\pi_j(t)\tau_{ij}^\rho(\mathrm{d}t) \right). \tag{4.2.3}$$

In particular, the excess growth rate of a portfolio π can be represented as a weighted average of the variances τ_{ii}^π of the individual stocks relative to the portfolio π, i.e.,

$$\gamma_\pi^*(\mathrm{d}t) = \frac{1}{2} \sum_{i=1}^d \pi_i(t)\tau_{ii}^\pi(\mathrm{d}t). \tag{4.2.4}$$

Furthermore, for any long-only portfolio π we get

$$\gamma_\pi^*(\mathrm{d}t) \geq 0.$$

Proof. Denoting

$$a_{\rho i}(\,\mathrm{d}t) = \sum_{j=1}^d \rho_j(t) a_{ij}(\mathrm{d}t) \quad \text{and} \quad a_{\rho\rho}(\,\mathrm{d}t) = \sum_{i,j=1}^d \rho_i(t)\rho_j(t)a_{ij}(\,\mathrm{d}t),$$

(4.1.20) yields

$$\begin{aligned} \sum_{i=1}^d \pi_i(t)\tau_{ii}^\rho(\,\mathrm{d}t) &= \sum_{i=1}^d \pi_i(t)\big(a_{ii}(\,\mathrm{d}t) - a_{\rho i}(\,\mathrm{d}t) - a_{\rho i}(\,\mathrm{d}t) + a_{\rho\rho}(\,\mathrm{d}t)\big) \\ &= \sum_{i=1}^d \pi_i(t)a_{ii}(\,\mathrm{d}t) - 2\sum_{i=1}^d \pi_i(t)a_{\rho i}(\,\mathrm{d}t) + a_{\rho\rho}(\,\mathrm{d}t), \end{aligned}$$

thanks to $\sum_{i=1}^d \pi_i(t) = 1$. Analogously,

$$\begin{aligned}\sum_{i=1}^d \sum_{j=1}^d \pi_i(t)\pi_j(t)\tau_{ij}^\rho(\,\mathrm{d}t) &= \sum_{i=1}^d \pi_i(t)\pi_j(t)\big(a_{ij}(\,\mathrm{d}t) - a_{\rho i}(\,\mathrm{d}t) - a_{\rho j}(\,\mathrm{d}t) + a_{\rho\rho}(\,\mathrm{d}t)\big)\\ &= \sum_{i,j=1}^d \pi_i(t)\pi_j(t)a_{ij}(\,\mathrm{d}t) - 2\sum_{i=1}^d \pi_i(t)a_{\rho i}(\,\mathrm{d}t) + a_{\rho\rho}(\,\mathrm{d}t).\end{aligned}$$

With the definition of the excess growth rate in (4.1.18), we thus infer

$$\begin{aligned}\gamma_\pi^*(\,\mathrm{d}t) &= \frac{1}{2}\left(\sum_{i=1}^d \pi_i(t)a_{ii}(\,\mathrm{d}t) - \sum_{i,j=1}^d \pi_i(t)\pi_j(t)a_{ij}(\,\mathrm{d}t)\right)\\ &= \frac{1}{2}\left(\sum_{i=1}^d \pi_i(t)\tau_{ii}^\rho(\,\mathrm{d}t) + 2\sum_{i=1}^d \pi_i(t)a_{\rho i}(\,\mathrm{d}t) - a_{\rho\rho}(\,\mathrm{d}t)\right)\\ &\quad - \frac{1}{2}\left(\sum_{i=1}^d \sum_{j=1}^d \pi_i(t)\pi_j(t)\tau_{ij}^\rho(\,\mathrm{d}t) + 2\sum_{i=1}^d \pi_i(t)a_{\rho i}(\,\mathrm{d}t) - a_{\rho\rho}(\,\mathrm{d}t)\right),\end{aligned}$$

which shows (4.2.3).

Equality (4.2.4) directly follows from (4.2.3) if we take $\rho \equiv \pi$, thanks to (4.1.21):

$$\begin{aligned}\gamma_\pi^*(\,\mathrm{d}t) &= \frac{1}{2}\left(\sum_{i=1}^d \pi_i(t)\tau_{ii}^\pi(\,\mathrm{d}t) - \sum_{i=1}^d \sum_{j=1}^d \pi_i(t)\pi_j(t)\tau_{ij}^\pi(\,\mathrm{d}t)\right)\\ &= \frac{1}{2}\left(\sum_{i=1}^d \pi_i(t)\tau_{ii}^\pi(\,\mathrm{d}t) - \sum_{i=1}^d \pi_i(t)\sum_{j=1}^d \pi_j(t)\tau_{ij}^\pi(\,\mathrm{d}t)\right)\\ &= \frac{1}{2}\sum_{i=1}^d \pi_i(t)\tau_{ii}^\pi(\,\mathrm{d}t).\end{aligned}$$

Since $\tau_{ii}^\pi(\,\mathrm{d}t) \geq 0$, due to Lemma 4.2.1, this implies $(\gamma_\pi^*)\,(\,\mathrm{d}t) \geq 0$ for any long-only portfolio π. □

Remark 4.2.3. We infer with the dynamics (4.1.22) that

$$\mathrm{d}\log\left(\frac{V^\pi(t)}{V^\mu(t)}\right) = (\pi(t) - \mu(t))\,\mathrm{d}\log S(t) + (\gamma_\pi^* - \gamma_\mu^*)(\mathrm{d}t). \tag{4.2.5}$$

On the other hand, using Lemma 4.1.9 and the associativity of the Stieltjes integral from [96, Theorem I.6 b] together with the associativity of the pathwise functional Itô integral, Theorem 3.3.1, we obtain

$$\frac{\pi(t)}{\mu(t)}\,\mathrm{d}\mu(t) = \big(\pi(t) - \mu(t)\big)\,\mathrm{d}\log S(t) - \gamma_\mu^*(\,\mathrm{d}t) + \frac{1}{2}\sum_{i=1}^d \pi_i(t)\tau_{ii}^\mu(\,\mathrm{d}t),$$

thanks to the fact that the portfolio weights sum up to one. Furthermore, applying the numéraire invariance property from Lemma 4.2.2 gives us

$$\frac{\pi(t)}{\mu(t)}\,\mathrm{d}\mu(t) = \big(\pi(t)-\mu(t)\big)\,\mathrm{d}\log S(t) - \gamma^*_\mu(\,\mathrm{d}t) + \frac{1}{2}\left(\sum_{i=1}^{d}\sum_{j=1}^{d}\pi_i(t)\pi_j(t)\tau^{\mu}_{ij}(\mathrm{d}t)\right) + \gamma^*_\pi(\,\mathrm{d}t).$$

Thus, we receive for any portfolio π the relative return formula

$$\mathrm{d}\log\left(\frac{V^\pi(t)}{V^\mu(t)}\right) = \frac{\pi(t)}{\mu(t)}\mathrm{d}\mu(t) - \frac{1}{2}\left(\sum_{i=1}^{d}\sum_{j=1}^{d}\pi_i(t)\pi_j(t)\tau^{\mu}_{ij}(\mathrm{d}t)\right), \tag{4.2.6}$$

or, equivalently,

$$\mathrm{d}\left(\frac{V^\pi(t)}{V^\mu(t)}\right) = \left(\frac{V^\pi(t)}{V^\mu(t)}\right)\frac{\pi(t)}{\mu(t)}\,\mathrm{d}\mu(t). \tag{4.2.7}$$

The unique solution of this Itô differential equation with initial condition $\frac{V^\pi(0)}{V^\mu(0)} = 1$, according to Theorem 3.4.1, is

$$\frac{V^\pi(t)}{V^\mu(t)} = \mathcal{E}\left(\int_0^{\cdot}\frac{\pi(s)}{\mu(s)}\,\mathrm{d}\mu(s)\right)(t). \tag{4.2.8}$$

We can now introduce the notion of *portfolio generating functionals*. These are smooth functionals that may depend on the entire past evolution of the trajectories $\mu_1,\dots,\mu_d$. We use the concepts and notation of functional pathwise Itô calculus from [34, 20] in the version of [82], as described in Chapter 3. Recall that Δ^d denotes the simplex in $\mathbb{R}^d$ and, similarly, $\Delta^d_+ = \{(\pi_1,\dots,\pi_d)^\top \in \Delta^d \mid \pi_1 > 0 \dots \pi_d > 0\}$. Also note that it will be convenient for us to use the following notation:

$$\nabla^2_X F(t,X^t,A^t) = \left(\partial^2_{ij}F(t,X^t,A^t)\right)_{i,j=1,\dots,d}.$$

Definition 4.2.4. Assume that we are given a non-anticipative functional $G : [0,T]\times\mathcal{V}^T\times\mathcal{W}^T_{BV} \mapsto (0,\infty)$, where $V \supset \Delta^d_+$ open, and $W \subset \mathbb{R}^m$ is a Borel subset of $\mathbb{R}^m$. We assume further that G is "sufficiently regular" in the sense that it is of class $\mathbb{C}^{1,2}([0,T])$ and satisfies the regularity conditions from Theorem 3.2.1. Then the portfolio π with weights

$$\pi_i(t) = \left[\partial_i\log G(t,\mu^t,A^t_\mu) + 1 - \sum_{j=1}^{d}\mu_j(t)\partial_j\log G(t,\mu^t,A^t_\mu)\right]\mu_i(t), \quad 1\le i\le d, \tag{4.2.9}$$

is called the *portfolio generated by G*. Here, A is a CBV^m-functional, i.e., the map $A : C([0,T],V) \ni X \mapsto A_X \in \mathcal{W}^T_{CBV}$ is such that $A_X(t),\ t\in[0,T]$, is a function of t and $(X(s))_{s\le t}$. An example would be the running maximum, $A_X(t) = \max_{s\le t}|X(s)|$.

The following questions arise:

- What is the relation between the wealth of the market portfolio and the functionally generated portfolio (4.2.9)?

- Is it possible to find descriptive conditions on the market structure that will allow us to construct portfolios with a "nice" behavior with respect to the market?

The next theorem provides the required tools for answering the above questions.

Theorem 4.2.5 (Functional Master equation)**.** *The relative wealth of the portfolio π generated by G, with respect to the market, is given by the following* functional master equation

$$\log\left(\frac{V^{\pi}(T)}{V^{\mu}(T)}\right) = \log\left(\frac{G(T,\mu^T,A_\mu^T)}{G(0,\mu^0,A_\mu^0)}\right) + \mathfrak{g}([0,T]) + \mathfrak{h}([0,T]), \quad 0 \le T < \infty, \tag{4.2.10}$$

where

$$\mathfrak{g}(\mathrm{d}t) := -\frac{1}{2G(t,\mu^t,A_\mu^t)}\sum_{i,j=1}^{d}\partial_{ij}^2 G(t,\mu^t,A_\mu^t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\mathrm{d}t) \tag{4.2.11}$$

is the second-order drift term, and

$$\mathfrak{h}(\mathrm{d}t) := -\mathcal{D}\log G(t,\mu^t,A_\mu^t)\nu(\mathrm{d}t) = -\frac{1}{G(t,\mu^t,A_\mu^t)}\mathcal{D}G(t,\mu^t,A_\mu^t)\nu(\mathrm{d}t) \tag{4.2.12}$$

is the horizontal drift term. Since A_μ has components of bounded variation, which correspond to finite measures ν_k, $k=1,\dots,m$, on $[0,T]$, we can write $\nu(\mathrm{d}s) := (\mathrm{d}s, A_{\mu,1}(\mathrm{d}s),\dots,A_{\mu,m}(\mathrm{d}s))^\top$. Thus, ν can be regarded as the measure associated to the function $A_\mu \in CBV^m([0,T])$.

Remark 4.2.6. Note that we do not have to know or estimate the volatility of the model in order to compute the second-order drift term: The master equation does it for us, in terms of quantities that are completely observable,

$$\mathfrak{g}([0,T]) = \log\left(\frac{V^{\pi}(T)G(0,\mu^0,A_\mu^0)}{V^{\mu}(T)G(T,\mu^T,A_\mu^T)}\right) - \mathfrak{h}([0,T]).$$

Also note that the weights (4.2.9) of the portfolio generated by G depend only on the trajectories of the market weights $\mu_1,\dots,\mu_d$, and not on the volatility structure of the model specified for the market. Thus, to implement the portfolio (4.2.9) we only have to know the evolution of the market weights until time t, and V^{π} is observable in time, only in terms of these market weights. That is, we do not need to estimate any parameters, and the portfolio generated by G is therefore not subject to uncertainty and the resulting model risk; in practice, this fact leads to an excellent performance of such portfolios. An analogous feature was also pointed out in [45] for the case of portfolio generating functions, however we emphasize that here we not only obtain as a result a path-by-path representation of the associated relative wealth, but all the corresponding derivations are completely probability-free.

Proof of Theorem 4.2.5. We introduce the notation

$$g_i(t) := \partial_i \log G(t,\mu^t,A_\mu^t) \quad \text{and} \quad N(t) := 1 - \sum_{j=1}^{d}\mu_j(t)g_j(t).$$

Then definition (4.2.9) becomes $\pi_i(t) = (g_i(t) + N(t))\mu_i(t),\ i = 1, \ldots, d$. Moreover, with $g(t) := (g_1(t), \ldots, g_d(t))^\top$, we obtain

$$\frac{\pi(t)}{\mu(t)}\,\mathrm{d}\mu(t) = g(t)\,\mathrm{d}\mu(t) + N(t)\cdot\,\mathrm{d}\left(\sum_{i=1}^{d}\mu_i(t)\right) = g(t)\,\mathrm{d}\mu(t)$$

and

$$\sum_{i=1}^{d}\sum_{j=1}^{d}\pi_i(t)\pi_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = \sum_{i=1}^{d}\sum_{j=1}^{d}g_i(t)g_j(t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t),$$

in conjunction with the elementary property (4.1.21). More precisely, we have

$$\begin{aligned}\sum_{i=1}^{d}\sum_{j=1}^{d}\pi_i(t)\pi_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) &= \sum_{i=1}^{d}\sum_{j=1}^{d}\big(g_i(t)+N(t)\big)\big(g_j(t)+N(t)\big)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t)\\ &= \sum_{i=1}^{d}\sum_{j=1}^{d}g_i(t)g_j(t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t),\end{aligned}$$

since (4.1.21) implies that

$$\sum_{i=1}^{d}\sum_{j=1}^{d}N(t)g_i(t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = \sum_{i=1}^{d}N(t)g_i(t)\mu_i(t)\sum_{j=1}^{d}\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = 0,$$

and analogously

$$\sum_{i=1}^{d}\sum_{j=1}^{d}N(t)g_j(t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = \sum_{j=1}^{d}N(t)g_j(t)\mu_j(t)\sum_{i=1}^{d}\mu_i(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = 0,$$

as well as

$$\sum_{i=1}^{d}\sum_{j=1}^{d}(N(t))^2\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = (N(t))^2\sum_{i=1}^{d}\mu_i(t)\sum_{j=1}^{d}\mu_j(t)\tau_{ij}^{\mu}(\,\mathrm{d}t) = 0.$$

Thus, formula (4.2.6) from Remark 4.2.3 gives us

$$\mathrm{d}\log\left(\frac{V^{\pi}(t)}{V^{\mu}(t)}\right) = g(t)\mathrm{d}\mu(t) - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}g_i(t)g_j(t)\mu_i(t)\mu_j(t)\tau_{ij}^{\mu}(\mathrm{d}t). \tag{4.2.13}$$

On the other hand, the chain rule for vertical derivatives from Lemma 3.3.3 yields

$$\begin{aligned}\partial_{ij}^2\log G(t,X^t,A^t) &= \partial_j\left(\partial_i\log G(t,X^t,A^t)\right) = \partial_j\left(\frac{\partial_i G(t,X^t,A^t)}{G(t,X^t,A^t)}\right)\\ &= \frac{\partial_{ij}^2 G(t,X^t,A^t)\cdot G(t,X^t,A^t) - \partial_i G(t,X^t,A^t)\cdot\partial_j G(t,X^t,A^t)}{(G(t,X^t,A^t))^2}\\ &= \frac{\partial_{ij}^2 G(t,X^t,A^t)}{G(t,X^t,A^t)} - \frac{\partial_i G(t,X^t,A^t)}{G(t,X^t,A^t)}\cdot\frac{\partial_j G(t,X^t,A^t)}{G(t,X^t,A^t)}\\ &= \frac{\partial_{ij}^2 G(t,X^t,A^t)}{G(t,X^t,A^t)} - \partial_i\log G(t,X^t,A^t)\cdot\partial_j\log G(t,X^t,A^t),\end{aligned}$$

which gives us

$$\partial^2_{ij} \log G(t,\mu^t,A^t_\mu) = \frac{\partial^2_{ij} G(t,\mu^t,A^t_\mu)}{G(t,\mu^t,A^t_\mu)} - \partial_i \log G(t,\mu^t,A^t_\mu)\, \partial_j \log G(t,\mu^t,A^t_\mu).$$

Thus, the change of variables formula, Theorem 3.2.1, implies

$$\begin{aligned}
\mathrm{d} \log G(t,\mu^t,A^t_\mu) &= \nabla_X \log G(t,\mu^t,A^t_\mu) \mathrm{d}\mu(t) + \frac{1}{2} \sum_{i,j=1}^{d} \partial^2_{ij} \log G(t,\mu^t,A^t_\mu)\, \mathrm{d}[\mu]_{ij}(t) \\
&\quad + \mathcal{D} \log G(t,\mu^t,A^t_\mu) \nu(\mathrm{d}t) \\
&= g(t)\mathrm{d}\mu(t) + \frac{1}{2} \sum_{i,j=1}^{d} \left(\frac{\partial^2_{ij} G(t,\mu^t,A^t_\mu)}{G(t,\mu^t,A^t_\mu)} - g_i(t) g_j(t) \right) \mu_i(t)\mu_j(t) \tau^\mu_{ij}(\mathrm{d}t) \\
&\quad + \mathcal{D} \log G(t,\mu^t,A^t_\mu) \nu(\mathrm{d}t),
\end{aligned} \tag{4.2.14}$$

in conjunction with (4.1.26). Using (4.2.13) we infer that (4.2.14) equals

$$\begin{aligned}
&\mathrm{d} \log \left(\frac{V^\pi(t)}{V^\mu(t)} \right) + \frac{1}{2G(t,\mu^t,A^t_\mu)} \sum_{i=1}^{d} \sum_{j=1}^{d} \partial^2_{ij} G(t,\mu^t,A^t_\mu) \mu_i(t)\mu_j(t) \tau^\mu_{ij}(\mathrm{d}t) \\
&\qquad + \mathcal{D} \log G(t,\mu^t,A^t_\mu) \nu(\mathrm{d}t) = \mathrm{d} \log \left(\frac{V^\pi(t)}{V^\mu(t)} \right) - \mathfrak{g}(\,\mathrm{d}t) - \mathfrak{h}(\,\mathrm{d}t),
\end{aligned}$$

which implies the assertion, since $\log\left(\frac{V^\pi(0)}{V^\mu(0)}\right) = 0$. □

4.3 Examples and backtests

Next we discuss simulation results in order to get some basic idea about the behavior of portfolio generating functionals and their associated portfolios. In the Appendix, we also provide the codes for the functional entropy weighting example. In the following, we will work with convex combinations $\widetilde{X}(t) := \alpha X(t) + (1-\alpha)\vartheta^*(t)$, where $\alpha \in (0,1)$, of the (underlying) path itself and its (modified) moving average ϑ^*, defined by

$$\vartheta^*_i(t) := \begin{cases} \frac{1}{\delta} \int\limits_0^t X_i(s)\,\mathrm{d}s + \frac{1}{\delta} \int\limits_{t-\delta}^0 X_i(0)\,\mathrm{d}s, & t \in [0,\delta), \\ \frac{1}{\delta} \int\limits_{t-\delta}^t X_i(s)\,\mathrm{d}s, & t \in [\delta, T], \end{cases} \qquad i = 1,\dots,d,$$

for $\delta > 0$. More precisely, letting $g(x,a)$ be a smooth function, we will consider generating functionals of the form $G(t,X^t,A^t) = g(X(t),A(t))$ for the special choice of $A(t) = \vartheta^*(t)$.

Example 4.3.1 (Geometric mean)**.** Consider the functional

$$G(t,X^t,A^t) = \begin{cases} \prod\limits_{k=1}^{d} \left[\alpha X_k(t) + (1-\alpha) \left(\frac{1}{\delta} \int\limits_0^t X_k(s)\,\mathrm{d}s - \frac{1}{\delta}(t-\delta) X_k(0) \right) \right]^{\frac{1}{d}}, & 0 \le t < \delta, \\ \prod\limits_{k=1}^{d} \left[\alpha X_k(t) + \frac{1-\alpha}{\delta} \int\limits_{t-\delta}^t X_k(s)\,\mathrm{d}s \right]^{\frac{1}{d}}, & 0 < \delta \le t. \end{cases}$$

Then, the vertical derivative is given by

$$\partial_i G(t, X^t, A^t) = \frac{\alpha}{d}\left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i}}^{d}\left(\widetilde{X}_k(t)\right)^{\frac{1}{d}} = \frac{\alpha}{d}\frac{G(t, X^t, A^t)}{\widetilde{X}_i(t)}.$$

Hence,

$$\partial_i \log G(t, X^t, A^t) = \frac{1}{G(t, X^t, A^t)}\partial_i G(t, X^t, A^t) = \frac{\alpha}{d}\frac{G(t, X^t, A^t)}{\widetilde{X}_i(t)}$$

and the portfolio generated by G has the weights

$$\begin{aligned}\pi_i(t) &= \left[\partial_i \log G(t, \mu^t, A^t_\mu) + 1 - \sum_{j=1}^{d}\mu_j(t)\partial_j \log G(t, \mu^t, A^t_\mu)\right]\mu_i(t)\\ &= \left[\frac{\alpha}{d\widetilde{\mu}_i(t)} + 1 - \sum_{j=1}^{d}\frac{\alpha\mu_j(t)}{d\widetilde{\mu}_j(t)}\right]\mu_i(t). \qquad (4.3.1)\end{aligned}$$

Furthermore, we compute for the second-order vertical derivative,

$$\begin{aligned}\partial^2_{ij} G(t, X^t, A^t) &= \begin{cases}\frac{\alpha}{d}\left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1}\frac{\alpha}{d}\left(\widetilde{X}_j(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i,j}}^{d}\left(\widetilde{X}_k(t)\right)^{\frac{1}{d}}, & j\neq i,\\ \frac{\alpha^2}{d}\left(\frac{1}{d}-1\right)\left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-2} \cdot \prod_{\substack{k=1\\k\neq i}}^{d}\left(\widetilde{X}_k(t)\right)^{\frac{1}{d}}, & j = i,\end{cases}\\ &= \begin{cases}\frac{\alpha^2}{d^2}\frac{1}{\widetilde{X}_i(t)\widetilde{X}_j(t)}G(t, X^t, A^t), & j\neq i,\\ \frac{\alpha^2}{d}\left(\frac{1}{d}-1\right)\frac{1}{\widetilde{X}^2_i(t)}G(t, X^t, A^t), & j = i.\end{cases}\end{aligned}$$

Therefore, the second-order drift term is given by

$$\begin{aligned}\mathfrak{g}(\mathrm{d}t) &= -\frac{1}{2G(t, \mu^t, A^t_\mu)}\sum_{i,j=1}^{d}\partial^2_{ij}G(t, \mu^t, A^t_\mu)\mu_i(t)\mu_j(t)\tau^\mu_{ij}(\mathrm{d}t)\\ &= \frac{\alpha^2}{2d}\left(\sum_{i=1}^{d}\frac{\mu^2_i(t)}{(\widetilde{\mu}_i(t))^2}\tau^\mu_{ii}(\mathrm{d}t) - \frac{1}{d}\sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\mu_i(t)\mu_j(t)}{\widetilde{\mu}_i(t)\widetilde{\mu}_j(t)}\tau^\mu_{ij}(\mathrm{d}t)\right).\end{aligned}$$

For the horizontal derivative, we have

$$\begin{aligned}\mathcal{D}G(t, X^t, A^t) &= \begin{cases}\sum_{i=1}^{d}\frac{1-\alpha}{\delta}(X_i(t-) - X_i(0))\frac{1}{d}\left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i}}^{d}\left(\widetilde{X}_k(t)\right)^{\frac{1}{d}}, & 0\le t<\delta,\\ \sum_{i=1}^{d}\frac{1-\alpha}{\delta}(X_i(t-) - X_i((t-\delta)-))\frac{1}{d}\left(\widetilde{X}_i(t)\right)^{\frac{1}{d}-1} \cdot \prod_{\substack{k=1\\k\neq i}}^{d}\left(\widetilde{X}_k(t)\right)^{\frac{1}{d}}, & 0<\delta\le t,\end{cases}\\ &= \begin{cases}\frac{1-\alpha}{\delta d}G(t, X^t, A^t)\sum_{i=1}^{d}(X_i(t-) - X_i(0))\frac{1}{\widetilde{X}_i(t)}, & 0\le t<\delta,\\ \frac{1-\alpha}{\delta d}G(t, X^t, A^t)\sum_{i=1}^{d}(X_i(t-) - X_i((t-\delta)-))\frac{1}{\widetilde{X}_i(t)}, & 0<\delta\le t.\end{cases}\end{aligned}$$

This implies that the horizontal drift term is given by

$$
\begin{aligned}
\mathfrak{h}(\mathrm{d}t) &= -\mathcal{D}\log G(t,\mu^t,A_\mu^t)\nu(\mathrm{d}t) = -\frac{1}{G(t,\mu^t,A_\mu^t)}\mathcal{D}G(t,\mu^t,A_\mu^t)\nu(\mathrm{d}t) \\
&= -\frac{1-\alpha}{d\delta}\sum_{i=1}^{d}\frac{\mathrm{d}t}{\widetilde{\mu}_i(t)}\cdot\begin{cases}\mu_i(t)-\mu_i(0), & 0\le t<\delta,\\ \mu_i(t)-\mu_i(t-\delta), & 0<\delta\le t.\end{cases}
\end{aligned}
$$

The master equation (4.2.10) then yields for the relative performance of this portfolio, with respect to the market,

$$
\begin{aligned}
\log\left(\frac{V^\pi(T)}{V^\mu(T)}\right) &= \log\left(\frac{\prod_{k=1}^{d}(\widetilde{\mu}_i(T))^{\frac{1}{d}}}{\prod_{k=1}^{d}(\mu_k(0))^{\frac{1}{d}}}\right)+\int_0^T\mathfrak{g}(\,\mathrm{d}t)+\int_0^T\mathfrak{h}(\,\mathrm{d}t), \qquad (4.3.2)\\
&= \log\left(\frac{\prod_{k=1}^{d}(\widetilde{\mu}_i(T))^{\frac{1}{d}}}{\prod_{k=1}^{d}(\mu_k(0))^{\frac{1}{d}}}\right)-\frac{\alpha^2}{2d^2}\sum_{\substack{i,j=1\\ j\neq i}}^{d}\int_{[0,T]}\frac{\mu_i(t)\mu_j(t)}{\widetilde{\mu}_i(t)\widetilde{\mu}_j(t)}\tau_{ij}^{\mu}(\mathrm{d}t)\\
&\quad -\frac{\alpha^2}{2d}\left(\frac{1}{d}-1\right)\sum_{i=1}^{d}\int_{[0,T]}\frac{\mu_i^2(t)}{(\widetilde{\mu}_i(t))^2}\tau_{ii}^{\mu}(\mathrm{d}t)-\frac{1-\alpha}{d\delta}\sum_{i=1}^{d}\int_{[0,\delta[}\frac{\mu_i(t)-\mu_i(0)}{\widetilde{\mu}_i(t)}\,\mathrm{d}t\\
&\quad -\frac{1-\alpha}{d\delta}\sum_{i=1}^{d}\int_{[\delta,T]}\frac{\mu_i(t)-\mu_i(t-\delta)}{\widetilde{\mu}_i(t)}\,\mathrm{d}t.
\end{aligned}
$$

The following two figures display the results of a simulation of such a geometrically weighted portfolio with the parameters $\delta = 60$ days and $\alpha = 0,7$. We used the stock data base from Reuters Datastream; our data included 10 years of daily values of the closing prices for the stocks that were in the DAX at the time point considered. In Figure 4.3.1 we see the relative performance of the portfolio (4.3.1) with respect to the DAX index. In Figure 4.3.2 we see the decomposition of the curve(s) in the left-hand panel according to the master equation. The blue curve is the change in the generating functional, while the red and the green ones are the respective drift terms. Each curve shows the cumulative value of the daily changes induced in the corresponding quantities by capital gains and losses. As can be seen, the cumulative second-order drift term was the dominant part over the period, with a total contribution of about 15 percentage points to the relative return. The second-order drift term was quite stable over the considered period, with an exception of the period around the financial crisis of 2008.

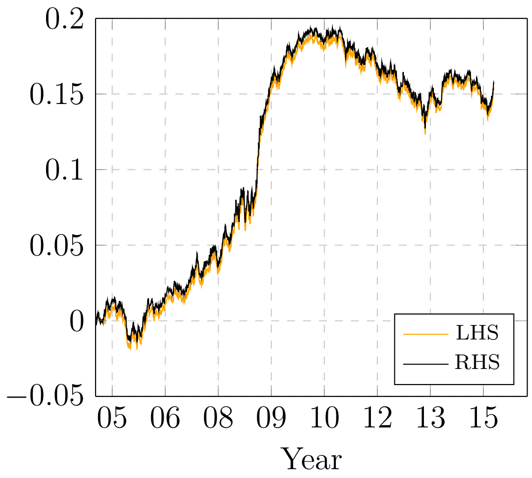

Figure 4.3.1: LHS vs. RHS of the master formula (4.2.10) for geometric weighting

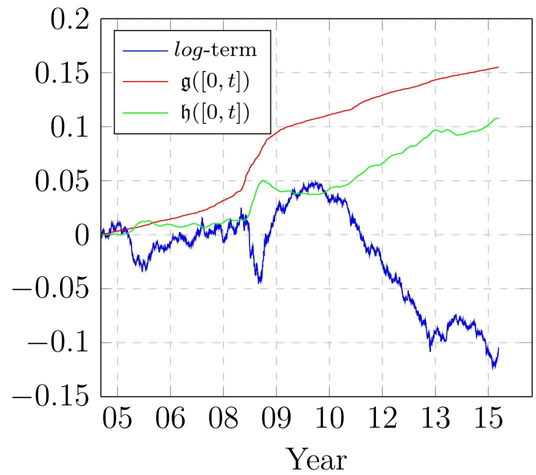

Figure 4.3.2: Componentwise representation of the RHS of (4.2.10) for geometric weighting

Example 4.3.2 (Functional Diversity weighting)**.** Consider the functional

$$G(t,X^t) := \begin{cases} \left[\sum\limits_{k=1}^{d}\left(\alpha X_k(t)+\dfrac{1-\alpha}{\delta}\left(\int\limits_0^t X_k(s)\,\mathrm{d}s-\int\limits_0^{t-\delta}X_k(0)\,\mathrm{d}s\right)\right)^p\right]^{\frac{1}{p}}, & 0\le t<\delta,\\ \left[\sum\limits_{k=1}^{d}\left(\alpha X_k(t)+\dfrac{1-\alpha}{\delta}\int\limits_{t-\delta}^{t}X_k(s)\,\mathrm{d}s\right)^p\right]^{\frac{1}{p}}, & 0<\delta\le t,\end{cases}$$

where $p\in(0,1)$. Then, the vertical derivative is given by

$$\partial_i G(t,X^t,A^t)=\alpha\left(\widetilde{X}_i(t)\right)^{p-1}\left[\sum_{k=1}^{d}\left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-1}=\alpha\left(\widetilde{X}_i(t)\right)^{p-1}\frac{G(t,X^t,A^t)}{\sum\limits_{k=1}^{d}\left(\widetilde{X}_k(t)\right)^p}.$$

Thus,

$$\partial_i\log G(t,X^t,A^t)=\frac{1}{G(t,X^t,A^t)}\partial_i G(t,X^t,A^t)=\frac{\alpha\left(\widetilde{X}_i(t)\right)^{p-1}}{\sum\limits_{k=1}^{d}\left(\widetilde{X}_k(t)\right)^p},$$

and the portfolio generated by G has the weights

$$\begin{aligned}\pi_i(t)&=\left[\partial_i\log G(t,\mu^t,A^t_\mu)+1-\sum_{j=1}^{d}\mu_j(t)\partial_j\log G(t,\mu^t,A^t_\mu)\right]\mu_i(t)\\&=\left[\frac{\alpha\left(\widetilde{\mu}_i(t)\right)^{p-1}}{\sum_{k=1}^{d}\left(\widetilde{\mu}_k(t)\right)^p}+1-\sum_{j=1}^{d}\frac{\alpha\mu_j(t)\left(\widetilde{\mu}_j(t)\right)^{p-1}}{\sum_{k=1}^{d}\left(\widetilde{\mu}_k(t)\right)^p}\right]\mu_i(t).\end{aligned}\qquad(4.3.3)$$

For the second-order vertical derivative, we compute

$$\partial^2_{ij} G(t, X^t, A^t) = \begin{cases} \alpha^2(1-p)\left(\widetilde{X}_i(t)\right)^{p-1}\left(\widetilde{X}_j(t)\right)^{p-1}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-2}, \\ \alpha^2(1-p)\left(\widetilde{X}_i(t)\right)^{p-1}\left(\widetilde{X}_i(t)\right)^{p-1}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-2} + (p-1)\alpha^2\left(\widetilde{X}_i(t)\right)^{p-2}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-1}. \end{cases}$$

Rearranging terms further gives

$$\partial^2_{ij} G(t, X^t, A^t) = \begin{cases} \alpha^2(1-p)\left(\widetilde{X}_i(t)\right)^{p-1}\left(\widetilde{X}_j(t)\right)^{p-1}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-2}, \\ \alpha^2(1-p)\left(\left(\widetilde{X}_i(t)\right)^{p-1}\left(\widetilde{X}_i(t)\right)^{p-1}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-2} - \left(\widetilde{X}_i(t)\right)^{p-2}\left[\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p\right]^{\frac{1}{p}-1}\right). \end{cases}$$

Thus, the second-order drift term is given by

$$\begin{aligned} \mathfrak{g}(\mathrm{d}t) &= -\frac{1}{2G(t,\mu^t,A^t_\mu)}\sum_{i,j=1}^d \partial^2_{ij}G(t,\mu^t,A^t_\mu)\mu_i(t)\mu_j(t)\tau^\mu_{ij}(\mathrm{d}t) \\ &= -\frac{\alpha^2(1-p)}{2\left[\sum_{k=1}^d (\widetilde{\mu}_k(t))^p\right]^{\frac{1}{p}}}\sum_{i,j=1}^d \mu_i(t)\mu_j(t)\left(\widetilde{\mu}_i(t)\right)^{p-1}\left(\widetilde{\mu}_j(t)\right)^{p-1}\left(\sum_{k=1}^d (\widetilde{\mu}_k(t))^p\right)^{\frac{1}{p}-2}\tau^\mu_{ij}(\mathrm{d}t) \\ &\quad + \frac{\alpha^2(1-p)}{2\left[\sum_{k=1}^d (\widetilde{\mu}_k(t))^p\right]^{\frac{1}{p}}}\sum_{i=1}^d \mu_i^2(t)\left(\widetilde{\mu}_i(t)\right)^{p-2}\left(\sum_{k=1}^d (\widetilde{\mu}_k(t))^p\right)^{\frac{1}{p}-1}\tau^\mu_{ii}(\mathrm{d}t) \\ &= \frac{\alpha^2(1-p)}{2}\left(\sum_{i=1}^d \frac{\mu_i^2(t)\left(\widetilde{\mu}_i(t)\right)^{p-2}}{\sum_{k=1}^d (\widetilde{\mu}_k(t))^p}\tau^\mu_{ii}(\mathrm{d}t) - \sum_{i=1}^d\sum_{j=1}^d \frac{\mu_i(t)\mu_j(t)\left(\widetilde{\mu}_i(t)\right)^{p-1}\left(\widetilde{\mu}_j(t)\right)^{p-1}}{\left(\sum_{k=1}^d (\widetilde{\mu}_k(t))^p\right)^2}\tau^\mu_{ij}(\mathrm{d}t)\right). \end{aligned}$$

For the horizontal derivative, we compute

$$\mathcal{D}G(t,X^t,A^t) = \begin{cases} \dfrac{1-\alpha}{\delta}\displaystyle\sum_{i=1}^d \frac{G(t,X^t,A^t)}{\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p}(X_i(t-) - X_i(0))\left(\widetilde{X}_i(t)\right)^{p-1}, & 0 \le t < \delta, \\ \dfrac{1-\alpha}{\delta}\displaystyle\sum_{i=1}^d \frac{G(t,X^t,A^t)}{\sum_{k=1}^d \left(\widetilde{X}_k(t)\right)^p}(X_i(t-) - X_i((t-\delta)-))\left(\widetilde{X}_i(t)\right)^{p-1}, & 0 < \delta \le t. \end{cases}$$

Hence, the horizontal drift term is given by

$$\begin{aligned} \mathfrak{h}(\mathrm{d}t) &= -\mathcal{D}\log G(t,\mu^t,A^t_\mu)\nu(\mathrm{d}t) = -\frac{1}{G(t,\mu^t,A^t_\mu)}\mathcal{D}G(t,\mu^t,A^t_\mu)\nu(\mathrm{d}t) \\ &= -\frac{1-\alpha}{\delta}\sum_{i=1}^d \frac{(\widetilde{\mu}_i(t))^{p-1}}{\sum_{k=1}^d (\widetilde{\mu}_k(t))^p}\,\mathrm{d}t \cdot \begin{cases} \mu_i(t) - \mu_i(0), & 0 \le t < \delta, \\ \mu_i(t) - \mu_i(t-\delta), & 0 < \delta \le t. \end{cases} \end{aligned}$$

Thus, the performance of the functionally generated diversity-weighted portfolio is given as follows, by the master formula,

$$
\begin{aligned}
\log\left(\frac{V^\pi(T)}{V^\mu(T)}\right) &= \log\left(\frac{\left(\sum_{k=1}^d (\widetilde{\mu}_k(T))^p\right)^{\frac{1}{p}}}{\left(\sum_{k=1}^d (\mu_k(0))^p\right)^{\frac{1}{p}}}\right) + \int_0^T \mathfrak{g}(\,\mathrm{d}t) + \int_0^T \mathfrak{h}(\,\mathrm{d}t) \\
= \log\left(\frac{\left(\sum_{k=1}^d (\widetilde{\mu}_k(T))^p\right)^{\frac{1}{p}}}{\left(\sum_{k=1}^d (\mu_k(0))^p\right)^{\frac{1}{p}}}\right) &- \frac{\alpha^2(1-p)}{2}\sum_{\substack{i,j=1\\ j\neq i}}^d \int_{[0,T]} \frac{\mu_i(t)\mu_j(t)\,(\widetilde{\mu}_i(t))^{p-1}\,(\widetilde{\mu}_j(t))^{p-1}}{\left(\sum\limits_{k=1}^d (\widetilde{\mu}_k(t))^p\right)^2}\tau^\mu_{ij}(\mathrm{d}t) \\
&- \int_{[0,T]} \frac{\alpha^2}{2}\sum_{i=1}^d \left(\frac{(p-1)\mu_i^2(t)\,(\widetilde{\mu}_i(t))^{p-2}}{\sum\limits_{k=1}^d (\widetilde{\mu}_k(t))^p} + \frac{(1-p)\mu_i^2(t)\,(\widetilde{\mu}_i(t))^{2p-2}}{\left(\sum\limits_{k=1}^d (\widetilde{\mu}_k(t))^p\right)^2}\right)\tau^\mu_{ii}(\mathrm{d}t) \\
&- \frac{1-\alpha}{\delta}\sum_{i=1}^d \int_{[0,\delta[} \frac{(\mu_i(t)-\mu_i(0))\,(\widetilde{\mu}_i(t))^{p-1}}{\sum\limits_{k=1}^d (\widetilde{\mu}_k(t))^p}\,\mathrm{d}t \\
&- \frac{1-\alpha}{\delta}\sum_{i=1}^d \int_{[\delta,T]} \frac{(\mu_i(t)-\mu_i(t-\delta))\,(\widetilde{\mu}_i(t))^{p-1}}{\sum\limits_{k=1}^d (\widetilde{\mu}_k(t))^p}\,\mathrm{d}t.
\end{aligned}
$$

To simulate such a diversity weighted portfolio with actual stocks, we used again Reuters Datastream to obtain our data base, containing now the monthly average prices of the stocks that were currently in the S&P 500 index for the period from 1973 to 2015. We filtered the data so as to consider only those stocks for which the prices are known at each time point of the considered time period. The results of a simulation of the portfolio (4.3.3) using the parameters $\delta = 12$ months, $\alpha = 0,6$, and $p = 0,1$ are presented below: Figure 4.3.3 shows the relative performance of this portfolio with respect to the filtered index, and Figure 4.3.4 shows its decomposition in the three components according to the master equation. Each curve represents the cumulative value of the monthly changes induced in the corresponding quantities by capital gains and losses, but contrarily to above, it is now the cumulative change in the generating functional that was the dominant part over the period, with a total contribution of about 70 percentage points to the relative return. The second-order drift term was quite stable over the period with a total contribution of about 30 percentage points, whereas the horizontal drift term can be viewed as the price we have to pay for more generality (see the discussion following the next example).

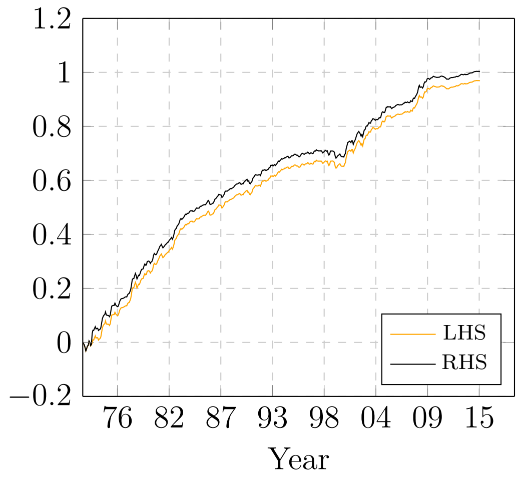

Figure 4.3.3: LHS vs. RHS of the master formula (4.2.10) for diversity weighting

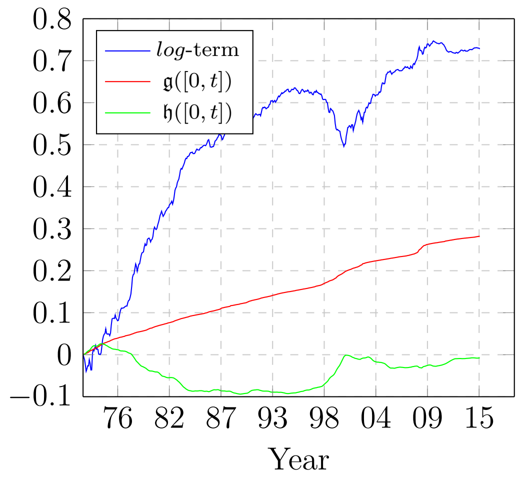

Figure 4.3.4: Componentwise representation of the RHS of (4.2.10) for diversity weighting

Example 4.3.3 (Functional Entropy weighting)**.** Consider the functional

$$G(t, X^t, A^t) := -\sum_{k=1}^{d} \widetilde{X}_k(t) \log\left(\widetilde{X}_k(t)\right). \tag{4.3.4}$$

We have that the vertical derivative is given by

$$\partial_i G(t, X^t, A^t) = -\alpha\left(\log\left(\widetilde{X}_i(t)\right) + 1\right).$$

Thus,

$$\partial_i \log G(t, X^t, A^t) = \frac{1}{G(t, X^t, A^t)} \partial_i G(t, X^t, A^t) = \frac{\alpha\left(\log\left(\widetilde{X}_i(t)\right) + 1\right)}{\sum\limits_{k=1}^{d} \widetilde{X}_k(t) \log\left(\widetilde{X}_k(t)\right)},$$

whence we infer that the weights in (4.2.9) are given by

$$\begin{aligned} \pi_i(t) &= \left[\frac{\alpha\left(\log\left(\widetilde{\mu}_i(t)\right) + 1\right)}{\sum_{k=1}^{d} \widetilde{\mu}_k(t) \log\left(\widetilde{\mu}_k(t)\right)} + 1 - \sum_{j=1}^{d} \frac{\alpha \mu_j(t)\left(\log\left(\widetilde{\mu}_j(t)\right) + 1\right)}{\sum_{k=1}^{d} \widetilde{\mu}_k(t) \log\left(\widetilde{\mu}_k(t)\right)}\right] \mu_i(t) \\ &= \left[\frac{\alpha \log\left(\widetilde{\mu}_i(t)\right)}{\sum_{k=1}^{d} \widetilde{\mu}_k(t) \log\left(\widetilde{\mu}_k(t)\right)} + 1 - \sum_{j=1}^{d} \frac{\alpha \mu_j(t) \log\left(\widetilde{\mu}_j(t)\right)}{\sum_{k=1}^{d} \widetilde{\mu}_k(t) \log\left(\widetilde{\mu}_k(t)\right)}\right] \mu_i(t). \end{aligned} \tag{4.3.5}$$

The corresponding code is given in Table A.7. Furthermore, computing

$$\partial^2_{ij} G(t, X^t, A^t) = \begin{cases} 0, & \text{if } j \neq i, \\ -\dfrac{\alpha^2}{\widetilde{X}_i(t)}, & \text{if } j = i, \end{cases}$$

yields for the second-order drift term:

$$\begin{aligned}\mathfrak{g}(\,\mathrm{d}t) &= -\frac{1}{2G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d \partial^2_{ii}G(t,\mu^t,A^t_\mu)\mu_i(t)\mu_i(t)\tau^\mu_{ii}(\mathrm{d}t)\\ &= \frac{\alpha^2}{2G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d \frac{\mu_i^2(t)}{\widetilde{\mu}_i(t)}\tau^\mu_{ii}(\mathrm{d}t) = \frac{\alpha^2}{G(t,\mu^t,A^t_\mu)}\widetilde{\gamma}^*_\mu(\,\mathrm{d}t),\end{aligned} \tag{4.3.6}$$

where

$$\widetilde{\gamma}^*_\mu(\,\mathrm{d}t) := \frac{1}{2}\sum_{i=1}^d \frac{\mu_i^2(t)}{\widetilde{\mu}_i(t)}\tau^\mu_{ii}(\mathrm{d}t). \tag{4.3.7}$$

For the horizontal derivative, we calculate

$$\mathcal{D}G(t,X^t,A^t) = \begin{cases} -\sum\limits_{i=1}^d \dfrac{1-\alpha}{\delta}(X_i(t-)-X_i(0))\left(\log\left(\widetilde{X}_i(t)\right)+1\right), & 0\le t<\delta,\\ -\sum\limits_{i=1}^d \dfrac{1-\alpha}{\delta}(X_i(t-)-X_i((t-\delta)-))\left(\log\left(\widetilde{X}_i(t)\right)+1\right), & 0<\delta\le t,\end{cases}$$

whence we infer that the horizontal drift term is given by

$$\begin{aligned}\mathfrak{h}(\mathrm{d}t) &= -\mathcal{D}\log G(t,\mu^t,A^t_\mu)\nu(\mathrm{d}t) = -\frac{1}{G(t,\mu^t,A^t_\mu)}\mathcal{D}G(t,\mu^t,A^t_\mu)\nu(\mathrm{d}t)\\ &= \frac{1-\alpha}{\delta G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d (\log(\widetilde{\mu}_i(t))+1)\ \mathrm{d}t\cdot\begin{cases}\mu_i(t)-\mu_i(0), & 0\le t<\delta,\\ \mu_i(t)-\mu_i(t-\delta), & 0<\delta\le t.\end{cases}\end{aligned} \tag{4.3.8}$$

Thus, the relative performance of functional entropy-weighting is given as follows, by the functional master equation,

$$\log\left(\frac{V^\pi(T)}{V^\mu(T)}\right) = \log\left(\frac{\sum_{k=1}^d \widetilde{\mu}_k(T)\log(\widetilde{\mu}_k(T))}{\sum_{k=1}^d \mu_k(0)\log(\mu_k(0))}\right) + \int_0^T \mathfrak{g}(\,\mathrm{d}t) + \int_0^T \mathfrak{h}(\,\mathrm{d}t)$$

$$\begin{aligned} &= \log\left(\frac{\sum_{k=1}^d \widetilde{\mu}_k(T)\log(\widetilde{\mu}_k(T))}{\sum_{k=1}^d \mu_k(0)\log(\mu_k(0))}\right) + \int_{[0,T]} \frac{\alpha^2}{2G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d \frac{\mu_i^2(t)}{\widetilde{\mu}_i(t)}\tau^\mu_{ii}(\mathrm{d}t)\\ &\quad + \int_{[0,\delta[} \frac{1-\alpha}{\delta G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d(\mu_i(t)-\mu_i(0))(\log(\widetilde{\mu}_i(t))+1)\ \mathrm{d}t\\ &\quad + \int_{[\delta,T]} \frac{1-\alpha}{\delta G(t,\mu^t,A^t_\mu)}\sum_{i=1}^d(\mu_i(t)-\mu_i(t-\delta))(\log(\widetilde{\mu}_i(t))+1)\ \mathrm{d}t\\ &= \log\left(\frac{\sum_{k=1}^d \widetilde{\mu}_k(T)\log(\widetilde{\mu}_k(T))}{\sum_{k=1}^d \mu_k(0)\log(\mu_k(0))}\right) + \int_0^T \frac{\alpha^2\widetilde{\gamma}^*_\mu(\,\mathrm{d}t)}{G(\widetilde{\mu}(t))} + \mathfrak{h}([0,T]) \qquad (4.3.9)\\ &\ge \log\left(\frac{\sum_{k=1}^d \widetilde{\mu}_k(T)\log(\widetilde{\mu}_k(T))}{\sum_{k=1}^d \mu_k(0)\log(\mu_k(0))}\right) + \frac{\alpha^2\widetilde{\gamma}^*_\mu([0,T])}{\log d} + \mathfrak{h}([0,T]). \qquad (4.3.10)\end{aligned}$$

We ran a simulation of this portfolio using the same data set as in the previous example, taking the parameters $\delta = 6$ months and $\alpha = 0,9$, which is presented in Figure 4.3.5 and Figure 4.3.6, respectively. The code for computing the left-hand side of (4.3.9) is given in Table A.8, while the code for computing its right-hand side is given in Table A.9, or equivalently, in Table A.10. We do not present in the appendix the codes for the tool functions, but the code generating the figures below is given in Table A.11.

Remark 4.3.4 (Intrinsic Volatility and Arbitrage)**.** Note that the positive Radon measure $\widetilde{\gamma}^*_\mu$ from (4.3.7) describes the market's "intrinsic" volatility, in some extended sense, since it is a weighted average of the variances of the individual stocks relative to the market. If we now assume that $\widetilde{\gamma}^*_\mu$ is a strictly positive measure, (4.3.10) yields $V^\pi(T) > V^\mu(T)$, under the additional condition that the horizontal drift term is significantly outperformed, i.e.,

$$\log\left(\frac{\sum_{k=1}^d \widetilde{\mu}_k(T)\log\left(\widetilde{\mu}_k(T)\right)}{\sum_{k=1}^d \mu_k(0)\log\left(\mu_k(0)\right)}\right) + \frac{\alpha^2\widetilde{\gamma}^*_\mu([0,T])}{\log d} + \mathfrak{h}([0,T]) > 0. \tag{4.3.11}$$

In this sense, "availability of intrinsic volatility" can be regarded as a property that admits strong relative arbitrage with respect to the market, provided that the horizontal drift term satisfies (4.3.11). Moreover, note that compared to the entropy-weighted portfolio in the classical Fernholz' setting the additional drift term $\mathfrak{h}([0,T])$ appears in (4.3.11), which can be either positive or negative. This can be considered as the price we have to pay in order to include running averages, and not only the current state of the market, in our consideration. Thus, the horizontal drift term $\mathfrak{h}([0,T])$ may be regarded as quantifying the *trade-off* between outperforming the market and including the past evolution in the consideration.

The above argument is supported by real market data, as can be seen in Figure 4.3.5 and Figure 4.3.6: Indeed, since the cumulative second-order drift term is continually increasing, we infer that the (extended) excess growth rate of the market $\widetilde{\gamma}^*_\mu$ is a strictly positive measure. Moreover, the horizontal drift term does not seem to have a large influence on the relative performance of the entropy-weighted portfolio, with a total contribution of less than 1 percentage point. Thus, entropy-weighting should significantly outperform the market on the considered time interval, which is confirmed in Figure 4.3.5.

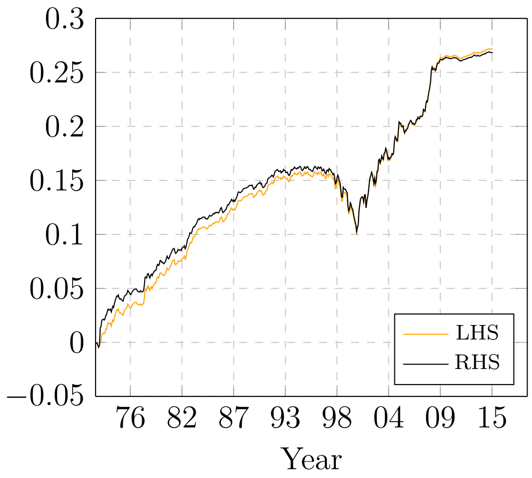

Figure 4.3.5: LHS vs. RHS of the master formula (4.2.10) for entropy weighting

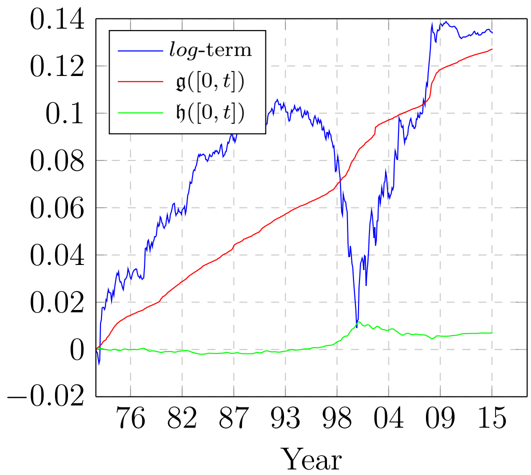

Figure 4.3.6: Componentwise representation of the RHS of (4.2.10) for entropy weighting

Note also that the above example is valid in a very general context, since we have not imposed in the above discussion *any* assumption on the volatility structure of the market model beyond the absolutely minimal condition – that the stock price vector S should admit continuous quadratic variation.

Remark 4.3.5. For the functionally generated entropy-weighted portfolio in (4.3.5), we have

$$\frac{\alpha^2}{G(\widetilde{\mu}(t))}\widetilde{\gamma}^*_\mu(\,\mathrm{d}t) = \mathfrak{g}(\,\mathrm{d}t) = \,\mathrm{d}\log\left(\frac{V^\pi(t)G(\mu(0))}{V^\mu(t)G(\widetilde{\mu}(t))}\right) \\ - \frac{1-\alpha}{\delta G(\widetilde{\mu}(t))}\sum_{i=1}^d \left(\log\left(\widetilde{\mu}_i(t)\right)+1\right)\,\mathrm{d}t\cdot\begin{cases}\mu_i(t)-\mu_i(0), & 0\le t<\delta,\\ \mu_i(t)-\mu_i(t-\delta), & 0<\delta\le t.\end{cases}$$

Hence, the distribution function of the measure $\widetilde{\gamma}^*_\mu$, which describes the generalized cumulative excess growth of the market, can be expressed as

$$\int_0^T \widetilde{\gamma}^*_\mu(\,\mathrm{d}t) = \frac{1}{\alpha^2}\int_0^T G(\widetilde{\mu}(t))\,\mathrm{d}\log\left(\frac{V^\pi(t)G(\mu(0))}{V^\mu(t)G(\widetilde{\mu}(t))}\right) \\ - \int_0^\delta \frac{1-\alpha}{\delta\alpha^2}\sum_{i=1}^d \left(\log\left(\widetilde{\mu}_i(t)\right)+1\right)\left(\mu_i(t)-\mu_i(0)\right)\mathrm{d}t \\ - \int_\delta^T \frac{1-\alpha}{\delta\alpha^2}\sum_{i=1}^d \left(\log\left(\widetilde{\mu}_i(t)\right)+1\right)\left(\mu_i(t)-\mu_i(t-\delta)\right)\mathrm{d}t, \tag{4.3.12}$$

in terms of quantities that are observable. The corresponding code is given in Table A.12. In Table A.13, we also give a supplementary code to compare the direct calculation of $\widetilde{\gamma}^*_\mu$ via its definition (4.3.7) and the computation (4.3.12).

Chapter 5

Pathwise no-arbitrage in a class of Delta-hedging strategies

Bick and Willinger [11] have developed a theory of hedging European-type options with payoff $H = h(S(T))$ for one-dimensional asset price trajectories $S = (S(t))_{0 \le t \le T}$ in the framework of Föllmer's pathwise Itô calculus [46]. In particular, Bick and Willinger [11] showed that if S is continuous, strictly positive and has the continuous quadratic variation $[S, S](t) = \int_0^t a(s, S(s))\,\mathrm{d}s$, where $a(s, x)$ is strictly positive, and if v solves the terminal-value problem

$$(\mathrm{TVP}_0) \qquad \begin{cases} v \in C^{1,2}([0,T) \times \mathbb{R}_+) \cap C([0,T] \times \mathbb{R}_+), \\ \frac{\partial v}{\partial t} + a\frac{\partial^2 v}{\partial x^2} = 0 \text{ in } [0,T) \times \mathbb{R}_+, \\ v(T,x) = h(x),\ x \in \mathbb{R}_+, \end{cases}$$

then $v(t, S(t))$ represents the portfolio value of a self-financing strategy that perfectly hedges the option with payoff $H = h(S(T))$, in a strictly pathwise sense. Thus, the initial investment $v(0, S(0))$ can be interpreted as the cost required to perfectly hedge the option H, which, in standard continuous-time finance, is usually regarded as an arbitrage-free price of H. This is, however, not so clear in our probability-free setting, since we first need to exclude the existence of arbitrage in a strictly pathwise sense.

In the present chapter, which is based on [83], we follow the approach from [11] and, in a first step, extend their results to a setting with a d-dimensional price trajectory, $S(t) = (S_1(t), \dots, S_d(t))^\top$, and an exotic derivative of the form $H = h(S(t_0), \dots, S(t_N))$, where $t_0 < t_1 < \cdots < t_N$ are the fixing times of daily closing prices and h is a certain function. In the second step, we then discuss the absence of strictly pathwise arbitrage within a class of strategies that are based on solutions of recursive schemes of terminal-value problems.

The chapter is structured as follows. In Section 5.1, we present a general framework for continuous-time trading based on Föllmer's pathwise Itô calculus [46]. Recalling Proposition 2.1.11, the basic main requirement on the price trajectories will be that they admit continuous quadratic variations and covariations in the sense of [46]. With this at hand, we will introduce the pathwise framework for hedging exotic derivatives in the spirit of [11]. In Section 5.2, we will introduce the class of strategies, comprising the natural Delta hedging strategies for path-dependent exotic options, for which sufficient conditions for no arbitrage can be formulated (Theorem 5.2.4). These results are extended to the functional setting in Section 5.3.

5.1 Strictly pathwise hedging of exotic derivatives

Following [83], we first describe a general approach to asset price modeling and to the hedging of (exotic) derivatives by means of Föllmer's pathwise Itô calculus (as motivated in Section 2.1). Let us assume that we wish to trade continuously in $d+1$ assets. The first one is a riskless bond, and we assume for simplicity that its price, $B(t)$, satisfies $B(t) = 1$ for all t, which can be justified by assuming that we are dealing here only with properly discounted asset prices. This implies in particular that we do not have to distinguish between forward prices and quoted prices of the underlying. The prices of the d risky assets will be described by continuous functions $S_1(t), \dots, S_d(t)$, where the time parameter t varies over a certain time interval $[0,T]$, during which our (frictionless) market is open for trade. Recalling the discussion in Chapter 2 and Chapter 4, for fixed $S = (S(t))_{0\le t\le T}$, a trading strategy will be described by a pair of functions $\xi = (\xi_1, \dots, \xi_d)^\top$ and η, where $\xi_i(t)$ corresponds to the number of shares held at time t in the i^{th} risky asset and $\eta(t)$ does the same for the riskless asset. The associated portfolio value of $(\xi(t), \eta(t))$ is given by

$$V(t) := \xi(t)\cdot S(t) + \eta(t)B(t) = \xi(t)\cdot S(t) + \eta(t), \qquad 0 \le t \le T. \tag{5.1.1}$$

Recall from Proposition 2.1.11 and the discussion following it that it is reasonable to require that price trajectories S of a risky asset possess all covariations $[S_i, S_j]$ in the sense that the limit in (2.1.12) exists for all $t \in [0,T]$ and all covariations $[S_i, S_j]$ are continuous functions in t. In particular, it is readily observed that for the class of "basic admissible integrands" ξ as defined subsequently, the Itô integral $\int_r^t \xi(s)\,\mathrm{d}S(s)$ (see (2.2.7)) exists for all $t \in [r,u] \subset [0,T]$ as the finite limit of Riemann sums; see [83] and [80, p. 86].

Definition 5.1.1 (Basic admissible integrands)**.** Let $0 \le r < u \le T$. We call an $\mathbb{R}^d$-valued function $[r,u] \ni t \mapsto \xi(t)$ a *basic admissible integrand* for $S \in QV^d$, if there are $m \in \mathbb{N}$, a continuous function $A : [r,u] \to \mathbb{R}^m$ with components of bounded variation, an open set $O \subset \mathbb{R}^m \times \mathbb{R}^d$ such that $(A(t), S(t)) \in O$ for all t, and a continuously differentiable function $f : O \to \mathbb{R}$ such that the function $x \to f(A(t), x)$ is for all t twice continuously differentiable on its domain, so that we can write

$$\xi(t) = \nabla_x f(A(t), S(t)).$$

Here, $\nabla_x f(a,x)$ is the gradient of the function $x \to f(a,x)$.

Typically, the exact dynamics of the actual realization S will be not known until time T, which corresponds to model uncertainty or *Knightian uncertainty* (see [61]). In the following, in order to explicitly take into account this model uncertainty, we fix not just one particular path S, but admit an entire class of possible dynamics. Specifically, we will consider the classes

$$\mathscr{S}_a := \Big\{ S \in QV^d \,\Big|\, [S_i, S_j](t) = \int_0^t a_{ij}(s, S(s))\,\mathrm{d}s,\ t \in [0,T],\ 1 \le i,j \le d \Big\}$$

and

$$\mathscr{S}_a^+ := \Big\{ S \in QV^d \,\Big|\, S_i(t) > 0,\ [S_i, S_j](t) = \int_0^t a_{ij}(s, S(s)) S_i(s) S_j(s)\,\mathrm{d}s \Big\},$$

where $a(t,x) = (a_{ij}(t,x))_{i,j=1,\dots,d}$ is a continuous function mapping $(t,x) \in [0,T] \times \mathbb{R}^d$ (respectively $(t,x) \in [0,T] \times \mathbb{R}_+^d$ in case of $\mathscr{S}_a^+$) into the set of positive definite symmetric $d \times d$-matrices. Additional assumptions on $a(t,x)$ will follow shortly. Here, $\mathbb{R}_+ := (0,\infty)$, and we will write $\mathbb{R}^d_{(+)}$ to denote the two possibilities, $\mathbb{R}^d$ and $\mathbb{R}^d_+$, according to whether we are considering $\mathscr{S}_a$ or $\mathscr{S}_a^+$. Similarly, we will use the notation $\mathscr{S}_a^{(+)}$ etc. Note that price paths in the class $\mathscr{S}_a^+$ can, for instance, arise as realizations of multi-dimensional local volatility models. In case of $d = 1$, the local volatility function $\sigma(\cdot) := \sqrt{a(\cdot)}$ is often chosen by calibrating to the market prices of liquid plain vanilla options [33] (see also [52]). Since in practice there are only finitely many given options prices, $\sigma(\cdot)$ is typically only determined on a finite grid [16], and so regularity assumptions on $\sigma(\cdot)$ can be made without loss of generality.

Our next goal is to introduce and characterize a class of self-financing trading strategies that may depend on the current value of the particular realization $S \in \mathscr{S}_a^{(+)}$ and includes candidates for hedging strategies of European derivatives. First, let us introduce some notation. By $C(D)$ we will denote the class of real-valued continuous functions on a set $D \subset \mathbb{R}^n$. For an interval $I \subset [0,T]$ with nonempty interior, $\mathring{I}$, we denote by $C^{1,2}(I \times \mathbb{R}^d_{(+)})$ the class of all functions in $C(I \times \mathbb{R}^d_{(+)})$ that are continuously differentiable in $(t,x) \in \mathring{I} \times \mathbb{R}^d_{(+)}$, twice continuously differentiable in x for all $t \in \mathring{I}$, and whose derivatives admit continuous extensions to $I \times \mathbb{R}^d_{(+)}$. We also introduce the following second-order differential operators,

$$\mathscr{L} := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} \qquad \text{and} \qquad \mathscr{L}^+ := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t,x) x_i x_j \frac{\partial^2}{\partial x_i \partial x_j}.$$

Proposition 5.1.2. *Let $0 \le r < u \le T$ and $v \in C^{1,2}([r,u] \times \mathbb{R}^d_{(+)})$. Then, the following two conditions are equivalent.*

(a) *For each $S \in \mathscr{S}_a^{(+)}$, there is a basic admissible integrand ξ^S on $[r,u]$ such that*

$$v(t, S(t)) = v(r, S(r)) + \int_r^t \xi^S(s)\ \mathrm{d}S(s) \qquad \text{for } t \in [r,u].$$

(b) *The function v satisfies the parabolic equation*

$$\frac{\partial v}{\partial t} + \mathscr{L}^{(+)} v = 0 \qquad \text{in } [r,u] \times \mathbb{R}^d_{(+)}. \tag{5.1.2}$$

Moreover, if these equivalent conditions are satisfied, then ξ^S *in* (a) *must necessarily be of the form*

$$\xi^S(t) = \nabla_x v(t, S(t)). \tag{5.1.3}$$

Proof. Using the pathwise Itô formula gives for $S \in \mathscr{S}_a^{(+)}$,

$$v(t, S(t)) = v(r, S(r)) + \int_r^t \nabla_x v(s, S(s))\, \mathrm{d}S(s) + \int_r^t \Big(\frac{\partial}{\partial t} v(s, S(s)) + \mathscr{L}^{(+)} v(s, S(s))\Big)\, \mathrm{d}s. \tag{5.1.4}$$

Thus, we infer immediately that (b) implies (a) and that (5.1.3) must hold.

Suppose now that (a) holds, whence we infer

$$\int_r^t \big(\xi^S(s) - \nabla_x v(s, S(s))\big)\, \mathrm{d}S(s) = \int_r^t \Big(\frac{\partial}{\partial t} v(s, S(s)) + \mathscr{L}^{(+)} v(s, S(s))\Big)\, \mathrm{d}s.$$

Because the right-hand side has vanishing quadratic variation (by [86, Proposition 2.2.2]), the left-hand side must also have vanishing quadratic variation. On the other hand, applying [80, Proposition 12] yields that the quadratic variation of the left-hand side is given by

$$\int_r^t \big(\xi^S(s) - \nabla_x v(s, S(s))\big)^\top a(s, S(s)) \big(\xi^S(s) - \nabla_x v(s, S(s))\big)\, \mathrm{d}s$$

in case of $S \in \mathscr{S}_a$. Differentiating with respect to time yields

$$\big(\xi^S(t) - \nabla_x v(t, S(t))\big)^\top a(t, S(t)) \big(\xi^S(t) - \nabla_x v(t, S(t))\big) = 0$$

for all t, and using the fact that the matrix $a(t, S(t))$ is positive definite we infer that (5.1.3) must be true. For $S \in \mathscr{S}_a^+$, the matrix $a(s, S(s))$ needs to be replaced by the matrix with components $a_{ij}(s, S(s)) S_i(s) S_j(s)$, and we obtain (5.1.3) by using the same argumentats as for the case $S \in \mathscr{S}_a$. Plugging (5.1.3) into (5.1.4) and using (a) yields that the rightmost integral in (5.1.4) is identically zero, which establishes (b) by again differentiatig with respect to the time parameter. □

Suppose now that we are given a continuous function $f : \mathbb{R}^d_{(+)} \to \mathbb{R}$ such that there exists a solution v to the following terminal-value problem,

$$(\mathrm{TVP}^{(+)}) \qquad \begin{cases} v \in C^{1,2}([0, T) \times \mathbb{R}^d_{(+)}) \cap C([0, T] \times \mathbb{R}^d_{(+)}), \\ \frac{\partial v}{\partial t} + \mathscr{L}^{(+)} v = 0 \text{ in } [0, T) \times \mathbb{R}^d_{(+)}, \\ v(T, x) = f(x) \text{ for } x \in \mathbb{R}^d_{(+)}. \end{cases}$$

For $S \in \mathscr{S}_a^{(+)}$ and $t \in [0, T)$, we can then define

$$\xi^S(t) := \nabla_x v(t, S(t)) \qquad \text{and} \qquad \eta^S(t) := v(t, S(t)) - \xi^S(t) \cdot S(t), \tag{5.1.5}$$

and Proposition 5.1.2 yields that

$$\xi^S(t) \cdot S(t) + \eta^S(t) = v(t, S(t)) = v(0, S(0)) + \int_0^t \xi^S(s)\, \mathrm{d}S(s). \tag{5.1.6}$$

Thus, (ξ^S, η^S) is a self-financing trading strategy with associated portfolio value $V^S(t) = v(t, S(t))$. Moreover, since the function v is continuous on $[0,T] \times \mathbb{R}^d_{(+)}$, the limit $V^S(T) := \lim_{t \uparrow T} V^S(t)$ exists, and we have

$$V^S(T) = f(S(T)) \qquad \text{for all } S \in \mathscr{S}_a^{(+)}.$$

In this sense, (ξ^S, η^S) can be regarded as a strictly pathwise hedging strategy for the derivative with (European-type) payoff structure $f(S(T))$.

Note that the preceding argument was first made by Bick and Willinger [11, Proposition 3], in a one-dimensional setting. It has several interesting consequences. For example, when considering the one-dimensional case with $a(t,x) = \sigma^2 x^2$ for some strictly positive constant $\sigma > 0$, the terminal value problem (TVP$^+$) reduces to the standard Black–Scholes equation, which can be solved for a large class of payoff functions f. The preceding argument then shows that the derivation of the Black–Scholes formula — which is nothing else than an explicit formula for the initial value $v(0, S_0)$ — does not require any probabilistic assumptions whatsoever. It follows in particular that the fundamental assumption underlying the Black–Scholes formula is not the log-normal distribution of asset price returns, but the fact that the quadratic variation of the asset prices is of the form $[S,S](t) = \sigma^2 \int_0^t S(s)^2 \, \mathrm{d}s$. Let us now state and show general existence results for solutions of (TVP) and (TVP$^+$), which in the case of (TVP) is taken from Janson and Tysk [56]. Recall that $a(t,x)$ is assumed to be positive definite for all t and x.

Theorem 5.1.3. *Assume that $f \in C(\mathbb{R}^d_{(+)})$ has at most polynomial growth in the following sense: $|f(x)| \le c_0(1 + |x|^p)$ for some constants $c_0, p > 0$. Then, under the following conditions,* (TVP$^{(+)}$) *admits a unique solution $v(t,x)$ within the class of functions that are of at most polynomial growth uniformly in t.*

(a) (Theorem A.14 in [56]) *In case of* (TVP)*, we assume that $a_{ij}(t,x)$ is locally Hölder continuous on $[0,T) \times \mathbb{R}^d$ and that $|a_{ij}(t,x)| \le c_1(1 + |x|^2)$ for a constant $c_1 \ge 0$, all $(t,x) \in [0,T] \times \mathbb{R}^d$, and all i, j.*

(b) *In case of* (TVP$^+$)*, we assume that $a_{ij}(t,x)$ is bounded and locally Hölder continuous on $[0,T) \times \mathbb{R}^d$ for all i,j.*

To prove Theorem 5.1.3 (b), we need some transformation lemmas.

Lemma 5.1.4. *For $x = (x_1, \dots, x_d)^\top \in \mathbb{R}^d$ let $\exp(x) := (e^{x_1}, \dots, e^{x_d})^\top \in \mathbb{R}^d_+$. Then $v(t,x)$ solves (TVP^+) if and only if $\widetilde{v}(t,x) := v(t, \exp(x))$ solves*

$$(\widetilde{TVP}) \qquad \begin{cases} \widetilde{v} \in C^{1,2}([0,T) \times \mathbb{R}^d) \cap C([0,T] \times \mathbb{R}^d), \\ \frac{\partial \widetilde{v}}{\partial t} + \widetilde{\mathscr{L}} \widetilde{v} = 0 \ \text{in } [0,T) \times \mathbb{R}^d, \\ \widetilde{v}(T,x) = \widetilde{f}(x) \ \text{for } x \in \mathbb{R}^d, \end{cases}$$

where $\widetilde{f}(x) = f(\exp(x))$ and

$$\widetilde{\mathscr{L}} := \frac{1}{2} \sum_{i,j=1}^d \widetilde{a}_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \widetilde{b}_i(t,x) \frac{\partial}{\partial x_i}, \qquad x \in \mathbb{R}^d, \tag{5.1.7}$$

for $\widetilde{a}_{ij}(t,x) := a_{ij}(t,\exp(x))$ *and* $\widetilde{b}_i(t,x) := -\frac{1}{2}a_{ii}(t,\exp(x))$.

Proof. Let us write $y_i := e^{x_i} > 0$, $i = 1,\dots,d$. We then have that $v(T,y) = f(y)$, $y \in \mathbb{R}^d_+$, if and only if $\widetilde{v}(T,x) = \widetilde{f}(x)$, $x \in \mathbb{R}^d$. Moreover, $\widetilde{v} \in C^{1,2}([0,T)\times\mathbb{R}^d)\cap C([0,T]\times\mathbb{R}^d)$ if and only if $v \in C^{1,2}([0,T)\times\mathbb{R}^d_+)\cap C([0,T]\times\mathbb{R}^d_+)$, and we have

$$\frac{\partial\widetilde{v}(t,x)}{\partial t} = \frac{\partial v(t,y)}{\partial t} \quad \text{and} \quad \frac{\partial\widetilde{v}(t,x)}{\partial x_i} = \frac{\partial v(t,y)}{\partial y_i}\cdot e^{x_i}.$$

For the second partial derivatives, we calculate

$$\frac{\partial^2\widetilde{v}(t,x)}{\partial x_i\partial x_j} = \begin{cases} \dfrac{\partial^2 v(t,y)}{\partial y_i\partial y_j}e^{x_i}e^{x_j}, & i \neq j,\\ \dfrac{\partial^2 v(t,y)}{\partial y_i^2}e^{x_i}e^{x_i} + \dfrac{\partial v(t,y)}{\partial y_i}\cdot e^{x_i}, & i = j. \end{cases}$$

Hence, it follows that

$$\begin{aligned}\frac{\partial\widetilde{v}(t,x)}{\partial t} + \widetilde{\mathscr{L}}\widetilde{v}(t,x) &= \frac{\partial v(t,y)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,y)y_iy_j\frac{\partial^2 v(t,y)}{\partial y_i\partial y_j}\\ &\quad + \frac{1}{2}\sum_{i=1}^d a_{ii}(t,y)\frac{\partial\widetilde{v}(t,x)}{\partial x_i} - \frac{1}{2}\sum_{i=1}^d a_{ii}(t,y)\frac{\partial\widetilde{v}(t,x)}{\partial x_i}\\ &= 0\end{aligned}$$

if and only if $\frac{\partial v(t,y)}{\partial t} + \mathscr{L}^+ v(t,y) = 0$, which leads to the conclusion. □

Next, the terminal-value problem ($\widetilde{\text{TVP}}$) will be once again transformed into another auxiliary terminal-value problem.

Lemma 5.1.5. *For* $p > 0$ *let* $g(x) := 1 + \sum_{i=1}^d e^{px_i}$. *Then* $\widetilde{v}(t,x)$ *solves* ($\widetilde{TVP}$) *if and only if* $\widehat{v}(t,x) := g(x)^{-1}\widetilde{v}(t,x)$ *solves*

$$(\widehat{TVP}) \qquad \begin{cases} \widehat{v} \in C^{1,2}([0,T)\times\mathbb{R}^d)\cap C([0,T]\times\mathbb{R}^d),\\ \frac{\partial\widehat{v}}{\partial t} + \widehat{\mathscr{L}}\widehat{v} = 0 \ \text{in } [0,T)\times\mathbb{R}^d,\\ \widehat{v}(T,x) = \widehat{f}(x) \ \text{for } x \in \mathbb{R}^d, \end{cases}$$

where $\widehat{f}(x) = \widetilde{f}(x)/g(x)$ *and*

$$\widehat{\mathscr{L}} := \frac{1}{2}\sum_{i,j=1}^d \widetilde{a}_{ij}(t,x)\frac{\partial^2}{\partial x_i\partial x_j} + \sum_{i=1}^d \widehat{b}_i(t,x)\frac{\partial}{\partial x_i} + \widehat{c}(t,x), \qquad x \in \mathbb{R}^d, \tag{5.1.8}$$

for

$$\widehat{b}_i(t,x) = \widetilde{b}_i(t,x) + pg(x)^{-1}\sum_{j=1}^d e^{px_j}\widetilde{a}_{ij}(t,x),$$

$$\widehat{c}(t,x) = \frac{p(p-1)}{2g(x)}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)e^{px_i}.$$

Proof. For the terminal condition, we have that $\widetilde{v}(T,x) = \widetilde{f}(x)$ if and only if $\widehat{v}(T,x) = \widehat{f}(x)$. Moreover, $\widehat{v} \in C^{1,2}([0,T)\times\mathbb{R}^d)\cap C([0,T]\times\mathbb{R}^d)$ if and only if $\widetilde{v} \in C^{1,2}([0,T)\times\mathbb{R}^d)\cap C([0,T]\times\mathbb{R}^d)$, and we calculate, using the quotient rule,

$$\frac{\partial \widehat{v}(t,x)}{\partial t} = \frac{1}{g(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial t}; \quad \frac{\partial \widehat{v}(t,x)}{\partial x_i} = \frac{1}{g(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_i} - \frac{1}{g^2(x)}\widetilde{v}(t,x)pe^{px_i};$$

$$\frac{\partial^2 \widehat{v}(t,x)}{\partial x_i \partial x_j} = \begin{cases} \frac{1}{g(x)}\cdot\frac{\partial^2 \widetilde{v}(t,x)}{\partial x_i\partial x_j} - \frac{1}{g^2(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_i}pe^{px_j} \\ -\frac{1}{g^2(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_j}pe^{px_i} + \frac{2}{g^3(x)}p^2e^{px_i}e^{px_j}\widetilde{v}(t,x), & i \neq j, \\ \frac{1}{g(x)}\cdot\frac{\partial^2 \widetilde{v}(t,x)}{\partial x_i^2} - \frac{1}{g^2(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_i}pe^{px_i} \\ -\frac{1}{g^2(x)}pe^{px_i}\left(\frac{\partial \widetilde{v}(t,x)}{\partial x_i} + p\widetilde{v}(t,x)\right) + \frac{2}{g^3(x)}p^2e^{2px_i}\widetilde{v}(t,x), & i = j. \end{cases}$$

Thus, we infer that

$$\begin{aligned}
\frac{\partial \widehat{v}}{\partial t}(t,x) + \widehat{\mathscr{L}}\widehat{v}(t,x) &= \frac{1}{g(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^d \widetilde{a}_{ij}(t,x)\Bigg(\frac{1}{g(x)}\cdot\frac{\partial^2 \widetilde{v}(t,x)}{\partial x_i\partial x_j} \\
&\quad - \frac{1}{g^2(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_i}pe^{px_j} - \frac{1}{g^2(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_j}pe^{px_i} + \frac{2}{g^3(x)}p^2e^{px_i}e^{px_j}\widetilde{v}(t,x)\Bigg) \\
&\quad - \frac{1}{2}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)\frac{1}{g^2(x)}p^2e^{px_i}\widetilde{v}(t,x) + \sum_{i=1}^d \widetilde{b}_i(t,x)\frac{1}{g(x)}\cdot\frac{\partial \widetilde{v}(t,x)}{\partial x_i} \\
&\quad + \sum_{i,j=1}^d \frac{p}{g^2(x)}\widetilde{a}_{ij}(t,x)e^{px_j}\frac{\partial \widetilde{v}(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)\frac{1}{g^2(x)}pe^{px_i}\widetilde{v}(t,x) \\
&\quad - \sum_{i,j=1}^d \widetilde{a}_{ij}(t,x)\frac{p^2}{g^3(x)}e^{px_i}e^{px_j}\widetilde{v}(t,x) + \frac{p(p-1)}{2g^2(x)}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)e^{px_i}\widetilde{v}(t,x) \\
&= \frac{1}{g(x)}\left(\frac{\partial \widetilde{v}(t,x)}{\partial t} + \widetilde{\mathscr{L}}\widetilde{v}(t,x)\right) \\
&\quad + \frac{p(1-p)}{2g^2(x)}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)e^{px_i}\widetilde{v}(t,x) + \frac{p(p-1)}{2g^2(x)}\sum_{i=1}^d \widetilde{a}_{ii}(t,x)e^{px_i}\widetilde{v}(t,x) \\
&= 0
\end{aligned}$$

if and only if $\frac{\partial \widetilde{v}(t,x)}{\partial t} + \widetilde{\mathscr{L}}\widetilde{v}(t,x) = 0$, which concludes the proof. □

Proof of Theorem 5.1.3. We will prove that ($\widetilde{\text{TVP}}$) admits a solution $\widetilde{v}$ if $|\widetilde{f}(x)| \leq c(1 + \sum_{i=1}^d e^{px_i})$ for some $p > 0$ and that $\widetilde{v}$ is uniquely given in the class of functions that satisfy a similar estimate uniformly in t. First, we observe that the coefficients of $\widehat{\mathscr{L}}$ satisfy the conditions of [56, Theorem A.14], i.e., $\widehat{a}(t,x) = \widetilde{a}(t,x)$ is positive definite, there exist constants c_1, c_2, c_3 such that for all t, x, and i,j we have $|\widetilde{a}_{ij}(t,x)| \leq c_1(1+|x|^2)$, $|\widehat{b}_i(t,x)| \leq c_2(1+|x|)$, $|\widehat{c}(t,x)| \leq c_3$, and the coefficients $\widetilde{a}_{ij}$, $\widehat{b}_i$, and $\widehat{c}$ are locally Hölder continuous in $[0,T)\times\mathbb{R}^d$.

Therefore, we infer that ($\widehat{\text{TVP}}$) admits a unique bounded solution $\widehat{v}$ whenever $\widehat{f}$ is bounded and continuous. But then $\widetilde{v}(t,x) := g(x)\widehat{v}(t,x)$ is a solution of the terminal value problem ($\widetilde{\text{TVP}}$) with terminal condition $\widetilde{f}(x) := g(x)\widehat{f}(x)$. Hence, ($\widetilde{\text{TVP}}$) has a solution whenever $|\widetilde{f}(x)| \le c(1+\sum_{i=1}^d e^{px_i})$ for some $p>0$. Applying Lemma 5.1.4 gives the existence of solutions to (TVP$^+$) under the assumption that the terminal condition is continuous and has at most polynomial growth. □

Remark 5.1.6. The preceding argument implies that if $|f(x)| \le c(1+|x|^p)$, then the corresponding solution v of the terminal value problem (TVP$^+$) will satisfy $|v(t,x)| \le \widetilde{c}(1+|x|^p)$ for a certain constant $\widetilde{c}$ and with the same exponent p.

In the next step, we will extend the preceding hedging argument to the case of a path-dependent *exotic option*, whose payoff is usually given by

$$H = h(S(t_0), \dots, S(t_N)), \tag{5.1.9}$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ denote the fixing times of the daily closing prices and h is a certain function.

Theorem 5.1.7. *Assume that the conditions of Theorem 5.1.3 are satisfied and h in* (5.1.9) *is a locally Lipschitz continuous function on $(\mathbb{R}^d_{(+)})^{N+1}$ with a Lipschitz constant that grows at most polynomially. That is, there exist $p \ge 0$ and $L \ge 0$ such that, for $|x_i|, |y_i| \le m$, we have*

$$\big|h(x_0,\dots,x_N) - h(y_0,\dots,y_N)\big| \le (1+m^p)L\sum_{i=0}^N |x_i - y_i|.$$

Then, letting

$$v_N(t,x_0,\dots,x_N,x) := h(x_0,\dots,x_N) \quad \text{for} \quad t\in[0,T],\ x\in\mathbb{R}^d_{(+)},$$

the following recursive scheme for functions $v_k : [t_k,t_{k+1}]\times(\mathbb{R}^d_{(+)})^{k+1}\times\mathbb{R}^d_{(+)} \to \mathbb{R}$, $k = 0,\dots,N-1$, is well-defined.

- *For $k = N-1, N-2, \dots, 0$, the function $f_{k+1}(x) := v_{k+1}(t_{k+1}, x_0,\dots,x_k,x,x)$ is continuous in x, and $(t,x) \mapsto v_k(t,x_0,\dots,x_k,x)$ solves the terminal value problem ($TVP^{(+)}$) with terminal condition f_{k+1} at time t_{k+1}.*

Remark 5.1.8. Note that the assumption concerning the local Lipschitz continuity of h in the preceding statement can sometimes be relaxed when it comes to more specific examples. Such instances are the (pathwise) versions of the Bachelier and Black–Scholes models, which arise when choosing $a_{ij}(t,x) = \widetilde{a}_{ij}$ for a constant positive definite matrix $(\widetilde{a}_{ij})$. For these special cases, the recursive scheme in Theorem 5.1.7 can be solved for large classes of payoff functions h without the assumption of local Lipschitz continuity. Moreover, even the continuity of h may be relaxed, which allows to account for discontinuous payoffs as, e.g., in barrier options. This also holds for the strictly pathwise hedging argument that will be formulated after the proof of Theorem 5.1.7.

Proof of Theorem 5.1.7. In a first step, we will prove the result for the terminal value problem (TVP). Clearly, the function v_k will be well-defined if f_{k+1} is continuous and has at most polynomial growth. These two properties will follow if v_{k+1} satisfies the following three conditions;

(i) $(x_0, \ldots, x_{k+1}, x) \mapsto v_{k+1}(t, x_0, \ldots, x_{k+1}, x)$ has at most polynomial growth;

(ii) $x \mapsto v_{k+1}(t_{k+1}, x_0, \ldots, x_{k+1}, x)$ is continuous for all $x_0, \ldots, x_{k+1}$;

(iii) $(x_0, \ldots, x_{k+1}) \mapsto v_{k+1}(t, x_0, \ldots, x_{k+1}, x)$ is locally Lipschitz continuous, uniformly in t and locally uniformly in x, with a Lipschitz constant that grows at most polynomially. More precisely, there exist $p \geq 0$ and $L \geq 0$ such that, for $|x|, |x_i|, |y_i| \leq m$ and $t \in [t_{k+1}, t_{k+2}]$,

$$\Big| v_{k+1}(t, x_0, \ldots, x_{k+1}, x) - v_{k+1}(t, y_0, \ldots, y_{k+1}, x) \Big| \leq (1 + m^p) L \sum_{i=0}^{k+1} |x_i - y_i|.$$

We will now show that v_k inherits the properties (i), (ii), and (iii) from v_{k+1}. Since v_N obviously satisfies these properties, backward induction will then complete the proof in the case of (TVP).

To show (i), let $p, c > 0$ be such that $\widetilde{f}_{k+1}(x) := c(|x_0|^p + \cdots + |x_k|^p + |x|^p + |x|^p)$ satisfies $-\widetilde{f}_{k+1} \leq f_{k+1} \leq \widetilde{f}_{k+1}$. Moreover, let $\widetilde{v}_k(t, x_0, \ldots, x_k, x)$ be the solution of (TVP) with terminal condition $\widetilde{f}_{k+1}$ at time t_{k+1}. Theorem 5.1.3, [56, Theorem A.7], and the linearity of solutions imply that $(x_0, \ldots, x_k, x) \mapsto \widetilde{v}_k(t, x_0, \ldots, x_k, x)$ has at most polynomial growth, while the maximum principle in the form of [56, Theorem A.5] implies that $-\widetilde{v}_k \leq v_k \leq \widetilde{v}_k$. This shows (i).

There is no need to show (ii), since solutions to (TVP) are continuous by construction.

To show (iii), let p and L be as in (iii) and x_i, y_i be given. We choose m so that $m \geq |x_i| \vee |y_i|$ for $i = 1, \ldots, k$ and let $\delta := L \sum_{i=0}^{k} |x_i - y_i|$. Then

$$-(1 + m^p + |x|^p)\delta \leq v_{k+1}(t_{k+1}, x_0, \ldots, x_k, x, x) - v_{k+1}(t_{k+1}, y_0, \ldots, y_k, x, x) \leq (1 + m^p + |x|^p)\delta.$$

Now we define $u(t, x)$ as the solution of (TVP) with terminal condition $u(t_{k+1}, x) = |x|^p$ at time t_{k+1}. Theorem 5.1.3 implies that u is well defined, and the maximum principle and [56, Theorem A.7] yield that $0 \leq u(t, x) \leq c|x|^p$ for some constant $c \geq 0$. Again applying the maximum principle we infer that

$$-(1 + m^p + u(t, x))\delta \leq v_k(t, x_0, \ldots, x_k, x) - v_k(t, y_0, \ldots, y_k, x) \leq (1 + m^p + u(t, x))\delta$$

for all t and x, which shows that (iii) holds for v_k with the same constant p and the new Lipschitz constant $(1 + c)L$.

Let us now turn to the proof for the terminal value problem (TVP$^+$). The proof of Theorem 5.1.3 (b) shows that (TVP$^+$) inherits the maximum principle from $(\widehat{\text{TVP}})$. Moreover, Remark 5.1.6 gives us that v_k inherits property (i) from v_{k+1}. So Remark 5.1.6 can be used instead of [56, Theorem A.7] in the preceding argument. Therefore, the proof for (TVP$^+$) can be carried out in the same way as for the terminal value problem (TVP). □

Now let H be an exotic option as in (5.1.9) and assume that the recursive scheme in Theorem 5.1.7 holds for functions v_k, $k = 0, \dots, N$. Denoting by $\nabla_x v_k$ the gradient of the function $x \mapsto v_k(t, x_0, \dots, x_k, x)$, we obtain that

$$\begin{aligned} \xi^S(t) &:= \nabla_x v_k(t, S(t_0), \dots, S(t_k), S(t)), \\ \eta^S(t) &:= v_k(t, S(t_0), \dots, S(t_k), S(t)) - \xi^S(t) \cdot S(t), \end{aligned} \qquad \text{for } t \in [t_k, t_{k+1}), \qquad (5.1.10)$$

is a self-financing trading strategy on each interval $[t_k, t_{k+1})$ in the following sense:

$$\xi^S(t) \cdot S(t) + \eta^S(t) = v_k(t, S(t_0), \dots, S(t_k), S(t)) = v_k(t_k, S(t_0), \dots, S(t_k), S(t_k)) + \int_{t_k}^t \xi^S(s)\, \mathrm{d}S(s).$$

By continuity of $t \mapsto v_k(t, S(t_0), \dots, S(t_k), S(t))$ we infer that the limit

$$\int_{t_k}^{t_{k+1}} \xi^S(s)\, \mathrm{d}S(s) := \lim_{t \uparrow t_{k+1}} \int_{t_k}^t \xi^S(s)\, \mathrm{d}S(s)$$

exists, which allows us to define

$$\int_0^t \xi^S(s)\, \mathrm{d}S(s) := \sum_{k=0}^{\ell-1} \int_{t_k}^{t_{k+1}} \xi^S(s)\, \mathrm{d}S(s) + \int_{t_\ell}^t \xi^S(s)\, \mathrm{d}S(s), \qquad t \in [0, T], \qquad (5.1.11)$$

where ℓ is the largest k such that $t_k < t$. Using these notations, we arrive at the following Delta hedging result.

Corollary 5.1.9. *Let H be an exotic option of the form* (5.1.9) *and suppose that the recursive scheme in Theorem 5.1.7 holds for functions v_k, $k = 0, \dots, N$. Then, for each $S \in \mathscr{S}_a^{(+)}$, the strategy* (5.1.10) *is self-financing in the above sense and satisfies*

$$\lim_{t \uparrow T} \xi^S(t) \cdot S(t) + \eta^S(t) = v_0(0, S(t_0)) + \int_0^T \xi^S(s)\, \mathrm{d}S(s) = h(S(t_0), \dots, S(t_N)).$$

In this sense, (ξ^S, η^S) is a strictly pathwise Delta hedging strategy for H.

The preceding corollary establishes a general, strictly pathwise hedging result for a large class of exotic options arising in practice. It also identifies $v_0(0, S(t_0))$ as the amount of cash that is required at time $t = 0$ in order to be able to perfectly replicate the payoff H for all price trajectories in the class $\mathscr{S}_a^{(+)}$. In standard continuous-time finance, this quantity is then usually referred to as an *arbitrage-free price* for H. In our strictly pathwise setting, however, an essential ingredient is missing for the identification of $v_0(0, S(t_0))$ as an arbitrage-free price: It is an open question whether our class of trading strategies is indeed arbitrage-free with respect to all possible price trajectories in $\mathscr{S}_a^{(+)}$. The next section is therefore devoted to investigating this question. The corresponding result, Theorem 5.2.4, yields sufficient conditions ensuring that strategies, as those in Corollary 5.1.9, do indeed not generate arbitrage in our strictly pathwise setting. Theorem 5.2.4 will be the main result of this chapter.

Remark 5.1.10 (Robustness of the hedging strategy)**.** Note that the strategy (5.1.10) gives a perfect hedge for the exotic option with payoff H only in case that the realized underlying trajectory, S, belongs to the class $\mathscr{S}_a^{(+)}$. In real world, however, the realized quadratic variation is often subject to risk and uncertainty, and therefore it may turn out *a posteriori* that S actually does not belong to $\mathscr{S}_a^{(+)}$. If S nevertheless belongs to the class QV^d, this issue is then termed volatility uncertainty. One possible approach to volatility uncertainty was introduced in [63], where, for the case of $H = h(S(T))$, the linear equation (TVP$^+$) is replaced by a certain nonlinear partial differential equation corresponding to a worst-case approach within a class of price trajectories whose realized volatility may vary within a prescribed set. Another approach to volatility uncertainty was proposed in [36] for the case $d = 1$. Although [36] uses a diffusion framework, it is straightforward to transfer the comparison result from [36, Theorem 6.2] to a strictly pathwise framework. For options of the form $H = h(S(T))$ with $h \geq 0$ convex, one then gets that the Delta hedge (5.1.10) is *robust* in the sense that it is still a superhedge as long as a overestimates the realized quadratic variation, i.e., $\int_r^t a(s, S(s))\,\mathrm{d}s \geq [S,S](t) - [S,S](r)$ for $0 \leq r \leq t \leq T$. Thus, if a Delta hedging strategy is robust, then its performance can be monitored by comparing $a(t, S(t))$ to the realized quadratic variation $[S,S]$. The paper [79] analyzed to what extent the preceding result can be generalized to exotic payoffs of the form $H = h(S(t_0), \ldots, S(t_N))$. It was shown that robustness then actually breaks down for a large class of relevant convex payoff functions h, but that it still holds if h satisfies the property of directional convexity (see [83, Remark 2.7]).

5.2 Absence of pathwise arbitrage

We will now study the absence of pathwise arbitrage in a class of trading strategies that is naturally suggested by the pathwise Delta hedging strategies, as constructed in Theorem 5.1.7 and Corollary 5.1.9. The reader is referred to the paragraph preceding Remark 5.1.10 for further details on the motivation of this subject. First, we introduce the class of strategies that it will make sense to consider.

Definition 5.2.1. Assume that $N \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_N = t_{N+1} = T$, and v_k $(k = 0, \ldots, N)$ are real-valued continuous functions on $[t_k, t_{k+1}] \times (\mathbb{R}^d_{(+)})^{k+1} \times \mathbb{R}^d_{(+)}$ such that, for $k = 0, \ldots, N-1$, the function $(t, x) \mapsto v_k(t, x_0, \ldots, x_k, x)$ solves (TVP$^{(+)}$) with terminal condition $f_{k+1}(x) := v_{k+1}(t_{k+1}, x_0, \ldots, x_k, x, x)$ at time t_{k+1}. For $S \in \mathscr{S}_a^{(+)}$, we then define ξ^S as in (5.1.10) and

$$V_\xi^S(t) := v_0(0, S(0)) + \int_0^t \xi^S(s)\,\mathrm{d}S(s), \tag{5.2.1}$$

where the pathwise Itô integral is understood as in (5.1.11). We use the notation $\mathscr{X}^{(+)}$ for the collection of all pairs $(v_0(0,\cdot), \xi^\cdot)$ arising in this way.

We note that, although Theorem 5.1.7 gives sufficient conditions for the existence of a family of functions (v_k) as in the preceding definition, these conditions are quite restrictive in nature and not necessary. In particular, as already mentioned above, in many applications it is possible to relax the local Lipschitz continuity of the terminal function v_N. Next, we proceed to defining our strictly pathwise notion of arbitrage.

Definition 5.2.2 ((Admissible) arbitrage opportunity)**.** We call a pair $(v_0(0,\cdot),\xi^{\cdot}) \in \mathscr{X}^{(+)}$ an *arbitrage opportunity* for $\mathscr{S}_a^{(+)}$ if the following conditions apply.

(a) $V_\xi^S(T) \geq 0$ for all $S \in \mathscr{S}_a^{(+)}$.

(b) There exists at least one $S \in \mathscr{S}_a^{(+)}$ for which $V_\xi^S(0) = v_0(0, S(0)) \leq 0$ and $V_\xi^S(T_0) > 0$ for some $T_0 \in (0, T]$.

Moreover, we call an arbitrage opportunity $(v_0(0,\cdot),\xi^{\cdot})$ *admissible* if also the following condition is satisfied.

(c) There exists a constant $c \geq 0$ such that $V_\xi^S(t) \geq -c$ for all $S \in \mathscr{S}_a^{(+)}$ and $t \in [0, T]$.

Remark 5.2.3. Before we state our main result, let us first make a few comments concerning the preceding definition (see [83, Section 3]). Condition (a) is tantamount to say that one can follow the strategy $(v_0(0,\cdot),\xi^{\cdot})$ up to time T without incurring the risk of ending up with negative wealth at the terminal time. Now, let S be as in condition (b). The initial spot value, $S_0 := S(0)$, will then satisfy $v_0(0, S_0) = V_\xi^S(0) \leq 0$, whence we infer that for any price trajectory $\widetilde{S} \in \mathscr{S}_a^{(+)}$ with $\widetilde{S}(0) = S_0$, only a nonpositive initial investment $v_0(0, S_0) = V^{\widetilde{S},\xi}(0)$ is required in order to end up with the nonnegative terminal wealth $V^{\widetilde{S},\xi}(T) \geq 0$. Beside this, for the particular price trajectory S, there exists a time T_0 at which one can make the strictly positive profit $V_\xi^S(T_0) > 0$. This profit can be locked in, e.g., by halting all trading from time T_0 onward. In this sense, the strategy $(v_0(0,\cdot),\xi^{\cdot})$ is indeed an arbitrage opportunity. The last condition, condition (c), is a constraint on strategies $(v_0(0,\cdot),\xi^{\cdot})$ that corresponds to the admissibility constraint usually imposed in continuous-time probabilistic models in order to exclude doubling-type strategies. Indeed, e.g., Dudley's result [32] implies that standard diffusion models typically admit arbitrage opportunities in the class of strategies whose value process is not bounded from below (see also the discussion in [57, Section 1.6.3]). For our pathwise setting, in Example 5.2.5 below, we construct an arbitrage opportunity that does not satisfy condition (c).

Theorem 5.2.4 (Absence of admissible arbitrage)**.** *Assume that $a(t, x)$ is continuous, bounded, and positive definite for all $(t, x) \in [0, \widetilde{T}] \times \mathbb{R}^d_{(+)}$, where $\widetilde{T} > T$. Then there are no admissible arbitrage opportunities in $\mathscr{X}^{(+)}$.*

Example 5.2.5 (A non-admissible arbitrage opportunity)**.** We assume that $d = 1$ and $a \equiv 2$. In this setting the assumptions of Theorem 5.2.4 are clearly satisfied. Moreover, $\mathscr{L} = \partial^2/\partial x^2$ and (TVP) is nothing else than the time-reversed Cauchy problem for the standard heat equation. There are many explicit examples of non-zero functions v that solve (TVP) with terminal condition $f \equiv 0$; see, e.g., [98, Section II.6]. Applying Widder's uniqueness theorem for nonnegative solutions of the heat equation, [98, Theorem VIII.2.2], it follows that any such function v must be unbounded from above and from below on every nontrivial strip $[t,T]\times\mathbb{R}$ with $t<T$. In particular, there must be $0\le t_0<t_1<T$ and $x_0,x_1\in\mathbb{R}$ such that $v(t_0,x_0)=0$ and $v(t_1,x_1)>0$. Using an appropriate time shift, we can assume without loss of generality that $t_0=0$. It follows easily that $\mathscr{S}_a$ contains trajectories that can connect the two points x_0 and x_1 within time t_1-t_0, which implies that the function v gives rise to an arbitrage opportunity (see also [83, Example 3.4]).

Proof of Theorem 5.2.4. First, we prove the result for the class $\mathscr{X}$. Suppose by way of contradiction that there exists an admissible arbitrage opportunity in $\mathscr{X}$, and let $0=t_0<t_1<\cdots<t_N=t_{N+1}=T$ and v_k be the corresponding time points and functions as in Definition 5.2.1.

Our assumptions ensure that the martingale problem for the operator $\mathscr{L}$ is well-posed (see [88]). We denote by $\mathbb{P}_{t,x}$ the corresponding Borel probability measures on $C([t,T],\mathbb{R}^d)$ under which the coordinate process, $(X(u))_{t\le u\le T}$, is a diffusion process with generator $\mathscr{L}$ and satisfies $X(t)=x$ $\mathbb{P}_{t,x}$-a.s. Thus in particular, $(X_i(u))_{t\le u\le T}$ is a continuous local $\mathbb{P}_{t,x}$-martingale for $i=1,\dots,d$. Moreover, the support theorem [89, Theorem 3.1] gives us that the law of $(X(u))_{t\le u\le T}$ under $\mathbb{P}_{t,x}$ has full support on $C_x([t,T],\mathbb{R}^d):=\{\omega\in C([t,T],\mathbb{R}^d)\,|\,\omega(t)=x\}$.

In a first step, we will use these facts to establish that all functions v_k are nonnegative. To this end, observe first that the support theorem implies that the law of $(X(t_1),\dots,X(t_N))$ under $\mathbb{P}_{0,x}$ has full support on $(\mathbb{R}^d)^N$. Since $\mathbb{P}_{0,x}$-a.e. trajectory in $C_x([0,T],\mathbb{R}^d)$ belongs to $\mathscr{S}_a$, it follows that the set $\big\{(S(t_1),\dots,S(t_N))\,|\,S\in\mathscr{S}_a,\ S(0)=x\big\}$ is dense in $(\mathbb{R}^d)^N$. Using condition (a) of Definition 5.2.2 and the continuity of v_N thus gives that $v_N(T,x_0,\dots,x_{N+1})\ge 0$ for all $x_0,\dots,x_{N+1}$. In the same way, we obtain from the admissibility of the arbitrage opportunity that $v_k(t,x_0,\dots,x_k,x)\ge -c$ for all k, $t\in[t_k,t_{k+1}]$ and $x_0,\dots,x_k,x\in\mathbb{R}^d$.

Let us for the moment fix $x_0,\dots,x_{N-1}$ and consider the function

$$u(t,x):=v_{N-1}(t,x_0,\dots,x_{N-1},x).$$

Let $Q\subset\mathbb{R}^d$ be a bounded domain whose closure is contained in $\mathbb{R}^d$ and denote by $\tau:=\inf\{s\,|\,X(s)\notin Q\}$ the first exit time from Q. Itô's formula, in conjunction with the fact that u solves (TVP), implies that we have $\mathbb{P}_{t,x}$-a.s. for $t\in[t_{N-1},T)$,

$$u(T\wedge\tau,X(T\wedge\tau))=u(t,x)+\int_t^{T\wedge\tau}\nabla_x u(s,X(s))\,\mathrm{d}X(s). \tag{5.2.2}$$

Since $\nabla_x u$ and the coefficients of $\mathscr{L}$ are bounded in the closure of Q, the stochastic integral on the right-hand side is a true martingale. Therefore,

$$u(t,x)=\mathbb{E}_{t,x}[\,u(T\wedge\tau,X(T\wedge\tau))\,]. \tag{5.2.3}$$

Now we take an increasing sequence $Q_1\subset Q_2\subset\cdots$ of bounded domains exhausting $\mathbb{R}^d$ and whose closures are contained in $\mathbb{R}^d$. Let τ_n denote the exit time from Q_n. Then, an application of (5.2.3) for each τ_n, Fatou's lemma in conjunction with the fact that $u\ge -c$, and the already established nonnegativity of $u(T,\cdot)$ yield

$$u(t,x)=\lim_{n\uparrow\infty}\mathbb{E}_{t,x}[\,u(T\wedge\tau_n,X(T\wedge\tau_n))\,]\ge\mathbb{E}_{t,x}[\,u(T,X(T))\,]\ge 0. \tag{5.2.4}$$

This shows the nonnegativity of v_{N-1} and in particular of the terminal condition f_{N-1} for v_{N-2}. The preceding argument may therefore be repeated for v_{N-2} and so forth. Hence, $v_k \geq 0$ for all k.

Now let $S \in \mathscr{S}_a$ and T_0 be such that $V_\xi^S(0) \leq 0$ and $V_\xi^S(T_0) > 0$, which exist according to the assumption made at the beginning of this proof. If k is such that $t_k < T_0 \leq t_{k+1}$ and $x_0 := S(0)$, then $v_0(0, x_0) = 0$ and $v_k(T_0, S(t_0), \ldots, S(t_k), S(T_0)) > 0$. By continuity, we actually have $v_k(T_0, \cdot) > 0$ in an open neighborhood $U \subset C_{x_0}([0,T], \mathbb{R}^d)$ of the path S.

Since $\mathbb{P}_{0,x_0}$-a.e. sample path belongs to $\mathscr{S}_a$, Itô's formula gives that $\mathbb{P}_{0,x_0}$-a.s.,

$$v_k(T_0, X(t_0), \ldots, X(t_k), X(T_0)) = v_0(0, x_0) + \int_0^{T_0} \xi^X(t)\, \mathrm{d}X(t).$$

Using a localization argument as in (5.2.4), in conjunction with the fact that $v_\ell \geq 0$ for all ℓ, implies

$$0 = v_0(0, x_0) \geq \mathbb{E}_{0,x_0}\big[\, v_k(T_0, X(t_0), \ldots, X(t_k), X(T_0))\,\big] \geq 0.$$

Another application of the support theorem now yields a contradiction to the fact that $v_k(T_0, \cdot) > 0$ in the open set U, which completes the proof for the class $\mathscr{X}$.

Now we turn to the proof for the class $\mathscr{X}^+$. In this case, the martingale problem for the operator $\widetilde{\mathscr{L}}$ defined in (5.1.7) is again well-posed, since the coefficients of $\widetilde{\mathscr{L}}$ are bounded and continuous (see [88]). These properties of the coefficients also ensure that the support theorem holds [89, Theorem 3.1]. Now, if $(\widetilde{\mathbb{P}}_{s,x}, \widetilde{X})$ is a corresponding diffusion process, we can consider the laws of $X(t) := \exp(\widetilde{X}(t))$ and, applying Lemma 5.1.4, we get a solution to the martingale problem for $\mathscr{L}^+$, which satisfies the support theorem with state space $\mathbb{R}^d_+$. The arguments from the proof for $\mathscr{X}$ can now simply be repeated in order to obtain the result for the class $\mathscr{X}^+$. □

5.3 Extension to functionally dependent strategies

Recall from (5.1.9) our representation $H = h(S(t_0), \ldots, S(t_N))$ of the payoff of an exotic option, based on asset prices sampled at the $N+1$ dates $0 = t_0 < t_1 < \cdots < t_N = T$. If N is large, it may be convenient to use a continuous-time approximation of the payoff H. For instance, the payoff $H = (\frac{1}{N}\sum_{n=1}^N S_1(t_n) - K)^+$ of an average-price Asian call option on the first asset, S^1, can be approximated by a call option based on a continuous-time average of asset prices, $H \approx (\frac{1}{T}\int_0^T S_1(t)\, dt - K)^+$. Approximations of this type may be easier to treat analytically and are standard in the textbook literature. In this section, we extend our preceding results to a situation that covers such continuous-time approximations of (5.1.9), i.e., we will consider payoffs of the form $H(S^T) = H(S)$, where S describes the entire path of the underlying price trajectory up to time T, and H is a suitable map from the Skorohod space $D([0,T], \mathbb{R}^d)$ to $\mathbb{R}$. The notation is as in Chapter 3. In particular, it follows from the continuous change of variables formula for non-anticipative functionals on path space, Theorem 3.2.1, that we can define "general admissible integrands" ξ in the following way, so as to ensure that the pathwise Itô integral $\int_r^t \xi(s)\, \mathrm{d}S(s)$ exists for all $t \in [r,u] \subset [0,T]$ as the finite limit of Riemann sums.

Definition 5.3.1 (General admissible integrands)**.** For $0 \le r < u \le T$, an $\mathbb{R}^d$-valued function $[r,u] \ni t \mapsto \xi(t)$ is called a *general admissible integrand* for $S \in QV^d$, if there exist $m \in \mathbb{N}$, a continuous function $A : [r,u] \to \mathbb{R}^m$ whose components are functions of bounded variation, and a non-anticipative left-continuous functional $F \in \mathbb{C}^{1,2}([r,u])$ with $\mathcal{D}F,\ \nabla_X F,\ \nabla_X^2 F \in \mathbb{B}$, so that

$$\xi(t) = \nabla_X F(t, S^t_{[r,u]}, A^t),$$

where $\nabla_X F$ is the vertical derivative of F with respect to X and $S_{[r,u]} := S\big|_{[r,u]}$ is the restriction of S to the interval $[r,u]$.

Note that in order to account for model uncertainty in a non-Markovian setting, we could have "lifted" the classes $\mathscr{S}_a$ and $\mathscr{S}_a^+$ from above as follows:

$$\widetilde{\mathscr{S}_a} := \left\{ S \in QV^d \,\middle|\, [S_i, S_j](t) = \int_0^t \widetilde{a}_{ij}(s, S^s)\,\mathrm{d}s,\ t \in [0,T],\ 1 \le i,j \le d \right\}$$

and

$$\widetilde{\mathscr{S}_a^+} := \left\{ S \in QV^d \,\middle|\, S_i(t) > 0,\ [S_i, S_j](t) = \int_0^t \widetilde{a}_{ij}(s, S^s) S_j(s) S_j(s)\,\mathrm{d}s \right\},$$

where $\widetilde{a}(t, X^t) = (\widetilde{a}_{ij}(t, X^t))_{i,j=1,\ldots,d}$ would be a non-anticipative functional mapping the path space $\bigcup_{t\in[0,T]} \mathcal{U}^t$ with $U = \mathbb{R}^d$ (respectively $\bigcup_{t\in[0,T]} \mathcal{U}^t$ with $U = \mathbb{R}^d_+$ in case of $\widetilde{\mathscr{S}_a^+}$) into the set of positive definite symmetric $d \times d$-matrices. Price trajectories in $\widetilde{\mathscr{S}_a^+}$ can, for instance, arise as sample paths of multi-dimensional *path-dependent* local volatility models. However, for our purpose, which is establishing conditions on the covariance structure of the price paths under which no admissible arbitrage opportunities exist, we must stick to the choice of Markovian volatility in order to be able to apply a support theorem later on.

In analogy to Proposition 5.1.2, we will now characterize self-financing trading strategies that may depend on the entire past evolution of the particular realization $S \in \mathscr{S}_a^{(+)}$. Let us denote by $\mathcal{R}^t_{(+)} \subset D([0,T], \mathbb{R}^d_{(+)})$ the set of $\mathbb{R}^d_{(+)}$-valued càdlàg paths stopped at time t. Analogously, $\mathcal{R}^t_{(+),I} \subset D(I, \mathbb{R}^d_{(+)})$ is the set of $\mathbb{R}^d_{(+)}$-valued càdlàg paths on I stopped at t. Recalling Definition 3.1.8, for an interval $I \subset [0,T]$, we denote by $\mathbb{C}^{1,2}(I)$ the class of all non-anticipative functionals on $\bigcup_{t\in I} \mathcal{R}^t_{(+),I}$ that are continuous at fixed times, locally uniformly, and admit left-continuous horizontal and first- and second-order vertical derivatives. Thus, lifting the second-order differential operators $\mathscr{L}$ and $\mathscr{L}^+$ yields the following operators on path space:

$$\mathscr{A} := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, X(t)) \nabla^2_{X,ij} \qquad \text{and} \qquad \mathscr{A}^+ := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, X(t)) X_i(t) X_j(t) \nabla^2_{X,ij}.$$

The next proposition establishes a BS-type PDE that the portfolio value of a self-financing strategy must satisfy, which can be regarded as a hedging PDE expressing the link between time decay and convexity.

Proposition 5.3.2. *Assume that $0 \le r < u \le T$ and let $F : [r,u] \times \mathcal{R}_{(+),[r,u]} \mapsto \mathbb{R}$ be a left-continuous non-anticipative functional of class $\mathbb{C}^{1,2}([r,u])$ with $\mathcal{D}F$, $\nabla_X F$, $\nabla_X^2 F \in \mathbb{B}$. Denote $S_{[r,u]}$ the restriction of S to the interval $[r,u]$. Then the following conditions are equivalent.*

(a) *For each $S \in \mathscr{S}_a^{(+)}$, there exists a general admissible integrand ξ^S on $[r,u]$ such that*

$$F(t, S^t_{[r,u]}) = F(r, S^r_{[r,u]}) + \int_r^t \xi^S(s)\,\mathrm{d}S(s) \qquad \text{for } t \in [r,u].$$

(b) *The functional F satisfies the path-dependent parabolic equation*

$$\mathcal{D}F + \mathscr{A}^{(+)}F = 0 \qquad \text{on } \mathscr{S}_a^{(+)}\big|_{[r,u]}. \tag{5.3.1}$$

Moreover, if these equivalent conditions hold, then ξ^S in (a) *must necessarily be of the form*

$$\xi^S(t) = \nabla_X F(t, S^t_{[r,u]}). \tag{5.3.2}$$

Proof. The proof is analogous to the proof of Proposition 5.1.2. The functional change of variables formula for continuous paths, in the form of Theorem 3.2.1, yields for $S \in \mathscr{S}_a^{(+)}$ that

$$F(t, S^t_{[r,u]}) = F(r, S^r_{[r,u]}) + \int_r^t \nabla_X F(s, S^s_{[r,u]})\,\mathrm{d}S(s) + \int_r^t \Big(\mathcal{D}F(s, S^s_{[r,u]}) + \mathscr{A}^{(+)}F(s, S^s_{[r,u]})\Big)\,\mathrm{d}s. \tag{5.3.3}$$

Thus, (b) implies (a) and (5.3.2) must hold.

Let us now assume that (a) holds. Then

$$\int_r^t \big(\xi^S(s) - \nabla_X F(s, S^s_{[r,u]})\big)\,\mathrm{d}S(s) = \int_r^t \Big(\mathcal{D}F(s, S^s_{[r,u]}) + \mathscr{A}^{(+)}F(s, S^s_{[r,u]})\Big)\,\mathrm{d}s.$$

Because the right-hand side has vanishing quadratic variation (see [86, Proposition 2.2.2]), the same must hold for the left-hand side. By Proposition 3.2.7, the quadratic variation of the left-hand side is given by

$$\int_r^t \big(\xi^S(s) - \nabla_X F(s, S^s_{[r,u]})\big)^\top a(s, S(s))\big(\xi^S(s) - \nabla_X F(s, S^s_{[r,u]})\big)\,\mathrm{d}s$$

in case of $S \in \mathscr{S}_a$. Differentiating with respect to t gives

$$\big(\xi^S(t) - \nabla_X F(t, S^t_{[r,u]})\big)^\top a(t, S(t))\big(\xi^S(t) - \nabla_X F(t, S^t_{[r,u]})\big) = 0$$

for all t, and using the fact that the matrix $a(t, S(t))$ is positive definite we infer that (5.3.2) must hold. For $S \in \mathscr{S}_a^+$, the matrix $a(s, S(s))$ has to be replaced by the matrix with components $a_{ij}(s, S(s))S_i(s)S_j(s)$, and we obtain (5.3.2) by the same arguments as for the case $S \in \mathscr{S}_a$. Plugging (5.3.2) into (5.3.3) we infer with (a) that the rightmost integral in (5.3.3) is identically zero, which establishes (b) by again taking the derivative with respect to t. □

Now suppose that for suitably given $H : \mathcal{R}^T_{(+)} \to \mathbb{R}$ there exists a solution F to the following path-dependent terminal-value problem,

$$\text{(FTVP}^{(+)}\text{)} \qquad \begin{cases} F \in \mathbb{C}^{1,2}([0,T]), & \text{satisfies the conditions from Definition 5.3.1,} \\ \mathcal{D}F + \mathscr{A}^{(+)}F = 0 & \text{in } \bigcup_{t\in[0,T)} \mathcal{R}^t_{(+)}, \\ F(T, X^T) = H(X^T) & \text{for } X^T \in \mathcal{R}^T_{(+)}. \end{cases}$$

Note that the terminal condition H needs to be defined on the Skorohod space $\mathcal{R}^T_{(+)} = D([0,T], \mathbb{R}^d_{(+)})$, instead of $C([0,T], \mathbb{R}^d_{(+)})$. This is due to the fact that the computation of pathwise derivatives requires applying discontinuous shocks.

Then, for $S \in \mathscr{S}^{(+)}_a$ and $t \in [0,T)$, we can define

$$\xi^S(t) := \nabla_X F(t, S^t) \qquad \text{and} \qquad \eta^S(t) := F(t, S^t) - \xi^S(t) \cdot S(t). \tag{5.3.4}$$

Applying Proposition 5.3.2 yields

$$\xi^S(t) \cdot S(t) + \eta^S(t) = F(t, S^t) = F(0, S^0) + \int_0^t \xi^S(s)\, \mathrm{d}S(s), \tag{5.3.5}$$

whence we infer that (ξ^S, η^S) is self-financing with portfolio value $V^S(t) = F(t, S^t)$. Moreover, the left-continuity of the functional F on $[0,T]$, in conjunction with the continuity of S, implies that the limit $V^S(T) := \lim_{t\uparrow T} V^S(t)$ exists and satisfies

$$V^S(T) = H(S) \qquad \text{for all } S \in \mathscr{S}^{(+)}_a.$$

Thus, (ξ^S, η^S) is a strictly pathwise Delta hedging strategy for H.

In the next step, we will derive conditions yielding the existence and uniqueness of solutions to (FTVP) and (FTVP$^+$). Path-dependent PDEs such as (5.3.1) are closely related to backward stochastic differential equations (BSDEs) generalizing the (functional) Feynman-Kac formula [34]. In [74], a one-to-one correspondence between a functional BSDE and a path-dependent PDE was established for the Brownian case. In [58], this result was generalized to the case of solutions of stochastic differential equations with functionally dependent drift and diffusion coefficients. We will now use [58, Theorem 20] to formulate conditions under which (FTVP) and (FTVP$^+$) admit unique solutions. To this end, we will need the following regularity conditions from [74, Definition 3.1].

Definition 5.3.3. The function $H : \mathcal{R}^T_{(+)} \mapsto \mathbb{R}$ is of class $C^2(\mathcal{R}^T_{(+)})$ if for all $X \in \mathcal{R}^T_{(+)}$ and $t \in [0,T]$, there exist $p_1 \in \mathbb{R}^d$ and $p_2 \in \mathbb{R}^d \times \mathbb{R}^d$ so that p_2 is symmetric and the following holds

$$H(X_{X^{t,h}}) - H(X) = p_1 \cdot h + \frac{1}{2} h^\top p_2 h + o(|h|^2), \quad h \in \mathbb{R}^d,$$

where $X_{X^{t,h}}(u) := X(u)\mathbb{1}_{[0,t)}(u) + (X(u)+h)\mathbb{1}_{[t,T]}(u)$. We denote $H'_{X^t}(X) := p_1$ and $H''_{X^t}(X) := p_2$. We say that $H : \mathcal{R}^T_{(+)} \mapsto \mathbb{R}$ is of class $C^2_{l,lip}(\mathcal{R}^T_{(+)})$ if $H'_{X^t}(X)$ and $H''_{X^t}(X)$ exist for all

$X \in \mathcal{R}^T_{(+)}$, $t \in [0,T]$, and there are constants $C, k > 0$ such that for all $X, Y \in \mathcal{R}^T_{(+)}$ (with $\|\cdot\|$ denoting the supremum norm),

$$|H(X) - H(Y)| \leq C(1 + \|X\|^k + \|Y\|^k)\|X - Y\|,$$

$$|H'_{X^t}(X) - H'_{Y^s}(Y)| \leq C(1 + \|X\|^k + \|Y\|^k)(|t-s| + \|X - Y\|),\ t, s \in [0,T],$$

$$|H''_{X^t}(X) - H''_{Y^s}(Y)| \leq C(1 + \|X\|^k + \|Y\|^k)(|t-s| + \|X - Y\|),\ t, s \in [0,T].$$

Theorem 5.3.4. *Suppose that the terminal condition H of* (FTVP$^{(+)}$) *is of class $C^2_{l,lip}(\mathcal{R}^T_{(+)})$. Then, under the following conditions,* (FTVP$^{(+)}$) *admits a unique solution $F \in \mathbb{C}^{1,2}([0,T])$.*

(a) (Theorem 20 in [58]) *In case of* (FTVP)*, we suppose that $a(t, X(t)) = \sigma(t, X(t))\sigma(t, X(t))^\top$ with a Lipschitz continuous volatility matrix σ.*

(b) *In case of* (FTVP$^+$)*, we suppose that $a(t, X(t)) = \sigma(t, X(t))\sigma(t, X(t))^\top$ with a Lipschitz continuous volatility matrix σ such that $a_{ii}(t, X(t))$ is also Lipschitz continuous.*

Note that, using [58, Theorem 20], it is possible to also formulate analogous conditions on the covariance structure for the case where these quantities are path-dependent. To prove the above theorem we will need the following transformation lemma, which is a straightforward extension of Lemma 5.1.4 to the functional setting. For X in the Skorohod space $\mathcal{R}^T = D([0,T], \mathbb{R}^d)$, we set $(\exp(X))^t = \exp(X^t) := (\exp(X(u \wedge t)))_{0 \leq u \leq T} \in \mathcal{R}^T_+$.

Lemma 5.3.5. *The functional $F(t, X^t)$ solves ($FTVP^+$) if and only if $\widetilde{F}(t, X^t) := F(t, \exp(X^t))$ solves*

$$(\widetilde{FTVP}) \qquad \begin{cases} \widetilde{F} \in \mathbb{C}^{1,2}([0,T]), \ \textit{satisfies the conditions from Definition 5.3.1}, \\ \mathcal{D}\widetilde{F} + \widetilde{\mathscr{A}}\widetilde{F} = 0 \qquad \textit{in } \bigcup_{t \in [0,T)} \mathcal{R}^t, \\ \widetilde{F}(T, X^T) = \widetilde{H}(X^T) \ \textit{for } X^T \in \mathcal{R}^T, \end{cases}$$

where $\widetilde{H}(X^T) = H(\exp(X^T))$ and

$$\widetilde{\mathscr{A}} := \frac{1}{2}\sum_{i,j=1}^d \widetilde{a}_{ij}(t, X(t))\nabla^2_{X,ij} + \sum_{i=1}^d \widetilde{b}_i(t, X(t))\partial_i, \qquad \textit{in} \bigcup_{t \in [0,T)} \mathcal{R}^t, \tag{5.3.6}$$

for $\widetilde{a}_{ij}(t, X(t)) := a_{ij}(t, \exp(X(t)))$ and $\widetilde{b}_i(t, X(t)) := -\frac{1}{2}a_{ii}(t, \exp(X(t)))$.

Proof. We write $\widetilde{F}(t, X^t)$ as $F(t, Y^t)$ with $Y^t = (\exp(X(u \wedge t)))_{0 \leq u \leq T}$. Thus, $F(T, Y^T) = H(Y^T)$, $Y^T \in \mathcal{R}^T_+$, if and only if $\widetilde{F}(T, X^T) = \widetilde{H}(X^T)$, $X^T \in \mathcal{R}^T$. By Lemma 3.3.3, $\widetilde{F} \in \mathbb{C}^{1,2}([0,T])$ and satisfies the conditions from Definition 5.3.1 on the path space of $\mathbb{R}^d$-valued càdlàg functions if and only if $F \in \mathbb{C}^{1,2}([0,T])$ and satisfies the conditions from Definition 5.3.1 on the path space of $\mathbb{R}^d_+$-valued càdlàg functions (because $\exp(X(t))$ is a sufficiently regular functional). Moreover, the chain rule for functional derivatives from Lemma 3.3.3 yields

$$\mathcal{D}\widetilde{F}(t, X^t) = \mathcal{D}F(t, Y^t) \quad \text{and} \quad \partial_i\widetilde{F}(t, X^t) = \partial_i F(t, Y^t)\exp(X_i(t)).$$

For the second-order vertical derivatives, we calculate

$$\nabla^2_{X,ij}\widetilde{F}(t,X^t) = \begin{cases} \nabla^2_{Y,ij}F(t,Y^t)\exp(X_i(t))\exp(X_j(t)), & i \neq j, \\ \nabla^2_{Y,ij}F(t,Y^t)\exp(X_i(t))\exp(X_i(t)) + \partial_i F(t,Y^t)\exp(X_i(t)), & i = j. \end{cases}$$

Hence, it follows that

$$\begin{aligned} \mathcal{D}\widetilde{F}(t,X^t) + \widetilde{\mathscr{A}}\widetilde{F}(t,X^t) &= \mathcal{D}F(t,Y^t) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,Y(t))Y_i(t)Y_j(t)\nabla^2_{Y,ij}F(t,Y^t) \\ &\quad + \frac{1}{2}\sum_{i=1}^{d} a_{ii}(t,Y(t))\partial_i\widetilde{F}(t,X^t) - \frac{1}{2}\sum_{i=1}^{d} a_{ii}(t,Y(t))\partial_i\widetilde{F}(t,X^t) \\ &= 0 \end{aligned}$$

if and only if $\mathcal{D}F(t,Y^t) + \mathscr{A}^+F(t,Y^t) = 0$, which leads to the conclusion. □

Proof of Theorem 5.3.4. Part (a) directly follows from [58, Theorem 20]. To prove part (b), note that the coefficients of $\widetilde{\mathscr{A}}$ satisfy the conditions of [58, Theorem 20], i.e., $\widetilde{a}(t,X(t))$ is positive definite and can be written as $\widetilde{\sigma}(t,X(t))\widetilde{\sigma}(t,X(t))^\top$ with a Lipschitz continuous volatility coefficient $\widetilde{\sigma}$, and $\widetilde{b}_i$ is Lipschitz. It therefore follows that $(\widetilde{\text{FTVP}})$ admits a unique solution $F \in \mathbb{C}^{1,2}([0,T])$ whenever $\widetilde{H} \in C^2_{l,lip}(\mathcal{R}^T)$. Applying Lemma 5.3.5 then gives the existence of solutions to (FTVP$^+$) if the terminal condition is of class $C^2_{l,lip}(\mathcal{R}^T_+)$. □

As in the previous section, the quantity $F(0,S^0)$ represents the amount of cash that is required at time $t=0$ so as to perfectly replicate the payoff H for all price trajectories in $\mathscr{S}_a^{(+)}$. In our situation, in order to interpret $F(0,S^0)$ as an arbitrage-free price, we have to know whether our class of trading strategies is indeed arbitrage-free with respect to all possible price trajectories in $\mathscr{S}_a^{(+)}$, which will be explored next. Theorem 5.3.7 is the functional counterpart of Theorem 5.2.4.

Definition 5.3.6. Suppose that the non-anticipative functional F satisfying the conditions from Definition 5.3.1 is the solution of the path-dependent heat equation

$$\mathcal{D}F + \mathscr{A}^{(+)}F = 0 \qquad \text{on} \bigcup_{t\in[0,T)} \mathcal{R}^t_{(+)} \cap C([0,T],\mathbb{R}^d_{(+)}).$$

For $S \in \mathscr{S}_a^{(+)}$, we then define ξ^S as in (5.3.2) and

$$V^S_\xi(t) := F(0,S^0) + \int_0^t \xi^S(s)\,\mathrm{d}S(s). \tag{5.3.7}$$

By $\mathscr{Y}^{(+)}$ we denote the collection of all pairs $(F(0,\cdot),\xi^{\cdot})$ that arise in this way.

The notion of an (admissible) arbitrage opportunity for $\mathscr{S}_a^{(+)}$ in the functional setting is defined in analogy to Definition 5.2.2, we only have to replace $\mathscr{X}^{(+)}$ by $\mathscr{Y}^{(+)}$.

Theorem 5.3.7. *Suppose that $a(t, X(t))$ is continuous, bounded, and positive definite for all $(t, X(t)) \in [0, \widetilde{T}] \times \mathbb{R}^d_{(+)}$, where $\widetilde{T} > T$. Then there are no admissible arbitrage opportunities in $\mathscr{Y}^{(+)}$.*

Proof. The proof is similar to the one of Theorem 5.2.4. We first consider the case of $\mathscr{Y}$. Let X and $\mathbb{P}_{t,x}$ ($0 \le t \le T$, $x \in \mathbb{R}^d$) be as in the proof of Theorem 5.2.4. For a path $Y \in C([0,T], \mathbb{R}^d)$, we define $\overline{\mathbb{P}}_{t,Y^t}$ as that probability measure on $C([0,T], \mathbb{R}^d)$ under which the coordinate process X satisfies $\overline{\mathbb{P}}_{t,Y^t}$-a.s. $X(s) = Y(s)$ for $0 \le s \le t$ and under which the law of $(X(u))_{t \le u \le T}$ is equal to $\mathbb{P}_{t,Y(t)}$. The support theorem [89, Theorem 3.1] then states that the law of $(X(u))_{0 \le u \le T}$ under $\overline{\mathbb{P}}_{t,Y^t}$ has full support on $C_{Y^t}([0,T], \mathbb{R}^d) := \{\omega \in C([0,T], \mathbb{R}^d) \,|\, \omega^t = Y^t\}$.

Now suppose by way of contradiction that there exists an admissible arbitrage opportunity arising from a non-anticipative functional F as in Definition 5.3.6. In a first step, we show that F is nonnegative on $[0,T] \times C([0,T], \mathbb{R}^d)$. As in the proof of Theorem 5.2.4, the support theorem implies that $\big\{(S(t))_{0 \le t \le T} \,|\, S \in \mathscr{S}_a,\ S(0) = x\big\}$ is dense in $C_x([0,T], \mathbb{R}^d)$. Condition (a) of Definition 5.2.2 and the left-continuity of F thus imply that $F(T, Y) \ge 0$ for all $Y \in C([0,T], \mathbb{R}^d)$. In the same way, we get from the admissibility of the arbitrage opportunity that $F(t, Y^t) \ge -c$ for all $t \in [0,T]$ and $Y \in C([0,T], \mathbb{R}^d)$. To show that actually $F(t, Y^t) \ge 0$, let $Q \subset \mathbb{R}^d$ be a bounded domain whose closure is contained in $\mathbb{R}^d$ and let $\tau := \inf\{s \,|\, X(s) \notin Q\}$ be the first exit time from Q. By the functional change of variables formula, in conjunction with the fact that F solves (FTVP) (on continuous paths), we obtain $\overline{\mathbb{P}}_{t,Y_t}$-a.s. for $t \in [0,T)$ that

$$F(T \wedge \tau, X^{T \wedge \tau}) = F(t, Y^t) + \int_t^{T \wedge \tau} \nabla_X F(s, X^s)\, \mathrm{d}X(s). \tag{5.3.8}$$

By Proposition 3.2.7, we have

$$\Big[\int_t^{\cdot \wedge \tau} \nabla_X F(s, X^s)\, \mathrm{d}X(s)\Big](T) = \int_t^{T \wedge \tau} \nabla_X F(s, X^s)^\top a(s, X(s)) \nabla_X F(s, X^s)\, \mathrm{d}s.$$

Since $\nabla_X F$ and the coefficients of $\mathscr{A}$ are bounded in the closure of Q, the stochastic integral on the right-hand side of (5.3.8) is a true martingale. Therefore,

$$F(t, Y^t) = \overline{\mathbb{E}}_{t,Y^t}[\, F(T \wedge \tau, X^{T \wedge \tau})\,]. \tag{5.3.9}$$

Now let us take an increasing sequence $Q_1 \subset Q_2 \subset \cdots$ of bounded domains exhausting $\mathbb{R}^d$ and whose closures are contained in $\mathbb{R}^d$. By τ_n we denote the exit time from Q_n. Then, an application of (5.3.9) for each τ_n, Fatou's lemma in conjunction with the fact that $F \ge -c$, and the already established nonnegativity of $F(T, \cdot)$ yield

$$F(t, Y^t) = \lim_{n \uparrow \infty} \overline{\mathbb{E}}_{t,Y^t}[\, F(T \wedge \tau_n, X^{T \wedge \tau_n})\,] \ge \overline{\mathbb{E}}_{t,Y^t}[\, F(T, X^T)\,] \ge 0. \tag{5.3.10}$$

This establishes the nonnegativity of F on $[0,T] \times C([0,T], \mathbb{R}^d)$.

Now let $S \in \mathscr{S}_a$ and T_0 be such that $V^S_\xi(0) \le 0$ and $V^S_\xi(T_0) > 0$. Since $V^S_\xi(t) = F(t, S^t)$ by Proposition 5.3.2, we have $F(0, S^0) = 0$ and $F(T_0, S^{T_0}) > 0$. By left-continuity of F, we actually have $F(T_0, \cdot) > 0$ in an open neighborhood $U \subset C_{S(0)}([0,T], \mathbb{R}^d)$ of the path S.

Since $\mathbb{P}_{0,S(0)}$-a.e. sample path belongs to $\mathscr{S}_a$, the functional change of variables formula gives that $\mathbb{P}_{0,S(0)}$-a.s.,

$$F(T_0, X^{T_0}) = F(0, S^0) + \int_0^{T_0} \xi^X(t)\, \mathrm{d}X(t).$$

Localization as in (5.3.10) and using the fact that $F \geq 0$ implies that

$$0 = F(0, S^0) \geq \mathbb{E}_{0,S(0)}\big[\, F(T_0, X^{T_0})\,\big] \geq 0.$$

Applying once again the support theorem now yields a contradiction to the fact that $F(T_0, \cdot) > 0$ in the open set U. This completes the proof for $\mathscr{Y}$.

The proof for $\mathscr{Y}^+$ is completed by an exponential transformation, as in the proof of Theorem 5.2.4. □

Appendix A

Matlab Simulations

Table A.1: Simulation of the portfolio value according to (4.1.11)*:*

```
function [ V ] = valueprocess_pi( p, pi )

L = size(p);
p(p==0) = 1;

V = zeros(1, L(2));
V(1) = 1;
xi = zeros(L(1), L(2)-1);
for t = 1:L(2)-1
    xi(:,t) = V(t) * pi(:,t) ./ p(:,t);
    V(t+1) = sum(xi(:,t) .* p(:,t+1));
end

end
```

Table A.2: Simulation of the market portfolio according to (4.1.14)*:*

```
function [ mu, summe ] = marketportfolio( p )

 L = size(p);

 STot = zeros(1,L(2));
 mu = zeros(L(1),L(2));
 for t = 1:L(2)
     STot(1,t) = sum(p(:,t));
     for i = 1:L(1)
         mu(i,t) = p(i,t) / STot(1,t);
```

```
     end
 end
 summe = sum(mu);
end
```

Table A.3: Simulation of the covariance structure of the individual stocks in the market according to (4.1.17)*:*

```
function [ kovariation ] = kovariation_matrix( p )

L = size(p);

kovariation = zeros(L(1),L(1),L(2)-1);
for t = 1:L(2)-1
    for i = 1:L(1)
        for j = 1:L(1)
            kovariation(i,j,t) = (log(p(i,t+1))-log(p(i,t)))*
            (log(p(j,t+1))-log(p(j,t)));
        end
    end
end

% kovariation(find(isnan(kovariation))) = 0;

end
```

Table A.4: Simulation of the portfolio excess growth according to (4.1.18)*:*

```
function [excessgrowth] = excessgrowth(pi, kovariation)

b = pi;
a = kovariation;
L = size(b);

for t = 1:L(2)-1
    for i = 1:L(1)
        teil1(i,t) = b(i,t) * a(i,i,t);
    end
end
teil1(find(isnan(teil1)))=0;
```

```
summe_teil1 = sum(teil1);

for t = 1:L(2)-1
   for i = 1:L(1)
       for j = 1:L(1)
           teil2(i,j,t) = b(i,t) * b(j,t) * a(i,j,t);
       end
   end
end
teil2(find(isnan(teil2)))=0;

for t = 1:L(2)-1
   for i = 1:L(1)
       summe1_teil2(i,t) = sum(teil2(i,:,t));
   end
end

for t = 1:L(2)-1
   summe_teil2(1,t) = sum(summe1_teil2(:,t));
end


excessgrowth = zeros(1,L(2)-1);
for t = 1:L(2)-1
    excessgrowth(1,t) = 0.5 * (summe_teil1(1,t) - summe_teil2(1,t));
end

end
```

Table A.5: Simulation of the relative covariances of the individual stocks in the market according to (4.1.19)*:*

```
function tau_pi=rel_cov(pi, a)

s=size(pi);
e=eye(s(1));

%preallocating:
tau_pi=zeros(s(1), s(1), s(2)-1);
```

```
for t=1:s(2)-1
     for i=1:s(1)
         for j=1:s(1)

            tau_pi(i,j,t)=(pi(:,t)-e(:,i))'*a(:,:,t)*(pi(:,t)-e(:,j));

         end
     end
end

%tau_pi(find(isinf(tau_pi))) = 0;

end
```

Table A.6: Simulation of the covariances of the individual stocks relative to the market according to (4.1.26)*:*

```
function [ tau ] = tau_mu( p )

L = size(p);
[mu, summe_mu] = marketportfolio(p);


tau = zeros(L(1),L(1),L(2)-1);
for t = 1:L(2)-1
    for i = 1:L(1)
    for j = 1:L(1)
    tau(i,j,t) = (log(mu(i,t+1))-log(mu(i,t)))*
    (log(mu(j,t+1))-log(mu(j,t)));
   %tau(i,j,t) = (1/(mu(i,t)*mu(j,t)))*
   % ((mu(i,t+1))-(mu(i,t)))*((mu(j,t+1))-(mu(j,t)));
    end
    end
end

end
```

Table A.7: Simulation of the portfolio weights of an entropy weighted portfolio according to (4.3.5)*:*

```
function [ pi , summe_pi ] = portfolio_H_alpha_minus (p, delta,alpha)
```

```
L = size(p);
pi = zeros(L(1),L(2));
helfer1 = zeros(L(1),L(2));
helfer2 = zeros(L(1),L(2));
[mu, summe_mu] = marketportfolio(p);


for t = 1:L(2)
    if delta >= t
       for j = 1:L(1)
helfer1(j,t) = -(alpha*mu(j,t)+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ...
((t-delta)/delta)* mu(j,1)))*log((alpha*mu(j,t)
+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1))));
        end
    else
        for j = 1:L(1)
helfer1(j,t) = -(alpha*mu(j,t) + (1-alpha)*(1/delta) * sum(mu(j, t-delta:t),2))...
*log(alpha*mu(j,t) + (1-alpha)*(1/delta) * sum(mu(j, t-delta:t),2));
        end
    end
end


for t = 1:L(2)
    if delta >= t
        for j = 1:L(1)
helfer2(j,t) = mu(j,t)*log((mu(j,t)+((1/delta)*sum(mu(j, 1:t),2) - ...
((t-delta)/delta)* mu(j,1))));
        end
    else
        for j = 1:L(1)
helfer2(j,t) = mu(j,t)*log(mu(j,t) + (1/delta)*sum(mu(j, t-delta:t),2)) ;
        end
    end
end

for i = 1:L(1)
    for t = 1:L(2)
   pi(i,t) = -alpha*helfer2(i,t) / sum(helfer1(:,t)) + mu(i,t) +
   alpha*mu(i,t)*sum(helfer2(:,t))/sum(helfer1(:,t));
```

```
    end
end

summe_pi = sum(pi);

end
```

Table A.8: Simulation of the left-hand side of the relative performance of entropy weighting as given in equation (4.3.9)*:*

```
function [ LHS ] = LHS_H_alpha_minus( p, delta,alpha )

[mu, summe_mu] = marketportfolio(p);
[pi, summe_pi] = portfolio_H_alpha(p,delta, alpha);

[V_mu] = valueprocess_pi(p,mu);
[V_pi] = valueprocess_pi(p,pi);

LHS = log ( V_pi ./ V_mu );

%plot(linkeSeite)

end
```

Table A.9: Simulation of the right-hand side of the relative performance of entropy weighting as given in equation (4.3.9)*:*

```
function [ RHS, a,b,c ] = RHS_H_alpha_minus( p , delta, alpha )

L = size(p);
[mu, summe_mu] = marketportfolio(p);
%a = kovariation_matrix(p);
% tau = tau_pi(mu,a);
tau = tau_mu(p);

helfer1 = -mu(1:end,1).*log(mu(1:end,1));
helfer2 = zeros(L(1),L(2));
helfer3 = zeros(L(1),L(2));
helfer4 = zeros(L(1),L(2)-1);
```

```
% log_term : Zähler

for t = 1:L(2)
        if delta >= t
            for j = 1:L(1)
   helfer2(j,t) = -(alpha*mu(j,t)+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) -...
    ((t-delta)/delta)* mu(j,1)))*log(alpha*mu(j,t)+(1-alpha)*((1/delta)*...
    sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)));
            end
        else
            for j = 1:L(1)
    helfer2(j,t) = -(alpha*mu(j,t) + ((1-alpha)/delta)*sum(mu(j, t-delta:t),2))...
    *log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2));
            end
        end
end

%log_term : Gesamt

log_term = log ( sum(helfer2) / sum(helfer1) );

% z_term : Zähler

for t = 1:L(2)
    if delta >= t
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,1))*(log(alpha*mu(j,t)...
+ (1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)))+1);
        end
    else
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,t-delta))...
 *(log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))+1);
        end
    end
end

% z_term : Gesamt

z_term = cumsum( sum(helfer3) ./ sum(helfer2) );
```

```
%tau_term

for t = 1:L(2)-1
    if delta >= t
        for j = 1:L(1)
helfer4(j,t) = (1/ (2*sum(helfer2(:,t))))*(alpha^2/(alpha*mu(j,t) +...
 (1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)))))*...
  mu(j,t)^2 * tau(j,j,t);
        end
    else
        for j = 1:L(1)
helfer4(j,t) = (1/(2* sum(helfer2(:,t))))*(alpha^2/(alpha*mu(j,t) +...
 (1-alpha)*((1/delta)* sum(mu(j, t-delta:t),2)))) *mu(j,t)^2 * tau(j,j,t);
        end
    end
end

tau_term = sum(cumsum(helfer4,2));

RHS = log_term(1,1:end-1) + z_term(1,1:end-1) + tau_term;

a = log_term(1,1:end-1);
b = tau_term;
c = z_term(1,1:end-1);
end
```

Table A.10: Alternative simulation of the right-hand side of the relative performance of entropy weighting as given in equation (4.3.9)*:*

```
function [ RHS, a,b,c ] = RHS_H_alpha( p , delta, alpha )

L = size(p);
[mu, summe_mu] = marketportfolio(p);
%a = kovariation_matrix(p);
% tau = tau_pi(mu,a);
tau = tau_mu(p);

helfer1 = mu(1:end,1).*log(mu(1:end,1));
helfer2 = zeros(L(1),L(2));
helfer3 = zeros(L(1),L(2));
helfer4 = zeros(L(1),L(2)-1);
```

```
% log_term : Zähler

for t = 1:L(2)
        if delta >= t
            for j = 1:L(1)
 helfer2(j,t) = (alpha*mu(j,t)+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) -...
  ((t-delta)/delta)* mu(j,1)))*log(alpha*mu(j,t)+(1-alpha)*((1/delta)*...
  sum(mu(j, 1:t),2)- ((t-delta)/delta)* mu(j,1)));
            end
        else
            for j = 1:L(1)
 helfer2(j,t) = (alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))...
 *log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2));
            end
        end
end

%log_term : Gesamt

log_term = log ( sum(helfer2) / sum(helfer1) );

% z_term : Zähler

for t = 1:L(2)
    if delta >= t
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,1))*(log(alpha*mu(j,t)...
+ (1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)))+1);
        end
    else
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,t-delta))...
*(log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))+1);
        end
    end
end

% z_term : Gesamt

z_term = cumsum( sum(helfer3) ./ sum(helfer2) );
```

```
%tau_term

for t = 1:L(2)-1
    if delta >= t
        for j = 1:L(1)
helfer4(j,t) = (1/ (2*sum(helfer2(:,t)))) *(alpha^2/(alpha*mu(j,t) +...
(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) -...
((t-delta)/delta)* mu(j,1)))) * mu(j,t)^2 * tau(j,j,t);
        end
    else
        for j = 1:L(1)
helfer4(j,t) = (1/(2* sum(helfer2(:,t)))) *(alpha^2/(alpha*mu(j,t) +...
(1-alpha)*((1/delta)* sum(mu(j, t-delta:t),2)))) *mu(j,t)^2 * tau(j,j,t);
        end
    end
end

tau_term = sum(cumsum(helfer4,2));

RHS = log_term(1,1:end-1) - z_term(1,1:end-1) - tau_term;

a = log_term(1,1:end-1);
b = - tau_term;
c = - z_term(1,1:end-1);
end
```

Table A.11: Plotting the graphs in Figure 4.3.5 and Figure 4.3.6:

```
function drawHalpha_minus_datum( p,delta,alpha,tt)

l = LHS_H_alpha_minus(p,delta,alpha);
[r,log_term,tau,z] = RHS_H_alpha_minus(p,delta,alpha);

t=tt(1:end-94);
l=l(1:end-1);

figure
%subplot(1,2,1);
hold
plot(t,l,'Color',[1,0.647,0])
```

```
plot(t,r,'k')

xlim([tt(1),tt(end-50)]);
datetick('x','yyyy','keeplimits','keepticks')

%title('LHS vs. RHS of the Master Formula')
xlabel('Year') % x-axis label
ylabel('Process values') % y-axis label
legend('LHS','RHS','Location','best')
set(gca, 'YGrid', 'off', 'XGrid', 'off','box','on');
gridxy(get(gca,'XTick'),get(gca,'YTick'),'Color',[0.8,0.8,0.8],'Linestyle',':');

figure
%subplot(1,2,2);
hold
plot(t,log_term)
plot(t,tau,'r')
plot(t,z,'g')

xlim([tt(1),tt(end-50)]);
datetick('x','yyyy','keeplimits','keepticks')

%title('Componentwise representation of the RHS')
xlabel('Year') % x-axis label
ylabel('Process values') % y-axis label
legend('log-term','g([0,t])','h([0,t])','Location','best')
set(gca, 'YGrid', 'off', 'XGrid', 'off','box','on');
gridxy(get(gca,'XTick'),get(gca,'YTick'),'Color',[0.8,0.8,0.8],'Linestyle',':');
end
```

Table A.12: Computing the cumulative generalized excess growth rate of the market according to (4.3.12)*:*

```
function [ eg ] = excessgrowth_H_alpha_minus( p , delta, alpha )

L = size(p);

[mu, summe_mu] = marketportfolio(p);
[pi, summe_pi] = portfolio_H_alpha_minus(p,delta, alpha);

[V_mu] = valueprocess_pi(p,mu);
```

```
[V_pi] = valueprocess_pi(p,pi);

helfer1 = -mu(1:end,1).*log(mu(1:end,1));
helfer2 = zeros(L(1),L(2));
helfer3 = zeros(L(1),L(2));
diff = zeros(1,L(2)-1);

% log_term : einzelne Summanden von G(tilde{mu}(t)) für t = 1...L(2)

for t = 1:L(2)
        if delta >= t
            for j = 1:L(1)
helfer2(j,t) = -(alpha*mu(j,t)+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) -...
((t-delta)/delta)* mu(j,1)))*log(alpha*mu(j,t)+(1-alpha)*((1/delta)*...
sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)));
            end
        else
            for j = 1:L(1)
helfer2(j,t) = -(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))...
*log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2));
            end
        end
end



for t = 1:L(2)-1
diff(1,t) = log((V_pi(t+1)*sum(helfer1))/(V_mu(t+1)*sum(helfer2(:,t+1))))...
-log((V_pi(t)*sum(helfer1))/(V_mu(t)*sum(helfer2(:,t))));
end


helfer2_1 = helfer2(:,1:end-1); %podpravitj dim
log_term = 1/alpha^2 * cumsum( sum(helfer2_1).*diff );




for t = 1:L(2)
```

```
    if delta >= t
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,1))*(log(alpha*mu(j,t)...
+ (1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)))+1);
        end
    else
        for j = 1:L(1)
helfer3(j,t) = ((1-alpha)/delta)*(mu(j,t) - mu(j,t-delta))...
*(log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))+1);
        end
    end
end



z_term = cumsum( sum(helfer3) / alpha^2 );


eg = log_term - z_term(1,1:end-1);


end
```

Table A.13: Comparing the computation (4.3.12) *with the definition of the generalized excess growth rate of the market in* (4.3.7)*:*

```
function [test] = excessgrowth_compare( p , delta, alpha )



[ eg ] = excessgrowth_H_alpha_minus( p , delta, alpha );

L = size(p);
[mu, summe_mu] = marketportfolio(p);
tau = tau_mu(p);

helfer1 = zeros(L(1),L(2));
helfer2 = zeros(L(1),L(2)-1);


for t = 1:L(2)
        if delta >= t
```

```
            for j = 1:L(1)
helfer1(j,t) = -(alpha*mu(j,t)+(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) -...
 ((t-delta)/delta)* mu(j,1)))*log(alpha*mu(j,t)+(1-alpha)*((1/delta)*...
 sum(mu(j, 1:t),2)- ((t-delta)/delta)* mu(j,1)));
            end
        else
            for j = 1:L(1)
helfer1(j,t) = -(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2))...
*log(alpha*mu(j,t) + ((1-alpha)/delta) * sum(mu(j, t-delta:t),2));
            end
        end
end

%tau_term

for t = 1:L(2)-1
    if delta >= t
        for j = 1:L(1)
helfer2(j,t) = (1/ (2*sum(helfer1(:,t)))) *(alpha^2/(alpha*mu(j,t) +...
(1-alpha)*((1/delta)*sum(mu(j, 1:t),2) - ((t-delta)/delta)* mu(j,1)))) *...
 mu(j,t)^2 * tau(j,j,t);
        end
    else
        for j = 1:L(1)
helfer2(j,t) = (1/(2* sum(helfer1(:,t)))) *(alpha^2/(alpha*mu(j,t) +...
 (1-alpha)*((1/delta)* sum(mu(j, t-delta:t),2)))) *mu(j,t)^2 * tau(j,j,t);
        end
    end
end

tau_term = sum(cumsum(helfer2,2));

G(1,:) = sum(helfer1);
test = tau_term .* G(1,1:end-1)/(alpha^2);

plot(eg)
hold
plot(test,'r')

end
```

Bibliography

[1] Acciaio B., Beiglböck, M., Penkner, F. and Schachermayer, W. (2013): A model-free version of the Fundamental Theorem of Asset Pricing and the super-replication theorem. *Mathematical Finance, 26(2):233–251.*

[2] Alvarez, A., Ferrando, S. and Olivares, P. (2013): Arbitrage and hedging in a non probabilistic framework. *Math. Financ. Econ., 7(1):1–28.*

[3] Ananova, A. and Cont, R. (2016): Pathwise integration with respect to paths of finite quadratic variation. *arXiv 1603.03305v2.*

[4] Banner, A. and Fernholz, D. (2007): Short-term arbitrage in volatility-stabilized markets. *Annals of Finance, Volume 4, Issue 4, pp 445-454.*

[5] Banner, A., Fernholz, R. and Karatzas, I. (2005): On Atlas models of equity markets. *Annals of Applied Probability, 15, 2296-2330.*

[6] Bayraktar, E. and Zhang, Y. (2013): Fundamental Theorem of Asset Pricing under transaction costs and model uncertainty. *arXiv 1309.1420v2.*

[7] Bayraktar, E., Zhang, Y. and Zhou, Z. (2014): A note on the Fundamental Theorem of Asset Pricing under model uncertainty. *Risks 2(4), 425-433.*

[8] Beiglböck, M., Cox, A. M. G., Huesmann, M., Perkowski, N. and Prömel, D. (2015): Pathwise super-replication via Vovk's outer measure. *arXiv: 1504.03644v1.*

[9] Bender, C., Sottinen, T. and Valkeila, E. (2008): Pricing by hedging and no-arbitrage beyond semimartingales. *Finance Stoch., 12(4):441-468.*

[10] Biagini, S., Bouchard, B., Kardaras, C. and Nutz, M. (2015): Robust fundamental theorem for continuous processes. *Mathematical Finance, preprint doi: 10.1111/mafi.12110.*

[11] Bick, A. and Willinger, W. (1994): Dynamic spanning without probabilities. *Stochastic Process. Appl., 50(2):349-374.*

[12] Bismut, J. M. (1973): Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications, 44, 384-404.*

[13] Bouchard, B. and Nutz, M. (2014): Consistent price systems under model uncertainty. *arXiv:1408.5510v1.*

[14] Bouchard, B. and Nutz, M. (2015): Arbitrage and duality in non-dominated discrete-time models. *Ann. Appl. Probab. 25(2), 823-859.*

[15] Brown, H., Hobson, D. and Rogers, L. C. G. (2001): Robust hedging of barrier options. *Math. Finance, 11(3):285-314.*

[16] Bühler, H. (2015): Pricing with a discrete smile. *SSRN preprint 2642630.*

[17] Burzoni, M., Frittelli. M. and Maggis, M. (2014): Universal arbitrage aggregator in discrete time markets under uncertainty. *arXiv:1407.0948, to appear in Finance Stoch.*

[18] Cont, R. (2006): Model uncertainty and its impact on the pricing of derivative instruments. *Math.Finance, 16(3):519-547.*

[19] Cont, R. and Fournié, D. A. (2010): A functional extension of the Itô formula. *C. R. Math. Acad. Sch. Paris, 348(1-2):5-61.*

[20] Cont, R. and Fournié, D. A. (2010): Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis 259, 1043-1072.*

[21] Cont, R. and Fournié, D. A. (2013): Functional Itô Calculus and stochstic integral representation of martingales. *The Annals of Probability Vol. 41, No. 1, 109-133.*

[22] Cosso, A. and Russo, F. (2014): A regularization approach to functional Itô calculus and strong-viscosity solutions for path-dependent PDEs. *arXiv:1401.5034.*

[23] Cox, A. M. G. and Oblój, J. (2011): Robust hedging of double touch barrier options. *SIAM J.Financial Math., 2:141-182.*

[24] Davis, M., and Hobson, D. (2007): The Range of Traded Option Prices. *Math. Finance 17(1), 1-14.*

[25] Davis, M., Oblój, J. and Raval, V. (2014). Arbitrage bounds for prices of weighted variance swaps. *Math. Finance, 24(4):821 – 854.*

[26] Di Girolami, C. and Russo, F. (2010): Infinite dimensional stochastic calculus via regularization and applications. *http://hal.archives-ouvertes.fr/inria-00473947/fr.*

[27] Di Girolami, C. and Russo, F. (2011): Clark-Okone type formula for non-semimartingales with finite quadratic variation. *C. R. Math. Acad. Sci. Paris, 349(3-4):209-214.*

[28] Di Girolami, C. and Russo, F. (2012): Generalized covariation and extended Fukushima decomposition for Banach space valued processes, applications to windows of Dirichlet processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top., 15(2):1250007, 50.*

[29] Di Girolami, C. and Fabbri, G., Russo, F. (2013): The covariation of Banach space valued processes and applications. *Metrika, preprint: http://hal.inria.fr/hal-00780430.*

[30] Di Girolami, C. and Russo, F. (2014): Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka Journal of Mathematics, 51(3).*

[31] Dolinsky, Y. and Soner, H. M. (2014): Robust hedging with proportional transaction costs. *Finance Stoch. 18(2), 327-347.*

[32] Dudley. R. M. (1977): Wiener functionals as Itô integrals. *Ann. Probability, 5(1):140–141.*

[33] Dupire, B. (1997): Pricing and hedging with smiles. *Mathematics of derivative securities (Cambridge, 1995), 103-111, Publ. Newton Inst., 15, Cambridge Univ. Press, Cambridge.*

[34] Dupire, B. (2009): Functional Itô calculus. *Portfolio Research Paper 2009-04, Bloomberg.*

[35] El Karoui, N. Peng, S. and Quenez, M. C. (1997): Backward stochastic differential equation in finance. *Mathematical Finance, 7(1), 1–71.*

[36] El Karoui, N., Jeanblanc-Picqué, M. and Shreve, S. E. (1998): Robustness of the Black and Scholes formula. *Math. Finance, 8(2):93–126.*

[37] Fernholz, E. R. (1999): On the diversity of equity markets. *Journal of Mathematical Economics 31, 393-417.*

[38] Fernholz, E. R. (1999.a): Portfolio generating functions. In M. Avellaneda (ed.), *Quantitative Analysis in Financial Markets, River Edge, NJ. World Scientific.*

[39] Fernholz, E. R. (2001): Equity portfolios generated by functions of ranked market weights. *Finance & Stochastics 5, 469-486.*

[40] Fernholz, E. R. (2002): Stochastic portfolio theory. *Springer-Verlag, New York.*

[41] Fernholz, E. R. and Karatzas, I. (2005): Relative arbitrage in volatility-stabilized markets. *Annals of Finance 1, 149-177.*

[42] Fernholz, E. R., Karatzas, I. and Kardaras, C. (2005): Diversity and arbitrage in equity markets. *Finance & Stochastics 9, 1-27.*

[43] Fernholz, E. R. and Karatzas, I. (2006): The implied liquidity premium for equities. *Annals of Finance 2, 87-99.*

[44] Fernholz, E. R., Karatzas, I. and Kardaras, C. (2005): Diversity and arbitrage in equity markets. *Finance & Stochastics 9, 1-27.*

[45] Fernholz, E. R. and Karatzas, I. (2009): Stochastic portfolio theory: an Overview. *Handbook of numerical analysis 15, pages 89-167.*

[46] Föllmer, H. (1979/80): Calcul d' Itô sans probabilités. *Seminar on Probability XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French), volume 850 of Lecture Notes in Math., pages 143-150. Springer, Berlin, 1981.*

[47] Föllmer, H. (2001): Probabilistic aspects of financial risk. *In European Congress of Mathematics, Vol. I (Barcelona, 2000), volume 201 of Progr. Math., pages 21-36. Birkhäuser, Basel.*

[48] Föllmer, H. and Schied, A. (2011) Stochastic finance. An introduction in discrete time. *Walter de Gruyter & Co., Berlin, 3rd revised and extended edition.*

[49] Föllmer, H. and Schied, A. (2013): Probabilistic aspects of finance. *Bernoulli 19(4):1306-1326.*

[50] Freedman, D. (1983): Brownian Motion and diffusion *Springer Verlag, New York, 2nd edition.*

[51] Gilboa, I. and Schmeidler, D. (1989): Maxmin expected utility with nonunique prior. *J. Math. Econom., 18(2):141-153.*

[52] Gyöngy, I. (1986): Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields 71 (1986), no. 4, 501-516.*

[53] Hansen, L. P. and Sargent, T. J. (2011): Robustness. *Princeton university press.*

[54] Hobson, D. (1998): Robust hedging of the look-back option. *Finance Stoch., 2(4):329–347.*

[55] Hobson, D. (2011): The Skorokhod embedding problem and model-independent bounds for option prices. *Paris-Princeton Lectures on Mathematical Finance 2010, volume 2003 of Lecture Notes in Math., pages 267–318. Springer, Berlin.*

[56] Janson, S. and Tysk, J. (2004): Preservation of convexity of solutions to parabolic equations. *J. Differential Equations, 206(1):182–226.*

[57] Jeanblanc, M. J. Yor, M. and Chesney, M. (2009): Mathematical methods for financial markets. *Springer Finance. Springer-Verlag London, Ltd., London.*

[58] Ji, S. and Yang, S. (2013): Classical solutions of path-dependent PDEs and functional Forward-Backward Stochastic Systems. *Mathematical Problems in Engineering, Hindawi, http://dx.doi.org/10.1155/2013/423101.*

[59] Karatzas, I. and Ruf, J. (2016): Trading strategies generated by Lyapunov functions. *http://www.oxford-man.ox.ac.uk/∼jruf/papers/Lyapunov.pdf.*

[60] Kardaras, C. (2010): Finitely additive probabilities and the Fundamental Theorem of Asset Pricing. *Contemporary Quantitative Finance, Berlin: Springer, pp 19-34.*

[61] Knight, F. (1921): Risk, uncertainty, and profit. *Houghton Mifflin, Boston.*

[62] Lipster, R. S. and Shiryaev, A. N. (1978): Statistics of Random Processes I. *Springer, Berlin.*

[63] Lyons, T. J. (1995): Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance, 2(2):117–133.*

[64] Lyons, T. J. (1998): Differential equations driven by rough signals. *Rev. Mat. Iberoam. 14, no. 2, 215-310.*

[65] Ma, J. and Yong, J. (1999): Forward-Backward Stochastic Differential Equations and their applications. *vol. 1702 of Lecture Notes in Mathematics, Springer, Berlin.*

[66] Maccheroni, F., Marinacci, M. and Rustichini, A. (2006): Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica, 74(6):1447-1498.*

[67] Mishura,Y. and Schied, A. (2016): Constructing functions with prescribed pathwise quadratic variation. *Journal of Mathematical Analysis and Applications (forthcoming).*

[68] Pal, S. and Wong, T. L. (2016): Energy, entropy, and arbitrage. *arXiv:1308.5376v2.*

[69] Pal, S. and Wong, T. L.: The geometry of relative arbitrage. *Mathematics and Financial Economics (forthcoming).*

[70] Pardoux, E. and Peng, S. (1990): Adapted Solutions of Backward Stochastic Equations. *Systems and Control Letters, 14, 55-61.*

[71] Pardoux, E. and Peng, S. (1992): Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Rozuvskii, B. L. and Sowers, R. B. (eds.) Stochastic partial differential equations and their applications. (Lecture notes Control Inf. Sci., Vol. 176, 200-217) Berlin Heidelberg New York: Springer.*

[72] Pardoux, E. and Tang, S. (1999): Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields, vol. 114, no. 2, pp. 123–150.*

[73] Peng, S. (1991): Probabilistic interpretations for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastics Reports, vol. 37, no. 1-2, pp. 61–74.*

[74] Peng, S. and Wang, F. (2011): BSDE, Path-dependent PDE and nonlinear Feynman-Kac formula. *arXiv:1108.4317.*

[75] Revuz, D. and Yor, M. (1990): Continuous martingales and Brownian Motion. *Springer Verlag.*

[76] Rogers, L. C. G. and Williams, D. (1987): Diffusions, markov processes and martingales. *2 Itô calculus, Wiley.*

[77] Riedel, F. (2011): Finance without probabilistic prior assumptions. *arXiv:1107.1078v1.*

[78] Russo, F. and Vallois, P. (2007): Elements of stochastic calculus via regularization. *Séminaire de Probabilités XL: Lecture Notes in Math., Vol. 1899, pp. 147-185, Springer Verlag.*

[79] Schied, A. and Stadje, M. (2007): Robustness of Delta hedging for path-dependent options in local volatility models. *J. Appl. Probab., 44(4):865-879.*

[80] Schied, A. (2013): Model-free CPPI. *Journal of Economic Dynamics and Control 40, 84-94, 2014.*

[81] Schied, A. (2015): On a class of generalized Takagi functions with a linear pathwise quadratic variation. *Journal of Mathematical Analysis and Applications 433, 947-990.*

[82] Schied, A. and Voloshchenko, I. (2016): The associativity rule in pathwise functional Itô calculus. University of Mannheim, *ArXiv: 1605.0886.*

[83] Schied, A. and Voloshchenko, I. (2016): Pathwise no-arbitrage in a class of Delta hedging strategies, *to appear in Probability, Uncertainty and Quantitative Risk.*

[84] A. Schied, L. Speiser, and I. Voloshchenko (2016): Model-free portfolio theory and its functional master formula. *arXiv: 1606.03325.*

[85] Shafer, G. and Vovk, V. (2001): Probability and finance: It's only a game! *Wiley, New York.*

[86] Sondermann, D. (2006): Introduction to stochastic calculus for finance. A new didactic approach, volume 579 of *Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, Berlin.*

[87] Strong, W. (2014): Generalizations of functionally generated portfolios with applications to statistical arbitrage. *SIAM J. Financial Math. Vol. 5, pp. 472-492.*

[88] Stroock, D. W. and Varadhan, S. R. S (1969): Diffusion processes with continuous coefficients. *Comm. Pure Appl. Math., 22:345–400, 479–530.*

[89] Strook, D. W. and Varadhan, S. R. S (1972): On the support of diffusion processes with applications to the strong maximum principle. *Proc. 6th Berkeley Symp. Math. Stat. Prob. III, pp. 333-359.*

[90] Vervuurt, A. and Karatzas, I. (2015): Diversity-weighted portfolios with negative parameter. *Annals of Finance, 11(3-4):411-432.*

[91] Vovk, V. (1993): A logic of probability, with application to the foundations of statistics (with discussion). *Journal of the Royal Statistical Society B, 55:317-351.*

[92] Vovk, V. (2011): Rough paths in idealized financial markets. *arXiv:1005.0279, 2011. Journal version: Lithuanian Mathematical Journal 51:274-285.*

[93] Vovk, V. (2012): Continuous-time trading and the emergence of probability. *Finance and Stochastics, 16:561-609.*

[94] Vovk, V. (2014): Itô calculus without probability in idealized financial markets. *arXiv:1108.0799v2.*

[95] Widder, D., V. (1941): The Laplace Transform. *Princeton Mathematical Series, v. 6. Princeton University Press, Princeton, N. J.*

[96] Widder, D. V. (1944): Positive temperatures on an infinite rod. *Trans. Amer. Math. Soc. 55, 85-95.*

[97] Widder, D. V. (1953): Positive temperatures on a semi-infinite rod. *Trans. Amer. Math. Soc. 75, 510-525.*

[98] Widder, D. V. (1975): The heat equation. *Academic Press, Pure and Applied Mathematics, Vol. 67.*

[99] Wong, T. L. (2015): Optimization of relative arbitrage. *Annals of Finance 11, no. 3, 345-382.*

[100] Wong, T. L. (2015): Universal portfolios in stochastic portfolio theory. *ArXiv e-prints.*