Non-anticipative functional calculus and applications to stochastic processes

Yi Lu

Université Pierre et Marie Curie

École Doctorale de Sciences Mathématiques
de Paris Centre

Laboratoire de Probabilités et Modèles Aléatoires

Calcul fonctionnel non-anticipatif et applications aux processus stochastiques

Non-anticipative functional calculus and applications to stochastic processes

Par Yi Lu

Thèse de doctorat de Mathématiques

Directeur de thèse: Rama Cont

Présentée et soutenue le 6 décembre 2017 devant un jury composé de:

M. Rama CONT	CNRS	Directeur
Mme. Nathalie EISENBAUM	CNRS	Examinatrice
M. Romuald ELIE	Université Paris-Est	Examinateur
M. Peter FRIZ	Technische Universitat Berlin	Rapporteur
M. Antoine LEJAY	INRIA Nancy Grand-Est	Rapporteur
M. David PROEMEL	University of Oxford	Examinateur

Abstract

This thesis focuses on various mathematical questions arising in the *non-anticipative functional calculus*, a differential calculus for functionals of (right-)continuous paths with finite quadratic variation and the application of this calculus to functionals of stochastic processes. This functional calculus was initially developed by Dupire [20] and Cont & Fournié [8, 9] using the notion of vertical derivative of a functional, a concept based on pathwise directional derivatives. In this thesis we extend the scope and results of this functional calculus to functionals which may not admit such directional derivatives, either through approximations (Part I) or by defining a notion of weak vertical derivative (Part II).

In the first part, we consider the representation of conditional expectations as non-anticipative functionals. Such functional representations may fail, in general, to admit directional derivatives. We show nevertheless that it is possible under very general conditions to approximate such functionals by a sequence of smooth functionals in an appropriate sense. Combined with the functional Itô calculus, this approach provides a systematic method for computing explicit approximations to martingale representations for a large class of Brownian functionals. We also derive explicit convergence rates of the approximations under some assumptions on the functionals. These results are then applied to the problem of sensitivity analysis and dynamic hedging of (path-dependent) contingent claims. Numerical illustrations are provided, which show that this approach is competitive with respect to other methods for computing sensitivities.

In the second part, we propose a concept of *weak vertical derivative* for non-anticipative functionals which may fail to possess directional derivatives. The definition of the weak vertical derivative is based on the notion of pathwise quadratic variation and makes use of the duality associated to the associated bilinear form. The weak vertical derivative operator with respect to a path of finite quadratic variation is shown to be the 'inverse' of the pathwise (Föllmer) integral with respect to this path. Our approach involves only pathwise arguments and does not rely on any probabilistic notions. When applied to functionals of a semimartingales, this notion of weak derivative coincides with the probabilistic weak derivative constructed by Cont and Fournié [10] in a martingale framework. Finally, we show that the notion of weak vertical derivative leads to a functional characterization of local martingales with respect to a reference process, and allows to define a concept of *pathwise* weak solution for path-dependent partial differential equations.

Résumé

Cette thèse est consacrée à l'étude du calcul fonctionnel non-anticipatif, un calcul différentiel pour des fonctionnelles sur l'espace des trajectoires à variation quadratique finie. Ce calcul fonctionnel est basé sur la notion de *dérivée verticale* d'une fonctionelle, qui est une dérivée directionnelle particulière. Dans cette thèse nous étendons le cadre classique du calcul fonctionnel non-anticipatif à des fonctionnelles ne possédant pas de dérivée directionnelle au sens classique (trajectoriel). Dans la première partie de la thèse nous montrons comment une classe importante de fonctionelles, définie par une espérance conditionnelle, peuvent être approchées de façon systématique par des fonctionnelles régulières. Dans la deuxième partie, nous introduisons une notion de dérivée verticale *faible* qui couvre une plus grande classe de fonctionnelles, et notamment toutes les martingales locales.

Dans la première partie, nous nous sommes intéressés à la représentation d'une espérance conditionnelle par une fonctionnelle non-anticipative. D'une manière générale, des fonctionnelles ainsi construites ne sont pas régulières. L'idée est donc d'approximer ces fonctionnelles par une suite des fonctionnelles régulières dans un certain sens. A l'aide du calcul d'Itô fonctionnel, cette approche fournit une façon systématique d'obtenir une approximation explicite de la représentation des martingales pour une grande famille de fonctionnelles Browniennes. Nous obtenons également un ordre de convergence explicite sous des hypothèses plus fortes. Quelques applications au problème de la couverture dynamique sont données à la fin de cette partie.

Dans la deuxième partie, nous étendons la notion de dérivée verticale d'une fonctionnelle non-anticipative, et nous proposons une notion de dérivée verticale faible pour des fonctionnelles qui n'admettent pas nécessairement de dérivée directionnelle. L'approche proposée est entièrement trajectorielle, et ne repose sur aucune notion probabiliste. Cependant, nous montrons que si l'on applique cette notion à un processus stochastique, elle coïncide avec la notion de dérivée faible proposée dans un cadre probabiliste par Cont et Fournié [10]. Cette notion nous permet également d'obtenir une caractérisation fonctionnelle d'une martingale locale par rapport à un processus de référence fixé, ce qui donne lieu à une notion de solution faible pour des équations aux dérivées partielles dépendant de la trajectoire.

Remerciements

Je tiens avant tout à remercier mon directeur de thèse, Rama Cont, d'avoir accepté d'encadrer mon doctorat. Il m'a prodigué de nombreux conseils, tout en m'accordant une grande indépendance au cours de mes recherches. Ses larges connaissances scientifiques et ses encouragements ont aussi été décisifs pour le bon découlement de ma thèse.

Je souhaite ensuite exprimer ma plus profonde gratitude à Peter Friz et Antoine Lejay, qui ont accepté d'être les rapporteurs de cette thèse. Je suis très honoré par leur lecture détaillée de ce manuscrit et leurs précieux commentaires. Ma reconnaissance va également à Nathalie Eisenbaum, Romuald Elie et David Proemel pour avoir accepté de faire partie du jury de ma soutenance.

Je suis très reconnaissant envers la Fondation Natixis pour avoir participé au financement de ma thèse.

Au LPMA, j'ai pu bénéficier d'excellentes conditions de travail. Je voudrais remercier tous les anciens et actuels membres du laboratoire pour une ambiance sympathique et agréable. Mes remerciements vont particulièrement à Alexandra, Alice, Candia, Chenguang, Eric, Minmin, Liping, Nina, Olga, Omar, Pamela, Saad, Sarah, Shen, Thibault, Wangru pour de bons moments passés ensemble au cours de ces dernières années. Je remercie également l'ensemble de l'équipe administrative du laboratoire pour leur gentillesse et leur disponibilité, qui rendent la vie des thésards beaucoup plus facile.

Merci également à Anis Al Gerbi, Pierre Blacque-Florentin, Jiatu Cai, Gaoyue Guo, Gang Liu, Zhenjie Ren, Sheng Fang, Dongli Wu, Junjian Yang, Xiaochuan Yang, avec qui les discussions sont toujours un plaisir.

Enfin, je remercie toute ma famille pour leur soutien et leurs encouragements constants, dans les moments de joie comme dans les moments de frustration.

Contents

Notations:

Acronyms and abbreviations:

càdlàg = right continuous with left limits

càglàd = left continuous with right limits

SDE = stochastic differential equation

BSDE = backward stochastic differential equation

PDE = partial differential equation

PPDE = path-dependent partial differential equation

s.t. = such that

a.s. = almost surely

a.e. = almost everywhere

e.g. = exempli gratia ≡ example given

i.e. = id est ≡ that is

etc. = et cetera ≡ and so on

Basic mathematical notations:

$\mathbb{M}_{d,n}(\mathbb{R})$ = set of $d \times n$ matrices with real coefficients

$\mathbb{M}_d(\mathbb{R}) = \mathbb{M}_{d,d}(\mathbb{R})$

$\mathcal{S}^d_+(\mathbb{R})$ = set of symmetric positive $d \times d$ matrices

$D([0,T],\mathbb{R}^d)$ = space of càdlàg paths from $[0,T]$ to $\mathbb{R}^d$, $d \in \mathbb{N}$

$C([0,T],\mathbb{R}^d)$ = space of continuous paths from $[0,T]$ to $\mathbb{R}^d$, $d \in \mathbb{N}$

$\cdot$ = scalar product in $\mathbb{R}^d$ (unless differently specified)

$\|\cdot\|_\infty$ = supremum norm in spaces of paths, e.g. $D([0,T],\mathbb{R}^d)$ or $C([0,T],\mathbb{R}^d)$

tr = trace operator, i.e. $\mathrm{tr}(A) = \sum_{i=1}^d A_{i,i}$ for $A \in \mathbb{M}_d(\mathbb{R})$.

${}^tA$ = transpose of a matrix A

$x(t)$ = value of x at time t for $x \in D([0,T],\mathbb{R}^d)$

$x(t-)$ = left limit of x at t, i.e. $\lim_{s\to t, s<t} x(s)$

$x(t+)$ = right limit of x at t, i.e. $\lim_{s\to t, s>t} x(s)$

$\Delta x(t)$ = jump of x at t, i.e. $x(t) - x(t-)$

x_t = $x(t \wedge \cdot) \in D([0,T],\mathbb{R}^d)$ the path of x 'stopped' at time t

x_{t-} = $x\mathbb{1}_{[0,t)} + x(t-)\mathbb{1}_{[t,T]} \in D([0,T],\mathbb{R}^d)$

x_t^e = $x_t + e\mathbb{1}_{[t,T]} \in D([0,T],\mathbb{R}^d)$ the vertical perturbation of a path $x \in D([0,T],\mathbb{R}^d)$ at time t by a vector $e \in \mathbb{R}^d$

Λ^d_T = space of d-dimensional càdlàg stopped paths

$\mathcal{W}^d_T$ = subspace of Λ^d_T of d-dimensional continuous stopped paths

$\mathcal{D}F$ = horizontal derivative of a non-anticipative functional F

$\nabla_\omega F$ = vertical derivative of a non-anticipative functional F

Introduction

Itô's stochastic calculus [39, 40, 46, 51, 19, 64, 60] is an important tool at the core of stochastic analysis and modern probability theory, which extends the methods of calculus to stochastic processes with irregular paths of infinite variation. It has many important applications in the analysis of phenomena with random, irregular evolution in time.

Itô calculus is suited for dealing with functions of stochastic processes which have non-smooth paths with infinite variation but finite quadratic variation in a probabilistic sense. Itô's stochastic integration theory allows to define integrals $\int_0^t H(s)dX(s)$ for a large class of non-anticipative integrands H with respect to a semimartingale X. It is easy to show that in this case, a naïve pathwise construction is impossible [60] due to the irregularity of the paths of X. Another important result is the Itô's formula, which is a change of variables formula for smooth functions $f(t, X(t))$ of the current state of a stochastic process X. An additional term linked to the notion of quadratic variation $[X]$ of X appears in the formula, which differentiates it from the standard differential calculus.

However, in many applications such as statistics of processes, physics or mathematical finance, uncertainty affects the current state of certain process even through its entire past history. In these cases, functionals, rather than functions, of a stochastic process are involved, i.e. quantities of the form

$$F(X_t) \quad \text{where } X_t := \{X(s), s \in [0,t]\}.$$

For example, in finance, the price of a path-dependent option can be seen as a functional of the entire past of the underlying. These functionals also arise naturally in the study of path-dependent stochastic equations and non-Markovian stochastic control problems.

Several approaches have been proposed to provide an analytical framework for the systematic study of such path-dependent functionals. One possible solution is to use the notion of Fréchet derivative for functions defined on a Banach space, for example, the space of paths $D([0,T])$ or $C([0,T])$, but path-dependent quantities in many applications may fail to be Fréchet-differentiable as it is a very strong notion of differentiability. When the underlying stochastic process is the Wiener process, the Malliavin calculus [70, 72, 3, 54, 50, 53] has proven to be a powerful tool for investigating various properties of Wiener functionals. Yet the construction of Malliavin derivative, which is a weak derivative for functionals on Wiener space, involves perturbations which apply to the whole path (both past and future) of the process. This leads to differential representations of Wiener functionals in terms of anticipative processes [6, 36, 54], whereas in many applications such as optimal control, or hedging in finance, it is more natural to consider non-anticipative, or causal versions of such quantities.

In a recent insightful work, inspired by methods used by practitioners to computer the sensitivity of path-dependent derivatives, Bruno Dupire [20] proposed a method to extend the Itô calculus to functionals of stochastic processes in a non-anticipative manner. He introduced two types of path perturbations, which allows to define two directional derivatives for functionals of the paths, called time and space derivatives, corresponding to the sensitivity of functionals to such perturbations.

More precisely, consider a (time-dependent) functional $F : [0,T]\times D([0,T]\times \mathbb{R})$ of a $\mathbb{R}$-valued càdlàg path $X : [0,T] \to \mathbb{R}$ such that the value $F(t,X)$ at time $t \in [0,T]$ depends only on the path X up to time t, i.e. $F(t,X) = F(t,X_t)$ where $X_t := \{X(s), s \in [0,t]\}$. Such functionals are called non-anticipative (or causal) functionals. The idea of Dupire is to analyze the influence of a smallperturbation of X on the functional F. For $t \in [0,T]$, he defined:

- X_t^h as X_t with the endpoint shifted by $h \in \mathbb{R}$:

$$X_t^h(s) = X(ss) \text{ for } s < t \text{ and } X_t^h(t) = x(t) + h.$$

- $X_{t,h}$ with $h > 0$ as an extension of X_t by freezing the endpoint over $[t, t+h]$:

$$X_{t,h}(s) = X(s) \text{ for } s \leq t \text{ and } X_{t,h}(s) = X(t) \text{ for } s \in [t, t+h].$$

Dupire defined the space derivative (or vertical derivative) of F as

$$\nabla_\omega F(t, X_t) := \lim_{h \to 0} \frac{F(t, X_t^h) - F(t, X_t)}{h},$$

(if the limit exists) and the time derivative (or horizontal derivative) of F as:

$$\mathcal{D}F(t, X_t) := \lim_{h \to 0^+} \frac{F(t+h, X_{t,h}) - F(t, X_t)}{h}.$$

Clearly $\nabla_\omega F$ still defines a non-anticipative functional, and thus we can similarly define the second order space derivative $\nabla_\omega^2 F$ as the space derivative of $\nabla_\omega F$.

It is important to emphasize that these derivatives are non-anticipative, i.e. they only depend on the underlying path up to the current time. Another remarkable observation is that the existence of these derivatives is weaker than requiring Fréchet or Gâteaux differentiability of F.

However, the main interest of these functional derivatives is that the knowledge of the second-order jet $(\mathcal{D}F, \nabla_\omega F, \nabla_\omega^2 F)$ allows to capture the behavior of the functional F along a given path if F and its derivatives satisfy in addition some continuity assumptions [8, 9, 10]. More precisely, we have a change of variable formula which is quite similar to the classical Itô formula: if X is continuous semimartingale, and we assume that the non-anticipative functional F is once time differentiable and twice space differentiable with F and its derivatives satisfying some continuity assumptions, then we have: for $t > 0$,

$$\begin{aligned} F(t, X_t) &= F(0, X_0) + \int_0^t \mathcal{D}F(s, X_s) ds + \int_0^t \nabla_\omega F(s, X_s) dX(s) \\ &\quad + \frac{1}{2} \int_0^t \nabla_\omega^2 F(s, X_s) d[X](s), \end{aligned}$$

where $[X]$ is the quadratic variation of the semimartingale X. This formula is called the functional Itô formula [8] as it extends the classical Itô formula to the case of a smooth functional F.

Following Dupire's idea, Cont and Fournié [8, 9, 10, 7] developed a rigorous mathematical framework for a path-dependent extension of the Itô calculus. While Dupire's original work [20] only considered functionals of stochastic processes and used probabilistic arguments in the proof of the functional Itô formula, Cont and Fournié [8] proposed a purely pathwise *non-anticipative functional calculus* without any reference to probability.

In the seminar paper *Calcul d'Itô sans probabilités* [27], Hans Föllmer proposed a non-probabilistic version of the Itô formula. The main concept is the quadratic variation of a path, which was identified by Föllmer as the relevant property of the path needed to derive the Itô formula. Combining this insight from Föllmer [27] with the ideas of Dupire [20], Cont and Fournié [8] constructed a pathwise functional calculus for non-anticipative functionals defined on the space of càdlàg paths.

Following the work of Dupire, Cont & Fournié, a lot of effort has been devoted to this functional calculus and its various applications, especially in the theory of path-dependent partial differential equations and applications to stochastic control and finance. One of the key topics in stochastic analysis is the deep link between Markov processes and partial differential equations. The development of the functional Itô calculus allows to extend this relation beyond the Markovian setting, leading to a new class of partial differential equations on the path space, commonly called path-dependent partial differential equations. The study of such equation, different notions of its solutions in particular, constitutes currently an active research topic, see for example [7, 21, 61, 15, 58, 22, 23, 14]. The pathwise functional calculus developed in [8] also provides naturally a model-free approach to the problem of continuous-time hedging and trading in finance, which allows to computer the gain of path-dependent trading strategies and analyze the robustness of such strategies in a pathwise manner, see for example [65, 68, 69, 59].

One of the main issues of this functional calculus is that although horizontal et vertical derivatives are directional derivatives, functionals in many applications are not horizontally or vertically differentiable in the sense of Dupire, and without these properties, the functional Itô formula may not be directly applicable [28]. For example, the price of some exotic (path-dependent) option, defined as the conditional expectation of the payoff, can be viewed as a non-anticipative functional of the underlying process, which is in general not horizontally or vertically differentiable (see [65] for some conditions on the payoff under which the conditional expectation admits a vertically differentiable functional representation). Another important example in which horizontal and vertical differentiabilities of a functional might be problematic comes from the theory of path-dependent partial differential equations. Such equations do not always admit a classical (smooth) solution even in the simple case of functional heat equations. This is also the reason various notions of solution are proposed for such equations.

This thesis is mainly devoted to dealing with functionals which are not necessarily horizontally or vertically differentiable. Motivated by applications in finance, we consider first functionals which represent conditional expectations. More precisely, let X be the solution of the following path-dependent stochastic differential equation:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d$$

where W is a standard d-dimensional Brownian motion, and X_t denotes the path of X up to time t. Clearly X is a non-Markovian process. We consider a functional $g : D([0, T], \mathbb{R}^d) \to \mathbb{R}$. The conditional expectation of $g(X_T)$ with respect to the natural filtration generated by the process X: $\mathbb{E}[g(X_T)|\mathcal{F}_t]$ can be viewed as a non-anticipative functional of X or W, i.e. there exist non-anticipative functionals F or G such that:

$$\mathbb{E}[g(X_T)|\mathcal{F}_t] = G(t, X_t) = F(t, W_t) \quad \text{a.s.} \tag{1}$$

If there exists some smooth functional F (or G) which satisfies (1), then by the functional Itô formula, we have:

$$g(X_T) = \mathbb{E}[g(X_T)] + \int_0^T \nabla_\omega F(t, W_t) \cdot dW(t)$$

since $\mathbb{E}[g(X_T)|\mathcal{F}_t]$ is a martingale. We have thus an explicit martingale representation formula for $g(X_T)$. In finance, the integrand $\nabla_\omega F(t, W_t)$ is closely related to the delta of the option. So in this case, the delta of the option is well defined and is explicit.

However, it is not always possible to construct or find smooth functionals F (or G) which satisfy (1) unless we put very strong assumptions on b, σ and g. Our idea is thus to construct a sequence of explicit smooth functionals $(F_n)_{n\geq 1}$ which approximates F in an appropriate sense, and approximate the integrand in the martingale representation of $g(X_T)$ by $\nabla_\omega F_n(\cdot, W_\cdot)$. The main advantages of this method are the following. First, we do not need any strong assumptions (for example differentiability conditions) on the coefficients b and σ or the functional g. Basically, a Lipschitz-type condition on these coefficients is sufficient. So our method applies in a very general framework. Another important convenience of this method is that the functionals constructed F_n are explicit and easy to implement and analyze, leading to an explicit control of the approximation error.

Another aspect we develop in this thesis to deal with functionals which are not necessarily smooth is to define a notion of weak derivative. A concept of weak functional derivative was initially proposed by Rama Cont and David-Antoine Fournié in [10] in a probabilistic setting. The main idea of their approach is the following. Let X be a square-integrable Brownian martingale. Consider now the space $\mathcal{C}_b^{1,2}(X)$ of all square-integrable martingales Y which admits a smooth functional representation of X, i.e. there exists a smooth functional F of class $\mathbb{C}_b^{1,2}$ (which is defined in the following chapter) such that $Y(t) = F(t, X_t)$ almost surely. They showed this space $\mathcal{C}_b^{1,2}(X)$ is dense in the space $\mathcal{M}^2(X)$ of all square-integrable martingales with initial value zero equipped with the norm $\|Y\| := \sqrt{\mathbb{E}|Y(T)|^2}$. This means that for any $Y \in \mathcal{M}^2(X)$, there exists a sequence of elements $(Y_n)_{n\geq 1}$ in $\mathcal{C}_b^{1,2}(X)$ such that

$\|Y_n - Y\| \underset{n\to\infty}{\to} 0$. Using the Itô isometry formula, Cont and Fournié showed that if F_n are smooth functionals such that $Y_n(t) = F_n(t, X_t)$ almost surely, then the sequence of processes $\nabla_\omega F_n(\cdot, X.)$ converges to some process ϕ in an appropriate sense. This allows to define ϕ as the weak derivative of Y with respect to X: $\phi := \nabla_X Y$.

Inspired by this idea, we propose a notion of weak derivative in a strictly pathwise framework using the pathwise calculus for non-anticipative functionals developed in [8]. More precisely, let x be a $\mathbb{R}^d$-valued continuous path defined in $[0, T]$, we would like to know for which functionals F we may define the weak derivative $\nabla F(\cdot, x.)$ of F along the path x. The idea is always to approximate F by a sequence of smooth functionals. However, compared to the construction of weak derivative in the probabilistic setting, several additional difficulties emerge in this pathwise construction. First, we need some kind of pathwise isometry formula to define the weak derivative. Using the notion of pathwise quadratic variation proposed by Hans Föllmer [27], we obtain a pathwise version of the Itô isometry formula which only holds for a subspace of all smooth functionals. Another obstacle comes from the fact that the set of paths with finite pathwise quadratic variation along a given sequence of partitions π does not form a vector space, which makes it hard to characterize the space of functionals which admit weak derivatives.

The notion of weak vertical derivative constructed in this thesis coincides with that proposed by Cont and Fournié when applied to a stochastic process. Moreover, this notion also allows to obtain a functional characterization of local martingales with respect to a reference process X, which enables us to define a notion of weak solution for path-dependent partial differential equations.

The thesis is structured as follows:

Chapter 1 The first chapter introduces the pathwise calculus for non-anticipative functionals developed by Rama Cont and David-Antoine Fournié in [8, 7]. We fix notations and recall important notions which will be used throughout the thesis, such as quadratic variation along a sequence of partitions, non-anticipative functionals, horizontal and vertical derivatives of such functionals, etc. The most important result of this chapter is a change of variable formula (theorem 1.16 [8]) which extends the pathwise Itô formula developed in [27] to the case of non-anticipative functionals.

Chapter 2 introduces the probabilistic counterpart of the pathwise functional calculus, also called the functional Itô calculus, following the work of Dupire [20] and Cont and Fournié [9, 10, 7]. We present in section 2.2 the weak functional calculus proposed in [10], which extends the vertical derivative operator with respect to a given square-integrable martingale X: ∇_X, to the space of all square-integrable martingales (theorem 2.9). This notion of weak derivative also allows to obtain a general martingale representation formula (theorem 2.10). We then recall briefly, in section 2.3, the relation between the functional Itô calculus and path-dependent partial differential equations, which extends the relation between Markov processes and partial differential equations to the path-dependent setting.

Chapter 3 deals with functionals which represent conditional expectations. Let W be a standard d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathcal{F}_t)$ its ($\mathbb{P}$-completed) natural filtration. Let X be the solution of a path-dependent stochastic differential equation:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d$$

where b and σ are two non-anticipative functionals, whose values at time t may depend on the path of X up to t. Let $g : D([0,T], \mathbb{R}^d) \to \mathbb{R}$ be a functional such that $g(X_T)$ is square-integrable. It is well known, by the martingale representation formula, that there exists a unique $(\mathcal{F}_t)$-predictable process ϕ with $\mathbb{E}[\int_0^T \mathrm{tr}(\phi(t)\, {}^t\phi(t))dt] < \infty$ such that:

$$g(X_T) = \mathbb{E}[g(X_T)] + \int_0^T \phi \cdot dW.$$

However, ϕ is in general not explicit unless in very specific cases where b, σ and g are assumed to satisfy strong conditions. The main objective of this chapter is to provide explicit constructive approximations for ϕ under a general framework. The idea is to represent the conditional expectation $\mathbb{E}[g(X_T)|\mathcal{F}_t]$ as a functional of W_t, i.e. $\mathbb{E}[g(X_T)|\mathcal{F}_t] = F(t, W_t)$ with F a non-anticipative functional . If F is smooth the Functional Itô formula yields

$$\phi(t) = \nabla_\omega F(t, W_t) \quad dt \times d\mathbb{P}\text{-a.e.}$$

where $\nabla_\omega F$ denotes the vertical derivative of F. If F is not smooth, we construct a sequence of smooth functionals (F_n) which converges to F in an appropriate sense, and such that $\phi_n(t) := \nabla_\omega F_n(t, W_t)$ approximates ϕ.

We construct the sequence of functionals (F_n) in section 3.3 (definition 3.6) using the Euler approximation scheme for path-dependent SDEs introduced in section 3.2. A first main result of this chapter is to show that F_n is smooth enough to apply the functional Itô formula (theorem 3.7 and theorem 3.8). We then establish, in section 3.4, several convergence results for our approximation method. In particular, we show in theorem 3.10 that we have an explicit rate of convergence under a slightly stronger assumption on g. Comparison with Malliavin calculus and some numerical aspects of this method are discussed respectively in section 3.5 and section 3.6. And finally in section 3.7, we provide some numerical examples to show that our method applies naturally to the problem of dynamic hedging of exotic options even in the case where the volatility might be path-dependent.

Chapter 4 is devoted to developing a notion of weak vertical derivative for functionals in a strictly pathwise setting. As we have mentioned previously, unlike the probabilistic construction of weak derivative proposed by Cont and Fournié [10], in absence of probabilistic assumptions, several supplementary difficulties need to be tackled in our construction.

First we establish in section 4.2 a pathwise isometry formula for cylindrical functionals (proposition 4.3) after explaining why this formula might not hold for general smooth functionals. The set $Q^\pi([0,T], \mathbb{R})$ of càdlàg paths with finite quadratic variation along a given sequence of partitions π is not a vector space [67]. So to tackle this problem, we propose in section 4.3 a notion of generalized quadratic variation along a given sequence of partitions (definition 4.5). The advantage of this notion is that the set $\widehat{Q}^{\pi}([0,T],\mathbb{R})$ of càdlàg paths with finite generalized quadratic variation along a sequence of partitions π forms now a vector space on which a semi-norm might be defined (proposition 4.4). We then introduce, in section 4.4, the notion of weak vertical derivative along a given path (proposition 4.5) and a characterization of this notion (proposition 4.6).

The notion of weak functional derivative along a single path x might not seem very interesting in itself as it provides little information on the functionals. However, this notion leads to interesting applications when applied to paths of a stochastic process, as shown in section 4.5 and section 4.6. The first question we are interested in is the relation between the notion of pathwise weak derivative proposed in this chapter and that of weak derivative constructed in a probabilistic framework by Cont and Fournié in [10]. The answer is given in subsection 4.5. Let X be a non-degenerate square-integrable martingale. We first show that every square-integrable $\mathcal{F}^X$-martingale M can be written as a functional of X: $M(t) = F(t, X_t)$ with F weakly vertically differentiable along almost all paths of X (proposition 4.7). Moreover, the pathwise weak derivative of F along $X(\omega, \cdot)$ coincides almost surely with the probabilistic weak derivative of M with respect to X defined in [10] (proposition 4.8).

In the last section 4.6, we examine the converse of the above result: given a Brownian martingale X, we seek to characterize functionals F such that $F(t, X_t)$ is a (local) martingale. When F is a smooth functional, it may be characterized as the solution of a path-dependent PDE [7]. Here we formulate a more general characterization using the notion of pathwise weak derivative (theorem 4.17). This result can also be extended to the case X is a continuous square-integrable semimartingale (proposition 4.11), which leads to a notion of weak solution for path-dependent PDEs (definition 4.18).

Chapter 1

Pathwise calculus for non-anticipative functionals

In this chapter, we review main concepts and results of the pathwise calculus for non-anticipative functionals developed in [8]. In his seminar paper *Calcul d'Itô sans probabilités* [27] in 1981, Föllmer proposed a non-probabilistic version of the Itô formula based on the notion of quadratic variation for paths which lie in the space $D([0,T],\mathbb{R}^d)$ of càdlàg paths along a certain sequence of partitions. In particular if $X=(X_t)_{t\in[0,T]}$ is a semimartingale [19, 51, 60], which is the classical setting for stochastic calculus, the paths of X have almost surely finite quadratic variation along a subsequence of such partitions. This shows clearly that the classical Itô calculus has actually a pathwise integral, and Itô integrals of the form $\int_0^T f(X(t-))dX(t)$ can be constructed as pathwise limits of Riemann sums for a certain class of functions f. A review of early results on this pathwise calculus is provided in [71].

Of course, the Itô integral with respect to a semimartingale X may be defined for a much larger class of adapted or non-anticipative integrands. So one of the main concerns to a functional extension of Föllmer's pathwise calculus is that this non-anticipativeness should be taken into account for functionals whereas it is automatically satisfied in the function case for $f(X(t-))$. This gives rise to the concept of non-anticipative functionals, which plays an important role in what follows.

The results in this chapter are entirely pathwise and do not make use of any probability measure. We start this chapter by recalling the notion of quadratic variation along a sequence of partitions for càdlàg paths and the concept of non-anticipative functionals, which are the cornerstones of this pathwise calculus. We then introduce, following Dupire [20], the horizontal and the vertical derivatives for non-anticipative functionals. Using these directional derivatives, we obtain the main result of this chapter, a functional change of variable formula [8] which shows that the variations of a functional along a càdlàg path with finite quadratic variation can be precisely described by these derivatives.

1.1 Quadratic variation along a sequence of partitions

Throughout the thesis, we denote by $D([0,T],\mathbb{R}^d)$ the space of càdlàg (right-continuous with left limits) paths defined on $[0,T]$ taking values in $\mathbb{R}^d$. Let X be the canonical process on $D([0,T],\mathbb{R}^d)$ and $(\mathcal{F}_t^0)$ the filtration generated by X.

Let $\pi = (\pi_m)_{m\geq 1}$ be a sequence of partitions of $[0,T]$ into intervals:

$$\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T).$$

$|\pi_m| := \sup\{|t_{i+1}^m - t_i^m|, i = 0, \cdots, k(m)-1\}$ will denote the mesh size of the partition. If in particular for any $n \geq m$, every interval $[t_i^n, t_{i+1}^n]$ of the partition π_n is included in one of the intervals of π_m, the sequence $(\pi_m)_{m\geq 1}$ is called a nested or refining sequence of partitions. In the following, we will always assume that π is a nested sequence of partitions unless otherwise specified.

In [27], Hans Föllmer proposed a notion of pathwise quadratic variation along a sequence of partitions π. Whereas Föllmer considered initially in his paper paths defined on $\mathbb{R}_+$, here we limit ourselves to paths defined on a finite time horizon $[0, T]$.

Definition 1.1 (Quadratic variation of a $\mathbb{R}$-valued path along a sequence of partitions [27])**.** *Let $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ be a sequence of partitions of $[0, T]$ with $|\pi_m| \underset{m\to\infty}{\to} 0$. A càdlàg path $x \in D([0, T], \mathbb{R})$ is said to have finite pathwise quadratic variation along π if the sequence of the discrete measures*

$$\xi_m := \sum_{i=0}^{k(m)-1} (x(t_{i+1}^m) - x(t_i^m))^2 \delta_{t_i^m}$$

where δ_t is the point mass at t, converges weakly to a Radon measure ξ such that $[x]_\pi(t) := \xi([0, t])$, $[x]_\pi$ has the following Lebesgue decomposition:

$$\forall t \in [0, T], \qquad [x]_\pi(t) = [x]_\pi^c(t) + \sum_{s \le t} |\Delta x(s)|^2 \tag{1.1}$$

with $[x]_\pi^c$ a continuous non-decreasing function and $\Delta x(s) := x(s) - x(s-)$. The non-decreasing function $[x]_\pi : [0, T] \to \mathbb{R}_+$ is then called the pathwise quadratic variation of x along the sequence of partitions $\pi = (\pi_m)_{m\ge1}$.

An intuitive characterization of this property which makes clear the link with the usual notion of quadratic variation is provided in [7] for continuous paths:

Lemma 1.2. *Let $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ be a sequence of partitions of $[0, T]$ with $|\pi_m| \underset{m\to\infty}{\to} 0$. A continuous path $x \in C^0([0, T], \mathbb{R})$ has finite pathwise quadratic variation along π if for any $t \in [0, T]$, the limit*

$$[x]_\pi(t) := \lim_{m\to\infty} \sum_{t_{i+1}^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2 < \infty \tag{1.2}$$

exists and the function $t \mapsto [x]_\pi(t)$ is a continuous increasing function.

This characterization is simpler since it only involves pointwise convergence of functions rather than weak convergence of measures.

Proof. If x has finite quadratic variation in the sense of definition 1.1, since $\xi([0,t])$ is continuous in $[0,T]$, we have, for any $t \in [0,T]$,

$$\xi_m([0,t]) = \sum_{t_i^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2 \underset{m\to\infty}{\to} \xi([0,t]).$$

Since $|\pi_m| \underset{m\to\infty}{\to} 0$, we have:

$$\lim_{m\to\infty} \sum_{t_{i+1}^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2 = \lim_{m\to\infty} \sum_{t_i^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2,$$

which implies $[x]_\pi(t) < \infty$ and $[x]_\pi(t) = \xi([0,t])$.

We assume now x satisfies (1.2). For $t \in [0,T]$, the cumulative distribution function of ξ_m at t converges to $[x]_\pi(t)$. Indeed, since $|\pi_m| \underset{m\to\infty}{\to} 0$, we have:

$$\lim_{m\to\infty} \xi_m([0,t]) = \lim_{m\to\infty} \sum_{t_i^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2 = \lim_{m\to\infty} \sum_{t_{i+1}^m \le t} (x(t_{i+1}^m) - x(t_i^m))^2 = [x]_\pi(t).$$

Now as $[x]_\pi$ is continuous and non-decreasing in $[0,T]$, there exists a Radon measure ξ such that (ξ_m) converges weakly to ξ, and we have $\xi([0,t]) = [x]_\pi(t)$. □

For paths which are only càdlàg, these two definitions are not equivalent. In fact for a path $x \in D([0,T],\mathbb{R})$, pointwise convergence in (1.2) implies that the partition 'exhausts' the jump times of x, i.e. if $\Delta x(s) \neq 0$ for some $s \in [0,T]$, then $s \in \pi_m$ for a certain m (thus for any $M \geq m$ since we assume π is a nested sequence of partitions). Otherwise $[x]_\pi(s)$ cannot have a Lebesgue decomposition of the form (1.1). This constraint becomes problematic if we want a set of càdlàg paths to have finite quadratic variation along a single fixed sequence of partitions.

However, since we work essentially with continuous paths throughout the thesis, these two definitions are precisely equivalent. We will use, in the following, the characterization (1.2) rather than Definition 1.1 as its formulation seems more natural to us. We denote by $Q^\pi([0,T],\mathbb{R})$ the set of $\mathbb{R}$-valued càdlàg paths x with finite pathwise quadratic variation along π.

Remark 1.3. *One may wonder why we take, in* (1.2), $\sum_{t_i^{m+1}\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$ *instead of* $\sum_{t_i^m\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$, *the cumulative distribution function* $\xi_m([0,t])$ *of* ξ_m *in Föllmer's definition. Other similar definitions which use, for example,* $\sum_{t_i^m\in\pi_m}(x(t_{i+1}^m\wedge t) - x(t_i^m\wedge t))^2$ *are also proposed in the literature. These definitions are equivalent in the case of continuous paths and define the same quadratic variation function since* $|\pi_m| \underset{m\to\infty}{\to} 0$. *However, the quantity considered in* (1.2) *seems more natural to us since the term* $\sum_{t_{i+1}^m\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$ *is both 'non-anticipative' and non-decreasing in* t *whereas* $\sum_{t_i^m\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$ *depends on the value of* x *after time* t *and* $\sum_{t_i^m\in\pi_m}(x(t_{i+1}^m\wedge t) - x(t_i^m\wedge t))^2$ *is not necessarily monotone in* t. *The property of being non-anticipative in the definition of* $[x]_\pi$ *is especially convenient if the partitions* $(t_i^m)_i$ *depend on* x, *i.e. they are stopping times with respect to the canonical filtration* $(\mathcal{F}_t^0)$. $\sum_{t_{i+1}^m\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$ *being non-decreasing is particularly interesting when* x *is continuous: the pointwise convergence of* $\sum_{t_{i+1}^m\leq t}(x(t_{i+1}^m) - x(t_i^m))^2$ *to* $[x]_\pi(t)$ *implies the uniform convergence in* t *in this case due to one of Dini's theorems. One interesting consequence of this property is the following:*

Corollary 1.1. *Let* $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ *be a sequence of partitions of* $[0,T]$ *with* $|\pi_m| \underset{m\to\infty}{\to} 0$, *and let* $x \in Q^\pi([0,T],\mathbb{R})$ *be a continuous path. For any continuous and bounded function* h, *we have:*

$$\sum_{t_i^{m+1}\leq t} h(t_i^m)(x(t_{i+1}^m) - x(t_i^m))^2 \underset{m\to\infty}{\to} \int_0^t h d[x]_\pi \tag{1.3}$$

uniformly in $t \in [0,T]$.

Proof. Observe first that (1.3) holds for any bounded function g of the form $g = \sum_j a_j \mathbb{1}_{[a_j,a_{j+1})}$. Consider now a continuous and bounded function h, we approximate h by a sequence of piecewise constant functions $(h^n)_{n\geq 1}$ such that h^n converges uniformly to h: $\|h - h^n\|_\infty \underset{n\to\infty}{\to} 0$. Since (h_n) is uniformly bounded, we have:

$$\sum_{t_i^{m+1}\leq t} h^n(t_i^m)(x(t_{i+1}^m) - x(t_i^m))^2 \underset{m\to\infty}{\to} \int_0^t h^n d[x]_\pi$$

uniformly in t and n. We conclude using the fact that:

$$\sum_{t_i^{m+1} \leq t} (h - h^n)(t_i^m)(x(t_{i+1}^m) - x(t_i^m))^2 \underset{m\to\infty}{\to} 0$$

and

$$\int_0^t (h - h^n) d[x]_\pi \underset{m\to\infty}{\to} 0$$

uniformly in t. $\square$

Note that the quadratic variation $[x]_\pi$ clearly depends on the sequence of partitions π. Indeed, as remarked in [7, Example 5.3.2], for any real-valued continuous path, we can construct a sequence of partitions along which that path has null quadratic variation.

Most stochastic processes have finite quadratic variation in the sense of Definition 1.1 along a certain sequence of partitions with probability equal to 1. Indeed, let X be a continuous semimartingale defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, it is well known that for any sequence of partitions $\pi = (\pi_m)_{m\geq 1}$ with $|\pi_m| \underset{m\to\infty}{\to} 0$, for any $t \in [0, T]$, the sequence

$$S_m^\pi(t) := \sum_{t_{i+1}^m \leq t} (X(t_{i+1}^m) - X(t_i^m))^2$$

converges in probability to the quadratic variation $[X, X](t)$ defined for semimartingales. So there exists a subsequence $(\pi_{\phi(m)})$ of (π_m) such that we have $\mathbb{P}$-almost sure convergence of $S_{\phi(m)}^\pi$, i.e.

$$\mathbb{P}(\{\omega \in \Omega, (X(.,\omega)) \text{ has pathwise quadratic variation along } \pi_\phi\}) = 1.$$

For a pathwise construction of functions with non-trivial quadratic variation along a sequence of dyadic partitions, we refer to the work of Mishura and Schied [52].

The notion of quadratic variation along a sequence of partitions is different from the p-variation of the path X for $p = 2$: the p-valuation involves taking a supremum over all partitions, not necessarily along a given sequence of partitions. Thus saying a path x is of finite 2-variation is stronger than assuming x has finite quadratic variation along a certain sequence of partitions.

In the sequel, we fix a nested sequence $\pi = (\pi_m)_{m\geq 1}$ of partitions with $|\pi_m| \underset{m\to\infty}{\to} 0$ and all limits will be considered along the same sequence π. We drop the subscript in $[x]_\pi$ whenever the context is clear.

The extension of this notion to vector-valued paths is somewhat subtle [27], since $Q^\pi([0,T],\mathbb{R})$ is not a vector space [67].

Definition 1.4 (Quadratic variation for vector-valued paths)**.** *A d-dimensional path $x = (x^1, \cdots, x^d) \in D([0,T],\mathbb{R}^d)$ is said to have finite pathwise quadratic variation along $\pi = (\pi_m)_{m\geq 1}$ if $x^i \in Q^\pi([0,T],\mathbb{R})$ and $x^i + x^j \in Q^\pi([0,T],\mathbb{R})$ for all $1 \leq i,j \leq d$. Then for $1 \leq i,j \leq d$ and $t \in [0,T]$, we have:*

$$\sum_{t_k^m \in \pi_m, t_{k+1}^m \leq t} (x^i(t_{k+1}^m) - x^i(t_k^m))(x^j(t_{k+1}^m) - x^j(t_k^m))$$
$$\underset{m\to\infty}{\to} [x]_{ij}(t) := \frac{[x^i + x^j](t) - [x^i](t) - [x^j](t)}{2}.$$

The matrix-valued function $[x] : [0,T] \to \mathcal{S}_d^+$ whose elements are given by:

$$[x]_{ij}(t) = \frac{[x^i + x^j](t) - [x^i](t) - [x^j](t)}{2}$$

is called the pathwise quadratic variation of the path x: for any $t \in [0,T]$,

$$\sum_{t_i^m \in \pi_m, t_{i+1}^m \leq t} (x(t_{i+1}^m) - x(t_i^m))\,{}^t(x(t_{i+1}^m) - x(t_i^m)) \underset{m\to\infty}{\to} [x](t) \in \mathcal{S}_d^+$$

and $[x]$ is non-decreasing in the sense of the order on positive symmetric matrices: for $h \geq 0$, $[x](t+h) - [x](t) \in \mathcal{S}_d^+$.

We denote by $Q^\pi([0,T],\mathbb{R}^d)$ the set of $\mathbb{R}^d-$valued càdlàg paths with finite pathwise quadratic variation along π.

Remark 1.5. *Note that in definition 1.4, we require that $x^i + x^j \in Q^\pi([0,T],\mathbb{R})$, which does not necessarily follow from requiring $x^i, x^j \in Q^\pi([0,T],\mathbb{R})$. This indicates that $Q^\pi([0,T],\mathbb{R})$ is not a vector space. Indeed, for $x, y \in Q^\pi([0,T],\mathbb{R})$,* *contrary to the definition of the variation for a path, here for a fixed* $t \in [0,T]$*, the sequence* (q_m) *defined by:*

$$q_m := \sum_{t_{i+1}^m \le t} (x(t_{i+1}^m) + y(t_{i+1}^m) - x(t_i^m) - y(t_i^m))^2 \tag{1.4}$$

is no longer non-decreasing in m *even for a nested sequence of partitions. Let* $\delta x_k^m = x(t_{k+1}^m) - x(t_k^m)$ *and* $\delta y_k^m = y(t_{k+1}^m) - y(t_k^m)$*. We have:*

$$|\delta x_k^m + \delta y_k^m|^2 = |\delta x_k^m|^2 + |\delta y_k^m|^2 + 2\delta x_k^m \delta y_k^m.$$

The cross-product terms may be positive, negative, or have an oscillating sign which may prevent the convergence of the sequence (q_m) *defined in* (1.4)*. An example of two paths* $x, y \in Q^\pi([0,T],\mathbb{R})$ *while* $x + y \notin Q^\pi([0,T],\mathbb{R})$ *is given in [67].*

The main purpose of this notion of pathwise quadratic variation is to introduce a non-probabilistic version of the Itô formula [27], in which the pathwise quadratic variation plays naturally the role of quadratic variation of a semimartingale in the classical Itô formula to describe the irregularity of a path, leading to a change of variable formula for paths which may have infinite variation.

Proposition 1.1 (Pathwise Itô formula for functions [27])**.** *Let* $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ *be a sequence of partitions of* $[0,T]$ *with* $|\pi_m| \underset{m\to\infty}{\to} 0$*, and* $x \in Q^\pi([0,T],\mathbb{R}^d)$ *a càdlàg path. Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a twice continuously differentiable function. We have,* $\forall t \in [0,T]$*,*

$$\begin{aligned} f(x(t)) \quad &= \quad f(x(0)) + \int_0^t \nabla f(x(s-)) \cdot d^\pi x(s) + \frac{1}{2}\int_0^t \operatorname{tr}(\nabla^2 f(x(s-)) d[x]_\pi^c(s)) \\ &\quad + \sum_{s \le t} [F(x(s)) - F(x(s-)) - \nabla f(x(s-)) \cdot \Delta x(s)] \end{aligned} \tag{1.5}$$

where ∇f *is the gradient of* f*,* $\nabla^2 f$ *its Hessian matrix, and the integral with respect to* $d^\pi x$ *is defined as the limit of non-anticipative Riemann sums along* π*:*

$$\int_0^t \nabla f(x(s-)) \cdot d^\pi x(s) := \lim_{m\to\infty} \sum_{t_i^m \in \pi_m, t_i^m \le t} \nabla f(x(t_i^m)) \cdot (x(t_{i+1}^m) - x(t_i^m)) \tag{1.6}$$

It is easy to deduce from (1.5) that if $x \in Q^\pi([0,T],\mathbb{R}^d)$ and $f \in C^2(\mathbb{R}^d,\mathbb{R})$, then $y = f \circ x \in Q^\pi([0,T],\mathbb{R})$ with

$$[y]_\pi(t) = \int_0^t \mathrm{tr}(\nabla^2 f(x(s-))d[x]_\pi^c(s)) + \sum_{s\le t} |f(x(s)) - f(x(s-))|^2.$$

This means that the class of paths with finite quadratic variation along a fixed sequence of partitions π is stable under C^2 transformations. The same result still holds if we only assume $f \in C^1(\mathbb{R}^d,\mathbb{R})$ (see for example [71]).

The pathwise Itô formula (1.5) was recently generalized by Cont and Fournié to the case where f is a path-dependent functional [8] using the non-anticipative functional calculus, which will be presented in the next section.

1.2 Non-anticipative functionals

As usual, we denote by $D([0,T],\mathbb{R}^d)$ the space of càdlàg paths on $[0,T]$ with values in $\mathbb{R}^d$. For a path $x \in D([0,T],\mathbb{R}^d)$, for any $t \in [0,T]$, we denote:

- $x(t) \in \mathbb{R}^d$ its value at time t;
- $x(t-) = \lim_{s\to t, s<t} x(s)$ its left limit at t;
- $\Delta x(t) = x(t) - x(t-)$ the jump size of x at t;
- $x_t = x(t \wedge \cdot) \in D([0,T],\mathbb{R}^d)$ the path of x stopped at time t;
- $x_{t-} = x\mathbb{1}_{[0,t)} + x(t-)\mathbb{1}_{[t,T]} \in D([0,T],\mathbb{R}^d)$;
- $\|x\|_\infty = \sup\{|x(t)|, t \in [0,T]\}$ the supremum norm of x on $[0,T]$.

The non-anticipative functional calculus is a functional calculus which applies to non-anticipative functionals of càdlàg paths with finite pathwise quadratic variation in the sense of definition 1.4. It was first introduced in [8], using the notion of directional derivatives proposed by Bruno Dupire [20]. We recall here some key concepts and results of this approach following [7].

Let X be the canonical process on $\Omega := D([0,T],\mathbb{R}^d)$, and $(\mathcal{F}_t^0)_{t\in[0,T]}$ the filtration generated by X. We are interested in non-anticipative functionals of X, that is, the functionals $F : [0,T]\times D([0,T],\mathbb{R}^d) \to \mathbb{R}$ such that:

$$\forall \omega \in \Omega, \quad F(t,\omega) = F(t,\omega_t). \tag{1.7}$$

The process $t \mapsto F(t,X)$ then only depends on the path of X up to time t and is $(\mathcal{F}_t^0)-$adapted.

It is sometimes convenient to define such functionals on the space of stopped paths [8, 7]: a stopped path is an equivalence class in $[0,T] \times D([0,T],\mathbb{R}^d)$ for the following equivalence relation:

$$(t,\omega) \sim (t',\omega') \Longleftrightarrow (t = t' \quad \text{and} \quad \omega_t = \omega'_{t'}). \tag{1.8}$$

The space of stopped paths is defined as the quotient of $[0,T]\times D([0,T],\mathbb{R}^d)$ by the equivalence relation (1.8):

$$\Lambda_T^d := \{(t,\omega(t\wedge .)), (t,\omega) \in [0,T]\times D([0,T],\mathbb{R}^d)\} = ([0,T]\times D([0,T],\mathbb{R}^d))/\sim .$$

We denote by $\mathcal{W}_T^d$ the subset of Λ_T^d consisting of continuous stopped paths. We endow Λ_T^d with a metric space structure by defining the following distance:

$$d_\infty((t,\omega),(t',\omega')) := \sup_{u\in[0,T]} |\omega(u\wedge t) - \omega'(u\wedge t')| + |t-t'| = \|\omega_t - \omega'_{t'}\|_\infty + |t-t'|.$$

(Λ_T^d, d_∞) is then a complete metric space.

Any map $F : [0,T]\times D([0,T],\mathbb{R}^d) \to \mathbb{R}$ satisfying the non-anticipativeness condition (1.7) can be equivalently viewed as a functional defined on the space Λ_T^d of stopped paths:

Definition 1.6. *A non-anticipative functional on $[0,T] \times D([0,T],\mathbb{R}^d)$ is a measurable map $F : (\Lambda_T^d, d_\infty) \to \mathbb{R}$ on the space (Λ_T^d, d_∞) of stopped paths.*

Using the metric structure of (Λ_T^d, d_∞), we denote by $\mathbb{C}^{0,0}(\Lambda_T^d)$ the set of continuous maps $F : (\Lambda_T^d, d_\infty) \to \mathbb{R}$. We can also define various weaker notions of continuity for non-anticipative functionals.

Definition 1.7. *A non-anticipative functional F is said to be:*

- *continuous at fixed times if for any $t \in [0,T]$, $F(t,.)$ is continuous with respect to the uniform norm $\|.\|_\infty$ in $[0,T]$, i.e. $\forall \omega \in D([0,T],\mathbb{R}^d)$, $\forall \epsilon > 0$, $\exists \eta > 0$, $\forall \omega' \in D([0,T],\mathbb{R}^d)$,*

$$\|\omega_t - \omega'_t\|_\infty < \eta \implies |F(t,\omega) - F(t,\omega')| < \epsilon.$$

- *left-continuous if $\forall (t,\omega) \in \Lambda_T^d$, $\forall \epsilon > 0$, $\exists \eta > 0$, $\forall (t',\omega') \in \Lambda_T^d$,*

$$(t' < t \quad \text{and} \quad d_\infty((t,\omega),(t',\omega')) < \eta) \implies |F(t,\omega) - F(t',\omega')| < \epsilon.$$

  We denote by $\mathbb{C}_l^{0,0}(\Lambda_T^d)$ the set of left-continuous functionals. Similarly, we can define the set $\mathbb{C}_r^{0,0}(\Lambda_T^d)$ of right-continuous functionals.

We also introduce a notion of local boundedness for functionals. We call a functional F boundedness-preserving if it is bounded on each bounded set of paths:

Definition 1.8. *A non-anticipative functional F is said to be boundedness-preserving if for any compact subset K of $\mathbb{R}^d$ and $t_0 < T$,*

$$\exists C(K,t_0) > 0, \forall t \in [0,t_0], \forall \omega \in D([0,T],\mathbb{R}^d), \quad \omega([0,t]) \subset K \implies F(t,\omega) < C(K,t_0).$$

We denote by $\mathbb{B}(\Lambda_T^d)$ the set of boundedness-preserving functionals.

Lemma 1.9 ([8])**.** *Several properties of regularity of the paths generated by non-anticipative functionals may be deduced from the regularities of such functionals:*

1. *If $F \in \mathbb{C}_l^{0,0}(\Lambda_T^d)$, then for all $\omega \in D([0,T],\mathbb{R}^d)$, the path $t \mapsto F(t,\omega_{t-})$ is left-continuous;*

2. *If $F \in \mathbb{C}_r^{0,0}(\Lambda_T^d)$, then for all $\omega \in D([0,T],\mathbb{R}^d)$, the path $t \mapsto F(t,\omega_t)$ is right-continuous;*

3. *If $F \in \mathbb{C}^{0,0}(\Lambda_T^d)$, then for all $\omega \in D([0,T],\mathbb{R}^d)$, the path $t \mapsto F(t,\omega_t)$ is càdlàg and continuous at each point where ω is continuous.*

4. *If $F \in \mathbb{B}(\Lambda_T^d)$, then for all $\omega \in D([0,T],\mathbb{R}^d)$, the path $t \mapsto F(t,\omega_t)$ is bounded.*

We now recall some notions of differentiability for non-anticipative functionals. For $e \in \mathbb{R}^d$ and $\omega \in D([0,T],\mathbb{R}^d)$, we define the vertical perturbation ω_t^e of (t,ω) as the càdlàg path obtained by shifting the path ω by e after t:

$$\omega_t^e := \omega_t + e\mathbb{1}_{[t,T]}.$$

Definition 1.10. *A non-anticipative functional F is said to be:*

- *horizontally differentiable at $(t,\omega) \in \Lambda_T^d$ if*

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

  exists. If $\mathcal{D}F(t,\omega)$ exists for all $(t,\omega) \in \Lambda_T^d$, then the non-anticipative functional $\mathcal{D}F$ is called the horizontal derivative of F.

- *vertically differentiable at $(t,\omega) \in \Lambda_T^d$ if the map:*

$$\begin{pmatrix} \mathbb{R}^d & \to & \mathbb{R} \\ e & \mapsto & F(t,\omega_t + e\mathbf{1}_{[t,T]}) \end{pmatrix}$$

  is differentiable at 0. Its gradient at 0 is called the vertical derivative of F at (t,ω):

$$\nabla_\omega F(t,\omega) := (\partial_i F(t,\omega), i = 1, \cdots, d) \in \mathbb{R}^d$$

  with

$$\partial_i F(t,\omega) := \lim_{h \to 0} \frac{F(t,\omega_t + he_i\mathbf{1}_{[t,T]}) - F(t,\omega_t)}{h}$$

  where $(e_i, i = 1, \cdots, d)$ is the canonical basis of $\mathbb{R}^d$. If F is vertically differentiable at all $(t,\omega) \in \Lambda_T^d$, $\nabla_\omega F : (t,\omega) \to \mathbb{R}^d$ defines a non-anticipative map called the vertical derivative of F.

We may repeat the same operation on $\nabla_\omega F$ and define similarly $\nabla_\omega^2 F$, $\nabla_\omega^3 F, \cdots$. This allows us to define the following classes of smooth functionals:

Definition 1.11 (Smooth functionals). *We define $\mathbb{C}_b^{1,k}(\Lambda_T^d)$ as the set of non-anticipative functionals $F : (\Lambda_T^d, d_\infty) \to \mathbb{R}$ which are:*

- *horizontally differentiable with $\mathcal{D}F$ continuous at fixed times;*
- *k times vertically differentiable with $\nabla_\omega^j F \in \mathbb{C}_l^{0,0}(\Lambda_T^d)$ for $j = 0, \cdots, k$;*
- *$\mathcal{D}F, \nabla_\omega F, \cdots, \nabla_\omega^k F \in \mathbb{B}(\Lambda_T^d)$.*

We denote $\mathbb{C}_b^{1,\infty}(\Lambda_T^d) = \bigcap_{k\geq 1} \mathbb{C}_b^{1,k}(\Lambda_T^d)$.

Remark 1.12. *Since horizontal and vertical derivatives are directional derivatives, the horizontal or vertical differentiability of a functional does not imply its continuity with respect to d_∞. Thus in definition 1.11, we still need to impose some continuity condition on F and its derivatives even if F is horizontally and vertically differentiable.*

However, many examples of functionals in applications may fail to be globally smooth, especially those involving exit times. Their derivatives may still be well behaved, except at certain stopping times. The following example illustrates a prototype of such functionals:

Example 1.13 ([28, Example 4.1]). *Let W be a one-dimensional Brownian motion, $b > 0$, and $M(t) = \sup_{0\leq s\leq t} W(s)$. Consider the $(\mathcal{F}_t^W)$-adapted martingale:*

$$Y(t) = \mathbb{E}[\mathbb{1}_{M(T)\geq b} | \mathcal{F}_t^W].$$

Then Y admits the functional representation $Y(t) = F(t, W_t)$ with F a non-anticipative functional defined as:

$$F(t,\omega) := \mathbb{1}_{\sup_{0\leq s\leq t}\omega(s)\geq b} + \mathbb{1}_{\sup_{0\leq s\leq t}\omega(s)< b}\left[2 - 2\Phi\left(\frac{b-\omega(t)}{\sqrt{T-t}}\right)\right],$$

where Φ is the cumulative distribution function of the standard normal variable. Observe that the functional $F \notin \mathbb{C}_l^{0,0}(\Lambda_T^d)$ since a path ω_t with $\omega(t) < b$ but $\sup_{0\leq s\leq t}\omega(s) = b$ can be approximated in the sup norm by paths with $\sup_{0\leq s\leq t}\omega(s) < b$. However, one can easily check that $\nabla_\omega F$, $\nabla_\omega^2 F$ and $\mathcal{D}F$ exist almost everywhere.

Fortunately, for such functionals, using a localization argument, we can still obtain a functional change of variable formula as will be presented in the following section. We now introduce the notion of locally smooth functionals:

Definition 1.14 (Locally smooth functionals [28]). *A non-anticipative functional F is said to be locally smooth of class $\mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ if there exists an increasing sequence $(\tau_n)_{n\geq 0}$ of stopping times with $\tau_0 = 0$ and $\tau_n \underset{n\to\infty}{\to} \infty$, and a sequence of functionals $F_n \in \mathbb{C}^{1,2}_b(\Lambda^d_T)$ such that:*

$$F(t,\omega) = \sum_{n\geq 0} F_n(t,\omega)\mathbb{1}_{[\tau_n(\omega),\tau_{n+1}(\omega))}(t).$$

Recall that a stopping time (or a non-anticipative random time) on $(\Omega, (\mathcal{F}^0_t)_{t\in[0,T]})$ is a measurable map $\tau : \Omega \to \mathbb{R}_+$ such that for any $t \geq 0$,

$$\{\omega \in \Omega, \tau(\omega) \leq t\} \in \mathcal{F}^0_t.$$

We end this section by giving a particularly important class of smooth functionals which are frequently used throughout this thesis, the so-called cylindrical non-anticipative functionals.

Definition 1.15 (Cylindrical functionals). *A non-anticipative functional F is said to be cylindrical if there exists $0 \leq t_1 < t_2 < \cdots < t_n \leq T$ such that for all $\omega \in D([0,T],\mathbb{R}^d)$,*

$$F(t,\omega) = h(\omega(t) - \omega(t_n-))\mathbb{1}_{t>t_n} g(\omega(t_1-),\omega(t_2-),\cdots,\omega(t_n-)) \tag{1.9}$$

for some continuous function $g \in \mathbb{C}(\mathbb{R}^{n\times d},\mathbb{R})$ and some twice differentiable function $h \in \mathbb{C}^2(\mathbb{R}^d,\mathbb{R})$ with $h(0) = 0$.

A cylindrical functional can be seen as the middle ground between a functional which depends on the whole path of ω and a function which at time t only depends on $\omega(t)$: the value of a cylindrical functional at time t depends only on the value of ω at t and a finite number of points in $[0,T]$ which are initially fixed. It is easy to check that the functional F defined by (1.9) is smooth, i.e. $F \in \mathbb{C}^{1,2}_b(\Lambda^d_T)$ with $\mathcal{D}F \equiv 0$ and for $j = 1,2$,

$$\nabla^j_\omega F(t,\omega) = \nabla^j h(\omega(t) - \omega(t_n-))\mathbb{1}_{t>t_n} g(\omega(t_1-),\omega(t_2-),\cdots,\omega(t_n-)).$$

1.3 Change of variable formula for functionals

In 2010, Cont and Fournié extended Föllmer's change of variable formula (1.5) to non-anticipative functionals defined on Λ_T^d [8]. A by-product of this formula is the definition of an analogue of Föllmer's integral (1.6) for certain class of functionals. When applying this formula to paths of a stochastic process with a properly chosen sequence of partitions, one obtains the functional Itô formula initially proposed in [20] with probabilistic arguments.

In the functional setting, the main difficulty compared to Föllmer's change of variable formula for functions (1.5) is to control the variation of a functional using its derivatives. In the case of a function, it is rather simple. It suffices to apply the Taylor expansion to the function as its value only depends on the value of the path on one point. However, a functional at t may depend on the whole path up to time t, which is clearly a quantity of infinite dimension. A direct application of the Taylor expansion is thus impossible.

The idea of Cont and Fournié [8] is to first approximate the path ω by a sequence of piecewise constant paths $(\omega^m)_{m\geq 0}$ which converges uniformly to ω. The variation of a non-anticipative functional F along ω^m can thus be decomposed into finite dimensional terms with its horizontal and vertical derivatives after a Taylor expansion. Then using the continuity assumption on F and its derivatives, we obtain a change of variable formula for F.

Let $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ be a nested sequence of partitions of $[0,T]$ with $|\pi_m| \to 0$, and $\omega \in Q^\pi([0,T],\mathbb{R}^d)$. Since ω has at most a countable set of jump times, we may always assume that the partition exhausts the jump times in the sense that:

$$\sup_{t\in[0,T]\setminus\pi_m} |\Delta\omega(t)| \underset{m\to\infty}{\to} 0. \tag{1.10}$$

We now define the piecewise constant approximations ω^m of ω by:

$$\omega^m(t) := \sum_{i=0}^{k(m)-1} \omega(t_{i+1}^m-)\mathbb{1}_{[t_i^m,t_{i+1}^m)}(t) + \omega(T)\mathbb{1}_{\{T\}}(t). \tag{1.11}$$

Clearly under the assumption (1.10), the sequence $(\omega^m)_{m\geq 0}$ converges uniformly to ω:

$$\sup_{t\in[0,T]} |\omega^m(t) - \omega(t)| \underset{m\to\infty}{\to} 0.$$

We now present the main result of this section, the change of variable formula for non-anticipative functionals of càdlàg paths:

Theorem 1.16 (Pathwise change of variable formula for smooth functionals [8])**.** *Let $\omega \in Q^\pi([0,T],\mathbb{R}^d)$ satisfying* (1.10)*. Then for any non-anticipative functional $F \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$, the limit*

$$\int_0^T \nabla_\omega F(t,\omega_{t-})\cdot d^\pi\omega := \lim_{m\to\infty} \sum_{i=0}^{k(m)-1} \nabla_\omega F(t_i^m, \omega_{t_i^m-}^{m,\Delta\omega(t_i^m)}) \cdot (\omega(t_{i+1}^m) - \omega(t_i^m)) \tag{1.12}$$

exists, and we have:

$$\begin{aligned} F(T,\omega_T) &= F(0,\omega_0) + \int_0^T \mathcal{D}F(t,\omega_{t-})dt \\ &\quad + \int_0^T \nabla_\omega F(t,\omega_{t-})\cdot d^\pi\omega + \frac{1}{2}\int_0^T \mathrm{tr}(\nabla^2_\omega F(t,\omega_{t-}) d[\omega]^c_\pi(t)) \\ &\quad + \sum_{t\in(0,T]} [F(t,\omega_t) - F(t,\omega_{t-}) - \nabla_\omega F(t,\omega_{t-})\cdot\Delta\omega(t)] \end{aligned} \tag{1.13}$$

The detailed proof of this theorem can be found in [8] under more general assumptions. Here we just provide some ideas of proof in the case ω is continuous.

First, using localization by a sequence of stopping times, we may assume that $F \in \mathbb{C}^{1,2}_b(\Lambda^d_T)$. We now decompose the variation of F between t_i^m and t_{i+1}^m for $0 \leq i \leq k(m)-1$ along two directions, the horizontal and the vertical:

$$\begin{aligned} F(t_{i+1}^m, \omega^m_{t_{i+1}^m-}) - F(t_i^m, \omega^m_{t_i^m-}) &= F(t_{i+1}^m, \omega^m_{t_{i+1}^m-}) - F(t_i^m, \omega^m_{t_i^m}) \\ &\quad + F(t_i^m, \omega^m_{t_i^m}) - F(t_i^m, \omega^m_{t_i^m-}) \end{aligned} \tag{1.14}$$

Since ω^m is piecewise constant, the first term of (1.14) can be written as an integral of the horizontal derivatives of F along ω^m:

$$F(t_{i+1}^m, \omega^m_{t_{i+1}^m-}) - F(t_i^m, \omega^m_{t_i^m}) = \int_{t_i^m}^{t_{i+1}^m} \mathcal{D}F(u, \omega^m_{t_i^m})du,$$

and the second term of (1.14) can be developed using a second-order Taylor expansion into terms involving first and second order vertical derivatives of F:

$$\begin{aligned} F(t_i^m, \omega^m_{t_i^m}) - F(t_i^m, \omega^m_{t_i^m-}) &= \nabla_\omega F(t_i^m, \omega^m_{t_i^m-}) \cdot \delta\omega_i^m \\ &+ \frac{1}{2}\mathrm{tr}(\nabla^2_\omega F(t_i^m, \omega^m_{t_i^m-})^t(\delta\omega_i^m)\delta\omega_i^m) + r_i^m \end{aligned}$$

where $\delta\omega_i^m := \omega(t_{i+1}^m) - \omega(t_i^m)$, and r_i^m is the reminder term in the Taylor expansion. We now sum all the terms from $i = 0$ to $k(m) - 1$. Using the continuity and boundedness preserving property of F, $\mathcal{D}F$, $\nabla_\omega F$ and $\nabla^2_\omega F$, each sum converges to the corresponding integral. The only difficulty might be the convergence of the term with $\nabla^2_\omega F$ for which we need a diagonal argument for weak convergence of measures:

Lemma 1.17 ([8]). *Let $(\mu_n)_{n\geq 1}$ be a sequence of Radon measures on $[0,T]$ converging vaguely to a Radon measure μ with no atoms, and let $(f_n)_{n\geq 1}$ and f be left-continuous functions defined on $[0,T]$ such that there exists $K > 0$, for all $t \in [0,T]$, $|f_n(t)| \leq K$ and $f_n(t) \underset{n\to\infty}{\to} f(t)$. Then we have:*

$$\forall t \in [0,T], \quad \int_0^t f_n d\mu_n \underset{n\to\infty}{\to} \int_0^t f d\mu.$$

1.4 Functionals defined on continuous paths

We end this chapter with a short discussion of functionals defined on the space of continuous paths. So far, we have been working with functionals F defined on the space of (stopped) càdlàg paths, which is a natural choice because even if a path ω is continuous, the definition of $\nabla_\omega F$ (definition 1.10) involves evaluating F on paths to which a jump perturbation has been added.

However, in some applications, we may only have access to the value of F along continuous paths. For example, if X is a continuous semimartingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we want to study the process

$F(t, X_t)$ (for instance when it is a local martingale). A priori, we have no information of F outside the space of (stopped) continuous paths:

$$\mathcal{W}_T^d := \{(t, \omega) \in \Lambda_T^d, \omega \in C([0,T], \mathbb{R}^d)\},$$

or the topological support of the law of X, i.e. F may take any value outside $\mathcal{W}_T^d$. In this case, the very definition of the vertical derivative $\nabla_\omega F$ becomes ambiguous.

Fortunately, due to a result of Cont and Fournié [7], the notion of vertical derivative is still well defined if F satisfies some regularity assumption:

Proposition 1.2 ([7]). *If $F^1, F^2 \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ coincide on continuous paths, i.e.*

$$\forall \omega \in C([0,T], \mathbb{R}^d), \forall t \in [0,T), F^1(t, \omega_t) = F^2(t, \omega_t),$$

then their first and second vertical derivatives also coincide on continuous paths: $\forall \omega \in C([0,T], \mathbb{R}^d)$, $\forall t \in [0,T)$,

$$\nabla_\omega F^1(t, \omega_t) = \nabla_\omega F^2(t, \omega_t) \quad \text{and} \quad \nabla_\omega^2 F^1(t, \omega_t) = \nabla_\omega^2 F^2(t, \omega_t).$$

One main interest of this proposition is that it allows us to define the class $\mathbb{C}_b^{1,2}(\mathcal{W}_T^d)$ of non-anticipative functionals F as the restriction of any functional $\tilde{F} \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ on $\mathcal{W}_T^d$ without having to extend the definition of F to the full space Λ_T^d:

$$F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T^d) \quad \Longleftrightarrow \quad \exists \tilde{F} \in \mathbb{C}_b^{1,2}(\Lambda_T^d), \quad \tilde{F}_{|\mathcal{W}_T^d} = F.$$

For such functionals, the change of variable formula (theorem 1.16) still holds if we only consider continuous paths ω:

Theorem 1.18 (Pathwise change of variable formula for functionals defined on continuous paths [7]). *For any $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T^d)$ and $\omega \in C([0,T], \mathbb{R}^d) \cap Q^\pi([0,T], \mathbb{R}^d)$, the limit*

$$\int_0^T \nabla_\omega F(t, \omega_t) \cdot d^\pi \omega := \lim_{m \to \infty} \sum_{i=0}^{k(m)-1} \nabla_\omega F(t_i^m, \omega_{t_i^m -}^m) \cdot (\omega(t_{i+1}^m) - \omega(t_i^m))$$

exists, and we have:

$$\begin{aligned} F(T,\omega_T) &= F(0,\omega_0) + \int_0^T \mathcal{D}F(t,\omega_t)dt + \int_0^T \nabla_\omega F(t,\omega_t)\cdot d^\pi\omega \\ &\quad + \frac{1}{2}\int_0^T \operatorname{tr}(\nabla_\omega^2 F(t,\omega_t)d[\omega]_\pi(t)) \end{aligned}$$

where ω^m *is the same piecewise constant approximation of* ω *as in* (1.11).

Chapter 2

Functional Itô calculus

In this chapter, we introduce a probability measure on the space of paths, and apply the non-anticipative calculus we have presented in the previous chapter under the probabilistic framework. In the initial work of Dupire [20], motivated by applications in mathematical finance, especially the problem of pricing and hedging for path-dependent options, he introduced two directional derivatives for functionals and obtained the functional Itô formula in a probabilistic setting.

Whereas Dupire's initial proof of this formula made use of probabilistic arguments, we have already seen in the previous chapter that the functional Itô formula has in fact a pathwise interpretation, with the quadratic variation for semimartingales replaced by the notion of pathwise quadratic variation (definition 1.4), and the stochastic integral replaced by the pathwise integral defined as the limit of non-anticipative Riemann sums (1.12).

Hence the pathwise change of variable formula (theorem 1.16 or theorem 1.18 in the continuous case) is stronger than the functional Itô formula in the sense that it directly implies the latter formula with a well chosen sequence of partitions. Moreover, it allows to deal with more general stochastic processes such as the so-called Dirichlet (or finite energy) processes which are defined as the sum of a semimartingale and a process with zero quadratic variation along a sequence of dyadic subdivisions.

So why are we particularly interested in a formula applied to semimartingales even when we already have a stronger pathwise version? First, we have at our disposal a whole set of stochastic tools when dealing with semimartingales. The notion of quadratic variation and that of stochastic integration are well known for semimartingales, which makes the functional Itô formula more natural and comprehensible than its pathwise counterpart. In particular, terms in this formula no longer depend on the sequence of partitions chosen.

More importantly, the space of martingales or semimartingales possess better space structure than the path space $D([0,T],\mathbb{R}^d)$ or $C([0,T],\mathbb{R}^d)$. For example, the space of square-integrable martingales is a Hilbert space when equipped with the L^2-norm at time T whereas the space of paths with finite quadratic variation along a sequence of partitions π: $Q^\pi([0,T],\mathbb{R}^d)$ is not even a vector space (see remark 1.5). The advantage of these structures is that they allow us to explore much more than a simple formula with semimartingales. For example, a notion of weak derivative can be defined for square-integrable martingales, leading to a general martingale representation formula, as we shall see in section 2.2 of this chapter.

Another key topic in stochastic analysis is the deep link between Markov processes and partial differential equations, which can also be extended beyond Markovian setting, leading to the so-called path-dependent partial differential equations. We review several properties of such equation and its relation with stochastic processes in section 2.3 following [7].

2.1 Functional Itô formula

Let X be a semimartingale defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, equipped with the natural filtration $\mathbb{F}=(\mathcal{F}_t)_{t\geq 0}$ of X. First we show briefly why the functional Itô formula can be implied directly by the pathwise change of variable formula (theorem 1.16) with a well-chosen sequence of partitions, as we have mentioned. To fix the idea, we only consider the case X is continuous. We distinguish the pathwise quadratic variation $[X(\omega,\cdot)]_\pi$ (or simply $[X]_\pi$) defined for a path $X(\omega)$ of X along a sequence of partitions π and $[X]$ the quadratic variation defined for continuous semimartingales. We start with a simple lemma which links these two notions of quadratic variation:

Lemma 2.1. *There exists a sequence of partitions $\pi = (\pi_m)_{m\geq 1}$ of $[0,T]$ with $|\pi_m| \underset{m\to\infty}{\to} 0$ such that the paths of X lie in $Q^\pi([0,T],\mathbb{R}^d)$ with probability 1, i.e.*

$$\mathbb{P}(\{\omega\in\Omega, X(\omega,\cdot)\in Q^\pi([0,T],\mathbb{R}^d)\}) = 1,$$

and

$$\mathbb{P}(\{\omega\in\Omega, [X(\omega,\cdot)]_\pi = [X](\omega)\}) = 1.$$

Proof. For any sequence of partitions $\pi = (\pi_m)_{m\geq 1}$ of $[0,T]$ with $|\pi_m|\to 0$, by definition of the quadratic variation defined for semimartingales, we have, for any $t\in[0,T]$,

$$\sum_{t_i^m\in\pi_m, t_{i+1}^m\leq t} (X(t_{i+1}^m) - X(t_i^m))^2 \underset{m\to\infty}{\to} [X](t)$$

in probability uniformly in t. We can thus extract a sub-sequence of π which achieves the result. □

This lemma means that with a well-chosen sequence of partitions, we have

$$\int_0^T \operatorname{tr}(\nabla_\omega^2 F(t,X_t) d[X]_\pi(t)) = \int_0^T \operatorname{tr}(\nabla_\omega^2 F(t,X_t) d[X](t))$$

$\mathbb{P}$-almost surely. What remains to show is that the pathwise integral

$$\int_0^T \nabla_\omega F(t,X_t)\cdot d^\pi X(t) := \lim_{m\to\infty} \sum_{i=0}^{k(m)-1} \nabla_\omega F(t_i^m, X^m_{t_i^m-})\cdot (X(t_{i+1}^m) - X(t_i^m)) \tag{2.1}$$

coincides with the stochastic integral $\int_0^T \nabla_\omega F(t,X_t)\cdot dX(t)$ with probability 1. We recall that X^m is the piecewise constant approximation of X defined by:

$$X^m(t) := \sum_{i=0}^{k(m)-1} X(t_{i+1}^m)\mathbb{1}_{[t_i^m,t_{i+1}^m)}(t) + X(T)\mathbb{1}_{\{T\}}(t).$$

Actually the right-hand side of equation (2.1) can also be written as:

$$\int_0^T f_m(t) \cdot dX(t)$$

with

$$f_m(t) := \sum_{i=0}^{k(m)-1} \nabla_\omega F(t_i^m, X_{t_i^m -}^m) \mathbb{1}_{(t_i^m, t_{i+1}^m]}(t).$$

Since the sequence of processes $f_m(t)$ converges $\mathbb{P}$-almost surely to the process $\nabla_\omega F(t, X_t)$ and (f_m) can be bounded independently of m by some X-integrable process using the boundedness-preserving property of $\nabla_\omega F$, the dominated convergence theorem for stochastic integrals [60] ensures that $\int_0^T f_m(t) \cdot dX(t)$ converges in probability to $\int_0^T \nabla_\omega F(t, X_t) \cdot dX(t)$. As $\int_0^T f_m(t) \cdot dX(t)$ converges almost surely to $\int_0^T \nabla_\omega F(t, X_t) \cdot d^\pi X(t)$ par definition of the pathwise integral, the two limits have to be the same, which ends the proof.

We obtain thus the functional Itô formula in the case X is a continuous semimartingale:

Theorem 2.2 (Functional Itô formula: continuous case [10])**.** *Let X be a $\mathbb{R}^d$-valued continuous semimartingale defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and $F \in \mathbb{C}_{loc}^{1,2}(\mathcal{W}_T^d)$. Then for any $t \in [0, T]$, we have:*

$$\begin{aligned} F(t, X_t) &= F(0, X_0) + \int_0^t \mathcal{D}F(s, X_s) ds + \int_0^t \nabla_\omega F(s, X_s) \cdot dX(s) \\ &+ \frac{1}{2} \int_0^t \mathrm{tr}(\nabla_\omega^2 F(s, X_s) d[X](s)) \end{aligned} \tag{2.2}$$

$\mathbb{P}$-almost surely. In particular, $Y(t) := F(t, X_t)$ is a continuous semimartingale: the class of continuous semimartingales is stable under transformation by $\mathbb{C}_{loc}^{1,2}(\mathcal{W}_T^d)$ functionals.

Actually the same functional Itô formula still holds for functionals whose vertical derivatives are right-continuous rather than left-continuous. We denote by $\mathbb{C}_{b,r}^{1,2}(\Lambda_T^d)$ the set of non-anticipative functionals F satisfying:

- F is horizontally differentiable with $\mathcal{D}F$ continuous at fixed times;

- F is twice vertically differentiable with $F \in \mathbb{C}_l^{0,0}(\Lambda_T^d)$ and $\nabla_\omega F, \nabla_\omega^2 F \in \mathbb{C}_r^{0,0}(\Lambda_T^d)$;

- $\mathcal{D}F, \nabla_\omega F, \nabla_\omega^2 F \in \mathbb{B}(\Lambda_T^d)$;

The localization is more delicate in this case, and we are not able to state a local version of the functional Itô formula by simply replacing $F_n \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ by $F_n \in \mathbb{C}_{b,r}^{1,2}(\Lambda_T^d)$ in definition 1.14 (see remark 4.2 in [28]). However if the stopping times τ_n are deterministic, then the functional Itô formula is still valid (proposition 2.4 and remark 4.2 in [28]).

Definition 2.3. *A non-anticipative functional is said to be locally smooth of class $\mathbb{C}_{loc,r}^{1,2}(\Lambda_T^d)$ if there exists an increasing sequence $(t_n)_{n\geq 0}$ of deterministic times with $t_0 = 0$ and $t_n \underset{n\to\infty}{\to} \infty$, and a sequence of functionals $F_n \in \mathbb{C}_{b,r}^{1,2}(\Lambda_T^d)$ such that:*

$$F(t,\omega) = \sum_{n\geq 0} F_n(t,\omega)\mathbb{1}_{[t_n,t_{n+1})}(t), \quad \forall (t,\omega) \in \Lambda_T^d.$$

Theorem 2.4 (Functional Itô formula: continuous case with right-continuous vertical derivatives [28])**.** *Let X be a $\mathbb{R}^d$-valued continuous semimartingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $F \in \mathbb{C}_{loc,r}^{1,2}(\Lambda_T^d)$. Then for any $t \in [0,T]$, we have:*

$$\begin{aligned} F(t,X_t) &= F(0,X_0) + \int_0^t \mathcal{D}F(s,X_s)ds \\ &\quad + \int_0^t \nabla_\omega F(s,X_s)\cdot dX(s) + \frac{1}{2}\int_0^t \mathrm{tr}(\nabla_\omega^2 F(s,X_s)d[X](s)) \end{aligned}$$

$\mathbb{P}$*-almost surely.*

The case when X is a càdlàg semimartingale is similar. By applying the pathwise change of variable formula (theorem 1.16) with a well-chosen sequence of partitions, we obtain the following càdlàg version of the functional Itô formula.

Theorem 2.5 (Functional Itô formula: càdlàg case [8]). *Let X be a $\mathbb{R}^d$-valued càdlàg semimartingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $F \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$. Then for any $t \in [0, T]$, we have:*

$$\begin{aligned} F(t, X_t) &= F(0, X_0) + \int_0^t \mathcal{D}F(s, X_{s-})ds \\ &+ \int_0^t \nabla_\omega F(s, X_{s-}) \cdot dX(s) + \frac{1}{2}\int_0^t \mathrm{tr}(\nabla^2_\omega F(s, X_{s-})d[X]^c(s)) \\ &+ \sum_{s\in(0,t]} [F(s, X_s) - F(s, X_{s-}) - \nabla_\omega F(s, X_{s-}) \cdot \Delta X(s)] \end{aligned} \tag{2.3}$$

$\mathbb{P}$-almost surely where $[X]^c$ denotes the continuous part of the quadratic variation of X. In particular, $Y(t) := F(t, X_t)$ is a semimartingale: the class of semimartingales is stable under transformation by $\mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ functionals.

Although these two formulas are implied by the stronger pathwise formula (theorem 1.16 and theorem 1.18), Cont and Fournié [10] also provided a direct probabilistic proof based on the classical Itô formula.

The main idea of the proof is quite similar to that of the pathwise change of variable formula. We first approximate X by a sequence of piecewise constant processes. We then decompose the increment of F into a horizontal and a vertical part, and to computer the vertical increment, instead of using a second order Taylor expansion, we apply directly the classical Itô formula. Finally to pass from its piecewise constant approximations to the process X itself, we again use the dominated convergence theorem, and its extension to stochastic integrals [60].

In [10], Cont and Fournié also introduced a second argument A in the non-anticipative functional F. A is an $\mathcal{S}^+_d$-valued process, which represents the Radon-Nikodym derivative of the quadratic variation of a continuous $\mathbb{R}^d$-valued semimartingale X with respect to the Lebesgue measure:

$$[X](t) = \int_0^t A(s)ds.$$

Although the process A is itself adapted to the natural filtration generated by X, thus can be considered as a functional of X, many interesting examples of functionals involving the quadratic variation of X cannot be represented as a continuous functional with respect to the supremum norm of X, and this continuity property is essential to apply the functional Itô formula. Thus introducing A as a second argument allows to control the regularity of the functional with respect to the quadratic variation of X by simply requiring continuity of F with respect to the pair (X, A).

We now review briefly the main results of this approach proposed by Cont and Fournié. More details can be found in their initial paper [10]. Let $\omega = (x, v) \in D([0, T], \mathbb{R}^d) \times D([0, T], \mathcal{S}_d^+)$, and we define

$$\|\omega\|_\infty := \|x\|_\infty + \|v\|_\infty.$$

We define the space $(\mathcal{S}_T^d, d_\infty)$ of stopped paths:

$$\mathcal{S}_T^d := \{(t, \omega(t \wedge \cdot)), (t, \omega) \in [0, T] \times D([0, T], \mathbb{R}^d) \times D([0, T], \mathcal{S}_d^+)\}$$

with

$$d_\infty((t, \omega), (t', \omega')) := \|\omega_t - \omega'_{t'}\|_\infty + |t - t'|.$$

Assumption 2.1. *Let $F : (\mathcal{S}_T^d, d_\infty) \to \mathbb{R}$ be a non-anticipative functional. We assume that F has predictable dependence with respect to the second argument:*

$$\forall (t, x, v) \in \mathcal{S}_T^d, \quad F(t, x, v) = F(t, x_t, v_{t-}). \tag{2.4}$$

With condition (2.4), the vertical derivatives of F with respect to the variable v are zero, so the vertical derivative on the product space coincides with the vertical derivative with respect to the variable x, and we continue to denote it as ∇_ω. The horizontal derivative of F can be defined similarly as in definition 1.10. We can thus define the corresponding class of smooth functionals $\mathbb{C}_{loc}^{1,2}(\mathcal{S}_T^d)$ by analogy with definition 1.14. We now state a change of variable formula for non-anticipative functionals which are allowed to depend on the path X and its quadratic variation.

Theorem 2.6 (Functional Itô formula with dependence on quadratic variation [10])**.** *Let X be a $\mathbb{R}^d$-valued continuous semimartingale defined on a* *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{S}^d_T)$ *a non-anticipative functional satisfying* (2.4)*. Then for any* $t \in [0, T]$*, we have:*

$$\begin{aligned} F(t, X_t, A_t) &= F(0, X_0, A_0) + \int_0^t \mathcal{D}F(s, X_s, A_s) ds \\ &+ \int_0^t \nabla_\omega F(s, X_s, A_s) \cdot dX(s) + \frac{1}{2} \int_0^t \operatorname{tr}(\nabla^2_\omega F(s, X_s, A_s) d[X](s)) \end{aligned}$$

$\mathbb{P}$*-almost surely. In particular,* $Y(t) := F(t, X_t, A_t)$ *is a continuous semimartingale.*

2.2 Weak functional calculus and martingale representation

As we have mentioned earlier, when working with semimartingles or martingales, we have at our disposal many stochastic tools which allows to explore more than when working with paths. One useful result in stochastic calculus is the Itô isometry formula which is particularly interesting in this case as it links a smooth functional with its vertical derivative.

More precisely, let X be a (continuous) square-integrable martingale, and we assume that $Y(t) := F(t, X_t)$ is also a square-integrable martingale with initial value zero where F is a smooth functional (for example $\in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$). A simple application of the functional Itô formula 2.2 yields:

$$Y(T) = \int_0^T \nabla_\omega F(t, X_t) \cdot dX(t).$$

We now apply the Itô isometry formula, and we obtain:

$$\|Y(T)\|_{L^2} = \|\nabla_\omega F(., X_.)\|_{L^2([X])} \tag{2.5}$$

where

$$\|\phi\|^2_{L^2([X])} := \mathbb{E}\left[\int_0^T \operatorname{tr}(\phi(t)\, {}^t\phi(t) d[X](t))\right].$$

Since the space $L^2([X])$ is a Hilbert space, the isometry relation (2.5) allows to extend the notion of vertical derivative to the closure with respect to the

L^2-norm at time T of the space of square-integrable martingales which admit a smooth functional representation, which turns out to be the space of all square-integrable martingales adapted to the filtration generated by X, as shown in [10].

In the following, we review in more detail this extension of the notion of vertical derivative developed by Cont and Fournié [10]. In [7], Cont carries this extension further, that is to the space of square-integrable semimartingales adapted to the filtration generated by X.

Let X be a continuous, $\mathbb{R}^d$-valued semimartingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Since all processes we consider are functionals of X, without loss of generality, we assume that Ω is the canonical space $D([0,T], \mathbb{R}^d)$ of càdlàg paths, $X(t,\omega) = \omega(t)$ is the coordinate process, and $\mathbb{P}$ a probability measure on Ω under which X is a continuous semimartingale. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ the filtration $(\mathcal{F}^X_{t+})_{t\in[0,T]}$ after $\mathbb{P}$-augmentation.

Again we assume that

$$[X](t) = \int_0^t A(s)ds, \quad \forall t \in [0,T] \tag{2.6}$$

for some $\mathcal{S}_d^+$-valued càdlàg process A.

Assumption 2.2 (Non-degeneracy of local martingale component)**.** *We assume that the process A defined in* (2.6) *is non-singular almost everywhere, i.e.*

$$\det(A(t)) \neq 0 \quad dt \times d\mathbb{P}\text{-}a.e. \tag{2.7}$$

Consider now an $\mathbb{F}$-adapted process Y which admits the following functional representation of X:

$$Y(t) = F(t, X_t), \quad \forall t \in [0,T] \tag{2.8}$$

with $F : \Lambda_T^d \to \mathbb{R}$ a non-anticipative functional. Clearly the functional representation (2.8) is not unique: if we modify F outside the topological support of $\mathbb{P}^X$, we obtain another non-anticipative functional satisfying (2.8). In particular, let $F^1, F^2 \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}_T^d)$ be such that, for any $t \in [0,T)$,

$$F^1(t, X_t) = F^2(t, X_t) \quad \mathbb{P}\text{-a.s.}$$

A priori, we do not know if $\nabla_\omega F^1(t, X_t)$ is equal to $\nabla_\omega F^2(t, X_t)$ up to an evanescent set as the vertical derivative $\nabla_\omega F$ seems to depend on the value of F on discontinuous paths of the form $X_t + e\mathbb{1}_{[t,T]}$.

However, thanks to the non-degeneracy condition (2.7), $\nabla_\omega F(., X_.)$ is uniquely defined up to an evanescent set, and is independent of the choice of $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ in the representation (2.8), which leads to the following definition:

Definition 2.7 (Vertical derivative of a process [10])**.** *We define $\mathcal{C}^{1,2}_{loc}(X)$ as the set of $\mathcal{F}$-adapted processes Y which admits a functional representation of X in $\mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$:*

$$\mathcal{C}^{1,2}_{loc}(X) := \left\{ Y : \exists F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T), \quad Y(t) = F(t, X_t) \quad dt \times d\mathbb{P}\text{−a.e.} \right\}. \quad (2.9)$$

Under assumption 2.2, for any $Y \in \mathcal{C}^{1,2}_{loc}(X)$, the predictable process

$$\nabla_X Y(t) := \nabla_\omega F(t, X_t)$$

is uniquely defined up to an evanescent set, independently of the choice of $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ in the representation (2.9)*. We call the process $\nabla_X Y$ the vertical derivative of Y with respect to X.*

This definition leads to the following representation for smooth local martingales.

Proposition 2.1 (Representation of smooth local martingales [7])**.** *For any local martingale $Y \in \mathcal{C}^{1,2}_{loc}(X)$, we have the following representation:*

$$Y(T) = Y(0) + \int_0^T \nabla_X Y \cdot dM$$

where M is the local martingale component of X.

Now we consider the case where X is a square-integrable Brownian martingale to explore the link between the notion of vertical derivative and the martingale representation theorem.

Let W be a standard d-dimensional Brownian motion with $\mathbb{F} = (\mathcal{F}_t^W)_{t\in[0,T]}$ its ($\mathbb{P}$-completed) natural filtration. Let X be a $\mathbb{R}^d$-valued process defined by:

$$X(t) = X(0) + \int_0^t \sigma(s)dW(s)$$

where σ is a $\mathbb{F}$-adapted process satisfying

$$\mathbb{E}\left[\int_0^T \|\sigma(t)\|^2 dt\right] < \infty \quad \text{and} \quad \det(\sigma(t)) \neq 0 \quad dt \times d\mathbb{P}\text{-a.e.} \tag{2.10}$$

Under these conditions, X is a square-integrable martingale with the predictable representation property [42, 64]: for any square-integrable $\mathcal{F}_T^W$-measurable random variable H, or equivalently, any square-integrable $\mathbb{F}$-martingale Y defined by $Y(t) := \mathbb{E}[H|\mathcal{F}_t^W]$, there exists a unique $\mathbb{F}$-predictable process ϕ with $\mathbb{E}\left[\int_0^T \operatorname{tr}(\phi(t)\,{}^t\phi(t)d[X](t))\right] < \infty$ such that:

$$H = Y(T) = \mathbb{E}[H] + \int_0^T \phi(t)\cdot dX(t).$$

As suggested by proposition 2.1, when the martingale Y admits a smooth functional representation of X, $\nabla_X Y$ is a natural candidate of ϕ in the martingale representation. We denote by $\mathcal{L}^2(X)$ the Hilbert space of $\mathbb{F}$-predictable processes such that:

$$\|\phi\|^2_{\mathcal{L}^2(X)} := \mathbb{E}\left[\int_0^T \operatorname{tr}(\phi(t)\,{}^t\phi(t)d[X](t)\right] < \infty.$$

Proposition 2.2 (Martingale representation formula for smooth martingales [10])**.** *Let $Y \in \mathcal{C}^{1,2}_{loc}(X)$ be a square-integrable martingale. Then $\nabla_X Y \in \mathcal{L}^2(X)$ and Y admits the following martingale representation:*

$$Y(T) = Y(0) + \int_0^T \nabla_X Y\cdot dX. \tag{2.11}$$

As we have mentioned earlier, since $\mathcal{L}^2(X)$ is a Hilbert space, using the Itô isometry formula, the operator $\nabla_X : \mathcal{C}^{1,2}_{loc}(X) \to \mathcal{L}^2(X)$ can be extended to a larger space by a density argument. We now formalize this idea.

Let $\mathcal{M}^2(X)$ be the space of $\mathbb{R}$-valued square-integrable $\mathbb{F}$-martingales with initial value zero equipped with the norm:

$$\|Y\|^2_{\mathcal{M}^2(X)} := \mathbb{E}[Y(T)^2].$$

By proposition 2.2, the vertical derivative ∇_X defines a continuous map

$$\nabla_X : D(X) := \mathcal{C}^{1,2}_{loc}(X) \cap \mathcal{M}^2(X) \to \mathcal{L}^2(X)$$

on the set $D(X)$ of square-integrable martingales with a smooth functional representation of X. Moreover, we have the following isometry property:

$$\|Y\|^2_{\mathcal{M}^2(X)} = \|\nabla_X Y\|^2_{\mathcal{L}^2(X)}, \quad \forall Y \in D(X). \tag{2.12}$$

As $\mathcal{L}^2(X)$ is a Hilbert space, the operator ∇_X can then be extended to the closure $\overline{D(X)}$ of $D(X)$ with respect to the norm $\mathcal{M}^2(X)$, and ∇_X still defines an isometry on $\overline{D(X)}$. What remains is to determine this space $\overline{D(X)}$.

Lemma 2.8 (Density of $\mathcal{C}^{1,2}_{loc}(X)$ in $\mathcal{M}^2(X)$ [10])**.** *$\{\nabla_X Y | Y \in D(X)\}$ is dense in $\mathcal{L}^2(X)$ and $D(X)$ is dense in $\mathcal{M}^2(X)$, i.e. $\overline{D(X)} = \mathcal{M}^2(X)$.*

We show in fact that the set $\{\nabla_X Y\}$ where $Y(t) = F(t, X_t) \in D(X)$ with F cylindrical functionals of the form:

$$F(t,\omega) = (\omega(t) - \omega(t_n-))\mathbb{1}_{t>t_n} g(\omega(t_1-), \omega(t_2-), \cdots, \omega(t_n-)), \quad g \in \mathbb{C}^\infty_b(\mathbb{R}^n, \mathbb{R})$$

is already dense in $\mathcal{L}^2(X)$. And the density of $\{\nabla_X Y | Y \in D(X)\}$ in $\mathcal{L}^2(X)$ entails the density of $D(X)$ in $\mathcal{M}^2(X)$ since the Itô integral with respect to X is a bijective isometry.

We can now extend the operator ∇_X to the whole space of square-integrable martingales $\mathcal{M}^2(X)$.

Theorem 2.9 (Extension of ∇_X to $\mathcal{M}^2(X)$ [10])**.** *The vertical derivative*

$$\nabla_X : D(X) \to \mathcal{L}^2(X)$$

admits a unique continuous extension to $\mathcal{M}^2(X)$, namely

$$\nabla_X : \mathcal{M}^2(X) \to \mathcal{L}^2(X)$$

which is a bijective isometry characterized by the following integration by parts formula: for $Y \in \mathcal{M}^2(X)$, $\nabla_X Y$ *is the unique element of* $\mathcal{L}^2(X)$ *such that, for any* $Z \in D(X)$,

$$\mathbb{E}[Y(T)Z(T)] = \mathbb{E}\left[\int_0^T \mathrm{tr}(\nabla_X Y(t)\, {}^t\nabla_X Z(t) d[X](t))\right].$$

∇_X *is the adjoint of the Itô stochastic integral*

$$\begin{aligned} I_X : \mathcal{L}^2(X) &\rightarrow \mathcal{M}^2(X) \\ \phi &\mapsto \int_0^{\cdot} \phi \cdot dX \end{aligned}$$

in the following sense: $\forall \phi \in \mathcal{L}^2(X)$, $\forall Y \in \mathcal{M}^2(X)$,

$$\mathbb{E}\left[Y(T)\int_0^T \phi \cdot dX\right] = \langle \nabla_X Y, \phi \rangle_{\mathcal{L}^2(X)}.$$

In particular, we have, for any $\phi \in \mathcal{L}^2(X)$,

$$\nabla_X \left(\int_0^{\cdot} \phi \cdot dX \right) = \phi.$$

This result leads to a general version of the martingale representation formula, valid for all square-integrable martingales.

Theorem 2.10 (Martingale representation formula: general case [10])**.** *For any square-integrable* $\mathbb{F}$*-martingale* Y*, we have, for any* $t \in [0,T]$,

$$Y(t) = Y(0) + \int_0^t \nabla_X Y \cdot dX \quad \mathbb{P}\text{-}a.s.$$

Cont [7] further extended this weak vertical derivative operator ∇_X to the space of all square-integrable $\mathbb{F}$-semimartingales. He also iterated this procedure to define a scale of 'Sobolev' spaces. Here we will not enter in detail into this extension.

2.3 Functional Kolmogorov equations

In this section, we review briefly the functional extension of main concepts and results which connect stochastic processes and partial differential equations developed in [7]. Clearly, under the functional framework, stochastic processes are no longer Markovian. One important class of non-Markovian processes is semimartingales which can be represented as the solution of a stochastic differential equation whose coefficients are allowed to be path-dependent.

Let W be a standard d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ its ($\mathbb{P}$-completed) natural filtration. We consider the following stochastic differential equation with path-dependent coefficients:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d \tag{2.13}$$

where $b : \Lambda_T^d \to \mathbb{R}^d$, $\sigma : \Lambda_T^d \to \mathbb{M}_d(\mathbb{R})$ are non-anticipative maps.

This class of processes is a natural path-dependent extension of diffusion processes. Various conditions, such as the functional Lipschitz property and boundedness conditions, may be imposed for the existence and uniqueness of the solution (see for example [60]). Here we provide one example of such conditions: b and σ are assumed to be Lipschitz continuous with respect to the distance d_∞ defined on Λ_T^d by:

$$d_\infty((t,\omega),(t',\omega')) := \|\omega_t - \omega'_{t'}\|_\infty + |t-t'|.$$

Assumption 2.3. *We assume that $b : (\Lambda_T^d, d_\infty) \to \mathbb{R}^d$, $\sigma : (\Lambda_T^d, d_\infty) \to \mathbb{M}_d(\mathbb{R})$ are Lipschitz continuous:*

$$\exists K_{Lip} > 0, \quad \forall t, t' \in [0,T], \forall \omega, \omega' \in D([0,T], \mathbb{R}^d),$$

$$|b(t,\omega) - b(t',\omega')| + \|\sigma(t,\omega) - \sigma(t',\omega')\| \leq K_{Lip}\, d_\infty((t,\omega),(t',\omega')).$$

Under assumption 2.3, (2.13) has a unique strong $\mathbb{F}$-adapted solution X.

Proposition 2.3. *Under assumption 2.3, there exists a unique $\mathbb{F}$-adapted process X satisfying* (2.13)*. Moreover for $p \geq 1$, we have:*

$$\mathbb{E}\left[\|X_T\|_\infty^{2p}\right] \leq C(1 + |x_0|^{2p})e^{CT} \tag{2.14}$$

for some constant $C = C(p, T, K_{Lip})$ depending on p, T and K_{Lip}.

Remark 2.11. *Assumption 2.3 might seem to be quite strong. Indeed, the previous proposition still holds under weaker conditions. For example, the joint Lipschitz condition in (t,ω) can be replaced by a Lipschitz condition only in ω together with a boundedness condition in t:*

$$\exists K_{Lip} > 0, \quad \forall t \in [0,T], \forall \omega, \omega' \in D([0,T],\mathbb{R}^d),$$

$$|b(t,\omega) - b(t,\omega')| + \|\sigma(t,\omega) - \sigma(t,\omega')\| \leq K_{Lip} \|\omega - \omega'\|_\infty$$

and

$$\sup_{t\in[0,T]} |b(t,\bar{0})| + \|\sigma(t,\bar{0})\| < \infty$$

where $\bar{0}$ denotes the path which takes constant value 0.

Proof. Existence and uniqueness of a strong solution follows from [60] (Theorem 7, Chapter 5). Let us just prove (2.14). Using the Burkholder-Davis-Gundy inequality and Hölder's inequality, we have:

$$\begin{aligned}
\mathbb{E}\left[\|X_T\|_\infty^{2p}\right] &\leq C(p)\left(|x_0|^{2p} + \mathbb{E}\left[\left(\int_0^T |b(t,X_t)|^2 dt\right)^p\right] + \mathbb{E}\left[\left(\int_0^T \|\sigma(t,X_t)\|^2 dt\right)^p\right]\right) \\
&\leq C(p,T)\left(|x_0|^{2p} + \mathbb{E}\left[\int_0^T |b(t,X_t)|^{2p} dt\right] + \mathbb{E}\left[\int_0^T \|\sigma(t,X_t)\|^{2p} dt\right]\right) \\
&\leq C(p,T)\left(|x_0|^{2p} + \mathbb{E}\left[\int_0^T \left(|b(0,\bar{0})| + \|\sigma(0,\bar{0})\| + K_{Lip}(t + \|X_t\|_\infty)\right)^{2p} dt\right]\right) \\
&\leq C(p,T,K_{Lip})\left(|x_0|^{2p} + 1 + \int_0^T \mathbb{E}\left[\|X_t\|_\infty^{2p}\right] dt\right).
\end{aligned}$$

We conclude using Gronwall's inequality. □

When considering the link between a diffusion process and the associated Kolmogorov equation, the domain of the partial differential equation is related to the support of the random variable $X(t)$. Analogously, we shall also take into account the support of the law of X on the space of paths when considering a functional extension of the Kolmogorov equation.

The topological support of a continuous process X is the smallest closed set $\mathrm{supp}(X)$ in $(C([0,T],\mathbb{R}^d), \|\cdot\|_\infty)$ such that $\mathbb{P}(X \in \mathrm{supp}(X)) = 1$. It can also be characterized by the following property: $\mathrm{supp}(X)$ is the set of paths

$\omega \in C([0,T],\mathbb{R}^d)$ for which every Borel neighborhood has strictly positive measure, i.e.

$$\mathrm{supp}(X) = \{\omega \in C([0,T],\mathbb{R}^d) | \, \forall \text{ neighborhood } V \text{ of } \omega, \mathbb{P}(X \in V) > 0\}.$$

We denote by $\Lambda_T(X)$ the set of stopped paths obtained from paths in $\mathrm{supp}(X)$:

$$\Lambda_T(X) := \{(t,\omega) \in \Lambda_T^d, \omega \in \mathrm{supp}(X)\}.$$

Clearly, for a continuous process X, $\Lambda_T(X) \subset \mathcal{W}_T^d$.

Similarly to the Markovian case, for a given process X, we are interested in the functionals which have the local martingale property, i.e. the functionals F such that $F(t, X_t)$ is a local martingale. We call such functionals $\mathbb{P}^X$-harmonic functionals:

Definition 2.12 ($\mathbb{P}^X$-harmonic functionals). *A non-anticipative functional F is called $\mathbb{P}^X$-harmonic if $Y(t) := F(t, X_t)$ is a $\mathbb{P}$-local martingale.*

Smooth $\mathbb{P}^X$-harmonic functionals can be characterized as solutions to the following functional Kolmogorov equation on the domain $\Lambda_T(X)$.

Theorem 2.13 (Functional Kolmogorov equation [7]). *If $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}_T^d)$ and $\mathcal{D}F \in \mathbb{C}^{0,0}_l(\mathcal{W}_T^d)$, then $Y(t) := F(t, X_t)$ is a local martingale if and only if F satisfies*

$$\mathcal{D}F(t,\omega_t) + b(t,\omega_t) \cdot \nabla_\omega F(t,\omega_t) + \frac{1}{2}\mathrm{tr}(\nabla^2_\omega F(t,\omega_t)\sigma \, {}^t\sigma(t,\omega_t)) = 0 \qquad (2.15)$$

for all $(t,\omega) \in \Lambda_T(X)$.

The proof of this theorem is based on the functional Itô formula and is provided in [7]. When $F(t,\omega) = f(t,\omega(t))$ and the coefficients b and σ are not path-dependent, this equation reduces to the well-known backward Kolmogorov equation [45].

When $X = W$ is a standard d-dimensional Brownian motion, the topological support of X is all continuous paths starting from 0, i.e. $\mathrm{supp}(X) = C_0([0,T],\mathbb{R}^d)$ with

$$C_0([0,T],\mathbb{R}^d) := \{\omega \in C([0,T],\mathbb{R}^d), \omega(0) = 0\},$$

and the functional Kolmogorov equation reduces to a functional heat equation:

Corollary 2.1. *If $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ and $\mathcal{D}F \in \mathbb{C}^{0,0}_l(\mathcal{W}^d_T)$, then $Y(t) := F(t, W_t)$ is a local martingale if and only if for any $t \in [0,T]$ and $\omega \in C_0([0,T], \mathbb{R}^d)$,*

$$\mathcal{D}F(t,\omega) + \frac{1}{2}\mathrm{tr}(\nabla^2_\omega F(t,\omega)) = 0. \tag{2.16}$$

By analogy with the classical finite-dimensional parabolic PDEs, we can also introduce the notions of sub-solution and super-solution of the path-dependent PDEs. A comparison principle is also established in [7] which allows to prove the uniqueness of solution.

Definition 2.14 (Sub-solutions and super-solutions)**.** *$F \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ is called a sub-solution (resp. super-solution) of* (2.15) *on a domain $U \subset \Lambda^d_T$ if for all $(t,\omega) \in U$,*

$$\mathcal{D}F(t,\omega) + b(t,\omega)\cdot\nabla_\omega F(t,\omega) + \frac{1}{2}(\nabla^2_\omega F(t,\omega)\sigma\, {}^t\sigma(t,\omega)) \geq 0\ (\textit{resp.} \leq 0).$$

We now state a comparison principle for path-dependent PDEs.

Theorem 2.15 (Comparison principle [7])**.** *Let $\underline{F} \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ be a sub-solution of* (2.15) *and $\overline{F} \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ be a super-solution of* (2.15) *such that for every $\omega \in supp(X)$, $\underline{F}(T,\omega) \leq \overline{F}(T,\omega)$, and*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\underline{F}(t,X_t) - \overline{F}(t,X_t)|\right] < \infty.$$

Then we have:

$$\forall (t,\omega) \in \Lambda_T(X), \quad \underline{F}(t,\omega) \leq \overline{F}(t,\omega).$$

A straightforward consequence of this comparison principle is the following uniqueness result for $\mathbb{P}$-uniformly integrable solutions of the functional Kolmogorov equation (2.15).

Theorem 2.16 (Uniqueness of solutions [7])**.** *Let* $H : (C([0,T],\mathbb{R}^d), \|\cdot\|_\infty) \to \mathbb{R}$ *be a continuous functional, and let* F^1, $F^2 \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ *be solutions of the functional Kolmogorov equation* (2.15) *satisfying*

$$\forall \omega \in supp(X), \quad F^1(T,\omega) = F^2(T,\omega) = H(\omega_T)$$

and

$$\mathbb{E}\left[\sup_{t\in[0,T]} |F^1(t,X_t) - F^2(t,X_t)|\right] < \infty.$$

Then they coincide on the topological support of X*, i.e.*

$$\forall (t,\omega) \in \Lambda_T(X), \quad F^1(t,\omega) = F^2(t,\omega).$$

Combining this uniqueness result with the characterization of the $\mathbb{P}^X$-harmonic functionals as solutions of a functional Kolmogorov equation leads to an extension of the well-known Feynman-Kac formula to the path-dependent case.

Theorem 2.17 (Feynman-Kac formula for path-dependent functionals [7])**.** *Let* $H : (C([0,T],\mathbb{R}^d), \|\cdot\|_\infty) \to \mathbb{R}$ *be a continuous functional, and let* $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ *be a solution of the functional Kolmogorov equation* (2.15) *satisfying*

$$\forall \omega \in C([0,T],\mathbb{R}^d), \quad \mathcal{F}(T,\omega) = H(\omega_T)$$

and

$$\mathbb{E}\left[\sup_{t\in[0,T]} |F(t,X_t)|\right] < \infty,$$

then F *has the following probabilistic representation:*

$$\forall \omega \in supp(X), \quad F(t,\omega) = \mathbb{E}[H(X_T)|X_t = \omega_t] = \mathbb{E}^{\mathbb{P}_{(t,\omega)}}[H(X_T)],$$

where $\mathbb{P}_{(t,\omega)}$ *is the law of the unique solution of the path-dependent SDE*

$$dX(t) = b(t,X_t)dt + \sigma(t,X_t)dW(t)$$

with initial condition $X_t = \omega_t$*. In particular,*

$$F(t,X_t) = \mathbb{E}[H(X_T)|\mathcal{F}_t] \quad dt \times d\mathbb{P}\text{-}a.s.$$

Clearly the above result assumes the existence of a solution to the functional Kolmogorov equation (2.15) with a given regularity, and then derive a probabilistic representation of this solution. In the case of theorem 2.17, this is equivalent to constructing, given a functional H, a smooth version of the conditional expectation $\mathbb{E}[H|\mathcal{F}_t]$.

Similarly to the case of finite-dimensional PDEs, such so-called strong or classical solutions, with the required differentiability, may fail to exist in many examples of interest, even for the functional heat equation (2.16). We refer to Peng and Wang [58] and Riga [65] for sufficient conditions under which a path-dependent PDE admits a classical solution.

Various notions of generalized solution have been proposed for such path-dependent equations. Cont [7] proposed a notion of Sobolev-type weak solution using the weak functional Itô calculus presented in section 2.2. Cosso and Russo [13] introduced a notion of strong-viscosity solution as the pathwise limit of classical solutions to semi-linear parabolic path-dependend PDEs. Ekren et al. [21] proposed a notion of viscosity solution for semi-linear parabolic path-dependent PDEs which allows to extend the non-linear Feynman-Kac formula to non-Markovian case. Ekren et al. [22] generalized this notion to deal with fully non-linear parabolic path-dependent PDEs. Different comparison results have also been established under this framework [23, 62, 63] (see also [61] for a review of these concepts). Cosso [12] extended the result of [23] to the case of a possibly degenerate diffusion coefficient in the forward process driving the BSDE.

Chapter 3

Weak approximation of martingale representations

In this chapter, we present a systematic method for computing explicit approximations to martingale representations for a large class of Brownian functionals. The approximations are obtained by computing a directional derivative of the weak Euler scheme (definition 3.6) and yield a consistent estimator for the integrand in the martingale representation formula for any square-integrable functional of the solution of an SDE with path-dependent coefficients. Explicit convergence rates are derived for functionals which are Lipschitz-continuous in the supremum norm (theorem 3.10). Our results require neither the Markov property, nor any differentiability conditions on the functional or the coefficients of the stochastic differential equations. We present, in section 3.7, several numerical applications of our method to the problem of dynamic hedging in finance.

The main results of this chapter were published in [11].

3.1 Introduction

Let W be a standard d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ its ($\mathbb{P}$-completed) natural filtration.

Then for any square-integrable $\mathcal{F}_T$-measurable random variable H, or equivalently, any square-integrable $\mathbb{F}$-martingale $Y(t) := \mathbb{E}[H|\mathcal{F}_t]$, there exists a unique $(\mathcal{F}_t)$-predictable process ϕ with $\mathbb{E}\left[\int_0^T \mathrm{tr}(\phi(t)\,{}^t\phi(t))dt\right] < \infty$ such that:

$$H = Y(T) = \mathbb{E}[H] + \int_0^T \phi(t) \cdot dW(t). \tag{3.1}$$

The classical proof of this representation result (see for example [64]) is non-constructive. However in many applications, such as stochastic control or mathematical finance, we are interested in an explicit expression for the process ϕ, which represents an optimal control or a hedging strategy.

Expressions for the integrand ϕ have been derived using a variety of methods and assumptions, for example Markovian techniques [17, 24, 56, 41, 26], integration by parts [5] or Malliavin calculus [6, 36, 2, 54, 43, 29, 53]. Some of these methods are limited to the case where Y is a Markov process; others require differentiability of H in the Fréchet or Malliavin sense [6, 54, 29, 30] or an explicit form for the density [5]. Almost all of these methods invariably involve an approximation step, either through the solution of an auxiliary partial differential equation (PDE) or the simulation of an auxiliary stochastic differential equation.

A systematic approach to obtaining martingale representation formulas has been proposed in [10], using the functional Itô calculus [20, 8, 9, 7] (theorem 2.10), for any square-integrable $\mathbb{F}$-martingale Y,

$$\forall t \in [0,T], \quad Y(t) = Y(0) + \int_0^t \nabla_W Y(s) \cdot dW(s) \quad \mathbb{P}\text{−a.s.}$$

where $\nabla_W Y$ is the weak vertical derivative of Y with respect to W, constructed as limit in $L^2([0,T] \times \Omega)$ of pathwise directional derivatives.

More precisely, let $\mathcal{M}^2$ be the space of $\mathbb{R}$-valued square-integrable $\mathbb{F}$-martingales with initial value zero equipped with the norm $\|Y\|^2_{\mathcal{M}^2} := \mathbb{E}[Y(T)^2]$. For any $Y \in \mathcal{M}^2$, as $D(W) := \mathcal{C}^{1,2}_{loc}(W) \cap \mathcal{M}^2$ is dense in $\mathcal{M}^2$ with respect to the norm $\|\cdot\|_{\mathcal{M}^2}$ (lemma 2.8), there exists a sequence $(Y_n)_{n\geq 1}$ of elements in $D(W)$ such that:

$$\|Y_n - Y\|_{\mathcal{M}^2} \underset{n\to\infty}{\to} 0$$

and

$$\|\nabla_W Y_n - \nabla_W Y\|_{\mathcal{L}^2(W)} \underset{n\to\infty}{\to} 0$$

where $\mathcal{L}^2(W)$ is the Hilbert space of $\mathbb{F}$-predictable processes such that:

$$\|\phi\|^2_{\mathcal{L}^2(W)} := \mathbb{E}\left[\int_0^T \operatorname{tr}(\phi(t)\,{}^t\phi(t))dt\right] < \infty.$$

Here contrary to $\phi = \nabla_W Y$ which is a weak vertical derivative, $\phi_n := \nabla_W Y_n$ is a classical pathwise vertical derivative:

$$\nabla_W Y_n(t) = \nabla_\omega F_n(t, W_t) = (\partial_i F_n(t, W_t), i = 1, \cdots, d) \in \mathbb{R}^d$$

with

$$\partial_i F_n(t, W_t) = \lim_{h\to 0} \frac{F_n(t, W_t + he_i \mathbb{1}_{[t,T]}) - F_n(t, W_t)}{h}$$

for any $F_n \in \mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$ such that $Y_n(t) = F_n(t, W_t) \quad dt \times d\mathbb{P}$-a.e. (definition 2.7). Hence ϕ_n is defined as a pathwise limit of finite-difference approximations, and thus readily computable path-by-path in a simulation setting, which makes it a natural explicit approximation of the integrand ϕ in the martingale representation (3.1).

This approach does not rely on any Markov property nor on the Gaussian structure of the Wiener space and is applicable to functionals of a large class of Itô processes. However, lemma 2.8 only ensures the existence of $(Y_n)_{n\geq 1}$ in $D(W)$ (or equivalently $(F_n)_{n\geq 1}$ in $\mathbb{C}^{1,2}_{loc}(\Lambda^d_T)$), it does not provide an explicit construction of $(Y_n)_{n\geq 1}$ (or $(F_n)_{n\geq 1}$) which approximates the initial martingale Y in $\|\cdot\|_{\mathcal{M}^2}$.

In this chapter, we propose a systematic method for constructing such a sequence of 'smooth' martingales $(Y_n)_{n\geq 1}$ in a general setting in which H is allowed to be a functional of the solution of a stochastic differential equation (SDE) with path-dependent coefficients:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t) \quad X(0) = x_0 \in \mathbb{R}^d \tag{3.2}$$

where $X_t = X(t \wedge \cdot)$ designates the path stopped at t and

$$b : \Lambda^d_T \to \mathbb{R}^d, \quad \sigma : \Lambda^d_T \to \mathbb{M}_d(\mathbb{R})$$

are continuous non-anticipative functionals.

For any square-integrable $\mathcal{F}_T$-measurable variable H of the form $H = g(X(t), 0 \le t \le T) = g(X_T)$ where $g : (D([0,T], \mathbb{R}^d), \|\cdot\|_\infty) \to \mathbb{R}$ is a continuous functional, we construct an explicit sequence of approximations ϕ_n for the integrand ϕ in (3.1). These approximations are constructed as vertical derivatives, in the sense of definition 1.10, of the weak Euler approximation F_n of the martingale Y, obtained by replacing X by the corresponding Euler scheme ${}_nX$:

$$\phi_n(t) = \nabla_\omega F_n(t, W), \quad \text{with} \quad F_n(t,\omega) := \mathbb{E}\left[g({}_nX(\omega \underset{t}{\oplus} W))\right] \tag{3.3}$$

where $\underset{t}{\oplus}$ is the concatenation of paths at t.

The main results of this chapter are the following. We first show that the functional F_n defined by (3.3) is horizontally differentiable and infinitely vertically differentiable (theorem 3.7), and F_n and its derivatives satisfy the necessary regularity conditions for applying the functional Itô formula (theorem 3.8), namely $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda^d_T)$ (definition 2.3). As $Y_n(t) := F_n(t, W_t)$ is a $\mathbb{F}$-martingale, i.e. $Y_n \in D(W)$, we establish the convergence of the approximations ϕ_n to the integrand ϕ in (3.1) in proposition 3.3. Under a Lipschitz assumption on g, we provide in theorem 3.10 an L^p error estimate for the approximation error. The proposed approximations are easy to compute and readily integrated in commonly used numerical schemes for SDEs. Some numerical aspects of this approximation method are discussed in section 3.6. And finally in section 3.7, we apply this method to the hedging problem of exotic options in finance.

Our approach requires neither the Markov property of the underlying processes nor the differentiability of coefficients b and σ of the path-dependent SDE (3.2), and is thus applicable to functionals of a large class of semimartingales. By contrast to methods based on Malliavin calculus [6, 36, 2, 54, 43, 29], it does not require Malliavin differentiability of the terminal variable H nor does it involve any choice of 'Malliavin weights', a delicate step in these methods.

Ideas based on the functional Itô calculus have also been recently used by Leão and Ohashi [47] for weak approximation of Wiener functionals, using a space-filtration discretization scheme. The approach of [47] essentially reduces to approximating the process using a (random-step) binomial tree and computing the estimators using the underlying tree. Unlike the approach proposed in [47], our approach is based on a Euler approximation on a fixed time grid, rather than a random time grid used in [47], which involves a sequence of first passage times. Our approach is thus much easier to implement and analyze and is readily integrated in commonly used numerical schemes for approximations of SDEs, which are typically based on fixed time grids.

3.2 Euler approximations for path-dependent SDEs

Let W be a standard d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ its ($\mathbb{P}$-completed) natural filtration. We consider the following stochastic differential equation with path-dependent coefficients:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d \tag{3.4}$$

where $b : \Lambda_T^d \to \mathbb{R}^d$ and $\sigma : \Lambda_T^d \to \mathbb{M}_d(\mathbb{R})$ are non-anticipative maps, assumed to be Lipschitz continuous with respect to the following distance d defined on Λ_T^d:

$$d((t,\omega),(t',\omega')) = \sup_{s\in[0,T]} |\omega(s\wedge t) - \omega'(s\wedge t')| + \sqrt{|t-t'|} = \|\omega_t - \omega'_{t'}\|_\infty + \sqrt{|t-t'|}.$$

Assumption 3.1. *We assume that $b : (\Lambda_T^d, d) \to \mathbb{R}^d$, $\sigma : (\Lambda_T^d, d) \to \mathbb{M}_d(\mathbb{R})$ are Lipschitz continuous:*

$$\exists K_{Lip} > 0, \quad \forall t, t' \in [0,T], \forall \omega, \omega' \in D([0,T], \mathbb{R}^d),$$

$$|b(t,\omega) - b(t',\omega')| + \|\sigma(t,\omega) - \sigma(t',\omega')\| \le K_{Lip}\, d((t,\omega),(t',\omega')).$$

Remark 3.1. *This Lipschitz condition with respect to the distance d is weaker than a Lipschitz condition with respect to the distance d_∞ as in assumption 2.3: it allows for a Hölder smoothness of degree $1/2$ in the t variable.*

Under assumption 3.1, (3.4) has a unique strong $\mathbb{F}$-adapted solution X.

Proposition 3.1. *Under assumption 3.1, there exists a unique $\mathbb{F}$-adapted process X satisfying (3.4). Moreover for $p \geq 1$, we have:*

$$\mathbb{E}\left[\|X_T\|_\infty^{2p}\right] \leq C(1+|x_0|^{2p})e^{CT}$$

for some constant $C = C(p, T, K_{Lip})$ depending on p, T and K_{Lip}.

The proof of this proposition is exactly the same as that of proposition 2.3. Again the assumption 3.1 might seem to be strong, and it can be replaced by the weaker condition given in remark 2.11. However assumption 3.1 is necessary for the convergence of the Euler approximation described later in this section, especially the results concerning its rate of convergence.

In the following, we always assume that assumption 3.1 holds. The strong solution X of equation (3.4) is then a semimartingale and defines a non-anticipative functional $X : \mathcal{W}_T^d \to \mathbb{R}^d$ given by the Itô map associated to (3.4).

3.2.1 Euler approximations as non-anticipative functionals

We now consider an Euler approximation for the SDE (3.4) and study its properties as a non-anticipative functional. Let $n \in \mathbb{N}$ and $\delta = \frac{T}{n}$. The Euler approximation ${}_nX$ of X on the grid $(t_j = j\delta, j = 0, \cdots, n)$ is defined as follows:

Definition 3.2 (Euler scheme)**.** *For $\omega \in D([0,T], \mathbb{R}^d)$, denote by ${}_nX(\omega) \in D([0,T], \mathbb{R}^d)$ the piecewise constant Euler approximation for (3.4) computed along the path ω, defined as follows: ${}_nX(\omega)$ is constant in each interval*

$[t_j, t_{j+1})$ *for any* $0 \le j \le n-1$ *with* ${}_nX(0,\omega) = x_0$, *and is defined recursively by: for* $0 \le j \le n-1$,

$$ {}_nX(t_{j+1},\omega) = {}_nX(t_j,\omega)+b(t_j, {}_nX_{t_j}(\omega))\delta+\sigma(t_j, {}_nX_{t_j}(\omega))(\omega(t_{j+1}-)-\omega(t_j-)), \tag{3.5}$$

where ${}_nX_t(\omega) = {}_nX(t\wedge\cdot,\omega)$ *the path of* ${}_nX$ *stopped at time* t *computed along* ω, *and by convention* $\omega(0-) = \omega(0)$.

When computed along the path of the Brownian motion W, ${}_nX(W)$ is simply the piecewise constant Euler-Maruyama scheme [57] for the stochastic differential equation (3.4).

By definition, the path ${}_nX(\omega)$ depends only on a finite number of increments of ω: $\omega(t_1-) - \omega(0), \cdots, \omega(t_n-) - \omega(t_{n-1}-)$. We can thus define a map:

$$p_n : \mathbb{M}_{d,n}(\mathbb{R}) \to D([0,T], \mathbb{R}^d)$$

such that for $\omega \in D([0,T]), \mathbb{R}^d)$,

$$ {}_nX(\omega) = p_n(\omega(t_1-)-\omega(0), \omega(t_2-)-\omega(t_1-), \cdots, \omega(t_n-)-\omega(t_{n-1}-)). \tag{3.6}$$

Clearly the map p_n only depends on the initial condition x_0 and the coefficients b and σ of the SDE (3.4). By a slight abuse of notation, in what follows, we denote by $p_t(y)$ the path $p_n(y)$ stopped at t for $y \in \mathbb{M}_{d,n}(\mathbb{R})$.

The map $p_n : \mathbb{M}_{d,n}(\mathbb{R}) \to (D([0,T], \mathbb{R}^d), \|\cdot\|_\infty)$ is then locally Lipschitz continuous, as shown by the following lemma.

Lemma 3.3. *For every* $\eta > 0$, *there exists a constant* $C(\eta, K_{Lip}, T)$ *such that for any* $y = (y_1, \cdots, y_n)$, $y' = (y'_1, \cdots, y'_n) \in \mathbb{M}_{d,n}(\mathbb{R})$,

$$\max_{1\le k\le n} |y_k| \vee |y'_k| \le \eta \quad \Longrightarrow \quad \|p_n(y) - p_n(y')\|_\infty \le C(\eta, K_{Lip}, T) \max_{1\le k\le n} |y_k - y'_k|.$$

Proof. As the two paths $p_n(y)$ and $p_n(y')$ are both piecewise constant by construction, it suffices to prove the inequality at times $(t_j)_{0\le j\le n}$. We prove by induction that, for any $0 \le j \le n$,

$$\|p_{t_j}(y) - p_{t_j}(y')\|_\infty \le C(\eta, K_{Lip}, T) \max_{1\le k\le j} |y_k - y'_k| \tag{3.7}$$

with some constant C which only depends on η, K_{Lip}, T (and n).

For $j = 0$, this is clearly the case as $p(y)(0) = p(y')(0) = x_0$. Assume that (3.7) is verified for some $0 \leq j \leq n-1$, consider now $\|p_{t_{j+1}}(y) - p_{t_{j+1}}(y')\|_\infty$, by definition of the map p_n, we have:

$$p_n(y)(t_{j+1}) = p_n(y)(t_j) + b(t_j, p_{t_j}(y))\delta + \sigma(t_j, p_{t_j}(y))y_{j+1}$$

and

$$p_n(y')(t_{j+1}) = p_n(y')(t_j) + b(t_j, p_{t_j}(y'))\delta + \sigma(t_j, p_{t_j}(y'))y'_{j+1}.$$

Thus

$$\begin{aligned}
& |p_n(y)(t_{j+1}) - p_n(y')(t_{j+1})| \\
\leq\ & |p_n(y)(t_j) - p_n(y')(t_j)| + |b(t_j, p_{t_j}(y)) - b(t_j, p_{t_j}(y'))|\delta \\
& + \|\sigma(t_j, p_{t_j}(y))\| \cdot |y_{j+1} - y'_{j+1}| + \|\sigma(t_j, p_{t_j}(y)) - \sigma(t_j, p_{t_j}(y'))\| \cdot |y'_{j+1}| \\
\leq\ & C(\eta, K_{Lip}, T) \max_{1\leq k\leq j} |y_k - y'_k| + K_{Lip} C(\eta, K_{Lip}, T) \max_{1\leq k\leq j} |y_k - y'_k|\delta \\
& + \left(\|\sigma(0, \bar{0})\| + K_{Lip}(\sqrt{t_j} + \|p_{t_j}(y))\|_\infty)\right) |y_{j+1} - y'_{j+1}| \\
& + K_{Lip} C(\eta, K_{Lip}, T)\eta \max_{1\leq k\leq j} |y_k - y'_k| \\
\leq\ & C(\eta, K_{Lip}, T) \max_{1\leq k\leq j+1} |y_k - y'_k|
\end{aligned}$$

(The constant C may differ from one line to another).

And consequently we have:

$$\|p_{t_{j+1}}(y) - p_{t_{j+1}}(y')\|_\infty \leq C(\eta, K_{Lip}, T) \max_{1\leq k\leq j+1} |y_k - y'_k|$$

for some different constant C depending only on η, K_{Lip} and T (and n). And we conclude by induction. □

3.2.2 Strong convergence

To simplify the notations, ${}_nX_T(W_T)$ will be simply noted as ${}_nX_T$ in the following. The following result, which gives a uniform estimate of the discretization error, $X_T - {}_nX_T$ extends similar results known in the Markovian case [57, 25, 37] to the path-dependent SDEs (3.4) (see also [34]):

Proposition 3.2. *Under assumption 3.1, we have the following estimate in L^{2p}-norm for the strong error of the piecewise constant Euler-Maruyama scheme:*

$$\mathbb{E}\left(\sup_{s\in[0,T]} \|X(s) - {}_nX(s)\|^{2p}\right) \leq C(x_0, p, T, K_{Lip}) \left(\frac{1+\log n}{n}\right)^p, \quad \forall p \geq 1$$

with C a constant depending only on x_0, p, T and K_{Lip}.

Proof. The idea is to construct a 'Brownian interpolation' ${}_n\hat{X}_T$ of ${}_nX_T$:

$${}_n\hat{X}(s) = x_0 + \int_0^s b\left(\underline{u}, {}_nX_{\underline{u}}\right) du + \int_0^s \sigma\left(\underline{u}, {}_nX_{\underline{u}}\right) dW(u)$$

where $\underline{u} = \left\lfloor \frac{u}{\delta} \right\rfloor \cdot \delta$ is the largest subdivision point which is smaller or equal to u.

Clearly ${}_n\hat{X}$ is a continuous semimartingale and $\| \sup_{s\in[0,T]} |X(s) - {}_nX(s)|\|_{2p}$ can be controlled by the sum of the two following terms:

$$\| \sup_{s\in[0,T]} |X(s) - {}_nX(s)|\|_{2p} \leq \| \sup_{s\in[0,T]} |X(s) - {}_n\hat{X}(s)|\|_{2p} + \| \sup_{s\in[0,T]} |{}_n\hat{X}(s) - {}_nX(s)|\|_{2p} \tag{3.8}$$

We start with the first term of (3.8) $\| \sup_{s\in[0,T]} |X(s) - {}_n\hat{X}(s)|\|_{2p}$. Using the Burkholder-Davis-Gundy inequality and Hölder's inequality, we have:

$$\begin{aligned}
\mathbb{E}\|X_T - {}_n\hat{X}_T\|_\infty^{2p} &= \mathbb{E}\left\|\int_0^{\cdot} \big(b(s,X_s) - b(\underline{s},{}_nX_{\underline{s}})\big)\,ds + \int_0^{\cdot} \big(\sigma(s,X_s) - \sigma(\underline{s},{}_nX_{\underline{s}})\big)\,dW(s)\right\|_\infty^{2p} \\
&\leq C(p)\Bigg(\mathbb{E}\left\|\int_0^{\cdot} \big(b(s,X_s) - b(\underline{s},{}_nX_{\underline{s}})\big)\,ds\right\|_\infty^{2p} \\
&\quad + \mathbb{E}\left\|\int_0^{\cdot} \big(\sigma(s,X_s) - \sigma(\underline{s},{}_nX_{\underline{s}})\big)\,dW(s)\right\|_\infty^{2p}\Bigg) \\
&\leq C(p)\Bigg(\mathbb{E}\left[\int_0^T |b(s,X_s) - b(\underline{s},{}_nX_{\underline{s}})|ds\right]^{2p} \\
&\quad + \mathbb{E}\left[\int_0^T \left\|\sigma(s,X_s) - \sigma(\underline{s},{}_nX_{\underline{s}})\right\|^2 ds\right]^p\Bigg) \\
&\leq C(p,T)\Bigg(\mathbb{E}\left[\int_0^T |b(s,X_s) - b(\underline{s},{}_nX_{\underline{s}})|^{2p}ds\right] \\
&\quad + \mathbb{E}\left[\int_0^T \left\|\sigma(s,X_s) - \sigma(\underline{s},{}_nX_{\underline{s}})\right\|^{2p} ds\right]\Bigg) \\
&\leq C(p,T,K_{Lip})\,\mathbb{E}\left[\int_0^T \big((s-\underline{s})^p + \|X_s - {}_nX_s\|_\infty^{2p}\big)\,ds\right] \\
&\leq C(p,T,K_{Lip})\left(\frac{1}{n^p} + \int_0^T \mathbb{E}\,\|X_s - {}_nX_s\|_\infty^{2p}\,ds\right)
\end{aligned}$$

We have used the fact that ${}_nX_{\underline{s}} = {}_nX_s$ since ${}_nX$ is piecewise constant by definition.

Consider now the second term of (3.8) $\|\sup_{s\in[0,T]} |{}_n\hat{X}(s) - {}_nX(s)|\|_{2p}$. First we observe that:

$${}_n\hat{X}(s) - {}_nX(s) = {}_n\hat{X}(s) - {}_n\hat{X}(\underline{s}) = b\left(\underline{s},{}_nX_{\underline{s}}\right)(s-\underline{s}) + \sigma\left(\underline{s},{}_nX_{\underline{s}}\right)(W(s) - W(\underline{s})),$$

thus we have:

$$\begin{aligned}
\|{}_n\hat{X}_T - {}_nX_T\|_\infty &\leq \sup_{0\leq s\leq T}\big(|b(\underline{s},{}_nX_{\underline{s}})| + |\sigma(\underline{s},{}_nX_{\underline{s}})|\big)\left(\frac{1}{n} + \sup_{s\in[0,T]}|W(s) - W(\underline{s})|\right) \\
&\leq C(K_{Lip},T)(1 + \|{}_nX_T\|_\infty)\left(\frac{1}{n} + \sup_{s\in[0,T]}|W(s) - W(\underline{s})|\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\|{}_n\hat{X}_T - {}_nX_T\|_\infty^{2p} &\leq C(p, K_{Lip}, T)\frac{1}{n^{2p}}\,\mathbb{E}\left[(1+\|{}_nX_T\|_\infty)^{2p}\right] \\
&\quad + C(p, K_{Lip}, T)\,\mathbb{E}\left[(1+\|{}_nX_T\|_\infty)^{2p}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{2p}\right] \\
&\leq C(p, K_{Lip}, T)\frac{1}{n^{2p}}\left(1+\mathbb{E}\|{}_nX_T\|_\infty^{2p}\right) \\
&\quad + C(p, K_{Lip}, T)\,\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{2p} \\
&\quad + C(p, K_{Lip}, T)\,\mathbb{E}\left[\|{}_nX_T\|_\infty^{2p}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{2p}\right]
\end{aligned}$$

The Cauchy-Schwarz inequality yields:

$$\begin{aligned}
\mathbb{E}\|{}_n\hat{X}_T - {}_nX_T\|_\infty^{2p} &\leq C(p, K_{Lip}, T)\frac{1}{n^{2p}}\left(1+\sqrt{\mathbb{E}\|{}_nX_T\|_\infty^{4p}}\right) \\
&\quad + C(p, K_{Lip}, T)\sqrt{\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{4p}} \\
&\quad + C(p, K_{Lip}, T)\sqrt{\mathbb{E}\|{}_nX_T\|_\infty^{4p}}\cdot\sqrt{\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{4p}} \\
&\leq C(p, K_{Lip}, T)\left(1+\sqrt{\mathbb{E}\|{}_nX_T\|_\infty^{4p}}\right) \\
&\quad \cdot\left(\frac{1}{n^{2p}}+\sqrt{\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{4p}}\right) \qquad (3.9)
\end{aligned}$$

We now show that $\mathbb{E}\|{}_nX_T\|_\infty^{4p}$ can be bounded independently of n. Using again the Burkholder-Davis-Gundy inequality and Hölder's inequality, we obtain:

$$\begin{aligned}
\mathbb{E}\|{}_nX_T\|_\infty^{4p} &\leq \mathbb{E}\|{}_n\hat{X}_T\|_\infty^{4p} \\
&\leq C(p)\left(x_0^{4p}+\mathbb{E}\left(\int_0^T|b(\underline{s},{}_nX_{\underline{s}})|ds\right)^{4p}+\mathbb{E}\left(\int_0^T\left\|\sigma(\underline{s},{}_nX_{\underline{s}})\right\|^2ds\right)^{2p}\right) \\
&\leq C(x_0,p,T)\left(1+\int_0^T\left(\mathbb{E}|b(\underline{s},{}_nX_{\underline{s}})|^{4p}+\mathbb{E}\left\|\sigma(\underline{s},{}_nX_{\underline{s}})\right\|^{4p}\right)ds\right) \\
&\leq C(x_0,p,T,K_{Lip})\left(1+\int_0^T\mathbb{E}\|{}_nX_s\|_\infty^{4p}\,ds\right)
\end{aligned}$$

And we deduce from Gronwall's inequality that $\mathbb{E}\|_n X_T\|_\infty^{4p}$ is bounded by a constant which depends only on x_0, p, T and K_{Lip}.

What remains is to control the term $\sqrt{\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{4p}}$ in (3.9). For this, we make use of the following result [55, p.209]:

$$\forall p>0, \quad \|\sup_{0\le s\le T}|W(s)-W(\underline{s})|\|_p \le C(W,p)\sqrt{\frac{T}{n}(1+\log n)}$$

where $C(W,p)$ is a constant which only depends on p and W. This result is a consequence of the following lemma [55, Lemma 7.1]:

Lemma 3.4. *Let $Y_1, \cdots, Y_n$ be non-negative random variables with the same distribution satisfying $\mathbb{E}\left(e^{\lambda Y_1}\right)<\infty$ for some $\lambda>0$. Then we have:*

$$\forall p>0, \quad \|\max(Y_1,\cdots,Y_n)\|_p \le \frac{1}{\lambda}(\log n + C(p,Y_1,\lambda))$$

for some constant $C(p,Y_1,\lambda)$ which only depends on p, Y_1 and λ.

We have thus:

$$\sqrt{\mathbb{E}\sup_{s\in[0,T]}|W(s)-W(\underline{s})|^{4p}} \le C(p,T)\left(\frac{1+\log n}{n}\right)^p.$$

And inequality (3.9) can now be written as:

$$\mathbb{E}\|_n\hat{X}_T - {}_nX_T\|_\infty^{2p} \le C(x_0,p,T,K_{Lip})\left(\frac{1+\log n}{n}\right)^p.$$

Finally (3.8) becomes:

$$\begin{aligned}
& \mathbb{E}\|X_T - {}_nX_T\|_\infty^{2p} \\
\le\ & C(p)\left(\mathbb{E}\|X_T - {}_n\hat{X}_T\|_\infty^{2p} + \mathbb{E}\|_n\hat{X}_T - {}_nX_T\|_\infty^{2p}\right) \\
\le\ & C(x_0,p,T,K_{Lip})\left(\left(\frac{1+\log n}{n}\right)^p + \int_0^T \mathbb{E}\|X_s - {}_nX_s\|_\infty^{2p} ds\right).
\end{aligned}$$

And we conclude by Gronwall's inequality. □

We may also deduce from proposition 3.2 an almost-sure rate of convergence.

Corollary 3.1. *Under assumption 3.1, we have the following almost-sure rate of convergence of the piecewise constant Euler-Maruyama scheme:*

$$\forall \alpha \in [0, \frac{1}{2}), \quad n^{\alpha} \| X_T - {}_n X_T \|_{\infty} \underset{n \to \infty}{\to} 0, \quad \mathbb{P}\text{−a.s.}$$

Proof. Let $\alpha \in [0, \frac{1}{2})$. For a p large enough, by proposition 3.2, we have:

$$\mathbb{E}\left[\sum_{n \geq 1} n^{2p\alpha} \| X_T - {}_n X_T \|_{\infty}^{2p} \right] < \infty.$$

Thus

$$\sum_{n \geq 1} n^{2p\alpha} \| X_T - {}_n X_T \|_{\infty}^{2p} < \infty, \quad \mathbb{P}\text{−a.s.}$$

and

$$n^{\alpha} \| X_T - {}_n X_T \|_{\infty} \underset{n \to \infty}{\to} 0, \quad \mathbb{P}\text{−a.s.}$$

□

3.3 Smooth functional approximations for martingales

Now return to the initial setting of our problem. We have a continuous semimartingale X which is the (strong) solution of the following path-dependent PDE:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d.$$

We have a functional $g : D([0,T], \mathbb{R}^d) \to \mathbb{R}$, and we would like to find an approximation of the martingale representation of the $\mathbb{F}_T$-measurable variable $g(X_T)$ or equivalently the martingale $Y(t) := \mathbb{E}[g(X_T)|\mathcal{F}_t]$.

First we impose a condition on the functional g to ensure the integrability of $g(X_T)$.

Assumption 3.2. *We assume that the functional* $g : (D([0,T], \mathbb{R}^d), \|\cdot\|_{\infty}) \to \mathbb{R}$ *is continuous with polynomial growth:*

$$\exists q \in \mathbb{N}, \exists C > 0, \forall \omega \in D([0,T], \mathbb{R}^d), \quad |g(\omega)| \leq C \left(1 + \|\omega\|_{\infty}^{q}\right).$$

Under this assumption, $g(X_T)$ (or the martingale Y) is square-integrable by proposition 3.1. This is still true if we replace X_T with its piecewise constant Euler-Maruyama scheme ${}_nX_T$ as we have shown in the proof of proposition 3.2 that $\mathbb{E}\|{}_nX_T\|_\infty^{4p}$ is bounded independently of n.

The (square-integrable) martingale Y may be represented as a non-anticipative functional of W:

$$Y(t) = F(t, W_t)$$

where the functional F is square-integrable but fails to be smooth a priori (see [65] for conditions under which Y admits a smooth functional representation of W). By theorem 2.10, we have:

$$g(X_T) = Y(T) = Y(0) + \int_0^T \nabla_W Y(s) \cdot dW(s) \quad \mathbb{P}-\text{a.s.}$$

where $\nabla_W Y$ is the weak vertical derivative of Y with respect to W (theorem 2.9), which cannot be computed directly as a pathwise directional derivative unless F is a smooth functional (for example $\in \mathbb{C}^{1,2}_{loc,r}(\Lambda_T^d)$).

The main idea is to approximate the martingale Y by a sequence of smooth martingales Y_n which, contrary to Y, admit a smooth functional representation $Y_n(s) = F_n(s, W_s)$ with $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda_T^d)$. Then by the functional Itô formula, we have:

$$\begin{aligned}\int_0^T \nabla_\omega F_n(s, W_s) \cdot dW(s) \quad &= \quad Y_n(T) - Y_n(0) \\ &\underset{n\to\infty}{\to} \quad Y(T) - Y(0) = \int_0^T \nabla_W Y(s) \cdot dW(s).\end{aligned}$$

We recall that the existence of such sequence $(Y_n)_{n\geq 1}$ is ensured by lemma 2.8. Here we focus on an explicit construction of $(Y_n)_{n\geq 1}$ (or $(F_n)_{n\geq 1}$). Once $(F_n)_{n\geq 1}$ is given, we then obtain the following estimator for $\nabla_W Y$:

$$Z_n(s) = \nabla_\omega F_n(s, W_s),$$

where the vertical derivative $\nabla_\omega F_n(s, W_s) = (\partial_i F_n(s, W_s), 1 \leq i \leq d)$ may be computed as a pathwise directional derivative

$$\partial_i F_n(s, W_s) = \lim_{h\to 0} \frac{F_n(s, W_s + he_i \mathbb{1}_{[s,T]}) - F_n(s, W_s)}{h},$$

yielding a concrete procedure for computing the estimator.

We will show in the following that the weak Euler approximation (definition 3.2) provides a systematic way of constructing such smooth functional approximations in the sense of definition 2.3.

Remark 3.5. *If we assume that the local martingale component of the process X is non-degenerate, i.e.* $\det(\sigma(t, X_t)) \neq 0$, *$dt \times d\mathbb{P}$-a.e. then X also admits the predictable representation property [42, 64]: $g(X_T) = Y(T)$ can also be represented as an integral with respect to the local martingale component M of X, which is an easy extension of theorem 2.10 to the case where X is a semimartingale:*

$$g(X_T) = Y(T) = Y(0) + \int_0^T \nabla_X Y(s) \cdot dM(s) \quad \mathbb{P}\text{−a.s.}$$

We can also attempt to construct a sequence of smooth martingales Y_n which admit a smooth functional representation with respect to X: $Y_n(s) = F_n(s, X_s)$. In some applications, this approach might seem to be more natural. For example, in mathematical finance, if X represents the price of an underlying asset and g is the payoff of some path-dependent option, then $\nabla_X Y$ represents the sensitivity of the option price with respect to the underlying, commonly called the delta of the option.

However in this case, it is more difficult to construct such smooth functionals F_n, which ofter requires some differentiability condition on the coefficients b and σ of the SDE (3.4) *satisfied by X. Indeed, as we shall see in the proof of theorem 3.7, in our approach, the smoothness of the functionals F_n relies on the smoothness of the density of increments of W. If we now work with functionals of X, as the density of increments of X is not necessarily smooth, (unless with some additional differentiability condition on b and σ), the smoothness of F_n might be problematic in this case.*

Now we define explicitly the functional F_n. Define first the concatenation of two càdlàg paths $\omega, \omega' \in D([0,T], \mathbb{R}^d)$ at time $s \in [0,T]$, which we note

$\omega \underset{s}{\oplus} \omega'$, as the following càdlàg path on $[0,T]$:

$$\omega \underset{s}{\oplus} \omega' = \omega_s \underset{s}{\oplus} \omega' = \begin{cases} \omega(u) & u \in [0,s) \\ \omega(s) + \omega'(u) - \omega'(s) & u \in [s,T] \end{cases}.$$

Observe that if $\omega, \omega' \in C([0,T], \mathbb{R}^d)$, then $\omega \underset{s}{\oplus} \omega'$ is also continuous. And for any $z \in \mathbb{R}^d$, we have:

$$\omega_s^z \underset{s}{\oplus} \omega' = (\omega_s \underset{s}{\oplus} \omega') + z\mathbb{1}_{[s,T]}.$$

Definition 3.6 (Weak Euler approximation)**.** *We define the (level-n) weak Euler approximation of F as the functional F_n by, for any $\omega \in D([0,T], \mathbb{R}^d)$,*

$$F_n(s, \omega_s) := \mathbb{E}\left[g\left({}_nX(\omega_s \underset{s}{\oplus} B_T) \right) \right] \tag{3.10}$$

where B is a Wiener process independent of W.

Applying this functional to the path of the Wiener process W, we obtain a $\mathbb{F}$-adapted process:

$$Y_n(s) := F_n(s, W_s).$$

Using independence of increments of W, we have:

$$\begin{aligned} Y_n(s) &= \mathbb{E}\left[g({}_nX(\omega_s \underset{s}{\oplus} B_T)) \right] |_{\omega_s = W_s} = \mathbb{E}\left[g({}_nX(W_s \underset{s}{\oplus} B_T)) | \mathcal{F}_s \right] \\ &= \mathbb{E}\left[g({}_nX(W_s \underset{s}{\oplus} W_T)) | \mathcal{F}_s \right] = \mathbb{E}\left[g({}_nX(W_T)) | \mathcal{F}_s \right] \end{aligned}$$

In particular Y_n is a square-integrable martingale, so is weakly differentiable in the sense of theorem 2.9. We will now show that F_n is in fact a smooth functional in the sense of definition 2.3, i.e. $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda_T^d)$.

Theorem 3.7. *Under Assumptions 3.1 and 3.2, the functional F_n defined in* (3.10) *is horizontally differentiable and infinitely vertically differentiable.*

Proof. First, note that under assumption 3.1, ${}_nX(\omega)$ is bounded by a polynomial in the variables $\omega(t_1-) - \omega(0), \omega(t_2-) - \omega(t_1-), \cdots, \omega(t_n-) - \omega(t_{n-1}-))$.

Combined with assumption 3.2, this implies that all expectations in the proof of this theorem are well defined.

Let $(s,\omega) \in \Lambda_T^d$ with $t_k \leq s < t_{k+1}$ for some $0 \leq k \leq n-1$. We start with the vertical differentiability of F_n at (s,ω), which is equivalent to the differentiability at 0 of the map $v : \mathbb{R}^d \to \mathbb{R}$ defined by:

$$v(z) := F_n(s,\omega_s^z) = \mathbb{E}\left[g({}_nX(\omega_s^z \underset{s}{\oplus} B_T))\right], \quad z \in \mathbb{R}^d.$$

The main idea of the proof is to absorb the dependence of the map v on z in the Gaussian density function when taking the expectation, which then implies smoothness in z since the Gaussian density function is infinitely differentiable.

As we have already shown, ${}_nX(\omega_s^z \underset{s}{\oplus} B_T)$ depends only on $(\omega_s^z \underset{s}{\oplus} B_T)(t_1-) - (\omega_s^z \underset{s}{\oplus} B_T)(0), \cdots, (\omega_s^z \underset{s}{\oplus} B_T)(t_n-) - (\omega_s^z \underset{s}{\oplus} B_T)(t_{n-1}-)$, which are all explicit using the definition of the concatenation. For $j < k$, we have:

$$(\omega_s^z \underset{s}{\oplus} B_T)(t_{j+1}-) - (\omega_s^z \underset{s}{\oplus} B_T)(t_j-) = \omega(t_{j+1}-) - \omega(t_j-).$$

In the case where $j = k$, we have:

$$\begin{aligned} & (\omega_s^z \underset{s}{\oplus} B_T)(t_{k+1}-) - (\omega_s^z \underset{s}{\oplus} B_T)(t_k-) \\ = \;& B(t_{k+1}) - B(s) + \omega(s) + z - \omega(t_k-) \\ = \;& B(t_{k+1}) - B(s) + z + \omega(s) - \omega(t_k-). \end{aligned}$$

And for $j > k$, we have:

$$\begin{aligned} & (\omega_s^z \underset{s}{\oplus} B_T)(t_{j+1}-) - (\omega_s^z \underset{s}{\oplus} B_T)(t_j-) \\ = \;& B(t_{j+1}) - B(s) + \omega(s) + z - (B(t_j) - B(s) + \omega(s) + z) \\ = \;& B(t_{j+1}) - B(t_j). \end{aligned}$$

Thus we have:

$$\begin{aligned} {}_nX(\omega_s^z \underset{s}{\oplus} B_T) = p_n \Big(\; & \omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), \\ & B(t_{k+1}) - B(s) + z + \omega(s) - \omega(t_k-), \\ & B(t_{k+2}) - B(t_{k+1}), \cdots, B(t_n) - B(t_{n-1}) \Big) \end{aligned}$$

where $p_n : \mathbb{M}_{d,n}(\mathbb{R}) \to D([0,T], \mathbb{R}^d)$ is the map defined in (3.6).

Observe from the previous expression of ${}_nX(\omega_s^z \underset{s}{\oplus} B_T)$ that, for a fixed z and s, the randomness of the piecewise constant stochastic process ${}_nX(\omega_s^z \underset{s}{\oplus} B_T)$ only comes from a finite number of Gaussian variables: $B(t_{k+1}) - B(s)$, $B(t_{k+2}) - B(t_{k+1})$, $\cdots$, $B(t_n) - B(t_{n-1})$. Since the joint distribution of these Gaussian variables is explicit, $v(z) = \mathbb{E}\left[g({}_nX_T(\omega_s^z \underset{s}{\oplus} B_T))\right]$ can be computed explicitly as an integral in finite dimension.

Let $y = (y_1, \cdots, y_{n-k}) \in \mathbb{M}_{d,n-k}(\mathbb{R})$.

$$
\begin{aligned}
v(z) &= \mathbb{E}\left[g({}_nX_T(\omega_s^z \underset{s}{\oplus} B_T))\right] \\
&= \mathbb{E}\Big[g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), \\
&\quad B(t_{k+1}) - B(s) + z + \omega(s) - \omega(t_k-), B(t_{k+2}) - B(t_{k+1}), \cdots, B(t_n) - B(t_{n-1})\Big)\Big] \\
&= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + z + \omega(s) - \omega(t_k-), \\
&\quad y_2, \cdots, y_{n-k})\Big) \Phi(y_1, t_{k+1} - s) \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy_1 dy_2 \cdots dy_{n-k} \\
&= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + \omega(s) - \omega(t_k-), \\
&\quad y_2, \cdots, y_{n-k})\Big) \Phi(y_1 - z, t_{k+1} - s) \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy_1 dy_2 \cdots dy_{n-k} \qquad (3.11)
\end{aligned}
$$

with $\delta = \frac{T}{n}$ and

$$\Phi(x,t) = (2\pi t)^{-\frac{d}{2}} \exp\left(-\frac{|x|^2}{2t}\right), \quad x \in \mathbb{R}^d$$

the density function of a d-dimensional $N(0, tI_d)$ random variable.

Since the only term which depends on z in the integrand of (3.11) is $\Phi(y_1 - z, t_{k+1} - s)$, which is a smooth function of z, thus v is differentiable at all $z \in \mathbb{R}^d$, in particular at 0. Hence F_n is vertically differentiable at

$(s,\omega) \in \Lambda_T^d$ with: for $1 \le i \le d$,

$$
\begin{aligned}
\partial_i F_n(s,\omega) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_k-)-\omega(t_{k-1}-),\\
&\quad y_1+\omega(s)-\omega(t_k-),y_2,\cdots,y_{n-k})\Big)\\
&\quad \frac{y_1\cdot e_i}{t_{k+1}-s}\Phi(y_1,t_{k+1}-s)\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1dy_2\cdots dy_{n-k}\\
&= \mathbb{E}\left[g({}_nX(\omega_s \underset{s}{\oplus} B_T))\frac{(B(t_{k+1})-B(s))\cdot e_i}{t_{k+1}-s}\right]. \qquad (3.12)
\end{aligned}
$$

Remark that when s tends towards t_{k+1}, $\nabla_\omega F_n(s,\omega)$ may tend to infinity because of the term $t_{k+1}-s$ in the denominator. However in the interval $[t_k,t_{k+1})$, $\nabla_\omega F_n(s,\omega)$ behaves well and is locally bounded.

Iterating this procedure, one can show that F_n is vertically differentiable for any order. For example, we have:

$$
\begin{aligned}
\partial_i^2 F_n(s,\omega) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_k-)-\omega(t_{k-1}-),\\
&\quad y_1+\omega(s)-\omega(t_k-),y_2,\cdots,y_{n-k})\Big)\left(\frac{(y_1\cdot e_i)^2}{(t_{k+1}-s)^2}-\frac{1}{t_{k+1}-s}\right)\\
&\quad \Phi(y_1,t_{k+1}-s)\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1dy_2\cdots dy_{n-k}.
\end{aligned}
$$

And for $i \neq j$:

$$
\begin{aligned}
\partial_{ij} F_n(s,\omega) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_k-)-\omega(t_{k-1}-),\\
&\quad y_1+\omega(s)-\omega(t_k-),y_2,\cdots,y_{n-k})\Big)\frac{(y_1\cdot e_i)(y_1\cdot e_j)}{(t_{k+1}-s)^2}\\
&\quad \Phi(y_1,t_{k+1}-s)\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1\cdots dy_{n-k}.
\end{aligned}
$$

The horizontal differentiability of F_n can be proved similarly. Consider the following map:

$$
w(h) := F_n(s+h,\omega_s) = \mathbb{E}\left[g({}_nX(\omega_s \underset{s+h}{\oplus} B_T))\right], \quad h>0.
$$

The objective is to show that w is right-differentiable at 0.

We assume again that $t_k \leq s < t_{k+1}$ for some $0 \leq k \leq n-1$, and we can always choose an $h > 0$ small enough such that $s+h < t_{k+1}$. Using the same argument as in the proof of the vertical differentiability of F_n and the fact that $\omega_s(s+h) = \omega(s)$, we have:

$$
\begin{aligned}
{}_nX(\omega_s \underset{s+h}{\oplus} B_T) &= p_n\Big(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), \\
&\quad B(t_{k+1}) - B(s+h) + \omega(s) - \omega(t_k-), \\
&\quad B(t_{k+2}) - B(t_{k+1}), \cdots, B(t_n) - B(t_{n-1})\Big).
\end{aligned}
$$

Let $y = (y_1, \cdots, y_{n-k}) \in \mathbb{M}_{d,n-k}(\mathbb{R})$. We compute explicitly $w(h)$:

$$
\begin{aligned}
w(h) &= \mathbb{E}\left[g({}_nX_T(\omega_s \underset{s+h}{\oplus} B_T))\right] \\
&= \mathbb{E}\Big[g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), \\
&\quad B(t_{k+1}) - B(s+h) + \omega(s) - \omega(t_k-), \\
&\quad B(t_{k+2}) - B(t_{k+1}), \cdots, B(t_n) - B(t_{n-1}))\Big)\Big] \\
&= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + \omega(s) - \omega(t_k-), \\
&\quad y_2, \cdots, y_n)\Big)\Phi(y_1, t_{k+1} - s - h)\prod_{l=2}^{n-k}\Phi(y_l, \delta)dy_1dy_2\cdots dy_{n-k}. \qquad (3.13)
\end{aligned}
$$

Again the only term which depends on h in the integrand of (3.13) is $\Phi(y_1, t_{k+1} - s - h)$, which is a smooth function of h. Therefore F_n is horizontally differentiable with:

$$
\begin{aligned}
\mathcal{D}F_n(s, \omega_s) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), \\
&\quad y_1 + \omega(s) - \omega(t_k-), y_2, \cdots, y_n)\Big)\left(\frac{d}{2(t_{k+1}-s)} - \frac{|y_1|^2}{2(t_{k+1}-s)^2}\right) \\
&\quad \Phi(y_1, t_{k+1} - s)\prod_{l=2}^{n-k}\Phi(y_l, \delta)dy_1\cdots dy_{n-k}. \qquad (3.14)
\end{aligned}
$$

□

The following result shows that the functional F_n and its derivatives satisfy the necessary regularity conditions for applying the functional Itô formula (theorem 2.4):

Theorem 3.8. *Under assumption 3.1 and assumption 3.2, $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda^d_T)$.*

Proof. We have already shown in Theorem 3.7 that F_n is horizontally differentiable and twice vertically differentiable. Using the expressions of $\mathcal{D}F_n$, $\nabla_\omega F_n$ and $\nabla^2_\omega F_n$ obtained in the proof of 3.7, and the assumption that g has at most polynomial growth with respect to $\|\cdot\|_\infty$ (assumption 3.2), we observe that in each interval $[t_k, t_{k+1})$ with $0 \leq k \leq n-1$, $\mathcal{D}F_n$, $\nabla_\omega F_n$ and $\nabla^2_\omega F_n$ satisfy the boundedness-preserving property. It remains to show that F_n is left-continuous, $\nabla_\omega F_n$ and $\nabla^2_\omega F_n$ are right-continuous, and $\mathcal{D}F_n$ is continuous at fixed times.

Let $s \in [t_k, t_{k+1})$ for some $0 \leq k \leq n-1$ and $\omega \in D([0,T], \mathbb{R}^d)$. We first prove that F_n is right-continuous at (s,ω), and is jointly continuous at (s,ω) for $s \in (t_k, t_{k+1})$. By definition of joint-continuity (or right-continuous) (definition 1.7), we want to show that: $\forall \epsilon > 0, \exists \eta > 0, \forall (s',\omega') \in \Lambda^d_T$ (for the right-continuity, we assume in addition that $s' > s$),

$$d_\infty((s,\omega),(s',\omega')) < \eta \Rightarrow |F_n(s,\omega) - F_n(s',\omega')| < \epsilon.$$

Let $(s',\omega') \in \Lambda^d_T$ (with $s' > s$ for the right-continuity). We assume that $d_\infty((s,\omega),(s',\omega')) \leq \eta$ with an η small enough such that $s' \in [t_k, t_{k+1})$ (this is always possible as for $s = t_k$, we are only interested in the right-continuity of F_n on (s,ω)). It suffices to prove that $|F_n(s,\omega) - F_n(s',\omega')| \leq C(s,\omega_s,\eta)$ with $C(s,\omega_s,\eta)$ a quantity depending only on s, ω_s and η (thus independent of (s',ω')), and $C(s,\omega_s,\eta) \underset{\eta\to 0}{\to} 0$.

Now we use the expression of F_n obtained in the proof of theorem 3.7. Let $y = (y_1, \cdots, y_{n-k}) \in \mathbb{M}_{d,n-k}(\mathbb{R})$. We have:

$$\begin{aligned} F_n(s,\omega) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + \omega(s) - \omega(t_k-), \\ &\qquad y_2, \cdots, y_n)\Big)\Phi(y_1, t_{k+1} - s)\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1 dy_2\cdots dy_{n-k} \end{aligned}$$

and

$$F_n(s',\omega') = \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega'(t_1-)-\omega'(0),\cdots,\omega'(t_k-)-\omega'(t_{k-1}-),y_1+\omega'(s')-\omega'(t_k-), y_2,\cdots,y_n)\Big)\Phi(y_1,t_{k+1}-s')\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1dy_2\cdots dy_{n-k}.$$

To simplify the notations, we set:

$$\tilde{p}(\omega,s,y)=p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_k-)-\omega(t_{k-1}-),y_1+\omega(s)-\omega(t_k-),y_2,\cdots,y_n)$$

and

$$\tilde{p}(\omega',s',y)=p_n(\omega'(t_1-)-\omega'(0),\cdots,\omega'(t_k-)-\omega'(t_{k-1}-),y_1+\omega'(s')-\omega'(t_k-),y_2,\cdots,y_n).$$

We denote by $\tilde{p}_t(\cdot)$ the path of $\tilde{p}(\cdot)$ stopped at time t. Since $\|\omega_s-\omega_{s'}\|_\infty\le\eta$, for any $0\le j\le k-1$, $|(\omega(t_{j+1}-)-\omega(t_j-))-(\omega'(t_{j+1}-)-\omega'(t_j-))|\le 2\eta$ and $|(y_1+\omega(s)-\omega(t_k-))-(y_1+\omega'(s')-\omega'(t_k-))|\le 2\eta$, using lemma 3.3, we have:

$$\|\tilde{p}(\omega,s,y)-\tilde{p}(\omega',s',y)\|_\infty\le C(s,\omega_s,y,K_{Lip},T)\eta.$$

We can now control the difference between $F_n(s,\omega)$ and $F_n(s',\omega')$.

$$\begin{aligned}
&|F_n(s,\omega)-F_n(s',\omega')|\\
\le\ & \int_{\mathbb{R}^{d\times(n-k)}}|g(\tilde{p}(\omega,s,y))\Phi(y_1,t_{k+1}-s)-g(\tilde{p}(\omega',s',y))\Phi(y_1,t_{k+1}-s')|\\
&\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1dy_2\cdots dy_{n-k}\\
\le\ & \int_{\mathbb{R}^{d\times(n-k)}}\Big(|g(\tilde{p}(\omega,s,y))-g(\tilde{p}(\omega',s',y))|\Phi(y_1,t_{k+1}-s')\\
&+|g(\tilde{p}(\omega,s,y))|\cdot|\Phi(y_1,t_{k+1}-s)-\Phi(y_1,t_{k+1}-s')|\Big)\\
&\times\prod_{l=2}^{n-k}\Phi(y_l,\delta)dy_1dy_2\cdots dy_{n-k}.
\end{aligned}\tag{3.15}$$

Observe that $|\Phi(y_1,t_{k+1}-s)-\Phi(y_1,t_{k+1}-s')|\le|s-s'|\cdot\rho(y_1,\eta)\le\rho(y_1,\eta)\cdot\eta$ with

$$\rho(y_1,\eta):=\sup_{t\in[t_{k+1}-s-\eta,\delta]}|\partial_t\Phi(y_1,t)|,$$

and we have:

$$\rho(y_1, \eta) \underset{\eta \to 0}{\to} \sup_{t \in [t_{k+1}-s, \delta]} |\partial_t \Phi(y_1, t)| = \sup_{t \in [t_{k+1}-s, \delta]} \left| \Phi(y_1, t) \left(\frac{|y_1|^2}{2t^2} - \frac{d}{2t} \right) \right| < \infty.$$

So the second part of (3.15) can be controlled by:

$$\begin{aligned} &\int_{\mathbb{R}^{d \times (n-k)}} |g(\tilde{p}(\omega, s, y))| \cdot |\Phi(y_1, t_{k+1} - s) - \Phi(y_1, t_{k+1} - s')| \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy \\ \leq\ & C(s, \omega_s, \eta) \end{aligned}$$

with

$$C(s, \omega_s, \eta) \underset{\eta \to 0}{\to} 0.$$

For the first part of (3.15), we use the continuity of the functional g. As g is continuous at $\tilde{p}(\omega, s, y)$, we have:

$$|g(\tilde{p}(\omega, s, y)) - g(\tilde{p}(\omega', s', y))| \leq C(s, \omega_s, y, \eta).$$

with

$$C(s, \omega_s, y, \eta) \underset{\eta \to 0}{\to} 0, \quad y \in \mathbb{M}_{d, n-k}(\mathbb{R}).$$

And $\Phi(y_1, t_{k+1} - s')$ can be bounded independently of s':

$$\Phi(y_1, t_{k+1} - s') \leq \sup_{t \in [t_{k+1}-s-\eta, \delta]} \Phi(y_1, t) < \infty.$$

An application of the dominated convergence theorem yields:

$$\begin{aligned} &\int_{\mathbb{R}^{d \times (n-k)}} |g(\tilde{p}(\omega, s, y)) - g(\tilde{p}(\omega', s', y))| \Phi(y_1, t_{k+1} - s') \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy \\ \leq\ & C(s, \omega_s, \eta) \end{aligned}$$

with

$$C(s, \omega_s, \eta) \underset{\eta \to 0}{\to} 0.$$

We conclude that $|F_n(s, \omega_s) - F_n(s', \omega'_{s'})| \leq C(s, \omega_s, \eta)$ with $C(s, \omega_s, \eta)$ depending only on s, ω_s and η, and $C(s, \omega_s, \eta) \underset{\eta \to 0}{\to} 0$, which proves the right-continuity of F_n and the joint-continuity of F_n at all $(s, \omega) \in \Lambda_T^d$ for $s \neq t_k$, $0 \leq k \leq n-1$.

The right-continuity of $\nabla_\omega F_n$, $\nabla^2_\omega F_n$ and the continuity at fixed times of $\mathcal{D}F_n$ can be deduced similarly from the expressions of $\nabla_\omega F_n$, $\nabla^2_\omega F_n$ and $\mathcal{D}F_n$ obtained in the proof of theorem 3.7. Now it remains to show that for $1 \leq k \leq n$ and $\omega \in D([0,T],\mathbb{R}^d)$, F_n is left-continuous at (t_k,ω). Let $(s',\omega') \in \Lambda^d_T$ with $s' < t_k$ such that $d_\infty((t_k,\omega),(s',\omega')) \leq \eta$. We can always choose an $\eta < \delta$ in order that $s' \in [t_{k-1},t_k)$. The objective is to show that $|F_n(t_k,\omega) - F_n(s',\omega')| \leq C(t_k,\omega_{t_k},\eta)$ for some $C(t_k,\omega_{t_k},\eta)$ depending only on t_k, ω_{t_k} and η with $C(t_k,\omega_{t_k},\eta) \underset{\eta\to 0}{\to} 0$.

We first decompose $|F_n(t_k,\omega) - F_n(s',\omega')|$ into two terms:

$$|F_n(t_k,\omega) - F_n(s',\omega')| \leq |F_n(t_k,\omega) - F_n(s',\omega_{s'})| + |F_n(s',\omega_{s'}) - F_n(s',\omega')|.$$

For the second part, since F_n is continuous at fixed time s' by the first part of the proof, and $\|\omega_{s'} - \omega'_{s'}\|_\infty \leq \eta$, we have $|F_n(s',\omega_{s'}) - F_n(s',\omega')| \leq C(t_k,\omega_{t_k},\eta)$ with $C(t_k,\omega_{t_k},\eta) \underset{\eta\to 0}{\to} 0$.

For the first part $|F_n(t_k,\omega) - F_n(s',\omega_{s'})|$, the difficulty is that s' and t_k no longer lie in the same interval of the level-n Euler approximation, thus one more integration appears in the expression of $F_n(s',\omega_{s'})$ compared to that of $F_n(t_k,\omega)$. Let $y = (y_1,\cdots,y_{n-k}) \in \mathbb{M}_{d,n-k}(\mathbb{R})$ and $y' \in \mathbb{R}^d$. Using again the expression of F_n obtained in the proof of theorem 3.7, we have:

$$\begin{aligned}
F_n(t_k,\omega) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-) - \omega(0),\cdots,\omega(t_k-) - \omega(t_{k-1}-),\\
&\qquad y_1 + \omega(t_k) - \omega(t_k-), y_2,\cdots,y_{n-k})\Big)\prod_{l=1}^{n-k}\Phi(y_l,\delta)dy
\end{aligned}$$

and

$$\begin{aligned}
F_n(s',\omega_{s'}) &= \int_{\mathbb{R}^{d\times(n-k+1)}} g\Big(p_n(\omega(t_1-) - \omega(0),\cdots,\omega(t_{k-1}-) - \omega(t_{k-2}-),\\
&\qquad y' + \omega(s') - \omega(t_{k-1}-), y_1,\cdots,y_n)\Big)\Phi(y',t_k - s')dy'\prod_{l=1}^{n-k}\Phi(y_l,\delta)dy\\
&= \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^{d\times(n-k)}} g(p(\omega(t_1-) - \omega(0),\cdots,\omega(t_{k-1}-) - \omega(t_{k-2}-),\\
&\qquad y' + \omega(s') - \omega(t_{k-1}-), y_1,\cdots,y_n))\prod_{l=1}^{n-k}\Phi(y_l,\delta)dy\Big)\Phi(y',t_k - s')dy'.
\end{aligned}$$

We now define $\zeta : \mathbb{R}^d \to \mathbb{R}$ by: for $y' \in \mathbb{R}^d$,

$$\begin{aligned}\zeta(y') &:= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_{k-1}-)-\omega(t_{k-2}-),\\ &\qquad y'+\omega(s')-\omega(t_{k-1}-),y_1,\cdots,y_n)\Big)\prod_{l=1}^{n-k}\Phi(y_l,\delta)dy.\end{aligned}$$

By lemma 3.3 and the continuity of g with respect to $\|\cdot\|_\infty$, the map

$$\begin{aligned}y' &\mapsto g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_{k-1}-)-\omega(t_{k-2}-),\\ &\qquad y'+\omega(s')-\omega(t_{k-1}-),y_1,\cdots,y_n)\Big)\end{aligned}$$

is continuous. Since g has at most polynomial growth with respect to $\|\cdot\|_\infty$ (assumption 3.2), by the dominated convergence theorem, ζ is also continuous, and has at most polynomial growth in y'. And as $t_k - s' \le \eta$, we have:

$$\begin{aligned}F_n(s',\omega_{s'}) &= \int_{\mathbb{R}^d}\zeta(y')\Phi(y',t_k-s')dy'\\ &= \int_{\mathbb{R}^d}(\zeta(y')-\zeta(0))\Phi(y',t_k-s')dy'+\zeta(0).\end{aligned}$$

with

$$\left|\int_{\mathbb{R}^d}(\zeta(y')-\zeta(0))\Phi(y',t_k-s')dy'\right| \le C(t_k,\omega_{t_k},\eta),$$

and $C(t_k,\omega_{t_k},\eta)\underset{\eta\to 0}{\to} 0$.

It remains to control the difference between $F_n(t_k,\omega)$ and $\zeta(0)$. Remark that:

$$\begin{aligned}\zeta(0) &= \int_{\mathbb{R}^{d\times(n-k)}} g\Big(p_n(\omega(t_1-)-\omega(0),\cdots,\omega(t_{k-1}-)-\omega(t_{k-2}-),\\ &\qquad \omega(s')-\omega(t_{k-1}-),y_1,\cdots,y_n)\Big)\prod_{l=1}^{n-k}\Phi(y_l,\delta)dy\\ &= \mathbb{E}\left[g\left({}_nX_T(\omega_{s'}\underset{t_k}{\oplus}B_T)\right)\right] = F_n(t_k,\omega_{s'}).\end{aligned}$$

As $\|\omega_{s'}-\omega_{t_k}\|_\infty \le \|\omega_{s'}-\omega'_{s'}\|_\infty+\|\omega'_{s'}-\omega_{t_k}\|_\infty \le 2\eta$, using again the continuity of F_n at fixed time t_k established in the first part of the proof, we obtain:

$$|F_n(t_k,\omega)-\zeta(0)| \le C(t_k,\omega_{t_k},\eta)$$

with $C(t_k, \omega_{t_k}, \eta) \underset{\eta \to 0}{\to} 0$.

We conclude that

$$|F_n(t_k, \omega) - F_n(s', \omega')| \leq C(t_k, \omega_{t_k}, \eta)$$

with $C(t_k, \omega_{t_k}, \eta) \underset{\eta \to 0}{\to} 0$, which proves the left-continuity of F_n at (t_k, ω) for $1 \leq k \leq n$. □

Corollary 3.2. *Under assumption 3.1 and assumption 3.2, we have:*

$$F_n(T, W_T) - F_n(0, W_0) = \int_0^T \nabla_\omega F_n(s, W_s) \cdot dW(s), \quad \mathbb{P}\text{−a.s.} \tag{3.16}$$

Proof. As $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda^d_T)$, we can apply the functional Itô formula (theorem 2.4), and we observe that the finite variation term is zero since $Y_n(s) = F_n(s, W_s) = \mathbb{E}\left[g({}_nX(W_T))|\mathcal{F}_s\right]$ is a martingale. □

Remark 3.9. *Using the expressions of $\mathcal{D}F_n$ and $\nabla^2_\omega F_n$ obtained in the proof of theorem 3.7, we can also verify by direct computation that the finite variation terms in* (3.16) *cancel each other. By the functional Itô formula (theorem 2.4), the finite variation term in* (3.16) *equals to $\mathcal{D}F_n(s, W_s) + \frac{1}{2}\mathrm{tr}(\nabla^2_\omega F_n(s, W_s))$. And for $(s, \omega) \in \Lambda^d_T$ with $s \in [t_k, t_{k+1})$, $0 \leq k \leq n-1$, we have:*

$$\begin{aligned}
& \mathrm{tr}\left(\nabla^2_\omega F_n(s, \omega)\right) = \sum_{i=1}^d \partial^2_i F_n(s, \omega) \\
= & \int_{\mathbb{R}^{d\times(n-k)}} g\left(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + \omega(s) - \omega(t_k-), y_2, \cdots, y_n)\right) \\
& \sum_{i=1}^d \left(\frac{(y_1 \cdot e_i)^2}{(t_{k+1} - s)^2} - \frac{1}{t_{k+1} - s}\right) \Phi(y_1, t_{k+1} - s) \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy_1 \cdots dy_{n-k} \\
= & \int_{\mathbb{R}^{d\times(n-k)}} g\left(p_n(\omega(t_1-) - \omega(0), \cdots, \omega(t_k-) - \omega(t_{k-1}-), y_1 + \omega(s) - \omega(t_k-), y_2, \cdots, y_n)\right) \\
& \left(\frac{|y_1|^2}{(t_{k+1} - s)^2} - \frac{d}{t_{k+1} - s}\right) \Phi(y_1, t_{k+1} - s) \prod_{l=2}^{n-k} \Phi(y_l, \delta) dy_1 \cdots dy_{n-k} \\
= & -2\mathcal{D}F_n(s, \omega_s)
\end{aligned}$$

which confirms that F_n is indeed a solution of the path-dependent heat equation (2.16)*:*

$$\mathcal{D}F_n(s, \omega_s) + \frac{1}{2}\mathrm{tr}\left(\nabla^2_\omega F_n(s, \omega_s)\right) = 0, \quad \omega \in D([0, T], \mathbb{R}^d).$$

3.4 Convergence and error analysis

In this section, we analyze the convergence and the rate of convergence of our approximation method. After having constructed a sequence of smooth functionals F_n defined by (3.10) (theorem 3.7 and theorem 3.8), we can now approximate $\nabla_W Y$ by:

$$Z_n(s) := \nabla_\omega F_n(s, W_s)$$

which, in contrast to the weak derivative $\nabla_W Y$, is computable as a pathwise directional derivative. In practice, $\nabla_\omega F_n(s, W_s)$ can be computed numerically via a finite difference method or a Monte-Carlo method using the expression (3.12) of $\nabla_\omega F_n$.

We can measure the error of our approximation method by the integral of $\nabla_W Y - Z_n$ with respect to W, i.e.

$$\int_0^T (\nabla_W Y - Z_n) \cdot dW = \int_0^T \nabla_W Y(s) \cdot dW(s) - \int_0^T \nabla_\omega F_n(s, W_s) \cdot dW(s).$$

By the martingale representation formula theorem 2.10 and corollary 3.2, we have $\mathbb{P}$-a.s.

$$\begin{aligned}\int_0^T (\nabla_W Y - Z_n) \cdot dW &= Y(T) - Y(0) - (Y_n(T) - Y_n(0)) \\ &= g(X_T) - g({}_nX(W_T)) - \mathbb{E}\left[g(X_T) - g({}_nX(W_T))\right]\end{aligned}$$

where ${}_nX$ is the path of the piecewise constant Euler-Maruyama scheme defined in definition 3.2. We first establish the almost sure convergence and the convergence in L^{2p} for $p \geq 1$ of our approximation.

Proposition 3.3. *Under assumption 3.1 and assumption 3.2, we have:*

$$\int_0^T (\nabla_W Y - Z_n) \cdot dW \underset{n\to\infty}{\to} 0, \quad \mathbb{P}\text{−a.s.}$$

Proof. Since g is assumed to be continuous with respect to $\|\cdot\|_\infty$ by assumption 3.2, and $\|X_T - {}_nX(W_T)\|_\infty \underset{n\to\infty}{\to} 0$ $\mathbb{P}$-a.s. by corollary 3.1, we have:

$$g(X_T) - g({}_nX(W_T)) \underset{n\to\infty}{\to} 0, \quad \mathbb{P}\text{−a.s.}$$

Moreover, under assumption 3.2, $\{(g({}_nX(W_T)))_{n\geq 1}, g(X_T)\}$ is bounded in L^2 thus uniformly integrable, therefore we also have:

$$\mathbb{E}\left[g(X_T) - g({}_nX(W_T))\right] \underset{n\to\infty}{\to} 0.$$

□

Corollary 3.3. *Under assumption 3.1 and assumption 3.2, we have:*

$$\left\|\int_0^T (\nabla_W Y - Z_n)\cdot dW\right\|_{2p} \underset{n\to\infty}{\to} 0, \quad \forall p \geq 1.$$

Proof. g has at most polynomial growth with respect to $\|\cdot\|_\infty$ by assumption 3.2, which ensures the uniform integrability of $|g({}_nX(W_T))|^{2p}$ for any $p \geq 1$. We conclude by applying the dominated convergence theorem. □

To obtain a rate of convergence of our approximation, we need some stronger Lipschitz-type assumption on g.

Theorem 3.10 (Rate of convergence of approximation)**.** *We assume $g : (D([0,T],\mathbb{R}^d), \|\cdot\|_\infty) \to \mathbb{R}$ is Lipschitz continuous:*

$$\exists g_{Lip} > 0, \quad \forall \omega, \omega' \in D([0,T],\mathbb{R}^d), \quad |g(\omega) - g(\omega')| \leq g_{Lip}\|\omega - \omega'\|_\infty.$$

Under assumption 3.1, the L^{2p}-error of the approximation Z_n of $\nabla_W Y$ along the path of W is bounded by:

$$\mathbb{E}\left\|\int_0^T (\nabla_W Y - Z_n)\cdot dW\right\|^{2p} \leq C(x_0, p, T, K_{Lip}, g_{Lip})\left(\frac{1+\log n}{n}\right)^p, \quad \forall p \geq 1$$

where the constant C depends only on x_0, p, T, K_{Lip} and g_{Lip}. In particular, we have:

$$\forall \alpha \in [0, \frac{1}{2}), \quad n^\alpha\left(\int_0^T (\nabla_W Y - Z_n)\cdot dW\right) \underset{n\to\infty}{\to} 0, \quad \mathbb{P}\text{−a.s.}$$

Proof. This result is a consequence of proposition 3.2 since

$$\begin{aligned}\left\|\int_0^T (\nabla_W Y - Z_n)\cdot dW\right\|_{2p} &\leq \|g(X_T) - g({}_nX(W_T))\|_{2p} + \|\mathbb{E}[g(X_T) - g({}_nX(W_T))]\|_{2p}\\ &\leq 2\,\|g(X_T) - g({}_nX(W_T))\|_{2p}\\ &\leq 2g_{Lip}\left\|\sup_{0\leq s\leq T}|X(s) - {}_nX(s)|\right\|_{2p}.\end{aligned}$$

□

We now provide an example to show how our result may be used to construct explicit approximations with controlled convergence rates for conditional expectation of non-smooth functionals:

Example 3.11. *Let X be the strong solution of the SDE* (3.4) *with b and σ satisfying assumption 3.1, and let*

$$g(\omega_T) := \phi\left(\omega(T), \sup_{0\le t\le T} \|\omega(t)\|\right)$$

where $\phi : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ is a continuous function with polynomial growth. Set $Y(t) := \mathbb{E}[g(X_T)|\mathcal{F}_t]$. Then g satisfies assumption 3.2, and our approximation method applies. Moreover, if ϕ is Lipschitz continuous, then theorem 3.10 provides an explicit control of the approximation error with a convergence rate of order $\sqrt{\frac{\log n}{n}}$.

3.5 Comparison with approaches based on the Malliavin calculus

The vertical derivative $\nabla_W Y$ which appears in the martingale representation formula may be viewed as the 'sensitivity' of the martingale Y to the Brownian motion W. Thus, our method can be compared to other methods proposed in the literature for 'sensitivity analysis' of Wiener functionals, in particular the Malliavin calculus.

Such methods can be roughly classified into two categories [5]: methods that differentiate paths and methods that differentiate densities. When the density of the functional is known, the sensitivity of an expectation with respect to some parameter is to differentiate directly the density function with respect to the parameter. However, as this is almost never the case in a general diffusion model, let alone a non-Markovian model, alternative methods are used. Usually the idea is to differentiate either the functional g, or the process with respect to the parameter under the expectation sign, then estimate the expectation with the Monte-Carlo method. To differentiate process, one requires the existence of the so-called first variation process, which requires the regularity of the coefficients of the SDE satisfied by X.

Sensitivity estimators for non-smooth functionals can be obtained using Malliavin calculus: this approach, proposed by Fournié et al. [29], and developed by Cvitanić, Ma and Zhang [16], Fournié et al. [30], Gobet and Kohatsu-Higa [33], Kohatsu-Higa and Montero [44], Davis and Johansson [18] and others, uses the Malliavin integration by parts formula on Wiener space in the case when g is not smooth. These methods require quite demanding regularity assumptions (differentiability and ellipticity condition on σ for example) on the coefficients of the initial SDE satisfied by X.

By contrast, the approximation method presented in this chapter allows for any continuous functional g with polynomial growth and requires only a Lipschitz continuity assumption on the functional σ, which also allows for degenerate coefficients. It is thus applicable to a wider range of examples than the Malliavin approach, while being arguably simpler from a computational viewpoint. Contrary to the Malliavin approach, which involves differentiating in the Malliavin sense, then discretizing the tangent process, our method involves discretizing then differentiating (the Euler-Maruyama scheme) which, as argued in [5], may have computational advantages.

In this chapter, we have shown that $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda^d_T)$ (theorem 3.8), which is sufficient for obtaining an approximation of martingale representation via the functional Itô formula. A natural question is whether the same type of result can be obtained using the Malliavin calculus, for example via a Clark-Haussmann-Ocone type formula. While under our quite general setting, the initial $\mathcal{F}_T$-measurable random variable $H := g(X_T)$ has no reason to be Malliavin differentiable, one may ask if the Malliavin calculus is applicable to the piecewise constant Euler approximation ${}_nX$ (definition 3.2). And in this case, whether the pathwise vertical derivative $\nabla_\omega F_n(t, W_t)$ leads to the same representation as in the Clark-Haussmann-Ocone formula.

Let $n \in \mathbb{N}$, and we define $H_n := g({}_nX(W_T))$ with ${}_nX$ the weak piecewise constant Euler-Maruyama scheme defined by (3.5). By the definition of ${}_nX$, the $\mathcal{F}_T$-measurable random variable H_n actually depends only on a finite number of Gaussian variables: $W(t_1)$, $W(t_2)-W(t_1)$, $\cdots$, $W(t_n)-W(t_{n-1})$, thus it can be written as:

$$H_n = h_n(W(t_1), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1}))$$

with $h_n : \mathbb{M}_{d,n}(\mathbb{R}) \to \mathbb{R}$ (h_n is actually the composition of g and p_n: $h = g \circ p_n$ with p_n defined by (3.6)).

Clearly if h_n is a smooth function with polynomial growth, then H_n is Malliavin differentiable ($H_n \in \mathbb{D}^{1,2}$) with Malliavin derivative [53]:

$$\mathbb{D}_t H_n = (\mathbb{D}_t^k H_n, 1 \le k \le d) \in \mathbb{R}^d, \quad t \in [0,T)$$

where

$$\mathbb{D}_t^k H_n = \sum_{j=0}^{n-1} \partial_{kj} h_n(W(t_1), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1})) \mathbb{1}_{[t_j, t_{j+1})}(t).$$

Using the Clark-Haussmann-Ocone formula [54, 53], we have thus:

$$H_n = \mathbb{E}[H_n] + \int_0^T \mathbb{E}[\mathbb{D}_t H_n | \mathcal{F}_t] \cdot dW(t).$$

The integrand $\mathbb{E}[\mathbb{D}_t H_n|\mathcal{F}_t]$ can be computed explicitly in this case. Assume that $t \in [t_j, t_{j+1})$ for some $0 \le j \le n-1$, we have, for $1 \le k \le d$,

$$\begin{aligned}
& \mathbb{E}[\mathbb{D}_t^k H_n | \mathcal{F}_t] \\
= \; & \mathbb{E}\left[\partial_{kj} h_n(W(t_1), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1})) | \mathcal{F}_t\right] \\
= \; & \mathbb{E}\big[\partial_{kj} h_n(\omega(t_1), \omega(t_2) - \omega(t_1), \cdots, \omega(t_j) - \omega(t_{j-1}), \omega(t) - \omega(t_j) + W(t_{j+1}) - W(t), \\
& W(t_{j+2}) - W(t_{j+1}), \cdots, W(t_n) - W(t_{n-1}))\big] |_{\omega_t = W_t} \\
= \; & \mathbb{E}\Big[\frac{\partial}{\partial h} h_n(\omega(t_1), \omega(t_2) - \omega(t_1), \cdots, \omega(t_j) - \omega(t_{j-1}), \omega(t) - \omega(t_j) + W(t_{j+1}) - W(t) + h e_k, \\
& W(t_{j+2}) - W(t_{j+1}), \cdots, W(t_n) - W(t_{n-1}))|_{h=0}\Big] |_{\omega_t = W_t} \\
= \; & \frac{\partial}{\partial h}\Big(\mathbb{E}\big[h_n(\omega(t_1), \omega(t_2) - \omega(t_1), \cdots, \omega(t_j) - \omega(t_{j-1}), \omega(t) - \omega(t_j) + W(t_{j+1}) - W(t) + h e_k, \\
& W(t_{j+2}) - W(t_{j+1}), \cdots, W(t_n) - W(t_{n-1}))\big] |_{\omega_t = W_t}\Big) |_{h=0}
\end{aligned}$$

which is none other than

$$\partial_k F_n(t, W_t) = \lim_{h \to 0} \frac{F_n(t, W_t + he_k \mathbb{1}_{[t,T]}) - F_n(t, W_t)}{h}$$

where F_n is the non-anticipative functional defined by (3.10). So in the case when h_n is smooth, our method provides the same representation as given by the Clark-Haussmann-Ocone formula. However, in our framework, as the functional g is only assumed to be continuous with polynomial growth, the function h_n may fail to be differentiable. So even in this simple case with the weak piecewise constant Euler-Maruyama scheme ${}_nX$, it is not clear whether the random variable H_n is differentiable in the Malliavin sense, and even if it is the case, it is difficult to obtain an explicit form for the conditional expectation $\mathbb{E}[\mathbb{D}_t H_n|\mathcal{F}_t]$ using the Malliavin calculus.

However our approximation method applies even in the cases when H_n is not differentiable in the Malliavin sense: indeed, as shown in section 3.3, as soon as H is square-integrable, the martingale $Y_n := \mathbb{E}[H_n|\mathcal{F}_t]$ has a smooth functional representation which is differentiable in the pathwise sense, even though H_n is not differentiable, neither in a pathwise nor in the Malliavin sense. This reveals one important difference between the Malliavin approach and our method. While the Malliavin approach requires in general conditions on the final $\mathcal{F}_T$-measurable random variable H or H_n (for example $H_n \in \mathbb{D}^{1,2}$), as the construction of Malliavin derivative involves perturbations which apply to the whole path of the process, our method, based on a non-anticipative calculus, focuses instead on the process $Y_n(t) := \mathbb{E}[H_n|\mathcal{F}_t]$ (or equivalently the functional F_n). Since the condition expectation of a random variable possesses in general better regularity than the variable itself (as shown in the proof of theorem 3.7), our method applies under quite general setting even when H_n might not be Malliavin differentiable.

Another interesting observation comes from the expression of $\nabla_\omega F_n(t, W_t)$ obtained in the proof of theorem 3.7. By (3.12), we have, for $t \in [t_k, t_{k+1})$,

$$\partial_i F_n(t, \omega) = \mathbb{E}\left[g({}_nX(\omega_t \underset{t}{\oplus} B_T)) \frac{(B(t_{k+1}) - B(t)) \cdot e_i}{t_{k+1} - t} \right].$$

Taking $\omega = W_t$ and using independence of Brownian increments, $\nabla_\omega F_n(t, W_t)$ can be written as:

$$\nabla_\omega F_n(t, W_t) = \mathbb{E}\left[g({}_nX(W_T))\frac{W(t_{k+1}) - W(t)}{t_{k+1} - t}|\mathcal{F}_t\right]. \tag{3.17}$$

The factor $\pi(t) := \frac{W(t_{k+1})-W(t)}{t_{k+1}-t}$ in (3.17) can be viewed as a weight function frequently used in Malliavin methods, which is similar to the weight proposed for Brownian models, i.e. $dX(t) = bdt + \sigma dW(t)$ with b and σ two constants (see for example [29, 30]). Indeed, in our method, we first approximate X by its piecewise constant Euler-Maruyama scheme, which consists of 'freezing' the coefficients of the SDE satisfied by X between two time grids. Even the initial coefficients b and σ of the SDE are general path-dependent functionals, by considering a piecewise constant approximation of X, b and σ become locally constant, which allows us to obtain a simple 'local' weight function as shown in (3.17).

3.6 Some numerical aspects of the approximation method

Unlike some applications, for example sensitivity analysis in finance, where the focus is to compute the sensitivity of an option price with respect to some parameter at a given time (usually the current time), our approximation method approximates $\nabla_W Y$ as a process along the whole interval $[0, T]$. Thus the approximation error we consider, as shown in section 3.4, is

$$\int_0^T (\nabla_W Y(t) - \nabla_\omega F_n(t, W_t)) \cdot dW(t),$$

not the pointwise error in $\nabla_\omega F_n(t, W_t)$, such as $|\nabla_W Y(t) - \nabla_\omega F_n(t, W_t)|$ for a given t or $\|\nabla_W Y - \nabla_\omega F_n(\cdot, W)\|_\infty$.

However in practice, it is only possible to simulate the process $\nabla_\omega F_n(\cdot, W)$ at a finite number of points in $[0, T]$, which consists of approximating $\nabla_\omega F_n(\cdot, W)$ by a piecewise constant process which coincides with $\nabla_\omega F_n(\cdot, W)$ at, for example, the time grid $(t_i)_{0\leq i\leq n}$. Here we are interested in this discretization error of $\nabla_\omega F_n(\cdot, W)$, i.e.

$$\int_0^T \nabla_\omega F_n(t, W_t) \cdot dW(t) - \sum_{i=0}^{n-1} \nabla_\omega F_n(t_i, W_{t_i}) \cdot (W(t_{i+1}) - W(t_i)). \quad (3.18)$$

Let $Z_n(t) := \nabla_\omega F_n(t, W_t)$ and $Y_n(t) := \mathbb{E}[g({}_nX(W_T))|\mathcal{F}_t]$ for $t \in [0, T]$. Recall that by (3.17), we have, for $t \in [t_k, t_{k+1})$ for some $0 \leq k \leq n-1$,

$$\begin{aligned} Z_n(t) &= \mathbb{E}\left[Y_n(T)\frac{W(t_{k+1}) - W(t)}{t_{k+1} - t}|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[Y_n(T)\frac{W(t_{k+1}) - W(t)}{t_{k+1} - t}|\mathcal{F}_{t_{k+1}}\right]|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[Y_n(t_{k+1})\frac{W(t_{k+1}) - W(t)}{t_{k+1} - t}|\mathcal{F}_t\right]. \end{aligned} \quad (3.19)$$

The expression (3.19) seems to be similar, for those who are familiar with the theory of backward stochastic differential equations (BSDE), to the commonly used numerical scheme of BSDEs [73, 4]. Indeed, a SDE can be viewed as a special case of BSDE. Let us consider the following decoupled path-dependent forward-backward stochastic differential equation (FBSDE):

$$\begin{cases} dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d \\ -dY(t) = f(t, X_t, Y(t), Z(t))dt - Z(t) \cdot dW(t), \quad Y(T) = g(X_T) \end{cases}$$

with b and σ non-anticipative functionals satisfying assumption 3.1, and $f : \Lambda_T^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ the driver of the BSDE. In the peculiar case where $f \equiv 0$, the solution (Y, Z) of the BSDE is explicit: $Y(t) = \mathbb{E}[g(X_T)|\mathcal{F}_t]$ et Z is the integrand in the martingale representation of Y.

With our method, we approximate X by the piecewise constant Euler scheme ${}_nX(W_T)$, Y by $Y_n = \mathbb{E}[g({}_nX(W_T))|\mathcal{F}_t]$, and Z by $Z_n = \nabla_W Y_n$ which admits also the representation (3.19). Since the above expression of (Y_n, Z_n) coincides with the numerical scheme of BSDEs [73, 4] at $(t_i)_{0\leq i\leq n}$, the results concerning (Y_n, Z_n) in the BSDE literature also apply here. We start with a simple result proven in [4]. Here we provide an elementary proof which does not make use of Malliavin calculus as in [4]. Recall that in our setting, $t_k = k \cdot \delta$ with $\delta = \frac{T}{n}$.

Lemma 3.12. *For all* $0 \leq k \leq n-1$, *we have:*

$$Z_n(t_k) = \frac{1}{\delta}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z_n(t)dt|\mathcal{F}_{t_k}\right]. \tag{3.20}$$

Proof. Let $0 \leq k \leq n-1$. Since Z_n is the integrand in the martingale representation of Y_n, we have:

$$Y_n(t_{k+1}) = Y_n(t_k) + \int_{t_k}^{t_{k+1}} Z_n(t) \cdot dW(t).$$

Now using the expression (3.19) of Z_n, we have:

$$\begin{aligned}
Z_n(t_k) &= \frac{1}{\delta}\mathbb{E}\left[Y_n(t_{k+1})(W(t_{k+1}) - W(t_k))|\mathcal{F}_{t_k}\right] \\
&= \frac{1}{\delta}\mathbb{E}\left[Y_n(t_k)(W(t_{k+1}) - W(t_k))|\mathcal{F}_{t_k}\right] \\
&\quad + \frac{1}{\delta}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} Z_n(t) \cdot dW(t)\right)\left(\int_{t_k}^{t_{k+1}} dW(t)\right)|\mathcal{F}_{t_k}\right] \\
&= \frac{1}{\delta}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z_n(t)dt|\mathcal{F}_{t_k}\right]
\end{aligned}$$

where we have used the (conditional) Itô isometry in the last equality. □

This lemma still holds if we consider now the value of Z_n at any $t \in [t_k, t_{k+1})$, i.e. for all $t \in [t_k, t_{k+1})$, we have:

$$Z_n(t) = \frac{1}{t_{k+1} - t}\mathbb{E}\left[\int_{t}^{t_{k+1}} Z_n(s)ds|\mathcal{F}_t\right].$$

. The property (3.20) of Z_n has an important numerical implication: $Z_n(t_k)$ is actually the best $\mathcal{F}_{t_k}$-measurable approximation of Z_n in $[t_k, t_{k+1})$ with respect to the L^2-norm:

Corollary 3.4. *For any* $0 \leq k \leq n-1$, $Z_n(t_k)$ *solves the following minimization problem:*

$$\min_{U:\mathcal{F}_{t_k}-\text{measurable}} \mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} (Z_n(t) - U) \cdot dW(t)\right)^2 |\mathcal{F}_{t_k}\right]. \tag{3.21}$$

Proof.

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}}(Z_n(t)-U)\cdot dW(t)\right)^2|\mathcal{F}_{t_k}\right]\\
= \ &\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\mathrm{tr}((Z_n(t)-U)\,{}^t(Z_n(t)-U))dt|\mathcal{F}_{t_k}\right]\\
= \ &\delta\cdot|U|^2-2\mathbb{E}\left[\int_{t_k}^{t_{k+1}}Z_n(t)dt|\mathcal{F}_{t_k}\right]\cdot U+\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_n(t)|^2dt|\mathcal{F}_{t_k}\right].
\end{aligned}$$

Clearly the $\mathcal{F}_{t_k}$-measurable random variable which minimizes (3.21) is given by:

$$U_{min}:=\frac{1}{\delta}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}Z_n(t)dt|\mathcal{F}_{t_k}\right]=Z_n(t_k).$$

□

What remains is to control explicitly the discretization error (3.18). We now state a result from [73, 4] adapted to our case.

Proposition 3.4. *We assume that the coefficients b and σ of the SDE satisfy assumption 3.1, and $g:(D([0,T],\mathbb{R}^d),\|\cdot\|_\infty)\to\mathbb{R}$ is Lipschitz continuous:*

$$\exists g_{Lip}>0,\quad\forall\omega,\omega'\in D([0,T],\mathbb{R}^d),\quad|g(\omega)-g(\omega')|\le g_{Lip}\|\omega-\omega'\|_\infty.$$

Under these assumptions, the L^2-discretization error of Z_n is bounded by:

$$\sum_{k=0}^{n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_n(t)-Z_n(t_k)|^2dt\right]\le C\left(\frac{1+\log n}{n}\right),\qquad(3.22)$$

where C is a constant which only depends on x_0, T, K_{Lip} and g_{Lip}.

Combining (3.22) with theorem 3.10, we obtain the L^2-approximation error of the martingale representation of $g(X_T)$ by the piecewise constant process $\sum_{k=0}^{n-1}Z_n(t_k)\mathbb{1}_{[t_k,t_{k+1})}$.

Corollary 3.5. *Under the assumptions of proposition 3.4, we have:*

$$\sum_{k=0}^{n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z(t)-Z_n(t_k)|^2dt\right]\le C\left(\frac{1+\log n}{n}\right),$$

where C is a constant which only depends on x_0, T, K_{Lip} and g_{Lip}, and Z the integrand in the martingale representation of $g(X_T)$, i.e. Z satisfies:

$$g(X_T) = \mathbb{E}[g(X_T)] + \int_0^T Z(t) \cdot dW(t).$$

The numerical computation of $Z_n(t_k)$, as we have shown earlier, can be performed via either a finite difference method: $Z_n(t_k) = (Z_n^i(t_k), 1 \leq i \leq d)$ with:

$$Z_n^i(t_k) = \lim_{h \to 0} \frac{F_n(t_k, W_{t_k} + he_i \mathbb{1}_{[t_k,T]}) - F_n(t_k, W_{t_k})}{h},$$

or a Monte-Carlo method using directly the expression (3.17) of Z_n.

Remark 3.13. *In some applications, such as sensitivity analysis in finance, we are more interested in the sensitivity of the martingale Y with respect to X, i.e. $\nabla_X Y$, the weak vertical derivative of Y with respect to X. As we have mentioned in remark 3.5, a direct construction of a sequence of smooth functionals of X which approximates Y is difficult since we have little information about the density of the law of X.*

However we can still obtain an approximation of $\nabla_X Y$ using our method since a jump of W and that of X are closely related if σ is non-degenerate. Let $M(t) := \int_0^t \sigma(s, X_s) dW(s)$ be the local martingale component of X. We have:

$$\begin{aligned}
& \int_0^T \nabla_\omega F_n(t, W_t) \cdot dW(t) \\
= & \int_0^T \nabla_\omega F_n(t, W_t) \cdot \sigma^{-1}(t, X_t) dM(t) \\
= & \int_0^T \left({}^t\sigma^{-1}(t, X_t) \nabla_\omega F_n(t, W_t) \right) \cdot dM(t).
\end{aligned}$$

So $\nabla_X Y(t)$ can be approximated by:

$${}^t\sigma^{-1}(t, X_t) \nabla_\omega F_n(t, W_t),$$

and the term ${}^t\sigma^{-1}(t, X_t)$ can again be approximated, for example, by ${}^t\sigma^{-1}(t, {}_nX_t)$ with ${}_nX$ the piecewise constant Euler approximation of X.

Remark 3.14. *Our method can also be applied to a general decoupled path-dependent forward-backward SDE:*

$$\begin{cases} dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d \\ -dY(t) = f(t, X_t, Y(t), Z(t))dt - Z(t) \cdot dW(t), \quad Y(T) = g(X_T) \end{cases}$$

with $f : \Lambda_T^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ a functional which at time t may depend on the path of X up to t.

We still approximate X by its piecewise constant Euler scheme ${}_nX$, and we pose $Y_n(T) := g({}_nX_T)$. For $1 \le k \le n$, we can show recursively:

- *$Y_n(t_k)$ depends only on $W(t_1)$, $W(t_2) - W(t_1)$, $\cdots$, $W(t_k) - W(t_{k-1})$.*
- *For $t \in [t_{k-1}, t_k)$, there exists $F_n^k \in \mathbb{C}_{b,r}^{1,2}(\Lambda_T^d)$ such that $\mathbb{E}[Y_n(t_k)|\mathcal{F}_t] = F_n^k(t, W_t)$. Thus $Z_n(t) := \nabla_\omega F_n^k(t, W_t)$ is well defined.*
- *$Y_n(t_{k-1})$ can be defined through the equation: $Y_n(t_{k-1}) = \mathbb{E}[Y_n(t_k)|\mathcal{F}_{t_{k-1}}] + f(t_{k-1}, {}_nX_{t_{k-1}}, Y_n(t_{k-1}), Z_n(t_{k-1}))\delta$.*
- *We can interpolate Y_n in the interval (t_{k-1}, t_k) by defining:*

  $$Y_n(t) := \mathbb{E}[Y_n(t_k)|\mathcal{F}_{t_{k-1}}] + f(t_{k-1}, {}_nX_{t_{k-1}}, Y_n(t_{k-1}), Z_n(t_{k-1}))(t_k - t), \quad t \in (t_{k-1}, t_k).$$

By the construction of Y_n and Z_n, (Y_n, Z_n) is actually the (exact) solution of the following equation:

$$\begin{cases} -dY_n(t) = f(\underline{t}, {}_nX_{\underline{t}}, Y_n(\underline{t}), Z_n(\underline{t}))dt - Z_n(t)dW(t) \\ Y_n(T) = g({}_nX_T) \end{cases}$$

with $\underline{t} := \lfloor \frac{t}{\delta} \rfloor \delta$, i.e. $\underline{t}$ is the largest subdivision point which is smaller or equal to t.

The advantage of this method is that it not only proposes a continuous numerical scheme of the BSDE, but also leads to a new way of simulating Z_n as the pathwise vertical derivative of Y_n with respect to W. Indeed, if we are able to simulate numerically $Y_n(t_k)$, using for example a regression-based method (see [35]), $Z_n(t_k)$ can be obtained via a finite difference method, or simply the derivative of $Y_n(t_k)$ with respect to $W(t_k) - W(t_{k-1})$ if $Y_n(t_k)$ is already approximated by an explicit smooth function of increments of W.

3.7 Applications to dynamic hedging of path-dependent options

As we have mentioned in the previous section, the weak martingale approximation method we proposed in section 3.3 allows to provide a systematic manner of hedging path-dependent options where the underlying security prices may following non-Markovian dynamics. Observe also that our method does not require strong assumptions (for example differentiability assumptions) on neither the coefficients of the SDE satisfied by the underlying nor the payoff functional.

We assume in the following that the financial market considered is complete, and there exists a risk-neutral measure (equivalent martingale measure) $\mathbb{P}$. Let $T > 0$, and $S = (S^i, 1 \leq i \leq d)$ be a d-dimensional process which represents the prices of d underlying assets. Let $r : [0,T] \to \mathbb{R}_+$ be a deterministic instantaneous interest rate function satisfying

$$\int_0^T r(t)dt < \infty.$$

We assume that S satisfies the following dynamics:

$$dS(t) = r(t)S(t)dt + \sigma(t, S_t)dW(t), \quad S(0) = s_0 \in \mathbb{R}^d,$$

where W is a d-dimensional $\mathbb{P}$-Brownian motion, and $\sigma : \Lambda_T^d \to \mathbb{M}_d(\mathbb{R})$ a non-anticipative functional which satisfies assumption 3.1. We assume in addition that σ is non-degenerate, i.e. $\det(\sigma(t, S_t)) \neq 0$, $dt \times d\mathbb{P}$-a.e. Let $g : D([0,T], \mathbb{R}^d) \to \mathbb{R}$ be the payoff functional of the option. In order to have an explicit control of the hedging error (theorem 3.10), we assume that g is Lipschitz continuous with respect to the supremum norm $\| \cdot \|$.

Denote by $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ the ($\mathbb{P}$-completed) natural filtration generated by W. Since the market is complete, the option can be perfectly hedged by a self-financing portfolio: the value of this portfolio V is equal to the price of the option at any time $t \in [0,T]$,

$$V(t) = \exp\left(-\int_t^T r(s)ds\right) \mathbb{E}\left[g(S_T)|\mathcal{F}_t\right].$$

Denote by $\widetilde{S}(t) := \exp(-\int_0^t r(s)ds)S(t)$ the discounted asset price, $\widetilde{V}(t) := \exp(-\int_0^t r(s)ds)V(t)$ the discounted portfolio value, and $\widetilde{g}(\widetilde{S}_T) := \exp(-\int_0^T r(s)ds)g(S_T)$. We have $\widetilde{V}(t) = \mathbb{E}[\widetilde{g}(\widetilde{S}_T)|\mathcal{F}_t]$. So by working with $(\widetilde{S}, \widetilde{V}, \widetilde{g})$ instead of (S, V, g), we may assume $r \equiv 0$ in the following without any loss of generality. Under this assumption, S and V are $\mathbb{P}$-martingales, and there exists a $\mathbb{F}$-adapted process ϕ such that:

$$g(S_T) = V(T) = V(0) + \int_0^T \phi(t) \cdot dS(t).$$

Let $n \in \mathbb{N}$, and $\delta = \frac{T}{n}$. We denote by ${}_nS$ the piecewise constant Euler approximation of S on the grid $(t_j = j\delta, j = 0, \cdots, n)$ defined in definition 3.2: for $\omega \in D([0,T], \mathbb{R}^d)$, ${}_nS(\omega)$ is constant in $[t_j, t_{j+1})$ for any $0 \le j \le n-1$ with ${}_nS(0,\omega) = s_0$, and

$${}_nS(t_{j+1}, \omega) = {}_nS(t_j, \omega) + \sigma(t_j, {}_nS_{t_j}(\omega))(\omega(t_{j+1}-) - \omega(t_j-)).$$

And similarly to definition 3.6, we define: for $\omega \in D([0,T], \mathbb{R}^d)$,

$$F_n(t, \omega_t) := \mathbb{E}\left[g\left({}_nS(\omega_t \underset{t}{\oplus} B_T)\right)\right], \quad t \in [0,T]$$

where B is a Wiener process independent of W. We have, for $t \in [0,T]$, $F_n(t, W_t) = \mathbb{E}[g({}_nS(W_T))|\mathcal{F}_t]$ $\mathbb{P}$-almost surely, and $F_n \in \mathbb{C}^{1,2}_{loc,r}(\Lambda^d_T)$ (theorem 3.8). In the following, we note simply ${}_nS$ for ${}_nS(W_T)$, so ${}_nS$ is the classical piecewise constant Euler-Maruyama scheme of S.

Here we are interested in the problem of dynamic hedging in discrete time, so the objective is to approximate ϕ by a piecewise constant process ϕ_n. As suggested by remark 3.13, a natural candidate of ϕ_n is: for $0 \le j \le n-1$,

$$\phi_n(t_j) := {}^t\sigma^{-1}(t_j, {}_nS_{t_j})\nabla_\omega F_n(t_j, W_{t_j}).$$

Here we assume in addition that $\det(\sigma(t, {}_nS_t)) \neq 0$, $dt \times d\mathbb{P}$-a.e.

To compute numerically $\nabla_\omega F_n(t_j, W_{t_j})$, we can either use directly the explicit expression of $\nabla_\omega F_n(t_j, W_{t_j})$ obtained in (3.17):

$$\nabla_\omega F_n(t_j, W_{t_j}) = \mathbb{E}\left[g({}_nS_T)\frac{W(t_{j+1}) - W(t_j)}{\delta}|\mathcal{F}_{t_j}\right],$$

and then compute the expectation using the Monte-Carlo method. We can also use the definition of vertical derivative as sensitivity of a functional with respect to a jump of the path, and thus approximate it via a finite difference method: $\nabla_\omega F_n(t_j, W_{t_j}) = (\partial_i F_n(t_j, W_{t_j}), 1 \le i \le d)$ with

$$\partial_i F_n(t_j, W_{t_j}) \simeq \frac{F_n(t_j, W_{t_j} + he_i \mathbb{1}_{[t_j,T]}) - F_n(t_j, W_{t_j})}{h}$$

for h small. The advantage of the first method is that it provides a unbiased estimator of $\nabla_\omega F_n(t_j, W_{t_j})$ contrary to the finite difference method. However since $\delta = \frac{T}{n}$ is in general quite small for a reasonable number of discretization steps, the term in the expectation can be quite considerable, leading to immense variance which makes the Monte-Carlo method less efficient. So in the following we adopt the second approach.

Let $h > 0$. Since F_n is infinitely vertically differentiable, we can use a centered difference scheme in order to have a more accurate approximation:

$$\widehat{\partial_i F_n}(t_j, W_{t_j}) := \frac{F_n(t_j, W_{t_j} + he_i \mathbb{1}_{[t_j,T]}) - F_n(t_j, W_{t_j} - he_i \mathbb{1}_{[t_j,T]})}{2h}.$$

Each of these two terms can be computed numerically via the Monte-Carlo method. For example:

$$\begin{aligned} & F_n(t_j, W_{t_j} + he_i \mathbb{1}_{[t_j,T]}) \\ = \ & \mathbb{E}\left[g\left({}_nS((W_{t_j} + he_i \mathbb{1}_{[t_j,T]}) \underset{t_j}{\oplus} B_T)\right)\right] = \mathbb{E}\left[g\left({}_nS((W_{t_j} \underset{t_j}{\oplus} B_T) + he_i \mathbb{1}_{[t_j,T]})\right)\right] \\ = \ & \mathbb{E}\left[g\left({}_nS(W_T + he_i \mathbb{1}_{[t_j,T]})\right) |\mathcal{F}_{t_j}\right]. \end{aligned}$$

In addition, for $\omega \in D([0,T], \mathbb{R}^d)$, a jump of size $\epsilon \in \mathbb{R}^d$ at time t_j of ω corresponds to a jump of size $\sigma(t_j, {}_nS_{t_j}(\omega))\epsilon \in \mathbb{R}^d$ at time t_{j+1} of ${}_nS(\omega)$ by definition of the Euler scheme ${}_nS$ (a jump at t_j of ω does not affect the value of ${}_nS(\omega)$ at t_j: ${}_nS$ has 'predictable' dependence with respect to ω). So a natural estimator of $\phi_n = (\phi_n^i(t_j), 1 \le i \le d)$ at time t_j is given by:

$$\widehat{\phi_n^i}(t_j) := \frac{\mathbb{E}\left[g\left({}_nS^{(t_{j+1}, he_i)}\right) |\mathcal{F}_{t_j}\right] - \mathbb{E}\left[g\left({}_nS^{(t_{j+1}, -he_i)}\right) |\mathcal{F}_{t_j}\right]}{2h}, \tag{3.23}$$

where ${}_nS^{(t_{j+1},he_i)}$ (resp. ${}_nS^{(t_{j+1},-he_i)}$) is the piecewise constant Euler-Maruyama scheme of S with a jump of size he_i (resp. $-he_i$) at time t_{j+1}, i.e. ${}_nS^{(t_{j+1},\pm he_i)}$ is constant in each interval $[t_k,t_{k+1})$ for $0\le k\le n-1$ with ${}_nS^{(t_{j+1},\pm he_i)}(0)=s_0$. For $k\neq j$,

$$ {}_nS^{(t_{j+1},\pm he_i)}(t_{k+1}) = {}_nS^{(t_{j+1},\pm he_i)}(t_k) + \sigma\left(t_k, {}_nS^{(t_{j+1},\pm he_i)}_{t_k}\right)(W(t_{k+1})-W(t_k)), $$

and

$$ {}_nS^{(t_{j+1},\pm he_i)}(t_{j+1}) = {}_nS^{(t_{j+1},\pm he_i)}(t_j) + \sigma\left(t_j, {}_nS^{(t_{j+1},\pm he_i)}_{t_j}\right)(W(t_{j+1})-W(t_j)) \pm he_i. $$

This might seem quite similar to the classical computation of the delta at time t of an option in Markovian case where we compute the sensitivity of the option price with respect to a small jump of S at t. Here we show that it can be extended to more general non-Markovian models, but with a slight difference: to compute the delta at time t_j, we perturb the path of S at a later time t_{j+1}. This point is actually crucial to prove that the 'delta' is well defined in our case, i.e. F_n is vertically differentiable (theorem 3.7).

As shown in the previous section, the error of approximation of ϕ by ϕ_n in discrete time is of order $n^{-\frac{1}{2}}$ (we neglect the term $\log n$ in the error). Let M be the number of simulations in the Monte-Carlo method. It is well known that the error due to the Monte-Carlo method combined with the centered difference scheme is of order $M^{-\frac{1}{2}}$ if we take the same sequence of random numbers to simulate $\mathbb{E}[g({}_nS^{(t_{j+1},he_i)})|\mathcal{F}_t]$ and $\mathbb{E}[g({}_nS^{(t_{j+1},-he_i)})|\mathcal{F}_t]$ and h is chosen to be of order $M^{-\frac{1}{4}}$ (see for example [31, 32, 49]). In summary if we take $M\simeq n\simeq h^{-\frac{1}{4}}$, then the replication error of the option should be of order $M^{-\frac{1}{2}}\simeq n^{-\frac{1}{2}}$.

3.7.1 Numerical examples

We end this section with several numerical examples. We start with a simple case where S is a $\mathbb{R}$-valued process which follows the Black-Scholes model. We always assume that the interest rate is zero, so S satisfies: for

$t \in [0, T]$,

$$\frac{dS(t)}{S(t)} = \sigma dW(t), \quad S_0 = s_0 \in \mathbb{R}$$

with $\sigma > 0$ and W a one-dimensional standard Brownian motion. We consider a lookback option of maturity T with payoff $g(S_T) = S(T) - \min_{t \in [0,T]} S(t)$. Clearly g is Lipschitz continuous with respect to the supremum norm $\| \cdot \|$.

Since the computational complexity for such dynamic hedging problems is of order $M \cdot n^2$, we will take n much smaller than M in our numerical implementations although from a theoretical point of view, n and M should be of the same order.

In our example, we assume that the underlying process S starts from $s_0 = 100$, and the maturity T of the option is equal to 1. First we take the number of time discretization $n = 100$, the number of Monte-Carlo simulation $M = 10^4$, and $h = 0.1$. We show in figure 3.1 the hedging error, for a given scenario, of the lookback option during its entire life with volatility σ equal to 0.2. We compute, at each time discretization, the Monte-Carlo price based on all past information, and the value of the portfolio which replicates the option, i.e. with initial value equal to the option price, and with variation between two time steps equal to the increments of S multiplied by the delta given by expression (3.23) (we assume $r = 0$).

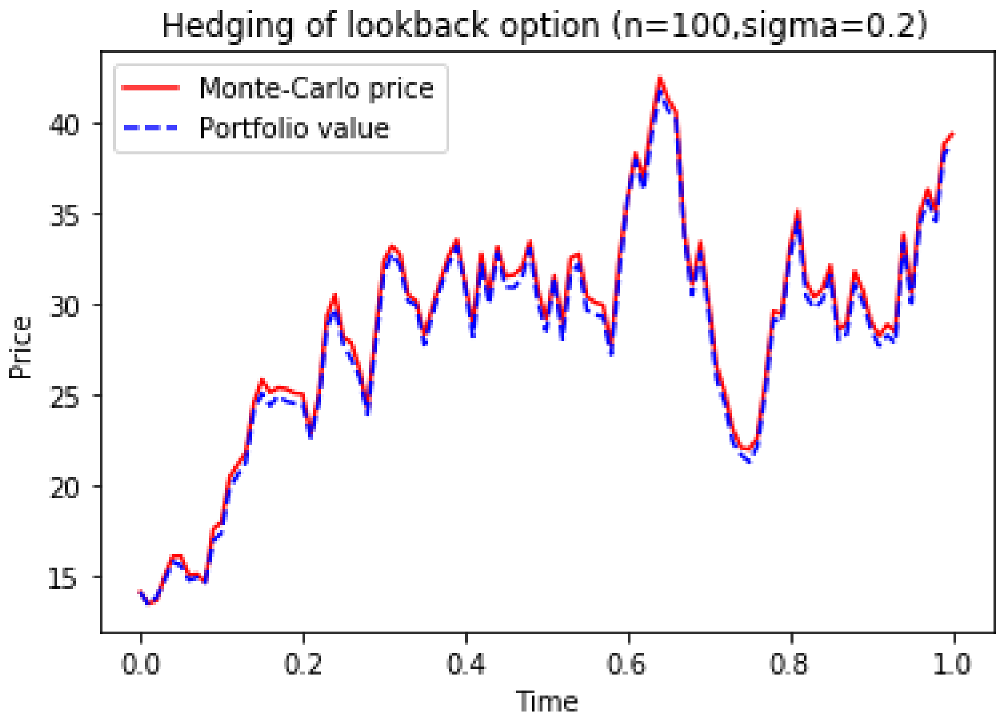

Figure 3.1: Hedging error of lookback option with $n = 100$

We observe from figure 3.1 that in this case, the two curves coincide well with each other, which means that the delta hedging strategy replicates well the option in this specific scenario. The following table provides the tracking error in this case:

	Tracking error at $T = 1$	Price of option	Relative error
$n = 100$	-0.5483	14.0994	3.89%

We can also increase the number of time discretization n to $n = 500$, and we obtain a similar result (figure 3.2). In this case, the terminal hedging error is equal to -0.3226, corresponding to a relative error of 2.22% with respect to the option price.

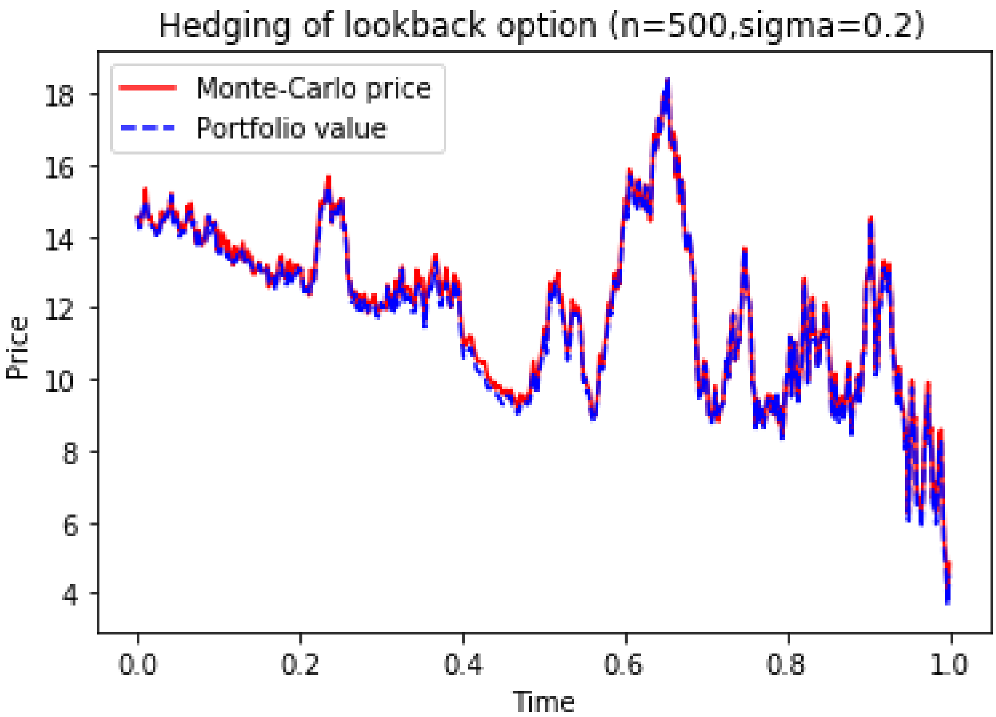

Figure 3.2: Hedging error with $n = 500$ and $\sigma = 0.2$

Clearly one specific path of scenario does not provide much information on the efficiency of our hedging strategy. We present in the following histogram (figure 3.3) the tracking error at time $T = 1$ of 100 paths of scenario. In these simulations, we take $n = 100$ and $\sigma = 0.2$. We also compute the mean and the variance of these tracking errors, which are given in the following table:

	Tracking errors
mean	0.0495
variance	1.1108

Table 3.1: Mean and variance of tracking error of 100 scenarios for hedging strategy computed using Euler and functional approximation

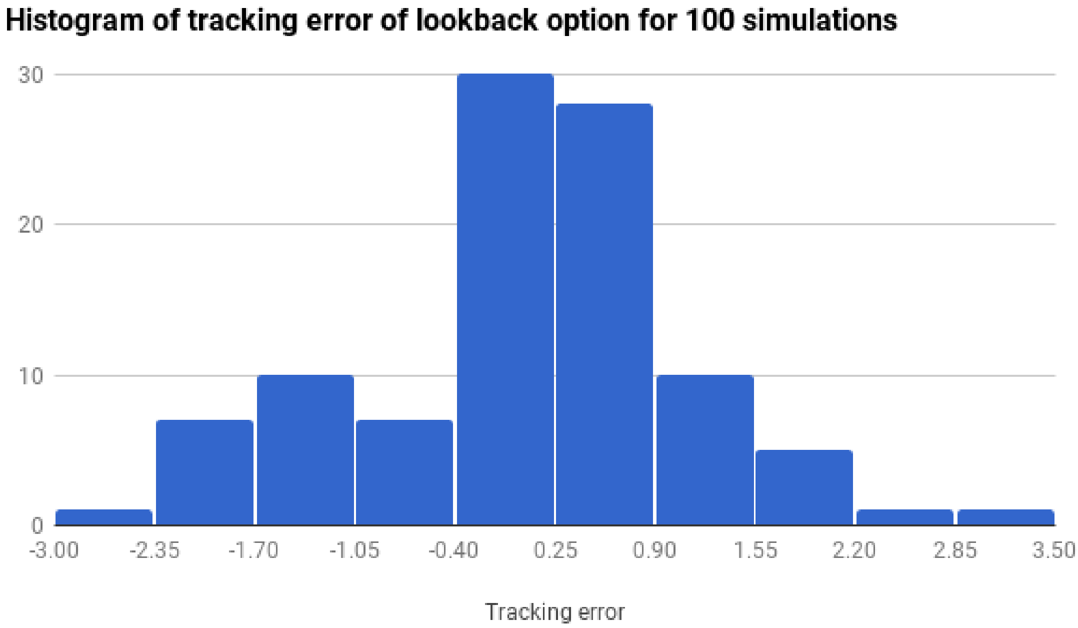

Figure 3.3: Histogram of tracking error for 100 simulations

We observe that even the variance of the tracking errors might seem quite important, our method works generally well in this simple case provided that the number of time discretization and the number of simulations in the Monte-Carlo method are not very significant.

We now move to a more sophisticated example in which the volatility at time t of the underlying S might depend on the whole path of S up to time t. Inspired by the path-dependent volatility model proposed by Hobson and Rogers [38], but for the reason of numerical simplicity, we assume that the underlying S follows the following dynamics (we always assume that $r = 0$):

$$dS(t) = \sigma_0 S(t)(1 + k\sigma(t, S_t, \lambda))dW(t), \quad S_0 = s_0 \in \mathbb{R}$$

with σ_0, k and λ positive constants, and the functional σ defined by:

$$\sigma(t, S_t, \lambda) := \frac{\int_0^t e^{\lambda s}(S(t) - S(s))ds}{S(t)\int_0^t e^{\lambda s} ds}, \text{ and } \sigma(0, S_0, \lambda) := 0. \tag{3.24}$$

So $\sigma(t, S_t, \lambda)$ is the weighted average of increments of S over $[0, t]$. Our model differs from the Hobson-Rogers model in several aspects, mainly for the purpose of simplification. First instead of considering an infinite window $(-\infty, t]$ in which the weighted average is taken as in the Hobson-Rogers model, we assume that only the path of S in the interval $[0, t]$ can affect the volatility at time t. Secondly, we consider directly the increments of S instead of those of $\log S$ in the functional σ.

We consider a rainbow option with two underlying assets S^1 and S^2. We assume further that there might be cross-dependency between the volatility of S^1 and that of S^2, i.e. the volatility of S^1 may depend on the path of S^2 and vice versa. Our model is the following:

$$\begin{cases} dS^1(t) = \sigma_0 S^1(t)(1 + k_1\sigma(t, S^1_t, \lambda) + k_2\sigma(t, S^2_t, \lambda))dW(t), \quad S^1_0 = s^1_0 \in \mathbb{R} \\ dS^2(t) = \sigma_0 S^2(t)(1 + k_1\sigma(t, S^2_t, \lambda) + k_2\sigma(t, S^1_t, \lambda))dB(t), \quad S^2_0 = s^2_0 \in \mathbb{R} \end{cases}$$

where σ_0, k_1, k_2 and λ are all positive constants, W and B two standard Brownian motions with a constant correlation ρ, i.e. $d\langle W, B\rangle(t) = \rho dt$, and σ the functional defined in (3.24). Clearly we can take a different σ_0, k_1, k_2 or λ for S^1 and S^2, but here for simplicity, we assume that the two underlyings share the same model parameters.

We decompose the Brownian motion B into two independent Brownian motions W and $W^\perp$: $B = \rho W + \sqrt{1-\rho^2}W^\perp$. Let $S := (S^1, S^2) \in \mathbb{R}^2$ and $\widetilde{W} := (W, W^\perp) \in \mathbb{R}^2$. Our initial model can be rewritten as:

$$dS(t) = \widetilde{\sigma}(t, S_t)d\widetilde{W}(t), \quad S_0 = s_0 = (s^1_0, s^2_0) \in \mathbb{R}^2$$

with $\widetilde{\sigma}(t, S_t) = \sigma_0 \times$

$$\begin{pmatrix} S^1(t)(1 + k_1\sigma(t, S^1_t) + k_2\sigma(t, S^2_t)) & 0 \\ \rho S^2(t)(1 + k_1\sigma(t, S^2_t) + k_2\sigma(t, S^1_t)) & \sqrt{1-\rho^2}S^2(t)(1 + k_1\sigma(t, S^2_t) + k_2\sigma(t, S^1_t)) \end{pmatrix}$$

(we omit the parameter λ in the functional σ).

Consider now a rainbow option on $S = (S^1, S^2)$ whose payoff g at maturity T is given by:

$$g(S_T) := \left(\max_{t\in[0,T]}(S^1(t) - S^2(t)) - K\right)_+$$

where $K > 0$ is the strike of the option.

In our simulation, we take $s_0^1 = s_0^2 = 100$, $T = 1$, $\sigma_0 = 0.2$, $\lambda = 10$, $k_1 = 1$, $k_2 = 0.4$, $K = 10$ and $\rho = 0.6$. We always take the number of time discretization $n = 100$, the number of simulations in the Monte-Carlo method $M = 10^4$, and $h = 0.1$. We show in figure 3.4 the tracking error of this option for one specific path of S. Again the two curves represent the option price obtained by the Monte-Carlo method and the value of the portfolio which replicates the option using the delta proposed in (3.23) at each time step.

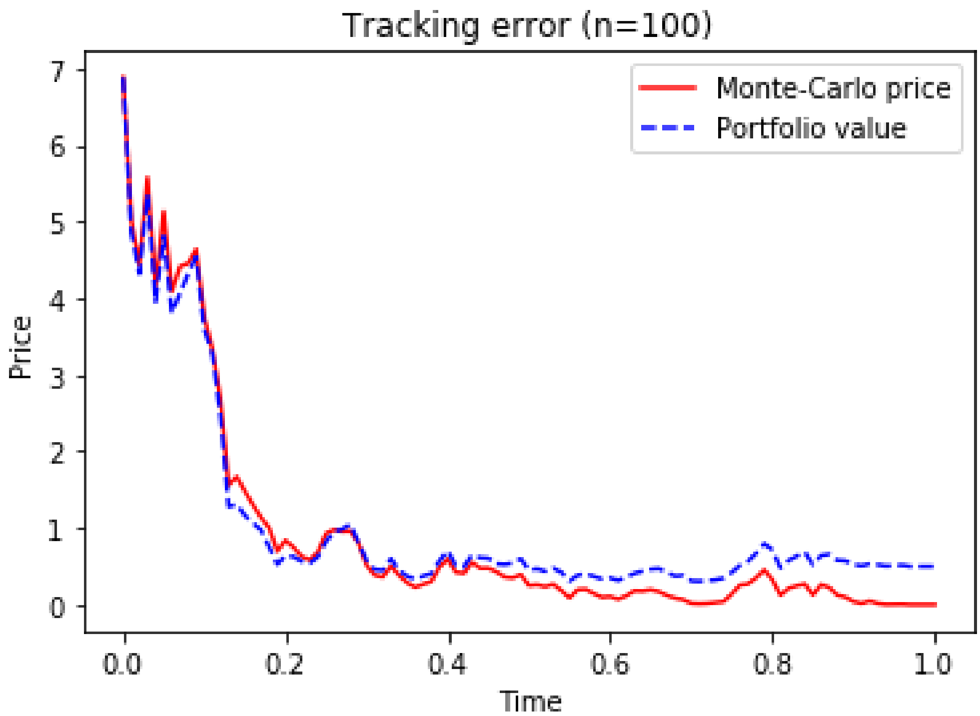

Figure 3.4: Hedging error of rainbow option with $n = 100$

From this figure, we observe that the gap between two curves is more considerable than in our first example with the Black-Scholes model, especially when close to the maturity. For this scenario, the tracking error at $T = 1$ is equal to 0.4999, corresponding to a relative error of 7.26% (the price of the option at $t = 0$ is equal to 6.8827).

We show in figure 3.5 another example of scenario with $n = 200$ now. The final tracking error in this case is 0.4132, corresponding to a relative error of 6.09%.

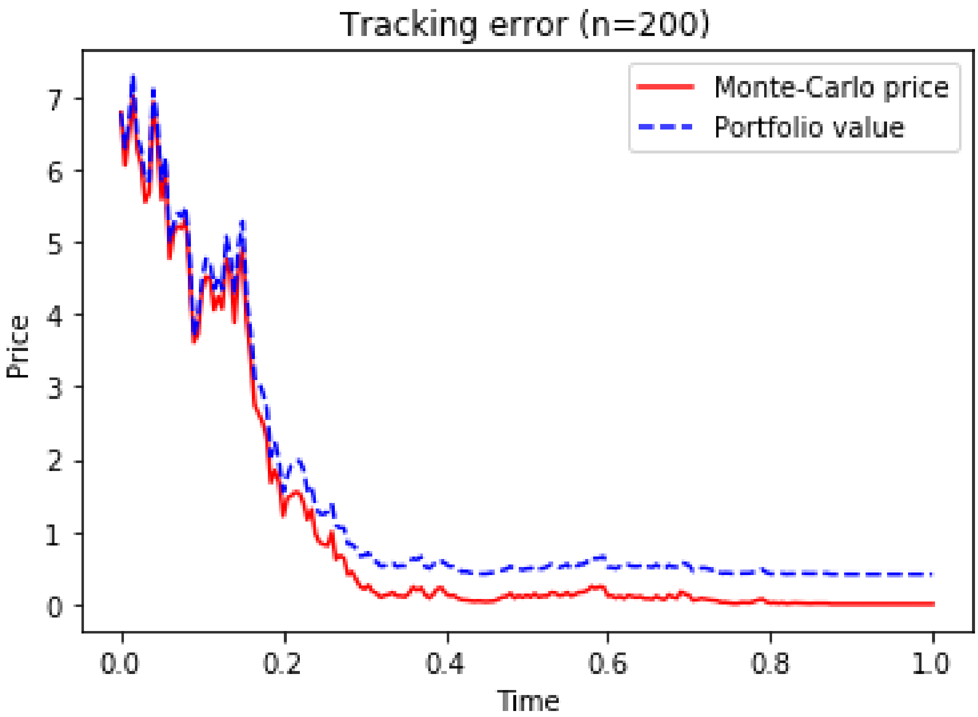

Figure 3.5: Hedging error of rainbow option with $n = 200$

We show finally, as we have done with the Black-Scholes model, the tracking error of 100 paths of scenario, presented in figure 3.6, as well as the mean and the variance of these tracking errors.

	Tracking errors
mean	-0.0377
variance	1.2287

Table 3.2: Mean and variance of tracking error of 100 scenarios for hedging strategy computed using Euler and functional approximation

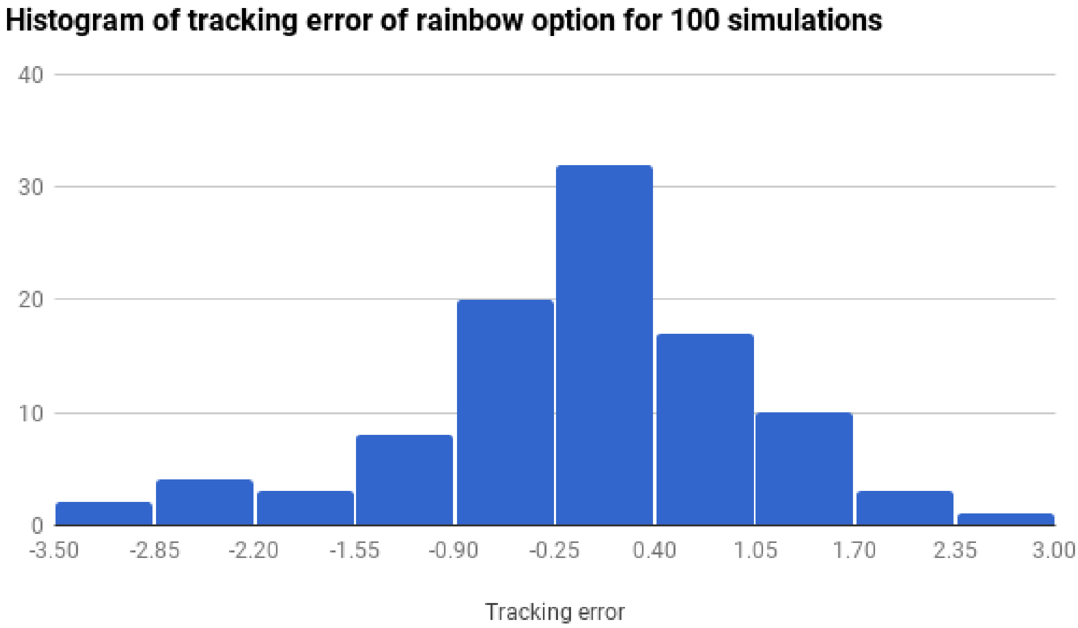

Figure 3.6: Histogram of tracking error for 100 simulations

We observe from these data that in this example, the variance of tracking errors is much more important (relative to the option price) compared to our first example with the Black-Scholes model. But in most of the cases, our method is still able to replicate the option with a reasonably small tracking error even the number of simulations in the Monte-Carlo method and the number of time discretization in our simulation are relatively small due to numerical complexity.

Chapter 4

Weak derivatives for non-anticipative functionals

In his seminal paper 'Calcul d'Itô sans probabilités' [27], Hans Föllmer proposed a pathwise derivation of the Itô formula, which was then extended to path-dependent functionals by Cont and Fournié [8], using a notion of pathwise directional derivative introduced by Dupire [20]. The associated functional calculus applies to functionals which possess certain directional derivatives in the strong sense i.e. at all paths in a certain set.

In this chapter, we introduce a notion of weak derivative for functionals which are not necessarily smooth in the sense of [8, 20]. To achieve this, we use the concept of pathwise quadratic variation along a sequence of partitions to define a bilinear form on the space of paths, and define a notion of weak derivative for functionals by duality with respect to this bilinear form. The whole approach involves only pathwise arguments and does not rely on any probabilistic notion. Nevertheless, we show that when applied to an Ito process, this notion of weak derivative coincides with the probabilistic weak derivative proposed by Cont and Fournié [10]. Our approach also provides a characterization of non-anticipative functionals which conserve the martingale property under a given probability measure, i.e. when applied to a given martingale, yield a martingale under the same probability measure.

4.1 Introduction

As discussed in section 2.2, Cont and Fournié introduced in [10] a notion of weak vertical derivative ∇_X (Theorem 2.9) with respect to a square-integrable martingale, in a probabilistic framework. This operator is shown to be the adjoint of the Itô stochastic integral with respect to X, and yields martingale representation formula (Theorem 2.10) [10].

The main ingredients of this construction can be thought of as a Sobolev-type construction on path space with respect to the reference measure $d[X](t) \times d\mathbb{P}_X$. First, the classical vertical derivative operator ∇_X defines an isometry (2.12) between two normed vector spaces (or more generally metric spaces) $(D(X), \|\cdot\|_{\mathcal{M}^2(X)})$ and $(\mathcal{L}^2(X), \|\cdot\|_{\mathcal{L}^2(X)})$:

$$\|Y\|^2_{\mathcal{M}^2(X)} = \|\nabla_X Y\|^2_{\mathcal{L}^2(X)}, \quad \forall Y \in D(X). \tag{4.1}$$

Now let $Y \in \overline{D(X)}$, the closure of $D(X)$ with respect to $\|\cdot\|_{\mathcal{M}^2(X)}$. By definition of $\overline{D(X)}$, there exists a sequence of $Y_n \in D(X)$ such that $\|Y_n - Y\|_{\mathcal{M}^2(X)} \underset{n\to\infty}{\to} 0$. In particular, $(Y_n)_{n\geq 1}$ is a Cauchy sequence, and so is $(\nabla_X Y_n)_{n\geq 1}$ using the isometry (4.1). As the space $\mathcal{L}^2(X)$ is complete, the sequence $(\nabla_X Y_n)_{n\geq 1}$ converges to some element of $\mathcal{L}^2(X)$, which is defined as the weak vertical derivative $\nabla_X Y$ of Y with respect to X. What remains is to determine the space $\overline{D(X)}$, which proves in this case to be the whole space $\mathcal{M}^2(X)$ (lemma 2.8).

The main objective of this chapter is to propose a notion of weak vertical derivative for non-anticipative functionals without intervention of any probability notions. More precisely, for a given path $x \in D([0,T], \mathbb{R}^d)$, we would like to extend the vertical derivative operator along x, $\nabla_\omega F(.,x)$ defined initially for vertically differentiable functionals F to more general functionals. The reason we consider vertical derivatives along a fixed path x is that it is more convenient to manipulate a path $\nabla_\omega F(.,x)$ of $[0,T]$ then a whole functional $\nabla_\omega F$ which requires much more information. This is in the same spirit of the construction in [10] in which the derivative operator ∇_X is extended for a given process X. We shall also consider, later in this chapter, functionals which are weakly vertically differentiable along a set of paths.

The probabilistic construction in [10] relies on the Ito isometry formula. Similarly, our construction relies on an isometry formula which makes use of the notion of pathwise quadratic variation (Definition 1.4) proposed by Föllmer [27]. However, contrary to the Itô isometry in probability, this pathwise isometry formula does not always hold for all (vertically differentiable) functionals. So we will need to limit ourselves to a subspace of smooth functionals. More importantly, the space $Q^\pi([0,T],\mathbb{R}^d)$ of paths with finite pathwise quadratic variation along a given sequence of partitions π is not a vector space (see Remark 1.5), which makes this pathwise construction of weak vertical derivatives technically more involved than its probabilistic counterpart. The concept of pathwise isometry for integrals of smooth functionals has been also studied by Ananova and Cont [1] in parallel with this work, but using a different approach and under stronger regularity assumptions.

The rest of this chapter is organized as follows. In section 4.2, using the notion of pathwise quadratic variation (definition 1.4), we extend the pathwise isometry formula initially proposed for functions (see for example [71]) to cylindrical functionals (definition 1.15), and show why it might not hold for general smooth functionals. We introduce, in section 4.3, a notion of generalized (pathwise) quadratic variation. The advantage of this notion is that, contrary to the space $Q^\pi([0,T],\mathbb{R}^d)$, the space $\hat{Q}^\pi([0,T],\mathbb{R}^d)$ of paths of finite generalized quadratic variation along π is a vector space, on which a semi-norm can be defined. We then construct, in section 4.4, weak pathwise vertical derivatives along a given path x for functionals which are not vertically differentiable but can be approximated by smooth functionals in some sense. In section 4.5, we prove that if X is square-integrable martingale, any square-integrable $\mathcal{F}^X$-martingale can be represented as a functional of X which is weakly vertically differentiable along almost all paths of X, and the weak vertical derivative of this functional along X coincides with the probabilistic weak derivative ∇_X of Cont and Fournié [10]. Finally, in Section 4.6, we obtain a characterization of functionals which conserve the martingale property when applied to a reference martingale.

4.2 Pathwise isometry formula

As we have shown in section 2.2, one of the main ingredients for extending the operator ∇_X to the space $\mathcal{M}^2(X)$ is the Ito isometry formula (4.1), which involves a probability measure $\mathbb{P}$. In a pathwise setting, we can no longer use the same norms as in (4.1) since the probability measure $\mathbb{P}$ is involved in both norms. However, the notion of pathwise quadratic variation defined in definition 1.4 seems relevant to our pathwise setting.

Throughout this section, we fix a sequence of partitions $\pi = (\pi_m)_{m\geq 1}$ with $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ and a continuous path $x \in C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$. Recall first a simple result of pathwise quadratic variation given in [66]:

Lemma 4.1. *Let $z \in C([0,T],\mathbb{R})$ such that $[z]_\pi(T) = 0$. Then for $x \in C([0,T],\mathbb{R})$ and $t \in [0,T]$, the quadratic variation $[x]_\pi(t)$ exists if and only if $[x+z]_\pi(t)$ exists. And in this case, we have $[x]_\pi(t) = [x+z]_\pi(t)$ for all $t \in [0,T]$.*

Consider now a function $f : \mathbb{R}^d \to \mathbb{R}$. The following result shows that if f is smooth, then the path $t \mapsto f(x(t))$ also has finite pathwise quadratic variation along the same sequence of partitions π.

Proposition 4.1 (Pathwise isometry formula for functions)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable function, and $x \in C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$. Then the path $f(x(\cdot))$ has finite quadratic variation which is given by: for $t \in [0,T]$,*

$$
\begin{aligned}
[f(x(\cdot))]_\pi(t) &= \left[\int_0^{\cdot} \nabla f(x(s)) \cdot d^\pi x(s)\right]_\pi (t) \\
&= \int_0^t \operatorname{tr}\left(\nabla(f(x(s))\, {}^t\nabla f(x(s)) d[x]_\pi(s)\right). \qquad (4.2)
\end{aligned}
$$

This result may be found in [71] for example; we provide a succinct proof below for the sake of completeness. In fact, the proposition holds even if f is only once continuously differentiable. However, we will still assume f to be at least twice continuously differentiable (and in the functional case $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$) in order to apply the Itô formula for later applications.

Proof. Recall first the pathwise Itô formula for functions (proposition 1.1) in continuous case: for $t \in [0,T]$, we have:

$$f(x(t)) = f(x(0)) + \int_0^t \nabla f(x(s)) \cdot d^\pi x(s) + \frac{1}{2}\int_0^t \mathrm{tr}\left(\nabla^2 f(x(s)) d[x]_\pi(s)\right). \tag{4.3}$$

Remark that the integral with respect to $d[x]_\pi$ in (4.3) has zero quadratic variation as it has finite variation. By lemma 4.1, the quadratic variation (if exists) of the path $f(x(\cdot))$ and that of $\int_0^{\cdot} \nabla f(x(s)) \cdot d^\pi x(s)$ shall be the same.

Let $m \in \mathbb{N}$. We note $\delta x_i^m := x(t_{i+1}^m) - x(t_i^m)$. By a Taylor expansion of second order, we have:

$$f(x(t_{i+1}^m)) - f(x(t_i^m)) = \nabla f(x(t_i^m)) \cdot \delta x_i^m + \frac{1}{2}\mathrm{tr}\left(\nabla^2 f(x(t_i^m))(\delta x_i^m)\, {}^t(\delta x_i^m)\right) + \epsilon_i^m |\delta x_i^m|^2 \tag{4.4}$$

with $\epsilon_i^m \underset{m\to\infty}{\to} 0$.

Now taking the square of both sides in (4.4), summing up for all $t_i^m \in \pi_m$ and taking the limit when m tends to infinity, we can readily check that the only term which does not tend to zero in the right-hand side of (4.4) is the square of $\nabla f(x(t_i^m)) \cdot \delta x_i^m$. So the quadratic variation of $f(x(\cdot))$ (if exists) is:

$$[f(x(.))]_\pi(t) = \lim_{m\to\infty} \sum_{t_{i+1}^m \le t} \mathrm{tr}\left(\nabla f(x(t_i^m))\, {}^t\nabla f(x(t_i^m))(\delta x_i^m)\, {}^t(\delta x_i^m)\right). \tag{4.5}$$

The limit on the right-hand side of (4.5) exists indeed, and is equal to $\int_0^t \mathrm{tr}(\nabla f(x(s))\, {}^t\nabla f(x(s)) d[x]_\pi(s))$ using lemma 1.17 and the definition of $[x]_\pi$. □

The extension of the pathwise isometry formula (4.2) to the case of non-anticipative functionals is subtle. A key ingredient in the proof of proposition

4.1 is a Taylor expansion. For a (twice) vertically differentiable functional $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$, we dispose of a Taylor expansion along piecewise-constant paths. Replacing x by a piecewise-constant approximation x^m along the partition π_m defined by

$$x^m(t) := \sum_{i=0}^{k(m)-1} x(t_{i+1}^m -)\mathbb{1}_{[t_i^m, t_{i+1}^m)}(t) + x(T)\mathbb{1}_{\{T\}}(t) \tag{4.6}$$

we obtain a convergent approximation:

Proposition 4.2. *Let $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ be a $\mathbb{R}-$valued non-anticipative functional. We have, for $t \in [0,T]$,*

$$\begin{aligned} &\sum_{t_{i+1}^m \le t} \Big((F(t_{i+1}^m, x^m_{t_{i+1}^m -}) - F(t_i^m, x^m_{t_i^m -}) \Big)^2 \\ &\underset{m\to\infty}{\to} \int_0^t \operatorname{tr}\left((\nabla_\omega F(s, x_s)\, {}^t\nabla_\omega F(s, x_s) d[x]_\pi(s) \right). \end{aligned} \tag{4.7}$$

Proof. First we proceed as in the proof of the pathwise change of variable formula (theorem 1.16) by decomposing $F(t_{i+1}^m, x^m_{t_{i+1}^m -}) - F(t_i^m, x^m_{t_i^m -})$ into two terms representing respectively the horizontal and vertical perturbations of a path:

$$\begin{aligned} & F(t_{i+1}^m, x^m_{t_{i+1}^m -}) - F(t_i^m, x^m_{t_i^m -}) \\ = \;& (F(t_{i+1}^m, x^m_{t_{i+1}^m -}) - F(t_i^m, x^m_{t_i^m})) + (F(t_i^m, x^m_{t_i^m}) - F(t_i^m, x^m_{t_i^m -})). \end{aligned} \tag{4.8}$$

The first term in (4.8) can be written as an integral of the horizontal derivative:

$$F(t_{i+1}^m, x^m_{t_{i+1}^m -}) - F(t_i^m, x^m_{t_i^m}) = \int_{t_i^m}^{t_{i+1}^m} \mathcal{D}F(u, x^m_{t_i^m}) du.$$

Let $\delta x_i^m = x(t_{i+1}^m) - x(t_i^m)$. Next, using a second order Taylor expansion of $e \mapsto F(t_i^m, x^m_{t_i^m -} + e 1_{[t_i^m, \infty)})$, the second part of (4.8) can be written as:

$$\begin{aligned} F(t_i^m, x^m_{t_i^m}) - F(t_i^m, x^m_{t_i^m -}) &= \nabla_\omega F(t_i^m, x^m_{t_i^m -}) \cdot \delta x_i^m + \frac{1}{2}\operatorname{tr}\left(\nabla_\omega^2 F(t_i^m, x^m_{t_i^m -})(\delta x_i^m)\, {}^t(\delta x_i^m) \right) \\ &\quad + \epsilon_i^m |\delta x_i^m|^2 \end{aligned}$$

with $\epsilon_i^m \to_{m\to\infty} 0$.

By the same argument as in the proof of proposition 4.1, the left-hand side term in (4.7) tends to (if the following limit exists):

$$\lim_{m\to\infty} \sum_{t_{i+1}^m \le t} \operatorname{tr}\left(\nabla_\omega F(t_i^m, x^m_{t_i^m-})\, {}^t\nabla_\omega F(t_i^m, x^m_{t_i^m-})(\delta x_i^m)\, {}^t(\delta\omega_i^m)\right).$$

Since $\nabla_\omega F(t, x^m_{t-}) \to_{m\to\infty} \nabla_\omega F(t, x_t)$ for all $t \in [0,T]$ as $\nabla_\omega F$ is left-continuous (definition 1.11), we obtain (4.7) using again lemma 1.17. □

Observe that the left-hand side term in (4.7) differs in general from the pathwise quadratic variation of the path $t \mapsto F(t, x_t)$. Indeed, the approximation of x by the piecewise constant path x^m works well to derive the functional Ito formula as in this case, we only require the sequence $(F(t, x^m_{t-}))_{m\ge 1}$ converges to $F(t, x_t)$ for a given t, i.e. the sequence of paths $F(\cdot, x^m_{\cdot-})$ converges pointwise to $F(\cdot, x_\cdot)$. This pointwise convergence does not imply, for example,

$$\sum_{t_i^m \in \pi_m} \left(F(t_i^m, x_{t_i^m}) - F(t_i^m, x^m_{t_i^m-})\right)^2 \underset{m\to\infty}{\to} 0. \tag{4.9}$$

So proposition 4.2 does not tell anything about the pathwise quadratic variation $[F(\cdot, x_\cdot)]_\pi$, and we do not even know if it exists for any $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$. However, if we assume in addition that the functional F satisfies (4.9), then the pathwise quadratic variation of $F(\cdot, x_\cdot)$ exists and is equal to the right-hand side term of (4.7).

Corollary 4.1. *Let $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ be a $\mathbb{R}-$valued non-anticipative functional satisfying* (4.9). *Then the path $F(\cdot, x_\cdot)$ has finite pathwise quadratic variation along π, given by: for $t \in [0,T]$,*

$$[F(\cdot, x_\cdot)]_\pi(t) = \left[\int_0^\cdot \nabla_\omega F(s, x_s)\cdot d^\pi x(s)\right]_\pi (t) = \int_0^t \operatorname{tr}\left(\nabla_\omega F(s, x_s)\, {}^t\nabla_\omega F(s, x_s) d[x]_\pi(s)\right). \tag{4.10}$$

Proof. We observe first that:

$$\begin{aligned} & F(t_{i+1}^m, x_{t_{i+1}^m}) - F(t_i^m, x_{t_i^m}) \\ = \;& (F(t_{i+1}^m, x_{t_{i+1}^m}) - F(t_{i+1}^m, x^m_{t_{i+1}^m-})) + (F(t_{i+1}^m, x^m_{t_{i+1}^m-}) - F(t_i^m, x^m_{t_i^m-})) \\ & + (F(t_i^m, x^m_{t_i^m-}) - F(t_i^m, x_{t_i^m})). \end{aligned} \tag{4.11}$$

The only term which contributes to $[F(\cdot, x.)]_\pi(t)$ (if it exists) in (4.11) is the sum of $(F(t_{i+1}^m, x^m_{t^m_{i+1}-}) - F(t_i^m, x^m_{t^m_i-}))^2$ along π_m: the terms with $(F(t_{i+1}^m, x_{t^m_{i+1}}) - F(t_{i+1}^m, x^m_{t^m_{i+1}-}))^2$ and $(F(t_i^m, x^m_{t^m_i-}) - F(t_i^m, x_{t_i^m}))^2$ tends to zero by assumption (4.9), and the cross-product terms also tend to zero due to the Cauchy-Schwarz inequality. We conclude using proposition 4.2. □

To the best of our knowledge, contrary to the case of smooth functions, we do not know if the path $F(\cdot, x.)$ has finite pathwise quadratic variation of the form (4.10) for any non-anticipative functional $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ (see also [1] for other conditions on F under which (4.10) holds). To tackle this problem, one possible solution, as suggested by proposition 4.2, is to define, for a non-anticipative functional F, a 'pseudo' quadratic variation along x defined as a 'diagonal' limit of squared increments along π_n computed at piecewise constant approximations along π_n:

Definition 4.2. *A non-anticipative functional $F : \Lambda_T \to \mathbb{R}$ is said to have finite pseudo quadratic variation along x if for any $t \in [0, T]$, the limit*

$$[\widetilde{F(\cdot, x.)}]_\pi(t) := \lim_{m\to\infty} \sum_{t^m_{i+1}\le t} \Big(F(t^m_{i+1}, x^m_{t^m_{i+1}-}) - F(t^m_i, x^m_{t^m_i-})\Big)^2 < \infty$$

exists, where x^m is the piecewise constant approximation of x defined by (4.6).

Using this notion of pseudo quadratic variation, proposition 4.2 can be reformulated as: for any $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$, for $t \in [0, T]$,

$$[\widetilde{F(\cdot, x.)}]_\pi(t) = \int_0^t \mathrm{tr}\left(\nabla_\omega F(s, x_s)\, {}^t\nabla_\omega F(s, x_s) d[x]_\pi(s)\right). \tag{4.12}$$

The isometry formula (4.12) holds for all smooth functionals in $\mathbb{C}_b^{1,2}(\Lambda_T^d)$ without any further conditions. However, one major drawback of this notion of pseudo quadratic variation is that its definition requires the knowledge of F evaluated not only at x but also at x^m. So this notion is not a property of the path $F(\cdot, x.)$, but a property of the functional F. In other words, this notion depends on the values taken by the functional F not just along the path x in the 'neighborhood' of x. For example, consider a stochastic process X

which admits a functional representation with respect to a Brownian motion W: $X(t) = F(t, W_t)$ for some non-anticipative functional F. In general the only thing we know about F is its value along W, while $[\widetilde{F(\cdot, W_\cdot)}]$ seems to depend on the values of F along discrete approximations of W.

In the following, we still use the notion of pathwise quadratic variation, but we limit ourselves to a subset of $\mathbb{C}_b^{1,2}(\Lambda_T^d)$ such that the pathwise isometry formula (4.10) holds for all functionals in this subset. Corollary 4.1 provides one such subset: smooth functionals satisfying (4.9). We now define another subset of $\mathbb{C}_b^{1,2}(\Lambda_T^d)$ which is simpler to characterize than condition (4.9): the space of cylindrical functionals, and we show that the pathwise isometry formula (4.10) holds for such functionals. Recall first the definition of cylindrical functionals:

Definition 4.3. *A non-anticipative functional F is said to be cylindrical if there exists $0 \leq t_1 < t_2 < \cdots < t_n \leq T$ such that for all $x \in D([0,T], \mathbb{R}^d)$,*

$$F(t,x) = h(x(t) - x(t_n-))\mathbb{1}_{t>t_n} g(x(t_1-), x(t_2-), \cdots, x(t_n-)) \qquad (4.13)$$

for some continuous function $g \in \mathbb{C}(\mathbb{R}^{n\times d}, \mathbb{R})$ and some twice differentiable function $h \in C^2(\mathbb{R}^d, \mathbb{R})$ with $h(0) = 0$.

The condition $h(0) = 0$ ensures that if x is a continuous path, then the path $F(\cdot, x_\cdot)$ is also continuous. It is clear that all cylindrical functionals are smooth (i.e. $\in \mathbb{C}_b^{1,2}(\Lambda_T^d)$) [7].

Lemma 4.4. *Let F be a non-anticipative cylindrical functional of the form (4.13). Then $F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ with $\mathcal{D}F \equiv 0$, and for $j = 1, 2$,*

$$\nabla_\omega^j F(t,x) = \nabla^j h(x(t) - x(t_n-))\mathbb{1}_{t>t_n} g(x(t_1-), x(t_2-), \cdots, x(t_n-)).$$

We denote by $\mathbb{S}(\Lambda_T^d)$ the linear span of all cylindrical functionals of the form (4.13). Clearly we have $\mathbb{S}(\Lambda_T^d) \subset \mathbb{C}_b^{1,2}(\Lambda_T^d)$. And for any $F \in \mathbb{S}(\Lambda_T^d)$, F can be written as:

$$F(t,x) = \sum_{j=1}^{k} h_j(x(t) - x(s_j-))\mathbb{1}_{t>s_j} g_j(x(s_1-), x(s_2-), \cdots, x(s_j-)) \quad (4.14)$$

with $g_j \in \mathbb{C}(\mathbb{R}^{j\times d},\mathbb{R})$, $h_j \in \mathbb{C}^2(\mathbb{R}^d,\mathbb{R})$, $h_j(0)=0$ and $0 \le s_1 < s_2 < \cdots < s_k \le T$. We now show that any functional F in $\mathbb{S}(\Lambda_T^d)$ satisfies the pathwise isometry formula (4.10).

Proposition 4.3 (Pathwise isometry formula for cylindrical functionals)**.** *Let x be a continuous path in $C([0,T],\mathbb{R}^d)\cap Q^\pi([0,T],\mathbb{R}^d)$. For $F\in\mathbb{S}(\Lambda_T^d)$, the path $F(\cdot,x.)$ has finite quadratic variation along π, and we have: for $t\in[0,T]$,*

$$[F(\cdot,x.)]_\pi(t) = \int_0^t \mathrm{tr}\left(\nabla_\omega F(s,x_s)\,{}^t\nabla_\omega F(s,x_s)d[x]_\pi(s)\right). \tag{4.15}$$

Proof. Assume that F takes the form (4.14). We calculate the increments of the path $F(\cdot,x.)$ between two consecutive time grids t_i^m and t_{i+1}^m of the partition π_m. We distinguish two cases:

- If t_i^m and t_{i+1}^m are in the same interval $[s_j,s_{j+1})$ for some $1\le j\le k$ (we assume $s_{k+1}=T$), then for $t\in[t_i^m,t_{i+1}^m]\subset[s_j,s_{j+1})$, $F(t,\omega)$ can be viewed as a smooth function of $x(t)$ as other terms are constant in this interval. So we can apply a second order Taylor expansion between t_i^m and t_{i+1}^m, and we obtain using the same argument as in proposition 4.1:

$$\left(F(t_{i+1}^m,x_{t_{i+1}^m})-F(t_i^m,x_{t_i^m})\right)^2 \simeq \int_{t_i^m}^{t_{i+1}^m}\mathrm{tr}\left(\nabla_\omega F(s,x_s)\,{}^t\nabla_\omega F(s,x_s)d[x]_\pi(s)\right).$$

  If we sum all i in this category, and as $|\pi_m|$ decreases to zero, we have:

$$\sum_{t_{i+1}^m\le t,[t_i^m,t_{i+1}^m]\subset[s_j,s_{j+1})}\left(F(t_{i+1}^m,x_{t_{i+1}^m})-F(t_i^m,x_{t_i^m})\right)^2 \underset{m\to\infty}{\to} \int_0^t \mathrm{tr}\left(\nabla_\omega F(s,x_s)\,{}^t\nabla_\omega F(s,x_s)d[x]_\pi(s)\right).$$

- There exists $1\le j\le k$ such that $t_i^m < s_j \le t_{i+1}^m$. Since the path $F(\cdot,x.)$ is continuous, and there exists only a finite number of such i (at most k), we have:

$$\sum_{t_{i+1}^m\le t,t_i^m<s_j\le t_{i+1}^m}\left(F(t_{i+1}^m,x_{t_{i+1}^m})-F(t_i^m,x_{t_i^m})\right)^2 \underset{m\to\infty}{\to} 0.$$

We conclude combining these two cases. □

As the space $\mathbb{S}(\Lambda_T^d)$ is a vector space, we can always define, for $F, G \in \mathbb{S}(\Lambda_T^d)$, the quadratic covariation of the paths $F(\cdot, x.)$ and $G(\cdot, x.)$ by: for $t \in [0, T]$,

$$[F(\cdot, x.), G(\cdot, x.)]_\pi(t) := \frac{[(F+G)(\cdot, x.)]_\pi(t) - [F(\cdot, x.)]_\pi(t) - [G(\cdot, x.)]_\pi(t)}{2}.$$

Using the expression of quadratic variation for cylindrical functionals (4.15), we obtain immediately:

$$[F(\cdot, x.), G(\cdot, x.)]_\pi(t) = \int_0^t \operatorname{tr}\left(\nabla_\omega F(s, x_s)\, {}^t\nabla_\omega G(s, x_s) d[x]_\pi(s)\right). \tag{4.16}$$

The aim of this chapter is to define, using the isometry formula (4.15), a notion of weak pathwise vertical derivatives for non-anticipative functionals which are not necessarily smooth. One natural approach is to consider the closure of $\mathbb{S}(\Lambda_T^d)$ with respect to the pathwise quadratic variation along x, i.e. we consider the space $\widetilde{H}^\pi(\Lambda_T^d, x)$ of non-anticipative functionals:

$$\widetilde{H}^\pi(\Lambda_T^d, x) := \left\{ G, \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, [(G - F_n)(\cdot, x.)]_\pi(T) \underset{n\to\infty}{\to} 0 \right\}.$$

By the isometry formula (4.15), the vertical derivatives of F_n along x: $\nabla_\omega F_n(\cdot, x.)$ will converge to some path ϕ in the sense that:

$$\int_0^T \operatorname{tr}\left((\nabla_\omega F_n(s, x_s) - \phi(s))\, {}^t(\nabla_\omega F_n(s, x_s) - \phi(s)) d[x]_\pi(s)\right) \underset{n\to\infty}{\to} 0.$$

ϕ may thus be defined as the weak vertical derivative of G along x. However, one of the problems of this approach is that since $Q^\pi([0, T], \mathbb{R})$ is not a vector space (see remark 1.5), in the definition of the space $\widetilde{H}^\pi(\Lambda_T^d, x)$, even if we assume G has finite quadratic variation along x, it might not be the case for $G - F_n$. So it might seem difficult to characterize precisely the space of functionals $\widetilde{H}^\pi(\Lambda_T^d, x)$. To tackle this problem, we introduce the notion of 'generalized' pathwise quadratic variation in the following section, whose domain of definition is a vector space.

4.3 Generalized quadratic variation

As we have mentioned in the previous section, one drawback of the notion of pathwise quadratic variation along a fixed sequence of partitions π is that the set of càdlàg paths with finite quadratic variation along π: $Q^\pi([0,T],\mathbb{R})$ is not a vector space [67]. Let $x, y \in Q^\pi([0,T],\mathbb{R})$. The reason $x+y$ may not have finite quadratic variation, as explained in remark 1.5, stems from the fact that the cross-product terms may have an oscillating sign which prevents the convergence of the sequence $(q_m)_{m\geq 1}$ defined by:

$$q_m := \sum_{t_i^m \in \pi_m} (\delta x_i^m + \delta y_i^m)^2$$

with $\delta x_i^m := x(t_{i+1}^m) - x(t_i^m)$ and similarly for δy_i^m.

However, a simple application of the Cauchy-Schwarz inequality shows that the cross-product term $\sum_i \delta x_i^m \delta y_i^m$ is bounded by:

$$\left(\sum_i \delta x_i^m \delta y_i^m\right)^2 \leq \left(\sum_i |\delta x_i^m|^2\right)\left(\sum_i |\delta y_i^m|^2\right)$$

which converges when m tends to infinity since $x, y \in Q^\pi([0,T],\mathbb{R})$. This means that if $x, y \in Q^\pi([0,T],\mathbb{R})$, the sequence $(q_m)_{m\geq 1}$ is bounded even if it does not converge. This observation inspires us to define:

Definition 4.5 (Generalized quadratic variation of paths along a sequence of partitions). *Let $\pi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ be a sequence of partitions of $[0,T]$ with $|\pi_m| \underset{m\to\infty}{\to} 0$. The generalized quadratic variation of a càdlàg path $x \in D([0,T],\mathbb{R})$ along π is defined by: for $t \in [0,T]$,*

$$\widehat{[x]}_\pi(t) := \limsup_{m\to\infty} \sum_{t_{i+1}^m \leq t} \left(x(t_{i+1}^m) - x(t_i^m)\right)^2 \in \mathbb{R}_+ \cup \{+\infty\}. \tag{4.17}$$

We may drop the subscript in $\widehat{[x]}_\pi$ when the context is clear. We denote by

$$\widehat{Q}^\pi([0,T],\mathbb{R}) = \{x \in D([0,T],\mathbb{R}), \quad \forall t \in [0,T], \quad \widehat{[x]}_\pi(t) < \infty\}$$

the set of $\mathbb{R}-$valued càdlàg paths with finite generalized quadratic variation along π.

Observe that the set $\widehat{Q}^\pi([0,T],\mathbb{R})$ is much larger than $Q^\pi([0,T],\mathbb{R})$. Indeed, for a càdlàg path x to have finite generalized quadratic variation, we only require the sequence $\sum(x(t_{i+1}^m)-x(t_i^m))^2$ to be bounded, which is much weaker than requiring it to be convergent. Moreover we no longer need that $\widehat{[x]}_\pi$ admits a Lebesgue decomposition of the form (1.1) as in the definition of the pathwise quadratic variation.

The main advantage of this notion of generalized quadratic variation is that, contrary to the set $Q^\pi([0,T],\mathbb{R})$, the set of paths with finite generalized quadratic variation along a given sequence of partitions $\widehat{Q}^\pi([0,T],\mathbb{R})$ is a vector space.

Proposition 4.4. *$\widehat{Q}^\pi([0,T],\mathbb{R})$ is a vector space, and $\|\cdot\|_{QV} : x \mapsto \sqrt{\widehat{[x]}_\pi(T)}$ defines a semi-norm on $\widehat{Q}^\pi([0,T],\mathbb{R})$.*

Proof. Observe first that for $x,y \in \widehat{Q}^\pi([0,T],\mathbb{R})$, we have:

$$\widehat{[x+y]}_\pi(t) \leq 2\left(\widehat{[x]}_\pi(t) + \widehat{[y]}_\pi(t)\right) < \infty, \quad \forall t \in [0,T].$$

This follows immediately from the inequalities $(a+b)^2 \leq 2(a^2+b^2)$ and $\limsup_m(a_m+b_m) \leq \limsup_m a_m + \limsup_m b_m$. So $x+y \in \widehat{Q}^\pi([0,T],\mathbb{R})$, and $\widehat{Q}^\pi([0,T],\mathbb{R})$ is a vector space.

Now we prove that $\|\cdot\|_{QV}$ defines a semi-norm. For $x,y \in \widehat{Q}^\pi([0,T],\mathbb{R})$, we shall show that:

$$\|x+y\|_{QV} \leq \|x\|_{QV} + \|y\|_{QV}.$$

Note again $\delta x_i^m = x(t_{i+1}^m) - x(t_i^m)$ and similarly for y. For $m \in \mathbb{N}$, using the Cauchy-Schwarz inequality, we have:

$$\begin{aligned} & \sum_i (\delta x_i^m + \delta y_i^m)^2 \\ = \; & \sum_i (\delta x_i^m)^2 + \sum_i (\delta y_i^m)^2 + 2\sum_i \delta x_i^m \delta y_i^m \\ \leq \; & \sum_i (\delta x_i^m)^2 + \sum_i (\delta y_i^m)^2 + 2\sqrt{\left(\sum_i (\delta x_i^m)^2\right)\left(\sum_i (\delta y_i^m)^2\right)}. \end{aligned}$$

So

$$\begin{aligned}
\|x+y\|_{QV}^2 &= \limsup_m \sum_i (\delta x_i^m + \delta y_i^m)^2 \\
&\leq \limsup_m \sum_i (\delta x_i^m)^2 + \limsup_m \sum_i (\delta y_i^m)^2 + 2 \limsup_m \sqrt{\left(\sum_i (\delta x_i^m)^2\right)\left(\sum_i (\delta y_i^m)^2\right)} \\
&= \|x\|_{QV}^2 + \|y\|_{QV}^2 + 2\sqrt{\limsup_m \left(\sum_i (\delta x_i^m)^2\right)\left(\sum_i (\delta y_i^m)^2\right)} \\
&\leq \|x\|_{QV}^2 + \|y\|_{QV}^2 + 2\|x\|_{QV}\|y\|_{QV} = (\|x\|_{QV} + \|y\|_{QV})^2
\end{aligned}$$

where we have used in the last inequality: $\limsup_m (a_m b_m) \leq \limsup_m a_m \limsup_m b_m$ for a_m, b_m two non-negative sequences. □

Corollary 4.2. *Let $(x_n)_{n\geq 1}$ be a sequence of paths in $\widehat{Q}^\pi([0,T],\mathbb{R})$ and $x \in \widehat{Q}^\pi([0,T],\mathbb{R})$. If $\widehat{[x_n - x]}_\pi(T) \to_{n\to\infty} 0$, then the sequence $\widehat{[x_n]}_\pi(T)$ converges and its limit is $\widehat{[x]}_\pi(T)$.*

Proof. $\widehat{[x_n - x]}_\pi(T) \to_{n\to\infty} 0$ implies $\|x_n - x\|_{QV} \to_{n\to\infty} 0$. Thus we have:

$$\left|\widehat{[x_n]}_\pi(T) - \widehat{[x]}_\pi(T)\right| = \left|\|x_n\|_{QV}^2 - \|x\|_{QV}^2\right| \leq \|x_n - x\|_{QV}(\|x_n\|_{QV} + \|x\|_{QV}) \underset{n\to\infty}{\to} 0$$

as $\|\cdot\|_{QV}$ is a semi-norm and $\|x_n\|_{QV} + \|x\|_{QV}$ can be bounded independently of n. □

The notion of generalized quadratic variation can be easily extended to paths of dimension $d > 1$.

Definition 4.6. *A d-dimensional path $x = (x^1, \cdots, x^d) \in D([0,T],\mathbb{R}^d)$ is said to have finite generalized quadratic variation along π if $x^i \in \widehat{Q}^\pi([0,T],\mathbb{R})$ for all $i = 1, \cdots, d$. The matrix-valued function $\widehat{[x]} : [0,T] \to \mathbb{M}_d$ whose elements are given by:*

$$\widehat{[x]}_{ij}(t) := \frac{\widehat{[x^i + x^j]}(t) - \widehat{[x^i]}(t) - \widehat{[x^j]}(t)}{2}$$

is called the generalized quadratic variation of the path x.

We denote by $\widehat{Q}^{\pi}([0,T],\mathbb{R}^d)$ the set of $\mathbb{R}^d-$valued càdlàg paths with finite generalized quadratic variation along π.

Remark 4.7. *There are several differences between this definition and definition 1.4 of $\mathbb{R}^d$-valued paths with finite quadratic variation. First of all, as $\widehat{Q}^{\pi}([0,T],\mathbb{R})$ is a vector space, we no longer need to require $x^i + x^j \in \widehat{Q}^{\pi}([0,T],\mathbb{R})$ which is now automatically implied by $x^i, x^j \in \widehat{Q}^{\pi}([0,T],\mathbb{R})$. Secondly, in definition 4.6, the matrix-valued function $\widehat{[x]} : [0,T] \to \mathbb{M}_d$ does not necessarily take values in $\mathcal{S}_d^+$ as $\limsup$ is not a linear operator. And $\widehat{[x]}_{ij}(t)$ does not necessarily equal to:*

$$\limsup_{m} \sum_{t_{i+1}^m \le t} (x^i(t_{i+1}^m) - x^i(t_i^m))(x^j(t_{i+1}^m) - x^j(t_i^m)).$$

So the notion of generalized quadratic covariation is not well defined. Nevertheless if $x \in Q^{\pi}([0,T],\mathbb{R}^d)$, then the two definitions coincide, and we have $[x]_\pi = \widehat{[x]}_\pi$.

4.4 Weak pathwise vertical derivatives

The generalized pathwise quadratic variation defines a bilinear form and a semi-norm on its domain, which is a vector space. We now use the duality structure associated with this bilinear form to define the notion of weak pathwise vertical derivatives for non-anticipative functionals which are not necessarily smooth. We always fix a sequence of partition $\pi = (\pi_m)_{m\ge1}$, and a continuous path $x \in Q^{\pi}([0,T],\mathbb{R}^d)$.

We denote by $L^2([0,T],[x]_\pi)$ the space of $\mathbb{R}^d-$valued paths ϕ such that:

$$\int_0^T \mathrm{tr}\left(\phi(t)\, {}^t\phi(t) d[x]_\pi(t)\right) < \infty.$$

For two elements ϕ, ψ in $L^2([0,T],[x]_\pi)$, we define the following equivalence relation:

$$\phi \sim \psi \Longleftrightarrow \int_0^T \mathrm{tr}\left((\phi(t)-\psi(t))\, {}^t(\phi(t)-\psi(t)) d[x]_\pi(t)\right) = 0. \qquad (4.18)$$

Let $\mathbb{L}^2([0,T],[x]_\pi)$ be the quotient of the space $L^2([0,T],[x]_\pi)$ by the equivalence relation (4.18):

$$\mathbb{L}^2([0,T],[x]_\pi) := L^2([0,T],[x]_\pi)/\sim .$$

Endowed with the inner product: for $\phi, \psi \in \mathbb{L}^2([0,T],[x]_\pi)$,

$$\langle \phi, \psi \rangle_{L^2_x} := \int_0^T \mathrm{tr}\left(\phi(t)\, {}^t\psi(t) d[x]_\pi(t)\right),$$

$\mathbb{L}^2([0,T],[x]_\pi)$ is a Hilbert space.

Let F be a $\mathbb{R}-$valued non-anticipative functional. We define:

$$\|F\|_{x,QV} := \|F(\cdot, x.)\|_{QV} \in \mathbb{R}_+ \cup \{+\infty\}. \tag{4.19}$$

Denote by $\widehat{Q}(\Lambda_T^d, x)$ the space of non-anticipative functionals F such that $\|F\|_{x,QV} < \infty$. By proposition 4.4, $\widehat{Q}(\Lambda_T^d, \omega)$ is also a vector space, and $\|\cdot\|_{x,QV}$ defines a semi-norm on $\widehat{Q}(\Lambda_T^d, x)$.

Consider now a non-anticipative functional $F \in \mathbb{S}(\Lambda_T^d)$, by proposition 4.3, $F \in \widehat{Q}(\Lambda_T^d, x)$ (the limit superior in the definition of the generalized quadratic variation becomes simply a limit), which implies that $\mathbb{S}(\Lambda_T^d) \subset \widehat{Q}(\Lambda_T^d, x)$. Moreover, we have $\nabla_\omega F(\cdot, x.) \in \mathbb{L}^2([0,T],[x]_\pi)$, and

$$\|F\|_{x,QV} = \|\nabla_\omega F(\cdot, x.)\|_{L^2_x},$$

which shows that the vertical derivative of a functional along x defined on the space of cylindrical functionals, $\nabla_\omega(\cdot, x.) : (\mathbb{S}(\Lambda_T^d), \|\cdot\|_{x,QV}) \to (\mathbb{L}^2([0,T],[x]_\pi), \|\cdot\|_{L^2_x})$ with:

$$\nabla_\omega(\cdot, x.)(F) := \nabla_\omega F(\cdot, x.)$$

is an isometry.

Since $\mathbb{L}^2([0,T],[x]_\pi)$ is a Hilbert space, this implies that the map $\nabla_\omega(\cdot, x.)$ admits a unique extension defined on the closure $H^\pi(\Lambda_T^d, x)$ of $\mathbb{S}(\Lambda_T^d)$ with respect to the semi-norm $\|\cdot\|_{x,QV}$:

$$H^\pi(\Lambda_T^d, x) := \left\{ G \in \widehat{Q}(\Lambda_T^d, x), \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, \|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0 \right\}. \tag{4.20}$$

$H^\pi(\Lambda_T^d, x)$ is also a vector space, and $\nabla_\omega(\cdot, x.) : (H^\pi(\Lambda_T^d, x), \|\cdot\|_{x,QV}) \to (\mathbb{L}^2([0,T],[x]_\pi), \|\cdot\|_{L^2_x})$ still defines an isometry.

Proposition 4.5. *Let G be a functional in $H^\pi(\Lambda_T^d, x)$. There exists a unique $\phi \in \mathbb{L}^2([0,T],[x]_\pi)$ satisfying the following property: for any sequence $F_n \in \mathbb{S}(\Lambda_T^d)$ such that $\|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0$, $\nabla_\omega F_n(\cdot, x.)$ converges to ϕ in $(\mathbb{L}^2([0,T],[x]_\pi), \|\cdot\|_{L_x^2})$. And we have:*

$$\|G\|_{x,QV} = \|\phi\|_{L_x^2}. \tag{4.21}$$

We define $\phi = \nabla_x^\pi G(\cdot, x.)$ as the weak vertical derivative of G along x.

Remark 4.8. *The weak vertical derivative along a path x can be viewed as the 'inverse' of the pathwise integral with respect to x defined in* (1.12)*. Indeed, consider a non-anticipative functional $F \in \mathbb{S}(\Lambda_T^d)$, we define another non-anticipative functional G such that for any $\omega \in Q^\pi([0,T],\mathbb{R}^d)$, $G(t,\omega) := \int_0^t \nabla_\omega F(s,\omega_s)\cdot d^\pi\omega(s)$ (the integral is well defined as $F \in \mathbb{S}(\Lambda_T^d) \subset \mathbb{C}_b^{1,2}(\Lambda_T^d)$). Then $G \in H^\pi(\Lambda_T^d, x)$, and $\nabla^\pi G(\cdot, x.) = \nabla_\omega F(\cdot, x.)$ in $\mathbb{L}^2([0,T],[x]_\pi)$ (we may take $F_n \equiv F$ in the definition of the space $H^\pi(\Lambda_T^d, x)$). And for $F \in \mathbb{S}(\Lambda_T^d)$, its weak vertical derivative along x coincides with its (strong) vertical derivative in the sense of Dupire, evaluated at x, i.e. $\nabla^\pi F(\cdot, x.) = \nabla_\omega F(\cdot, x.)$ in $\mathbb{L}^2([0,T],[x]_\pi)$.*

As we shall see in the following section, in most of the cases, the lim sup in the definition of the generalized quadratic variation reduces to a limit, and $\|\cdot\|_{x,QV}^2$ is nothing but the quadratic variation of a functional along x.

Since lim sup is not a linear operator, the notion of generalized quadratic covariation is not well defined for all pairs of paths in $\widehat{Q}^\pi([0,T],\mathbb{R})$. However, as shown in (4.16), the generalized quadratic covariation along x is well defined for cylindrical functionals as all limits superior in this case are simply limits. And since $H^\pi(\Lambda_T^d, x)$ is the closure of $\mathbb{S}(\Lambda_T^d)$ with respect to the semi-norm $\|\cdot\|_{x,QV}$, we may also equip $H^\pi(\Lambda_T^d, x)$ with a degenerate inner product $\langle\cdot,\cdot\rangle_{x,QV}$ defined by: for $G_1, G_2 \in H^\pi(\Lambda_T^d, x)$ with two sequences $(F_n^{(1)})$ and $(F_n^{(2)})$ in $\mathbb{S}(\Lambda_T^d)$ such that $\|G_1 - F_n^{(1)}\|_{x,QV} \to 0$ and $\|G_2 - F_n^{(2)}\|_{x,QV} \to 0$,

$$\langle G_1, G_2\rangle_{x,QV} := \lim_{n\to\infty} \left[F_n^{(1)}(\cdot, x.), F_n^{(2)}(\cdot, x.)\right]_\pi (T). \tag{4.22}$$

(4.22) is well defined as

$$[F_n^{(1)}(\cdot,x.),F_n^{(2)}(\cdot,x.)]_\pi(T)=\frac{\|F_n^{(1)}+F_n^{(2)}\|^2_{x,QV}-\|F_n^{(1)}\|^2_{x,QV}-\|F_n^{(2)}\|^2_{x,QV}}{2},$$

and by corollary 4.2, when taking the limit when n tends to infinity, we obtain:

$$\langle G_1,G_2\rangle_{x,QV}=\frac{\|G_1+G_2\|^2_{x,QV}-\|G_1\|^2_{x,QV}-\|G_2\|^2_{x,QV}}{2}.$$

This shows that the limit on the right-hand side of (4.22) exists and is independent of the sequences $(F_n^{(1)})$ and $(F_n^{(2)})$ chosen. Moreover, $\langle\cdot,\cdot\rangle_{x,QV}$ is a semi-definite bilinear form on $H^\pi(\Lambda_T^d,x)$, which leads to the following characterization of weak vertical derivatives:

Proposition 4.6. *Let $G\in H^\pi(\Lambda_T^d,x)$. The weak vertical derivative of G along x: $\nabla^\pi G(\cdot,x.)$ is the unique element of $\mathbb{L}^2([0,T],[x]_\pi)$ which satisfies:*

$$\forall F\in\mathbb{S}(\Lambda_T^d),\quad \langle F,G\rangle_{x,QV}=\langle\nabla_\omega F(\cdot,x.),\nabla^\pi G(\cdot,x.)\rangle_{L^2_x}. \tag{4.23}$$

Proof. Clearly using the definition (4.22) of $\langle\cdot,\cdot\rangle_{x,QV}$ and proposition 4.5, $\nabla^\pi G(\cdot,x.)$ satisfies (4.23). To prove the uniqueness, it suffices to show that $\{\nabla_\omega F(\cdot,x.),F\in\mathbb{S}(\Lambda_T^d)\}$ is dense in $\mathbb{L}^2([0,T],[x]_\pi)$. Let $C\in\mathbb{R}^d$ and $t_0\in[0,T]$. We consider functionals of the form: $F(t,\omega_t):=C\cdot(\omega(t)-\omega(t_0))\mathbb{1}_{t>t_0}$, for $\omega\in D([0,T],\mathbb{R}^d)$. Clearly such functionals are cylindrical, and $\nabla_\omega F(\cdot,x.)=C\mathbb{1}_{t>t_0}$. Since the set of simple functions is dense in $\mathbb{L}^2([0,T],[x]_\pi)$, $\{\nabla_\omega F(\cdot,x.),F\in\mathbb{S}(\Lambda_T^d)\}$ is also dense in $\mathbb{L}^2([0,T],[x]_\pi)$, which proves the uniqueness of $\nabla^\pi G(\cdot,x.)$. □

So far we have defined a weak vertical derivative of a functional G along a given path x. Actually the construction of this weak derivative only involves the value of G along x, i.e. the path $t\mapsto G(t,x_t)$. So this weak derivative is a local property, i.e.

$$F,G\in H^\pi(\Lambda_T^d,x),\quad G(.,x)=F(.,x)\quad\Rightarrow\nabla^\pi G(\cdot,x.)=\nabla^\pi F(\cdot,x.).$$

This echoes an anologous property of the (strong) vertical derivative [7, Sec. 5.4].

We shall now exploit this 'locality' property to define a weak derivative along a (closed) *set* of paths.

4.5 Application to functionals of stochastic processes

4.5.1 Weak differentiability along a set of paths

To apply the concepts above to path-dependent functionals of a stochastic process, we need to require the property of weak differentiability along any typical sample path of the process. For this purpose, we consider functionals which are weakly vertically differentiable along a set of paths A, which will then be chosen to be a set with full measure with respect to the law of some process.

Let $\pi = (\pi_m)_{m\geq 1}$ be a sequence of partitions of $[0,T]$. We still limit the paths along which the weak derivative is defined to continuous paths with finite quadratic variation along π, i.e. $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$.

Definition 4.9. *Let $\pi = (\pi_m)_{m\geq 1}$ be a sequence of partitions and $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$. We define $H^\pi(\Lambda_T^d, A)$ as the space of non-anticipative functionals which can be approximated by cylindrical functionals in quadratic variation along all paths in A:*

$$H^\pi(\Lambda_T^d, A) := \left\{ G \in \widehat{Q}^\pi(\Lambda_T^d, A), \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, \forall x \in A, \|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0 \right\} \tag{4.24}$$

where

$$\widehat{Q}^\pi(\Lambda_T^d, A) := \cap_{x\in A} \widehat{Q}^\pi(\Lambda_T^d, x).$$

Remark that $H^\pi(\Lambda_T^d, A)$ is a vector space, and the condition $G \in H^\pi(\Lambda_T^d, A)$ is stronger than G being weakly vertically differentiable along any $x \in A$ as in (4.24) we require the approximating $(F_n)_{n\geq 1}$ to be independent of $x \in A$.

4.5.2 Weak derivatives and martingale representation

As discussed in section 2.2, Cont and Fournié [10] developed a notion of weak vertical derivatives in a probabilistic framework, and derived a constructive martingale representation formula (Theorem 2.10). In this section, we show that the weak pathwise derivative we have constructed in Section 4.4, when applied to a martingale X, coincides with Cont and Fournié's concept of weak derivative with respect to X. However in our construction we do not require the martingale property or square-integrability of the canonical process, which are used in an essential way in [10].

We will work under the same framework as in section 2.2. Let W be a standard $d-$dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} := (\mathcal{F}_t^W)_{t\geq 0}$ its ($\mathbb{P}-$completed) natural filtration. Let X be a $\mathbb{R}^d-$valued Brownian martingale defined by: for all $t \in [0,T]$,

$$X(t) = X(0) + \int_0^t \sigma(s) \cdot dW(s), \tag{4.25}$$

where $\sigma : [0,T] \to \mathbb{M}_d(\mathbb{R})$ is a $\mathbb{F}-$adapted process satisfying (2.10), i.e.

$$\mathbb{E}\left[\int_0^T \|\sigma(t)\|^2 dt\right] < \infty \quad \text{and} \quad \det(\sigma(t)) \neq 0 \quad dt \times d\mathbb{P}\text{-a.e.} \tag{4.26}$$

Similarly to definition 2.7, we define $\mathcal{S}(X)$ the set of $\mathbb{F}-$adapted processes Y which admits a functional representation $Y(t) = F(t, X_t)$ with respect to X with $F \in \mathbb{S}(\Lambda_T^d)$:

$$\mathcal{S}(X) := \left\{Y : \exists F \in \mathbb{S}(\Lambda_T^d), \quad Y(t) = F(t, X_t) \quad dt \times d\mathbb{P}\text{-a.e.}\right\} \tag{4.27}$$

Since $\mathcal{S}(X) \subset \mathbb{C}_{loc}^{1,2}(X)$, in Assumption (4.26), for $Y \in \mathcal{S}(X)$, $\nabla_X Y$ is uniquely defined up to an evanescent set, independently of the choice of $F \in \mathbb{S}(\Lambda_T^d)$ in (4.27).

Recall that we denote by $\langle X \rangle$ the quadratic variation of the martingale X (to distinguish from the pathwise quadratic variation $[x]$ of a path x). Let $\mathcal{L}^2(X)$ be the space of $\mathbb{F}-$predictable process ϕ such that:

$$\|\phi\|_{\mathcal{L}^2(X)}^2 := \mathbb{E}\left[\int_0^T \operatorname{tr}\left(\phi(t)\, {}^t\phi(t) d\langle X\rangle(t)\right)\right] < \infty,$$

and $\mathcal{M}^2(X)$ the space of $\mathbb{R}$-valued square-integrable $\mathbb{F}-$martingales Y with initial value zero equipped with the norm:

$$\|Y\|^2_{\mathcal{M}^2(X)} := \mathbb{E}[Y(T)^2].$$

$\mathcal{L}^2(X)$ and $\mathcal{M}^2(X)$ are both Hilbert spaces. Let $D(X) := \mathcal{S}(X) \cap \mathcal{M}^2(X)$. Using the same arguments as in the proof of lemma 2.8, we still have the density of $D(X)$ in $\mathcal{M}^2(X)$.

Let $M \in \mathcal{M}^2(X)$. We will show that there exists a sequence of partitions π, a set $A \in C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$, and a non-anticipative functional $F \in H^\pi(\Lambda^d_T, A)$ such that $\mathbb{P}(X \in A) = 1$ and $M(t) = F(t, X_t)$ up to an evanescent set. We start with some preliminary results.

Lemma 4.10. *There exists a sequence $(M_n)_{n\geq 1}$ of elements in $D(X)$ such that:*

$$\langle M - M_n\rangle(T) \underset{n\to\infty}{\to} 0 \quad \mathbb{P}\text{-}a.s. \tag{4.28}$$

Proof. As $D(X)$ is dense in $\mathcal{M}^2(X)$, there exists a sequence $(M_n)_{n\geq 1}$ in $D(X)$ such that $\|M-M_n\|_{\mathcal{M}^2(X)} \underset{n\to\infty}{\to} 0$, which implies that $\mathbb{E}[\langle M-M_n\rangle(T)] \to_{n\to\infty} 0$. Thus there exists a sub-sequence $(M_{l(n)})_{n\geq 1}$ such that $\langle M - M_{l(n)}\rangle(T)$ converges to zero $\mathbb{P}$-almost surely. □

Consider now a sequence $(M_n)_{n\geq 1}$ of elements in $D(X)$ satisfying (4.28). By definition of the space $D(X)$, there exists a sequence of non-anticipative functional $(F_n)_{n\geq 1}$ in $\mathbb{S}(\Lambda^d_T)$ such that $M_n(t) = F_n(t, X_t)$ up to an evanescent set. Let $M(t) = F(t, X_t)$ for some non-anticipative functional F (clearly such functional F is not unique). We shall now construct a sequence of partitions π and a set $A \in C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ such that $\mathbb{P}(X \in A) = 1$, and for any $x \in A$, $F \in \widehat{Q}^\pi(\Lambda^d_T, x)$ and $\|F_n - F\|_{x,QV} \underset{n\to\infty}{\to} 0$ using simply the definition (4.24) of the space $H^\pi(\Lambda^d_T, A)$. This is equivalent to finding a sequence of partitions π which satisfies:

C.1 $\mathbb{P}(X \in Q^\pi([0,T],\mathbb{R}^d)) = 1$.

C.2 $\mathbb{P}(M \in \widehat{Q}^\pi([0,T],\mathbb{R})) = 1$.

C.3 $\mathbb{P}\left(\widehat{[M_n - M]}_\pi(T) \to 0\right) = 1$.

The idea is thus to construct a sequence of partitions π such that $\mathbb{P}$-almost surely, $X \in Q^\pi([0,T],\mathbb{R}^d)$ and $M, M_n - M \in Q^\pi([0,T],\mathbb{R})$ for any $n \geq 1$. If this is the case, then we have immediately, for any $n \geq 1$, $[M]_\pi = \langle M \rangle$ and $[M_n - M]_\pi = \langle M_n - M \rangle$ $\mathbb{P}$-almost surely. Since the generalized quadratic variation coincides with the standard quadratic variation in this case and $(M_n)_{n\geq 1}$ verifies (4.28), **C.1**, **C.2** and **C.3** are satisfied.

Let $\phi_m = (0 = t_0^m < t_1^m < \cdots < t_{k(m)}^m = T)$ be a sequence of partitions of $[0,T]$ with $|\phi_m| \underset{m\to\infty}{\to} 0$. For $Z \in \mathcal{M}^2(X)$, we define:

$$\langle Z \rangle_m^\phi(t) := \sum_{t_{i+1}^m \leq t} (Z(t_{i+1}^m) - Z(t_i^m))^2.$$

By the definition of quadratic variation for a square-integrable martingale, we have: $\langle Z \rangle_m^\phi(t) \underset{m\to\infty}{\to} \langle Z \rangle(t)$ in probability uniformly in $t \in [0,T]$. We can thus extract a sub-sequence of partitions $(\phi_{l(m)})_{m\geq 1}$ from $(\phi_m)_{m\geq 1}$ such that $\langle Z \rangle_{l(m)}^\phi(t)$ converges to $\langle Z \rangle(t)$ almost surely uniformly in $t \in [0,T]$.

Here as we are working with a sequence of martingales $(M_n)_{n\geq 1}$, a natural question is whether we can still find a sub-sequence $(\phi_{l(m)}))_{m\geq 1}$ of $(\phi_m)_{m\geq 1}$ such that for a sequence $(Z_n)_{n\geq 1}$ of elements in $\mathcal{M}^2(X)$, $\langle Z_n \rangle_{l(m)}^\phi(t)$ converges to $\langle Z_n \rangle(t)$ $\mathbb{P}$-almost surely uniformly in $t \in [0,T]$ along this same sub-sequence for all $n \geq 1$. This is ensured by the following lemma.

Lemma 4.11. *Let $(R_{m,n})_{m,n\geq 1}$ and $(R_n)_{n\geq 1}$ be random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $R_{m,n} \to_{m\to\infty} R_n$ in probability for all $n \in \mathbb{N}$. Then there exists a sub-sequence $l(m)$ such that $R_{l(m),n} \to_{m\to\infty} R_n$ almost surely for all $n \in \mathbb{N}$.*

Proof. We use a 'diagonal' argument. We first construct a sequence of sub-sequences $l_i(m)$ by induction: for $i = 1$, there exists a sub-sequence $l_1(m)$ such that $R_{l_1(m),1}$ converges almost surely to R_1. Having constructed $l_i(m)$, we extract from $l_i(m)$ a sub-sequence $l_{i+1}(m)$ (i.e. $\{l_{i+1}(m), m \in \mathbb{N}^*\} \subset \{l_i(m), m \in \mathbb{N}^*\}$) such that $R_{l_{i+1}(m),i+1}$ converges almost surely to R_{i+1}. By the construction of $l_i(m)$, for all $j \leq i$, $R_{l_i(m),j}$ converges almost surely to R_j. Now we define $l(m) := l_m(m)$. For any $n \in \mathbb{N}^*$, for $m \geq n$, $R_{l(m),n} = R_{l_m(m),n}$ converge almost surely to R_n since $\{l_m(m), m \geq n\} \subset \{l_n(m), m \geq 1\}$. $\square$

Applying lemma 4.11, we can extract from $(\phi_m)_{m\geq 1}$ a sub-sequence of partitions $\pi_m := \phi_{l(m)}$ such that $\mathbb{P}$-almost surely, for any $t \in [0,T]$ and for any $n \geq 1$,

$$\langle X\rangle^\pi_m(t) \underset{m\to\infty}{\to} \langle X\rangle(t), \qquad \langle M\rangle^\pi_m(t) \underset{m\to\infty}{\to} \langle M\rangle(t)$$
$$\text{and} \quad \langle M_n - M\rangle^\pi_m(t) \underset{m\to\infty}{\to} \langle M_n - M\rangle(t),$$

which is equivalent to $\mathbb{P}$-almost surely, for any $n \geq 1$,

$$[X]_\pi = \langle X\rangle, [M]_\pi = \langle M\rangle \quad \text{and} \quad [M_n - M]_\pi = \langle M - M_n\rangle.$$

Such sequence of partitions π clearly satisfies **C.1**, **C.2** and **C.3**.

Proposition 4.7. *Let $M \in \mathcal{M}^2(X)$. There exists a sequence of partitions π, a set of paths $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$, and a non-anticipative functional $F \in H^\pi(\Lambda^d_T, A)$ such that $M(t) = F(t, X_t)$ up to an evanescent set.*

Remark 4.12. *We can actually be more precise about the set of paths A. If σ in Definition* (4.25) *is a non-anticipative functional of X, i.e. $\sigma(t) := \sigma(t, X_t)$ for any $t \in [0,T]$, then A is in fact a subset of*

$$C_\sigma := \left\{ x \in C([0,T],\mathbb{R}^d), [x]_\pi(t) = \int_0^t \left(\sigma(s,x_s)\,{}^t\sigma(s,x_s)\right) ds, \quad \forall t \in [0,T] \right\}.$$

We refer to Mishura and Schied [52] for a pathwise construction of C_σ.

By proposition 4.7, the weak vertical derivative of the functional F along X: $\nabla^\pi F(\cdot, X_\cdot)$ is well defined $\mathbb{P}$-almost surely. We now show that the process $\nabla^\pi F(\cdot, X_\cdot)$ has an intrinsic character, independent of the functional F and the sequence of partitions π chosen in proposition 4.7. More precisely, it coincides with the weak derivative $\nabla_X M$ defined in theorem 2.9.

Proposition 4.8. *Let $M \in \mathcal{M}^2(X)$. Let π be a sequence of partitions of $[0,T]$, $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$ and $F \in H^\pi(\Lambda_T^d, A)$ such that $M(t) = F(t,X)$. Then the weak pathwise vertical derivative of F along $X(.,\omega)$ is a version of the the (probabilistic) weak derivative $\nabla_X M$of M with respect to X as defined in [10], i.e.*

$$\nabla^\pi F(t,X_t) = \nabla_X M(t) \quad dt \times d\mathbb{P}\text{-a.e.}$$

Proof. Since $F \in H^\pi(\Lambda_T^d, A)$, there exists a sequence of functionals $(F_n)_{n\geq 1}$ in $\mathbb{S}(\Lambda_T^d)$ such that for any $x \in A$, $\|F - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0$. Let $S_n(t) := F_n(t, X_t)$ for $t \in [0,T]$. By the functional Itô formula (theorem 2.2), S_n is a continuous semimartingale, and its (local) martingale part is $M_n(t) := \int_0^t \nabla_\omega F_n(s,X_s) \cdot dX(s)$.

Without any loss of generality, we may assume that $[X]_\pi = \langle X \rangle$, and for any $n \geq 1$, $[M - M_n]_\pi = \langle M - M_n \rangle$ $\mathbb{P}$-almost surely (otherwise we can always extract from π a sub-sequence of partitions ϕ such that these conditions hold, and we have $\nabla^\phi F(t,X_t) = \nabla^\pi F(t,X_t)\ dt \times d\mathbb{P}$-a.e.). Now using the extension of the isometry formula to $H^\pi(\Lambda_T^d, x)$ (4.21), $x \in A$, we have:

$$\begin{aligned}
& [M - M_n]_\pi(T) \\
= & \int_0^T \mathrm{tr}\left((\nabla^\pi F(t,X_t) - \nabla_\omega F_n(t,X_t))\, {}^t(\nabla^\pi F(t,X_t) - \nabla_\omega F_n(t,X_t)) d[X]_\pi(t) \right) \\
& \underset{n\to\infty}{\to} 0 \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{4.29}$$

On the other hand, since $M \in \mathcal{M}^2(X)$, by the martingale representation formula (theorem 2.10), we have, for any $t \in [0,T]$, $M(t) = \int_0^t \nabla_X M(s) \cdot dX(s)$ $\mathbb{P}$-almost surely, which implies:

$$\begin{aligned}
& \langle M - M_n \rangle(T) \\
= & \int_0^T \mathrm{tr}\left((\nabla_X M(t) - \nabla_\omega F_n(t,X_t))\, {}^t(\nabla_X M(t) - \nabla_\omega F_n(t,X_t)) d\langle X \rangle(t) \right) \\
= & \int_0^T \mathrm{tr}\left((\nabla_X M(t) - \nabla_\omega F_n(t,X_t))\, {}^t(\nabla_X M(t) - \nabla_\omega F_n(t,X_t)) d[X]_\pi(t) \right) \\
& \underset{n\to\infty}{\to} 0 \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{4.30}$$

Combining (4.29) with (4.30), and using the assumption (4.26) on σ, we obtain:

$$\nabla^\pi F(t, X_t) = \nabla_X M(t) \quad dt \times d\mathbb{P}\text{-a.e.}$$

$\square$

Corollary 4.3 (Martingale representation formula with weak pathwise derivatives). *Let $M \in \mathcal{M}^2(X)$. π be a sequence of partitions of $[0,T]$, $A \subset C([0,T], \mathbb{R}^d) \cap Q^\pi([0,T], \mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$, and $F \in H^\pi(\Lambda_T^d, A)$ such that $M(t) = F(t, X)$. Then*

$$\forall t \in [0,T], \quad M(t) = \int_0^t \nabla^\pi F(s, X_s) \cdot dX(s) \quad \mathbb{P}\text{-}a.s.$$

4.6 Pathwise characterization of martingale functionals

Let X be the canonical process on $\Omega = C^0([0,T] \times \mathbb{R}^d)$ and $\mathbb{P}$ be a probability measure on Ω such that X is a square-integrable $\mathbb{P}-$martingale satisfying (4.25) with integrand σ satisfying (4.26) and $\mathbb{F} := (\mathcal{F}_t^X)_{t\geq 0}$ the ($\mathbb{P}-$completed) natural filtration of X.

We have shown, in Proposition 4.7, that every square-integrable martingale $M \in \mathcal{M}^2(X)$ admits a functional representation in $H^\pi(\Lambda_T^d, A)$ with respect to X with A a set of paths satisfying $\mathbb{P}(X \in A) = 1$ for some sequence of partitions π.

Now we would like to study the converse of this problem. Given a non-anticipative functional F, can we give conditions on F such that the process $F(\cdot, X.)$ is a $\mathbb{P}$-(local) martingale? We call such a functional F an ($\mathbb{P}-$) harmonic functional (or simply harmonic functional in the case where $\mathbb{P}$ is the Wiener measure). Smooth $\mathbb{P}-$) harmonic functionals i.e. satisfying $F \in \mathbb{C}^{1,2}(\Lambda_T)$ may be characterized as solutions of the path-dependent PDE [7]:

$$\mathcal{D}F(t, x_t) + \frac{1}{2}\text{tr}(\nabla_\omega^2 F(t, x_t)\sigma(t, x)^t \sigma(t, x)) = 0$$

on the space of continuous functions. However, examples abound of functionals which have the martingale property but which fail to have the required directional derivatives [28]. In this section we present a different approach to the characterization of harmonic functionals which bypasses the smoothness requirement.

4.6.1 Martingale-preserving functionals

Let π be a sequence of partitions, and $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ such that $\mathbb{P}(X \in A) = 1$. Consider first functionals in $H^\pi(\Lambda_T^d, X)$. Clearly the process $F(\cdot, X.)$ is not necessarily a (local) martingale for an arbitrary choice of $F \in H^\pi(\Lambda_T^d, A)$. Indeed, since $\|\cdot\|_{x,QV}$ defined by (4.19) is only a semi-norm, if for some functional $F \in H^\pi(\Lambda_T^d, A)$, $F(\cdot, X.)$ is a (local) martingale, we can always add to F a non-anticipative functional G such that $\|G\|_{x,QV} = 0$ for any $x \in A$. By definition of the space $H^\pi(\Lambda_T^d, A)$, $F + G$ is still in $H^\pi(\Lambda_T^d, A)$. However, the process $(F+G)(\cdot, X.)$ is no longer a (local) martingale as it may possess a zero quadratic variation component given by G. This means for any sequence of partitions π, and for any set of paths A satisfying $\mathbb{P}(X \in A) = 1$, the space $H^\pi(\Lambda_T^d, A)$ is too large for $F(\cdot, X.)$ to be a (local) martingale for any $F \in H^\pi(\Lambda_T^d, A)$.

The idea is to find a subspace of $H^\pi(\Lambda_T^d, A)$ such that for functionals F in this subspace, when applied to the square-integrable martingale X, the process $F(\cdot, X.)$ is no longer allowed to have a zero quadratic variation part. Clearly a condition which only controls the (generalized) quadratic variation of F by a sequence of cylindrical functionals, as in the definition of the space $H^\pi(\Lambda_T^d, A)$ or $H^\pi(\Lambda_T^d, x)$, is not sufficient to eliminate the zero quadratic variation part of $F(\cdot, X.)$. We need another condition on F which allows to control the supremum norm of F in $[0,T]$, which motivates the following definition.

Definition 4.13 (Martingale-preserving functionals)**.** *Let π be a sequence of partitions of $[0,T]$. For a continuous path $x \in Q^\pi([0,T],\mathbb{R}^d)$, we define*

$I^\pi(\Lambda_T^d, x)$ as the space of non-anticipative functionals which can be approximated in (generalized) quadratic variation and in supremum norm along x by a sequence of cylindrical functionals:

$$I^\pi(\Lambda_T^d, x) := \Big\{ \quad G \in \widehat{Q}(\Lambda_T^d, x), \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, \|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0,$$
$$\text{and} \quad \int_0^{\cdot} \nabla_\omega F_n(t, x_t) \cdot d^\pi x(t) \underset{n\to\infty}{\to} G(\cdot, x.) \quad \text{in} \quad \|\cdot\|_\infty \Big\}. \tag{4.31}$$

We can easily check that $I^\pi(\Lambda_T^d, x)$ is a vector space, and $I^\pi(\Lambda_T^d, x) \subset H^\pi(\Lambda_T^d, x)$. We also observe that the space $I^\pi(\Lambda_T^d, x)$ is not empty. For example, assume that $d[x]_\pi$ is absolutely continuous with $\frac{d[x]_\pi}{dt} := a(t) \in \mathcal{S}_d^+(\mathbb{R})$. Consider now a functional $F \in \mathbb{S}(\Lambda_T^d)$ which satisfies $\mathcal{D}F(t, x_t) + \frac{1}{2}\mathrm{tr}(\nabla_\omega^2 F(t, x_t) a(t)) = 0$ for all $t \in [0, T]$ (this equation only needs to be satisfied along x). Then $F \in I^\pi(\Lambda_T^d, x)$ (we may take $F_n \equiv F$).

Remark 4.14. *The idea behind Definition 4.13 of the space $I^\pi(\Lambda_T^d, x)$ is that, intuitively, for $G \in I^\pi(\Lambda_T^d, x)$, $G(\cdot, x.)$ can be viewed as the pathwise integral of its weak pathwise derivative $\nabla^\pi(\cdot, x.)$ with respect to x, i.e.*

$$''G(t, x_t) = \int_0^t \nabla^\pi G(s, x_s) \cdot d^\pi x(s)''$$

even though the latter pathwise integral is not defined as a limit of Riemann sums. Nevertheless, we will see that when applied to the martingale X, the functional in such space is indeed equal to the stochastic integral of its weak derivative with respect to X.

Now let $A \in C([0, T], \mathbb{R}^d) \cap Q^\pi([0, T], \mathbb{R}^d)$ be a set of paths. Similarly to the definition 4.9 of the space $H^\pi(\Lambda_T^d, A)$, we define $I^\pi(\Lambda_T^d, A)$ as the space of non-anticipative functionals G such that we can find a sequence of functionals $(F_n)_{n\geq 1}$ in $\mathbb{S}(\Lambda_T^d)$ which satisfies (4.31) for any path $x \in A$:

$$I^\pi(\Lambda_T^d, A) := \Big\{ \quad G \in \widehat{Q}(\Lambda_T^d, A), \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, \forall x \in A, \|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0,$$
$$\text{and} \quad \Big\| \int_0^{\cdot} \nabla_\omega F_n(t, x_t) \cdot d^\pi x(t) - G(\cdot, x.)\Big\|_\infty \underset{n\to\infty}{\to} \quad 0. \Big\}.$$

Clearly $I^\pi(\Lambda_T^d, A) \subset H^\pi(\Lambda_T^d, A)$. We have shown in proposition 4.7 that every martingale in $\mathcal{M}^2(X)$ admits a functional representation in $H^\pi(\Lambda_T^d, A)$ for some sequence of partitions π and some set of paths A satisfying $\mathbb{P}(X \in A) = 1$. We now prove that the same result still holds if we replace the space $H^\pi(\Lambda_T^d, A)$ by $I^\pi(\Lambda_T^d, A)$. First we establish a useful result for smooth functionals.

Lemma 4.15. *Let $G \in \mathbb{C}_b^{1,2}(\Lambda_T^d)$ and π a sequence of partitions satisfying:*

$$\mathbb{P}\left(\quad \{\omega \in \Omega, \forall t \in [0,T], \quad [X(.,\omega)]_\pi(t) = \langle X\rangle(t,\omega)\right) = 1. \tag{4.32}$$

Then:

$$\mathbb{P}\text{-}a.s. \quad \forall t \in [0,T], \int_0^t \nabla_\omega G(s, X_s) \cdot d^\pi X(s) = \int_0^t \nabla_\omega G(s, X_s) \cdot dX(s).$$

Remark first that such a sequence of partitions π satisfying (4.32) exists. We observe also the difference between the two integrals. The first integral is a pathwise integral defined as limit of the non-anticipative Riemann sums along π evaluated on X, which is well defined $\mathbb{P}$-almost surely as $\mathbb{P}(X \in Q^\pi([0,T], \mathbb{R}^d)) = 1$ by assumption (4.32). And the second integral is the classical stochastic integral with respect to the martingale X. This lemma shows that these two integrals coincide $\mathbb{P}$-almost surely.

Proof. Using the pathwise Itô formula for functionals (theorem 1.18) and functional Itô formula (theorem 2.2), we have:

$$\begin{aligned} G(t, X_t) &= G(0, X_0) + \int_0^t \mathcal{D}G(s, X_s)ds + \int_0^t \nabla_\omega G(s, X_s) \cdot d^\pi X(s) \\ &\quad + \frac{1}{2}\int_0^t \operatorname{tr}\left(\nabla_\omega^2 G(s, X_s) d[X]_\pi(s)\right) \quad \mathbb{P}\text{-a.s.} \\ &= G(0, X_0) + \int_0^t \mathcal{D}G(s, X_s)ds + \int_0^t \nabla_\omega G(s, X_s) \cdot dX(s) \\ &\quad + \frac{1}{2}\int_0^t \operatorname{tr}\left(\nabla_\omega^2 G(s, X_s) d\langle X\rangle(s)\right) \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

And we conclude using assumption (4.32). $\square$

We can now state a refined version of Proposition 4.7:

Proposition 4.9. *Let $M \in \mathcal{M}^2(X)$. There exists a sequence of partitions π, a set of paths $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$, and a non-anticipative functional $F \in I^\pi(\Lambda_T^d, A)$ such that $M(t) = F(t, X_t)$ up to an evanescent set.*

Proof. The proof is quite similar to that of proposition 4.7. The only thing we shall prove in addition is that $F(\cdot, x_\cdot)$ is the limit in the supremum norm of a sequence of pathwise integrals of cylindrical functionals $(F_n)_{n\geq 1}$ along x for all $x \in A$, i.e. the process $F(\cdot, X_\cdot)$ is the almost sure limit in the supremum norm of $\int_0^\cdot \nabla_\omega F_n(t, X_t) \cdot d^\pi X(t)$ which is well defined for π satisfying (4.32).

Let $(M_n)_{n\geq 1}$ be a sequence of elements in $D(X) := \mathcal{S}(X) \cap \mathcal{M}^2(X)$ which converges to M in $\|\cdot\|_{\mathcal{M}^2(X)}$, and $(F_n)_{n\geq 1}$ a sequence of non-anticipative functionals in $\mathbb{S}(\Lambda_T^d)$ such that $M_n(t) = F_n(t, X_t)$ up to an evanescent set. Consider a sequence of partitions π which satisfies (4.32). Since $F_n(\cdot, X_\cdot)$ is a martingale, by lemma 4.15, we have: for any $t \in [0,T]$,

$$F_n(t, X_t) = \int_0^t \nabla_\omega F_n(s, X_s) \cdot dX(s) = \int_0^t \nabla_\omega F_n(s, X_s) \cdot d^\pi X(s) \quad \mathbb{P}\text{-a.s.}$$

Thus the additional condition on F of convergence in supremum norm is equivalent to, in terms of probability, the almost sure convergence of $\sup_{0\leq t\leq T} |M_n(t) - M(t)|$ to 0. Such sequence $(M_n)_{n\geq 1}$ (or equivalently $(F_n)_{n\geq 1}$) exists since $\sup_{0\leq t\leq T} |M_n(t) - M(t)|$ converges to 0 in $\mathbb{L}^2$ by Doob's martingale inequality. □

We can also state a 'local' version of proposition 4.9 by introducing a sequence of stopping times in the definition of the space $I^\pi(\Lambda_T^d, A)$. Let π be a sequence of partitions, and $A \in C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ a set of paths, we define $I^\pi_{loc}(\Lambda_T^d, A)$ as the space of non-anticipative functionals G such that there exists an increasing sequence $(\tau_k)_{k\geq 1}$ of stopping times (with respect to the filtration generated by the canonical process on $C([0,T],\mathbb{R}^d)$) with $\tau_k \underset{k\to\infty}{\to} \infty$ such that the functional G stopped at τ_k: G^{τ_k} belongs to $I^\pi(\Lambda_T^d, A)$ for any $k \geq 1$. We denote by $\mathcal{M}_{loc}(X)$ the space of continuous

$\mathbb{F}$-local martingales with initial value zero. Then every local martingale in $\mathcal{M}_{loc}(X)$ admits a functional representation in $I^{\pi}_{loc}(\Lambda^d_T, A)$ with respect to X for some sequence of partitions π and some set of paths A satisfying $\mathbb{P}(X \in A) = 1$.

Corollary 4.4. *Let $M \in \mathcal{M}_{loc}(X)$. There exists a sequence of partitions π, a set of paths $A \subset C([0,T],\mathbb{R}^d) \cap Q^{\pi}([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$, and a non-anticipative functional $F \in I^{\pi}_{loc}(\Lambda^d_T, A)$ such that $M(t) = F(t, X_t)$ up to an evanescent set.*

Proof. Denote by $\mathbb{F}^c := (\mathcal{F}^c_t)_{t\geq 0}$ the filtration generated by the canonical process on $C([0,T],\mathbb{R}^d)$. Recall that $\mathbb{F} = (\mathcal{F}^W_t)_{t\geq 0} = (\mathcal{F}^X_t)_{t\geq 0}$ is the $\mathbb{P}$-completed natural filtration of X (or W). Since M is a continuous $\mathbb{F}$-local martingale, there exists a sequence of $\mathbb{F}$-stopping times $(\mu_k)_{k\geq 1}$ with $\mu_k \underset{k\to\infty}{\to} \infty$ such that $M^{\mu_k} \in \mathcal{M}^2(X)$ for all $k \geq 1$. By proposition 4.9, there exists a sequence of partitions π^k, a set of paths $A^k \in C([0,T],\mathbb{R}^d) \cap Q^{\pi^k}([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A^k) = 1$, and a non-anticipative functional $F^k \in I^{\pi^k}(\Lambda^d_T, A^k)$ such that $M^{\mu_k}(t) = F^k(t, X_t)$ up to an evanescent set. Moreover, since μ_k are $\mathbb{F}$-stopping times, there exists a sequence of $\mathbb{F}^c$-stopping times $(\tau_k)_{k\geq 1}$ such that $\mu_k = \tau_k$ $\mathbb{P}$-almost surely for all $k \geq 1$.

Without loss of generality, we may assume that for any $k \geq 1$, π^{k+1} is a sub-sequence of π^k, i.e. $\{\pi^{k+1}_m, m \in \mathbb{N}^*\} \subset \{\pi^k_m, m \in \mathbb{N}^*\}$. Now define $\pi := (\pi_m)_{m\geq 1} = (\pi^m_m)_{m\geq 1}$, and $B = \cap_{k\geq 1} A^k$. Clearly we have $\mathbb{P}(X \in B) = 1$, and we can readily check that $B \in C([0,T],\mathbb{R}^d) \cap Q^{\pi}([0,T],\mathbb{R}^d)$ and $F^k \in I^{\pi}(\Lambda^d_T, B)$ for any $k \geq 1$. For $x \in A$, set $F(t,x) := \lim_{k\to\infty} F^k(t,x)$ if the limit exists, and $F(t,x) = 0$ otherwise. By definition of the functionals F^k, $F(t, X_t) = \lim_{k\to\infty} F^k(t, X_t) = \lim_{k\to\infty} M^{\mu_k}(t) = M(t)$ $\mathbb{P}$-almost surely, we have $F(t, X_t) = M(t)$ up to an evanescent set. Moreover, since $\mu_k = \tau_k$ $\mathbb{P}$-almost surely, $F^{\tau_k}(t, X_t) = M^{\mu_k}(t) = F^k(t, X_t)$ up to an evanescent set. Let $A \subset B$ be the set of x such that $\lim_{k\to\infty} F^k(t,x)$ exists and $F^{\tau_k}(t,x) = F^k(t,x)$ for any $t \in [0,T]$. We have $\mathbb{P}(X \in A) = 1$ and $F \in I^{\pi}_{loc}(\Lambda^d_T, A)$. $\square$

Now we prove a converse result to Proposition 4.7 (more precisely, a converse to Corollary 4.4):

Proposition 4.10. *If there exists a sequence of partitions π such that $F \in I^{\pi}_{loc}(\Lambda^d_T, A)$ for some $A \subset C([0,T],\mathbb{R}^d) \cap Q^{\pi}([0,T],\mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$ then the process $F(\cdot, X.)$ is a local martingale.*

Proof. Let π be a sequence of partitions, $A \subset C([0,T],\mathbb{R}^d) \cap Q^{\pi}([0,T],\mathbb{R}^d)$ a set of paths satisfying $\mathbb{P}(X \in A) = 1$, and F a non-anticipative functional in $I^{\pi}_{loc}(\Lambda^d_T, A)$. By definition of the space $I^{\pi}_{loc}(\Lambda^d_T, A)$, there exists a sequence of $\mathbb{F}^c$-stopping times $(\tau_k)_{k\geq 1}$ with $\tau_k \underset{k\to\infty}{\to} \infty$ (which are also $\mathbb{F}$-stopping times) such that $F^{\tau_k} \in I^{\pi}(\Lambda^d_T, A)$ for any $k \geq 1$. And if $F^{\tau_k}(\cdot, X.)$ is a local martingale for any $k \geq 1$, then it is a classical result of stochastic calculus that $F(\cdot, X.)$ is itself a local martingale. So in the following we may assume without loss of generality that $F \in I^{\pi}(\Lambda^d_T, A)$.

First observe that, up to extraction of a sub-sequence of π, we may always assume that the sequence of partitions π satisfies condition (4.32). The main idea of the proof is to show that: for any $t \in [0,T]$,

$$F(t, X_t) = \int_0^t \nabla^{\pi} F(s, X_s) \cdot dX(s) \quad \mathbb{P}\text{-a.s.}$$

Remark that the previous stochastic integral is well defined since by definition, the weak vertical derivative of F along the paths of X: $\nabla^{\pi} F(\cdot, X.)$ belongs to the space $\mathbb{L}^2([0,T],[X]_{\pi}) = \mathbb{L}^2([0,T],\langle X\rangle)$ $\mathbb{P}$-a.s.

By definition of the space $I^{\pi}(\Lambda^d_T, A)$ and the fact that $\mathbb{P}(X \in A) = 1$, there exists a sequence of functionals $(F_n)_{n\geq 1}$ in $\mathbb{S}(\Lambda^d_T)$ such that:

$$\|F - F_n\|_{X,QV} \underset{n\to\infty}{\to} 0 \quad \mathbb{P}\text{-a.s.} \tag{4.33}$$

and

$$\int_0^{\cdot} \nabla_{\omega} F_n(t, X_t) \cdot d^{\pi} X(t) \underset{n\to\infty}{\to} F(\cdot, X.) \quad \text{in} \quad \|\cdot\|_{\infty} \quad \mathbb{P}\text{-a.s.} \tag{4.34}$$

(4.33) implies, by the extension of the isometry formula to the space $H^{\pi}(\Lambda^d_T, x)$

(4.21) for $x \in A$, that:

$$\begin{aligned}&\int_0^T \operatorname{tr}\left((\nabla^\pi F - \nabla_\omega F_n)(t, X_t)\, {}^t(\nabla^\pi F - \nabla_\omega F_n)(t, X_t) d[X]_\pi(t)\right)\\ =\ &\int_0^T \operatorname{tr}\left((\nabla^\pi F - \nabla_\omega F_n)(t, X_t)\, {}^t(\nabla^\pi F - \nabla_\omega F_n)(t, X_t) d\langle X\rangle(t)\right) \underset{n\to\infty}{\to} 0 \quad \mathbb{P}\text{-a.s.}\end{aligned} \tag{4.35}$$

Let $M_n(t) := \int_0^t (\nabla^\pi F - \nabla_\omega F_n)(t, X_t) \cdot dX(t)$ be a sequence of local martingales. By (4.35), we have $\langle M_n\rangle(T) \underset{n\to\infty}{\to} 0$ $\mathbb{P}$-almost surely. On the other hand, combining (4.34) with lemma 4.15, we have:

$$\int_0^\cdot \nabla_\omega F_n(t, X_t) \cdot dX(t) \underset{n\to\infty}{\to} F(\cdot, X_\cdot) \quad \text{in} \quad \|\cdot\|_\infty \quad \mathbb{P}\text{-a.s.}$$

So if we are able to show that the sequence of martingales $(M_n)_{n\geq 1}$ converges uniformly in $[0, T]$ to 0 in some sense, by identification of the limit, the process $F(\cdot, X_\cdot)$ is necessarily equal to $\int_0^\cdot \nabla^\pi F(t, X_t) \cdot dX(t)$ $\mathbb{P}$-almost surely, thus is a local martingale. This is ensured by the following result:

Lemma 4.16. *Let $(M_n)_{n\geq 1}$ be a sequence of continuous local martingales starting at zero. If $\langle M_n\rangle(T)$ converges to 0 in probability, then $\sup_{t\in[0,T]} |M_n(t)|$ converges to 0 in probability.*

This is a classical exercise of stochastic calculus, which can be found in, for example [64], or [48] for more details. In our case, $\langle M_n\rangle(T)$ converges to 0 $\mathbb{P}$-almost surely (thus in probability). By lemma 4.16, $\sup_{t\in[0,T]} |M_n(t)|$ converges to 0 in probability. Since $M_n(t) = \int_0^t \nabla^\pi F(s, X_s) \cdot dX(s) - \int_0^t \nabla_\omega F_n(s, X_s) \cdot dX(s)$ converges to $\int_0^t \nabla^\pi F(s, X_s) \cdot dX(s) - F(t, X_t)$ in $\|\cdot\|_\infty$ $\mathbb{P}$-almost surely, we have, $\mathbb{P}$-almost surely, $F(t, X_t) = \int_0^t \nabla^\pi F(s, X_s) \cdot dX(s)$ for any $t \in [0, T]$. $\square$

Combining corollary 4.4 with proposition 4.10, we arrive at our main result:

Theorem 4.17 (Characterization of martingale-preserving functionals)**.** *Let X be a square-integrable martingale of the form* (4.25) *with σ satisfying condition* (4.26)*, and let F be a non-anticipative functional. The process $F(\cdot, X_\cdot)$ is a $(\mathbb{P}, \mathbb{F})$-local martingale starting at* 0 *if and only if there exists a sequence of partitions π such that $F \in I^\pi_{loc}(\Lambda^d_T, A)$ for some $A \subset C([0,T], \mathbb{R}^d) \cap Q^\pi([0,T], \mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$.*

This result can be viewed as a functional characterization of local martingales with respect to the filtration $\mathbb{F}$ of a square-integrable martingale X. This characterization is not entirely pathwise but depends on the probability measure $\mathbb{P}$ only through its null sets.

4.6.2 Extension to the case of Itô processes

This result can also be extended to the case X is a square-integrable Itô process. Let $(X, \mathbb{P})$ be the weak solution to the following path-dependent SDE:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t), \quad X(0) = x_0 \in \mathbb{R}^d \tag{4.36}$$

Let b and σ two non-anticipative functionals assumed to satisfy conditions for (4.36) to admit a unique weak solution $\mathbb{P}$ on the canonical space. We assume in addition that σ satisfies (4.26).

The objective is to characterize non-anticipative functionals F such that $F(\cdot, X_\cdot)$ is a $\mathbb{F}$-local martingale. If we still take $F \in I^\pi_{loc}(\Lambda^d_T, A)$ for some sequence of partitions π and some set of paths $A \subset C([0,T], \mathbb{R}^d) \cap Q^\pi([0,T], \mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$, using the same argument as in the case X is a square-integrable martingale, for $t \in [0,T]$, $F(t, X_t) = \int_0^t \nabla^\pi F(s, X_s) \cdot dX(s)$ $\mathbb{P}$-almost surely which is clearly not a local martingale.

The idea is to eliminate the finite variation part of X in the definition of $I^\pi_{loc}(\Lambda^d_T, A)$. Let π be a sequence of partitions, and $A \subset C([0,T], \mathbb{R}^d) \cap$

$Q^\pi([0,T],\mathbb{R}^d)$ a set of paths. We define the space of non-anticipative functionals $I_b^\pi(\Lambda_T^d, A)$ as follows:

$$I_b^\pi(\Lambda_T^d, A) = \Big\{ G \in \widehat{Q}(\Lambda_T^d, A), \exists (F_n) \in \mathbb{S}(\Lambda_T^d)^{\mathbb{N}}, \forall x \in A, \|G - F_n\|_{x,QV} \underset{n\to\infty}{\to} 0,$$
$$\int_0^{\cdot} \nabla_\omega F_n(t, x_t) \cdot d^\pi x(t) - \int_0^{\cdot} \nabla_\omega F_n(t, X_t) \cdot b(t, x_t) dt \underset{n\to\infty}{\to} G(\cdot, x.) \quad \text{in} \quad \|\cdot\|_\infty \Big\}.$$

Let $M(t) := \int_0^t \sigma(s, X_s) dW(s)$ be the martingale part of X. Using the same argument as in the proof of proposition 4.10, for $F \in I_b^\pi(\Lambda_T^d, A)$ with $\mathbb{P}(X \in A) = 1$, we have, $\mathbb{P}$-almost surely, $F(t, X_t) = \int_0^t \nabla^\pi F(s, X_s) \cdot dM(s)$ for any $t \in [0,T]$. We define similarly a local version $I_{loc,b}^\pi(\Lambda_T^d, A)$ of the space $I_b^\pi(\Lambda_T^d, A)$, and we obtain the following functional characterization of local martingales with respect to X.

Proposition 4.11. *Let* $(X, \mathbb{P})$ *be a semimartingale defined by* (4.36) *with* σ *satisfying the integrability and non-singularity condition* (4.26)*, and* F *be a non-anticipative functional. The process* $F(\cdot, X.)$ *is a* $\mathbb{F}$*-local martingale starting at* 0 *if and only if there exists a sequence of partitions* π *and a set of paths* $A \subset C([0,T],\mathbb{R}^d) \cap Q^\pi([0,T],\mathbb{R}^d)$ *satisfying* $\mathbb{P}(X \in A) = 1$ *such that* $F \in I_{loc,b}^\pi(\Lambda_T^d, A)$.

4.6.3 A weak solution concept for path-dependent PDEs

The above characterization of local martingales gives rise to a notion of weak solution for linear path-dependent PDEs. Consider the following path-dependent PDE:

$$\mathcal{D}F(t, x_t) + b(t, x_t) \cdot \nabla_\omega F(t, x_t) + \tfrac{1}{2}\mathrm{tr}\left(\sigma(t, x_t)\, {}^t\sigma(t, x_t) \nabla_\omega^2 F(t, x_t)\right) = 0,$$
$$t \in [0,T], x \in C([0,T],\mathbb{R}^d), \tag{4.37}$$

where b and σ are two non-anticipative functionals which satisfy assumption 3.1, and $\det(\sigma(t, x_t)) \neq 0$ for any $t \in [0,T]$ and $x \in C([0,T],\mathbb{R}^d)$. Let W be a standard d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $(X, \mathbb{P})$ a semimartingale defined by (4.36). Under assumption

3.1, X is square-integrable (proposition 3.1) and we have:

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^T \|\sigma(t, X_t)\|^2 dt\right] < \infty.$$

Definition 4.18 (Weak solution of path-dependent PDEs)**.** *A non-anticipative functional F is a weak solution of the path-dependent PDE* (4.37) *if there exists a sequence of partitions π, and a set of paths $A \subset C([0,T], \mathbb{R}^d) \cap Q^\pi([0,T], \mathbb{R}^d)$ satisfying $\mathbb{P}(X \in A) = 1$ such that $F \in I^\pi_{loc,b}(\Lambda^d_T, A)$.*

Under technical assumptions on σ and b, the choice of the set A depends only on σ, not on b.

The following result shows that the above Definition extends the notion of $\mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ (strong) solution of the path-dependent PDE (4.37): any $\mathbb{C}^{1,2}_b$ solution is also a weak solution in the sense of definition 4.18:

Proposition 4.12. *If $F \in \mathbb{C}^{1,2}_{loc}(\mathcal{W}^d_T)$ is a strong solution of the path-dependent PDE* (4.37) *with $\mathcal{D}F \in \mathbb{C}^{0,0}_l(\mathcal{W}^d_T)$, then F is also a weak solution of* (4.37) *in the sense of Definition 4.18.*

Proof. This is an immediate consequence of theorem 2.13 and proposition 4.11 since the topological support of the semimartingale X is the set of all continuous paths starting from x_0 as we have assumed $\det(\sigma(t, x_t)) \neq 0$ for any $t \in [0,T]$ and $x \in C([0,T], \mathbb{R}^d)$. □

Our notion of weak solution is actually at least as weak as most of the other notions of solution for linear path-dependent PDEs such as the viscosity solution proposed in [21] or the Sobolev-type weak solution proposed in [7]. Indeed, in our definition, for F to be a weak solution of (4.37), we only require $F(\cdot, X.)$ is a local martingale by proposition 4.11.

One of the obstacles to the use of strong solutions is that the conditional expectation of a functional $H : C^0([0,T],\mathbb{R}^d) \mapsto \mathbb{R}$ does not necessarily have the (continuous) directional derivatives, even when H is smooth [65, 58]. Our concept of weak solution bypasses this difficulty. The versatility of Definition 4.18 is illustrated by the following existence result, which is a reformulation of 4.11:

Proposition 4.13 (Existence of weak solutions). *Let b and σ two non-anticipative functionals such that* (4.36) *admits a unique weak solution* $(X,\mathbb{P})$ *on the canonical space, and σ further satisfies* (4.26)*. Then for any functional $H : C^0([0,T],\mathbb{R}^d) \mapsto \mathbb{R}$ such that $H(X_T) \in L^1(\Omega,\mathcal{F}_T^X,\mathbb{P})$, the path-dependent PDE*

$$\begin{aligned} \mathcal{D}F(t,x_t) + b(t,x_t)\cdot\nabla_\omega F(t,x_t) + \frac{1}{2}\mathrm{tr}\left(\sigma(t,x_t)\,{}^t\sigma(t,x_t)\nabla_\omega^2 F(t,x_t)\right) &= 0, \\ F(T,x) = H(x), \qquad t\in[0,T], x &\in C([0,T],\mathbb{R}^d). \end{aligned}$$

admits a weak solution. Furthermore, if F_1, F_2 are two such weak solutions then

$$F_1(t,X) = F_2(t,X) \quad dt\times d\mathbb{P} - a.e.$$

Bibliography

[1] Anna Ananova and Rama Cont. Pathwise integration with respect to paths of finite quadratic variation. *Journal de Mathématiques Pures et Appliquées*, 107(6):737–757, June 2017. URL https://doi.org/10.1016/j.matpur.2016.10.004.

[2] Jean-Michel Bismut. A generalized formula of Itô and some other properties of stochastic flows. *Z. Wahrsch. Verw. Gebiete*, 55(3):331–350, 1981. ISSN 0044-3719. doi: 10.1007/BF00532124. URL http://dx.doi.org/10.1007/BF00532124.

[3] Jean-Michel Bismut. Calcul des variations stochastique et processus de sauts. *Z. Wahrsch. Verw. Gebiete*, 63(2):147–235, 1983. ISSN 0044-3719. doi: 10.1007/BF00538963. URL http://dx.doi.org/10.1007/BF00538963.

[4] Bruno Bouchard and Nizar Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111(2):175–206, 2004. ISSN 0304-4149. doi: 10.1016/j.spa.2004.01.001. URL http://dx.doi.org/10.1016/j.spa.2004.01.001.

[5] Nan Chen and Paul Glasserman. Malliavin Greeks without Malliavin calculus. *Stochastic Process. Appl.*, 117(11):1689–1723, 2007. ISSN 0304-4149. doi: 10.1016/j.spa.2007.03.012. URL http://dx.doi.org/10.1016/j.spa.2007.03.012.

[6] J. M. C. Clark. The representation of functionals of Brownian motion by stochastic integrals. *Ann. Math. Statist.*, 41:1282–1295, 1970. ISSN 0003-4851. doi: 10.1214/aoms/1177696903. URL `http://dx.doi.org/10.1214/aoms/1177696903`.

[7] Rama Cont. Functional Itô calculus and functional Kolmogorov equations. In *Stochastic integration by parts and functional Itô calculus*, Advanced Courses in Mathematics (CRM Barcelona), pages 115–207. Birkhäuser/Springer, 2016.

[8] Rama Cont and David-Antoine Fournié. Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.*, 259(4): 1043–1072, 2010. ISSN 0022-1236. doi: 10.1016/j.jfa.2010.04.017. URL `http://dx.doi.org/10.1016/j.jfa.2010.04.017`.

[9] Rama Cont and David-Antoine Fournié. A functional extension of the Ito formula. *C. R. Math. Acad. Sci. Paris*, 348(1-2):57–61, 2010. ISSN 1631-073X. doi: 10.1016/j.crma.2009.11.013. URL `http://dx.doi.org/10.1016/j.crma.2009.11.013`.

[10] Rama Cont and David-Antoine Fournié. Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.*, 41(1): 109–133, 2013. ISSN 0091-1798. doi: 10.1214/11-AOP721. URL `http://dx.doi.org/10.1214/11-AOP721`.

[11] Rama Cont and Yi Lu. Weak approximation of martingale representations. *Stochastic Process. Appl.*, 126(3):857–882, 2016. ISSN 0304-4149. doi: 10.1016/j.spa.2015.10.002. URL `http://dx.doi.org/10.1016/j.spa.2015.10.002`.

[12] Andrea Cosso. Viscosity solutions of path-dependent pdes and non-markovian forward-backward stochastic equations. 2013. Preprint, `https://arxiv.org/abs/1202.2502`.

[13] Andrea Cosso and Francesco Russo. Strong-viscosity solutions: Semilinear parabolic pdes and path-dependent pdes. 2015. Preprint, `https://arxiv.org/abs/1505.02927`.

[14] Andrea Cosso and Francesco Russo. Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 19(4): 1650024, 44, 2016. ISSN 0219-0257.

[15] Andrea Cosso, Cristina Di Girolami, and Francesco Russo. Calculus via regularizations in Banach spaces and Kolmogorov-type path-dependent equations. In *Probability on algebraic and geometric structures*, volume 668 of *Contemp. Math.*, pages 43–65. Amer. Math. Soc., Providence, RI, 2016. URL `http://dx.doi.org/10.1090/conm/668/13396`.

[16] Jakša Cvitanić, Jin Ma, and Jianfeng Zhang. Efficient computation of hedging portfolios for options with discontinuous payoffs. *Math. Finance*, 13(1):135–151, 2003. ISSN 0960-1627. doi: 10.1111/1467-9965.00010. URL `http://dx.doi.org/10.1111/1467-9965.00010`. Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001).

[17] M. H. A. Davis. Functionals of diffusion processes as stochastic integrals. *Math. Proc. Cambridge Philos. Soc.*, 87(1):157–166, 1980. ISSN 0305-0041. doi: 10.1017/S0305004100056590. URL `http://dx.doi.org/10.1017/S0305004100056590`.

[18] Mark H. A. Davis and Martin P. Johansson. Malliavin Monte Carlo Greeks for jump diffusions. *Stochastic Process. Appl.*, 116(1):101–129, 2006. ISSN 0304-4149. doi: 10.1016/j.spa.2005.08.002. URL `http://dx.doi.org/10.1016/j.spa.2005.08.002`.

[19] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland

Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978. ISBN 0-7204-0701-X.

[20] Bruno Dupire. Functional Itô calculus. *Bloomberg Portfolio Research paper*, (2009-04), 2009.

[21] Ibrahim Ekren, Christian Keller, Nizar Touzi, and Jianfeng Zhang. On viscosity solutions of path dependent PDEs. *Ann. Probab.*, 42(1):204–236, 2014. ISSN 0091-1798. doi: 10.1214/12-AOP788. URL `http://dx.doi.org/10.1214/12-AOP788`.

[22] Ibrahim Ekren, Nizar Touzi, and Jianfeng Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I. *Ann. Probab.*, 44(2):1212–1253, 2016. ISSN 0091-1798. doi: 10.1214/14-AOP999. URL `http://dx.doi.org/10.1214/14-AOP999`.

[23] Ibrahim Ekren, Nizar Touzi, and Jianfeng Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part II. *Ann. Probab.*, 44(4):2507–2553, 2016. ISSN 0091-1798. doi: 10.1214/15-AOP1027. URL `http://dx.doi.org/10.1214/15-AOP1027`.

[24] Robert J. Elliott and Michael Kohlmann. A short proof of a martingale representation result. *Statist. Probab. Lett.*, 6(5):327–329, 1988. ISSN 0167-7152. doi: 10.1016/0167-7152(88)90008-9. URL `http://dx.doi.org/10.1016/0167-7152(88)90008-9`.

[25] Olivier Faure. *Simulation du mouvement brownien et des diffusions*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1992.

[26] P. J. Fitzsimmons and B. Rajeev. A new approach to the martingale representation theorem. *Stochastics*, 81(5):467–476, 2009. ISSN 1744-2508. doi: 10.1080/17442500802343417. URL `http://dx.doi.org/10.1080/17442500802343417`.

[27] H. Föllmer. Calcul d'Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.

[28] David-Antoine Fournié. *Functional Itô calculus and applications.* PhD thesis, Columbia University, 2010.

[29] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance Stoch.*, 3(4):391–412, 1999. ISSN 0949-2984. doi: 10.1007/s007800050068. URL `http://dx.doi.org/10.1007/s007800050068`.

[30] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, and Pierre-Louis Lions. Applications of Malliavin calculus to Monte-Carlo methods in finance. II. *Finance Stoch.*, 5(2):201–236, 2001. ISSN 0949-2984. doi: 10.1007/PL00013529. URL `http://dx.doi.org/10.1007/PL00013529`.

[31] Paul Glasserman and David D. Yao. Some guidelines and guarantees for common random numbers. *Manage. Sci.*, 38(6):884–908, June 1992. ISSN 0025-1909. doi: 10.1287/mnsc.38.6.884. URL `http://dx.doi.org/10.1287/mnsc.38.6.884`.

[32] P. W. Glynn. Optimization of stochastic systems via simulation. In *1989 Winter Simulation Conference Proceedings*, pages 90–105, Dec 1989. doi: 10.1109/WSC.1989.718667.

[33] Emmanuel Gobet and Arturo Kohatsu-Higa. Computation of Greeks for barrier and look-back options using Malliavin calculus. *Electron. Comm. Probab.*, 8:51–62, 2003. ISSN 1083-589X. doi: 10.1214/ECP.v8-1069. URL `http://dx.doi.org/10.1214/ECP.v8-1069`.

[34] Emmanuel Gobet and Stéphane Menozzi. Discrete sampling of functionals of Itô processes. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 355–374. Springer, Berlin, 2007. doi: 10.

1007/978-3-540-71189-6_19. URL http://dx.doi.org/10.1007/978-3-540-71189-6_19.

[35] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005. ISSN 1050-5164. doi: 10.1214/105051605000000412. URL http://dx.doi.org/10.1214/105051605000000412.

[36] U. G. Haussmann. On the integral representation of functionals of Itô processes. *Stochastics*, 3(1):17–27, 1979. ISSN 0090-9491. doi: 10.1080/17442507908833134. URL http://dx.doi.org/10.1080/17442507908833134.

[37] Desmond J. Higham, Xuerong Mao, and Andrew M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40(3):1041–1063, 2002. ISSN 0036-1429. doi: 10.1137/S0036142901389530. URL http://dx.doi.org/10.1137/S0036142901389530.

[38] David G. Hobson and L. C. G. Rogers. Complete models with stochastic volatility. *Math. Finance*, 8(1):27–48, 1998. ISSN 0960-1627. doi: 10.1111/1467-9965.00043. URL http://dx.doi.org/10.1111/1467-9965.00043.

[39] Kiyosi Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519–524, 1944. URL http://projecteuclid.org/euclid.pja/1195572786.

[40] Kiyosi Itô. On a stochastic integral equation. *Proc. Japan Acad.*, 22 (nos. 1-4):32–35, 1946. ISSN 0021-4280. URL http://projecteuclid.org/euclid.pja/1195572371.

[41] Jean Jacod, Sylvie Méléard, and Philip Protter. Explicit form and robustness of martingale representations. *Ann. Probab.*, 28(4):1747–

1780, 2000. ISSN 0091-1798. doi: 10.1214/aop/1019160506. URL `http://dx.doi.org/10.1214/aop/1019160506`.

[42] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991. ISBN 0-387-97655-8. doi: 10.1007/978-1-4612-0949-2. URL `http://dx.doi.org/10.1007/978-1-4612-0949-2`.

[43] Ioannis Karatzas, Daniel L. Ocone, and Jinlu Li. An extension of Clark's formula. *Stochastics Stochastics Rep.*, 37(3):127–131, 1991. ISSN 1045-1129.

[44] Arturo Kohatsu-Higa and Miquel Montero. Malliavin calculus in finance. In *Handbook of computational and numerical methods in finance*, pages 111–174. Birkhäuser Boston, Boston, MA, 2004.

[45] A. Kolmogoroff. über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.*, 104(1):415–458, 1931. ISSN 0025-5831. doi: 10.1007/BF01457949. URL `http://dx.doi.org/10.1007/BF01457949`.

[46] Hiroshi Kunita and Shinzo Watanabe. On square integrable martingales. *Nagoya Math. J.*, 30:209–245, 1967. ISSN 0027-7630. URL `http://projecteuclid.org/euclid.nmj/1118796812`.

[47] Dorival Leão and Alberto Ohashi. Weak approximations for Wiener functionals. *Ann. Appl. Probab.*, 23(4):1660–1691, 2013. ISSN 1050-5164.

[48] Jean-François Le Gall. *Mouvement brownien, martingales et calcul stochastique*, volume 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2013. ISBN 978-3-642-31897-9; 978-3-642-31898-6. doi: 10.1007/978-3-642-31898-6. URL `http://dx.doi.org/10.1007/978-3-642-31898-6`.

[49] Pierre L'Ecuyer and Gaétan Perron. On the convergence rates of IPA and FDC derivative estimators. *Oper. Res.*, 42(4):643–656, 1994. ISSN 0030-364X. doi: 10.1287/opre.42.4.643. URL `http://dx.doi.org/10.1287/opre.42.4.643`.

[50] Paul Malliavin. *Stochastic analysis*, volume 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. ISBN 3-540-57024-1. doi: 10.1007/978-3-642-15074-6. URL `http://dx.doi.org/10.1007/978-3-642-15074-6`.

[51] P. A. Meyer. Un cours sur les intégrales stochastiques. pages 245–400. Lecture Notes in Math., Vol. 511, 1976.

[52] Yuliya Mishura and Alexander Schied. Constructing functions with prescribed pathwise quadratic variation. *J. Math. Anal. Appl.*, 442(1):117–137, 2016. ISSN 0022-247X. doi: 10.1016/j.jmaa.2016.04.056. URL `http://dx.doi.org/10.1016/j.jmaa.2016.04.056`.

[53] David Nualart. *Malliavin calculus and its applications*, volume 110 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2009. ISBN 978-0-8218-4779-4. doi: 10.1090/cbms/110. URL `http://dx.doi.org/10.1090/cbms/110`.

[54] Daniel Ocone. Malliavin's calculus and stochastic integral representations of functionals of diffusion processes. *Stochastics*, 12(3-4):161–185, 1984. ISSN 0090-9491. doi: 10.1080/17442508408833299. URL `http://dx.doi.org/10.1080/17442508408833299`.

[55] Gilles Pagès. Introduction to numerical probability for finance, 2016. Lecture notes of Master 2 Probabilités et Finance, `http://www.proba.jussieu.fr/pageperso/pages`.

[56] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lect. Notes Control Inf. Sci.*, pages 200–217. Springer, Berlin, 1992. doi: 10.1007/BFb0007334. URL `http://dx.doi.org/10.1007/BFb0007334`.

[57] É. Pardoux and D. Talay. Discretization and simulation of stochastic differential equations. *Acta Appl. Math.*, 3(1):23–47, 1985. ISSN 0167-8019. doi: 10.1007/BF01438265. URL `http://dx.doi.org/10.1007/BF01438265`.

[58] ShiGe Peng and FaLei Wang. BSDE, path-dependent PDE and nonlinear Feynman-Kac formula. *Sci. China Math.*, 59(1):19–36, 2016. ISSN 1674-7283.

[59] Nicolas Perkowski and David J. Prömel. Pathwise stochastic integrals for model free finance. *Bernoulli*, 22(4):2486–2520, 2016.

[60] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. ISBN 3-540-00313-4. doi: 10.1007/978-3-662-10061-5. URL `http://dx.doi.org/10.1007/978-3-662-10061-5`. Second edition. Version 2.1, Corrected third printing.

[61] Zhenjie Ren, Nizar Touzi, and Jianfeng Zhang. An overview of viscosity solutions of path-dependent PDEs. In *Stochastic analysis and applications 2014*, volume 100 of *Springer Proc. Math. Stat.*, pages 397–453. Springer, Cham, 2014. doi: 10.1007/978-3-319-11292-3_15. URL `http://dx.doi.org/10.1007/978-3-319-11292-3_15`.

[62] Zhenjie Ren, Nizar Touzi, and Jianfeng Zhang. Comparison of viscosity solutions of semi-linear path-dependent pdes. 2014. Preprint, `https://arxiv.org/abs/1410.7281`.

[63] Zhenjie Ren, Nizar Touzi, and Jianfeng Zhang. Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent pdes. 2015. Preprint, https://arxiv.org/abs/1511.05910.

[64] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7. doi: 10.1007/978-3-662-06400-9. URL http://dx.doi.org/10.1007/978-3-662-06400-9.

[65] Candia Riga. *Pathwise functional calculus and applications to continuous-time finance*. PhD thesis, Université Pierre et Marie Curie and Scuola Normale Superiore di Pisa, 2015.

[66] Alexander Schied. Model-free CPPI. *J. Econom. Dynam. Control*, 40: 84–94, 2014. ISSN 0165-1889. doi: 10.1016/j.jedc.2013.12.010. URL http://dx.doi.org/10.1016/j.jedc.2013.12.010.

[67] Alexander Schied. On a class of generalized Takagi functions with linear pathwise quadratic variation. *J. Math. Anal. Appl.*, 433(2):974–990, 2016. ISSN 0022-247X. doi: 10.1016/j.jmaa.2015.08.022. URL http://dx.doi.org/10.1016/j.jmaa.2015.08.022.

[68] Alexander Schied and Iryna Voloshchenko. Pathwise no-arbitrage in a class of delta hedging strategies. *Probability, Uncertainty and Quantitative Risk*, 1(1):3, 2016. ISSN 2367-0126. doi: 10.1186/s41546-016-0003-2. URL http://dx.doi.org/10.1186/s41546-016-0003-2.

[69] Alexander Schied, Leo Speiser, and Iryna Voloshchenko. Model-free portfolio theory and its functional master formula. 2016. Preprint, https://arxiv.org/abs/1606.03325.

[70] Ichiro Shigekawa. Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.*, 20(2):263–289,

1980. ISSN 0023-608X. doi: 10.1215/kjm/1250522278. URL http://dx.doi.org/10.1215/kjm/1250522278.

[71] Dieter Sondermann. *Introduction to stochastic calculus for finance*, volume 579 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-34836-8; 3-540-34836-0. A new didactic approach.

[72] Daniel W. Stroock. The Malliavin calculus, a functional analytic approach. *J. Funct. Anal.*, 44(2):212–257, 1981. ISSN 0022-1236. doi: 10.1016/0022-1236(81)90011-2. URL http://dx.doi.org/10.1016/0022-1236(81)90011-2.

[73] Jianfeng Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 2004. ISSN 1050-5164. doi: 10.1214/aoap/1075828058. URL http://dx.doi.org/10.1214/aoap/1075828058.