# TWO TOPICS IN FINANCIAL MATHEMATICS:

# Forward Utility and Consumption Functions

# &

# Hedging with Variance Swaps in Infinite Dimensions

FRANCOIS BERRIER

Wolfson College & Statistical Laboratory

THIS DISSERTATION IS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN

PURE MATHEMATICS AND MATHEMATICAL STATISTICS

SEPTEMBER 2009

**DECLARATION OF ORIGINALITY**

The whole content of my thesis is based on work which I have done jointly with my supervisor, Dr. Michael Tehranchi.

In Chapter 1, the idea of adding consumption in a time-consistent manner to the definition of forward utility functions (i.e. in the manner described in Definition 1.1.1) is novel and to the best of my knowledge has not been exploited before. Therefore, all the results of Chapter 1 which relate to forward utility and consumption functions are original. The last section of Chapter 1, especially Theorem 1.4.1 and its corollaries on the characterization of decreasing forward dynamic utilities (without consumption) as integral transforms of positive measures, is also original and has given rise to the preprint [1].

The main results of Chapter 2, i.e. Theorem 2.5.2, Theorem 2.6.15 and Theorem 2.6.18 and the associated examples are based on similar results of Carmona and Tehranchi ([7], [8]) on infinite dimensional interest rates models, but are derived in the quite different setup of variance swaps modelling, and are therefore new.

**ABSTRACT**

Financial Mathematics is often presented as being composed of two main branches: one dealing with investment and consumption, with the aim of answering the now ancient question of how people should invest and spend their money, and the other dealing with the pricing and hedging of derivative instruments. This distinction between both branches of Financial Mathematics is reflected in my thesis, which is a compilation of two very different subjects on which I have worked during the past three years.

The first chapter, entitled “Forward Utility and Consumption Functions”, contributes to the investment branch of Financial Mathematics. Forward utilities have been introduced (under different names) a few years ago by Musiela and Zariphopoulou on the one hand, and by Henderson and Hobson on the other hand. Their idea is to define families (indexed by time and randomness) of utility functions which make the investment decisions of agents consistent over time. The contribution of this chapter is to extend the definition of forward utilities by adding consumption into the story and by giving explicit ways of constructing consumption functions from utilities and vice versa. The last part of this first chapter characterizes, in a Laplace integral form, the decreasing forward utilities (without consumption, and subject to some regularity conditions).

The second chapter, entitled “Hedging with Variance Swaps in Infinite Dimensions”, contributes to the derivatives pricing and hedging branch of Financial Mathematics. It is at the interface between the works of Buehler, who has shown that one could apply the HJM framework to model (forward) variance swaps curves, and the works of Carmona and Tehranchi, who have proved that infinite dimensional interest rates models can display theoretically nice features which are absent from their finite dimensional counterpart, such as uniqueness and maturity-specific properties of hedging portfolios for contingent claims. After an introductory section on terminology and after explaining the Buehler-HJM framework, I give a concrete example of finite dimensional model and show its (theoretical) shortcomings. I then port some results of Carmona and Tehranchi from interest rates modelling to variance swaps modelling in infinite dimensions and

finally give a concrete example of model and of classical payoffs to which the results apply.

Because many results and prerequisites to this chapter are quite technical, I have added a short appendix, giving modest introductions to infinite dimensional stochastic analysis, Malliavin calculus and SPDEs in Hilbert spaces.

## ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Michael Tehranchi, for his patience and support during the past three years. He has given me very interesting and challenging subjects to work on. I admire very much his intelligence, his incredible knowledge and understanding of Mathematical Finance and Mathematics in general, and his kindness.

I am dedicating this thesis to my family: my parents Sylvie and Gerard Berrier, my brother Constant, my sister in law Muriel and my niece bebe Jeanne, my grand parents Mata and Jacques Faure, my late Mamie Ambroisine and Grand-Pere Louis Berrier, my uncle and aunt Stephan and Avona Rebours, my family in law Thay Chieu, Co Thanh and Em Nguyen Binh Dinh.

Foremost, I dedicate it to my loved and loving little wife Thanh Nha who has supported me enormously all along, and especially when I thought I would never see the light at the end of the PhD tunnel.

Finally, I would also like to thank Karen Romani, Ulrich Jetzek and Markus Spahn from Ericsson for their support and understanding.

# Contents

# Chapter 1

# Forward Utility and Consumption Functions

**Abstract:** Recently, the notion of time-consistent utility functions has appeared in the mathematical finance literature, independently introduced by Zariphopoulou and Musiela (under the original name of "forward dynamic utilities", and later renamed into "forward performance") and by Henderson and Hobson (under the name of "horizon-unbiased utility functions"). To summarize, their idea has been to define utility functions $U(t, x, \omega)$, depending on time, on an agent's level of (discounted) wealth and on randomness, and for which the classical problem of finding the optimal strategy $\pi$, which maximizes the expected utility of wealth $\mathbb{E}_t\Big[U(T, X_T^{(\pi)}(\omega), \omega)\Big]$, gives results which are independent of the horizon $T$.

In this chapter, we broaden their definition by introducing consumption in the story: our agent does not only invest in a financial market, but also consumes a part of her wealth at each instant. This gives rise to the definition of forward utility and consumption functions. We aim then at studying and finding pairs of utility and consumption functions $U$ and $U^{(c)}$ for which the optimal strategy and consumption $(\pi, c)$ which maximize the quantity:

$$\mathbb{E}_t\Big[U(T, X_T^{(\pi,c)}) + \int_{u=t}^{T} U^{(c)}(u, c_u)du\Big]$$

are independent of the horizon $T$.

The plan is as follows: the first section serves as an introduction, setting up the investment world, and stating a few assumptions on our market model that will hold all along. We then give our definition of forward utility and consumption functions and show that this definition is sound, in the sense that pairs of functions that satisfy this definition indeed lead to solutions for the utility maximization problem which are independent of the horizon considered. We finish the introductory section by showing that utility and consumption functions do exist so that it makes sense to study them.

The second section gives a sufficient condition on a consumption function $U^{(c)}$ that one could check and that guarantees the existence of an associated utility function $U$. We then build some examples via this sufficient condition.

The third section takes the opposite view compared to section 2: we start from a utility function $U$ and give a sufficient condition that guarantees the existence of an associated consumption $U^{(c)}$. Here again, we give examples built via this sufficient condition.

Finally, the fourth section deals with utility functions without consumption. We characterize in a Laplace integral form all the decreasing forward utility functions (subject to some smoothness assumptions (i.e. $C^{1,3}$) and satisfying the Inada conditions).

## 1.1 Introduction and setup

### 1.1.1 Motivation

In the classical investment setup, as in the case of the famous lifetime portfolio selection problem of Merton [31], an agent invests in a financial market and consumes part of her wealth up to an horizon $T$, supposed to represent the time at which she retires. In this setup, we postulate that the utility derived by the agent is the sum of her consumption utility between time 0 and time $T$ and of her terminal wealth's utility. The total utility $\mathcal{U}$ derived is thus given by an expression of the form:

$$\mathcal{U} = \mathbb{E}\Big[\int_{u=0}^{T} U^{(c)}(u, c_u)du + U(T, X_T^{(\pi,c)})\Big] \tag{1.1.1}$$

where $\pi$ and $c$ are the investment and consumption strategies between times 0 and $T$, $X_T^{(\pi,c)}$ is the agent's wealth at time $T$, and where the map $x \mapsto U(T, x)$ and each of the maps $x \mapsto U^{(c)}(t, x)$ for $t \in [0, T]$ are utility functions (i.e. strictly increasing and concave functions). Solving the investment/consumption problem consists then in looking for the optimal investment/consumption strategy up to time $T$ so as to maximize $\mathcal{U}$. That is, we try to find $(\pi, c)$ which maximize $\mathcal{U}$. Solving this problem is either done directly making use of the dynamic programming principle which leads to the famous HJB equation or by making use of the duality theory where the maximization problem over investment/consumption processes is replaced by a minimization problem over density processes of equivalent martingale measures. For an account of the first method, the reader is referred to Merton's seminal paper [31] and Karatzas and Shreve [26]. For an account of the second method, the reader is referred to Karatzas and Shreve [26] again or to Kramkov and Schachermayer [25].

In many cases, this classical approach is certainly fine, but there are also problems of interest where the introduction of an horizon time $T$ seems rather artificial. For instance, a fund manager may just aim at having her portfolio's value grow gradually as time passes, and consume a part of it (e.g. for her salary), but without having any terminal date $T$ in mind. In such a case (and others described in [33] and [21] for instance), it

may be better to have a framework in which no horizon date $T$ plays a particular role nor affects the problem's solution.
Indeed, if we consider our fund manager's problem, we see easily that a priori (unless the functions $U^{(c)}$ and $U$ have some particular time consistency properties), the choice of the horizon date $T$ would affect the solution. For instance, if our fund manager was to solve the investment/consumption problem from year to year, each time fixing $T$ one year ahead, or if she was to solve the problem by periods of two years at a time, she would probably end up taking different decisions.
This is the motivation for introducing forward utility and consumption functions. Namely, we are interested in pairs of (random) functions $U^{(c)}$ and $U$ for which the optimal arguments that maximize (1.1.1) are independent of the horizon $T$.

### 1.1.2 Setup

#### Investment world

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space supporting an adapted $N$-dimensional Brownian motion $W_t := (W_t^{(1)}, ..., W_t^{(N)})$. We assume that $\mathcal{F}_t$ satisfies the usual conditions and that $\mathcal{F}_0$ is trivial. The investment world is composed of $N+1$ assets $(P^{(0)}, P^{(1)}, ..., P^{(N)})$ whose prices are strictly positive continuous semi-martingales. $P_t^{(0)}$ is taken as a numeraire, in units of which all other prices, wealths and consumptions are expressed. To simplify notation, but without loss of generality, we can therefore (and do) consider that:

$$P_t^{(0)} = 1$$

The $N$ other "risky" assets $P_t^{(i)}$, $i \in \{1, ..., N\}$ have the following Ito decomposition:

$$dP_t^{(i)} = P_t^{(i)}\Big(\mu_t^{(i)} dt + \sum_{j=1}^{N} \sigma_t^{(i,j)} dW_t^{(j)}\Big)$$

where the coefficients $\{\mu_t^{(i)}\}_{i\in\{1,...,N\}}$ and $\{\sigma_t^{(i,j)}\}_{(i,j)\in\{1,...,N\}^2}$ are $\mathcal{F}_t$-adapted processes satisfying suitable conditions for the above SDEs to have well defined positive solutions.
Notice that we do not assume that the market is complete, i.e. $\mathcal{F}_t$ may well be (strictly) larger than $\mathcal{F}_t^{(W)}$, the filtration generated by the Brownian motion $W$.

We further assume that for all $t$, almost surely, the matrix $\sigma_t := (\sigma_t^{(i,j)})_{(i,j)\in\{1,...,N\}^2}$ is invertible, and we define the "market price of risk" vector:

$$\Theta_t = \sigma_t^{-1}\mu_t$$

where $\mu_t := (\mu_t^{(1)}, ..., \mu_t^{(N)})$. We will assume that $||\Theta_t||$ is bounded from above uniformly in $t$ and $\omega$. We will later use the notation $A_t(\omega) := \frac{1}{2}\int_{u=0}^t ||\Theta_u(\omega)||^2 du$ and $A'_t(\omega)$ to denote the derivative of $A$ with respect to $t$, i.e. $A'_t(\omega) = \frac{1}{2}||\Theta_t(\omega)||^2$. We will finally assume that, almost surely, $\lim_{t\to\infty} A_t = \infty$. This guarantees that the change of time $t(\omega) \mapsto \tau(\omega)$ that we will do at a later point by defining $\tau(\omega) := A_t(\omega)$ is bijective from $[0, \infty)$ into itself. It is interesting to note that all along, we will never change numeraire. $P^{(0)}$ is and remains our "reference" asset. For an analysis of what happens under change of numeraire, the reader is invited to refer to El Karoui and M'rad [16] and Musiela and Zariphopoulou [33].

**Equivalent martingale measure**

We denote by $Z$ the martingale:

$$Z_t = \exp\Big( - \int_{u=0}^t \Theta_u^T dW_u - \frac{1}{2}\int_{u=0}^t ||\Theta_u||^2 du\Big) =: \mathcal{E}( - \int_{u=0}^t \Theta_u^T dW_u)$$

where we use $\mathcal{E}$ as a notation for the Doleans exponential local-martingale. $Z_t$ is the density process of an equivalent martingale measure. Notice that our assumption that $\Theta$ is bounded implies that $Z$ is a true martingale, and not only a local-martingale. Notice also that $Z$ corresponds to the Follmer-Schweizer minimal martingale measure (see [19]). We do not make any completeness assumption, and therefore $Z$ is not necessarily the unique (density process of an) equivalent martingale measure.

**Investment strategies, consumptions, wealths and admissible strategies**

We identify any $\mathbb{R}^N$-valued, measurable, adapted and $P$-integrable process $(\pi_u)_{u\geq 0}$ and any measurable, adapted, scalar positive integrable process $(c_u)_{u\geq 0}$ with, respectively, an investment strategy and a consumption process. Their interpretation is as follows: for $i \in \{1, ..., N\}$, $\pi_u^{(i)}$ represents the quantity of asset $P^{(i)}$ held by the agent at time $u$, and for any times $t, T$ with $t \leq T$, $\int_t^T c_u du$ represents the quantity of wealth consumed by the agent between time $t$ and time $T$. We define the wealth process of an agent following the investment strategy $\pi$ and consuming $c$ by her initial wealth $X_0$, and then by the relation: $X_t := X_0 + \int_{u=0}^t \pi_u dP_u - \int_{u=0}^t c_u du$. Although notationally heavier, we generally write $X_t^{(\pi,c)}$ to stress the dependency of $X_t$ on the strategy and consumption followed by the agent.

For any two times $t, T$ with $0 \leq t \leq T \leq \infty$, any $x > 0$ representing the wealth of the agent at time $t$, we denote by $\mathcal{A}_{t,T}^{(x)}$ the set of investment strategies $(\pi_u)_{u\in[t,T]}$ and of consumptions $(c_u)_{u\in[t,T]}$ which lead to positive wealths at all times, i.e. such that, almost surely for all $\tau \in [t, T]$:

$$X_\tau^{(\pi,c)} = x + \int_{u=t}^{\tau} \pi_u dP_u - \int_{u=t}^{\tau} c_u du \geq 0$$

Not only such a restriction makes sense from an economic point of view (the agent is not allowed to go in debt), but it is also mathematically convenient (because the agent cannot implement doubling strategies in finite time). Finally we can remark that $\mathcal{A}_{t,T}^{(x)}$ is a convex set for all $t, T, x$, by linearity of Riemann and Ito integrals.

### 1.1.3 Definition

#### Forward utility and consumption functions

**Definition 1.1.1** *Two measurable functions $U$ and $U^{(c)}$ from $\Big([0,\infty)\times(0,\infty)\times\Omega, \mathcal{B}[0,\infty)\times \mathcal{B}(0,\infty)\times\mathcal{F}\Big)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ are called forward utility and consumption functions (associated with each other) if:*

1. *For each $x > 0$, $(t, \omega) \mapsto U(t, x, \omega)$ and $(t, \omega) \mapsto U^{(c)}(t, x, \omega)$ are stochastic processes adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$.*

2. *The maps $x \mapsto U(t, x, \omega)$ and $x \mapsto U^{(c)}(t, x, \omega)$ are strictly increasing and strictly concave for $Leb[0,\infty) \times \mathbb{P}$-almost all $(t, \omega) \in [0,\infty) \times \Omega$*

3. *For all $x > 0$, all $T \geq t \geq 0$ and all $(\pi, c) \in \mathcal{A}_{t,T}^{(x)}$:*

$$U(t, x) \geq \mathbb{E}_t\Big[U(T, X_T^{(\pi,c)}) + \int_{u=t}^{T} U^{(c)}(u, c_u)du\Big]$$

4. *For all $x > 0$, all $T \geq t \geq 0$, there exists an optimal $(\pi^*, c^*) \in \mathcal{A}^{(x)}_{t,T}$ such that:*

$$U(t,x) = \mathbb{E}_t\Big[U(T, X_T^{(\pi^*,c^*)}) + \int_{u=t}^{T} U^{(c)}(u, c_u^*)du\Big]$$

**Uniqueness of optimal strategy, consumption and wealth**

From the above definition, it is not entirely obvious that the optimal solution to (1.1.1) is independent of $T$. That was what motivated us to define forward utility and consumption functions though, therefore we ought to check that this is indeed the case, and state precisely in what sense this is true:

**Theorem 1.1.2** *Uniqueness of optimal solution*
*Let $(U, U^{(c)})$ be associated utility and consumption functions. Then, for all $x > 0$, all $t \geq 0$, there exist an optimal strategy/consumption pair $\big((\pi_u^*)_{u\geq t}, (c_u^*)_{u\geq t}\big) \in \mathcal{A}^{(x)}_{t,\infty}$, and an associated optimal wealth process $(X_u^*)_{u\geq t} := (X_u^{(\pi^*,c^*)})_{u\geq t}$ such that the process:*

$$M_T := U(T, X_T^*) + \int_{u=t}^{T} U^{(c)}(u, c_u^*)du$$

*is an $(\mathcal{F}_T)_{T\geq t}$-martingale.*

*Moreover, if $(X_T^{**})_{T\geq t}, (c_T^{**})_{T\geq t}, (\pi_T^{**})_{T\geq t}$ are wealth, consumption and investment strategy processes for which the same property is true, then it holds that:*

- *$(X_T^*)_{T\geq t}$ and $(X_T^{**})_{T\geq t}$ are indistinguishable,*
- *$c^* = c^{**}$ $\mathbb{P} \times Leb[t, \infty)$ almost surely, and*
- *$\pi^* = \pi^{**}$ $\mathbb{P} \times Leb[t, \infty)$ almost surely.*

<u>Proof of Theorem 1.1.2</u>:
Let $t, T_1$ and $T_2$ be three times satisfying $t \leq T_1 \leq T_2$. Let $x > 0$. Let $\pi^{(1)}$ and $c^{(1)}$ denote an optimal strategy and consumption in $\mathcal{A}^{(x)}_{t,T_1}$, and let $\pi^{(2)}$ and $c^{(2)}$ denote an optimal

strategy and consumption in $\mathcal{A}^{(x)}_{t,T_2}$. Such optima exists by (4) of Definition 1.1.1.
Then by optimality of $\pi^{(1)}$ and $c^{(1)}$, it holds that:

$$\begin{aligned}\mathbb{E}_t\Big[U(T_1, X_{T_1}^{(\pi^{(2)},c^{(2)})}) + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(2)})du\Big] &\leq \mathbb{E}_t\Big[U(T_1, X_{T_1}^{(\pi^{(1)},c^{(1)})}) + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(1)})du\Big] \\ &= U(t,x)\end{aligned}$$

But we also have, by the tower property of conditional expectations, that:

$$\begin{aligned}U(t,x) &= \mathbb{E}_t\Big[U(T_2, X_{T_2}^{(\pi^{(2)},c^{(2)})}) + \int_{u=t}^{T_2} U^{(c)}(u, c_u^{(2)})du\Big] \\ &= \mathbb{E}_t\Big[\mathbb{E}_{T_1}[U(T_2, X_{T_2}^{(\pi^{(2)},c^{(2)})}) + \int_{u=t}^{T_2} U^{(c)}(u, c_u^{(2)})du]\Big] \\ &= \mathbb{E}_t\Big[\mathbb{E}_{T_1}[U(T_2, X_{T_2}^{(\pi^{(2)},c^{(2)})}) + \int_{u=T_1}^{T_2} U^{(c)}(u, c_u^{(2)})du] + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(2)})du\Big] \\ &\leq \mathbb{E}_t\Big[U(T_1, X_{T_1}^{(\pi^{(2)},c^{(2)})}) + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(2)})du\Big]\end{aligned}$$

Combining both inequalities, we get that:

$$\begin{aligned}U(t,x) &= \mathbb{E}_t\Big[U(T_1, X_{T_1}^{(\pi^{(2)},c^{(2)})}) + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(2)})du\Big] \\ &= \mathbb{E}_t\Big[U(T_1, X_{T_1}^{(\pi^{(1)},c^{(1)})}) + \int_{u=t}^{T_1} U^{(c)}(u, c_u^{(1)})du\Big]\end{aligned}$$

Therefore, $(X_{T_1}^{(\pi^{(1)},c^{(1)})}, \pi^{(1)}, c^{(1)})$ and $(X_{T_1}^{(\pi^{(2)},c^{(2)})}, \pi^{(2)}, c^{(2)})$ are both optimum arguments which maximize the quantity:

$$\operatorname{esssup}_{\xi,\pi,c}\, \mathbb{E}_t\Big[U(T_1, \xi) + \int_{u=t}^{T_1} U^{(c)}(u, c_u)du\Big]$$

with the following constraints: $\xi(\pi, c) = x + \int_{u=t}^{T_1} \pi_u dP_u - \int_{u=t}^{T_1} c_u du$ and $(\pi, c) \in \mathcal{A}^{(x)}_{t,T_1}$.
This implies that $X_{T_1}^{(1)} = X_{T_1}^{(2)}$ almost surely, for otherwise, by convexity of the set $\mathcal{A}^{(x)}_{t,T_1}$, the agent could use the admissible strategy/consumption $(\pi^*, c^*) = \frac{1}{2}\Big((\pi^{(1)}, c^{(1)}) + (\pi^{(2)}, c^{(2)})\Big)$ which would be a strictly better optimizer (for the above optimization problem) than both strategies/consumption $(\pi^{(1)}, c^{(1)})$ and $(\pi^{(2)}, c^{(2)})$, a contradiction.
The continuity of wealth processes and the representation $X_T = x + \int_{u=t}^{T} \pi_u dP_u - \int_{u=t}^{T} c_u du$ yield then the uniqueness statement. ■

### 1.1.4 Do forward utility and consumption functions exist at all?

We may indeed wonder whether this is ever possible to construct associated forward utility and consumption functions. The answer is fortunately yes. Actually, it is not too difficult to find examples based on the setup of Merton's lifetime portfolio and consumption selection problem (see [31]) where the stock is assumed to be evolving as a geometric Brownian motion, i.e. $P$ satisfies the constant coefficients SDE: $dP_t = P_t\mu dt + P_t\sigma dW_t$ where $W$ is a scalar Brownian motion. We can consider without loss that the short interest rate $r$ is constant equal to 0. It is then possible to verify that for any $\gamma \in (-\infty, 1)$, $\gamma \neq 0$, the following utility and consumption functions are forward utility and consumption functions associated with one another:

$$U(t,x) = U^{(c)}(t,x) = e^{-\rho t}\frac{x^\gamma}{\gamma}$$

where $\rho$ is given by:

$$\rho = 1 - \gamma - \frac{\gamma}{2(\gamma-1)}\Theta^2 \tag{1.1.2}$$

where $\Theta := \frac{\mu}{\sigma}$ is the (constant) market price of risk. We can indeed look for pairs of associated utility and consumption functions in that setup, of the form $U(t,x) = e^{-\rho t}\frac{x^\gamma}{\gamma}$ and $U^{(c)}(t,c) = e^{-\eta t}\frac{c^\delta}{\delta}$, for some constants $\rho, \eta, \gamma$ and $\delta$. We can then proceed as in the original paper from Merton by deriving and solving the HJB equation. It leads us to the conclusion that for $U$ and $U^{(c)}$ to be forward utility and consumption associated with each other, the only possibility is that $\rho = \eta$, $\gamma = \delta$. Moreover, $\rho$ and $\gamma$ have to be related through the above equation (1.1.2). Notice that condition (1.1.2) implies that, in order to keep the time consistency between utilities at different times, we cannot choose arbitrarily the time decay coefficient $\rho$ as is generally done in the case of traditional utility and consumption functions. $\rho$ has to be related to the market price of risk $\Theta$ and the risk aversion coefficient $\gamma$. This was already observed by Henderson and Hobson in ([21]).

Examples without consumption (i.e. $U^{(c)} \equiv 0$) have also been constructed and studied by Henderson and Hobson [21] in the context of an asset sale problem and by Musiela and Zariphopoulou ([33], [34]).

## 1.2 Constructing $U$ from $U^{(c)}$

### 1.2.1 Using Merton's infinite horizon problem

Given some consumption function $U^{(c)}$, does there exist an associated utility $U$ and can we build it from $U^{(c)}$? We will partially answer this question in this section, by giving a sufficient condition on $U^{(c)}$ which guarantees the existence of an associated $U$, and tells us how to construct it. The idea is based on Merton's infinite horizon problem. Indeed, if we could let $T$ go to infinity in point (4) of the definition of forward utility and consumption functions, and if we had that $\lim_{T\to\infty} \mathbb{E}_t U(T, X_T^{(\pi^*,c^*)}) = 0$ then we would have a direct relation between $U$ and $U^{(c)}$: the utility of wealth $U(t, X_t)$ at time $t$ would simply be the total utility of (optimal) consumption over the infinite horizon $[t, \infty)$. We use this observation to state the following proposition:

**Proposition 1.2.1** *Constructing U from* $U^{(c)}$
*Let* $U^{(c)}$ *be a measurable function from* $([0,\infty)\times(0,\infty)\times\Omega, \mathcal{B}[0,\infty)\times\mathcal{B}(0,\infty)\times\mathcal{F})$ *to* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*:* $(t, x, \omega) \to U^{(c)}(t, x, \omega)$*, strictly increasing and strictly concave in x for all t and all* $\omega$*, such that* $U^{(c)}(., x, .)$ *is adapted for all* $x > 0$*, and such that for all* $t \geq 0$*, all* $x > 0$*, there exists* $(\pi^*, c^*) \in \mathcal{A}_{t,\infty}^{(x)}$ *such that:*

$$\mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^*)du = \operatorname{esssup}_{(\pi,c)\in\mathcal{A}_{t,\infty}^{(x)}} \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u)du < \infty$$

*Then, the function U defined by:*

$$U(t, x) := \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^*)du$$

*is a forward utility associated with* $U^{(c)}$*.*

Proof of Proposition 1.2.1:

Property (1) of Definition 1.1.1 is true by hypothesis for $U^{(c)}$ and by the definition of conditional expectations for $U$. We have to show that properties (2), (3) and (4) of Definition 1.1.1 are satisfied: we begin by proving the monotonicity of the map $x \mapsto U(t, x)$: let $0 < x_1 < x_2$, $t \geq 0$. Let us denote by $\pi^{(1)}, c^{(1)}$ and by $\pi^{(21)}, c^{(21)}$ the optimal trading strategy/consumption pairs associated with initial wealth $x_1$ and $x_2 - x_1$. Clearly, we have $\mathcal{A}_{t,\infty}^{(x_1)} + \mathcal{A}_{t,\infty}^{(x_2-x_1)} \subseteq \mathcal{A}_{t,\infty}^{(x_2)}$. Therefore:

$$U(t, x_2) \geq \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^{(1)} + c_u^{(21)})du$$

Then, by the strict monotonicity of $U^{(c)}$, we get that:

$$U(t, x_2) > \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^{(1)})du = U(t, x_1)$$

We continue the proof of point (2) by proving the concavity of the map $x \mapsto U(t, x)$: let $0 < x_1 < x_2$, $t > 0$, $\lambda \in (0, 1)$. Let us denote by $\pi^{(1)}, c^{(1)}$ and by $\pi^{(2)}, c^{(2)}$ the optimal trading strategy/consumption pairs associated with initial wealth $x_1$ and $x_2$. Clearly, we have $\lambda\mathcal{A}_{t,\infty}^{(x_1)} + (1 - \lambda)\mathcal{A}_{t,\infty}^{(x_2)} \subseteq \mathcal{A}_{t,\infty}^{(\lambda x_1+(1-\lambda)x_2)}$. Therefore,

$$U(t, \lambda x_1 + (1 - \lambda)x_2) \geq \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, \lambda c_u^{(1)} + (1 - \lambda)c_u^{(2)})du$$

Then by the strict concavity of $U^{(c)}$, we get that:

$$\begin{aligned} U(t, \lambda x_1 + (1 - \lambda)x_2) &> \lambda\mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^{(1)})du + (1 - \lambda)\mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^{(2)})du \\ &= \lambda U(t, x_1) + (1 - \lambda)U(t, x_2) \end{aligned}$$

We now prove points (3) and (4) of the definition: let $x > 0$, $T \geq t \geq 0$. Let $(\pi, c) \in \mathcal{A}_{t,T}^{(x)}$. Let us denote by $(\pi^*, c^*)$ the optimal strategy and consumption between times $t$ and $\infty$. Then we have that:

$$U(t, x) = \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u^*)du = \mathbb{E}_t\Big[ \int_{u=t}^{T} U^{(c)}(u, c_u^*)du + \mathbb{E}_T \int_{u=T}^{\infty} U^{(c)}(u, c_u^*)du\Big]$$

But it is easy to see that $(\pi^*, c^*)$ is a better strategy than $(\pi, c)\mathbb{1}_{[t,T]} + (\pi^{(T*)}, c^{(T*)})\mathbb{1}_{[T,\infty)}$, where $(\pi^{(T*)}, c^{(T*)})$ is the optimal strategy in $\mathcal{A}_{T,\infty}^{(X_T^{(\pi,c)})}$. Therefore, we have that:

$$\begin{aligned} U(t, x) &\geq \mathbb{E}_t\Big[ \int_{u=t}^{T} U^{(c)}(u, c_u)du + \mathbb{E}_T \int_{u=T}^{\infty} U^{(c)}(u, c_u^{(T*)})du\Big] \\ &= \mathbb{E}_t\Big[ \int_{u=t}^{T} U^{(c)}(u, c_u)du + U(T, X_T^{(\pi,c)})\Big] \end{aligned}$$

Obviously, if we take the optimal $(\pi^*, c^*)$ on $[t, \infty)$, we get that:

$$U(t,x) = \mathbb{E}_t\Big[\int_{u=t}^{T} U^{(c)}(u, c_u^*)du + U(T, X_T^{(\pi^*,c^*)})\Big] \ \blacksquare$$

### 1.2.2 Classical CRRA examples

Using the previous Proposition 1.2.1, we can build examples based on the classical CRRA family of utility functions: let $\alpha_t$ be a scalar process satisfying suitable conditions so that: $M_t := \mathcal{E}\big(-\int_{u=0}^{t}\alpha_u\theta_u^T dW_u\big)$ is a martingale. Let $f : [0,\infty) \to (0,\infty)$ be an integrable function, i.e. satisfying $\int_{u=0}^{\infty} f(u)du < \infty$, and let us denote by $F$ the opposite of its antiderivative (up to a constant), i.e. we set $F : t \mapsto \int_{s=t}^{\infty} f(s)ds$. The following examples then satisfy the hypothesis of Proposition 1.2.1:

- $U_{\log}^{(c)}(t,c) := M_t f(t)\log(c)$
- $U_{\text{pow}}^{(c)}(t,c) := Z_t^{\gamma} M_t^{1-\gamma} f(t)^{1-\gamma}\frac{c^{\gamma}}{\gamma}$

where $\gamma < 1$, $\gamma \neq 0$. Indeed, let $t$ and $x = X_t$ be given. We can then check that the optimum consumption, wealth and trading strategy are given for the above functions by:

- $c_u^* = x\frac{Z_t M_u f(u)}{Z_u M_t F(t)}$
- $X_u^* = x\frac{Z_t M_u F(u)}{Z_u M_t F(t)}$
- $\pi_u^* = X_u^*(1-\alpha_u)(\sigma_u^T)^{-1}\Theta_u$

The associated forward utilities are given by:

- $U_{\log}(t,x) = M_t F_t \log(x) + \mathbb{E}_t \int_{s=t}^{\infty} M_s f(s)\log\big(\frac{Z_t M_s f(s)}{Z_s M_t F_t}\big)ds$
- $U_{\text{pow}}(t,x) = \frac{F(t)^{1-\gamma}}{f(t)^{1-\gamma}} U_{\text{pow}}^{(c)}(t,x)$

### 1.2.3 Combining utilities by convex duality

We now show how to combine different already known utility/consumption pairs to build new ones. If $U_1^{(c)}$ and $U_2^{(c)}$ are two consumption functions satisfying the assumption of Proposition 1.2.1, we could define $\tilde{U}(t,x) := \text{esssup}_{(\pi,c)\in\mathcal{A}_{t,\infty}^{(x)}} \mathbb{E}_t \int_{s=t}^{\infty}\Big[U_1^{(c)}(s,c_s) +$

$U_2^{(c)}(s,c_s)\Big]ds$. However, because the optimal consumptions for $U_1^{(c)}$ and $U_2^{(c)}$ have no reason to be the same, there is little hope that $\tilde{U}$ so constructed be a utility associated with $U^{(c)} := U_1^{(c)} + U_2^{(c)}$. In order to combine different already known utilities, we will make use of duality arguments: let us suppose that $U$ and $U^{(c)}$ are a pair of associated utility and consumption functions satisfying the assumptions of Proposition 1.2.1, and that they are $C^1$ in $x$. Let us denote by $\mathcal{Z}_t$ the set of positive martingales $(Y_s)_{s\geq t}$ such that $(X_T Y_T + \int_{s=t}^T c_s Y_s ds)_{T\geq t}$ is a positive super-martingale for all admissible consumption and wealth processes $(c, X)$. We have then the dual inequality, valid for all $y > 0$:

$$V(t,y) \leq \inf_{Y\in\mathcal{Z}_t} \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, y\frac{Y_u}{Y_t})du \tag{1.2.1}$$

where $V$ and $V^{(c)}$ are the convex conjugates of $U$ and $U^{(c)}$. If, in addition, there exists an optimizer $Y^* \in \mathcal{Z}_t$ such that, for all $y > 0$:

- $(\tilde{c}^*_s)_{s\geq t} := (-V_y^{(c)}(s, y\frac{Y^*_s}{Y^*_t}))_{s\geq t}$ is an admissible consumption process (with wealth starting from some $x$ at time $t$), and
- $-V_y(t,y) = \mathbb{E}_t \int_{s=t}^{\infty} \tilde{c}^*_s \frac{Y^*_s}{Y^*_t} ds$

then we have that $Y^*$ gives equality in the dual inequality (1.2.1), i.e. we have that:

$$V(t,y) = \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, y\frac{Y^*_u}{Y^*_t})du$$

Based on these observations, we can give sufficient conditions on the convex conjugate $V^{(c)}$ of $U^{(c)}$, which may be in some cases more straightforward to check than their primal counterparts, and which allow us to guarantee the existence of an associated utility $U$:

**Proposition 1.2.2** *Constructing U from $U^{(c)}$ - dual formulation*
*Let $U^{(c)}$ be a measurable function from $([0,\infty) \times (0,\infty) \times \Omega, \mathcal{B}[0,\infty) \times \mathcal{B}(0,\infty) \times \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$: $(t,x,\omega) \rightarrow U^{(c)}(t,x,\omega)$, strictly increasing, strictly concave and of class $C^1$ in x for all t and all $\omega$, such that $U^{(c)}(.,x,.)$ is adapted for all $x > 0$. We denote by $V^{(c)}(t,y)$ its convex conjugate. Let us suppose that for all $t \geq 0$, $y > 0$ and almost all $\omega$, there is an optimizer $Y^* \in \mathcal{Z}_t$ such that:*

$$V(t,y) := \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, y\frac{Y^*_u}{Y^*_t})du = \mathrm{essinf}_{Y\in\mathcal{Z}_t} \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, y\frac{Y_u}{Y_t})du$$

*and that this infimum is well defined (i.e. neither $\infty$ nor $-\infty$). Let us also assume that for all $x > 0$, $T \geq t \geq 0$, almost all $\omega$, we can find $y > 0$ such that:*

- $V_y(t, y) = \mathbb{E}_t \int_{u=t}^{\infty} \frac{Y_u^*}{Y_t^*} V_y^{(c)}(u, y\frac{Y_u^*}{Y_t^*})du = -x$
- $(\tilde{c}_u)_{u\in[t,T]} := ( - V_y^{(c)}(u, y\frac{Y_u^*}{Y_t^*}))_{u\in[t,T]}$ *and* $\tilde{X}_T := -V_y(T, y\frac{Y_T^*}{Y_t^*})$ *are admissible consumption and (time $T$) wealth, starting from $t, x$.*

*Then $U$, the concave conjugate of $V(t, y) := \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, y\frac{Y_u^*}{Y_t^*})du$ is a utility function associated to $U^{(c)}$.*

Proof of Proposition 1.2.2: Let $x > 0$ and $t, T \geq 0, t \leq T$ be given, and $(c_u, \pi_u)_{u\in[t,T]} \in \mathcal{A}_{t,T}^{(x)}$ and $(X_u)_{u\in[t,T]}$ the associated admissible wealth. Let $y := -U_x(t, x)$. By definition of convex/concave conjugation, it holds that:

$$
\begin{aligned}
U(t, x) - xy = V(t, y) &= \mathbb{E}_t\Big[ \int_{u=t}^{T} V^{(c)}(u, y\frac{Y_u^*}{Y_t^*})du + \int_{u=T}^{\infty} V^{(c)}(u, y\frac{Y_u^*}{Y_t^*})du\Big] \\
&\geq \mathbb{E}_t\Big[ \int_{u=t}^{T} (U^{(c)}(u, c_u) - c_u y\frac{Y_u^*}{Y_t^*})du + V(T, y)\Big] \\
&\geq \mathbb{E}_t\Big[ \int_{u=t}^{T} (U^{(c)}(u, c_u) - c_u y\frac{Y_u^*}{Y_t^*})du + U(T, X_T) - X_T y\frac{Y_T^*}{Y_t^*}\Big]
\end{aligned}
$$

From the budget constraint, we now get that:

$$
U(t, x) \geq \mathbb{E}_t\Big[ \int_{u=t}^{T} U^{(c)}(u, c_u)du + U(T, X_T)\Big]
$$

Finally, we can obtain equality in place of the inequalities in all of the above if we use the optimal consumption and the optimal wealth given in the proposition (and which are admissible by assumption). ■

Now we can see that in the case of our earlier CRRA examples, although they do not have the same "primal" optimizer $c^*$ (i.e. different values of $\gamma$ correspond to different optimal consumptions), they on the other hand have the same "dual" optimizer $Y^*$, namely the minimal martingale $Z$. That means that we can combine our CRRA-like consumptions by taking convex combinations of their convex conjugate, then define $V$, the appropriate associated convex conjugate utility as in Proposition 1.2.2 and finally define

$U$ the concave conjugate of $V$. The following family of forward utility and consumption functions is based on this method:

**Proposition 1.2.3** *First parametric family of forward utility and consumption functions Let:*

- *$\nu$ be a positive Borel measure such that $\int_{r>0} y^{-r}\nu(dr) < \infty$ for all $y > 0$.*
- *$\delta$ be a function from $(0, \infty)$ into itself, bounded away from $0$.*
- *$M_t^{(\alpha)}$ be a strictly positive martingale started at $1$ given by $dM_t^{(\alpha)} = -M_t^{(\alpha)}\alpha_t\Theta_t^T dW_t$, for some scalar process $\alpha_t$.*

*Then, the two functions:*

$$V(t, y) := \int_{r>0} M_t \frac{1 - (\frac{y}{Z_t})^{1-r}}{\delta(r)(1-r)} e^{-\delta(r)t}\nu(dr)$$

*and*

$$V^{(c)}(t, y) := \int_{r>0} M_t \frac{1 - (\frac{y}{Z_t})^{1-r}}{1-r} e^{-\delta(r)t}\nu(dr)$$

*are the convex conjugates of associated forward utility and consumption functions.*

Proof of Proposition 1.2.3:

The functions $V$ and $V^{(c)}$ defined as above are smooth functions of $y$, and are strictly decreasing and convex in $y$. We can define $U$ and $U^{(c)}$ the concave conjugates of $V$ and $V^{(c)}$ which are then strictly increasing and concave in $x$. Remains to prove properties (3) and (4) of Definition 1.1.1.

Fubini's theorem, as everything is explicit, allows us to show that, for all $z > 0$:

$$V(t, z) = \mathbb{E}_t \int_{u=t}^{\infty} V^{(c)}(u, z\frac{Z_u}{Z_t})du$$

This, plus the budget constraint over $[t, \infty)$ and the definition of convex conjugates yield, for any $y, x > 0$ and any admissible consumption $(c_u)_{u \geq t}$:

$$V(t, y) + yx \geq \mathbb{E}_t \int_{u=t}^{\infty} (V^{(c)}(u, y\frac{Z_u}{Z_t}) + y\frac{Z_u}{Z_t}c_u)du \geq \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u)du$$

Taking the infimum over $y > 0$, we get:

$$U(t,x) \geq \mathbb{E}_t \int_{u=t}^{\infty} U^{(c)}(u, c_u) du$$

To get equality in place of all of the above inequalities, we can take the optimal consumption:

$$c_u^* = -V_y^{(c)}(u, y\frac{Z_u}{Z_t}) = \frac{M_u}{Z_u} \int_{r>0} \frac{y^{-r}}{Z_t^{-r}} e^{-\delta(r)u} \nu(r)$$

where $y$ is the unique solution of:

$$x = -V_y(t,y)$$

and take the optimal wealth

$$X_u^* = \frac{M_u}{Z_u} \int_{r>0} \frac{y^{-r}}{Z_t^{-r}} \frac{e^{-\delta(r)u}}{\delta(r)} \nu(dr)$$

which is attainable with consumption $c^*$ as above along with the following optimal trading strategy:

$$\pi_u^* = X_u^*(1-\alpha_u)(\sigma_u^{-1})^T \Theta_u^T$$

■

## 1.3 Recovering $U^{(c)}$ from $U$

We now look at the inverse problem compared to the previous section, i.e. given a utility function $U$, does there exists an associated consumption function $U^{(c)}$, and can we recover it from $U$? Here again, we partially answer the question, by giving sufficient conditions under which we can recover an associated consumption function $U^{(c)}$ from $U$. We will place ourselves in a particular setting where we assume $U$ to be two times continuously differentiable in $x$ and once continuously differentiable in $t$, and where $U(t, ., \omega)$ is assumed to satisfy the Inada conditions for all $t$ and $\omega$ (i.e. $\lim_{x\to 0} U_x(t, x, \omega) = \infty$ and $\lim_{x\to\infty} U_x(t, x, \omega) = 0$).

We begin by showing that under these assumptions, any pair $(U, U^{(c)})$ of associated utility and consumption functions must satisfy a random non-linear partial differential equation (PDE), and that this PDE becomes linear if we change $U$ and $U^{(c)}$ into their respective convex conjugate $V$ and $V^{(c)}$. We start with the random non linear PDE:

**Proposition 1.3.1** *Non-Linear PDE*
*Let $(U, U^{(c)})$ be associated utility and consumption functions. Suppose further that U is two times continuously differentiable with respect to x and one time continuously differentiable with respect to t. Then, for all $x > 0$ and all $t \geq 0$, the following random non-linear PDE must hold $\mathbb{P}$-almost surely:*

$$U_t(t, x) - \frac{1}{2}\|\Theta(t, \omega)\|^2 \frac{U_x^2(t, x)}{U_{xx}(t, x)} + V^{(c)}(t, U_x(t, x)) = 0 \tag{1.3.1}$$

*where $V^{(c)}$ is the convex conjugate of $U^{(c)}$.*

Proof of Proposition 1.3.1:

The variation of our agent's wealth, taking into account the budget and self-financing constraints, is equal to:

$$dX_u^{(\pi,c)} = \Big[\pi_u^T \sigma_u \Theta_u - c_u\Big] du + \pi_u^T \sigma_u dW_u \tag{1.3.2}$$

Thus, using the generalized Ito's formula (see Theorem 3.3.1 of Kunita's book [28] for instance), we get:

$$
\begin{aligned}
U(T, X_T^{(\pi,c)}) &= U(t,x) + \int_{u=t}^{T} U_t(u, X_u^{(\pi,c)})du \\
&+ \int_{u=t}^{T} U_x(u, X_u^{(\pi,c)})\Big[(\pi_u^T \sigma_u \Theta_u - c_u)du + \pi_u^T \sigma_u dW_u\Big] \\
&+ \frac{1}{2}\int_{u=t}^{T} U_{xx}(u, X_u^{(\pi,c)})\|\pi_u^T \sigma_u\|^2 du.
\end{aligned}
$$

By definition of forward utility and consumption functions, $U(t, X_t^{(\pi,c)}) + \int_{u=0}^{t} U^{(c)}(u, c_u)du$ is a super-martingale for all $(\pi, c)$, and a martingale for $(\pi^*, c^*)$, which together with the above decomposition, imply that:

(1) For all $(\pi, c)$

$$
\begin{aligned}
&\int_{u=t}^{T} \Big[U_t(u, X_u^{(\pi,c)})du + U_x(u, X_u^{(\pi,c)})(\pi_u^T \sigma_u \Theta_u - c_u) \\
&+ \ \frac{1}{2} U_{xx}(u, X_u^{(\pi,c)})\|\pi_u^T \sigma_u\|^2 + U^{(c)}(u, c_u)\Big]du \ \leq \ 0
\end{aligned}
$$

(2) For the optimal $(\pi^*, c^*)$

$$
\begin{aligned}
&\int_{u=t}^{T} \Big[U_t(u, X_u^*)du + U_x(u, X_u^*)((\pi_u^*)^T \sigma_u \Theta_u - c_u^*) \\
&+ \ \frac{1}{2} U_{xx}(u, X_u^*)\|(\pi_u^*)^T \sigma_u\|^2 + U^{(c)}(u, c_u^*)\Big]du \ = \ 0
\end{aligned}
\tag{1.3.3}
$$

We now define a locally optimal solution by taking, for times $u$ between $t$ and $T$:

$$
\begin{aligned}
c_u^+ &= (U_x^{(c)})^{-1}(u, U_x(u,x))\mathbb{1}_{\{X_u^+ \geq 0\}} \\
\pi_u^+ &= -(\sigma_u^T)^{-1}\Theta_u \frac{U_x(u,x)}{U_{xx}(u,x)}\mathbb{1}_{\{X_u^+ \geq 0\}}
\end{aligned}
$$

By definition (thanks to the indicator function), this strategy/consumption pair is admissible. The super-martingale property for $U(t, X_t^+) + \int_{u=0}^{t} U^{(c)}(u, c_u^+)du$ yields:

$$
\begin{aligned}
&\int_{u=t}^{T} \Big[U_t(u, X_u^+) + U_x(u, X_u^+)\Big( - \frac{U_x(u,x)}{U_{xx}(u,x)} \Theta_u^T \sigma_u^{-1} \sigma_u \Theta_u \mathbb{1}_{\{X_u^+ \geq 0\}}\Big) \\
&+ \ \frac{1}{2} U_{xx}(u, X_u^+)\Big\|\frac{U_x(u,x)}{U_{xx}(u,x)} \Theta_u^T \sigma_u^{-1} \sigma_u \mathbb{1}_{\{X_u^+ \geq 0\}}\Big\|^2 \\
&- \ (U_x^{(c)})^{-1}(u, U_x(u,x))\mathbb{1}_{\{X_u^+ \geq 0\}} U_x(u, X_u^+) \\
&+ \ U^{(c)}(u, (U_x^{(c)})^{-1}(u, U_x(u,x))\mathbb{1}_{\{X_u^+ \geq 0\}})\Big]du \leq 0
\end{aligned}
$$

Letting now $T$ tend to $t$ and by continuity of all the quantities involved, we get that:

$$\begin{aligned} & U_t(t,x) - \frac{1}{2}\frac{U_x^2(t,x)}{U_{xx}(t,x)}\|\Theta_t\|^2 \\ & - \ (U_x^{(c)})^{-1}(t, U_x(t,x))U_x(t,x) + U^{(c)}\Big(t, (U_x^{(c)})^{-1}(t, U_x(t,x))\Big) \ \leq \ 0 \end{aligned} \tag{1.3.4}$$

This last inequality must be true for all $t \geq 0$ and all $x > 0$.

Making use of this, applied to the optimal strategy/consumption and combining it with the equality (1.3.3), we get that:

$$\begin{aligned} 0 \leq & \int_{u=t}^{T} \Big\{\frac{1}{2}\|\Theta_u\|^2 \frac{U_x^2(u,X_u^*)}{U_{xx}(u,X_u^*)} + U_x(u,X_u^*)(\pi_u^*)^T\sigma_u\Theta_u \\ & + \frac{1}{2}U_{xx}(u,X_u^*)\|(\pi_u^*)^T\sigma_u\|^2 \\ & + U^{(c)}(u,c_u^*) - c_u^* U_x(u,X_u^*) - \Big[U^{(c)}\Big(u, (U_x^{(c)})^{-1}(u, U_x(u,X_u^*))\Big) \\ & - (U_x^{(c)})^{-1}(u, U_x(u,X_u^*))U_x(u,X_u^*)\Big]\Big\}du \end{aligned}$$

which is after close inspection a sum of two negative terms, and that finally shows that the locally optimal strategy was actually the optimal one, i.e. we have:

$$\begin{aligned} c_u^* &= (U_x^{(c)})^{-1}(u, U_x(u,X_u^*)) \\ \pi_u^* &= -(\sigma_u^T)^{-1}\Theta_u \frac{U_x(u,X_u^*)}{U_{xx}(u,X_u^*)} \end{aligned}$$

This optimal solution is the only one which gives equality in (1.3.4), and this yields the PDE announced in the proposition. ■

We now linearize the above PDE by introducing $V$, the convex conjugate of $U$. This gives rise to the following proposition:

**Proposition 1.3.2** *Linear PDE*

*Let $(U, U^{(c)})$ be two utility and consumption functions, such that $U$ is twice continuously differentiable in $x$ and once continuously differentiable in $t$, and such that $U(t, ., \omega)$ satisfies the Inada conditions for all $t$ and $\omega$.*

*Then, for all $y > 0$ and all $t \geq 0$, the following random PDE must hold $\mathbb{P}$-almost surely:*

$$V_t(t, y, \omega) + \frac{1}{2} y^2 \|\Theta(t, \omega)\|^2 V_{yy}(t, y, \omega) + V^{(c)}(t, y, \omega) = 0 \tag{1.3.5}$$

*where $V$ and $V^{(c)}$ are the convex conjugates of, respectively, $U$ and $U^{(c)}$.*

<u>Proof of Proposition 1.3.2</u>:

We do the following change of variable in the random PDE of Proposition 1.3.2:

$(t, x, \omega) \mapsto (t, y := U_x(t, x, \omega), \omega)$

Notice that this change of variable is bijective from $[0, \infty) \times (0, \infty) \times \Omega$ into itself because of the Inada conditions assumption.

Then, we can make the observation that:

$$-\frac{1}{y^2} = \frac{\partial \frac{1}{y}}{\partial y} = \frac{\partial \frac{1}{U_x(t,x)}}{\partial x} \frac{\partial x}{\partial y} = \frac{U_{xx}(t, x)}{U_x^2(t, x)} V_{yy}(t, y)$$

and thus:

$$-\frac{U_x^2(t, x)}{U_{xx}(t, x)} = y^2 V_{yy}(t, y) \tag{1.3.6}$$

We also have that:

$$V(t, y) = U(t, U_x^{-1}(t, y)) - x U_x^{-1}(t, x)$$

Differentiating with respect to $t$, we get:

$$\begin{aligned} V_t(t, y) &= U_t(t, x) + \frac{\partial U_x^{-1}(t, x)}{\partial t} U_x(t, U_x^{-1}(t, x)) - x \frac{\partial U_x^{-1}(t, x)}{\partial t} \\ &= U_t(t, x) + x U_x(t, U_x^{-1}(t, x)) - x U_x(t, U_x^{-1}(t, x)) \end{aligned}$$

And therefore:

$$V_t(t, y) = U_t(t, x) \tag{1.3.7}$$

Replacing (1.3.6) and (1.3.7) in the random non-linear PDE of the previous section, we get Proposition 1.3.2. ■

The PDE of Proposition 1.3.2 is a necessary condition for $U$ and $U^{(c)}$ to be forward utility and consumption functions (provided they satisfy the conditions of the proposition). Combined with conditions which ensure that we can solve the optimization problem (1.1.1), the PDE can be used as a sufficient condition to recover $U^{(c)}$ from a suitable function $U$, as the next proposition shows:

**Proposition 1.3.3** *Recovering $U^{(c)}$ from U*
*Let U be a measurable function from $([0,\infty)\times(0,\infty)\times\Omega, \mathcal{B}[0,\infty)\times\mathcal{B}(0,\infty)\times\mathcal{F})$ to $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ satisfying the assumptions of Proposition 1.3.2. Suppose that, in addition, U satisfies the following conditions (where V denotes the convex conjugate of $U$):*

1. $V^{(c)}(t,y) := -V_t(t,y) - \frac{1}{2}y^2\|\Theta_t\|^2 V_{yy}(t,y)$ *is strictly decreasing and strictly convex in y.*

2. $(V(u,y\frac{Z_u}{Z_t}) + \int_{s=t}^u V^{(c)}(s,y\frac{Z_s}{Z_t})ds)_{u\geq t}$ *is an $(\mathcal{F}_u)_{u\geq t}$-martingale for all $y>0$, $t\geq 0$.*

3. *for all $T\geq t\geq 0$ and $x>0$, there exists $(\pi,c)\in\mathcal{A}_{t,T}^{(x)}$ such that*

   - $X_T^{(\pi,c)} = -V_y(T, U_x(t,x)\frac{Z_T}{Z_t})$
   - $c_u = -V_y^{(c)}(u, U_x(t,x)\frac{Z_u}{Z_t})$
   - $\mathbb{E}_t\Big[X_T^{(\pi,c)}Z_T + \int_{u=t}^T c_u Z_u du\Big] = xZ_t$

*Then, $U^{(c)}$, the concave conjugate of $V^{(c)}$ is a forward consumption function associated with $U$.*

Proof of Proposition 1.3.3:
We can define $U^{(c)}$ the concave conjugate of $V^{(c)}$. $U$ and $U^{(c)}$ have the correct monotonicity and concavity properties by assumption. We need only check the super-martingale and martingale properties (3) and (4) of Definition 1.1.1. Let then $x>0$, $T\geq t\geq 0$ and $(\pi,c)\in\mathcal{A}_{t,T}^{(x)}$ be given. For any $y>0$, condition 2 of the proposition, along with the budget constraint give us that:

$$V(t,y) + xy \geq \mathbb{E}_t\Big[V(T,y\frac{Z_T}{Z_t}) + \int_{u=t}^T V^{(c)}(u,y\frac{Z_u}{Z_t})du + yX_T^{(\pi,c)}\frac{Z_T}{Z_t} + \int_{u=t}^T yc_u\frac{Z_u}{Z_t}du\Big]$$

By definition of convex conjugates, the right hand side (RHS) is larger than:

$$RHS \geq \mathbb{E}_t\Big[U(T, X_T^{(\pi,c)}) + \int_{u=t}^{T} U^{(c)}(u, c_u)du\Big]$$

Now, taking the infimum of $V(t, y) + xy$ over $y > 0$, we get finally that:

$$U(t, x) \geq \mathbb{E}_t\Big[U(T, X_T^{(\pi,c)}) + \int_{u=t}^{T} U^{(c)}(u, c_u)du\Big]$$

Taking the optimal $(\pi, c)$ as in condition 3 of the proposition, which is admissible by assumption, we get equality in place of the inequalities in all of the above and the proof is complete. ■

Using Proposition 1.3.3, we can build the following example of a family of forward utility and consumption functions:

**Proposition 1.3.4** *Second parametric family of forward utility and consumption functions*

*Let:*

- *$\nu$ be a positive Borel measure such that $\int_{r>0} y^{-r}\nu(dr) < \infty$ for all $y > 0$.*
- *$\delta$ be a function from $(0, \infty)$ into itself, of polynomial growth as $r \to \infty$ and such that $\lim_{r\to 1} \frac{\delta(r)}{1-r} < \infty$.*

*Then, the two functions:*

$$V(t, y) := \int_{r>0} \frac{1}{1-r}\Big(1 - y^{1-r} e^{[r(1-r)-\delta(r)]A_t(\omega)}\Big)\nu(dr)$$

*and*

$$V^{(c)}(t, y) := \int_{r>0} \frac{-y^{1-r}}{1-r} A_t'\delta(r) e^{[r(1-r)-\delta(r)]A_t(\omega)}\nu(dr)$$

*are the convex conjugates of associated forward utility and consumption functions.*

Proof of Proposition 1.3.4:

The functions $U$ and $U^{(c)}$, conjugate convexes of $V$ and $V^{(c)}$ satisfy all the conditions of Proposition 1.3.3, and therefore are forward utility and consumption functions associated with each other. ■

## 1.4 A characterization of decreasing forward utility functions

Finally, in this last section, we are interested in answering the following question: given some strictly increasing and strictly concave function $u_0$ on $[0,\infty)$, does there exist a forward utility $U$ (without consumption) such that $U(0,x) = u_0(x)$? We partially answer this question by giving a family of functions $u_0$ which correspond exactly to (all of) the time 0 values of forward utilities of a certain family (i.e. the forward utility functions which are $C^1$ in $t$, $C^3$ in $x$ and which satisfy the Inada conditions). For more details and full proofs of the propositions of this section, the reader is referred to the preprint [1] by the author and Rogers and Tehranchi. Notice that the differentiability in $t$ of $U(t,x)$ a.s. implies that $U(t,x)$ is a decreasing function of $t$, a.s. Indeed, from the definition, $U(t,x)$ is a super-martingale for each $x$ and therefore has a decreasing drift term in its semi-martingale decomposition. The differentiability assumption implies that the local-martingale term is zero and therefore implies that a.s. the paths of $U(t,x)$ are decreasing.

We place ourselves again in the setup of the previous section ($U$ three times continuously differentiable in $x$, differentiable in $t$ and satisfying the Inada conditions). Proposition 1.3.2 still holds, but here of course with $V^{(c)} = 0$. However in this case, it is possible to characterize exactly the decreasing and convex solutions of this (simpler) random linear PDE. We get the following Theorem and its corollaries:

**Theorem 1.4.1** *Characterization of Decreasing Forward Utilities*
*Let $U$ be a forward utility, $C^1$ in $t$, $C^3$ in $x$ and satisfying the Inada conditions. Then, there exists a positive measure $\nu$ on $(0,\infty)$ such that $\int_{r>0} y^{-r}\nu(dr) < \infty\ \forall y > 0$ and a constant $C$ such that:*

$$V(t,y,\omega) = \int_{r\in(0,\infty)} \frac{1}{1-r}\Big(1 - y^{1-r}e^{r(1-r)A_t(\omega)}\Big)\nu(dr) + C$$

**Corollary 1.4.2** *Time* 0 *Utilities*
*A utility function $u_0$ is the time 0 value of a forward utility $U$ satisfying the conditions*

*of Theorem 1.4.1 if and only if:*

$$v_0(y) = \int_{r\in(0,\infty)} \frac{1-y^{1-r}}{1-r}\nu(dr) + C$$

*for some finite positive Borel measure $\nu$ on $(0,\infty)$ and a constant $C$, where $v_0$ is the convex conjugate of $u_0$.*

Notice that these time 0 utilities also correspond to the time 0 utilities of the parametric family of Proposition 1.3.4. Therefore, any function $u_0$ as above can be seen as the time 0 of a FDU with or without consumption.

**Corollary 1.4.3** *Dual Formulation*
*Let $U$ be a forward utility function satisfying the conditions of Theorem 1.4.1, then it holds that, for all $y > 0$:*

$$\mathbb{E}_t\Big[V(T, y\frac{Z_T}{Z_t})\Big] = V(t,y)$$

*where $V$ is the convex conjugate of $U$*

**Corollary 1.4.4** *Mutual Fund*
*Let $U$ be a forward utility satisfying the conditions of Theorem 1.4.1, then an optimal investment strategy is given by:*

$$\pi_s^* = (\sigma_s^{-1})^T \Theta_s^T \int_{r>0} r\Big(y\frac{Z_s}{Z_t}\Big)^{-r} e^{r(1-r)A_s}\nu(dr)$$

*and the optimal wealth is then equal to:*

$$X_s^* = \int_{r>0} \Big(y\frac{Z_s}{Z_t}\Big)^{-r} e^{r(1-r)A_s}\nu(dr)$$

*where $y$ is as earlier the conjugate of $x$ which is given by $V(t,y) = U(t,x) - xy$.*

<u>Outline of proof of Theorem 1.4.1</u>:
$V$ satisfies the PDE $V_t(t,y) + \frac{1}{2}y^2 V_{yy}(t,y)||\Theta_t||^2 = 0$. We do the following bijective change of variable $(t,y,\omega) \mapsto (\tau, z, \rho)$ from $[0,\infty)\times(0,\infty)\times\Omega$ into $[0,\infty)\times\mathbb{R}\times\Omega$:

- $\tau = A_t(\omega)$

- $z = \log(y) + A_t(\omega)$
- $\rho = \omega$

and we set $W(\tau, z, \rho) := V(t, y, \omega)$

Then it is easy to see that $W$ satisfies the Backward Heat Equation:

$$W_{zz}(\tau, z, \rho) + W_\tau(\tau, z, \rho) = 0$$

and so does $W_z$, which in addition has to be negative because of the decreasing monotonicity of $V$. Therefore, by Widder's characterization of positive solutions of the Backward Heat Equation (see [44] or [45]), it must hold that:

$$W_z(\tau, z) = \int_{r\in\mathbb{R}} -e^{rz-r^2\tau}\nu(dr) + D$$

for some positive finite Borel measure $\nu$ and some constant $D$. Notice that to be precise, from the equation above, there is no reason why $\nu$ should not depend on $\rho$. However, properties (3) and (4) of Definition 1.1.1 imply that $U(t, x)$ is a super-martingale, thus adapted. As we have assumed $\mathcal{F}_0$ to be trivial, $U(0, x)$ has to be independent of $\omega$, and so have to be $V(0, y)$ and $W(0, z)$. Therefore, $\nu$ has to be the identical across $\rho \in \Omega$. Integrating with respect to $z$ and taking into account the fact that $W$ must also be solution of the Backward Heat Equation, we get:

$$W(\tau, z) = \int_{r\in\mathbb{R}} \frac{1}{r}(1 - e^{rz-r^2\tau})\nu(dr) + C$$

for some constant $C$. Going back to $V$ and observing that the Inada conditions and convexity of $V$ can hold only if $\nu$ is null on $[1, \infty)$, we get that it must hold that

$$V(t, y) = \int_{r\in(0,\infty)} \frac{1}{1-r}\left(1 - y^{1-r}e^{r(1-r)A_t(\omega)}\right)\eta(dr) + C$$

for some positive measure $\eta$, defined on $[0, \infty)$ such that $\eta(0) = 0$.

It remains now to verify that such a $V$ is the convex conjugate of a forward utility. Let us then take $V$ as in the Theorem. It is obvious that $V$ is strictly decreasing and convex. Remains to verify properties (3) and (4) of Definition 1.1.1: the check goes as in the previous cases with consumption. (3) is checked by appealing to the definition of convex/concave conjugates and Corollary 1.4.3, which can be proved by direct computation given that everything is explicit. Finally, (4) is checked by using the optimal wealth and admissible trading strategy given in Corollary 1.4.4. ■

# Chapter 2

# Hedging with Variance Swaps in Infinite Dimensions

**Abstract:** It has been shown recently by Buehler that variance swaps can be modelled, jointly with the stock on which they are written, in a manner strikingly similar to the HJM interest rates framework. We apply this technique to model the (forward) variance swaps curve by a stochastic partial differential equation (SPDE) in a Hilbert space, and apply the tools of Malliavin calculus to give an explicit representation of the hedging portfolio for a class of exotic contingent claims written on the stock and variance instruments (variance swaps or forward variance swaps). We also show that under suitable conditions on the SPDE and SDE satisfied respectively by the forward variance swaps curve and by the stock price, the (self-financing and admissible) hedging portfolio is unique and satisfies a maturity-specific property similar to the one proved by Carmona and Tehranchi for interest rates contingent claims in the context of infinite dimensional models.

## 2.1 Introduction

### 2.1.1 Motivation

One of the goals financial mathematicians have been trying to achieve in recent years has been to build and study stochastic arbitrage-free market models, where they can jointly model, in a consistent manner, the prices of an underlying (say a stock for instance) and liquid derivatives written on this underlying. Much research has focused on call (or put) options market models (see for instance Cont, Fonseca and Durrleman [12] who studied empirical and statistical features of the call surface to model the implied volatility, Davis [14] who studied complete market models of stochastic volatility, Wissel [47] and Schweizer and Wissel [41] and [42] who studied arbitrage-free market models for call surfaces, and Carmona and Nadtochiy (see [5] and [6]) who studied local volatility market models.) However, call surface market models are very difficult to study because of consistency conditions which are imposed on the different assets we are trying to model (the stock should be recovered from the call with strike 0, there are boundary conditions at maturity of the calls, as maturity tends to infinity, as the strike tends to infinity, etc). Buehler [4], on the other hand, focused his attention on another type of market models: variance swaps models. It turns out that the modelling is not only much simpler than in the call surface case (which is not too surprising because there is only the maturity dimension, versus maturity and strike dimensions for call surface models), but it is also very close to the famous HJM framework for forward rates modelling. In this chapter, we apply Buehler's modelling idea to model a stock and (forward) variance swaps written on this stock, and look at the problem of hedging discretely monitored exotic European options with the stock and (forward) variance swaps. Very much as Carmona and Tehranchi [7] did in the case of interest rates contingent claims, we focus on genuine infinite dimensional models, which have two interesting features which are absent in classical finite dimensional models (provided some assumptions on the model's parameters): the hedging portfolio for a given contingent claim is unique, and it makes only use of variance swaps maturing on, or before the maturity of the option. Let us assume for instance that we are trying to hedge a look-back option, paying the difference between the maximum and the minimum of the stock, observed daily at the close of market, over a period of one year. Such an option seems to be a measure in some sense of the volatility of the stock over the period of one year from now, so it certainly seems like a reasonable thing to do to use variance swaps to hedge it (we do not worry about transaction costs here). It would however, seem unnatural to use as hedging instruments variance swaps maturing in 2 years, 3 years, etc, because there should be (at least intuitively) more correlation between the look-back option we are trying to hedge and variance swaps maturing within a year than between the look-back option and variance swaps whose maturity exceeds that of the look-back option. Finite dimensional models, because they are complete (provided we can trade as many variance swaps as there are independent Brownian motions underlying the model), do not preclude such an unnatural choice of hedging instruments. Infinite dimensional models can be made to preclude it.

### 2.1.2 Organization of this chapter

We now shortly describe how this chapter is organized: in the second section, we start by giving the definition of variance swaps and explain why they have become actively traded. We also give a short overview of Buehler's quite recent discovery that variance swaps can be modelled in an HJM framework, similarly to what is done in interest rates when one models the forward curve. In the third section we will have a short look at an example of finite dimensional stochastic volatility model where the market can be completed by trading in the stock and in a finite number of arbitrary chosen variance swaps. This shows that although of practical importance, finite dimensional models have the unsatisfying theoretical property that the hedging instruments can be chosen independently of the claim to hedge. This leads us, in the fourth section, to introduce variance swaps models in infinite dimensional spaces. From that point on we follow very closely the work of Carmona and Tehranchi on interest rates modelling in infinite dimensions: we start by listing the assumptions we need to make on the function spaces in which the (forward) variance curve will live and on the stochastic equation that governs its evolution. In section 5, we discuss the meaning we should give to portfolios in this infinite dimensional setup and explain how the classical self-financing condition can be extended in order to be consistent with common sense. We show that under some conditions on the SPDE satisfied by the forward variance swaps curve, portfolios leading to a given wealth at some fixed time in the future are unique.

In section 6, we use Malliavin calculus and in particular the Clark-Ocone formula to derive an expression for the hedging portfolios of contingent claims written on the stock and the variance instruments. We show that under additional assumptions on the volatility operator of the forward variance swaps curve and on the correlation between the stock and the (forward) variance swaps, the unique hedging portfolio satisfies a maturity-specific property which was absent in finite dimensional models. Finally in the seventh and last section, we give a concrete example of model that satisfies all the assumptions we listed in sections 4, 5 and 6. We also give some examples of classical payoffs involving the stock and the variance instruments and which fit in the framework we have presented.

An important appendix gives short introductions to the mathematical tools we have used in the main text, namely stochastic analysis in infinite dimensions, Malliavin calculus for Hilbert space valued random variables and existence, uniqueness and Malliavin differentiability of mild solutions to SPDEs in Hilbert spaces. The propositions and theorems discussed there are important, but rather technical, so we chose to relegate them to the appendix in order to lighten the main text. Although these results can be found in the literature, we thought it convenient to have them detailed here, so that the chapter is reasonably self-contained.

## 2.2 Variance swaps and consistent HJM models

### 2.2.1 Variance swaps and forward variance swaps

All along we will consider a risky asset $P$, whose price process $(P_t)_{t\geq 0}$ is assumed to be a continuous positive local-martingale of the form:

$$P_t = P_0 \mathcal{E}\Big( \int_{u=0}^{t} \sqrt{\varsigma_u} dW_u \Big) \tag{2.2.1}$$

where $(W_t)_{t\geq 0}$ is a scalar Brownian motion defined on a complete filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ satisfying the usual conditions, where the square root of the variance process $(\varsigma_t)_{t\geq 0}$ is $W$-stochastically integrable, and where $\mathcal{E}$ denotes the Doleans exponential local-martingale. Notice that $\{\mathcal{F}_t\}_{t\geq 0}$ may be larger than the augmented filtration generated by $W$. We make the assumption that interest rates are constant and equal to 0.

For any partition $\mathcal{T} := \{t_j\}_{j=0}^N$ of the time interval $[0, T]$, i.e. such that $0 = t_0 \leq ... \leq t_N = T$ and for each $n \leq N$, we define the variance swap $V_{\mathcal{T}}^{(t_n)}, n \leq N$ to be the financial contract paying at time $t_n$ the quantity:

$$V_{\mathcal{T}}(t_n, t_n) := \sum_{j=1}^{n} \Big( \log \frac{P_{t_j}}{P_{t_{j-1}}} \Big)^2$$

and will write $V_{\mathcal{T}}(t, t_n)$ for the value of this contract at any time $t \leq t_n$. The average payoff $\frac{1}{t_n} V_{\mathcal{T}}(t_n, t_n)$ is an estimator of $\frac{1}{t_n}\mathbb{E}\int_{u=0}^{t_n} \varsigma_u du$ and is generally referred to as the "realized variance". Variance swaps, such as we have just defined them, are quite actively traded on some major stocks and indices. In reality however, they pay the difference between $V_{\mathcal{T}}(t_n, t_n)$ and some pre-agreed constant value, which may for instance be chosen in such a way that the original value of the variance swap is null. From a mathematical standpoint however, the addition of this constant does not make any difference so that we will ignore it to make things simpler. $\mathcal{T}$ is generally running through the trading days (Monday to Friday) of a given period ( e.g. three months, a year, etc). Other products which have become popular are the forward variance swaps, which pay the realized variance between two dates in the future. We define then for any $n, m \in \{0, ..., N\}$, with $n < m$, the forward variance swap $v_{\mathcal{T}}^{(t_n,t_m)}$ as the contract paying at time $t_m$ the quantity:

$$v_{\mathcal{T}}(t_m, t_n, t_m) := \frac{1}{t_m - t_n} \sum_{j=n}^{m-1} \Big( \log \frac{P_{t_{j+1}}}{P_{t_j}} \Big)^2.$$

Similarly to variance swaps, we denote its value at time $t \leq t_m$ by $v_{\mathcal{T}}(t, t_n, t_m)$. Of course it is easy to observe that having the prices of the family of variance swaps $\left\{ V_{\mathcal{T}}^{(t_n)} \right\}_{n=1}^{N}$, or those of the family of forward variance swaps $\left\{ v_{\mathcal{T}}^{(t_n,t_m)} \right\}_{n<m \in \{0,...,N\}}$ are equivalent, because of the relations:

$$V_{\mathcal{T}}^{(t_n)} = \sum_{j=0}^{n-1} (t_{j+1} - t_j) v_{\mathcal{T}}^{(t_j,t_{j+1})} \tag{2.2.2}$$

$$v_{\mathcal{T}}^{(t_n,t_{n+1})} = \frac{V_{\mathcal{T}}^{(t_{n+1})} - V_{\mathcal{T}}^{(t_n)}}{t_{n+1} - t_n} \tag{2.2.3}$$

which are just respectively a discrete integral and a discrete derivative.

Let us now consider a family $\mathcal{T}^{(n)} := \{t_j\}_{j=0}^{N^{(n)}}$ of increasing partitions of the time interval $[0, T]$, and let us denote by $\|\mathcal{T}\|$ the maximum step of partition $\mathcal{T} := \{t_j\}_{j=0}^{N}$, that is:

$$\|\mathcal{T}\| := max_{j \in \{1,...,N\}} |t_j - t_{j-1}|$$

It is well known that if

$$\lim_{n\to\infty} \|\mathcal{T}^{(n)}\| = 0,$$

then:

$$\lim_{n\to\infty} V_{\mathcal{T}^{(n)}}(\tau, \tau) =< \log(P) >_{\tau},$$

this last limit being understood in the sense of ucp convergence (i.e. uniformly on compacts, in probability). See Protter [38] p. 66 for instance for this result. $\tau$ in the above can be chosen to be any point in $\mathcal{T}^{(\infty)} := \bigcup_{n=0}^{\infty} \mathcal{T}^{(n)}$, which is dense in $[0, T]$ by the assumptions that $\|\mathcal{T}^{(n)}\| \mapsto 0$ and that the partitions are increasing.

Because of these remarks, we will consider as an approximation of the real world that for any two instants $t \leq T$, we can trade at time $t$ a product paying at time $T$ the quadratic variation of the log of the stock price: $< \log(P) >_T$. We will continue to call this product the variance swap maturing at time $T$ and denote its price at time $t$ by $V(t, T)$. From a practical perspective, we can remark that the approximation is only sensible for large values of $T - t$, but we shall ignore this issue.

Obviously for all $t \geq 0$, the curve $T \mapsto V(t, T)$, defined on $[t, \infty)$, is increasing, and therefore is differentiable Lebesgue-almost everywhere. If $V(t, .)$ is differentiable at $T_0$, and to keep consistency with our earlier definition of real forward variance swaps, we will use the notation: $v(t, T_0) := \frac{\partial V(t,T)}{\partial T}\Big|_{T_0}$ and call this quantity the ($t$-time value of the) forward variance swap maturing at time $T_0$. In particular, and by analogy with interest rates theory, we will call $v(t, t)$ the short variance at time $t$.

Neuberger [35] showed that in the context of continuous models of the form (2.2.1), which are fairly general, one can replicate the quadratic variation of the log of the stock by trading continuously in the stock, in a model-independent manner, and by taking a long position in an option paying the logarithm of the stock. Indeed, an application of Ito's formula gives that:

$$\int_{u=0}^{T} \varsigma_u du = 2 \int_{u=0}^{T} \frac{dP_u}{P_u} + 2[\log(P_0) - \log(P_T)]$$

This formula, combined with the well known fact that European options depending only on the terminal value $P_T$ (namely $\log(P_T)$) can be replicated by a model-independent static position in call and put options, is probably one of the reasons that explain why variance swaps have become popular. Indeed in the case of the log contract, we can show using integration by parts that we have the following identity, holding for any $K > 0$:

$$\log(P_T) = \log(K) + \frac{P_T - K}{K} - \int_{k=0}^{K} \frac{(u - P_T)_+}{u^2} du - \int_{k=K}^{\infty} \frac{(P_T - u)_+}{u^2} du$$

which means that one can replicate statically the log contract by holding $\log(K)$ of the bond, $1/K$ of the forward contract struck at $K$, selling $1/u^2$ of the put option struck at $u$

for each $u \leq K$ and selling $1/u^2$ of the call option struck at $u$ for each $u \geq K$.

As pointed out by Neuberger [35] however, traders were already betting on variance even before the appearance of variance swaps, generally using other options like straddles, or via delta-hedging strategies, which were providing a less direct and less perfect exposure to realized variance (see also the survey article on volatility trading by Carr and Madan [9]).

### 2.2.2 Buehler-HJM market models

Buehler, in his PhD thesis and subsequent papers such as [4], has shown that the HJM methodology could be applied to variance swaps modelling. Suppose that we start with a stock $(P_t)_{t\geq 0}$ modelled as a strictly positive and continuous local-martingale of the form:

$$P_T = P_0 \exp\Big( \int_{u=0}^{T} \sqrt{\varsigma_u} dW_u - \frac{1}{2} \int_{u=0}^{T} \varsigma_u du \Big) \tag{2.2.4}$$

for some Brownian motion $W$ and some variance process $\varsigma$ (which may not be adapted to the filtration generated by $W$, for instance if we wanted to have an incomplete market model). As explained in the previous section, we can then define the (approximated) variance swaps at time $t$ maturing at time $T$ by:

$$V(t, T) := \mathbb{E}_t\Big[ \int_{u=0}^{T} \varsigma_u du \Big]$$

and under some mild regularity conditions, it may make sense to define the forward variance swaps as:

$$v(t, T) := \partial_T \mathbb{E}_t\Big[ \int_{u=0}^{T} \varsigma_u du \Big]$$

which would also be strictly positive local-martingales (i.e. for each $T \geq 0$, $\big(v(t, T)\big)_{t\in[0,T]}$ is a local-martingale). We can remark then that $v(t, t) = \varsigma_t$. The discovery of Buehler is that we could go backward instead: we could start by defining the forward variance swaps as a family (indexed by $T$) of (say) strictly positive martingales $\big(v(t, T)\big)_{t\in[0,T]}$ via a family of SDEs. Then we could define the variance swaps by integration, and finally define the stock via equation (2.2.4), where $W$ is an arbitrary Brownian motion (whose dependence on the randomness driving the forward variance swaps may be chosen as we wish). This construction guarantees that all assets' prices are (local) martingales, and we have of course that for all $t \leq T < \infty$, $V(t,T) = < \log P >_T$. This means that we are building a market model, where we describe the joint evolution of a stock and a family of derivatives written on this stock, in a consistent and arbitrage-free manner.

The advantage of this methodology is that it allows one to find prices of complex exotic options in terms of a stock and some actively traded derivatives on the stock (i.e. variance swaps). This may therefore be of interest to traders because it allows them to work out how to hedge exotic options not only by trading in the stock, but also by trading in variance swaps (although some may argue that transaction costs on variance swaps may rule out such practice. We do not address these practical issues in this chapter however). From the point of view of an economist finally, it is also satisfying to have models which are incomplete, with potentially an arbitrary number of random factors, but where the additional factors which render the market complete can be related to quantities directly observable in the market (i.e. the variance swaps prices), rather than having, like in more conventional stochastic volatility models, factors which are non observable or which at the very best can only be estimated.

### 2.2.3 Musiela's time to maturity notation

We finish this introductory section by a point of notation: exactly as in the interest rates world, it may be convenient (and we will actually do it in the sequel) to work in terms of "time to maturity" rather than in terms of "time of maturity". We will therefore use Musiela's notation:

$$V_t(x) := V_t(T-t) := V(t,T)$$

$$v_t(x) := v_t(T-t) := v(t,T)$$

where $x$ denotes therefore the time to maturity. The advantages (time-independent curve domain) and disadvantages (have generally to deal with SPDEs instead of SDEs) of both conventions are well discussed in the literature, and the interested reader may for instance consult [32].

## 2.3 Shortcomings of finite dimensional models

This section, on continuous finite dimensional variance swaps models, serves as a justification for introducing infinite dimensional models. Instead of looking at generic finite dimensional models however, we give a particular example in which many computations can be done in closed form. The reader can then generalize to other finite dimensional models which for obvious reasons (there are more traded assets than there are underlying (scalar) Brownian motions) cannot guarantee uniqueness of hedging portfolios for contingent claims and therefore cannot exhibit any maturity-specific risk feature.

### 2.3.1 Concrete example of a complete finite dimensional model

Buehler [4] (following Filipovic [18] very closely) studies finite dimensional realizations of variance swaps HJM models. Precisely, he is interested in knowing if there exist models of the whole forward variance swaps curve $x \mapsto v_t(x)$ which can be written as $v_t(x) = G(Z_t; x)$ for some smooth function $G$ and some finite dimensional diffusion $Z_t$, which may represent factors (observable or not, tradable or not). He found that a necessary condition is that $G$ satisfies a PDE (unsurprisingly close to the heat equation) whose coefficients are linked to those of the SDE satisfied by $Z_t$. In addition, he gives numerous examples of such models, which fit in the so-called polynomial-exponential family. The famous stochastic volatility model by Heston [22] is probably the simplest and most well known example.

We show in this section a model where, given any set of $N$ variance swaps of different maturities, we can replicate any European (possibly path dependent) option by trading in the stock and these $N$ variance swaps. Moreover, for a very large class of discretely monitored options, the holding in each hedging instrument can be computed explicitly. From a practical perspective this may be interesting because the parameters of such models could be optimized to (best) fit an initial set of variance swaps, call and put options, and then used to hedge exotic options with the available variance swaps. Very recently, Heston et al have studied a two-dimensional version of this same model [10] and have shown that it improves on the 1-dimensional case with regard to capturing features of the volatility smile (i.e. it is possible to fit better the smile's shape and deformation over time).

**The model**

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a probability space supporting a $N+1$-dimensional Brownian motion $W = (W_t)_{t\geq 0}$, and where $(\mathcal{F}_t)_{t\geq 0}$ is the augmented filtration generated by $W$ and $\mathcal{F} = \mathcal{F}_\infty$. We look at $W$ as living in the Euclidean space $\mathbb{R}^{N+1}$, and we will write its components as $(W_t^{(0)}, ..., W_t^{(N)})$

We define the stock $(P_t)_{t\geq 0}$ and the $N$ short volatility components $\{(B_t^{(n)})_{t\geq 0}\}_{n\in\{1,...,N\}}$ to be the strong solutions (starting respectively at the positive values $P_0$,$B_0^{(1)}$,...,$B_0^{(N)}$) of the following SDEs:

$$dP_t = P_t \sum_{n=1}^{N} \sqrt{B_t^{(n)}} d\tilde{W}_t^{(n)} \tag{2.3.1}$$

*and*

$$\forall n \in \{1, ..., N\}, \quad dB_t^{(n)} = k^{(n)}(\Theta^{(n)} - B_t^{(n)})dt + \sigma^{(n)} \sqrt{B_t^{(n)} \vee 0}\, dW_t^{(n)}$$

where $\tilde{W}_t^{(1)} := \rho W_t^{(0)} + \sqrt{1-\rho^2} W_t^{(1)}$, for some constant $\rho \in (-1, 1)$, where $\{k^{(n)}\}_{n=\{1,...N\}}$, $\{\Theta^{(n)}\}_{n=\{1,...N\}}$, and $\{\sigma^{(n)}\}_{n=\{1,...N\}}$, are positive constants, and $\tilde{W}_t^{(n)} = W_t^n, n \geq 2$ (the constant $\rho$ gives the correlation between the variance and the stock). We will in addition make the following assumptions on the coefficients of the SDEs:

**Assumption 2.3.1** *The values of the coefficients $\{k^{(n)}\}_{n\in\{1,...,N\}}$ are all distinct from each other.*

**Assumption 2.3.2** $\forall n \in \{1, ..., N\}$, $(\sigma^{(n)})^2 < 2k^{(n)}\Theta^{(n)}$

Assumption 2.3.2 is the well known "Feller condition" that ensures that the processes $(B_t^{(n)})_{t\geq 0}, n \in \{1, ..., N\}$ remain strictly positive at all times, almost surely (see for instance [17]). That the SDEs for $\{(B_t^{(n)})_{t\geq 0}\}_{n\in\{1,...,N\}}$ have strong unique solutions is far from being obvious, especially because of the square root term, which prevents us from using classical theorems on SDEs with Lipschitz coefficients. However, a proof of that fact can be found in Ikeda and Watanabe [23], Theorem IV.2.3 page 173 and Theorem IV.3.2 page 182 as well as Example IV.8.2 page 235. A proof of the Feller condition can be found in Example 8.2 page 235 and is based on a technical result on explosion times of paths of diffusions which can be found in Theorem VI.3.1 page 447.

#### Completeness

We make Assumption 2.3.1 for the following reason:

**Theorem 2.3.3** *Let $\{T_m\}_{m\in\{1,\dots,N\}}$ be $N$ given times, distinct from one another and let us denote by $T$ their minimum. Then the market composed of the stock $P_t$ and the $N$ variance swaps $\{V_t(T_m - t)\}_{m\in\{1,\dots,N\}}$ maturing at times $\{T_m\}_{m\in\{1,\dots,N\}}$ is complete on $[0,T)$ if and only if Assumption 2.3.1 holds.*

Notice that by complete on $[0,T)$ we mean that any contingent claim $\xi_\tau$ which is $\mathcal{F}_\tau$-measurable for some $\tau \in [0,T)$ can be replicated by trading in the above mentioned assets.

Proof of Theorem 2.3.3:

Let us first remark that the SDE for the stock $P_t$ can be rewritten as:

$$dP_t = P_t\sqrt{v_t^0}dZ_t$$

where $v_t^0 := \sqrt{\sum_{n=1}^N B_t^{(n)}}$ is the short variance, and where $Z_t$, defined by

$$dZ_t := \frac{\sum_{n=1}^N \sqrt{B_t^{(n)}}d\tilde{W}_t^{(n)}}{\sqrt{\sum_{n=1}^N B_t^{(n)}}}$$

is easily seen to be a 1-dimensional Brownian motion, by Levy's characterization of Brownian motion.

This allows us to compute quite easily the variance swaps prices (a rigorous computation can be found in the paper by Potter [37] who studies a class of 2-dimensional complete stochastic volatility models where he assumes both the stock and an additional option (closely related to a variance swap) to be tradable): the result, in our notation, is that for $t, x \geq 0$, the time $t$ price $V_t(x)$ of the variance swap maturing at time $t + x$ is given by:

$$V_t(x) = \int_{u=0}^{t} v_u^0 du \; + \; \sum_{n=1}^{N} \Big[ \Theta^{(n)} x + (B_t^{(n)} - \Theta^{(n)}) \frac{1 - \exp(-k^{(n)} x)}{k^{(n)}} \Big] \tag{2.3.2}$$

**Remark 2.3.4** *It can be noticed that the above depends on* $V_t(0) = \int_{u=0}^{t} v_u^0 du$*, which makes the dependence of* $V_t(x)$ *on the factors of our SDEs non Markovian. For concreteness, one would need to keep track of the accumulated variance from* 0 *to* $t$ *in order to compute the value of* $V_t(x)$ *by the above formula. This is one of the reasons why Potter [37] prefers to use another option to complete the market and derive hedging portfolios, rather than using a variance swap. On the other hand, the option he uses is an approximation of a variance-related object that requires the calls of all strikes to be traded which is also a bit unrealistic, and this has to be added to the fact that the variance-related objects that we are considering are already themselves approximations of real-world payoffs (the real-world variance swaps are computed using discretely monitored values of the stock).*

**Remark 2.3.5** *Notice that equation (2.3.2) can of course be written as*

$$V_t(x) = G(V_t(0), B_t; x)$$

*for some suitable smooth function* $G$*. The multi-dimensional Heston model belongs therefore to the finite dimensional HJM realizations family.*

From (2.3.2) and from the SDE (2.3.1) satisfied by the stock $P$, we get that the vector of assets $(P_t, V_t(T_1 - t), ..., V_t(T_N - 1))$ satisfies the SDE:

$$\begin{pmatrix} dP_t \\ dV_t(T_1 - t) \\ . \\ . \\ . \\ dV_t(T_N - t) \end{pmatrix} =$$

$$\begin{pmatrix} P_t\rho\sqrt{B_t^{(1)}} & P_t\sqrt{1-\rho^2}\sqrt{B_t^{(1)}} & P_t\sqrt{B_t^{(2)}} & . & P_t\sqrt{B_t^{(N)}} \\ 0 & \frac{1-\exp(-k^{(1)}(T_1-t))}{k^{(1)}}\sigma^{(1)}\sqrt{B_t^{(1)}} & . & . & \frac{1-\exp(-k^{(N)}(T_1-t))}{k^{(N)}}\sigma^{(N)}\sqrt{B_t^{(N)}} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & \frac{1-\exp(-k^{(1)}(T_N-t))}{k^{(1)}}\sigma^{(1)}\sqrt{B_t^{(1)}} & . & . & \frac{1-\exp(-k^{(N)}(T_N-t))}{k^{(N)}}\sigma^{(N)}\sqrt{B_t^{(N)}} \end{pmatrix}\begin{pmatrix} dW_t^{(0)} \\ dW_t^{(1)} \\ . \\ . \\ . \\ dW_t^{(N)} \end{pmatrix}$$

$$=: A_t dW_t$$

We can now observe that the matrix $A_t$ is invertible for all $t < T := \inf_{n\leq N}\{T_n\}$ and for all $\omega \in \Omega$ if and only if no two coefficients $k^{(n)}, n \in \{1, ..., N\}$ are equal. Indeed, doing a Laplace expansion of $A_t$ along the first column, we see that $A_t$ is invertible if and only if:

$$\begin{pmatrix} \frac{1-\exp(-k^{(1)}x_1)}{k^{(1)}}\sigma^{(1)}\sqrt{B_t^{(1)}} & . & . & . & \frac{1-\exp(-k^{(N)}x_1)}{k^{(N)}}\sigma^{(N)}\sqrt{B_t^{(N)}} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{1-\exp(-k^{(1)}x_N)}{k^{(1)}}\sigma^{(1)}\sqrt{B_t^{(1)}} & . & . & . & \frac{1-\exp(-k^{(N)}x_N)}{k^{(N)}}\sigma^{(N)}\sqrt{B_t^{(N)}} \end{pmatrix}$$

is, where we have used $x_n$ as a short-hand notation for $T_n - t$. Simplifying factors which are common across lines or columns, we see that $A_t$ is invertible if and only if:

$$\begin{pmatrix} 1-\exp\left(-k^{(1)}(x_1)\right) & . & . & . & 1-\exp\left(-k^{(N)}(x_1)\right) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1-\exp\left(-k^{(1)}(x_N)\right) & . & . & . & 1-\exp\left(-k^{(N)}(x_N)\right) \end{pmatrix}$$

is. And finally, this means that $A_t$ is invertible if and only if the following generalized Vandermonde matrix is itself invertible, a classical fact (see Gantmacher [20] p. 87).

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \exp\left(-k^{(1)}x_1\right) & . & . & . & \exp\left(-k^{(N)}x_1\right) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \exp\left(-k^{(1)}x_N\right) & . & . & . & \exp\left(-k^{(N)}x_N\right) \end{pmatrix}$$

■

One can easily observe that the matrix $A_t$ becomes singular if $t = T$ as one of the column (the one corresponding to $n$ such that $T_n = T$) will converge to a multiple of the first column. On the other hand, if we let the maturities go to $\infty$, then the matrix becomes also singular because the terms in $\frac{1-\exp\left(-k^{(n)}(T_n-t)\right)}{k^{(n)}}$ would converge to $\frac{1}{k^{(n)}}$. What this means is that this model allows one to hedge any contingent claim $\xi_\tau$ with the stock plus any $N$ variance swaps of maturities strictly greater than $\tau$. However, if the maturities are too far away from $\tau$ then the hedging portfolio would most likely become really risky and unnatural, i.e. with very large positive and negative holdings in each instrument due to the matrix becoming singular.

To conclude on this example of finite dimensional model, we see that it may be useful from a practical perspective because it gives us a complete market model where one can hedge fairly general European options with the stock and some variance swaps. As we have seen, the hedging portfolio may become very unnatural if we choose as hedging instruments variance swaps whose maturities are much larger than the maturity of the option to be hedged, which is in some sense a weak maturity-specific risk property. However, we would like to have models with a stronger maturity-specific risk property, i.e. in which it is not possible to replicate an $\mathcal{F}_T$-measurable contingent claims by trading in variance swaps maturing later than $T$. Only by moving to infinite dimensional models do we have some hope of finding such models.

Indeed, quite informally, if we choose models of the form: $dV(.,t) = \sigma dW_t$ where now $W$ is an infinite dimensional Brownian motion living in a space $\mathbb{G}$ (say), the curves $T \mapsto V(T,t)$ live in an infinite dimensional space of curves $\mathbb{F}$, and $\sigma$ is an appropriate operator taking $\mathbb{G}$ into $\mathbb{F}$, we can see that we now have the possibility of choosing $\sigma$ with adjoint operator of trivial kernel. Indeed, still informally, what we would like to be able to write is that, for a trading strategy $(\phi_t)_{t\in[0,T]}$, if $\phi_t\sigma_t = 0$, then $\phi_t = 0$ (which is exactly saying that the kernel of $\sigma^*$ is trivial!). That would indeed guarantee that two trading strategies $(\psi_t)_{t\in[0,T]}$ and $(\eta_t)_{t\in[0,T]}$ giving the same trading gains on the interval $[0,T]$, i.e. satisfying $\int_{t=0}^T \psi_t\sigma_t dW_t = \int_{t=0}^T \eta_t\sigma_t dW_t$, would satisfy $\psi = \eta$.

We will see however that the task is not so easy. Firstly, we have to make sense mathematically of trading strategies $(\phi_t)_{t\in[0,T]}$ making use of infinitely many instruments and check that the quantity $\int_{t=0}^T \phi_t\sigma_t dW_t$ can indeed be interpreted as the trading gain of following strategy $\phi$. Secondly, the obvious choice of taking $\sigma$ bijective from $\mathbb{G}$ to $\mathbb{F}$ is not available to us, because infinite dimensional stochastic integration theory tells us that $\sigma$ should be an Hilbert-Schmidt operator from $\mathbb{G}$ to $\mathbb{F}$ and hence cannot be bijective. Fortunately, we can still make $\sigma$ dense-range, which is good enough for us, as this means the adjoint $\sigma^*$ will have trivial kernel.

## 2.4 The infinite dimensional setup

### 2.4.1 Probability space and sources of randomness

Let $\mathbb{G}$ be a separable infinite dimensional Hilbert space, and $(W_t)_{t\geq 0} := (\{W_t^{(n)}\}_{n=1}^{\infty})_{t\geq 0}$ a $\mathbb{G}$-cylindrical Brownian motion, defined on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\geq 0}$ is assumed to be the completion of the filtration generated by $W$, and where $\mathcal{F} := \mathcal{F}_\infty$.

For any $\mathbb{G}^*$-valued adapted process $(\lambda_t)_{t\geq 0}$ of norm 1, we will denote by $W^{(\lambda)}$ the scalar Brownian motion defined by $W_t^{(\lambda)} := \int_{u=0}^t \lambda_u dW_u$.

Notice that the nature of the Hilbert space $\mathbb{G}$ is completely irrelevant to us. The only thing that will play a role in the following is that it is separable and genuinely infinite dimensional. For simplicity and concreteness, the reader may consider without loss that $\mathbb{G} = l_2$, the set of square integrable real-valued sequences. We will denote by $\{g_n\}_{n=1}^{\infty}$ an orthonormal basis of $\mathbb{G}$.

Finally some notation that will be used quite often: for any metric space $\mathbb{S}$, we denote by $\mathcal{B}(\mathbb{S})$ its Borel $\sigma-$algebra. For any $\mathbb{F}$ and $\mathbb{G}$ Hilbert spaces, we denote by $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$ the Hilbert space of Hilbert-Schmidt operators from $\mathbb{G}$ to $\mathbb{F}$, and by $\mathbb{F} \otimes \mathbb{G}$ the tensor product Hilbert space of $\mathbb{F}$ and $\mathbb{G}$. Notice that $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$ and $\mathbb{F} \otimes \mathbb{G}$ may be identified with each other.

### 2.4.2 State spaces

We will model the variance swaps curve $V_t(.)$ and the forward variance swaps curve $v_t(.)$ as stochastic processes valued respectively in $\tilde{\mathbb{F}}$ and $\mathbb{F}$, two separable Hilbert spaces of continuous functions and satisfying some assumptions that we will introduce shortly. Notice that it could seem at first sight more natural to have our curve of assets valued in $C := C([0, \infty))$, the space of continuous functions on the interval $[0, \infty)$ (see for

instance the paper by De-Donno and Pratelli [15] and the older paper by Bjork et al [3]). However, $C$ (endowed with the supremum norm) is (only) a Banach space, so we will find it more convenient and easier to consider instead Hilbert space valued assets. We really only observe in the market a discrete (and indeed finite) subset of the variance swaps curves (or of bond prices/yield curves in interest rate theory for instance), so that the additional smoothness assumptions imposed on the curves to make the state spaces Hilbert spaces is not really making the model less realistic.
Notice that we may restrict ourselves to modelling $V_t$ without having to think about $v_t$ but it may not be easy to construct directly a solution $V_t$ that has the required properties (especially $V_t(.)$ must be a positive and increasing function for all $t$). We therefore follow instead the approach of starting, and indeed working almost exclusively with a model for $v_t(.)$ (which only need be positive), and simply keeping in mind that the corresponding model for $V_t$ can be easily obtained via integration.
All along, we will impose the following on the state spaces $\mathbb{F}$ and $\tilde{\mathbb{F}}$:

**Assumption 2.4.1** *$\mathbb{F}$ is a separable Hilbert space, and a subset of $C$, the sets of continuous functions on $\mathbb{R}_+$.*

**Assumption 2.4.2** *The family of left shifts $\{S_t\}_{t\geq 0}$, defined by $S_t f(.) := f(t+.)$, forms a strongly continuous semigroup on $\mathbb{F}$. We will denote the infinitesimal generator of $(S)_{t\geq 0}$ by $A$. Notice that whenever $f \in \mathbb{F}$ is differentiable, then $f' = Af$.*

**Assumption 2.4.3** *The evaluation functionals $\delta_x : f \mapsto f(x)$ are continuous linear functionals on $\mathbb{F}$. The set $\{\|\delta_x\|_{\mathbb{F}}\}_{x\geq 0}$ is uniformly bounded by some constant $K$.*

We define then $\tilde{\mathbb{F}}$, the space in which $V_t$ will be valued, as: $\tilde{\mathbb{F}} := \{f \in C^1([0,\infty)), f' \in \mathbb{F}\}$, and we endow $\tilde{\mathbb{F}}$ with the norm: $\|f\|^2_{\tilde{\mathbb{F}}} := f(0)^2 + \|f'\|^2_{\mathbb{F}}$. It can be seen with this definition that the following properties hold (the reader can look at the concrete example of state space $\mathbb{F}$ we give later, for which we prove these facts):

**Property 2.4.4** *$\tilde{\mathbb{F}}$ so defined is also a Hilbert space, subset of $C^1$, the sets of continuously differentiable functions on $\mathbb{R}_+$, and the left shift operator family $\{S_t\}_{t\geq 0}$ also forms a strongly continuous semigroup on $\tilde{\mathbb{F}}$.*

**Property 2.4.5** *The integration functional $J$ defined by:*

$$\begin{aligned} J : \mathbb{F} &\rightarrow C^1 \\ g &\mapsto \left(x \mapsto \int_{u=0}^{x} g(u)du\right) \end{aligned}$$

*is a bijective bounded linear operator from $\mathbb{F}$ to $\tilde{\mathbb{F}}_0$, the closed subspace of $\tilde{\mathbb{F}}$ constituted of functions started at* 0*.*

**Property 2.4.6** *The differentiation operator $D$ defined by:*

$$\begin{aligned} D : \tilde{\mathbb{F}} &\rightarrow C \\ g &\mapsto g' \end{aligned}$$

*is a bounded linear operator from $\tilde{\mathbb{F}}$ into $\mathbb{F}$.*

We will denote by $\mathbb{F}_+$ the Borel subset of $\mathbb{F}$ of strictly positive functions, and by $\tilde{\mathbb{F}}_{\nearrow}$ the Borel subset of $\tilde{\mathbb{F}}$ of strictly increasing functions. That these two subsets are Borel subsets come from the separability of the real numbers, the continuity of the evaluation functionals $\{\delta_x\}_{x\geq 0}$ on $\mathbb{F}$ (Assumption 2.4.3) and the continuity of the differentiation operator $D$ from $\tilde{\mathbb{F}}$ to $\mathbb{F}$ (Property 2.4.6).

As a final remark on these state spaces, it is worth noticing that because $\mathbb{F}$ is a subset of $C$ and is likely to contain $\mathcal{S}$, the space of functions of rapid decrease (see Reed and Simon [39]), we can look at $\mathbb{F}^*$ as a space of distribution (i.e. a subset of the tempered distributions) which contains the finite signed measures (i.e. which contains $C^*$).

### 2.4.3 The model

**Dynamics of the forward variance swaps curve process $(v_t)_{t\geq 0}$**

Let $v_0$ be an element of $\mathbb{F}_+$ and $\sigma$ a measurable map from $\left(\mathbb{R}_+ \times \Omega \times \mathbb{F}_+, \mathcal{P} \times \mathcal{B}(\mathbb{F}_+)\right)$ into $\left(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}), \mathcal{B}(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}))\right)$, where $\mathcal{P}$ denotes the usual predictable sigma algebra of $\mathbb{R}_+ \times \Omega$. We assume that there exists a constant $K$ such that:

**Assumption 2.4.7** *$\sigma$ is globally Lipschitz of coefficient K: for all $t \geq 0$, $\omega \in \Omega$ and all $(f_1, f_2) \in \mathbb{F}_+^2$:*

$$\|\sigma(t, \omega; f_1) - \sigma(t, \omega; f_2)\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} \leq K\|f_1 - f_2\|_{\mathbb{F}}$$

**Assumption 2.4.8** *$\sigma$ has linear growth: for all $t \geq 0$, $\omega \in \Omega$ and all $f \in \mathbb{F}_+$:*

$$\|\sigma(t, \omega; f)\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} \leq K(1 + \|f\|_{\mathbb{F}})$$

**Assumption 2.4.9** *For all $t, x \geq 0$, $\omega \in \Omega$ and all $f \in \mathbb{F}_+$:*

$$\|\delta_x \sigma(t, \omega; f)\|_{\mathbb{G}^*} \leq K f(x)$$

**Remark 2.4.10** *Unconventional notation:*
*Notice that we use the notation $\|\delta_x \sigma(t, \omega; f)\|_{\mathbb{G}^*}$, which is of course equivalent to the more conventional $\|\sigma(t, \omega; f)^* \delta_x\|_{\mathbb{G}}$.*

We define $v_t$, the forward variance swaps curve, to be the unique continuous mild solution in $\mathbb{F}$, started at $v_0$ (see Proposition C.0.3 in the appendix) of the stochastic evolution equation:

$$dv_t = Av_t dt + \sigma(t, \omega; v_t) dW_t$$

We recall here that this means that $(v_t)_{t\geq 0}$ is the (unique continuous) solution of the equation:

$$v_t = S_t v_0 + \int_{u=0}^{t} S_{t-u} \sigma(u, \omega; v_u) dW_u \tag{2.4.1}$$

Let us give some motivation behind the assumptions we have made on $\sigma$:

- Assumption 2.4.7 and Assumption 2.4.8 guarantee that we can apply Proposition C.0.3 and therefore that $v_t$ is well defined, continuous and has moments of all orders.

- Assumption 2.4.9 guarantees that, almost surely, the curve $v_t(.)$ remains positive for all $t \geq 0$. Indeed, the assumption allows us to define, for each $T \geq 0$, the (positive) Doleans exponential local martingale: $(\mathcal{E}(\int_{u=0}^{t} \frac{\delta_{T-u}\sigma_u}{v_u(T-u)} dW_u))_{t\in[0,T]}$, as the integrand $\frac{\delta_{T-u}\sigma_u}{v_u(T-u)}$ is, in norm, smaller than $K$. Using Ito's formula shows easily that this Doleans exponenatial satisfies the same SDE as $(v_t(T-t))_{t\in[0,T]}$.

Notice that we will often use the notation $(\sigma_t)_{t\geq 0}$ for the (now well defined) stochastic process $(t,\omega) \mapsto \sigma(t,\omega; v_t(\omega))$. However, this does not mean that we consider $\sigma$ to be non random, nor independent of $v_t$.

**Remark 2.4.11** *The trivial case where $\sigma = 0$ and where $v_0(.)$ is a constant function corresponds to the Black and Scholes model. This choice of $\sigma$ satisfies of course the three assumptions above, but it is a 1-dimensional model. Recasting the (1-dimensional) Heston model in the above formulation would lead to $\delta_x\sigma$ being proportional to $\sqrt{v_t(x)}$ and therefore would not satisfy Assumption 2.4.9. We will see later that the model we propose here has finite moments at all times, i.e. for any $T \geq 0$ and any $p \geq 1$, $\mathbb{E}P_T^p < \infty$, which is not the case of Heston's model. Empirical evidence seem to favor models which display moments explosions (see for instance Keller-Ressel's paper [27] on moments explosion for affine stochastic volatility models), so that our model may be unrealistic. However, our assumption is mathematically convenient.*

**Definition and dynamics of the variance swaps curve process $(V_t)_{t\geq 0}$**

For all $(t,x) \geq 0$, we define the variance swaps by:

$$V_t(x) := \int_{s=0}^{t} v_s(0)ds + \int_{s=0}^{x} v_t(s)ds$$

and we claim the following:

**Proposition 2.4.12** *The variance swaps curve process $(V_t)_{t\geq 0}$ is an $\tilde{\mathbb{F}}$-valued stochastic process, unique continuous mild solution of the stochastic equation:*

$$V_0(.) = \int_{s=0}^{.} v_0(s)ds$$

$$dV_t = AV_tdt + \Sigma(t,\omega; V_t)dW_t$$

*where the map $\Sigma$, defined by:*

$$\Sigma(u,\omega;V) := J\sigma(u,\omega;DV) \tag{2.4.2}$$

*is measurable from* $\left(\mathbb{R}_+ \times \Omega \times \tilde{\mathbb{F}}_{\nearrow}, \mathcal{P} \times \mathcal{B}(\tilde{\mathbb{F}}_{\nearrow})\right)$ *into* $\left(\mathcal{L}_{HS}(\mathbb{G},\tilde{\mathbb{F}}), \mathcal{B}(\mathcal{L}_{HS}(\mathbb{G},\tilde{\mathbb{F}}))\right)$ *and satisfies, mutatis mutandis, Assumption 2.4.7, Assumption 2.4.8 and Assumption 2.4.9.*

Proof of Proposition 2.4.12:

By definition of $V$, we have for all $t \geq 0$ and all $x \geq 0$:

$$V_t(x) = \int_{s=0}^{t} v_s(0)ds + \int_{s=0}^{x} v_t(s)ds$$

So that replacing $v_s(0)$ and $v_t(s)$ by their expression as stochastic integrals gives:

$$\begin{aligned}
V_t(x) &= \int_{s=0}^{t} \Big((S_s v_0)(0) + \delta_0 \int_{u=0}^{s} S_{s-u}\sigma_u dW_u\Big)ds \\
&+ \int_{s=0}^{x} \Big((S_t v_0)(s) + \delta_s \int_{u=0}^{t} S_{t-u}\sigma_u dW_u\Big)ds \\
&= \int_{s=0}^{t} v_0(s)ds + \int_{s=0}^{x} v_0(t+s)ds \\
&+ \int_{s=0}^{t}\Big(\int_{u=0}^{s} \delta_0 S_{s-u}\sigma_u dW_u\Big)ds + \int_{s=0}^{x}\Big(\int_{u=0}^{t} \delta_s S_{t-u}\sigma_u dW_u\Big)ds \\
&= \int_{s=0}^{t+x} v_0(s)ds + \int_{u=0}^{t}\Big(\int_{s=u}^{t} \delta_{s-u}\sigma_u ds\Big)dW_u + \int_{u=0}^{t}\Big(\int_{s=0}^{x} \delta_{t+s-u}\sigma_u ds\Big)dW_u \\
&= (S_t V_0)(x) + \int_{u=0}^{t}\Big(\int_{s=0}^{t+x-u} \delta_s\sigma_u ds\Big)dW_u \\
&= \delta_x(S_t V_0) + \delta_x \int_{u=0}^{t} S_{t-u} J\sigma(u,\omega;DV_u)dW_u \\
&= \delta_x\Big(S_t V_0 + \int_{u=0}^{t} S_{t-u}\Sigma_u dW_u\Big)
\end{aligned}$$

Notice that we have made use of the Stochastic Fubini theorem in the above calculation, and this is justified by Assumption 2.4.8 which, together with the finiteness of $\sup_{t\leq T} \mathbb{E}\|v_t\|^2$, imply that for all $T \geq 0$:

$$\mathbb{E}\int_{u=0}^{T} \|\sigma_u\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} du \; < \; \infty$$

The Lipschitz property of $\Sigma$ can be easily seen by the following computation, which makes use of Property 2.4.6 which guarantees the boundedness of the derivative operator

$D$ and of Property 2.4.5 which guarantees the boundedness of the integration functional $J$: for all $(V_1, V_2) \in \tilde{\mathbb{F}}$, all $t \in \mathbb{R}_+$ and all $\omega \in \Omega$, we have:

$$\begin{aligned}
\|\Sigma(t,\omega;V_1) - \Sigma(t,\omega;V_2)\|_{\mathcal{L}_{HS}(G,\tilde{\mathbb{F}})} &\le \|J\|_{\mathcal{L}(\mathbb{F},\tilde{\mathbb{F}})}\|\sigma(t,\omega;DV_1) - \sigma(t,\omega;DV_2)\|_{\mathcal{L}_{HS}(G,\mathbb{F})} \\
&\le K\|J\|_{\mathcal{L}(\mathbb{F},\tilde{\mathbb{F}})}\|D\|_{\mathcal{L}(\tilde{\mathbb{F}},\mathbb{F})}\|V_1 - V_2\|_{\tilde{\mathbb{F}}} \\
&=: K_\Sigma\|V_1 - V_2\|_{\tilde{\mathbb{F}}}
\end{aligned}$$

The boundedness of $\Sigma$ in the sense of Assumption 2.4.8 is also easy to show. We indeed have for all $(V, t, \omega) \in \tilde{\mathbb{F}} \times \mathbb{R}_+ \times \Omega$ that:

$$\begin{aligned}
\|\Sigma(t,\omega;V)\| &\le \|J\|_{\mathcal{L}(\mathbb{F},\tilde{\mathbb{F}})}K(1 + \|D\|_{\mathcal{L}(\tilde{\mathbb{F}},\mathbb{F})}\|V\|_{\tilde{\mathbb{F}}}) \\
&\le \tilde{K}_\Sigma(1 + \|V\|_{\tilde{\mathbb{F}}})
\end{aligned}$$

where $\tilde{K}_\Sigma$ is any constant greater than $K\|J\|_{\mathcal{L}(\mathbb{F},\tilde{\mathbb{F}})}$ and $K\|J\|_{\mathcal{L}(\mathbb{F},\tilde{\mathbb{F}})}\|D\|_{\mathcal{L}(\tilde{\mathbb{F}},\mathbb{F})}$.

We now show that Assumption 2.4.9 holds, almost surely, for $\Sigma$: let $t, T, x > 0$ be given (Note the strict inequality which was not necessary in the case of $\sigma$, because $v_t$ was strictly positive almost surely, whereas $V_0(0) = 0$). Let us remark that $V_T(x) = \int_{s=0}^T v_s(0)ds + \int_{s=0}^x v_T(s)ds \ge \int_{s=0}^x v_T(s)ds = \delta_x J v_T$. Therefore, it holds almost surely that:

$$\begin{aligned}
\frac{\|\delta_x\Sigma_t\|_{\mathbb{G}^*}}{\delta_x V_t} \le \frac{\|\delta_x J\sigma_t\|_{\mathbb{G}^*}}{\delta_x J v_t} &= \frac{\|\int_{s=0}^x \sigma_t(s)ds\|_{\mathbb{G}^*}}{\int_{s=0}^x v_t(s)ds} \\
&\le \frac{\int_{s=0}^x \|\delta_s\sigma_t\|_{\mathbb{G}^*}ds}{\int_{s=0}^x \delta_s v_t ds} \le K
\end{aligned}$$

where $K$ is as in Assumption 2.4.9.■

Let us make two additional remarks on $\Sigma$: firstly it is obvious that $\Sigma_t g \in \tilde{\mathbb{F}}_0$ for all $g \in \mathbb{G}$, which means that $\Sigma$ is actually valued in the space $(\mathcal{L}_{HS}(\mathbb{G}, \tilde{\mathbb{F}}_0))$. Secondly, if for all $(t, \omega, v) \in \mathbb{R}_+ \times \Omega \times \mathbb{F}_+$, $\sigma(t, \omega; v)$ is dense-range in $\mathbb{F}$ (this assumption will actually appear naturally later in order to get uniqueness of hedging portfolios), then $\Sigma(t, \omega; V)$ is dense-range in $\tilde{\mathbb{F}}_0$ for all $(t, \omega, V) \in \mathbb{R}_+ \times \Omega \times \tilde{\mathbb{F}}_{\nearrow}$.

**Remark 2.4.13** *Let us finally remark that by construction, $V$ has the correct properties to preclude arbitrage, since $(V_t(T-t))_{t\in[0,T]}$ is a martingale for all $T \geq 0$. The construction also guarantees that the variance swaps curve $V_t(.)$ is increasing for all $t$.*

### Definition and dynamics of the stock process $(P_t)_{t\geq 0}$

Finally, we define the stock $P_t$ as follows: let $P_0 > 0$ be the initial value of the stock, and let $(\lambda_t)_{t\geq 0}$ be a $\mathbb{G}^*$-valued stochastic process, adapted and of norm 1. We then define the stock price process $(P_t)_{t\geq 0}$ to be the local-martingale Doleans exponential:

$$\begin{aligned} P_t &= P_0 \mathcal{E}\Big( \int_{u=0}^{t} \sqrt{v_u(0)} \lambda_u dW_u \Big) \\ &=: P_0 \mathcal{E}\Big( \int_{u=0}^{t} \sqrt{v_u(0)} dW_u^{(\lambda)} \Big) \end{aligned}$$

**Remark 2.4.14** *So far, there is no reason to believe that $(P_t)_{t\geq 0}$ is a martingale. However, notice that it is a well defined positive local-martingale. Indeed, from Theorem C.0.3 and Assumption 2.4.3, we get that for all $T \geq 0$:*

$$\mathbb{E} \int_{u=0}^{T} v_u(0) du \;\leq\; \|\delta_0\|_{\mathbb{F}} \sqrt{T A_{T,2}} \;<\; \infty.$$

### Rationale behind the Hilbert space valued SPDE

For the reader unfamiliar with Musiela's notation and SPDE's in Hilbert spaces, it may be unclear why we chose to model $v_t(.)$ via an equation of the form (2.4.1). We explain here, informally, the rationale behind this model: as we are working directly under a pricing measure, we could start off with a model where for any $T$, $(v(t,T))_{t\geq 0}$ is a local-martingale, constant after $T$. This could be achieved by taking $v(t,T)$ solution of a family (indexed by $T$) of stochastic differential equations (SDE), where for each $T$, $v(.,T)$ is solution of:

$$v(t,T) = v(0,T) \;+\; \int_{u=0}^{t\wedge T} \mu_u^{(T)} dW_u \tag{2.4.3}$$

for some finite dimensional or infinite dimensional Brownian motion $W$ defined (cylindrically) in a (separable) Hilbert space $\mathbb{G}$.

Of course, requirements would have to be imposed on the family (indexed by $T$) of processes $\mu^{(T)}$ to ensure that $v(t,T)$ has the desired properties (positive at all times, absolutely continuous in $T$, etc). Another approach is to see $(v(t,.))_{t\geq 0}$ as a stochastic process evolving in some Hilbert space $\mathbb{F}$ of appropriate functions (i.e. for each $t$, $v(t,.)$ is a curve). We can then rewrite (2.4.3) as a stochastic differential equation in $\mathbb{F}$, under the form:

$$v(t,.) = v(0,.) \ + \ \int_{u=0}^{t} \eta_u dW_u \tag{2.4.4}$$

where $\eta_u$ is a suitable Hilbert-Schmidt operator from $\mathbb{G}$ to $\mathbb{F}$ (readers not familiar with stochastic integration in infinite dimension may consult the appendix). Under some suitable assumptions of course, we could identify $\mu_u^{(T)}$ with $\delta_T \eta_u$ (the composition of $\delta_T$, the evaluation functional at $T$, and $\eta$, the volatility of the swap curve, so that in some sense, both ways of doing (family of SDEs versus SDE in a Hilbert space) are equivalent.

Let us consider that we start from (2.4.4), where $\eta_u$ is an appropriate integrand. We now would like to see what equation should $v_t$ satisfy, where $v_t(.)$ is $v(t,.)$ expressed in the Musiela notation introduced earlier: $v_t(x) := v(t, t+x)$. Notice that, because for each $T$, $v(.,T)$ is constant after $T$, we can pass without loss of information from the family $(v_t)_{t\geq 0}$ to the family $(v(t,.))_{t\geq 0}$ and vice versa. Getting from $v(t,.)$ to $v_t(.)$ is done easily via the left shift operator $S_t$ defined by $S_t f(.) := f(t+.)$. Getting the other way can be done via the right shift operators $(S_{-t})_{t\geq 0}$ defined by $(S_{-t}f)(x) = f(x-t)$ for $x \leq t$ and $(S_{-t}f)(x) = f(0)$ if $x < t$, and then modifying the (irrelevant) parts of $t \mapsto v(t,T)$ for $t \in [0,T]$ so that each $v(.,T)$ is constant after $T$.

The fact that $v(.,T)$ is constant for $t \geq T$ also implies that $\delta_v \eta_u = 0$ for all $v \leq u$. This means that we can, without loss, apply $S_{-u} \circ S_u$ to $\eta_u$ (which is in general not true, because doing this would cut off the initial part of the curve to which we apply $S_{-u} \circ S_u$). This observation justifies the following calculation and the definition of $\sigma$:

$$\begin{aligned} v_t(.) = S_t v(t,.) = S_t v(0,.) + \int_{u=0}^{t} S_t \eta_u dW_u \\ = S_t v_0(.) + \int_{u=0}^{t} S_t S_{-u} S_u \eta_u dW_u \\ = S_t v_0(.) + \int_{u=0}^{t} S_{t-u} \sigma_u dW_u \end{aligned}$$

where we define:

$$\sigma_u := S_u \eta_u$$

Therefore, the equivalent of (2.4.4) in Musiela's notation is:

$$v_t(.) = S_t v_0(.) + \int_{u=0}^{t} S_{t-u} \sigma_u dW_u \tag{2.4.5}$$

■

## 2.5 Uniqueness of hedging portfolio and market incompleteness

### 2.5.1 Trading strategies and the self-financing condition

Our assumption that for any time $t \geq 0$ there exists a variance swap maturing at any later time $T$ lead us to model directly the whole variance swaps curve as the solution to a SDE in a function space $\tilde{\mathbb{F}}$. Continuing further with this functional analysis approximation of the real world, we will assume that traders can not only hold "atomic" portfolios, consisting of a finite number of variance swaps maturing at different times in the future, but that they can also hold more general portfolios valued in the dual space of $\tilde{\mathbb{F}}$. Notice that this is the same kind of trick, introduced for mathematical convenience, as the one we generally follow when passing from trading strategies valued in the space of simple integrands to trading strategies valued in the space of predictable square integrable processes. Of course in the real world, only simple integrands make sensible trading strategies, but that space is not complete and is therefore not well suited to perform mathematical analysis.

We have assumed throughout that interest rates are null, and therefore, in addition to trading in the variance swaps curve $V(t, .)$ and in the stock $P_t$, we can also hold a certain amount of cash, which plays the role of the usual bank account (paying no interest). We will therefore authorize ourselves to hold portfolios of the form $\phi = (\phi^{(C)}, \phi^{(P)}, \phi^{(V)})$ valued in $\mathbb{R} \times \mathbb{R} \times \tilde{\mathbb{F}}^*$, by which we mean that we hold $\phi^{(C)}$ in cash, $\phi^{(P)}$ units of the stock $P$ and $\phi^{(V)}$ of the variance swaps curve $V$. The wealth associated with portfolio $\phi$ is then defined by: $X^{(\phi)} := \phi^{(C)} + \phi^{(P)} P + < \phi^{(V)}, V >_{\tilde{\mathbb{F}}}$, or more shortly: $X^{(\phi)} =< \phi, (1, P, V) >_{\mathbb{H}}$, where the Hilbert space product in $\mathbb{H} := \mathbb{R} \times \mathbb{R} \times \tilde{\mathbb{F}}^*$ is defined in the obvious way.

We now would like to define the self-financing condition in such a way that we can talk about $\mathbb{H}^*$-valued self-financing trading strategies and that it is consistent with common sense. That is, our definition should mean that whenever the trading strategy is an atomic measure of $\mathbb{H}^*$ (i.e. we hold cash, the stock and a finite number of variance

swaps) and is a simple integrand, then our portfolio's value is changing only through changes in the assets' value, not because of any additional expense or income.
We already know from classical theory what the self-financing condition is for continuous time trading in a scalar asset (the stock $P$), so we can focus on the condition for trading in the variance swaps curve alone. Now, by linearity of the infinite dimensional stochastic integral, we see that it is clearly sufficient to take three times $T_1 < T_2 < T$ and to define the self-financing condition in the case of the atomic simple strategy $\phi_t^{(V)} = \mathbb{1}_{[T_1,T_2]}(t)\delta_{T-t}$. This strategy consists simply in holding between times $T_1$ and $T_2$ the variance swap that matures at time $T$. In this case, the self-financing condition means:

$$\begin{aligned}
X_{T_2} - X_{T_1} &= V_{T_2}(T-T_2) - V_{T_1}(T-T_1) \\
&= \delta_{T-T_2}(S_{T_2}V_0 + \int_{u=0}^{T_2} S_{T_2-u}\Sigma_u dW_u) - \delta_{T-T_1}(S_{T_1}V_0 + \int_{u=0}^{T_1} S_{T_1-u}\Sigma_u dW_u) \\
&= V_0(T) + \int_{u=0}^{T_2} \delta_{T-T_2}S_{T_2-u}\Sigma_u dW_u - V_0(T) - \int_{u=0}^{T_1} \delta_{T-T_1}S_{T_1-u}\Sigma_u dW_u \\
&= \int_{u=0}^{T_2} \delta_{T-u}\Sigma_u dW_u - \int_{u=0}^{T_1} \delta_{T-u}\Sigma_u dW_u \\
&= \int_{u=T_1}^{T_2} \delta_{T-u}\Sigma_u dW_u = \int_{u=0}^{T} \phi_u^{(V)}\Sigma_u dW_u
\end{aligned}$$

This leads us to the following definition that we adopt in order to extend to our setup the classical notions of trading strategies and of self-financing portfolios:

**Definition 2.5.1** *An $\mathbb{H}^*$-valued stochastic process $(\phi_t)_{t\leq 0} = (\phi_t^{(C)}, \phi_t^{(P)}, \phi_t^{(V)})_{t\leq 0}$ is called an admissible trading strategy if:*

- *$\phi$ is predictable in the sense that, as a mapping from $(\mathbb{R}_+ \times \Omega)$ to $\mathbb{H}^*$, $(t,\omega) \mapsto \phi(t,\omega)$ is $\mathcal{P}/\mathcal{B}(\mathbb{H}^*)$ measurable.*
- *$\phi$ is stochastically integrable in the (strong) sense that for all $T \geq 0$:*

$$\mathbb{E}\int_{u=0}^{T} (\phi_u^{(P)})^2 P_u^2 v_u(0) du + \mathbb{E}\int_{u=0}^{T} \|\phi_u^{(V)}\Sigma_u\|_{\mathbb{G}^*}^2 du \ < \ \infty.$$

*$(\phi_t)_{t\leq 0}$ is in addition said to be self-financing if, for all $T \geq 0$:*

$$< \phi_T, (1, P_T, V_T) >_{\mathbb{H}} = < \phi_0, (1, P_0, V_0) >_{\mathbb{H}} + \int_{u=0}^{T} \phi_u^{(P)} P_u \sqrt{v_u(0)} dW_u^{(\lambda)} + \int_{u=0}^{T} \phi_u^{(V)}\Sigma_u dW_u.$$

*and when this holds, we define the wealth $X_T^{(\phi)}$ at time $T$ of an investor following the self-financing trading strategy $\phi$ by:*

$$X_T^{(\phi)} :=< \phi_T, (1, P_T, V_T) >_{\mathbb{H}} .$$

It is interesting to notice that from a mathematical perspective, the definition of the stochastic integral $\int_{u=0}^T \phi_u^{(V)} \Sigma_u dW_u$ does not require that $\phi_u^{(V)}$ be $\tilde{\mathbb{F}}^*$-valued. It would indeed be enough that $\phi_u^{(V)}$ be in the closure of the space of $\tilde{\mathbb{F}}^*$-valued adapted processes (the closure being understood with respect to the norm $|||\phi_u^{(V)}|||_{\Sigma,T} := \int_{u=0}^T ||\phi_u^{(V)} \Sigma_u||_{\mathbb{G}^*}^2 du$). This requirement is rather justified by us wanting to interpret $< \phi_T, (1, P_T, V_T) >_{\mathbb{H}}$ as the time $T$ wealth of a trader following strategy $\phi$, and therefore, we need to be able to make sense of $< \phi_u^{(V)}, V_u >_{\tilde{\mathbb{F}}}$ as a real number.

### 2.5.2 Equivalence of trading strategies in $v_t$ and $V_t$

Notice that we have chosen for our state space $\tilde{\mathbb{F}}$ a relatively smooth space, and in particular smooth enough that the point-wise differential operators are continuous. This means that we authorize ourselves to hold portfolios such as $\frac{\partial}{\partial_x}\Big|_{x=x_0} V_t(.)$, which amounts to holding the forward variance swap $v_t(x_0)$.

In our setup, it actually turns out that the notions of trading strategies in $V_t$ and in $v_t$ are equivalent, so that we will in the sequel examine everything as if we were holding portfolios of forward variance swaps. Indeed, let us suppose that $\phi_t^{(V)}$ is an admissible self-financing trading strategy in the variance swaps curve. Then we can simply define $\phi_t^{(v)} := \phi_t^{(V)} J$, which can easily be seen to be an admissible self-financing strategy in the forward variance swaps curve, and leading to the same wealth at all times. Conversely, if we start with an admissible self-financing trading strategy $\phi_t^{(v)}$ in the forward variance curve, then we can define $\phi_t^{(V)} := \phi_t^{(v)} D$, etc.

### 2.5.3 Uniqueness of self-financing strategies

In finite dimension, we had no chance of having uniqueness of the self-financing hedging portfolio for a given claim, because the adjoint $\sigma^*$ of the volatility of the variance swaps curve has a kernel which is necessarily non trivial. In infinite dimensions, the situation is very different as the following theorem, which is the first main result of this chapter, shows (see Proposition 6.6 in Carmona and Tehranchi [8] for the equivalent proposition in infinite dimensional interest rates modelling).

**Theorem 2.5.2** ***Uniqueness of hedging portfolios:***
*Suppose that the following assumptions hold:*

***Assumption 2.5.3*** *For almost all $(t,\omega) \in (\mathbb{R}_+ \times \Omega)$:*

$$\ker\left[\sigma(t,\omega;v_t(\omega))^*\right] = \{0_{\mathbb{F}}\}.$$

***Assumption 2.5.4*** *For almost all $(t,\omega) \in (\mathbb{R}_+ \times \Omega)$:*

$$\lambda_t(\omega)^* \notin [\ker\sigma(t,\omega;v_t(\omega))]^{\perp}.$$

*Let $\phi_t$ and $\psi_t$ be two self-financing trading strategies such that, for some $T > 0$:*

$$X_T^{(\phi)} = X_T^{(\psi)} almost\ surely.$$

*Then it must hold for almost all $(t,\omega) \in [0,T] \times \Omega$, that:*

$$\phi_t(\omega) = \psi_t(\omega).$$

Proof of Theorem 2.5.2:

Let us consider the self-financing trading strategy $\zeta_t(\omega) := \phi_t(\omega) - \psi_t(\omega)$. Then we have that:

$$X_T^{(\zeta)} = \int_{t=0}^{T} \zeta_t^{(P)} P_t \sqrt{v_t(0)} dW_t^{(\lambda)} + \int_{t=0}^{T} \zeta_t^{(v)} \sigma_t dW_t = 0 \tag{2.5.1}$$

This implies, by Ito's isometry, that for almost all $(t, \omega)$:

$$\zeta_t^{(P)} P_t \sqrt{v_t(0)} \lambda_t + \zeta_t^{(v)} \sigma_t = 0 \tag{2.5.2}$$

But because of the assumption that for almost all $(t, \omega)$, $\lambda_t(\omega)^* \notin [\ker \sigma(t, \omega; v_t(\omega))]^\perp$, we can find an element $g_t(\omega) \in \mathbb{G}$ such that $\sigma_t(\omega) g_t(\omega) = 0$, and $\lambda_t(\omega) g_t(\omega) > 0$. This implies that, for almost all $(t, \omega)$:

$$\zeta_t^{(P)} P_t \sqrt{v_t(0)} \lambda_t g_t(\omega) = 0$$
$$P_t \sqrt{v_t(0)} \lambda_t g_t(\omega) > 0$$

and thus $\zeta_t^{(P)}(\omega) = 0$ for almost all $(t, \omega)$. Equation (2.5.2) now gives that $\zeta_t^{(v)} \sigma_t = 0$ for almost all $(t, \omega)$, which finally together with the assumption that $\ker [\sigma(t, \omega; v_t(\omega))^*] = \{0_{\mathbb{F}}\}$ for almost all $(t, \omega)$ means that $\zeta_t^{(v)}(\omega) = 0$ a.s.. ∎

### 2.5.4 Incompleteness

The uniqueness of hedging portfolios discussed above has strong implications in terms of market completeness of course. Unlike the finite dimensional case, where the market can always be completed by adding new instruments (most likely, $d$ instruments in a $d$-dimensional continuous model will lead to a complete model, for instance like we have seen a variance swap along with the stock will complete a Heston model), the infinite dimensional model we have here cannot be complete, no matter how many variance swaps (of different maturities) we choose as tradable instruments. Indeed, even if we could trade $N$ different variance swaps $\{v_t(T_n - t)\}_{n \in \{1,\ldots,N\}}$, for some (possibly very large) integer $N$, by the uniqueness of trading strategies, we can always choose $T$ different from all the $T_n$'s, and the time $T$ payoff $\xi := v_T(T)$ is not replicable by any trading strategy in terms of the other $N$ variance swaps. In that respect, continuous infinite dimensional models are more realistic than continuous finite dimensional models. Of course, in infinite dimensional setups, one generally considers that agents can hold portfolios valued in the dual space of the space in which the asset (curve) is valued, so we do not limit ourselves to atomic portfolios. The question of which family of contingent claims can be replicated is then complicated and in general there is no market completeness, but only approximate completeness. The reader is referred to the papers by Taflin [43] and De-Donno and Pratelli [15] for instance, for a definition and a detailed discussion of approximate completeness.

## 2.6 Characterization of hedging portfolios

In this section, we derive an explicit representation of the hedging portfolio for a class of contingent claims written on the stock and variance instruments, in the specific case where the model is Markovian. We also show that under some conditions on the volatility operator $\sigma$ and on the correlation vector $\lambda$, the (unique) hedging portfolio satisfies a maturity-specific property which we have seen was lacking in finite dimensional models. Notice that we will limit our study to contingent claims $\xi$ of the form:

$$\xi = g(P_{T_1}, ..., P_{T_n}, v_{T_{n+1}}, ..., v_{T_{n+m}}) \tag{2.6.1}$$

where $n$ and $m$ are integers, $\{T_j\}_{j=1}^{n+m}$ is a sequence of times, and where the function $g$ is measurable from $\left(\mathbb{R}_+^n \times \mathbb{F}_+^m, \mathcal{B}(\mathbb{R}_+^n) \times \mathcal{B}(\mathbb{F}_+^m)\right)$ into $\left(\mathbb{R}, \mathcal{B}(\mathbb{R})\right)$ and sufficiently well behaved (i.e. Lipschitz).

**Remark 2.6.1** *Because many (real world) payoffs of interest are more naturally expressed in terms of the variance swaps curve $V_T$, we will also explain how the formula we derive for hedging claims of the form (2.6.1) can be modified to hedge options of the form:*

$$\zeta = h(P_{T_1}, ..., P_{T_n}, V_{T_{n+1}}, ..., V_{T_{n+m}}) \tag{2.6.2}$$

As we intend to make use of the Clark-Ocone formula on $\xi$, we see that we will have first to prove that the $T$-time prices of the different assets $P_T$, $v_T$ and $V_T$ are Malliavin differentiable, and then use some chain rule to differentiate $\xi$. In our setup, it actually turns out that $P_T$, $v_T$ and $V_T$ belong to the appropriate $\mathbb{D}^{1,\infty}$ spaces. This makes things simpler when it comes to expressions involving products or powers of such quantities, as those will also automatically belong to $\mathbb{D}^{1,\infty}$. Of course, the reason why we choose to specialize to Markovian coefficients $\sigma$ and $\lambda$ is that it allows us to make use of chain rules for the Malliavin derivative. Precisely, we make the assumption that:

$$\sigma_u = \sigma(u; v_u), \ \lambda_u = \lambda(u; v_u).$$

We will also make the further assumption that $\lambda$ is globally Lipschitz, in the sense that:

**Assumption 2.6.2** *There exists a constant $L \geq 0$ such that for all $(t, v_1, v_2) \in \mathbb{R}_+ \times \mathbb{F}_+^2$:*

$$\Big\|\lambda(t, v_1) - \lambda(t, v_2)\Big\|_{\mathbb{G}^*} \;\leq\; L\Big\|v_1 - v_2\Big\|_{\mathbb{F}}.$$

### 2.6.1 Malliavin differentiability of the variance instruments

We start with two lemmas that will be useful later when we will prove that $q$-powers of variance swaps and forward variance swaps are Malliavin differentiable. Notice that the exponent $q$ in these two propositions can take negative values as well as positive.

**Lemma 2.6.3** *For any $T, x \geq 0$:*

$$\sup_{t \leq T} \mathbb{E}[v_t(x)^q] \;<\; \infty, \;\forall\, q \in \mathbb{R}$$

**Lemma 2.6.4** *For any $T \geq 0$ and any $y > 0$:*

$$\sup_{t \leq T} \mathbb{E}[V_t(y)^q] \;<\; \infty, \;\forall\, q \in \mathbb{R}$$

<u>Proofs of Lemma 2.6.3 and Lemma 2.6.4</u>:

For any $T \geq t \geq 0$, we have that:

$$v_t(T - t) = v_0(T) + \int_{u=0}^{t} \delta_{T-u}\sigma_u dW_u$$

Because of Assumption 2.4.9, we can rewrite $v_t(T - t)$ in a more convenient form, which shows by the way its positivity:

$$v_t(T - t) = v_0(T)\mathcal{E}\Big(\int_{u=0}^{t} \frac{\delta_{T-u}\sigma_u}{v_u(T - u)} dW_u\Big)$$

where $\mathcal{E}$ denotes the usual Doleans exponential local-martingale.

By the Cauchy-Schwarz inequality, we get:

$$\begin{aligned}\mathbb{E}[v_t(T-t)^q] &\leq v_0(T)^q \sqrt{\mathbb{E}\mathcal{E}\Big(2q\int_{u=0}^{t}\frac{\delta_{T-u}\sigma_u}{v_u(T-u)}dW_u\Big)} \\ &\cdot \sqrt{\mathbb{E}\exp\Big\{q(2q-1)\int_{u=0}^{t}\frac{\|\delta_{T-u}\sigma_u\|^2_{\mathbb{G}^*}}{v_u(T-u)^2}du\Big\}} \\ &=: v_0(T)^q\sqrt{\mathbb{E}Q_t^{(T)}}\sqrt{\mathbb{E}R_t^{(T)}}\end{aligned}$$

The right hand side is composed of three terms. The first one is constant, the second one is smaller than 1 given that $(Q_t^{(T)})_{t\in[0,T]}$ is a positive local-martingale (and thus a super-martingale) started at 1. The last term is bounded, as $R_t^{(T)}$ is itself bounded by $\exp\left\{|q|(2|q|+1)TK^2\right\}$ where $K$ is as in Assumption 2.4.9. It is now enough to set $T = t+x$ in the above to get that, for an arbitrary real number $q$ and an arbitrary $t \geq 0$:

$$\mathbb{E}[v_t(x)^q] \leq v_0(t+x)^q \exp\Big[\frac{1}{2}|q|(2|q|+1)tK^2\Big]$$

Consequently, we get that, for $q$ and an arbitrary $T \geq 0$:

$$\sup_{t\leq T}\mathbb{E}[v_t(x)^q] \leq \max_{t\leq T}\{v_0(t+x)^q\}\exp\Big[\frac{1}{2}|q|(2|q|+1)TK^2\Big] < \infty$$

The proof of Lemma 2.6.4 is similar. ■

The following proposition and its corollaries give us formulae for the Malliavin derivatives of forward variance swaps related quantities in terms of the operator $\sigma$:

**Proposition 2.6.5** *For all $T \geq 0$, $v_T \in \mathbb{D}^{1,\infty}(\mathbb{F})$, and $D_t v_T$ is given by the formula:*

$$D_t v_T = \mathbb{1}_{\{t\leq T\}} Y_{t,T}.\sigma_t$$

*where $(Y_{t,T})_{0\leq t\leq T}$ is the family (indexed by $t, T$) of strong $\mathcal{L}(\mathbb{F})$-valued random variables, solutions of the family of equations:*

$$Y_{t,T} = S_{T-t} + \int_{u=t}^{T} S_{T-u}\nabla\sigma_u Y_{t,u}.dW_u \tag{2.6.3}$$

*Moreover, for any $T \geq t \geq 0$, $p \geq 2$ and $f_t \in L^p(\Omega; \mathbb{F})$, it holds that:*

$$\mathbb{E} \sup_{u \in [t,T]} \left\| Y_{t,u} f_t \right\|_{\mathbb{F}}^p \leq C_{T,p} \left\| f_t \right\|_{\mathbb{F}}^p$$

*for some constant $C_{T,p}$ depending only on $T$ and $p$.*

**Corollary 2.6.6** *For all $T, x \geq 0$ and all $q \in \mathbb{R}$, $[v_T(x)]^q \in \mathbb{D}^{1,\infty}$, and $D_t[v_T(x)]^q$ is given by the formula:*

$$D_t[v_T(x)]^q \;=\; \mathbb{1}_{\{t \leq T\}} q [v_T(x)]^{q-1} \delta_x Y_{t,T} \sigma_t$$

**Lemma 2.6.7** *For all $T \geq 0$, $\mathcal{V}_T := \int_{u=0}^T v_u du \in \mathbb{D}^{1,\infty}(\mathbb{F})$, and $D_t \mathcal{V}_T$ is given by the formula:*

$$D_t \mathcal{V}_T \;=\; \mathbb{1}_{\{t \leq T\}} \Big( \int_{u=t}^T Y_{t,u}.du \Big) \sigma_t$$

**Corollary 2.6.8** *For all $T, x \geq 0$, $\mathcal{V}_T(x) = \int_{u=0}^T v_u(x) du \in \mathbb{D}^{1,\infty}$, and $D_t \mathcal{V}_T(x)$ is given by the formula:*

$$D_t \mathcal{V}_T(x) \;=\; \mathbb{1}_{\{t \leq T\}} \Big( \delta_x \int_{u=t}^T Y_{t,u}.du \Big) \, \sigma_t$$

**Proposition 2.6.9** *For all $(T, x) \geq 0$, $\mathcal{A}_T(x) := \int_{u=0}^T \sqrt{v_u(x)} dW_u^{(\lambda)} \in \mathbb{D}^{1,\infty}$, and $D_t \mathcal{A}_T(x)$ is given by the formula:*

$$D_t \mathcal{A}_T(x) = \sqrt{v_t(x)} \lambda_t + \Big( \int_{u=t}^T \frac{1}{2\sqrt{v_u(x)}} \delta_x Y_{t,u}.dW_u^{(\lambda)} \Big) \sigma_t + \Big( \int_{u=t}^T \sqrt{v_u(x)} \triangledown \lambda_u Y_{t,u}.dW_u \Big) \sigma_t$$

We have of course a proposition and a corollary similar to 2.6.5 and 2.6.6 for the variance swaps curve:

**Proposition 2.6.10** *For all $T \geq 0$, $V_T \in \mathbb{D}^{1,\infty}(\tilde{\mathbb{F}})$, and $D_t V_T$ is given by the formula:*

$$D_t V_T = \mathbb{1}_{\{t \leq T\}} Z_{t,T} \sigma_t$$

*where $(Z_{t,T})_{0 \leq t \leq T}$ is the family (indexed by $t, T$) of strong $\mathcal{L}(\mathbb{F}, \tilde{\mathbb{F}})$-valued random variables, solutions of the family of equations:*

$$Z_{t,\tau} = S_{T-t} J + \int_{u=t}^{\tau} S_{T-u} J \nabla \sigma_u Z_{t,u} dW_u \tag{2.6.4}$$

**Corollary 2.6.11** *For all $T \geq 0$, $y > 0$ and $q \in \mathbb{R}$, $[V_T(y)]^q \in \mathbb{D}^{1,\infty}$, and $D_t[V_T(y)]^q$ is given by the formula:*

$$D_t[V_T(y)]^q = \mathbb{1}_{\{t \leq T\}} q [V_T(y)]^{q-1} \delta_y Z_{t,T} \sigma_t$$

Proof of Proposition 2.6.5:
This proposition is a direct application of Theorem C.0.4, which can be found in the appendix. ■

Proof of Corollary 2.6.6: Let $T, x \geq 0$ and $q \in \mathbb{R}$ be given. It is clear that the candidate for the Malliavin derivative of $[v_T(x)]^q$ is $q[v_T(x)]^{q-1}\delta_x D_t v_T$. By Proposition B.3.2, this candidate will indeed be the Malliavin derivative of $[v_T(x)]^q$ if $\mathbb{E}\left\|[v_T(x)]^{q-1}\delta_x D_t v_T\right\|_{\mathbb{F}\otimes\mathbb{U}} < \infty$ (the reader may consult Appendix B for definitions and notations related to Malliavin calculus, that we will be using in this section, for instance for a definition of the space $\mathbb{U}$). That this holds follows from Holder's inequality and the finiteness of $\mathbb{E}[v_T(x)]^q$ for any real number $q$, and of $\mathbb{E}\left\|D_t v_T\right\|_{\mathbb{F}}^p$ for any $p \geq 1$. ■

Proofs of Lemma 2.6.7 and Corollary 2.6.8: Let $T \geq 0$ be given. We are interested in taking the Malliavin derivative of the quantity $\mathcal{V}_T := \int_{u=0}^{T} v_u du$. We first have to show that this quantity makes sense. For this Bochner integral to be well defined, it is clearly sufficient to show that $\int_{u=0}^{T} \|v_u\|_{\mathbb{F}} du$ is finite $\mathbb{P}$-almost surely. But this is true because of (see Theorem C.0.3) the existence of a constant $A_{T,2}$ such that $\sup_{u\leq T}\mathbb{E}\Big\{\|v_u\|_{\mathbb{F}}^2\Big\}\leq A_{T,2}$. The following computation shows that $(\mathcal{V}_t)_{t\leq T}$ is the unique continuous mild solution of a stochastic evolution equation in $\mathbb{F}$, and will help us identify the Malliavin derivative of $\mathcal{V}_T$:

$$\begin{aligned}\mathcal{V}_T &= \int_{u=0}^{T}\Big[S_u v_0 + \int_{t=0}^{u} S_{u-t}\sigma_t dW_t\Big]du\\ &= \int_{u=0}^{T} S_u v_0 du + \int_{t=0}^{T}\int_{u=0}^{T-t} S_u\sigma_t du\, dW_t\end{aligned}$$

Because of the uniform boundedness of the shift operators $S_u, u\leq T$, it is clear that we can apply exactly the same reasoning as for Proposition 2.6.5 to deduce that $\mathcal{V}_T$ belongs to $\mathbb{D}^{1,\infty}(\mathbb{F})$ and to deduce the formula giving $D_t\mathcal{V}_T$. Corollary 2.6.8 is then obvious as $\mathcal{V}_T(x)=\delta_x\mathcal{V}_T$. ■

Proof of Proposition 2.6.9:

We want to make use of Proposition B.2.1 on $\mathcal{A}_T(x) := \int_{t=0}^{T}\sqrt{v_t(x)}\lambda_t dW_t$. Let us fix $T\geq 0$ and $p\geq 2$. We have already seen that for any $q\in\mathbb{R}$, $\sup_{t\leq T}\mathbb{E}\sqrt{v_t(x)}^q<\infty$. Moreover, $\|\lambda_t\|_{\mathbb{G}^*}=1$. Therefore, by a simple application of Cauchy-Schwartz's inequality, it is clear that $\mathbb{E}\int_{t=0}^{T}\Big\|\sqrt{v_t(x)}\lambda_t\Big\|_{\mathbb{G}^*}^p dt<\infty$. It remains only to prove that:

$$\mathbb{E}\int_{t=0}^{T}\Big\|D(\sqrt{v_t(x)}\lambda_t)\Big\|_{\mathbb{G}^*\otimes\mathbb{U}}^p dt<\infty \tag{2.6.5}$$

By the product rule of the Malliavin derivative, and the already established facts that $\sqrt{v_t(x)}\in\mathbb{D}^{1,\infty}$ and that $v_t\in\mathbb{D}^{1,\infty}$ along with formulae for their derivatives, we get that for $u\leq t$:

$$D_u(\sqrt{v_t(x)}\lambda_t) = \sqrt{v_t(x)}\nabla\lambda_t Y_{u,t}\sigma_u + \frac{1}{2\sqrt{v_t(x)}}\delta_x Y_{u,t}\sigma_u\lambda_t$$

The bound (2.6.5) follows then from the uniform boundedness, in $(u,t)\in[0,T]^2, u\leq t$, of the quantities $\sqrt{v_t(x)}$, $\frac{1}{\sqrt{v_t(x)}}$ , $\lambda_t$ and $Y_{u,t}\sigma_u$, in the appropriate $L^p$-norms, from the Lipschitz bound $L$ on $\nabla\lambda_t$, from the boundedness of $\delta_x$ and from the application again of Cauchy-Schwartz's inequality. ■

Proofs of Proposition 2.6.10 and Corollary 2.6.11:

The proofs are exactly similar to the ones of Proposition 2.6.5 and Corollary 2.6.6, using the fact that

$$V_T = S_T J v_0 + \int_{u=0}^{T} S_{T-u} J \sigma_u dW_u$$

■

## 2.6.2 Malliavin differentiability of the stock price

From now on, we assume that in addition to the conditions already listed previously, $\sigma$ satisfies the following:

**Assumption 2.6.12** *For all* $(T, p) \geq 0$*:*

$$\mathbb{E} \exp \Big[ p \int_{u=0}^{T} \int_{t=0}^{T} \|\delta_t \sigma_u\|^2_{\mathbb{G}^*} dt du \Big] < \infty$$

Under the above additional assumption, we will start by proving that the stock has all its (positive) moments finite:

**Lemma 2.6.13** *For all* $T, p \geq 0$

$$\mathbb{E}[P_T^p] < \infty$$

The following proposition then proves that the stock is Malliavin differentiable and gives us a convenient formula for its derivative in terms of the operators $\lambda$ and $\sigma$:

**Proposition 2.6.14** *Under it holds that for all* $T \geq 0$*,* $P_T \in \mathbb{D}^{1,\infty}$*, and* $D_t P_T$ *is given by the formula:*

$$\begin{aligned} D_t P_T &= \mathbb{1}_{\{t \leq T\}}(t, T) P_T \Big[ \sqrt{v_t(0)} \lambda_t + \int_{u=t}^{T} \frac{1}{2\sqrt{v_u(0)}} \delta_0 Y_{t,u} \sigma_t dW_u^{(\lambda)} \\ &\quad + \int_{u=t}^{T} \sqrt{v_u(0)} \nabla \lambda_u Y_{t,u} \sigma_t dW_u - \frac{1}{2} \int_{u=t}^{T} \delta_0 Y_{t,u} \sigma_t du \Big] \end{aligned} \tag{2.6.6}$$

Proof of Lemma 2.6.13:

The $p$-power of the stock at time $T$ is given by:

$$P_T^p = P_0^p \exp\Big[p\int_{u=0}^T \sqrt{v_u(0)}dW_u^{(\lambda)} - \frac{p}{2}\int_{u=0}^T v_u(0)du\Big]$$

so that, using the Cauchy-Schwarz inequality and the same introduction of a super-martingale started at 1 as earlier for the finiteness of $\mathbb{E}v_T(x)^q$, we get:

$$\begin{aligned}\mathbb{E}P_T^p &\leq P_0^p\sqrt{\mathbb{E}\exp\Big[2p^2\int_{u=0}^T v_u(0)du\Big]}\\ &= P_0^p\sqrt{\mathbb{E}\exp\Big[2p^2\int_{u=0}^T (v_0(u) \ + \ 2p^2\int_{t=0}^u \delta_{u-t}\sigma_t dW_t)du\Big]}\end{aligned}$$

Using the fact that $\int_{u=0}^T v_0(u)du = V_0(T)$, and stochastic Fubini's theorem for the second term, justified by Assumption 2.6.12 and Theorem 4.18 of Da-Prato and Zabczyk [13], we get therefore that:

$$\mathbb{E}P_T^p \ \leq \ P_0^p\exp\{p^2V_0(T)\}\sqrt{\mathbb{E}\exp\Big[2p^2\int_{t=0}^T\int_{u=t}^T \delta_{u-t}\sigma_t du dW_t\Big]}$$

which, using again the super-martingale trick to get rid of the stochastic integral and Jensen's inequality to justify us putting the norm inside the integrals, is smaller than:

$$P_0^p\exp\left[p^2V_0(T)\right]\Big[\mathbb{E}\exp\Big((8p^4+4p^2)\int_{t=0}^T\int_{u=0}^{T-t}\|\delta_u\sigma_t\|_{\mathbb{G}}^2 du dt\Big)\Big]^{1/4}$$

Finally, to make things simpler, we can find an upper bound by letting $u$ go from 0 to $T$ in the inner integral instead of 0 to $T-t$, and we get as final word:

$$\mathbb{E}P_T^p \ \leq \ P_0^p\exp\left[p^2V_0(T)\right]\Big[\mathbb{E}\exp\Big((8p^4+4p^2)\int_{t=0}^T\int_{u=0}^{T}\|\delta_u\sigma_t\|_{\mathbb{G}}^2 du dt\Big)\Big]^{1/4}$$

and the right hand side is finite by Assumption 2.6.12. ■

Proof of Proposition 2.6.14: Because of Theorem B.3.2, and of Lemma 2.6.13, it is sufficient to show that $(\log P_T)/P_T$ is Malliavin differentiable and of finite $L^p(\Omega;\mathbb{U}^*)$-norm for all $p \geq 2$.

But we have that:

$$\begin{aligned}(\log P_T)/P_T &= \int_{u=0}^T \sqrt{v_u(0)}d^\lambda W_u - \frac{1}{2}\int_{u=0}^T v_u(0)du\\ &= \mathcal{A}_T(0) - \frac{1}{2}\mathcal{V}_T(0).\end{aligned}$$

It is now enough to recall Corollary 2.6.8 and Proposition 2.6.9 (where the quantities $\mathcal{A}_T(x)$ and $\mathcal{V}_T(x)$ are defined) to conclude. ■

### 2.6.3 Explicit hedging portfolio formula

We now state the second main result of this chapter, which gives the explicit hedging portfolio for contingent claims of the type (2.6.1):

**Theorem 2.6.15** ***Explicit characterization of the hedging portfolio for discretely monitored claims:***
*Let $(v_t)_{t\geq 0}$ and $(P_t)_{t\geq 0}$ be the forward variance swaps curve process and its associated stock price process, defined as in the previous sections. Let $\xi$ be a contingent claim of the form:*

$$\xi = g(P_{T_1}, ...P_{T_n}, v_{T_{n+1}}, ..., v_{T_{n+m}}) \tag{2.6.7}$$

*for some positive integers n and m, some function g measurable from $(\mathbb{R}_+^n \times \mathbb{F}_+^m, \mathcal{B}(\mathbb{R}_+^n \times \mathbb{F}_+^m))$ into $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and some positive times $T_1, ..., T_{n+m}$, the maximum of which we will denote by $T$.*
*Assume that g either satisfies:*

- *(i) the conditions of Proposition B.3.1.*
- *(ii) the conditions of Proposition B.3.2 and $\nabla g$ is a well defined $(\mathbb{R}^n \times \mathbb{F}^m)^*$-valued random variable with finite (positive) moments.*

*In either case, $\xi \in \mathbb{D}^{1,\infty}$ and there exists an $(\mathbb{R}^n \times \mathbb{F}^m)^*$-valued random variable $\nabla g$ with finite moments such that:*

$$D_t\xi = \nabla g D_t(P_{T_1}, ..., P_{T_n}, v_{T_{n+1}}, ..., v_{T_{n+m}})$$

*Then, the unique hedging portfolio for $\xi$ is given at any time $t \in [0,T]$ by:*

$$\phi_t^{(P)} = \sum_{j=1}^{n} \mathbb{1}_{\{t\leq T_j\}} \frac{1}{P_t} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},j)} P_{T_j}\Big], \text{ and} \tag{2.6.8}$$

$$\begin{aligned}\phi_t^{(v)} &= \sum_{j=1}^{n} \mathbb{1}_{\{t\leq T_j\}} \mathbb{E}_t\Big[P_{T_j}\nabla g\big|_{(\mathbb{R},j)} \int_{u=t}^{T_j} \Big(\frac{1}{2\sqrt{v_u(0)}}\delta_0 Y_{t,u}.\lambda_u + \sqrt{v_u(0)}\nabla\lambda_u Y_{t,u}.\Big)dW_u \\ &- P_{T_j}\nabla g\big|_{(\mathbb{R},j)} \frac{1}{2}\int_{u=t}^{T_j} \delta_0 Y_{t,u}.du\Big] + \sum_{j=1}^{m} \mathbb{1}_{\{t\leq T_{n+j}\}} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{F},j)} Y_{t,T_{n+j}}\Big].\end{aligned} \tag{2.6.9}$$

*where $(Y_{t,T})_{0\leq t\leq T<\infty}$ is the family of strong $\mathcal{L}(\mathbb{F})$-valued random operators defined previously by equation (2.6.3), and where the operators $\nabla g\big|_{(\mathbb{R},j)}$, $j \leq n$ and $\nabla g\big|_{(\mathbb{F},j)}$, $j \leq m$, valued in $\mathbb{R}$ and $\mathbb{F}^*$ are defined by:*

$$\nabla g =: \big(\nabla g\big|_{(\mathbb{R},1)}, ..., \nabla g\big|_{(\mathbb{R},n)}, \nabla g\big|_{(\mathbb{F},1)}, ..., \nabla g\big|_{(\mathbb{F},m)}\big)$$

**Remark 2.6.16** *Remark on the condition on g in Theorem 2.6.15: The reason for allowing g not to be Lipschitz is that many payoffs are not globally Lipschitz, some not even locally (i.e. at* 0*). For instance, many contracts are written on the realized volatility, the square root of realized variance, such as call options on realized volatility. We want our theorem to apply in these important cases too.*

Proof of Theorem 2.6.15:

In order to simplify the notation slightly, but without loss, we prove the theorem in the case where $n = m = 1$, and $\mathbb{E}(\xi) = 0$. By the Clark-Ocone formula for Malliavin differentiable Hilbert space valued random variables (see Theorem 4.1 of Carmona and Tehranchi [8]), we have that:

$$\xi = \mathbb{E}(\xi) + \int_{t=0}^{T} \mathbb{E}_t\Big[D_t g(P_{T_1}, v_{T_2})\Big] dW_t$$

thus:

$$\xi = \int_{t=0}^{T} \mathbb{E}_t\Big[\nabla g D_t(P_{T_1}, v_{T_2})\Big] dW_t \tag{2.6.10}$$

We now recall Proposition 2.6.14 and Proposition 2.6.5 which tell us what the Malliavin derivatives of $P_{T_1}$ and $v_{T_2}$ are. Plugging these expressions in equation (2.6.10), we get:

$$\begin{aligned}
\xi &= \int_{t=0}^{T} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)} D_t P_{T_1}\Big] dW_t + \int_{t=0}^{T} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{F},1)} D_t v_{T_2}\Big] dW_t \\
&= \int_{t=0}^{T_1} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)} P_{T_1}/P_t\Big] P_t \sqrt{v_t(0)}\lambda_t dW_t \\
&+ \int_{t=0}^{T_1} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)} \int_{u=t}^{T_1} \frac{1}{2\sqrt{v_u(0)}} \delta_0 Y_{t,u} dW_u\Big] \sigma_t dW_t && (2.6.11) \\
&+ \int_{t=0}^{T_1} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)} \int_{u=t}^{T_1} \sqrt{v_u(0)} \nabla\lambda_u Y_{t,u} dW_u\Big] \sigma_t dW_t && (2.6.12) \\
&- \int_{t=0}^{T_1} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)} \frac{1}{2} \int_{u=t}^{T_1} \delta_0 Y_{t,u} du\Big] \sigma_t dW_t && (2.6.13) \\
&+ \int_{t=0}^{T_2} \mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{F},1)} Y_{t,T_2}\Big] \sigma_t dW_t && (2.6.14) \\
&=: \int_{t=0}^{T_1} \phi_t^{(P)} dP_t + \int_{t=0}^{T_2} \phi_t^{(v)} \sigma_t dW_t
\end{aligned}$$

where we set, by definition, $\phi_t^{(P)}$ and $\phi_t^{(v)}$ as in the theorem. All that remains to prove is that $\phi^{(P)}$ and $\phi^{(v)}$ indeed define an admissible trading strategy. We start with $\phi^{(P)}$: several applications of Jensen's and Cauchy-Schwartz's inequalities, along with the uniform boundedness of the moments of $v_t$ on $[0, T_1]$ and the finiteness of all moments of $P_{T_1}$ justify the following calculation:

$$\begin{aligned}
\mathbb{E}\int_{t=0}^{T_1} (\phi_t^{(P)})^2 P_t^2 v_t(0) dt &= \mathbb{E}\int_{t=0}^{T_1} \Big[\mathbb{E}_t \nabla g\big|_{(\mathbb{R},1)} P_{T_1}/P_t\Big]^2 P_t^2 v_t(0) dt \\
&\leq \mathbb{E}\int_{t=0}^{T_1} \Big[\mathbb{E}_t \nabla g\big|_{(\mathbb{R},1)}^2 P_{T_1}^2/P_t^2\Big] P_t^2 v_t(0) dt \\
&\leq \sqrt{\mathbb{E}\int_{t=0}^{T_1} (\mathbb{E}_t \nabla g\big|_{(\mathbb{R},1)}^2 P_{T_1}^2)^2 dt}\ \sqrt{\mathbb{E}\int_{t=0}^{T_1} v_t(0)^2 dt} \\
&\leq \sqrt{\mathbb{E}\int_{t=0}^{T_1} \mathbb{E}_t \nabla g\big|_{(\mathbb{R},1)}^4 P_{T_1}^4 dt}\ \sqrt{\mathbb{E}\int_{t=0}^{T_1} v_t(0)^2 dt} \\
&\leq T_1 \Big(\mathbb{E}\big\|\nabla g\big|_{(\mathbb{R},1)}\big\|^8\Big)^{1/4} \Big(\mathbb{E} P_{T_1}^8\Big)^{1/4} \Big(\sup_{u\leq T_1} \mathbb{E} v_u(0)^2\Big)^{1/2} \ < \ \infty
\end{aligned}$$

As for $(\phi_t^{(v)})_{t\leq T_2}$, what we have to check is that it is a well defined adapted $\mathbb{F}^*$-valued stochastic process, and that $\mathbb{E}\int_{u=0}^{T_2} \|\phi_u^{(v)}\sigma_u\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} du < \infty$. The contribution to $\phi^{(v)}$ of term (2.6.14) is the easiest to deal with: if $f \in \mathbb{F}$ and $t \leq T_2$, then using the constant

$C_{T_2,4}$ defined in Theorem C.0.4, we have that:

$$\begin{aligned}\left|\mathbb{E}_t\left[\nabla g|_{(\mathbb{F},1)} Y_{t,T_2}\right] f\right| &\leq \sqrt{\mathbb{E}_t\left|\nabla g|_{(\mathbb{F},1)} Y_{t,T_2} f\right|^2} \\ &\leq \sqrt[4]{\mathbb{E}_t\left\|\nabla g|_{(\mathbb{F},1)}\right\|_{\mathbb{F}^*}^4}\sqrt[4]{\mathbb{E}_t\left\|Y_{t,T_2} f\right\|_{\mathbb{F}}^4} \\ &\leq M_t C_{T_2,4}^{1/4}\|f\|_{\mathbb{F}} < \infty\end{aligned}$$

where $M_t$ is the almost surely finite random variable $\sqrt[4]{\mathbb{E}_t\|\nabla g|_{(\mathbb{F},1)}\|^4}$. This shows that $\mathbb{E}_t\left[\nabla g\big|_{(\mathbb{F},1)} Y_{t,T_2}\right] \in \mathbb{F}^*$ a.s. We also have that:

$$\mathbb{E}\int_{u=0}^{T_2}\left\|\mathbb{E}_u\left[\nabla g\big|_{(\mathbb{F},1)} Y_{u,T_2}\right]\sigma_u\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}^2 du \leq M_0^2 K^2 C_{T_2,4}^{1/2}\sqrt{\mathbb{E}\int_{u=0}^{T_2}\|v_u\|_{\mathbb{F}}^4 du}$$

where $K$ is the Lipschitz bound on $\sigma$ and $M_0$ is defined as $M_t$ a few lines earlier, evaluated at $t = 0$. Finally, the uniform boundedness of the fourth moment of $v_u$ on $[0, T_2]$ by Theorem C.0.3 concludes. Similar arguments allow us to deal with the three other terms (2.6.11), (2.6.12) and (2.6.13), because the bounds $A_{T_1,p}$ and $C_{T_1,p}$ of Theorem C.0.3 and Theorem C.0.4 apply for all values of $p \geq 2$ and uniformly on $[0, T_1]$. Specifically, this implies, after some computations that we skip, that:

- $\left|\mathbb{E}_t\left[\nabla g\big|_{(\mathbb{R},1)}\int_{u=t}^{T_1}\sqrt{v_u(0)}\nabla\lambda_u Y_{t,u} dW_u\right] f\right|$
  $\leq L\|\delta_0\|^{1/2}\left(\mathbb{E}_t\|\nabla g\|^2\right)^{1/2}\left(\mathbb{E}_t\int_{u=t}^{T_1}\|v_u\|_{\mathbb{F}}^2 du\right)^{1/4} T_1^{1/4} C_{T_1,4}^{1/4}\|f\|_{\mathbb{F}}$
- $\left|\mathbb{E}_t\left[\nabla g\big|_{(\mathbb{R},1)}\frac{1}{2}\int_{u=t}^{T_1}\delta_0 Y_{t,u} du\right] f\right|$
  $\leq \frac{T_1}{2}\|\delta_0\|\left(\mathbb{E}_t\|\nabla g\|^2\right)^{1/2} C_{T_1,2}^{1/2}\|f\|_{\mathbb{F}}$
- $\left|\mathbb{E}_t\left[\nabla g\big|_{(\mathbb{R},1)}\int_{u=t}^{T_1}\frac{1}{2\sqrt{v_u(0)}}\delta_0 Y_{t,u} dW_u\right] f\right|$
  $\leq \frac{T_1^{1/4}}{2}\|\delta_0\|\left(\mathbb{E}_t\|\nabla g\|^2\right)^{1/2}\left(\mathbb{E}_t\int_{u=t}^{T_1} v_u(0)^{-2} du\right)^{1/4} C_{T_1,4}^{1/4}\|f\|_{\mathbb{F}}$

and

- $$\mathbb{E}\int_{t=0}^{T_1}\Big\|\mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)}\int_{u=t}^{T_1}\sqrt{v_u(0)}\nabla\lambda_u Y_{t,u}dW_u\Big]\sigma_t\Big\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}dt$$
$$\leq T_1^{3/2}B_{T_1,4}^{1/2}L^2\big\|\delta_0\big\|K^2C_{T_1,8}^{1/4}\Big(\mathbb{E}\big\|\nabla g\big\|^4\Big)^{1/2}\Big(\sup_{u\leq T}\mathbb{E}\big\|v_u\big\|^4\Big)^{1/4}\Big(\sup_{u\leq T}\mathbb{E}\big\|v_u\big\|^8\Big)^{1/4}<\infty$$
- $$\mathbb{E}\int_{t=0}^{T_1}\Big\|\mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)}\frac{1}{2}\int_{u=t}^{T_1}\delta_0 Y_{t,u}du\Big]\sigma_t\Big\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}dt$$
$$\leq\frac{T_1^3}{4}\big\|\delta_0\big\|^2\Big(\mathbb{E}\big\|\nabla g\big\|^4\Big)^{1/2}K^2C_{T_1,4}^{1/2}\Big(\sup_{u\leq T}\mathbb{E}\big\|v_u\big\|^4\Big)^{1/2}<\infty$$
- $$\mathbb{E}\int_{t=0}^{T_1}\Big\|\mathbb{E}_t\Big[\nabla g\big|_{(\mathbb{R},1)}\int_{u=t}^{T_1}\frac{1}{2\sqrt{v_u(0)}}\delta_0 Y_{t,u}dW_u\Big]\sigma_t\Big\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}dt$$
$$\leq\frac{1}{4}T_1^{3/2}B_{T_1,4}^{1/2}L^2\big\|\delta_0\big\|^2K^2C_{T_1,8}^{1/4}\Big(\mathbb{E}\big\|\nabla g\big\|^4\Big)^{1/2}\Big(\sup_{u\leq T}\mathbb{E}\frac{1}{v_u(0)^4}\Big)^{1/4}\Big(\sup_{u\leq T}\mathbb{E}\big\|v_u\big\|^8\Big)^{1/4}<\infty$$

where $L$ is a Lipschitz bound on $\lambda$ and the Burkholder constants $B_{T,p}$ are introduced in appendix A.3. ■

**Remark 2.6.17** *We have seen earlier that trading strategies in terms of the variance swaps and in terms of the forward variance swaps are equivalent, so it would seem a bit redundant to consider as well contingent claims depending on variance swaps, as we already have a formula to hedge contingent claims depending on the stock and the forward variance swaps. Here are two ways of looking at the problem of hedging a claim of the form $\xi_T = g(V_T)$:*

- *(1) We can start by approximating $\xi_T$ by a function of points on the curve $V_T$, i.e. $\xi_T \approx \tilde{g}(V_T(y_1), ...., V_T(y_n))$ for some points $y_1, ..., y_n$. We can therefore focus on the hedging of contingent claims of the form $\psi = g(V_T(y))$ for some $y \geq 0$. We have by definition that $V_T(y) = V_T(0) + \int_0^y v_T(u)du \approx \sum_{n=1}^N \left(\log\left(P_{T_{n+1}}/P_{T_n}\right)^2\right) + l(v_T)$ for some partition $\{T_n\}_{n=1}^{N-1}$ of $[0,T]$ and some functional $l \in \mathbb{F}^*$ so that $\xi_T$ can be rewritten in the form: $\xi_T \approx h(P_{T_1}, ..., P_{T_N}, v_T)$, which makes possible to use Theorem 2.6.15 directly.*

- *(2) Without doing any approximation, we can use the expression for $D_tV_T$ given* *in Proposition 2.6.10 in terms of the family of strong random operators $Z_{t,u}$. This leads to a formula similar to that given in (2.6.9).*

### 2.6.4 Analysis of the hedging portfolio and maturity-specific risk

Not surprisingly, we see that if the contingent claim is a function of the variance instruments only (i.e. $g(x_1, v) = g(x_2, v)$ for all $x_1$, $x_2$ and $v$), then the hedging portfolio does not involve trading in the stock (i.e. $\phi_t^{(P)} = 0$). This is of course because we are modelling the forward variance swaps curve as an autonomous process (the equation governing the evolution of $v$ does not involve the stock price). This feature would of course have no reason to be true if we had taken $\sigma$ to be also a function of the stock $P$. We now turn to the more interesting question of maturity-specific risk. We have seen earlier that in finite dimensional models, one can hedge options with a certain (and finite) number of arbitrarily chosen variance swaps. For instance, the Heston model would allow us to hedge a look-back option, paying the maximum of the stock between now and a year in the future, by trading in the stock and a variance swap maturing in 10 years! We also pointed out earlier that we could make this counterintuitive feature disappear by using infinite dimensional models. We now make this claim precise: under some conditions on $\sigma$ and $\lambda$, the hedging portfolio $\phi_t^{(v)}$ satisfies the same "maturity-specific" feature as the one proved by Carmona and Tehranchi for interest rates infinite dimensional models [7]. The following theorem is the third main result of this chapter:

**Theorem 2.6.18** ***Maturity-specific property of hedging portfolios:***
*Let $\xi$ be a contingent claim of the form $\xi = g(P_{T_1}, v_{T_2})$ satisfying the conditions of Theorem 2.6.15. In addition, we assume that this contingent claim depends only on instruments maturing on or before $T$, that is, it holds that $T_1 \leq T$, $T_2 \leq T$ and:*

$$\Big\{(v_1(y) = v_2(y) \;\; \forall y \leq T - T_2)\Big\}$$

$$\Rightarrow$$

$$\Big\{g(x, v_1) = g(x, v_2) \;\; \forall x > 0\Big\}$$

*We also assume that the "volatility" of the forward variance swaps curve $\sigma$ satisfies, in addition to all the previous conditions, the following "no-maturity-mixing" condition:*

***Assumption 2.6.19*** *For all $x \geq 0$, it holds that:*

$$\left\{v_1(y) = v_2(y)\ \forall\ y\ \leq\ x\right\}$$
$$\Rightarrow$$
$$\left\{\delta_x\sigma(t; v_1) = \delta_x\sigma(t; v_2)\ \ \forall t \geq 0\right\}$$

*Finally, we assume that the stock is correlated to the forward variance swaps curve only through the short variance, that is, we make the following assumption:*

***Assumption 2.6.20*** *$\lambda$ is of the form:*

$$\lambda_u = \lambda(u, v_u(0))$$

*Then the holding $\phi_t^{(v)}$ in the forward variance swaps curve (recall that the unique hedging portfolio $\phi_t := (\phi_t^{(P)}, \phi_t^{(v)})$ is given by the formula of Theorem 2.6.15) has the following property:*

$$support(\phi_t^{(v)}) \subseteq [0, T - t],\ \ \mathbb{P} \times Leb\ \ a.s.$$

Proof of Theorem 2.6.18:

For any $x \geq 0$, we denote by $\mathbb{F}_x$ the subset of $\mathbb{F}$ of functions null on $[0, x]$. The continuity of the evaluation functionals on $\mathbb{F}$ implies that $\mathbb{F}_x$ is a closed subspace of $\mathbb{F}$. We can identify (by the Riesz representation lemma) the orthogonal complement of $\mathbb{F}_x$ with the subset of $\mathbb{F}^*$ of bounded linear functional on $\mathbb{F}$ which have support included in $[0, x]$. We will denote this orthogonal complement by $\mathbb{F}_x^{\perp}$. What we want to prove is that for any $t \leq T_0 := \max\{T_1, T_2\}$, we have $\phi_t^{(v)} \in \mathbb{F}_{T-t}^{\perp}$. It was proved by Carmona and Tehranchi (see Theorem 6.6 in [8]) that the family of strong random operators $(Y_{t,u})_{0\le t\le u}$ satisfies the following property:

$$\begin{aligned} \forall (x,t,u) \le 0,\ t \le u,\ \forall f \in \mathbb{F}_{x+u-t}, \\ Y_{t,u} f \in \mathbb{F}_x, \end{aligned} \tag{2.6.15}$$

We do not repeat the proof here, but we just mention that this is proved by doing an induction reasoning on the Picard iterations $\{Y_{t,u}^{(n)}\}_{n=1}^{\infty}$ that we introduce in Section C.2. Clearly this property means that for any $u \in [t, T_0]$, $Y_{t,u} f \in \mathbb{F}_x$ whenever $f \in \mathbb{F}_{x+T-t}$, because $T - t \ge T_0 - u$. Therefore, for any $f \in \mathbb{F}_{x+T-t}$, it will hold that:

$$\begin{aligned} &\int_{u=t}^{T_0} \frac{1}{2\sqrt{v_u(0)}} \delta_0 Y_{t,u} f dW_u^{(\lambda)} = 0 \\ &\int_{u=t}^{T_0} \sqrt{v_u(0)} \nabla \lambda_u Y_{t,u} f dW_u = 0 \\ &\int_{u=t}^{T_0} \delta_0 Y_{t,u} f du = 0 \end{aligned}$$

where the second assertion comes from the fact that $\nabla \lambda_u$ can be written as $\alpha_u \delta_0$, for some scalar process $\alpha$, by the assumption that $\lambda_u$ depends on $v_u$ only via the short variance $v_u(0)$. But the three terms above are exactly the ones involved in the expressions (2.6.11), (2.6.12) and (2.6.13) of $\phi_t^{(v)}$, so that $\phi_t^{(v)} f = 0$ whenever $f \in \mathbb{F}_{x+T-t}$ and this proves the theorem. ■

## 2.7 Concrete examples

### 2.7.1 Examples of state spaces and infinite dimensional model

We now give an example of a model and of state spaces that satisfy all of the assumptions listed previously. That is, $\mathbb{F}$ satisfies Assumption 2.4.1, Assumption 2.4.2, Assumption 2.4.3, $\sigma$ satisfies Assumption 2.4.7, Assumption 2.4.8, Assumption 2.4.9, Assumption 2.5.3, Assumption 2.6.12 and Assumption 2.6.19, and $\lambda$ satisfies Assumption 2.5.4, Assumption 2.6.2 and Assumption 2.6.20.

We take for $\tilde{\mathbb{F}}$ the weighted Sobolev spaces of continuous functions introduced by Filipovic (see [18] for instance): let $w$ be a continuous, positive, increasing function such that:

$$\int_{u=0}^{\infty} w(u)^{-1} du < \infty$$

(in particular $w$ diverges), and let us define $\mathbb{F}_w$ as the set of all absolutely continuous functions $f$ defined on $\mathbb{R}_+$, whose weak derivative will be denoted by $f'$ and which satisfy:

$$\int_{u=0}^{\infty} |f'(u)|^2 w(u) du < \infty$$

endowed with the product

$$< f, g >_{\mathbb{F}_w} := f(0)g(0) + \int_{u=0}^{\infty} f'(u)g'(u)w(u)du$$

To lighten the notation, we will drop from now on the subscript in $\mathbb{F}_w$ and denote this space, as in the previous subsections, simply as $\mathbb{F}$. We have then the following proposition:

**Proposition 2.7.1** *$\mathbb{F}$ satisfies Assumption 2.4.1, Assumption 2.4.2 and Assumption 2.4.3.*

<u>Proof of Proposition 2.7.1</u>:

By definition, elements of $\mathbb{F}$ are continuous functions. It is easily seen that, once endowed with the products $< ., . >_{\mathbb{F}}$, $\mathbb{F}$ turns into separable Hilbert space. It is well known that the left shift operators form a strongly continuous semigroup of contractions on this space (see Filipovic [18]). We now prove the uniform boundedness of the evaluation operators: for $x \geq 0$ and $f \in \mathbb{F}$ given, we have that:

$$
\begin{aligned}
f(x)^2 &\leq 2f(0)^2 + 2\Big(\int_{u=0}^{\infty} \big|f'(u)\big| du\Big)^2 \\
&\leq 2\Big[f(0)^2 + \int_{u=0}^{\infty} f'(u)^2 w(u) du \int_{u=0}^{\infty} w(u)^{-1} du\Big]
\end{aligned}
$$

and thus:

$$
\big|\delta_x f\big| \leq K \big\|f\big\|_{\mathbb{F}}
$$

where $K^2 := 2\max(1, \int_{u=0}^{\infty} w(u)^{-1} du)$.

Notice that because the boundedness in the above is uniform in $x$, this implies that elements of $\mathbb{F}$ are bounded functions and we will use for any element $f \in \mathbb{F}$ the usual notation $\|f\|_\infty$ to denote the finite supremum of $\{|f(x)|, x \geq 0\}$. Notice actually that $f(\infty) := \lim_{x\to\infty} f(x)$ is well defined for all $f \in \mathbb{F}$, as it holds that $f(\infty) = f(0) + \int_0^{\infty} f'(u)du$. Moreover, the same reasoning as above also implies that $\delta_\infty : f \in \mathbb{F} \mapsto \lim_{x\to\infty} f(x) \in \mathbb{R}$ is a linear bounded operator with norm smaller than the $K$ defined above. ■

As in Subsection 2.4.2, we can define $\tilde{\mathbb{F}}$ as the set of all continuously differentiable functions $\tilde{f}$ defined on $\mathbb{R}_+$, whose derivative $\tilde{f}'$ is in $\mathbb{F}$. $\tilde{\mathbb{F}}$ is then endowed with the product:

$$
< \tilde{f}, \tilde{g} >_{\tilde{\mathbb{F}}_w} := \tilde{f}(0)\tilde{g}(0) + < \tilde{f}', \tilde{g}' >_{\mathbb{F}}
$$

We now verify as promised earlier that the properties of $\tilde{\mathbb{F}}$ listed in Subsection 2.4.2 hold:

*Proof that Property 2.4.5 holds*:

For any $f \in \mathbb{F}$, we have:

$$
\begin{aligned}
\big\|Jf\big\|_{\tilde{\mathbb{F}}}^2 &= (Jf)(0)^2 + (Jf)'(0)^2 + \int_{u=0}^{\infty} (Jf)''(u)^2 w(u) du \\
&= 0 + f(0)^2 + \int_{u=0}^{\infty} f'(u)^2 w(u) du = \big\|f\big\|_{\mathbb{F}}^2
\end{aligned}
$$

That is, $J$ is an isometric bijection from $\mathbb{F}$ to $\tilde{\mathbb{F}}_0$.

*Proof that Property 2.4.6 holds*:
For all $\tilde{f} \in \tilde{\mathbb{F}}$:

$$\left\|D\tilde{f}\right\|_{\mathbb{F}}^2 = \tilde{f}'(0)^2 + \int_{u=0}^{\infty} \tilde{f}''(u)^2 w(u) du \leq \left\|\tilde{f}\right\|_{\tilde{\mathbb{F}}}^2$$

■

Although this was not a requirement on the space $\mathbb{F}$, we will find it useful to know that our example of $\mathbb{F}$ is stable by multiplication and that there exists a bound on the ratio $\frac{||fg||_{\mathbb{F}}}{||f||_{\mathbb{F}}||g||_{\mathbb{F}}}$, uniformly on $(f, g) \in \mathbb{F}^2$.
*Proof*: let $(f, g) \in \mathbb{F}^2$. Let $\epsilon > 0$ be given. $f$ and $g$ being absolutely continuous, we can find $\delta > 0$ such that:

$$\sum_{n=1}^{\infty} |f(b_n) - f(a_n)| < \epsilon/(2||g||_{\infty})$$

and

$$\sum_{n=1}^{\infty} |g(b_n) - g(a_n)| < \epsilon/(2||f||_{\infty})$$

whenever the series of intervals $\{[a_n, b_n]\}_{n=1}^{\infty}$ satisfies $\sum |b_n - a_n| < \delta$. Thus for any such family of intervals, we have that:

$$\begin{aligned}\sum |fg(b_n) - fg(a_n)| &\leq \sum |f(b_n)||g(b_n) - g(a_n)| + \sum |g(a_n)||f(b_n) - f(a_n)| \\ &\leq ||f||_{\infty} \sum |g(b_n) - g(a_n)| + ||g||_{\infty} \sum |f(a_n) - f(b_n)| \leq \epsilon\end{aligned}$$

This proves that $fg$ is absolutely continuous, and in this case, we know that the weak derivative of $fg$ is given by the usual $gf' + fg'$. Finally, we can compute the square norm of the product $||fg||_{\mathbb{F}}^2$:

$$\begin{aligned}||fg||_{\mathbb{F}}^2 &= f(0)^2 g(0)^2 + \int_{u=0}^{\infty} (fg' + gf')(u)^2 w(u) du \\ &\leq ||f||_{\mathbb{F}}||g||_{\mathbb{F}} + 2||f||_{\infty}||g||_{\mathbb{F}} + 2||g||_{\infty}||f||_{\mathbb{F}} \\ &\leq (1 + 4K)||f||_{\mathbb{F}}||g||_{\mathbb{F}}\end{aligned}$$

where $K$ is as before (upper bound on $\|f\|_\infty/\|f\|_{\mathbb{F}}$). This concludes the proof. ■

We now give a concrete example of operators $\sigma_t$ and $\lambda_t$ which work nicely with the above example of state space $\mathbb{F}$. We recall that $\mathbb{G}$ is endowed with a complete orthonormal system which we denote by $\{g_n\}_{n=1}^{\infty}$. We need first however to make an assumption on the infinite-end of the initial volatility curve $v_0(\infty) := \lim_{x\to\infty} v_0(x)$ which will guarantee that our example of operator $\sigma$ is a.s. dense range:

**Assumption 2.7.2** $v_0(\infty) > 0$.

**Proposition 2.7.3** *Let $B$ be an arbitrary dense-range Hilbert-Schmidt operator from $\mathbb{G}$ to $\mathbb{F}$, such that $g_1 \in \ker B$, and let $l$ be an arbitrary positive function of $C^1(\mathbb{R}_+)$, bounded and with bounded derivative, started at 0, and such that $(x \mapsto l(f(x))) \in \mathbb{F}_+$ for all $f \in \mathbb{F}_+$. We then define the operator $\sigma := \sigma_{l,B}$ by:*

$$\begin{aligned}\sigma_{l,B} : &\mathbb{F}_+ \to \mathcal{L}_{HS}(\mathbb{G},\mathbb{F})\\ &f \mapsto \left[g \mapsto \left(x \mapsto l(f(x)).(Bg)(x)\right)\right]\end{aligned}$$

*Then $\sigma$ as defined above is a measurable map from $(\mathbb{F}_+, \mathcal{B}(\mathbb{F}_+))$ into $\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})$ and satisfies Assumption 2.4.7, Assumption 2.4.8, Assumption 2.4.9, Assumption 2.6.12 and Assumption 2.6.19. Under Assumption 2.7.2, $\sigma$ also satisfies Assumption 2.5.3.*

**Remark 2.7.4** *Remark that in Proposition 2.7.3, $l(f)$ is an element of $\mathbb{F}$ by hypothesis (on $l$), and that $l(f).Bg$ is an element of $\mathbb{F}$ because of the stability of $\mathbb{F}$ by multiplication, which we showed earlier.*

**Proposition 2.7.5** *Let $\lambda_t$ be a constant unit vector $\lambda$ of $\mathbb{G}^*$, such that $\lambda g_1$ is non zero (i.e. the first component of $\lambda^*$, the Riesz identification of $\lambda$ in $\mathbb{G}$, is non zero). This choice of $\lambda_t$ satisfies Assumption 2.5.4, Assumption 2.6.2 and Assumption 2.6.20.*

**Remark 2.7.6** *To be even more concrete in our example, notice that we could take $B = \sum_{j=1}^{\infty} b_j f_j \otimes g_{j+1}$ for some $b \in l_2$, with $b_j \neq 0$ for all $j$, and $l(.) = arctan(.)$.*

Proof of Proposition 2.7.3:

Let us first of all show that $\sigma(f) \in \mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$ for all $f$:

$$\sum_{n=1}^{\infty} \left\|\sigma g_n\right\|_{\mathbb{F}}^2 = \sum_{n=1}^{\infty} \left\|l(f)Bg_n\right\|_{\mathbb{F}}^2 \leq \sum_{n=1}^{\infty} K^2\left\|l(f)\right\|_{\mathbb{F}}^2\left\|Bg_n\right\|_{\mathbb{F}}^2 \leq K^2\left\|l(f)\right\|_{\mathbb{F}}^2\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}^2 < \infty$$

*Proof that Assumption 2.4.8 holds*: It is obvious from the above inequality, and the fact that:

$$\left\|l(f)\right\|_{\mathbb{F}} \leq \left\|l'\right\|_{\infty}\left\|f\right\|_{\mathbb{F}}.$$

*Proof that Assumption 2.4.7 holds*: A similar computation to the one just above, simply replacing $f$ by $f - g$ gives us

$$\left\|\sigma(f) - \sigma(g)\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} \leq K\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}\left\|l(f-g)\right\|_{\mathbb{F}} \leq K\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}\left\|l'\right\|_{\infty}\left\|f-g\right\|_{\mathbb{F}}$$

which shows that $\sigma$ is Lipschitz.

*Proof that Assumption 2.4.9 holds*: For $x \geq 0$, $f \in \mathbb{F}_+$, we have:

$$\begin{aligned}
\left\|\delta_x\sigma(f)\right\|_{\mathbb{G}^*} &= \sup_{g\in\mathbb{G},\|g\|=1} l(f)(x)\left|Bg(x)\right| \\
&\leq \sup_{g\in\mathbb{G},\|g\|=1} l(f)(x)\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}\left\|\delta_x\right\|_{\mathbb{F}}\left\|g\right\|_{\mathbb{G}} \\
&\leq K\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}\left\|l'\right\|_{\infty}f(x)
\end{aligned}$$

where $K$ is a uniform bound on $\|\delta_x\|_{\mathbb{F}}$, $x \geq 0$, and where the last inequality follows from the facts that $l(0) = 0$ and that $l'$ is bounded, by hypothesis.

*Proof that Assumption 2.6.12 holds*: for all $(t, u) \leq T$:

$$\left\|\delta_t\sigma(v_u)\right\|_{\mathbb{G}^*} \leq \sup_{g\in\mathbb{G},\|g\|=1} l(v_u)(t)(Bg)(t) \leq \left\|l\right\|_{\infty}\left\|B\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}K$$

This implies that:

$$\mathbb{E}\exp\Big(p\int_{t=0}^{T}\int_{u=0}^{T}\big\|\delta_t\sigma_u\big\|_{\mathbb{G}^*}dudt\Big) < \exp\Big(pT^2\big\|l\big\|_\infty\big\|B\big\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}K\Big) < \infty$$

*Proof that Assumption 2.6.19 holds*: the "no-maturity-mixing" condition is clear, because $\sigma(f)$ makes use of the curve $f$ simply through pointwise multiplication ($f$ being here of course the $\mathbb{F}$-valued argument).

*Proof that Assumption 2.5.3 holds*: Finally we prove that for almost all $\omega$, for all $t \geq 0$, $\sigma(v_t(\omega))$ is dense-range in $\mathbb{F}$: let $h \in \mathbb{F}$ and $\epsilon > 0$ be given. We want to find $g \in \mathbb{G}$ such that $\|\sigma(v_t)g - h\|_{\mathbb{F}} < \epsilon$.
Let us first remark that almost surely, $v_t(\infty) := \lim_{x\to\infty} v_t(x)$ exists and is strictly positive, which along with the continuity and positivity of $v_t$ means that almost surely, for all $t \geq 0$, $v_t$ is bounded away from 0 (as a function of $x$). Indeed, from the stochastic equation satisfied by $v_t$ and the fact that the operator $\delta_\infty$ is a bounded linear functional on $\mathbb{F}$, we can see that $v_t(\infty)$ is a well defined a.s. strictly positive continuous martingale started at $v_0(\infty) > 0$ (thanks to Assumption 2.7.2), and therefore a.s. for all $t \geq 0$, $v_t(\infty) > 0$. This implies that for almost all $(t,\omega)$, $h/l(v_t(\omega)) \in \mathbb{F}$. Indeed, $(h/l(v))' = h'/l(v) + hv'l'(v)/l(v)^2 < M(h' + v')$, where $M$ is a suitable bound, whose existence is easy to prove as $v$ is bounded from below, and therefore $l(v)$ is as well, and as $l'$ is bounded as well. Therefore, we have that:

$$\big\|l(v_t(\omega))Bg - h\big\|_{\mathbb{F}} \leq \big\|l(v_t(\omega))\big\|_{\mathbb{F}}\big\|Bg - h/l(v_t(\omega))\big\|_{\mathbb{F}}$$

By the dense-range property of $B$, we can find $g \in \mathbb{G}$ such that:

$$\big\|Bg - h/l(v_t(\omega))\big\|_{\mathbb{F}} < \epsilon/\big\|l(v_t(\omega))\big\|_{\mathbb{F}}$$

and this completes the proof. ■
We can notice that without the assumption that $v_0(\infty) > 0$, i.e. if $v_0(\infty) = 0$, then $v_t(\infty) = 0$ a.s. and for all elements $h$ of $\mathbb{F}$ such that $h(\infty) \neq 0$, $h/l(v_t)$ would not be an

element of $\mathbb{F}$ (as it would diverge), and therefore, $\sigma$ could not be dense-range.

Proof of Proposition 2.7.5: That Assumption 2.6.2 and Assumption 2.6.20 hold is clear. The validity of Assumption 2.5.4 comes from the facts that $< \lambda^*, g_1 >_{\mathbb{G}^*} \neq 0$ while $g_1 \in \ker \sigma_t$. $\blacksquare$

### 2.7.2 Black & Scholes and Heston models as forward variance swaps curves models

It is interesting to remark that we can recast the usual models in terms of forward variance swaps models. This is the case of the Black and Scholes model (although it leads to a trivial formulation), and of common stochastic volatility models, like the multi-dimensional model we have presented earlier.

The Black and Scholes model corresponds naturally to the case where the operator $\sigma$ is equal to zero, and the initial curve $v_0$ is a constant function equal to the constant short variance (in the terminology we have used so far). This choice of $\sigma$ gives rise to the trivial solution $V_t(x) = (t+x)v_0(0)$, and the stock is given by the usual geometric Brownian motion: $dP_t = P_t\sqrt{v_0(0)}dW_t$.

The case of the multi-dimensional Heston model is more interesting so we look at it in more details: differentiating equation (2.3.2) with respect to $x$ leads to:

$$v_t(x) = \sum_{n=1}^{N}\Theta^{(n)} + \sum_{n=1}^{N}(B_t^{(n)} - \Theta^{(n)})\exp(-k^{(n)}x) =: \theta + \sum_{n=1}^{N}(B_t^{(n)} - \Theta^{(n)})\exp(-k^{(n)}x)$$

which, using Ito's formula and simplifying slightly, gives:

$$v_t(x) = v_0(x+t) + \delta_x \int_{u=0}^{t} S_{t-u}\Big(\sum_{n=1}^{N}\exp\big(-k^{(n)}(.)\big)\sigma^{(n)}\sqrt{B_u^{(n)}} \otimes g_n\Big)dW_u$$

This means that $v_t$ lives in $\mathbb{F}_N$, the $N+1$-dimensional subspace of $\mathbb{F}$ spanned by the functions $x \mapsto \exp\big(-k^{(n)}x\big)$, $n \le N$ and the constant function $x \mapsto 1$. We can finally invert equation (2.3.2) to express each of the $B_u^{(n)}$ in terms of $v_u$, say as: $B_u^{(n)} = f^{(n)}(v_u)$. This shows that the multi-dimensional Heston model corresponds to the following choice of $\sigma$:

$$\sigma_{MultiHeston}(t;v) = \sum_{n=1}^{N}\exp\big(-k^{(n)}(.)\big)\sigma^{(n)}\sqrt{f^{(n)}(v)} \otimes g_n$$

As we have already seen, there is no unique way of inverting the dependency of $B_t$ in terms of the curve $v_t$, and in particular, any arbitrarily chosen $N$ points of that curve are enough to recover $B_t := \{B^{(n)}\}_{n=1}^{N}$. This means that, seen as a (forward) variance swaps curve model, the multi-dimensional Heston model satisfies Assumption 2.6.19. Indeed,

$x > 0$ being given, if two curves $v_1$ and $v_2$ of $\mathbb{F}_N$ are identical up to $x$, then they are identical on $[0, \infty)$ and thus $\delta_x\sigma(v_1) = \delta_x\sigma(v_2)$. However, if we insist that only $N$ given, and fixed, forward variance swaps $v(t, (T_n), n \leq N$ be traded, and we express the model's dynamic as:

$$dv(t, T_n) = \sigma_n(t, v(t, T_1), ..., v(t, T_N))dW_t$$

then it does not satisfy the finite dimensional equivalent of Assumption 2.6.19, which would require that $\sigma_n$ depend only on $v(t, T_1), ..., v(t, T_n)$, but not on $v(t, T_{n+1}), ...v(t, T_N)$. This cannot be the case as the matrix in equation (2.3.2) giving the dependency of $\{v_t\}_{n=1}^N$ in terms of $B_t$ is not triangular, so its inverse cannot be either.

### 2.7.3 Examples of payoffs

We now give some concrete examples of (classical) payoffs for which Theorem 2.6.15 holds.

#### Options on the stock alone

The following contingent claims satisfy the assumption of the theorem:

- Polynomial options: $\xi = g(P_{T_1}, ..., P_{T_n})$ where $g$ is a polynomial in $\mathbb{R}^n$, of any degree. We can notice that $\nabla g\big|_{\mathbb{R},j}$ is also a polynomial for each $j \leq n$, and therefore by the boundedness of all moments of the stock, $g$ satisfies hypothesis (ii) of the theorem.

  In particular, it is interesting to look at the case $\xi = P_T$, just to do a sanity check on the formula of Theorem 2.6.15. By uniqueness, we should have in that case that $\phi_t^{(P)} = 1$ and $\phi_t^{(v)} = 0$. It is obvious by inspection of the formula that $\phi_t^{(P)} = 1$. However, the second assertion is far from obvious, but the following calculation shows that this is indeed the case:

$$\begin{aligned}\phi_t^{(v)} &= \mathbb{E}_t\Big[\frac{P_T}{2}\delta_0\Big(\int_{u=t}^{T}\frac{1}{\sqrt{v_u(0)}}Y_{t,u}dW_u^{(\lambda)} - \int_{u=t}^{T}\sqrt{v_u(0)}Y_{t,u}du\Big)\\ &+ P_T\int_{u=t}^{T}\sqrt{v_u(0)}\nabla\lambda_u Y_{t,u}dW_u\Big]\\ &= \mathbb{E}_t\Big[\frac{P_T}{2}\delta_0\int_{u=t}^{T}\frac{1}{\sqrt{v_u(0)}}Y_{t,u}dW_u^{(\lambda,P)}\\ &+ \Big(P_t + \int_{u=t}^{T}P_u\sqrt{v_u(0)}\lambda_u dW_u\Big)\int_{u=t}^{T}\sqrt{v_u(0)}\nabla\lambda_u Y_{t,u}dW_u\Big]\\ &= P_t\mathbb{E}_t^{\mathbb{Q}}\Big[\frac{1}{2}\delta_0\int_{u=t}^{T}\frac{1}{\sqrt{v_u(0)}}Y_{t,u}dW_u^{(\lambda,P)}\Big] + P_t\mathbb{E}_t\Big[\int_{u=t}^{T}\sqrt{v_u(0)}\nabla\lambda_u Y_{t,u}dW_u\Big]\\ &+ \mathbb{E}_t\Big[\int_{u=t}^{T}v_u(0)P_u\nabla\lambda_u^*\lambda_u Y_{t,u}d_u\Big] = 0\end{aligned}$$

  where $\mathbb{Q}$ is the probability measure equivalent to $\mathbb{P}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_u} = \frac{P_u}{P_0}$ and where $W_t^{(\lambda,P)} := W_t^{(\lambda)} - \int_{u=0}^{t}\sqrt{v_u(0)}du$ is a $\mathbb{Q}$-Brownian motion. Notice that the last term on the last line in the above computation vanishes because $\lambda$ is of norm 1, and therefore $\nabla\lambda_u^*\lambda_u = 0$.

- Call options: $\xi = g(P_T) = (P_T - K)_+$. $g$ is Lipschitz of coefficient 1 and therefore satisfies assumption (i) of the theorem.

- Forward start options: $\xi = (P_{T_2} - P_{T_1})_+$, for $0 \leq T_1 \leq T_2$.

- Spread options: let $T \geq 0$ be given, and $\{T_j\}_{j=1}^n$ be some partition of $[0, T]$. We then define the spread option $\xi = g(P_{T_1}, ..., P_{T_n}) = \max_{j\in\{1,...n\}} P_{T_j} - \min_{j\in\{1,...n\}} P_{T_j}$. That $g$ is Lipschitz can be seen from the following observation: Let us define $h$ by $h(P_{T_1}, ..., P_{T_n}) := \max_{j\in\{1,...n\}} P_{T_j}$. let $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ be two vectors in the positive hortant of $\mathbb{R}^n$. Then $|h(x) - h(y)| = |x_k - y_l|$ for some $k, l \in \{1, ..., n\}$. Let us suppose without loss of generality that $x_k \geq y_l$ so that $|h(x) - h(y)| = x_k - y_l$. Now, we clearly have that $y_l \geq y_k$, so that $|h(x) - h(y)|$ is smaller than $x_k - y_k$ and thus in any case, we have that $|h(x) - h(y)| \leq \sum_{j=1}^n |x_j - y_j|$. We can deal with the min part of $g$ in the same way, so that $g$ is Lipschitz as sum of two Lipschitz functions.

**Options on the variance instruments alone**

The following contingent claims, depending only on the variance swaps and forward variance swaps curves, satisfy the assumptions of the theorem:

- Powers of variance instruments, such as $\xi = [v_T(x)]^q$ or $\xi = [V_T(x)]^q$ for any real $q$.

- Call options on volatility or variance swaps, of the form: $\xi = (V_T(x) - K)_+$ or $\xi = (\sqrt{V_T(x)} - K)_+$

- Forward start on variance or volatility swaps, of the form $\xi = (\frac{V_{T_2}(x_2)}{T_2+x_2} - \frac{V_{T_1}(x_1)}{T_1+x_1})_+$ or $\xi = (\sqrt{\frac{V_{T_2}(x_2)}{T_2+x_2}} - \sqrt{\frac{V_{T_1}(x_1)}{T_1+x_1}})_+$

They all satisfy hypothesis (ii) of the theorem by the finiteness of all moments (positive and negative) of the variance swaps and the forward variance swaps.

**Options depending on both the variance instruments and the stock**

Finally, it is possible to design payoffs depending on both the stock and the variance instruments, and which will satisfy (ii) in the theorem. An example could be a:

- Weighted call option on a volatility swap: $\xi = (\sqrt{V_T(x)} - K)_+ h(P_{T_1}, ..., P_{T_n})$, where $\{T_j\}_{j=1}^n$ and $T$ are arbitrary times and $h$ is a polynomially bounded function.

# Appendix A

# A short introduction to Hilbert space valued random variables and infinite dimensional stochastic analysis

For a complete exposition of stochastic analysis in infinite dimensional spaces (i.e. Banach and Hilbert spaces), the reader is invited to refer to Da-Prato and Zabczyk [13]. The (older) article by Yor [48] gives also a good overview of Brownian motions and diffusions in Hilbert spaces. We give here the shortest of introductions to stochastic analysis in Hilbert spaces in order to make the main text self-contained.

## A.1 Gaussian measures in Hilbert spaces

Let $\mathbb{H}$ be a separable (i.e. having a countable dense subset, or equivalently a countable basis) Hilbert space with an orthonormal basis $\{h_n\}_{n=1}^{\infty}$. We can endow this space with its Borel $\sigma$-algebra $\mathcal{B}(\mathbb{H})$, turning $(\mathbb{H}, \mathcal{B}(\mathbb{H}))$ into a measurable space. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Any map $X$, from $\Omega$ into $\mathbb{H}$, which is $\mathcal{B}(\mathbb{H})/\mathcal{F}$-measurable, is called an $\mathbb{H}$-valued random variable. We can remark that $\mathcal{L}_X$, the law of $X$, defined as a map from $\mathcal{B}(\mathbb{H})$ into $\mathbb{R}$ by the usual $\mathcal{L}_X(B) := \mathbb{P}(X^{-1}(B))$ for $B \in \mathcal{B}(\mathbb{H})$ is a probability measure on $(\mathbb{H}, \mathcal{B}(\mathbb{H}))$. Very often, we can ignore the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and equivalently look at $(\mathbb{H}, \mathcal{B}(\mathbb{H}), \mathcal{L}_X)$ directly as the probability space of interest.

By extension to the Euclidean case, we can define Gaussian measures on the measurable space $(\mathbb{H}, \mathcal{B}(\mathbb{H}))$ by the following (notice that we cannot define the measure by its density as there is no equivalent of the Lebesgue measure in infinite dimensional spaces!):

**Definition A.1.1** *A measure $\mu$ on $(\mathbb{H}, \mathcal{B}(\mathbb{H}))$ is said to be Gaussian if for all $h \in \mathbb{H}$, the map*

$$x \mapsto < h, x >_{\mathbb{H}} : \mathbb{H} \to \mathbb{R}$$

*is a (real-valued) Gaussian random variable, i.e. if there exists $m_h \in \mathbb{R}$ and $q_h \in \mathbb{R}_+$ such that:*

$$\forall y \in \mathbb{R},\ \mu\{x \in \mathbb{H}, < h, x >_{\mathbb{H}} \leq\ y\} \ = \ \int_{u=-\infty}^{y} \frac{\exp \frac{-(u-m_h)^2}{2q_h^2}}{\sqrt{2\pi q_h^2}} du$$

It is not too difficult to prove that the map $h \mapsto m_h$ is linear and bounded and therefore, by the Riesz representation lemma, there exists an element $m \in \mathbb{H}$ such that

$$m_h =< m, h >_{\mathbb{H}} \ \ \forall h \in \mathbb{H}.$$

$m$ is called the mean of the Gaussian measure $\mu$. Similarly, one can show the existence of a unique symmetric, non negative and trace class operator $Q$ such that:

$$< Qh, h >_{\mathbb{H}}= q_h^2 \ \ \forall h \in \mathbb{H}.$$

$Q$ is called the covariance operator of the Gaussian measure $\mu$ and it can be easily observed that it is consistent with the usual covariance matrix definition in the finite dimensional case. There exists thus a basis in $\mathbb{H}$, which we still denote by $\{h_n\}_{n=1}^{\infty}$ and a sequence $\lambda = \{\lambda_n\}_{n=1}^{\infty}$ of positive real numbers, belonging to $l_1$, such that $Qh_n = \lambda_n h_n$. We can therefore regard Gaussian measures as equivalent to a pair $(m, Q)$ and write $\mu^{(m,Q)}$.

**Remark A.1.2** *It may seem obscure why we require $\mathbb{H}$ to be separable. One of the many reasons is that, in order to define for instance simple objects such as the expected value* *of* $\mathbb{E}X :==: \int_{\omega\in\Omega} X(\omega)d\mu(\omega)$*, where* $X$ *is an* $\mathbb{H}$*-valued random variable, the Lebesgue-like strategy of approximating* $X(\omega)$ *by a series of simple random variables* $\{X_n\}_{n=1}^{\infty}$ *may not work if the space is not separable. For if it is not, then we have no guarantee that we can construct a sequence* $\{X_n\}_{n=1}^{\infty}$ *such that* $\|X_n - X\| \downarrow 0$ *almost surely. For interesting discussions of why separability is generally assumed for the state space of random variables (this state space being generally in the weakest case a separable metric space), the reader is referred to for instance Chapter 2 of Ledoux and Talagrand [29], Chapter 1 of Ikeda and Watanabe [23] or Chapter 1 of Da-Prato and Zabczyk [13].*

## A.2 $Q$-Wiener processes and cylindrical Wiener processes

**Definition A.2.1** *Let* $Q$ *be a symmetric, non negative, trace class operator on* $\mathbb{H}$*. A continuous adapted stochastic process* $W_t^{(Q)}$ *taking values in* $\mathbb{H}$*, defined on a filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$*, with independent and identically distributed increments, and such that the law of* $W_1$ *is* $\mu^{(0,Q)} =: \mu^{(Q)}$*, is called a* $Q$*-Wiener process in* $\mathbb{H}$*.*

Given $\mu^{(Q)}$, it is actually easy to construct a $Q$-Wiener process. Indeed, the process defined by $W_t := \sum_{n=1}^{\infty} \sqrt{\lambda_n} W_t^{(n)} h_n$, for some family of independent scalar Brownian motions $\{W_t^{(n)}\}_{n=1}^{\infty}$ is easily shown to be a $Q$-Wiener process.

If we start from a symmetric, non negative and bounded operator $Q$ but which is not trace class (i.e. with $trace(Q) = \infty$), we can still informally consider the process $W$ given by the formula $W_t := \sum_{n=1}^{\infty} \sqrt{\lambda_n} W_t^{(n)} h_n$, where now $\lambda \in l_\infty$ but not necessarily $\lambda \in l_1$. However, for all $t \geq 0$, $W_t$ is almost surely not valued in $\mathbb{H}$, or even: $(W_t \notin \mathbb{H}\ \forall t \geq 0), a.s.$. $W$ would have to be seen as taking values in a larger Hilbert space in which $\mathbb{H}$ is embedded. Notice that this can for instance be achieved by defining a new space $\mathbb{H}^{(\gamma)}$ whose basis is defined as $\{h_n^{(\gamma)}\}_{n=1}^{\infty} := \{\gamma_n^{-1} h_n\}_{n=1}^{\infty}$ for an arbitrary $\gamma := \{\gamma_n\}_{n=1}^{\infty} \in l_2$. Indeed, $\mathbb{H}$ is embedded in $\mathbb{H}^{(\gamma)}$, for if $x = \sum_{n=1}^{\infty} x_n h_n \in \mathbb{H}$, we have that $\|x\|_{\mathbb{H}^{(\gamma)}}^2 = \sum_{n=1}^{\infty} x_n^2 \gamma_n^2 < \infty$, so that $x \in \mathbb{H}^{(\gamma)}$, but now in this new space $\mathbb{H}^{(\gamma)}$, we have that $\mathbb{E}||W_t||^2_{\mathbb{H}(\gamma)} = \mathbb{E}\sum_{n=1}^{\infty}||W_t^{(n)}\sqrt{\lambda_n}h_n||^2_{\mathbb{H}(\gamma)} = t\sum_{n=1}^{\infty}\lambda_n\gamma_n^2 \leq t||\lambda||_\infty||\gamma||_{l_2} < \infty$. However, even in the original space $\mathbb{H}$, it is still the case that for any $h \in \mathbb{H}$, the product $< W_t, h >_{\mathbb{H}}$, defined in the obvious way, is a well defined scalar Brownian motion (because of the martingale convergence theorem in $L_2(\Omega)$). Probably for this reason, $W$ so defined is generally called a cylindrical Wiener process. Notice that the case where $Q$ is the identity operator $I$ defines a cylindrical Wiener process in the above sense, and this is actually the only case we consider outside this section: whenever we talk about a $\mathbb{G}$-cylindrical Wiener process (or Brownian motion) $(W_t)_{t\geq 0}$, we mean $Q = I$ and it is to be understood implicitly that $W_t$ is not really valued in $\mathbb{G}$.

## A.3 Stochastic integral against a cylindrical Wiener process

Let $W$ :==: $(W_t)_{t\geq 0}$ be a $\mathbb{G}$-cylindrical Brownian motion, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We wish to define Ito-like stochastic integrals with respect to $W$, say of the form $(\int_{u=0}^T \phi_u dW_u)_{T\geq 0}$, and would like this to be a square integrable $\mathbb{F}$-valued martingale, where $\mathbb{F}$ is also a separable Hilbert space. The question is: what is an appropriate class of integrands $\phi := (\phi_u)_{u\geq 0}$? It turns out that the answer depends only on $\mathbb{G}$ and $\mathbb{F}$ (recall that this is not a priori obvious, because $W$ does not really live in $\mathbb{G}$!). Fortunately, we can take $\phi$ to be a predictable square integrable stochastic process valued in $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$, the space of Hilbert-Schmidt operators from $\mathbb{G}$ to $\mathbb{F}$, i.e. such that for all $T \geq 0$:

$$|||\phi|||_T := \sqrt{\mathbb{E}\int_{u=0}^T ||\phi_u||^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} du} < \infty \tag{A.3.1}$$

The reason why we can define the integral for this family of integrands is that we can approximate $\phi$ (in the topology induced by the norm $|||.|||_T$) by simple integrands, i.e. by linear combinations of integrands of the form $X_s\mathbb{1}_{(s,t]}(u)f \otimes g$, where $s \leq t$, $X_s$ is an $\mathcal{F}_s$-measurable real value square integrable random variable and $f$ and $g$ are vectors in $\mathbb{F}$ and $\mathbb{G}$. It is therefore enough to decree that:

$$\int_{u=0}^T [X_s\mathbb{1}_{(s,t]}(u)f \otimes g]dW_u := X_s < g, W_{t\wedge T} - W_{s\wedge T} >_{\mathbb{G}} f \tag{A.3.2}$$

The generalization to $\phi$ satisfying (A.3.1) is then straightforward once we notice the Ito isometry:

$$\mathbb{E}\left\| \int_{u=0}^{T} [X_s \mathbb{1}_{(s,t]}(u) f \otimes g] dW_u \right\|_{\mathbb{F}}^2 = \left|\left|\left| X_s \mathbb{1}_{(s,t]}(u) f \otimes g \right|\right|\right|_T^2$$

Let us finally notice that the Burkholder's inequalities are available for Ito infinite dimensional stochastic integrals, and along with Jensen's inequality implies that for all $T \geq 0$ and $p \geq 2$, there exists a constant $B_{T,p}$ depending only on $T$ and $p$ such that:

$$\mathbb{E}\Big[\sup_{t \leq T} \left\| \int_{u=0}^{t} \phi_u dW_u \right\|^p\Big] \leq B_{T,p}\ \mathbb{E} \int_{u=0}^{T} \left\| \phi_u \right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}^p du$$

for all sufficiently integrable processes $(\phi_u)_{u \geq T}$. A proof of this inequality can be found in Lemma 7.2 of Da-Prato and Zabczyk [13].

## A.4 Stochastic convolutions

Let now $(S_T)_{T \geq 0}$ be a family of bounded linear operators on $\mathbb{F}$, and suppose further that for all $T \geq 0$, $(S_{T-u}\phi_u)_{u \in [0,T]}$ is a well defined predictable $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$-valued stochastic process on $[0, T]$ satisfying the condition $|||(S_{T-u}\phi_u)_{u \in [0,T]}|||_T < \infty$. Then clearly we can still define, for any fixed value of $T \geq 0$, the stochastic integral:

$$\int_{u=0}^{T} S_{T-u}\phi_u dW_u \tag{A.4.1}$$

However, because $S_{T-u}$ depends now on the upper bound of integration $T$, there is no reason why $(\int_{u=0}^{T} S_{T-u}\phi_u dW_u)_{T \geq 0}$ should be a martingale (and it is generally not). Integrals such as (A.4.1) are known as stochastic convolutions. In financial modelling where the asset is valued in a Hilbert space, like it can be the case in interest rates modelling, or like in the previous chapter, stochastic convolutions are often used with $(S_T)_{T \geq 0}$ being the left shift operators. This trick allows one to pass from the "time *of* maturity" notation to Musiela's "time *to* maturity" notation. See subsection 2.4.3 for more details on this.

Although $(\int_{u=0}^{T} S_{T-u}\phi_u dW_u)_{T \geq 0}$ is no longer (a priori) a martingale, under the additional assumption that $(S_t)_{t \geq 0}$ is a strongly continuous semigroup on $\mathbb{F}$, we still have moments inequalities, i.e. for any $T \geq 0$ and any $p > 2$, there exists a constant $C_{T,p}$ depending only on $T$ and $p$, such that:

$$\mathbb{E}\Big[\sup_{t\leq T}\Big\|\int_{u=0}^{t} S_{t-u}\phi_u dW_u\Big\|^p\Big] \leq C_{T,p}\ \mathbb{E}\int_{u=0}^{T}\big\|\phi_u\big\|^p_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}du$$

for all sufficiently integrable processes $(\phi_u)_{u\geq T}$. See Proposition 7.3 of Da-Prato and Zabczyk [13].

# Appendix B

# A short introduction to Malliavin calculus for Hilbert space valued random variables

In this section, we give an overview of some classical theorems and results on Malliavin calculus which have been used in the main text. There is of course no claim of originality in any of this and these results which in essence can always be found in Nualart's book [36] or Malliavin and Thalmaier's book [30] are just restated here for convenience and to make this thesis more self-contained than it would otherwise be. Some results may also be difficult to find in exactly the same setup (infinite dimensional Brownian motion underlying the isonormal process, Hilbert space valued random variables, $L^p$ spaces with $p \geq 2$) so that it makes sense to detail them here.

## B.1 The derivative operator and the $\mathbb{D}^{1,p}(\mathbb{F})$ spaces

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$ be a filtered probability space supporting a Brownian motion $(W_t)_{t\leq T}$ cylindrically defined on a separable infinite dimensional Hilbert space $\mathbb{G}$, which can be identified without loss of generality with $l_2$. We denote by $\mathbb{F}$ another separable Hilbert space in which the random variables that we will consider are valued. We denote by $\mathbb{U}$ the space of square integrable functions on $[0, T]$, valued in $\mathbb{G}^*$, that is, $u = (u_t)_{t\leq T} \in \mathbb{U}$

means that $\left\|u\right\|_{\mathbb{U}}^{2} := \int_{t=0}^{T} \|u_t\|_{\mathbb{G}^*}^2 dt < \infty$. We define a process, also called $W$, which takes elements of $\mathbb{U}$ into $L^2(\Omega)$, the space of square integrable real-valued random variables, according to the rule: $W(u) := \int_{t=0}^{T} u_t dW_t$. This last integral is to be understood as an Ito stochastic integral against a cylindrical Brownian motion as constructed earlier. $W(.)$ is usually called an isonormal process (because it is an isometry that transforms elements of $\mathbb{U}$ into Gaussian random variables). Specifically, the Ito isometry is equivalent to the following scalar product conservation:

$$< W(u), W(v) >_{L^2(\Omega)} = < u, v >_{\mathbb{U}}$$

As a final remark on this setup, we can observe that $\mathbb{U}$ is itself a separable Hilbert space. Given CONS $\{g_n\}_{n=1}^{\infty}$ and $\{l_n\}_{n=1}^{\infty}$ of respectively $\mathbb{G}^*$ and $L^2[0,T]$, we can see easily that $\{l_n.g_m\}_{n,m=1}^{\infty}$ is a countable CONS of $\mathbb{U}$ that we can rewrite, after some reordering, with a single index as $\{u_n\}_{n=1}^{\infty}$.

We now define the Malliavin derivative operator $D$ as the linear unbounded operator from $\bigcap_{p\geq 1} L^p(\Omega;\mathbb{F})$ into $\bigcap_{p\geq 1} L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$ as follows:

$$D[W(u)^n f] := nW(u)^{n-1} f \otimes u \tag{B.1.1}$$

for any $n \geq 1$, $u \in \mathbb{U}$ and $f \in \mathbb{F}$. Let us call any finite sum of random variables of the form $W(u)^n f$ a "smooth random variable", and observe that we can define their image by imposing that $D$ be linear. It is well known that $D$ so defined is well defined and is closable from $L^p(\Omega;\mathbb{F})$ into $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$ for any $p \geq 1$, and we can therefore denote by $\mathbb{D}^{1,p}(\mathbb{F})$ the closure of $L^p(\Omega;\mathbb{F})$ for the (classical graph) norm $\|.\|_{1,p,\mathbb{F}}$:

$$\|X\|_{1,p,\mathbb{F}}^p := \|X\|_{L^p(\Omega;\mathbb{F})}^p + \|DX\|_{L^p(\Omega;\mathbb{F}\otimes\mathbb{U})}^p$$

We are using the usual notation $\otimes$ to denote the tensor product. For a definition of tensor products of Hilbert spaces, the reader can consult "Functional Analysis" by Reed and Simon [39].

If we were very cautious, we ought to denote the closed extension of $D$ (as defined on the core of smooth random variables) to $\mathbb{D}^{1,p}$ as $D_p$. However, we can remark easily that the $\|.\|_{1,p,\mathbb{F}}$ norms are increasing with $p$, and thus the domains $\mathbb{D}^{1,p}(\mathbb{F})$ are shrinking with $p$ increasing. By the closability of $D$ in the different $L^p(\Omega, \mathbb{F})$ spaces, we see that if an $\mathbb{F}$-valued random variable $X$ belongs to $\mathbb{D}^{1,p}(\mathbb{F})$ and $\mathbb{D}^{1,q}(\mathbb{F})$ then we have that $D_pX = D_qX$ and therefore we can ignore the subscript and just write $D$ for the different extensions of $D$ into the different $L^p(\Omega; \mathbb{F})$ spaces.

We will also denote, following Nualart's notation in [36], by $\mathbb{D}^{1,\infty}(\mathbb{F})$ the intersection $\bigcap_{p\geq 1} \mathbb{D}^{1,p}(\mathbb{F})$. That this space is non empty is obvious as smooth random variables certainly belong to it.

**Remark B.1.1** *It is interesting to remark that a well known basis in $L^2[0,T]$ is the so called Walsh orthonormal system, given by: $l_1 := 1/\sqrt{T}.\mathbb{1}_{[0,T]}$, $l_2 := 1/\sqrt{T}.\mathbb{1}_{[0,T/2)} - 1/\sqrt{T}.\mathbb{1}_{[T/2,T]}$, $l_3 := 1/\sqrt{T}.\mathbb{1}_{[0,T/4)} - 1/\sqrt{T}.\mathbb{1}_{[T/4,T/2)} + 1/\sqrt{T}.\mathbb{1}_{[T/2,3T/4)} - 1/\sqrt{T}.\mathbb{1}_{[3T/4,T]}$, etc...*
*Looking at the Malliavin derivative of smooth random variables of the form $\xi_{n,m} = f \int_{t=0}^{T} l_n(t) \otimes g_m dW_t$ for some $f \in \mathbb{F}$ and some $m, n \in \mathbb{N}$ gives us a very intuitive picture of what the Malliavin derivative is. This is actually the approach which has been chosen in the book by Malliavin and Thalmaier [30]: it shows us that $D_t\xi$ simply gives the variations of the random variable $\xi$ resulting from a change in the Brownian motion's increment at time $t$.*

The following proposition gives us a means to check (well, sometimes) that a random variable belongs to $\mathbb{D}^{1,p}(\mathbb{F})$. The proof we give here is based on a simple extension of Lemma 1.2.3 of Nualart [36] from the Hilbert space case ($p = 2$) to general $p > 1$, instead of using the more complicated proof given in Lemma 1.5.3 of that same book [36] (which in essence is the same, but requires the introduction of other operators).

**Proposition B.1.2** *Let $X_n$ be a series of random variables converging in $L^p(\Omega; \mathbb{F})$ for some $p > 1$ to some random variable $X$. Then:*

$$\left\{\left\|DX_n\right\|_{1,p,\mathbb{F}}\right\}_{n=1}^{\infty} \in l_\infty \;\Rightarrow\; \left(X \in \mathbb{D}^{1,p}(\mathbb{F}) \;\; and \;\; DX_n \to DX\right)$$

Proof of Proposition B.1.2:

Let us denote by $M$ an upper bound for the sequence $\left\{\left\|DX_n\right\|_{L^p(\Omega;\mathbb{F}\otimes\mathbb{U})}\right\}_{n=1}^{\infty}$. Let us pick a countable dense set $\{\xi_m\}_{m=1}^{\infty}$ in $L^q(\Omega;\mathbb{F}\otimes\mathbb{U})$, where $q$ is the adjoint of $p$, given by the relation $p^{-1}+q^{-1}=1$. Notice that this space is the dual (Banach space) of $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$. Also notice that this would not be possible if we had allowed $p=1$ as it is well known that the dual space of a $L^1$ space is $L^\infty$, which is, except in degenerate cases, not separable.

Let us now observe that for any $m\in\mathbb{N}$, the sequence $\{\mathbb{E}\xi_m DX_n\}_{n=1}^{\infty}$ is well defined, by Holder's inequality, and bounded by $\|\xi_m\|M$. It is therefore easy by a diagonal argument to construct a subsequence $\{q_n\}_{n=1}^{\infty}$ such that, for each $m$, the sequence $\{\mathbb{E}\xi_m DX_{q_n}\}_{n=1}^{\infty}$ is converging to a value that we will call $c_m$. We can remark that $|c_m|$ is bounded by $\|\xi_m\|M$ and therefore the map which to $\xi_m$ associates the value $c_m$ can uniquely be extended to the whole of $L^q(\Omega;\mathbb{F}\otimes\mathbb{U})$ to a bounded linear functional that we will call $G$ (i.e., $G$ is a random variable in the space $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$). What we have just done here is to construct a random variable $G$ towards which $DX_{q_n}$ converges weakly in $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$. (Notice that we could also have used directly the Banach-Alaoglu theorem, but the separable Banach space structure makes it unnecessary. The argument we have given here is the one given by Banach himself in his book [2], see paragraph 4, Theorem 2 and Theorem 3 from Chapter 8).

Finally, for each $n$, we can construct $\tilde{X}_n\in C_n$, the closure (in the $\|.\|_{1,p,\mathbb{F}}$ norm) of the convex hull of $(X_{q_1},...,X_{q_n})$, and such that $\{\tilde{X}_n\}_{n=1}^{\infty}$ converges in the (strong) $L^p(\Omega;\mathbb{F})$-norm to $X$ and such that $\{D\tilde{X}_n\}_{n=1}^{\infty}$ converges in the (strong) $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$-norm to $G$ (in short, we can consider $(\tilde{X},\tilde{G})$ the element of $C:=\overline{\bigcup_{k=1}^{\infty}C_k}$ which minimizes the $L^p(\Omega;\mathbb{F})\times L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$-distance to $(X,G)$. That such a distance minimizer element exists and actually belongs to $C$ comes from the fact that $C$ is closed and convex by construction. It is then not too difficult to argue that the weak convergences of $\{X_{q_n}\}_{n=1}^{\infty}$ to $X$ and of $\{DX_{q_n}\}_{n=1}^{\infty}$ to $G$ imply that $(X,G)$ and $(\tilde{X},\tilde{G})$ have to be identical, for instance by using the contrapositive of the separating hyperplane theorem). This shows of course that $X$ is an element of $\mathbb{D}^{1,p}(\mathbb{F})$ and that $DX=G$. ∎

## B.2 Derivative of stochastic integrals

In the main text, when differentiating (in the Malliavin sense) the stock price, we had to take the derivative of a stochastic integral. The following proposition gives us conditions under which we can do so and tells us what the derivative of this integral is. Notice that the formula we get would make sense without the two adjoint symbols "*" but would be wrong. This is because for a fixed $t$ and if, say, $\phi_u$ is valued in $\mathbb{F} \otimes \mathbb{G}$ then $D_t\phi_u$ is valued in $\mathbb{F} \otimes \mathbb{G} \otimes \mathbb{G}$, so that writing something like $\int_{u=t}^{T} D_t\phi_u dW_u$ would make perfect sense, but this is not what we want to do! Despite what we just said, we drop the "*" in most sections, in order to lighten slightly the notation which is already heavy enough.

**Proposition B.2.1** *Let $T \geq 0$ and $p \geq 2$ be given. Let $(\phi_t)_{t\in[0,T]}$ be an $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$-valued stochastic process, predictable and satisfying the integrability condition:*

$$\mathbb{E} \int_{t=0}^{T} \|\phi_t\|^p_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} dt \ < \ \infty.$$

*Suppose in addition that for all $t \in [0, T]$, $\phi_t \in \mathbb{D}^{1,p}(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}))$ and that:*

$$\mathbb{E} \int_{t=0}^{T} \|D\phi_t\|^p_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})\otimes\mathbb{U}} dt \ < \ \infty$$

*Then, the $\mathbb{F}$-valued random variable $\int_{t=0}^{T} \phi_t dW_t$ belongs to $\mathbb{D}^{1,p}(\mathbb{F})$ and:*

$$D \int_{u=0}^{T} \phi_u dW_u \ = \ \phi(. \wedge T) + \left\{ \int_{u=0}^{T} (D\phi_u)^* dW_u \right\}^*$$

Proof of Proposition B.2.1:

Let us show the proposition for elementary integrands of the form

$$\psi_u = W(h)^n f \mathbb{1}_{(s,t]}(u) \otimes g$$

where $h(u) = 0$ for $u \geq s$, where $s \leq t$, $n \in \mathbb{N}$ and $(f, g) \in \mathbb{F} \times \mathbb{G}$:

$$
\begin{aligned}
D \int_{u=0}^{T} \psi_u dW_u &= D\Big[ W(h)^n f < g, W_{t\wedge T} - W_{s\wedge T} >_{\mathbb{G}} \Big] \\
&= W(h)^n f D < g, W_{t\wedge T} - W_{s\wedge T} >_{\mathbb{G}} + D[W(h)^n] f < g, W_{t\wedge T} - W_{s\wedge T} >_{\mathbb{G}} \\
&= W(h)^n f \mathbb{1}_{(s,t]}(. \wedge T) \otimes g + nW(h)^{n-1} < g, W_{t\wedge T} - W_{s\wedge T} >_{\mathbb{G}} f \otimes h \\
&= W(h)^n f \mathbb{1}_{(s,t]}(. \wedge T) \otimes g + \int_{u=0}^{T} nW(h)^{n-1} f \otimes g \mathbb{1}_{(s,t]}(u) dW_u \otimes h \\
&= \psi(. \wedge T) + \Big\{ \int_{u=0}^{T} h \otimes [nW(h)^{n-1} f \otimes g \mathbb{1}_{(s,t]}(u)] dW_u \Big\}^* \\
&= \psi(. \wedge T) + \Big\{ \int_{u=0}^{T} [(nW(h)^{n-1} f \otimes g \mathbb{1}_{(s,t]}(u)) \otimes h]^* dW_u \Big\}^* \\
&= \psi(. \wedge T) + \Big\{ \int_{u=0}^{T} (D\psi_u)^* dW_u \Big\}^* \qquad \text{(B.2.1)}
\end{aligned}
$$

Now, by definition of the stochastic integral, there exists a sequence $\{\phi^{(n)}\}_{n=1}^{\infty}$ of simple integrands of the form:

$$
\phi_{(n)} = \sum_{k=0}^{N_n - 1} \xi_k^{(n)} \mathbb{1}_{(t_k^{(n)}, t_{k+1}^{(n)}]}(t)
$$

and which converge to $\phi$ in the sense: $\lim_{n\to\infty} \mathbb{E} \int_{t=0}^{T} \|\phi_t - \phi_t^{(n)}\|^p_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} dt = 0$. For each $n \in \mathbb{N}$, $N_n$ is an integer, $0 =: t_0^{(n)} < .... < t_{N_n}^{(n)} := T$ is a partition of $[0, T]$ and the $\xi_k^{(n)}$'s are $\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})$-valued random variables which are $\mathcal{F}_{t_k^{(n)}}$-measurable and of finite $L^p(\Omega; \mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}))$-norm. By linearity and density arguments, we can extend the above formula (B.2.1) and get:

$$
D \int_{t=0}^{T} \phi_t^{(n)} dW_t = \phi_t^{(n)}(. \wedge T) + \Big\{ \int_{u=0}^{T} (D\phi_u^{(n)})^* dW_u \Big\}^*
$$

Finally, the last line converges as $n$ goes to infinity to $\phi(. \wedge T) + \int_{t=0}^{T} (D\phi_t) dW_t$ in the $L^p(\Omega; \mathbb{F} \otimes \mathbb{U})$-norm because of the hypothesis of the proposition and the convergence of $\phi^{(n)}$ towards $\phi$. ■

## B.3 Chain rules for the Malliavin derivative

Let $\mathbb{F}$ and $\mathbb{H}$ be two separable Hilbert spaces and let $X$ be an $\mathbb{F}$-valued random variable. In many applications, we are interested in differentiating (in the Malliavin sense) a random variable $Y$ of the form $Y = g(X)$, where $g$ is a measurable function from $(\mathbb{F}, \mathcal{B}(\mathbb{F}))$ into $(\mathbb{H}, \mathcal{B}(\mathbb{H}))$. The following proposition is then very useful. (Notice that we still have in the background $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$, $\mathbb{G}$, $\mathbb{U}$ and $W$ which have the same definitions as in the previous section).

**Proposition B.3.1** *Let $p > 1$. If $X \in \mathbb{D}^{1,p}(\mathbb{F})$, and if $g$ is globally Lipschitz on $\mathbb{F}$ in the sense that:*

$$\exists K \geq 0,\ \forall (x, y) \in \mathbb{F}^2,\ \|g(x) - g(y)\|_{\mathbb{H}} \leq K\|x - y\|_{\mathbb{F}}$$

*then $g(F) \in \mathbb{D}^{1,p}(\mathbb{H})$, and there exists $\nabla g$, an $\mathcal{L}(\mathbb{F}, \mathbb{H})$-valued random variable of norm a.s. smaller than $K$, such that:*

$$Dg(X) = \nabla g DX \ \ a.s.$$

Proof of Proposition B.3.1:

The proof is similar to that of Proposition 5.2 in Carmona and Tehranchi [8], but with $p$ arbitrarily strictly larger than 1. ■

In many interesting cases, $g$ is unfortunately not globally Lipschitz, so that the previous proposition does not apply. This is for instance the case when $X$ is a real-valued random variable and $g$ is one of the following: $g(x) = \exp(x)$, $g(x) = x^p, p > 1$ (not Lipschitz at $\infty$) or $g(x) = x^p, 0 < p < 1$ (not Lipschitz at the origin). However, the following proposition may give us a way out in these cases:

**Proposition B.3.2** *Let $p > 1$. Let $X$ be a real-valued random variable belonging to $\mathbb{D}^{1,p}$ and taking values in some open interval $I := (a, b)$, where $a$ and $b$ are possibly $-\infty$ and $\infty$. Let $g$ be a positive, increasing, and $C^1$ function on $I$, not necessarily of bounded* *derivative.*

*If* $\mathbb{E}\|g^{'}(X)DX\|^p_{U^*} < \infty$*, then it holds that:*

$$g(X) \in \mathbb{D}^{1,p}, \text{ and}$$

$$Dg(X) = g'(X)DX$$

Proof of Proposition B.3.2: Let $n_0 \in \mathbb{N}$ and $x_0 \in I$ be such that $|g(x_0)| + |g'(x_0)| < n_0$. Such values $n_0$ and $x_0$ exist by the assumption that $g$ is $C^1$. We then define for each $n \geq n_0$ $I^{(n)} := (a^{(n)}, b^{(n)})$, where $\{a^{(n)}\}_{n=1}^{\infty}$ ( respectively $\{b^{(n)}\}_{n=1}^{\infty}$) is a decreasing (resp. increasing) sequence, bounded above (resp. below) by $x_0$, converging to $a$ (resp. $b$) and such that on each interval $I^{(n)}$, $g$ together with its derivative $g'$ are bounded by $n$. The existence of such sequences is also guaranteed by the fact that $g$ is $C^1$. Finally, we define the truncated approximating functions $g_n, n \geq 0$ by:

$$g_n(x) = g(a^{(n)})\mathbb{1}_{\{x \leq a^{(n)}\}}(x) + g(x)\mathbb{1}_{\{x \in [a^{(n)}, b^{(n)}]\}}(x) + g(b^{(n)})\mathbb{1}_{\{x \geq b^{(n)}\}}(x)$$

By definition, $g$ and $g_n$ coincide with each other on $\mathcal{I}^{(n)}$, and $\mathcal{I}^{(n)} \uparrow I$ as $n \uparrow \infty$. Therefore $g_n(X)$ converges to $g(X)$ almost surely, and by the bounded convergence theorem, the convergence is also true in $L^p$. Also by construction, each of the $g_n$ is a Lipschitz function of coefficient $n$. Therefore, $g_n(X) \in \mathbb{D}^{1,p}$. Finally, we have that:

$$\lim_{n \to \infty} \mathbb{E}\|g'(X)DX - g_n'(X)DX\|^p_{U^*} = 0$$

by monotone convergence because $\|g'(X)DX - g_n'(X)DX\|$ is non zero only if $X(\omega) \notin I^{(n)}$ and when it is non zero, it is equal to $\|g'(X)DX\|_{U^*}$. This proves the proposition. ■

# Appendix C

# Existence, uniqueness and Malliavin differentiability of mild solutions to SPDEs in Hilbert spaces

The following theorem is a particular case of Theorem 7.4 of Da-Prato and Zabczyk [13] and has been used in the main text to define the forward variance swaps curve. We detail the proof here for completeness and adapt it to our specific case. Notice that, in essence, this is still the same usual story as for ordinary differential equations, i.e. using the fixed point theorem for contractions.

**Theorem C.0.3** *Let $\mathbb{G}$ and $\mathbb{F}$ be two separable Hilbert spaces and $(W_t)_{t\geq 0}$ a $\mathbb{G}$-cylindrical Wiener process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Let $(S_t)_{t\geq 0}$ be a strongly continuous semi-group on $\mathbb{F}$, with infinitesimal generator $A$. Let $\sigma$ be a (measurable) map from $(\mathbb{R}_+ \times \Omega \times \mathbb{F}, \mathcal{P} \times \mathcal{B}(\mathbb{F}))$ into $(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}), \mathcal{B}(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F})))$ satisfying the global Lischitz and growth conditions:*

$$
\begin{aligned}
\|\sigma(t,\omega;f_1) - \sigma(t,\omega;f_2)\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} &\leq K\|f_1 - f_2\|_{\mathbb{F}}, \ \forall (t,\omega,f_1,f_2) \in \mathbb{R}_+ \times \Omega \times \mathbb{F}^2 \\
\|\sigma(t,\omega;f)\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})} &\leq K(1 + \|f\|_{\mathbb{F}}), \ \forall (t,\omega,f) \in \mathbb{R}_+ \times \Omega \times \mathbb{F}
\end{aligned}
$$

*for some constant $K \geq 0$.*

*Then, for any $v_0 \in \mathbb{F}$, there exists a unique (up to indistinguishability) continuous $\mathbb{F}$-*

*valued stochastic process* $(v_t)_{t\geq 0}$ *starting at* $v_0$ *which satisfies the evolution equation:*

$$v_T = S_T v_0 + \int_0^T S_{T-u}\sigma(u,\omega;v_u)dW_u$$

*In addition, for any* $p \geq 2$, $T \geq 0$, *there is a positive constant* $A_{T,p}$ *such that:*

$$\sup_{t\in[0,T]} \mathbb{E}\{\|v_T\|^p\} \leq A_{T,p}$$

*In particular, this implies that* $v_T$ *belongs to* $\bigcap_{n=1}^{\infty} L^p(\Omega;\mathbb{F})$.

The above solution is said to be a mild solution to the stochastic partial differential equation:

$$dv_t = Av_t dt + \sigma(t,\omega;v_t)dW_t$$

with initial condition $v_0$. Let us remark that the term "mild" is here opposed to "strong" in a sense similar to the distinction made between strong and weak solutions of PDEs (see Reed and Simon [39] p. 149 for instance), but not in the probabilistic sense. In the probabilistic sense (see for instance Karatzas and Shreve [24] p. 285), our mild solution introduced above is a strong solution, i.e. the solution $v_t$ to the equation is adapted to the augmented filtration generated by the cylindrical Brownian motion. This means (see for instance Williams's "Probability with Martingales" [46]) that there exists a measurable map $\mathcal{M}$ such that $v_t = \mathcal{M}((W_u)_{u\leq t})$. The reason why we need to make use of mild solutions here is that, although it may be impossible to find a solution $v$ such that

$$v_T = v_0 + \int_{t=0}^T Av_t dt + \int_{t=0}^T \sigma(t,\omega;v_t)dW_t$$

(because we cannot guarantee that $v_t$ will remain in the domain of $A$ which is likely to be an unbounded operator), it may still be possible to find $v$ satisfying the mild form ($S_t$, unlike $A$, is a linear bounded operator defined on the whole of $\mathbb{F}$).

Proof of Theorem C.0.3:

We outline the proof here for convenience. A full proof with more details can be found in Da-Prato and Zabczyk [13] p186 to 193. We start by proving uniqueness. Let us suppose that we have two solutions $(v_t^{(1)})_{t\geq 0}$ and $(v_t^{(2)})_{t\geq 0}$. For $j \in \{1,2\}$, and arbitrary

$0 \le t \le T$, it holds that:

$$\mathbb{E}\int_{u=0}^{t} \|\sigma(u,\omega;v_u^{(j)})\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}du \le \mathbb{E}\int_{u=0}^{T} K^2(1+\|v_u^{(j)}\|^2) \le K^2T(1+A_{T,2}) < \infty$$

So it must hold that $\int_{u=0}^{T} \|\sigma(u,\omega;v_u^{(j)})\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}du \ < \ \infty$ a.s. We can therefore define for any $M \ge 0$ the stopping time $\tau_M := \inf\{t \le T, \exists j \in \{1,2\}, \int_{u=0}^{t} \|\sigma(u,\omega;v_u^{(j)})\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}du \ge M\}$. Let us denote by $\tilde{v}_t^{(j)}$, $j \in \{1,2\}$ the processes killed at the stopping time $\tau_M$: $(\tilde{v}_t^{(j)}) := \mathbb{1}_{\{t\le\tau_M(\omega)\}}(t,\omega)v_t^j$. Then we have for any $t \le T$:

$$\begin{aligned}
&\mathbb{E}\|\tilde{v}_t^{(2)} - \tilde{v}_t^{(1)}\|^2_{\mathbb{F}} \\
&\le \mathbb{E}\{\mathbb{1}_{\{t\le\tau_M(\omega)\}}(t,\omega)\int_{u=0}^{t} \|\mathbb{1}_{\{u\le\tau_M(\omega)\}}(u,\omega)S_{t-u}(\sigma(u,\omega;v_u^{(2)}) - \sigma(u,\omega;v_u^{(1)}))\|^2_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}du\} \\
&\le \ B_T^2K^2\mathbb{E}\{\int_{u=0}^{t} \|\tilde{v}_u^{(2)} - \tilde{v}_u^{(1)}\|^2_{\mathbb{F}}du\}
\end{aligned}$$

where $B_T$ denotes a bound on $\sup_{t\le T}\left\|S_t\right\|_{\mathcal{L}(\mathbb{F})}$. This implies by Gronwall's Lemma (and the finiteness of all quantities in the above inequalities (bounded by a multiple of $M$)) that $\mathbb{E}\|\tilde{v}_t^{(2)} - \tilde{v}_t^{(1)}\|^2_{\mathbb{F}} = 0$ for all $t \le T$. As this holds for arbitrarily large $M$, we deduce that $\mathbb{E}\|v_t^{(2)} - v_t^{(1)}\|^2_{\mathbb{F}} = 0$ for any $t \le T$ as well. Finally, using the continuity of $v_t^{(j)}$, that implies that both solutions are indistinguishable.

As for the existence of a continuous solution, it can be proved by making use of the Picard iterations, i.e. introducing: $v_t^{(0)} := S_t v_0$ and then recursively, for $n \ge 1$ and all $t \ge 0$:

$$v_t^{(n+1)} := S_t v_0 \ + \ \int_{u=0}^{t} S_{t-u}\sigma(u,\omega;v_u^{(n)})dW_u$$

By induction, we can remark that $(v_t^{(n)})_{t\ge0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\ge0}$. We have then that for any $n \ge 1$ and any $T \ge 0$:

$$\sup_{t\le T}\mathbb{E}\left\|v_t^{(n+1)} - v_t^{(n)}\right\|^p \le K^pB_T^pB_{T,p}\sup_{t\le T}\mathbb{E}\left\|v_t^{(n)} - v_t^{(n-1)}\right\|^p$$

As we can choose a particular value of $T$ (say $\tilde{T}$) that makes the above constant $K^pB_T^pB_{T,p}$ strictly less than 1, we have therefore convergence of $v_t^n$ on $[0,\tilde{T}]$ to a limit that we call $v_t$ and which is seen to satisfy the equation, and is bounded in the $\left(\sup_{t\le\tilde{T}}\mathbb{E}\|v_t\|^p\right)^{1/p}$-norm. Notice that we can then patch solutions on intervals of the form $[0,\tilde{T}]$, $[\tilde{T},2\tilde{T}]$, etc, to construct the solution on the original interval of interest $[0, T]$. $\blacksquare$

We needed in the main text to take the Malliavin derivative of the mild solution to a SPDE in Hilbert space. We made use of the following theorem:

**Theorem C.0.4** *Let $\mathbb{G}$, $\mathbb{F}$, $(W_t)_{t\geq 0}$, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, $(S_t)_{t\geq 0}$, $A$ and $\sigma$ be defined as in the previous Theorem C.0.3. We assume in addition that $\sigma$ does not depend on $\omega$ in a direct manner, but only through $v_t$, i.e. $\sigma = \sigma(t; v_t(\omega))$. We denote by $(v_t)_{t\geq 0}$ the unique continuous mild solution to the equation $dv_t = Av_t dt + \sigma(t; v_t)dW_t$ with initial condition $v_0$.*

*Then for any $T \geq 0$, $v_T \in \mathbb{D}^{1,\infty}(\mathbb{F})$ and for any $t \geq 0$, there exists a unique strong $\mathcal{L}(\mathbb{F})$-valued random variable $Y_{t,T}$ such that $D_t v_T = Y_{t,T}\sigma_t$. The family of random variables $(Y_{t,T})_{0\leq t\leq T<\infty}$ is the unique solution to the family of equations:*

$$Y_{t,T} = S_{T-t} + \int_{u=t}^{T} S_{T-u} \nabla\sigma_u Y_{t,u} dW_u$$

*Moreover, for any $t, T \geq 0$ with $t \leq T$, any $p \geq 2$, and any $\mathbb{F}$-valued integrable random variable $f_t$, there exists a positive constant $C_{T,p}$ depending only on $T$ and $p$ such that:*

$$\sup_{u\in[t,T]} \mathbb{E}_t \left\| Y_{t,u} f_t \right\|_{\mathbb{F}}^{p} \leq C_{T,p} \left\| f_t \right\|_{\mathbb{F}}^{p} \tag{C.0.1}$$

Proof of Theorem C.0.4:

We give below the proof of the theorem in the general case $p \geq 2$. The particular case $p = 2$ is treated in Carmona and Tehranchi [8]. Our proof is divided in three parts: we first show that $v_T \in \mathbb{D}^{1,\infty}(\mathbb{F})$ for all $T$, and give the $\mathbb{F} \otimes \mathbb{U}$-valued SPDE to which $Dv_T$ is solution. Then we explain where the family $Y$ of strong operators is coming from, and finally we derive the bound (C.0.1). During the proof, we make again use of the Picard iterations, that is, we define:

$$v_T^{(0)} := S_T v_0$$

and then, by induction, we define for $n \geq 0$

$$v_T^{(n+1)} := S_T v_0 + \int_{u=0}^{T} S_{T-u}\sigma(u; v_u^{(n)})dW_u =: S_T v_0 + \int_{u=0}^{T} \phi_{T,u}^{(n)} dW_u$$

We have already seen that these iterations are properly defined, and that for any $T \geq 0$ and any $p \geq 1$, there is uniform convergence on $[0, T]$ of $v_t^{(n)}$ to $v_t$ in the $L^p(\Omega; \mathbb{F})$-norm. In particular, there exists a constant $c_{T,p}$ such that $\sup_{n\in\mathbb{N}} \sup_{t\leq T} \mathbb{E}\left\|v_t^{(n)}\right\|_{\mathbb{F}}^p < c_{T,p}$.

## C.1 Malliavin differentiability of $v_T$ and SPDE

We now show that the series $\{Dv_T^{(n)}\}_{n=1}^{\infty}$ is bounded in the $L^p(\Omega; \mathbb{F} \otimes \mathbb{U})$-norm so that we can make use of Proposition B.1.2. We proceed by induction, showing that for each $n$, we can use the formula of Proposition B.2.1 to get the Malliavin derivative of $v_T^{(n)}$ and to compute a bound on the $L^p(\Omega; \mathbb{F}\otimes\mathbb{U})$-norm of $Dv_T^{(n)}$. Let us start by noticing that the case $n = 0$ is trivial as $v_u^{(0)}$ is non random and thus, $Dv_u^{(0)}$ exists and is equal to 0 for all $u \leq T$. Let us now suppose that $n$ is such that $n \geq 0$, $v_u^{(n)} \in \mathbb{D}^{1,\infty}(\mathbb{F})$, $\phi_{T,u}^{(n)} \in \mathbb{D}^{1,\infty}(\mathcal{L}_{HS}(\mathbb{G}, \mathbb{F}))$ for all $u \leq T$, and that $\int_{u=0}^{T} \left\|D\phi_{T,u}^{(n)}\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})\otimes\mathbb{U}}^p du < \infty$ for all $p \geq 2$. This assumption allows us to use Proposition B.2.1 on $v_T^{(n+1)}$ to get that:

$$D_t v_T^{(n+1)} = S_{T-t}\sigma(t; v_t^{(n)}) + \int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)} dW_u \tag{C.1.1}$$

and therefore, for any $p \geq 2$:

$$\begin{aligned}
\mathbb{E}\int_{t=0}^{T} \left\|D_t v_T^{(n+1)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p dt &= \mathbb{E}\int_{t=0}^{T} \left\|S_{T-t}\sigma(t; v_t^{(n)}) + \int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)} dW_u\right\|_{\mathbb{F}\otimes\mathbb{G}}^p dt \\
&\leq \mathbb{E}\int_{t=0}^{T} 2^p \left\|S_{T-t}\sigma(t; v_t^{(n)})\right\|_{\mathbb{F}\otimes\mathbb{G}}^p dt \\
&\quad + \mathbb{E}\int_{t=0}^{T} 2^p \left\|\int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)} dW_u\right\|_{\mathbb{F}\otimes\mathbb{G}}^p dt \\
&\leq 2^p M^p \mathbb{E}\int_{t=0}^{T} K^p (1 + \left\|v_t^{(n)}\right\|_{\mathbb{F}})^p dt \\
&\quad + 2^p B_{T,p} \mathbb{E}\int_{t=0}^{T}\int_{u=t}^{T} \left\|S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p du dt
\end{aligned}$$

where we have used $M$ to denote the quantity $\sup_{u\leq T} \left\|S_u\right\|_{\mathbb{F}}$, $K$ to denote the linear and Lipschitz bounds on $\sigma$ and where $B_{T,p}$ is a bound depending only on $p$ and $T$ that satisfies:

$$\mathbb{E}\left\|\int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)} dW_u\right\|_{\mathbb{F}\otimes\mathbb{G}}^p \leq B_{T,p} \mathbb{E}\int_{u=t}^{T} \left\|S_{T-u}\nabla\sigma(u; v_u^{(n)}) D_t v_u^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p du$$

We continue simplifying the above inequalities to get that:

$$\begin{aligned}
&\mathbb{E}\int_{t=0}^{T}\left\|D_t v_T^{(n+1)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p dt \\
&\leq (4MK)^p T(1+C_{T,p}) + B_{T,p}(2MK)^p \mathbb{E}\int_{u=0}^{T}\int_{t=0}^{u}\left\|D_t v_u^{(n)}\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F})}^p dt du
\end{aligned}$$

where we have used Fubini's theorem to interchange the right hand side double integral. Letting finally $a_{T,p}$ be the maximum of $(4MK)^p T(1+C_{T,p})$ and $B_{T,p}(2MK)^p$, and noticing that $a_{T,p}$ can be chosen to be an increasing function of $T$, we obtain that for $t \leq T$:

$$\mathbb{E}\int_{u=0}^{t}\left\|D_u v_t^{(n+1)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p du \leq a_{T,p}(1+\int_{u=0}^{t}\int_{\tau=0}^{u}\left\|D_\tau v_u^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p d\tau du)$$

which allows us to prove by induction that for all $n \in \mathbb{N}$, $T \geq 0$:

$$\sup_{t\leq T}\mathbb{E}\int_{u=0}^{t}\left\|D_u v_t^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p du \leq a_{T,p}\exp(Ta_{T,p}) \tag{C.1.2}$$

Finally, we remark that by Jensen's inequality:

$$\left\|D_u v_t^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{U}}^p \leq T^{p/2-1}\mathbb{E}\int_{u=0}^{t}\left\|D_u v_t^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{G}}^p du$$

so that:

$$\sup_{t\leq T}\mathbb{E}\left\|D_u v_t^{(n)}\right\|_{\mathbb{F}\otimes\mathbb{U}}^p \leq T^{p/2-1}a_{T,p}\exp(Ta_{T,p}) \tag{C.1.3}$$

Using Proposition B.1.2, we conclude that $v_T \in \mathbb{D}^{1,\infty}(\mathbb{F})$ as desired. To conclude the induction that we have started, we need to establish as well that:

$$\int_{u=0}^{T}\left\|D\phi_{T,u}^{(n+1)}\right\|_{\mathcal{L}_{HS}(\mathbb{G},\mathbb{F}))\otimes\mathbb{U}}^p du \ < \ \infty$$

for all $p \geq 2$, but this is now clear from the inequality (C.1.3).

## C.2 The family of operators *Y*

We now turn to the second part of the theorem, introducing the family of strong operators $Y_{t,T}$, and establishing the bound (C.0.1). We can let $n$ go to infinity in equation (C.1.1), because all three quantities involved in that equation converge in $L^p(\Omega;\mathbb{F}\otimes\mathbb{U})$

(or in $L^p(\Omega; \mathbb{F})$ if we look at this equation for a fixed value of $t$). We get therefore that, for all $t \leq T$:

$$D_t v_T = S_{T-t}\sigma(t; v_t) + \int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u) D_t v_u dW_u$$

Let us assume for now that we can define the family $Y$ of strong $\mathcal{L}(\mathbb{F})$-valued operators $(Y_{t,T})_{0 \leq t \leq T < \infty}$, which satisfies for all $f \in \mathbb{F}$, and all $t \leq T$:

$$Y_{t,T} f = S_{T-t} f + \int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u) Y_{t,u} f dW_u$$

Then we can see by uniqueness of the solution to equation (C.1.3) that $D_t v_T$ and $Y_{t,T}\sigma(t; v_t)$ have to be identical, hence the formula of the proposition. It remains only to show that we can define the family $Y$ and that it satisfies the announced bound: proceeding similarly to earlier when we proved the uniform (for $t \in [0, T]$) convergence of the Picard iteration $v_t^{(n)}$ to $v_t$ in the $L^p(\Omega; \mathbb{F})$-norm, we can define for any $f \in \mathbb{F}$ the following Picard iterations:

$$Y_{t,T}^{(0)} f := S_{T-t} f$$

for $0 \leq t \leq T < \infty$, and then for all $n \in \mathbb{N}$ we define by induction

$$Y_{t,T}^{(n+1)} f := S_{T-t} f + \int_{u=t}^{T} S_{T-u}\nabla\sigma(u; v_u) Y_{t,u}^{(n)} f dW_u$$

Exactly by similar arguments to those given earlier, the series $\{\mathbb{E} \sup_{\tau \in [t,T]} \left\| Y_{t,\tau}^{(n)} f \right\|_{\mathbb{F}}^p\}_{n=1}^{\infty}$ is Cauchy for arbitrary $p \geq 2$, which allows us to define $(Y_{t,\tau} f)_{\tau \in [t,T]}$ as the limit of $(Y_{t,\tau}^{(n)} f)_{\tau \in [t,T]}$. We can finally observe that $Y_{t,\tau}$ is linear in $f$ because each Picard iteration is, and therefore defines a strong operator valued process, solution to the equation:

$$Y_{t,\tau} = S_{\tau-t} + \int_{u=t}^{\tau} S_{\tau-u}\nabla\sigma(u; v_u) Y_{t,u} dW_u$$

provided of course that we set by definition:

$$\Big(\int_{u=t}^{\tau} S_{\tau-u}\nabla\sigma_u Y_{t,u} dW_u\Big)(f) := \int_{u=t}^{\tau} S_{\tau-u}\nabla\sigma_u Y_{t,u} f dW_u.$$

## C.3 Bound on $Y$

Finally, let us proceed by induction to show the bound (C.0.1): let $0 \leq t \leq u \leq T < \infty$ and $f_t$ an $\mathcal{F}_t$-measurable, integrable $\mathbb{F}$-valued random variable. Then, with still $M$ denoting $\sup_{t \leq T} \left\|S_t\right\|_{\mathbb{F}}$, we have that:

$$\mathbb{E}_t\left\|Y^{(0)}_{t,u} f_t\right\|^p_{\mathbb{F}} = \mathbb{E}_t\left\|S_{t-u} f_t\right\|^p_{\mathbb{F}} \leq M^p \|f_t\|^p_{\mathbb{F}}$$

and then by induction for $n \geq 0$:

$$\mathbb{E}_t\left\|Y^{(n+1)}_{t,u} f_t\right\|^p_{\mathbb{F}} \leq (2M)^p \left\|f_t\right\|^p_{\mathbb{F}} + (2MK)^p B_{T,p} \mathbb{E}_t \int_{u=t}^{T} \left\|Y^{(n)}_{t,u} f_t\right\|^p_{\mathbb{F}} du$$

By induction again, we deduce that for all $n \in \mathbb{N}$:

$$\sup_{u \in [t,T]} \mathbb{E}_t\left\|Y^{(n)}_{t,u} f_t\right\|^p_{\mathbb{F}} \leq \left\|f_t\right\|^p_{\mathbb{F}} b_{T,p} \exp\left(b_{T,p}\right)$$

where $b_{T,p}$ is a constant larger than $(2M)^p$ and $(2MK)^p B_{T,p}$. ■

# Bibliography


[1] Berrier F., Rogers L.C.G. and Tehranchi M.R., A characterization of forward utility functions, preprint, 2007.

[2] Banach S., Theorie des operations lineaires, Monografie Matematyczne 1, 1932.

[3] Bjork T., Di-Masi G., Kabanov Y. and Runggaldier W., Towards a general theory of bond markets, Finance and Stochastics, Vol. 1, pp. 141-174, 1997.

[4] Buehler H., Consistent Variance Curve Models, Finance and Stochastics, Vol. 10, No. 2, 2006.

[5] Carmona R. and Nadtochiy S., An infinite dimensional stochastic analysis approach to local volatility dynamic models, Communications on Stochastic Analysis, Vol. 2, No. 1, pp. 109-123, 2008.

[6] Carmona R. and Nadtochiy S., Local volatility dynamic models, Finance and Stochastics, Vol. 13, No. 1, pp. 1-48, 2009.

[7] Carmona R. and Tehranchi M., A characterization of hedging portfolios for interest rates contingent claims, The annals of Applied Probability, Vol. 14, No. 3, pp. 1267-1294, 2004.

[8] Carmona R. and Tehranchi M., Interest Rates Models, an Infinite Dimensional Stochastic Analysis Perspective, Springer Finance, 2006.

[9] Carr P. and Madan D., Towards a theory of volatility trading, Option pricing, interest rates and risk management, pp. 458-476, Handb. Math. Finance, Cambridge Univ. Press, Cambridge, 2001.

[10] Christoffersen P., Heston S., and Jacobs K., The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well, working paper, 2009.

[11] Cont R., Modeling Term Structure Dynamics: An Infinite Dimensional Approach, International Journal of Theoretical and Applied Finance, Vol. 8, No. 3, pp. 357-380, 2005.

[12] Cont R., Da Fonseca J., and Durrleman V., Stochastic Models Of Implied Volatility Surfaces, Economic Notes by Banca Monte dei Paschi di Siena SpA, Vol. 31, No. 2, pp. 361-377, 2002.

[13] Da-Prato G. and Zabczyk J., Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and Its Applications 44, Cambridge University Press, 1992.

[14] Davis M.H.A., Complete-market models of stochastic volatility, Stochastic analysis with applications to mathematical finance, Proceedings of the Royal Society, Lond. Ser. A Math. Phys. Eng. Sci, Vol. 460, No. 2041, pp. 11-26, 2004.

[15] De-Donno M. and Pratelli M., On the use of measure-valued strategies in bond markets, Finance and Stochastics, Vol. 8, pp. 87-109, 2004.

[16] El Karoui N. and M'rad M., Forward Dynamic Utility, powerpoint presentation accessed from http://stats.lse.ac.uk/cetin/istanbulworkshop/ElKaroui.pdf on the 28/08/2009.

[17] Feller W., Two Singular Diffusion Problems, Annals of Mathematics, Vol. 54, No. 1, July 1951.

[18] Filipovic D., Consistency problems for Heath-Jarrow-Morton interest rate models, Lecture Notes in Math. 1760, Springer-Verlag, 2001.

[19] Follmer H. and Schweizer M., Hedging of Contingent Claims under Incomplete Information, M. H. A. Davis and R. J. Elliott (eds.), Applied Stochastic Analysis,

Stochastics Monographs, Vol. 5, Gordon and Breach, London/New York , pp. 389-414, 1991.

[20] Gantmacher F.R., Matrizenrechnung II, VEB Deutscher Verlag der Wissenschaften, Berlin, 1959.

[21] Henderson V. and Hobson D., Horizon-unbiased utility functions, Stochastic Processes and their Application, Vol. 117, No. 11, pp. 1621-1641, 2007.

[22] Heston S.L., A Closed Form Solution For Options with Stochastic Volatility, with Applications to Bond and Currency Options, The review of financial studies, Vol. 6, 327-343, 1993.

[23] Ikeda N. and Watanabe S., Stochastic Differential Equations and Diffusions Processes, Second Edition, North-Holland Publishing Company, 1989.

[24] Karatzas I. and Shreve S.E., Brownian Motion and Stochastic Calculus, Springer, 1998.

[25] Kramkov D. and Schachermayer W., The Asymptotic Elasticity Of Utility Functions And Optimal Investment In Incomplete Markets, Annals of Applied Probabilities, Vol. 9, No. 3, pp. 904-950, 1999.

[26] Karatzas I. and Shreve S.E., Methods of Mathematical Finance, Springer-Verlag New York Inc, 2001.

[27] Keller-Ressel M., Moment Explosion and Long Term Behavior of Affine Stochastic Volatility Models, Preprint, 2009.

[28] Kunita H. Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, 1997.

[29] Ledoux M., Talagrand M., Probability in Banach spaces: isoperimetry and processes, Springer, 1991.

[30] Malliavin P. and Thalmaier A., Stochastic Calculus of Variations in Mathematical Finance, Springer Finance, 2006.

[31] Merton R. C., Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case, The Review of Economics and Statistics, Vol. 51, No. 3, pp. 247-257, 1969.

[32] Musiela M., Stochastic PDEs and term structure models, Preprint, University of New South Wales, 1993.

[33] Musiela M. and Zariphopoulou T., Backward and forward utilities and the associated indifference pricing systems: The case study of the binomial model, Indifference Pricing, R. Carmona (ed.), Princeton University Press, 2005.

[34] Musiela M. and Zariphopoulou T., Portfolio choice under space-time monotone performance criteria, working paper, 2008.

[35] Neuberger A., Volatility Trading, London Business School working paper, 1990.

[36] Nualart D., The Malliavin Calculus and Related Topics, Springer, 2006.

[37] Potter C.W., Complete Stochastic Volatility Models With Variance Swaps, Working paper, 2004.

[38] Protter P., Stochastic Integration and Differential Equations, Second Edition, Springer, 2004.

[39] Reed M. and Simon B., Functional Analysis, Academic Press, 1980.

[40] Rogers L.C.G. and Williams D., Diffusions, Markov Processes and Martingales, Volume 1, Cambridge University Press, 2nd Edition, 2000.

[41] Schweizer M. and Wissel J., Arbitrage-free market models for option prices: the multi-strike case, Finance and Stochastics, Vol. 12, No. 4, pp. 469-505, 2008.

[42] Schweizer M. and Wissel J., Term structures of implied volatilities: absence of arbitrage and existence results, Mathematical Finance, Vol. 18, No. 1, pp. 77-114, 2008.

[43] Taflin E., Bond market completeness and attainable contingent claims, Finance and Stochastics, Vol. 9, No. 3, pp. 429-452, 2005.

[44] Widder D.V., The Heat Equation, Academic Press Inc., U.S., 1976.

[45] Widder D.V., The Role of the Appell Transformation in the Theory of Heat Conduction, Transactions of the American Mathematical Society, Vol. 109, No. 1, pp. 121-134., 1963.

[46] Williams D., Probability With Martingales, Cambridge University Press, 1991.

[47] Wissel J., Arbitrage-free market models for option prices, National Centre of Competence in Research Financial Valuation and Risk Management, Working paper no. 428, 2007.

[48] Yor M., Existence et unicite de diffusions a valeur dans un espace de Hilbert, Annales de l'I.H.P., section B, tome 10, No. 1, pp. 55-88, 1974.