



**HAL**  
open science

# Hedging contingent claims by convex local risk-minimization

Nicolas Millot

► **To cite this version:**

Nicolas Millot. Hedging contingent claims by convex local risk-minimization. Probability [math.PR]. Ecole Centrale Paris, 2012. English. NNT : . tel-00722225v1

**HAL Id: tel-00722225**

**<https://theses.hal.science/tel-00722225v1>**

Submitted on 1 Aug 2012 (v1), last revised 4 Jul 2014 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hedging Contingent Claims by Convex Local Risk-Minimization

Nicolas Millot  
MAS Department  
École Centrale Paris

A thesis submitted for the degree of  
*PhD in Applied Mathematics*

February 2012

---

À mes parents.

## Résumé

On s'intéresse dans cette thèse à la couverture des produits dérivés dans des marchés incomplets. L'approche choisie peut se voir comme une extension des travaux de M. Schweizer sur la minimisation locale du risque quadratique. En effet, tout en restant dans le cadre de la modélisation des actifs par des semimartingales, notre méthode consiste à remplacer le critère de risque quadratique par un critère de risque plus général, sous la forme d'une fonctionnelle convexe du coût local. Nous obtenons d'abord des résultats d'existence, d'unicité et de caractérisation des stratégies optimales dans un marché sans friction, en temps discret et en temps continu. Puis nous explicitons ces stratégies dans le cadre de modèles de diffusion avec et sans sauts. Nous étendons également notre méthode au cas où la liquidité n'est plus infinie. Enfin nous montrons par le biais de simulations numériques les effets du choix de la fonctionnelle de risque sur la constitution du portefeuille optimal.

## Abstract

This thesis deals with the issue of hedging contingent claims in incomplete markets. The way we tackle this issue may be seen as an extension of M. Schweizer's work on quadratic local risk-minimization. Indeed, while still modelling assets as semimartingales, our method relies on the introduction of a convex function of the local costs to assess risk, thus relaxing the quadratic assumption. The results we obtain are existence and uniqueness results first and characterizations of optimal strategies in a frictionless market, both in discrete and continuous time settings. We then make those strategies explicit by using diffusion models with and without jumps. We further extend our approach in the case when liquidity is given through a stochastic supply curve. Finally we show the effect of the choice of different risk functions on the optimal portfolio by numerically solving the optimality equations.

## Remerciements

Je tiens tout d'abord à exprimer ma sincère gratitude envers Frédéric Abergel pour m'avoir proposé ce sujet, je le remercie chaleureusement pour ses nombreux conseils fructueux et surtout pour ses encouragements, sa compréhension et sa gentillesse.

Un grand merci également à Lionel Gabet qui m'a tenu au courant des avancées de la création de la chaire de finance quantitative après que je lui ai fait part de mon désir d'effectuer une thèse.

Je remercie Bruno Bouchard et Denis Talay qui m'ont fait l'honneur d'accepter d'être les rapporteurs de cette thèse et j'adresse mes sincères remerciements aux membres du jury, Huyên Pham et Martin Schweizer.

Je remercie également le personnel et les doctorants du laboratoire MAS de l'École Centrale Paris pour leur accueil sympathique ; je remercie plus particulièrement Abhijeet Gaikwad, Aymen Jedidi, Olaf Torné et Riadh Zaatour pour leur aide à l'accomplissement de ce travail. Toute ma gratitude va aussi à Laurent Series grâce à qui j'ai pu avoir accès aux supercalculateurs du mésocentre, soirs et weekends compris !

Enfin, mes plus vifs remerciements vont à ma famille et à mes amis dont le soutien sans faille aura été plus que nécessaire pour mener à bien cette thèse.

# Contents

<b>Notations</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Pricing and Hedging Derivatives Products . . . . .	1
<b>2 Introduction</b>	<b>7</b>
2.1 Couverture et évaluation des produits dérivés . . . . .	7
<b>3 Quadratic Local Risk Minimization</b>	<b>19</b>
3.1 Discrete Time . . . . .	19
3.1.1 Definitions . . . . .	19
3.1.2 Local Risk-Minimization . . . . .	22
3.2 Continuous Time Setting . . . . .	26
3.2.1 Assumptions and Definitions . . . . .	26
3.2.2 Local Risk-Minimization . . . . .	28
3.2.3 Explicit Characterization of Locally Risk-Minimizing Strategies .	31
<b>4 Convex Local Risk Minimization</b>	<b>33</b>
4.1 Measuring Risk . . . . .	33
4.2 Discrete time . . . . .	35
4.2.1 Definitions and Assumptions . . . . .	35
4.2.2 Local $f$ -Risk Minimization . . . . .	35
4.3 Continuous Time Setting . . . . .	40
4.3.1 Definitions and Assumptions . . . . .	40
4.3.2 Local $f$ -Risk-Minimization . . . . .	42
4.3.3 The $f$ -Costs Process . . . . .	42



## CONTENTS

---

4.3.4	Pseudo-Optimal Strategies . . . . .	46
4.4	$g$ -Martingales and Orthogonality . . . . .	46
<b>5</b>	<b>Application to Stochastic Volatility Models</b>	<b>57</b>
5.1	Model Assumptions . . . . .	57
5.2	Quadratic PDE . . . . .	58
5.2.1	Existence and Uniqueness Results . . . . .	60
5.2.2	Complete markets case . . . . .	62
5.3	Quadratic FBSDE . . . . .	62
<b>6</b>	<b>Application to Jump Diffusion Models</b>	<b>65</b>
<b>7</b>	<b>Liquidity</b>	<b>69</b>
7.1	Motivations . . . . .	69
7.2	Liquidity costs and risk process . . . . .	69
7.3	Optimal and pseudo-optimal strategies . . . . .	72
7.4	Continuous time setting . . . . .	75
7.4.1	The $f$ -Costs Process (inclusive of liquidity costs) . . . . .	75
7.4.2	The supply price process . . . . .	79
7.5	Application to stochastic volatility models . . . . .	80
7.5.1	PDE formulation . . . . .	81
7.5.2	The minimization problem . . . . .	83
7.6	Application to stochastic volatility/jump diffusion models . . . . .	86
7.6.1	PIDE formulation . . . . .	87
<b>8</b>	<b>Numerical Results and Comparisons</b>	<b>91</b>
8.1	Motivations . . . . .	91
8.2	Benchmark Stochastic Volatility Models . . . . .	92
8.2.1	Solving the Quadratic Forward-Backward Stochastic Differential Equation . . . . .	93
8.2.2	Solving the Nonlinear Partial Differential Equation . . . . .	96
8.2.3	Discrete Time Approximation . . . . .	96
8.2.4	Numerical Results . . . . .	98
8.3	Mean Costs . . . . .	98

<b>References</b>	<b>101</b>
-------------------	------------

## CONTENTS

---

# Notations

Given a filtration  $(\mathcal{F}_t)$  the *conditional covariance* of two random variables  $W$  and  $Z$  with respect to a probability measure  $P$  is defined as

$$\text{cov}(W, Z|\mathcal{F}_t) := E[WZ|\mathcal{F}_t] - E[W|\mathcal{F}_t]E[Z|\mathcal{F}_t] \quad (1)$$

provided that the conditional expectations and their difference make sense. Likewise, we define the *conditional variance* of  $W$  under  $P$ :

$$\text{var}(W|\mathcal{F}_t) := E[W^2|\mathcal{F}_t] - E[W|\mathcal{F}_t]^2 \quad (2)$$

$$= \text{cov}(W, W|\mathcal{F}_t) \quad (3)$$

A sequence of processes  $(X^n)_{n \in \mathbb{N}}$  indexed by  $n$  is said to converge to  $X$  in *ucp* if, for each  $t > 0$ ,  $\sup_{0 \leq s \leq t} |X_s^n - X_s|$  converges to 0 in probability.

Given a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^1(\mathbb{R}^n, \mathbb{R})$ ,  $Df$  is the gradient of  $f$ , the vector of first order derivatives:

$$Df(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \quad (4)$$

If  $f \in C^2(\mathbb{R}^n, \mathbb{R})$ ,  $D^2f$  is the Hessian of  $f$ , the symmetric matrix of second order derivatives

$$D^2f(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \quad (5)$$

## 0. NOTATIONS

---

# 1

## Introduction

### 1.1 Pricing and Hedging Derivatives Products

This thesis is focusing on pricing and hedging contingent claims in incomplete markets. The emphasis will rather be put on the latter part which stands as essential once we discuss models and contingent claims which do not qualify for the complete markets paradigm. We recall that complete markets are markets where assets are modelled in such a way that contingent claims written on those assets and satisfying some integrability requirements can be exactly reproduced by trading in the underlying assets with self-financing strategies. In the usual terminology they are called “attainable” or “redundant”. The hypothesis to rely on self-financing strategies for qualifying as a redundant claim is essential in the pricing methodology which follows: self-financing means that there are no further inputs or withdrawals of money from the hedging portfolio, thus with the no-arbitrage assumption the price at which the contingent claim should be sold has to be the initial value of the hedging portfolio. The theory dates back to the seminal work of Black and Scholes (5) and Merton (41) who studied a particular type of market where two assets are traded: a risk-free asset representing the bank account and a risky asset modelled as a geometric Brownian motion with a drift. They show that this market is indeed complete and for that they rely on solving a PDE for the value of the hedging portfolio which once solved in turn gives the perfect delta hedging. Since their work in this setting, the theory of complete markets was thoroughly developed and given a sound mathematical background with the work of Harrison and Kreps (23) and Harrison and Pliska (24) who for that purpose introduced

## 1. INTRODUCTION

---

the notion of equivalent martingale measure, a measure that turns discounted assets into martingales. Their results identify complete markets as those markets which possess a single equivalent martingale measure. Their setting goes much beyond the initial settings of Black, Scholes and Merton since it includes rather general semimartingales. The pricing is then done through the computation of an expectation under the risk-free measure whereas the hedging strategy is obtained thanks to a representation formula. The most recent work in the area of qualifying complete markets is the paper of Delbaen and Schachermayer (16) who gives the precise no arbitrage condition (NFLVR: No Free Lunch with Vanishing Risk) so as to preclude dubious arbitrage strategies while allowing for the most general semimartingale models.

Yet going as far as the work of Mandelbrot (38, 39), evidences against the simpler models relying on Brownian motion, such as the Black, Scholes and Merton model (5, 41), accumulated and realistic models for describing the underlying assets do not qualify for complete markets. One such evidence is that the log returns of stock prices are not distributed normally but rather exhibit fat tails, which may be accounted for by the micro-structure nature of price formation as explained in Abergel et Al (10, 11). A number of models have then been proposed in the literature to account for this observation among which the most popular are the Heston model (27) which postulate that the volatility driving the Brownian motion of the stock is itself stochastic, and models using Levy processes as in Cont and Tankov (13). One common characteristic of these models is that they feature non hedgeable risk by adding sources of risk which cannot be traded with market assets. Hedging is therefore a much more involved task in these settings and it cannot be done unambiguously as was the case with complete markets. Same goes for pricing. One of the first works which address this question of hedging and pricing in incomplete markets is the paper of Föllmer and Sondermann (51). In order to protect against the intrinsic risk of a given contingent claim they proposed a sequential regression technique in a discrete time setting. Their approach is thus concerned with the backward minimization of the quadratic deviation between the option payout and its hedging portfolio. This approach was further extended in Föllmer and Schweizer (46) who gave results of existence of risk-minimizing strategies in a martingale setting, and was recently revisited by Cerny and Kallsen (9). It is noticeable that this approach relied on *mean self-financing* strategies, strategies which on average have zero costs but which are no longer self-financing.

## 1.1 Pricing and Hedging Derivatives Products

---

So as to remain within the self-financing world, Schweizer in (49) introduced the concept of mean-variance hedging which proposes to measure the riskiness of a strategy at inception by considering the quadratic deviation between the contingent claim value at expiry and the hedging portfolio. The procedure for finding optimal strategies then consists in minimizing the quadratic criterion among all suitable strategies. In contrast with the previous approach for which the minimization was carried out backward and sequentially in time, the optimal strategies are to be found at inception through a global minimization. Thus in essence the latter criterion is very close to the well-developed theory of stochastic optimal control. Using this theory gave rise to a number of other similar approach, still using self-financing strategies but turning to utility maximization instead of risk-minimization. Pricing is then achieved through indifference valuation, which means that the fair price of the contingent claim is the amount of money which leaves the utility unchanged from the situation where the portfolio manager would not have sold the product. In the usual case of exponential utility, this gives rise to a formulation in terms of forward backward stochastic differential equations with drivers of quadratic growth for which Imkeller, Reis and Zhang (28) is a good reference. We will see that our approach shares the same kind of formulation. Other approaches, though this is certainly not an exhaustive list, include risk-minimization using coherent risk measures, a thorough account of it which can be found in Barrieu and El Karoui (3), local utility maximization as in Kallsen (30), superhedging as originally introduced by Davis and Clark in (15), which provides bounds on prices so that the hedging portfolio always dominates the contingent claim.

Most of the time, each of these approaches turn to SDE (Stochastic Differential Equations) to model the underlying assets so that the natural tools for deriving hedging and pricing equations are PDE (Partial Differential Equations) and FBSDE (Forward Backward Stochastic Differential Equations). Usually these are solved numerically and for that purpose in high dimension the latter representation will be preferred. The closest approach to ours in this list, from the point of view of the techniques used, would be Kallsen's utility maximization. Actually it can be seen as an "orthogonal" methodology since instead of minimizing the risk of a trading strategy meant to produce the payoff of a given contingent claim, it proposes to maximize the utility of the gains of a trading strategy. To effectively deal with contingent claim pricing and hedging, the trading strategy should include one unit of the derivative which then produces its payoff



## 1. INTRODUCTION

---

at the terminal date. Derivatives prices are then determined through an equilibrium or neutral/indifferent pricing argument. Especially in this approach, the author uses a limiting process because he wants to maximize a local utility instead of the usual one, and we will see that we as well need a limiting process to define a special kind of optimality. We also need to insist on the fact that most of the theory, be it in complete or incomplete markets, relies on the use of stochastic integrals to model trading gains. This is the direct extension of the discrete time formula  $G_k = \sum_{j=1}^k \delta_j \Delta S_k$  which explains why Ito's formula and more generally tools from stochastic calculus have proved so useful in financial mathematics. This representation however does not really extend easily whenever imperfections in the market have to be taken into account. In 2004, Cetin, Jarrow and Protter (8) introduced liquidity costs in the theory of self-financing strategies and they derived an expression for the value  $V$  of a self-financing portfolio when trading on a stock which has a stochastic supply curve.

Having presented schematically the different approaches for pricing and hedging contingent claims in incomplete markets, we are now in a position to describe succinctly our methodology. For that purpose we step back a little and putting ourselves in the position of a trader who sold an option we make the simple remark that at the expiration of that option he will be bound to deliver the cash or asset corresponding to the contingent claim. If he were able to hedge against this unpredictable payment only once, at the initial date, then a natural way to build his portfolio is to consider the costs incurred at expiration date from adjusting the portfolio value to the contingent claim value. Since he is hedging, he would look for the initial composition of the portfolio that will leave him, on average, with the minimum costs at expiration date. Now the criterion he chooses to transform costs at expiration date into a risk function will have to weight, on average over all possible scenarios, losses and gains from adjusting the portfolio. Arguably the risk is more important if losses are suffered rather than gains. Yet because he is a trader and not a portfolio manager, he will bear another constraint in that the initial costs of setting up the portfolio should be minimal or at least as close as possible to the market price of the option, if there is any available. This naturally leads to choosing a function  $f$  of the costs which will be positive, convex, for the usual reason that we want to have  $\mathbb{E}(f(x)) > f(\mathbb{E}(x))$ , a way of mathematically specifying risk-aversion, null at zero, and asymmetric, favouring gains over losses, so with  $f(x) > f(-x)$  for  $x > 0$ . With just one hedging date the formulation obviously

does not differ whether we are performing a global minimization, meaning that we minimize the risk over all possible strategies from inception date to expiry date and local minimization when we minimize the risk over strategies which are perturbations of the optimal strategy at inception date only. Neither does the concept of local risk, being the risk due to costs incurred between two trading dates, and global risk which is the risk of the total costs accumulated from start date until end date.

In this thesis we will consider only local minimization of local risk. We chose local risk over global risk mainly for tractability reason and also because the approach then generalizes more easily to options of American type and to include some market inefficiencies such as liquidity costs. As well we chose local minimization because it is rather fruitful in terms of optimal strategies characterization, and also to avoid the time inconsistencies which would probably occur otherwise. We also think that this approach would yield more steady strategies in case of a change of regime in the market for instance.

In chapter 3 we give as an introduction to the mathematical setting and notations an overview of the main results that have been found by Schweizer (47, 48) since his initial thesis work on the hedging of options in general incomplete markets by means of quadratic hedging methods.

In chapter 4 we first introduce our method which generalizes the latter by considering a discrete time setting. This allows us to already give several equivalent characterizations of the optimal strategies which will prove useful when tackling the continuous time setting which is the object of the second part of the chapter.

In chapters 5 and 6 we exemplify our method in two usual settings in continuous time: the stochastic volatility model and the jump-diffusion model.

In chapter 7 we introduce liquidity costs in the modelling of the trading costs and go again through the characterization of optimal strategies in both discrete and continuous time settings.

Finally in chapter 8 we consider some specific Markovian models to describe the evolution of underlying assets and solve for optimal strategies by means of numerical methods for which we discuss the different schemes. This allows us to compare our approach with the quadratic framework of Schweizer.

## 1. INTRODUCTION

---

## 2

# Introduction

## 2.1 Couverture et évaluation des produits dérivés

Cette thèse s'intéresse à la problématique de la couverture des produits dérivés dans les marchés incomplets. La complétude des marchés est une notion introduite par Harrison et Kreps et c'est une propriété à la fois du modèle utilisé pour décrire l'évolution des actifs à risque et du produit dérivé que l'on cherche à couvrir. Cette notion exprime le fait que les produits contingents satisfaisant certaines hypothèses de régularité sont répliqués exactement par une stratégie d'achat/vente d'actifs à risque, sans apport ni retrait d'argent. Ces deux points sont essentiels pour résoudre du même coup la problématique de l'évaluation et de la couverture d'un produit contingent. La complétude des marchés est donc liée à la notion essentielle de stratégie auto-finançante, sans apport ni retrait d'argent, et elle fut d'abord exprimée dans un cadre discret puis généralisée au cas continu. L'outil mathématique pour parvenir à cette généralisation est l'intégrale stochastique d'Itô : en temps discret les gains de trading accumulés lors de l'application d'une stratégie auto-finançante sont donnés par  $G_k = \sum_{j=1}^k \delta_j (S_{j+1} - S_j)$  et par  $G_t = \int_0^t \delta_s dS_s$  en temps continu. On voit alors que si l'on dispose d'une telle stratégie auto-finançante pour répliquer un produit contingent, le prix du produit dérivé correspondant doit être la valeur initiale du portefeuille de couverture, par un simple raisonnement d'arbitrage, si le marché est à l'équilibre. Trouver une telle stratégie pour n'importe quel produit contingent, et donc montrer que le marché est complet, peut s'effectuer à l'aide de théorèmes de représentation ou par le biais de résolution d'équations aux dérivées partielles. Cette dernière approche fut celle employée par

## 2. INTRODUCTION

---

Black, Merton et Scholes pour l'évaluation et la couverture d'options européennes dans un marché où l'actif à risque est modélisé par un mouvement Brownien géométrique. Dans un marché incomplet, il existe par définition des produits contingents qui ne sont pas répliquables, ou, autrement dit, pour lesquels on ne peut pas éliminer totalement le risque en exécutant n'importe quelle stratégie de couverture auto-finançante. Le risque est une mesure de l'écart entre la valeur du produit contingent à l'échéance (pour un produit de type européen, le seul type de produits envisagé dans cette thèse) et la valeur du portefeuille de couverture. En effet, pour une stratégie auto-finançante, une fois le capital initial et la stratégie d'investissement en actif à risque décidés, il n'existe plus de degré de liberté pour ajuster la valeur du portefeuille au cours de la vie de l'option. Il peut donc être intéressant de relâcher l'hypothèse d'auto-financement afin de pouvoir utiliser d'autres critères d'optimalité. Si l'on ne travaille donc plus qu'avec des stratégies auto-finançantes, le coût de la stratégie n'est plus seulement déterminé à l'échéance de l'option mais peut être observé tout au long de sa vie. En fait le coût apparait alors comme un processus, non trivial (dans le cas d'une stratégie auto-finançante il est nul jusqu'à la maturité où le saut est alors appelé erreur de réplification), adapté à la filtration et qui dépend alors des deux composantes de la stratégie : la valeur du portefeuille et la quantité d'actifs à risque détenue. Des critères d'optimalité peuvent alors porter sur le coût local, c'est-à-dire le coût d'application de la stratégie entre deux dates de couverture, ou bien sur le coût global ou terminal, soit le coût d'application de la stratégie jusqu'à l'échéance. Dans le premier cas on s'intéresse donc au processus

$$\Delta C_k = \Delta V_k - \delta_k \Delta S_k$$

alors que dans le dernier cas, c'est le processus

$$C_k = \sum_{i \geq k} \Delta C_i$$

qu'on regarde. Dans nos travaux de thèse, les résultats obtenus se rapportent au premier cas : on minimise localement le risque local, qui est l'espérance conditionnelle d'une fonction convexe du coût local. Comme on le note dans le premier chapitre consacré à la récapitulation des résultats obtenus dans le cas quadratique par Schweizer, coût global et coût local sont équivalents dans le sens où ils donnent lieu aux mêmes stratégies optimales. Ce n'est plus vrai lorsque l'on s'intéresse à des fonctions convexes

## 2.1 Couverture et évaluation des produits dérivés

---

non quadratiques. Si l'on s'est également penché sur le cas du risque global, il nous a semblé plus fructueux de porter notre attention sur le coût local, celui-ci permettant notamment d'étendre naturellement les résultats obtenus par Schweizer et ce par le biais de techniques similaires.

On a donc commencé par regarder la formulation du problème en temps discret, ce qui fait l'objet de la première partie du deuxième chapitre de la thèse. Naturellement on obtient un programme de minimisation rétrograde puisque la stratégie optimale est connue à la date terminale du contrat. Par la propriété de convexité de la fonctionnelle de risque, on a alors le théorème (1) relatif à l'existence et l'unicité de la solution du programme de minimisation. On note par ailleurs qu'il faut que la condition de non-dégénérescence du processus de prix de l'actif risqué soit vérifiée pour obtenir l'unicité, ce qui était attendu au vue du même résultat dans le cas quadratique. Dans le but d'étendre les résultats de caractérisation des stratégies optimales au cas du temps continu, on reformule les conditions d'optimalité du premier ordre satisfaites par les deux composantes de la stratégie de la manière suivante : le processus des  $f$ -coûts  $\sum_{i \leq k} f'(\Delta C_i)$  est une martingale orthogonale au processus de prix  $S$ . En effet, en temps continu, d'une part il est crucial de bien définir la minimisation locale, en particulier par le choix de l'espace des perturbations admissibles, et d'autre part le choix du processus à minimiser n'est pas évident. En l'occurrence, une première approche de la minimisation locale du risque local en temps continu pourrait être de définir le risque local en temps continu comme la limite des incréments de risque  $\sum_{i \leq k} f(\Delta C_i)$  sur une partition tendant vers l'identité. Pourtant il est facile de voir que les stratégies optimales obtenues par cette approche, dans le cas où le processus de prix est continu, sont les mêmes que les stratégies optimales du cas quadratique et ne tiennent donc pas compte de l'asymétrie de la fonction  $f$ . En fait, comme s'attache à le démontrer la deuxième partie du deuxième chapitre, le bon processus à considérer pour un passage à la limite est bien le processus des  $f$ -coûts.

La deuxième partie du deuxième chapitre s'attache donc d'une part à définir précisément l'optimalité en temps continu et d'autre part à relier cette notion à celle de martingalité et d'orthogonalité à la limite du processus des  $f$ -coûts dans un cadre très général où le processus de prix est donné par une semimartingale quelconque. L'objet du théorème (2) est justement de donner un résultat d'existence de ce processus, et de le caractériser

## 2. INTRODUCTION

---

à partir du processus de prix et de la fonctionnelle de risque. L'expression obtenue est la suivante :

$$\begin{aligned} C_t^f(\phi) &= f''(0) \left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) \\ &+ \frac{f^{(3)}(0)}{2} \left( [V, V]_t^c - 2 \int_{0+}^t \delta_{s-} d[V, X]_s^c + \int_{0+}^t \delta_{s-}^2 d[X, X]_s^c \right) \\ &+ \sum_{0 < s \leq t} f'(\Delta V_s - \delta_{s-} \Delta X_s) - f''(0)(\Delta V_s - \delta_{s-} \Delta X_s) \end{aligned}$$

Ceci permet de d'introduire la notion de pseudo-optimalité, par analogie avec le temps discret et également de manière analogue à ce qui a été fait dans le cas quadratique. On qualifie donc de pseudo-optimale une stratégie dont le processus des  $f$ -coûts défini ci-dessus est une martingale orthogonale au processus de prix.

La question qui se pose alors naturellement est de savoir si l'on peut relier la notion de pseudo-optimalité à une notion d'optimalité en temps continu. On introduit alors une notion d'optimalité en temps continu. Pour ce faire on définit d'abord le coût  $C_t$  d'une stratégie  $\phi$  en temps continu:

$$C_t(\phi) := V_t(\phi) - \int_0^t \delta_u dX_u, \quad (0 \leq t \leq T)$$

Puis on introduit le risque local étant donnée une partition  $\tau = \{0 = t_0, t_1, \dots, t_k = T\}$  de l'intervalle  $[0, T]$  correspondant aux dates de couverture:

$$\Delta R_{t_i}(\phi) := \mathbb{E} \left( f(\Delta C_{t_{i+1}}(\phi)) \mid \mathcal{F}_{t_i} \right)$$

Et enfin la notion de petite perturbation : une stratégie de trading  $\phi = (\beta, \delta)$  bornée et telle que  $\beta_T = 0$  et  $\delta_T = 0$ .

Ces définitions en place on introduit le  $f$ -quotient de risque pour un produit contingent  $H$ , une stratégie de trading  $\phi$  simulant  $H$ , une partition  $\tau = \{0 = t_0, t_1, \dots, t_k = T\}$  de  $[0, T]$  et une petite perturbation  $\Delta$

$$r_f^\tau[\phi, \Delta](t, \omega) = \sum_{t_i, t_{i+1} \in \tau} \frac{\Delta R_{t_i}(\phi + \Delta|_{(t_i, t_{i+1}]})(\omega) - \Delta R_{t_i}(\phi)(\omega)}{t_{i+1} - t_i} \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

Et finalement on dira qu'une stratégie de trading  $\phi$  simulant  $H$  est optimale si pour toute petite perturbation  $\Delta$  n'importe quelle séquence de partitions  $(\tau_n)_{n \in \mathbb{N}}$  tendant vers l'identité, on a

$$\liminf_{n \rightarrow \infty} r_f^{\tau_n}[\phi, \Delta] \geq 0 \quad P - a.e.$$

## 2.1 Couverture et évaluation des produits dérivés

---

Afin d'étudier le lien entre optimalité et pseudo-optimalité, on introduit les notions de  $g$ -martingale et de  $g$ -martingale orthogonale à une autre martingale. Un processus adapté  $Y$  est une  $g$ -martingale (orthogonale à une martingale  $M$ ) s'il existe une martingale  $M^\perp$  (orthogonale à  $M$ ) telle que

$$Y = \sum_{j=1}^k g(\Delta M_j^\perp)$$

Ces définitions nous permettent en particulier d'étendre la notion de décomposition de Föllmer-Schweizer. En outre, moyennant une hypothèse sur l'espace de probabilité  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$  (existence d'une base de décomposition orthogonale de  $L^2(P)$  et régularité de la base), on obtient le théorème (6) qui montre qu'une stratégie pseudo-optimale est toujours optimale.

Les troisième et quatrième parties de la thèse sont consacrées à l'application des résultats théoriques de la deuxième partie dans le cadre des modèles à volatilité stochastique avec et sans sauts.

Le choix de ces modèles de marché incomplet est bien sûr lié à la popularité des modèles de Heston et de Bates pour expliquer les déviations observées entre la théorie de Black, Scholes et Merton et les données de marché pour les options liquides (smile/skew de volatilité).

On fait les hypothèses nécessaires et suffisantes pour obtenir existence et unicité du processus solution de l'équation différentielle stochastique régissant l'évolution de l'actif à risque qui, dans le cas de la volatilité stochastique, s'écrit

$$\begin{aligned} dX_s &= a(s, X_s, Y_s)ds + b(s, X_s, Y_s)dW_s^1 \\ dY_s &= c(s, X_s, Y_s)dt + d(s, X_s, Y_s) \left( \rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2 \right) \end{aligned}$$

où  $(W^1, W^2)$  est un processus de Wiener standard bidimensionnel sous la mesure  $P$ . Ceci nous permet de nous placer dans un cadre Markovien pour les stratégies optimales et donc de les chercher sous la forme paramétrique suivante

$$\begin{cases} \delta_t &= \delta(t, X_t, Y_t) \\ V_t &= V(t, X_t, Y_t) \end{cases}$$

où  $X$  est l'actif à risque actualisé et  $Y$  est la volatilité de la diffusion associée.

Dès lors on peut obtenir une expression du processus des  $f$ -coûts en fonction de la stratégie et en appliquant les critères de pseudo-optimalité on arrive aux équations



## 2. INTRODUCTION

---

aux dérivées partielles satisfaites par la stratégie optimale. Dans le cas de la volatilité stochastique, on obtient l'EDP quadratique suivante

$$\frac{\partial V}{\partial u} + \Lambda V = \frac{a}{b} \left( \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho d \right) + \alpha \left( \sqrt{1 - \rho^2} \frac{\partial V}{\partial Y} d \right)^2$$

où  $\Lambda$  est le générateur infinitésimal de la diffusion de paramètres  $a$ ,  $b$ ,  $c$  et  $d$ , et  $\alpha = -\frac{1}{2} \frac{f^{(3)}(0)}{f''(0)}$ . Le ratio de couverture  $\delta$  vérifie l'équation

$$\delta b = \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho$$

L'existence et l'unicité sont obtenus pour l'EDP quadratique au sens des solutions de viscosité grâce à des résultats acquis dans un cadre plus général. Pour obtenir un résultat plus fort sur l'optimalité de la solution donnée par l'EDP, on utilise la caractérisation de la solution optimale par une équation différentielle stochastique rétrograde

$$\begin{aligned} -dV_s &= g(s, X_s, Y_s, V_s, Z_s^1, Z_s^2) ds - Z_s^1 dW_s^1 - Z_s^2 dW_s^2 \\ V_T &= H \end{aligned}$$

avec  $g(s, X, Y, Z^1, Z^2) = -\frac{a}{b} Z^1 - \alpha (Z^2)^2$ , with  $V = \beta + \delta X$  et  $Z = (\delta b, \frac{\partial V}{\partial Y} d \sqrt{1 - \rho^2})$ . On sait alors que la solution de cette EDSR donne une solution de viscosité pour l'EDP quadratique. Or l'EDSR est également la formulation de l'optimalité par la  $g$ -martingalité.

Dans le cas où l'on ne considère plus seulement une évolution de l'actif à risque avec des trajectoires continues, on utilise la modélisation suivante

$$\begin{aligned} dX_s &= a(s, X_{s-}, Y_{s-}) ds + b(s, X_{s-}, Y_{s-}) dW_s^1 + k dN_s \\ dY_s &= c(s, X_{s-}, Y_{s-}) ds + d(s, X_{s-}, Y_{s-}) \left( \rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2 \right) \end{aligned}$$

où  $(W^1, W^2)$  est un processus de Wiener standard bidimensionnel sous la mesure  $P$  et  $N_t$  est un processus de Poisson d'intensité  $\lambda$  et l'amplitude des sauts associés  $k$  a une densité de probabilité  $K$ .

On obtient moins de résultat théorique dans ce cadre, mais on est au moins capable, en se plaçant toujours dans un cadre Markovien, d'écrire les équations de pseudo-optimalité.

## 2.1 Couverture et évaluation des produits dérivés

---

La condition de martingalité du processus des  $f$ -coûts donne l'EIDP suivante

$$\begin{aligned}
 f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_{u-} a_u \right) \\
 + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
 - f^{(3)}(0) \delta_{u-} \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_{u-}^2 b_u^2 \\
 + \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u) K(k) dk \lambda_u = 0
 \end{aligned}$$

avec la condition terminale  $V_T = H$ .

Tandis que la condition d'orthogonalité nous permet de relier le ratio de couverture optimale à la valeur du portefeuille

$$\left( \frac{\partial V}{\partial X} - \delta_{u-} \right) b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u + \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u) k K(k) dk \lambda_u = 0$$

La cinquième partie de la thèse s'attache à étendre les résultats des trois parties précédentes au cadre plus général de la liquidité finie, ou plus précisément au cas où le prix d'achat ou de vente de l'actif à risque dépend de la quantité.

On revient donc au cas discret afin d'examiner la notion d'optimalité qu'il est possible d'obtenir dans ce cadre. On s'intéresse toujours à minimiser séquentiellement une fonction des coûts de couverture, ces derniers prenant la forme suivante

$$\Delta C_k(\phi) = \mathcal{L}((\delta_{k+1} - \delta_k), X_k, t_k) + (\beta_k - \beta_{k-1}) \quad \forall k \in \{k = 1, \dots, T\}$$

où la fonction  $\mathcal{L}$  donne les coûts liés à l'ajustement de la stratégie sur la partie actif à risque. C'est cette fonction qui a pour but de modéliser les effets de liquidité. Elle a donc certaines propriétés et est en particulier strictement croissante et convexe. La fonction à minimiser s'exprime alors comme l'espérance conditionnelle des coûts

$$\Delta R_k^f(\phi) = \mathbb{E}_k(f(\Delta C_{k+1}))$$

On a toujours l'existence d'une solution au programme de minimisation mais non son unicité. Une stratégie optimale doit vérifier les conditions du premier ordre qui se traduisent par le système d'équations suivant

$$\mathbb{E}_k(f'(\Delta C_{k+1}(\phi^*))) = 0 \tag{2.1}$$

$$\mathbb{E}_k(f'(\Delta C_{k+1}(\phi^*))) l(\delta_{k+2} - \delta_{k+1}) X_{k+1} = 0 \tag{2.2}$$

## 2. INTRODUCTION

---

où  $l$  est la dérivée partielle de la fonction de liquidité  $\mathcal{L}$  par rapport à sa première variable (la quantité d'actifs).

Afin d'étendre la notion de minimisation locale du risque au temps continu, de mme que pour le cas de la liquidité infinie, on réinterprète les équations d'optimalité en terme d'orthogonalité de processus. Pour ce faire on définit le processus des  $f$ -coûts  $(C_k^f)_k$  par  $C_k^f = \sum_{i=1}^k f'(\Delta C_i)$ ,  $C_0^f = 0$ , et le processus de prix ajusté de la liquidité  $(X_k^S)_k$  par  $X_k^S = X_0 + \sum_{i=1}^k (l(\Delta \delta_i) X_i - l(0) X_{i-1}) = X_0 + \sum_{i=1}^k (l(\Delta \delta_i) X_i - X_{i-1})$ ,  $X_0^S = X_0$ . Ainsi on dira qu'une stratégie est pseudo-optimale si le processus  $C^f$  est une martingale fortement orthogonale à la partie martingale du processus  $X^S$ .

Pour le cas continu on s'attache tout d'abord à étudier l'existence des processus limites définis ci-dessus dans un cadre général où l'actif à risque est une semimartingale.

On obtient des résultats d'existence et des caractérisations explicites des deux processus  $C^f$  et  $X^S$  de manière très similaire au cas de la liquidité infinie.

Le processus des  $f$ -coûts s'écrit

$$\begin{aligned} C_t^f(\phi) &= f''(0) \left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) + f''(0) l'(0) \left( \frac{1}{2} \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \right) \\ &+ \frac{f^{(3)}(0)}{2} \left( [V, V]_t^c - 2 \int_{0+}^t \delta_{s-} d[V, X]_s^c + \int_{0+}^t \delta_{s-}^2 d[X, X]_t^c \right) \\ &+ \sum_{0 < s \leq t} f'(\Delta V_s - \delta_{s-} \Delta X_s + \mathcal{L}(\Delta \delta_s, X_s) - \Delta \delta_s X_s) \\ &- \sum_{0 < s \leq t} f''(0) (\Delta V_s - \delta_{s-} \Delta X_s) \end{aligned}$$

alors que le processus de prix ajusté de la liquidité s'écrit

$$\begin{aligned} X_t^S(\phi) &= X_t + l'(0) \left( \delta_t X_t - \delta_0 X_0 - \int_{0+}^t \delta_{s-} dX_s \right) + \frac{1}{2} l''(0) \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \\ &+ \sum_{0 < s \leq t} (l((\Delta \delta_s) - 1) X_s - l'(0) \Delta \delta_s X_s) \end{aligned}$$

On applique d'abord les résultats qui précèdent aux cas de la volatilité stochastique avec et sans sauts afin d'obtenir une caractérisation des stratégies pseudo-optimales.

Avec les mmes hypothèses sur le processus d'actif à risque que dans le cas de la liquidité infinie, on obtient que les composantes d'une stratégie pseudo-optimale doivent

## 2.1 Couverture et évaluation des produits dérivés

---

satisfaire le système d'EDP parabolique et hyperbolique suivant

$$\left\{ \begin{array}{l} \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a + \frac{\partial V}{\partial Y} b + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b d = \\ \delta a + \alpha \left( \left( \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho d - \delta b \right)^2 + (1 - \rho^2) \left( \frac{\partial V}{\partial Y} \right)^2 d^2 \right) \\ \quad + l'(0) \frac{X}{2} \left( \left( \frac{\partial \delta}{\partial X} b + \frac{\partial \delta}{\partial Y} \rho d \right)^2 + (1 - \rho^2) \left( \frac{\partial \delta}{\partial Y} \right)^2 d^2 \right) \\ \left( \frac{\partial V}{\partial X} - \delta \right) \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) b^2 + \frac{\partial V}{\partial Y} \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) \rho b d \\ \quad + \left( \frac{\partial V}{\partial X} - \delta \right) \frac{\partial \delta}{\partial Y} l'(0) X \rho b d + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0) X d^2 = 0 \end{array} \right.$$

avec  $V_T = H$ . Toujours dans ce cadre de volatilité stochastique sans sauts, on parvient à montrer l'équivalence entre pseudo-optimalité et optimalité, lorsque l'optimalité est définie en ne considérant que des stratégies de type "buy and hold", c'est à dire des stratégies constantes sur les intervalles sur lesquels on mesure le risque local. Le passage à la limite en temps permet donc d'obtenir un résultat de caractérisation plus fort que dans le cas du temps discret.

En revanche on n'obtient pas de résultat de ce type pour le cas des processus à trajectoires discontinues et on donne donc simplement la caractérisation due au critère de pseudo-optimalité. Le ratio de couverture  $\delta$  est solution de l'équation non-linéaire suivante

$$\begin{aligned} X_t^S(\phi) - \mathbb{E}(X_t^S(\phi)) &= \int_0^t \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) b_u dW_u^1 + \int_0^t l'(0) X \frac{\partial \delta}{\partial Y} d_u dW_u^2 \\ &\quad + \int_0^t \int_{\mathbb{R}} ((l(\Delta \delta_u) - 1) X_u + k) K(k) dk d\tilde{N}_u \end{aligned}$$

où  $\tilde{N}$  est le processus de Poisson compensé de  $N$ . La valeur théorique du portefeuille  $V$  est solution de l'EIDP suivante

$$\begin{aligned} &\left( \frac{\partial V}{\partial X} - \delta \right) \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) b^2 + \frac{\partial V}{\partial Y} \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) \rho b d \\ &\quad + \left( \frac{\partial V}{\partial X} - \delta \right) \frac{\partial \delta}{\partial Y} l'(0) X \rho b d + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0) X d^2 \\ &+ \int_{\mathbb{R}} f'(\Delta V_u - \delta_u - \Delta X_u + \mathcal{L}(\Delta \delta_u, X_u) - \Delta \delta_u X_u) ((l(\Delta \delta_u) - 1) X_u + k) K(k) dk \lambda_u = 0 \end{aligned}$$

Tout comme dans le cas volatilité stochastique pure, les résultats d'existence et d'unicité pour ces systèmes d'équations aux dérivées partielles couplées restent à établir.

## 2. INTRODUCTION

---

Enfin la dernière partie de la thèse aborde la question de la résolution numérique de quelques unes des équations obtenues dans les parties théoriques précédentes. On compare également les approximations obtenues avec la résolution numérique du problème en temps discret.

Pour ces études, on se place dans un cadre de volatilité stochastique où les équations d'évolution de l'actif à risque sont

$$\begin{aligned}\frac{dX_t}{X_t} &= \mu(t, Y_t)dt + Y_t dW_t \\ dY_t &= a(t, Y_t)dt + b(t, Y_t)dW'_t\end{aligned}$$

Les paramètres choisis sont

Modèle	Volatilité stochastique $Y$	Taux de rendement $\mu$
Stein	$dY_t = \kappa(\theta - Y_t)dt + k dW_t$	$\mu(t, Y_t) = \Delta Y_t$
Heston	$d(Y_t)^2 = \kappa(\theta - (Y_t)^2)dt + \Sigma Y_t dW'_t$	$\mu(t, Y_t) = \Delta Y_t$

(2.3)

On s'intéresse dans un premier temps à la résolution de l'équation différentielle stochastique rétrograde suivante

$$\begin{aligned}\frac{dX_t}{X_t} &= \mu(t, Y_t)dt + Y_t dW_t^1 \\ dY_t &= a(t, Y_t)dt + b(t, Y_t)(\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2) \\ -dV_s &= g(s, X_s, Y_s, V_s, Z_s^1, Z_s^2)ds - Z_s^1 dW_s^1 - Z_s^2 dW_s^2 \\ V_T &= h(X_T)\end{aligned}$$

avec  $W = (W^1, W^2)$  un mouvement Brownien bi-dimensionnel standard et  $g(s, S, \sigma, Y, Z^1, Z^2) = -\frac{\mu}{\sigma} Z^1 - \alpha(Z^2)^2$ .

A cause de la croissance quadratique du générateur  $g$ , on doit utiliser un schéma de troncature. On a alors la convergence du schéma de troncature sous certaines conditions vérifiées dans le cadre de nos hypothèses de diffusion. On applique ensuite à l'EDSR dont le générateur auquel est appliquée la troncature est alors Lipschizien une méthode de résolution numérique basée sur des régressions.

Dans un deuxième temps on cherche à résoudre numériquement l'EDP quadratique associée

$$\frac{\partial V}{\partial u} + \Lambda V = \frac{\mu}{Y} \left( \frac{\partial V}{\partial X} XY + \frac{\partial V}{\partial Y} \rho b \right) + \alpha \left( \sqrt{1 - \rho^2} \frac{\partial V}{\partial Y} b \right)^2$$

## 2.1 Couverture et évaluation des produits dérivés

---

avec la condition terminale  $V_T = H(X_T) = (X_T - K)_+$  et le générateur correspondant à nos hypothèses de diffusion

$$\Delta V_u = \frac{\partial V}{\partial X} \mu_u X + \frac{\partial V}{\partial Y} a_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} X^2 Y^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} b_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u X Y$$

On utilise pour sa résolution un code d'éléments finis et l'on choisit d'appliquer les conditions de Dirichlet  $V(X_{min}) = K - X_{min}$  et  $V(X_{max}) = 0$  sur une grille assez large en  $X$ , et des conditions de frontière libre en  $Y$ .

Enfin on compare avec la résolution du problème en temps discret. Pour cette résolution, discrétise d'abord l'EDS afin de simuler l'actif à risque et la volatilité aux dates de couverture. Puis on approxime la fonction de risque  $f$  par un développement polynomial à l'ordre 3. Ceci permet d'écrire les équations d'optimalité du premier ordre sous forme d'un système d'équations polynomiales en les composantes de la stratégie où les coefficients sont des espérances conditionnelles qui ne dépendent que de valeurs connues. On calcule donc ces espérances conditionnelles par le biais de régressions, ce qui permet alors de résoudre simplement le système d'équations polynomiales.

On peut alors comparer nos trois méthodes et constater la bonne convergence pour différents types d'options européennes.

## 2. INTRODUCTION

---

# 3

## Quadratic Local Risk Minimization

In this chapter we recall the main results obtained by Schweizer (47, 48) and Föllmer and Schweizer (21) for the method of quadratic hedging with local risk-minimization. This allows us to introduce our notations for the classical problem of hedging contingent claims in incomplete markets. We start this review with the discrete time case before moving onto the continuous time setting.

### 3.1 Discrete Time

#### 3.1.1 Definitions

Let  $X = (X_t)_{t=0, \dots, T}$  be a stochastic process defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0, \dots, T})$ .  $T \in \mathbb{N}$  is a fixed and finite time horizon.  $(\mathcal{F}_t)_{0 \leq t \leq T}$  is assumed to satisfy the usual hypothesis, meaning that  $\mathcal{F}_0$  is complete. We also assume that  $\mathcal{F}_0$  is trivial, *i.e.*  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ , so that random variables adapted to  $\mathcal{F}_0$  are treated like constants. The process  $X$  describes the price evolution of a risky asset and as we assume the existence of a (locally) risk-free asset (the bank account), strictly positive at all time.  $X$  will actually stand for the discounted price process. The bank account, discounted, is then worth 1 at all time. Furthermore we assume that  $X$  is adapted (*i.e.*  $X_k$  is  $\mathcal{F}_k$ -measurable) and that it is a square-integrable process (*i.e.*  $X_k \in L^2(P) \quad \forall k \in \{0, \dots, T\}$ ). We use the notation  $\Delta X_k := X_k - X_{k-1}$  for  $k \in \{1, \dots, T\}$ .

With these two assets, and with the aim of hedging a claim contingent on the value



### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

of  $X$ , we build a portfolio consisting of  $\delta$  shares of the risky asset and  $\beta$  shares of the risk-free one. This consideration leads to the following definition:

**Definition 1.** A trading strategy  $\phi$  is a pair of two stochastic processes  $(\beta, \delta)$  such that

$$\beta = (\beta_k)_{k=0, \dots, T} \text{ is adapted.} \quad (3.1)$$

$$\delta = (\delta_k)_{k=1, \dots, T} \text{ is a predictable process.} \quad (3.2)$$

$$\sum_{j=1}^k \delta_j \Delta X_j \in L^2(P) \text{ for } k \in \{1, \dots, T\}. \quad (3.3)$$

$$\delta_k X_k + \beta_k \in L^2(P) \text{ for } k \in \{1, \dots, T\}. \quad (3.4)$$

**Definition 2.** The (discounted) value process  $V$  of  $\phi = (\beta, \delta)$  is defined as

$$V_0 := \beta_0$$

$$\text{and } V_k := \beta_k + \delta_k \cdot X_k \text{ for } k \in \{1, \dots, T\}$$

For a trading strategy, condition (3.4) then says that the value process  $V$  has to be square-integrable.

With these definitions, the trading strategy of the trader is thus interpreted as choosing  $\beta_k$  at time  $k$  after having observed the value  $X_k$ .  $\delta_k$  was chosen in the previous period. Knowing  $\beta$  at time  $k$  is thus equivalent to knowing  $V$ . Therefore with this trading strategy the trader is assured to meet the contingent claim requirements at time  $T$  by adjusting only  $\beta_T$ .

**Definition 3.** For a trading strategy  $\phi = (\beta, \delta)$ , the (cumulative) gains process  $G$  accumulated up to time  $k$  by investing into the risky asset is given by

$$G_k := \sum_{j=1}^k \delta_j \cdot \Delta X_j, \quad k = 1, \dots, T \text{ and } G_0 := 0.$$

The (cumulative) costs process is then defined as the difference

$$C_k := V_k - G_k, \quad k = 0, \dots, T \quad (3.5)$$

between the value process  $V$  and the gains process  $G$ .

We now introduce the definition of a contingent claim, which will be considered throughout this thesis of European type.

**Definition 4.** A contingent claim  $H$  is an  $\mathcal{F}_T$ -adapted square-integrable random variable, i.e.  $H \in L^2(P)$ .

A classical example is a European call option of strike  $K$  which is represented by  $H = (X_T - K)^+$ . We will only focus on strategies which can replicate the contingent claim at terminal time  $T$ , we thus have the following definition of admissibility:

**Definition 5.** A trading strategy  $\phi$  is called  $H$ -admissible if

$$V_T(\phi) = H \text{ } P - a.s.$$

Since we always want to replicate our contingent claim, we had to relax the predictability assumption on the whole strategy  $\phi$  by allowing  $\beta$  to be adapted while keeping  $\delta$  predictable. But then an obvious  $H$ -admissible strategy would be to do nothing until the terminal date and then let  $\beta_T = H$ . In order to preclude such strategies Schweizer in (47) after Föllmer and Sondermann (51) introduced a criterion based on the costs from trading so that while we do not restrict ourselves to self-financing strategies, reasonable strategies will still have to satisfy the weaker property of being *mean self-financing*.

**Definition 6.** A trading strategy  $\phi$  is called mean self-financing if its costs process  $C(\phi)$  is a square-integrable martingale.

Lemma I.1 of Schweizer (47) then shows that there is a bijective correspondence between the set of all mean self-financing  $H$ -admissible trading strategies and the set of all predictable process  $\delta$  satisfying (3.3). It is given by  $\delta \mapsto \phi = (\delta, \beta)$  with

$$\beta_k := \mathbb{E} \left[ H - \sum_{j=k+1}^T \delta_j \Delta X_j \middle| \mathcal{F}_k \right] - \delta_k X_k.$$

**Remark 1.** This bijection will prove a major ingredient for the derivation of optimal strategies once it has been proven that we can indeed focus only on those mean self-financing strategies. In that respect it proves that the approach used by Schweizer is very much an extension of the traditional self-financing framework in complete markets. We will also see in the next chapter that the absence of this bijection will slightly complicate things for a non-quadratic measure of risk.

### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

#### 3.1.2 Local Risk-Minimization

The aim is to exhibit strategies which bear minimum risks in some sense while still allowing the perfect replication of our contingent claim  $H$ . The measure of riskiness proposed by Schweizer is defined through this definition of the risk process:

**Definition 7.** *The (global) risk process  $R(\phi)$  of a trading strategy  $\phi$  is defined by*

$$R_k(\phi) := \mathbb{E} \left[ (C_T(\phi) - C_k(\phi))^2 \middle| \mathcal{F}_k \right], \quad k \in \{0, \dots, T\}.$$

The previous remark allows one to directly associate a risk process to every predictable process  $\delta$  satisfying (3.3) for a given contingent claim  $H$ . So as to select strategies which minimize the risk  $R$  over all  $H$ -admissible strategies, we have to introduce a notion of perturbations. Schweizer chose local perturbations over global perturbations since for the latter a (global) risk-minimizing strategies might just fail to exist (see example at the end of section I.2 in Schweizer (47)). This calls for the following definition:

**Definition 8.** *Let  $\phi$  be a trading strategy and  $k$  a trading date. An admissible local variation of  $\phi$  at  $k$  is a trading strategy  $\Delta = (\xi, \eta)$  such that*

$$\begin{aligned} \xi_j &= 0 \text{ for } 1 \leq j \leq k \text{ and } k+2 \leq j \leq T \\ \eta_j &= 0 \text{ for } 0 \leq j \leq k-1 \text{ and } k+1 \leq j \leq T-1 \end{aligned}$$

and

$$V_T(\phi + \Delta) = V_T(\phi) \quad P - a.s. \quad (3.6)$$

This concept of local perturbations amounts to consider varying only  $\delta_{k+1}$  and  $\beta_k$  while leaving the rest of the strategy  $\phi$  unchanged. (3.6) then determines  $\beta_T$  from  $\delta_T$ :  $\beta_T = -\delta_T X_T$  so that  $\Delta = 0$  for  $k = T$ . For all other  $k < T$ ,  $\Delta$  is uniquely determined by  $\delta_{k+1}$  and  $\beta_k$ . With this concept in place, we can introduce the definition of a locally risk-minimizing strategy:

**Definition 9.** *A trading strategy  $\phi$  is called locally risk-minimizing if for any trading date  $k$  and any admissible local variation  $\Delta$  of  $\phi$  at  $k$  we have*

$$R_k(\phi + \Delta) - R_k(\phi) \geq 0 \quad P - a.s.$$

Local risk-minimization corresponds to a backward sequential regression algorithm: since admissible local variations vary only the components of the strategy at the considered time  $k$  and together with the terminal condition which enforce that the perturbed  $H$ -admissible strategies remain  $H$ -admissible, we see that we have to start from the terminal date  $T$  when  $V_T$  is known and then proceed backward with the minimization. Now that we have introduced most notations, we give the results Schweizer got in his PhD thesis (47).

**Lemma 1.** *Let  $\phi$  be a trading strategy. If  $\phi$  is locally risk-minimizing, then  $\phi$  is mean self-financing.*

*Proof.* See lemma I.7 of Schweizer (47). The proof relies on constructing an admissible local variation which varies only the risk-less component by adding the conditional expectation of the future costs of the strategy. It is then readily seen that such a local variation would decrease the risk, unless the costs process of the strategy is a martingale.  $\square$

To give more insights on the structure of the strategy we need the following definition:

**Definition 10.** *Two adapted processes  $U$  and  $Y$  are called strongly orthogonal with respect to  $P$  if the conditional covariances*

$$\text{cov}(U_{t+1} - U_t, Y_{t+1} - Y_t | \mathcal{F}_t), \quad t = 0, \dots, T - 1$$

*are well-defined and vanish  $P$ -almost surely.*

In the particular case where either  $U$  or  $Y$  is a  $P$ -martingale, their conditional covariance reduces to

$$\text{cov}(U_{t+1} - U_t, Y_{t+1} - Y_t | \mathcal{F}_t) = E[(U_{t+1} - U_t)(Y_{t+1} - Y_t) | \mathcal{F}_t]$$

The next proposition gives the procedure to actually find the locally risk-minimizing strategy associated with a contingent claim  $H$ . We make use of the following notation

$$\sigma_k = \sqrt{\text{Var}_{\mathcal{F}_{k-1}}(\Delta X_k)}.$$

**Proposition 1.** *Let  $H$  be a contingent claim and  $\phi = (\delta, \beta)$  an  $H$ -admissible trading strategy. The following statements are equivalent*

1.  $\phi$  is locally risk-minimizing.

### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

2.  $\phi$  is mean self-financing and

$$\text{Cov}_{\mathcal{F}_{k-1}}(\Delta C_k(\phi), \Delta X_k) = 0 \text{ } P - a.s., 1 \leq k \leq T$$

3.  $\phi$  is given by

$$\delta_k = \frac{\text{Cov}_{\mathcal{F}_{k-1}}\left(H - \sum_{j=k+1}^T \delta_j \Delta X_j, \Delta X_k\right)}{\text{Var}_{\mathcal{F}_{k-1}}(\Delta X_k)} \cdot I_{\sigma_k \neq 0}, 1 \leq k \leq T \quad (3.7)$$

and

$$\beta_k = \mathbb{E} \left[ H - \sum_{j=k+1}^T \delta_j \Delta X_j | \mathcal{F}_k \right] - \delta_k X_k, 0 \leq k \leq T \quad (3.8)$$

*Proof.* See proposition 8 of Schweizer (47). The proof relies on a backward induction together with classical result on optimal linear prediction.  $\square$

**Remark 2.** Originally, the criterion used is a global criterion as is noticed from the definition. It was however already noted in Schweizer (47) that it is equivalent with using the following local criterion:

**Definition 11.** The (local) risk process  $R^l(\phi)$  of a trading strategy  $\phi$  is defined by

$$R_k(\phi) := \mathbb{E} \left[ (C_{k+1}(\phi) - C_k(\phi))^2 | \mathcal{F}_k \right], k \in \{0, \dots, T-1\}.$$

To see that the local criterion is equivalent to the global one, a backward induction shows that the optimal solution is given by

$$V_k = \mathbb{E} [V_{k+1} - \delta_{k+1} \Delta X_{k+1} | \mathcal{F}_k]$$

$$\delta_{k+1} = \frac{\text{Cov}_{\mathcal{F}_k}(V_{k+1}, \Delta X_{k+1})}{\text{Var}_{\mathcal{F}_k}(\Delta X_{k+1})} \cdot I_{\sigma_{k+1} \neq 0}$$

But then from the definition of  $C_k$  this is seen to be equivalent to

$$\mathbb{E}[\Delta C_{k+1} | \mathcal{F}_k] = 0$$

$$\text{Cov}_{\mathcal{F}_k}(\Delta C_{k+1}, \Delta X_{k+1}) = 0$$

which just says that  $(C_k)$  is a martingale strongly orthogonal to the martingale part of  $(X_k)$ .

This result shows how to construct a candidate for the locally risk-minimizing strategy but assumes that  $\phi$  is an  $H$ -admissible strategy. Schweizer gives a sufficient condition on  $X$  so that the candidate defined only by relations (3.7) and (3.8) is indeed admissible. It further assumes that there exists  $c \in \mathbb{R}$  with  $0 \leq c < 1$  so that

$$(\mathbb{E} [\Delta X_k | \mathcal{F}_{k-1}])^2 \leq c \mathbb{E} [\Delta X_k^2 | \mathcal{F}_{k-1}] \quad P - a.s., \quad 1 \leq k \leq T$$

This condition is essentially a non-degeneracy condition for the martingale part of  $X$  since using a Doob decomposition of  $X = M + A$  with  $M$  a martingale and  $A$  a predictable process, it can be rephrased  $\mathbb{E} [\Delta M_k^2 | \mathcal{F}_{k-1}] \geq K \Delta A_k^2$  for  $K > 0$ . This condition is also better known as the fact that  $X$  has a *bounded mean-variance trade-off*, the *mean-variance process* of  $X$  being defined as

$$\sum_{j=1}^k \frac{(\mathbb{E} [\Delta X_j | \mathcal{F}_{j-1}])^2}{\text{Var}_{\mathcal{F}_{j-1}}(\Delta X_j)}$$

The following result characterizes the existence of locally risk-minimizing strategies in terms of a decomposition of the claim  $H$ .

**Corollary 1.** *There exists a locally risk-minimizing strategy if and only if  $H$  admits a decomposition*

$$H = c + \sum_{j=1}^T \delta_j \cdot \Delta X_j + L_T \quad P - a.s.,$$

where  $c$  is a constant,  $\delta$  is a predictable process such that

$$\delta_j \cdot \Delta X_j \in L^2(P) \quad \text{for all } j,$$

and where  $L$  is a square integrable  $P$ -martingale which is strongly orthogonal to  $X$  and satisfies  $L_0 = 0$ . In this case, the locally risk-minimizing strategy  $(\hat{\beta}, \hat{\delta})$  is given by  $\hat{\delta} = \delta$  and by the adapted process  $\hat{\beta}$  defined by  $\hat{\beta}_0 = c$  and

$$\hat{\beta}_k = c + \sum_{j=1}^k \delta_j \cdot \Delta X_j + L_k - \delta_j \cdot X_j, \quad j = 1, \dots, T.$$

Moreover the decomposition is unique in the sense that the constant  $c$  and the martingale  $L$  are uniquely determined.

*Proof.* See remark at the end of chapter I of Schweizer (47) and corollary 10.14 of Föllmer and Schied (20). □

### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

A decomposition of this form will be called the *orthogonal decomposition* of the contingent claim  $H$  with respect to the process  $X$ . If  $X$  is itself a  $P$ -martingale, then the orthogonal decomposition reduces to the *Kunita-Watanabe decomposition*. We next move to transferring the same local risk-minimization ideas to the continuous time setting.

## 3.2 Continuous Time Setting

### 3.2.1 Assumptions and Definitions

We work with a probability space  $(\Omega, \mathcal{F}, P)$  now equipped with a continuous time filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$  where  $T \in \mathbb{R}$  is still a fixed and finite time horizon. As usual we assume that  $(\mathcal{F}_t)$  satisfies the usual conditions of right-continuity and completeness. We also assume that  $\mathcal{F}_0$  is trivial and that  $\mathcal{F}_T = \mathcal{F}$ .  $X$  is still the discounted price process of the risky asset and is modelled now as a càdlàg (right continuous with limits from the left) adapted process. Furthermore we assume that

$$X = (X_t)_{(0 \leq t \leq T)} \text{ is a semimartingale with a decomposition } X = X_0 + M + A$$

such that

$$M = (M_t)_{(0 \leq t \leq T)} \text{ is a square-integrable martingale with } M_0 = 0$$

and

$$A = (A_t)_{(0 \leq t \leq T)} \text{ is a predictable process of finite variation } |A| \text{ with } A_0 = 0.$$

$M$  has a sharp bracket (predictable variation) process  $\langle M \rangle$  with respect to  $P$ , and we denote by  $P_M$  the measure  $P \times \langle M \rangle$  on the product space  $\bar{\Omega} := \Omega \times [0, T]$  with the  $\sigma$ -algebra of predictable sets. We now introduce the concept of trading strategy in continuous time:

**Definition 12.** A trading strategy  $\phi$  is a pair of processes  $\delta = (\delta_t)_{(0 \leq t \leq T)}$ ,  $\beta = (\beta_t)_{(0 \leq t \leq T)}$  satisfying the following conditions

1.  $\delta$  is predictable
2. The process  $\int_0^T \delta_u dX_u$  ( $0 \leq t \leq T$ ) is a semimartingale of class  $S^2$ , the class of  $P$ -square integrable processes

3.  $\beta$  is adapted

4. The process  $V(\phi)$  defined by  $V_t(\phi) := \delta_t \cdot X_t + \beta_t$ , ( $0 \leq t \leq T$ ) is right-continuous and satisfies  $V_t(\phi) \in L^2(P)$ , ( $0 \leq t \leq T$ )

The integrability condition 2. is equivalent to

$$E \left[ \int_0^T \delta_u^2 d \langle M \rangle_u + \left( \int_0^T |\delta_u| d|A|_u \right)^2 \right] < \infty$$

which means that

$$\delta \in L^2(P_M) \text{ and } \int_0^T |\delta_u| d|A|_u \in L^2(P) \quad (3.9)$$

In accordance with the usual terminology, the process  $V(\phi)$  is called the *value process* of  $\phi$  and the right-continuous square-integrable process  $C_\phi$  defined by

$$C_t(\phi) := V_t(\phi) - \int_0^t \delta_u dX_u, \quad 0 \leq t \leq T$$

the (cumulative) *costs process* of  $\phi$ . We have the same definition of mean self-financing strategies as in discrete time:

**Definition 13.** A trading strategy  $\phi$  is called mean self-financing if its costs process  $C(\phi)$  is a martingale.

A *contingent claim*  $H$  is intended to model the payout at time  $T$  of some financial instrument. In mathematical terms, a contingent claim is a random variable  $H \in L^2(P)$ . We will concentrate on strategies which are  $H$ -admissible in the sense that:

**Definition 14.** A trading strategy  $\phi$  is called  $H$ -admissible for a contingent claim  $H$  if  $V_T(\phi) = H$   $P$ -a.s.

The trading strategy  $\phi$  is then said to *generate*  $H$ . Just like in discrete time we note that an  $H$ -admissible strategy always exists: we can simply choose  $\delta \equiv 0$  and  $\beta \equiv 0$  except for  $\beta_T = H$ . Schweizer in (47) gives the same result which allows to identify predictable processes satisfying integrability condition (2) with  $H$ -admissible mean self-financing strategies.



### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

**Lemma 2.** *Let  $H$  be a contingent claim. Then there exists a bijective correspondence between the set of all mean self-financing  $H$ -admissible trading strategies  $\phi$  and the set of all predictable processes  $\delta$  satisfying (2). It is given by  $\delta \mapsto \phi = (\delta, \beta)$  with*

$$\beta_t := \mathbb{E} \left[ H - \int_t^T \delta_u dX_u \middle| \mathcal{F}_t \right] - \delta_t X_t, \quad 0 \leq t \leq T$$

where we choose right-continuous versions for both the martingale given by the expectation and the stochastic integral.

*Proof.* See lemma 1 of chapter II of Schweizer (47). □

**Remark 3.** *Again we emphasize that this bijection is actually essential for deriving the optimality equation satisfied by locally risk-minimizing strategies.*

#### 3.2.2 Local Risk-Minimization

**Definition 15.** *As a measure of riskiness, we introduce for each strategy the conditional mean square error process*

$$R_t(\phi) := E \left[ (C_T(\phi) - C_t(\phi))^2 \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T$$

defined as a right-continuous version.

We now introduce the concept of a locally  $R$ -minimizing trading strategy in continuous time. Being an infinitesimal concept, it will involve limit considerations, and under suitable assumptions on the price process, the required limit actually exists. This will enable to prove that a trading strategy is locally  $R$ -minimizing if and only if it is mean-self-financing and satisfies a stochastic optimality equation.

**Definition 16.** *A trading strategy  $\Delta = (\delta, \beta)$  is called a small perturbation if it satisfies the following conditions*

1.  $\delta$  is bounded.
2.  $\int_0^T |\delta_u| d|A|_u$  is bounded.
3.  $\delta_T = \beta_T = 0$ .

As the idea is to introduce the notion of a local variation of a trading strategy, we consider partitions  $\tau = (t_i)_{(0 \leq i \leq N)}$  of the interval  $[0, T]$ . Such partitions will always satisfy

$$0 = t_0 < t_1 < \dots < t_N = T,$$

### 3.2 Continuous Time Setting

---

and their mesh size will be defined by  $|\tau| := \max_{1 \leq i \leq N} (t_i - t_{i-1})$ . A sequence  $(\tau_n)_{n \in \mathbb{N}}$  of partitions will be called *increasing* if  $\tau_n \subseteq \tau_{n+1}$  for all  $n$ . It will be called *0-convergent* if it satisfies

$$\lim_{n \rightarrow \infty} |\tau_n| = 0$$

If  $\Delta$  is a small perturbation and  $(s, t]$  is a subinterval of  $[0, T]$ , we define the small perturbation

$$\Delta|_{(s,t]} := (\delta|_{(s,t]}, \beta|_{[s,t)})$$

by setting

$$\begin{aligned} \delta|_{(s,t]}(\omega, u) &:= \delta_u(\omega) \cdot I_{(s,t]}(u) \\ \beta|_{[s,t)}(\omega, u) &:= \beta_u(\omega) \cdot I_{[s,t)}(u) \end{aligned}$$

if  $t < T$  and

$$\begin{aligned} \delta|_{(s,t]}(\omega, u) &:= \delta_u(\omega) \cdot I_{(s,t]}(u) \\ \beta|_{[s,t)}(\omega, u) &:= \beta_u(\omega) \cdot I_{[s,t)}(u) \end{aligned}$$

if  $t = T$ .

The asymmetry is stemming from the fact that  $\delta$  is predictable while  $\beta$  is merely adapted. This small perturbation allows to define in continuous time the equivalent concept of risk-minimization as in discrete time, given a partition  $\tau$  of  $[0, T]$ .

**Definition 17.** *Let  $\phi$  be a trading strategy,  $\Delta$  a small perturbation and  $\tau$  a partition of  $[0, T]$ . Then we can define the risk quotient*

$$r^\tau[\phi, \Delta](\omega, t) := \sum_{t_i \in \tau} \frac{R_{t_i}(\phi + \Delta|_{(t_i, t_{i+1}]}) - R_{t_i}(\phi)}{E[\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i} | \mathcal{F}_{t_i}]}(\omega) \cdot I_{(t_i, t_{i+1}]}(t)$$

The strategy  $\phi$  is called locally  $R$ -minimizing if

$$\liminf_{n \rightarrow \infty} r^{\tau_n}[\phi, \Delta] \geq 0 \quad P_M - a.e.$$

for every small perturbation  $\Delta$  and every increasing 0-convergent sequence  $(\tau_n)$  of partitions of  $[0, T]$ .

### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

$r^\tau[\phi, \Delta]$  is a stochastic process which is well defined  $P_M$ -a.e. on  $\Omega \times [0, T]$ . It can be interpreted as a measure for the total change of riskiness if  $\phi$  is locally perturbed by  $\Delta$  along the partition  $\tau$ . The denominator describes the appropriate time scale for these measurements. We make an additional assumption on  $X$ :

**Assumption 1.** *For  $P$ -almost all  $\omega$ , the measure on  $[0, T]$  induced by  $\langle M \rangle(\omega)$  has the whole interval  $[0, T]$  as its support.*

Equivalently we could postulate that  $\langle M \rangle(\omega)$  is strictly increasing  $P$ -a.s (see Schweizer (47, 50)). This non-degeneracy condition prevents the martingale  $M$  from being locally constant. The following lemma shows that it is enough to look for optimal strategies among mean self-financing ones:

**Lemma 3.** *Under assumptions (3.9) and (1), if a trading strategy  $\phi$  is locally risk-minimizing then it is mean self-financing.*

*Proof.* See lemma 2.1 of Schweizer (48). □

**Remark 4.** *This last result together with Remark 3 is the key ingredient for the derivation of the characterization of optimal strategies for the quadratic risk. It is indeed essential since it allows to concentrate on mean self-financing strategies to find locally risk-minimizing ones, by varying only the  $\delta$  component and using martingale techniques from the property of  $C(\phi)$ . The absence of this decoupling makes things considerably harder in the general case of a convex risk measure as we will see in the next chapter.*

The next result which gives the characterization of locally risk-minimizing strategies requires these additional technical assumptions on  $X$ :

**Assumption 2.** 1.  *$A$  is continuous*

2.  *$A$  is absolutely continuous with respect to  $\langle M \rangle$  with a density  $\alpha$  satisfying*  

$$E_M[|\alpha| \log^+ |\alpha|] < \infty$$

**Proposition 2.** *Assume that  $X$  satisfies conditions (3.9) - (2) and let  $H$  be a contingent claim and  $\phi$  and  $H$ -admissible trading strategy. Then the following statements are equivalent*

1.  *$\phi$  is locally risk-minimizing*
2.  *$\phi$  is mean self-financing and the martingale  $C(\phi)$  is (strongly) orthogonal to  $M$*

*Proof.* See proposition 2.3 of Schweizer (48). □

### 3.2.3 Explicit Characterization of Locally Risk-Minimizing Strategies

Finally, to complete this chapter review of results in the quadratic case we discuss the practical ways of explicitly computing locally risk-minimizing strategies. We present two methods, one using the so-called *minimal equivalent local martingale measure (minimal ELMM)*, the other using a forward backward stochastic differential equation (FB-SDE). Both these methods further assume that  $X$  is continuous.

#### Minimal ELMM

The minimal ELMM (for uniqueness, we refer to theorem 3.5 of Föllmer and Schweizer (21))  $\hat{P}$  has the property of turning  $X$  into a martingale while preserving strong orthogonality with respect to  $M$ , *i.e.* if a  $P$ -martingale  $N$  is strongly orthogonal to  $M$ , then under  $\hat{P}$  it is also a martingale, strongly orthogonal to  $M$ . This property and the fact that it minimizes the relative entropy with respect to the original measure  $P$  explains the terminology. For the successful use of the minimal ELMM, continuity of  $X$  is required so that the expectation of  $H$  under  $\hat{P}$ , denoted by  $V_t^{H,\hat{P}} := \hat{E}[H|\mathcal{F}_t]$ ,  $0 \leq t \leq T$  is a continuous local  $\hat{P}$ -martingale then admits a *Galtchouk-Kunita-Watanabe decomposition* with respect to  $X$  as the following decomposition

$$V_t^{H,\hat{P}} = V_0^{H,\hat{P}} + \int_0^t \delta_u^{H,\hat{P}} dX_u + L_t^{H,\hat{P}} \quad (3.10)$$

Then  $\delta_u = \delta_u^{H,\hat{P}}$  and  $\beta_u$  given according to the formula

$$\beta_u = \hat{E}[H] + \int_0^t \delta_u dX_u + L_t^{H,\hat{P}} - \delta_u X_u$$

is pseudo-optimal, hence locally risk-minimizing from proposition (2). So the basic idea is to find the minimal ELMM and then write explicitly the decomposition (3.10) for  $H$ . This is especially easy in a Markovian framework for a European contingent claim. We refer to the article of Heath, Platen and Schweizer (26) for the full derivation.

#### Forward Backward SDE

In the quadratic case, and for  $(\mathcal{F}_t)$  the completed Brownian filtration generated by  $X$ , the formulation of local risk-minimization stems straightforwardly from the Follmer Schweizer decomposition in continuous time:

### 3. QUADRATIC LOCAL RISK MINIMIZATION

---

**Definition 18.** An  $\mathcal{F}_T$ -measurable random variable  $Y \in L^2(P)$  admits a Föllmer-Schweizer decomposition if it can be written as

$$Y = Y_0 + \int_0^T \nu_s^Y dX_s + L_T^Y P - a.s.$$

where  $Y_0 \in L^2(P)$  is  $\mathcal{F}_0$ -measurable,  $\nu^Y \in \Theta_S$  where

$$\Theta_S = \left\{ \theta \text{ predictable process such that } E \left[ \int_0^T \theta_s^2 d \langle M \rangle_s + \left( \int_0^T |\theta_s| d|A|_s \right)^2 \right] \right\}$$

so that  $\int_0^T \nu_s^Y dX_s$  is well-defined and a semimartingale in  $\mathbb{S}^2$ , and the process  $L^Y = (L_t^Y)$  is a right-continuous square integrable martingale null at 0 and strongly orthogonal to  $M$ .

Then proposition 5.2 of Schweizer (50) or proposition 2.24 of Föllmer and Schweizer (21) states that when  $X$  is continuous or satisfies the so-called structured condition (requiring that  $A$  be absolutely continuous with respect to  $\langle M \rangle$  and the mean-variance tradeoff process be finite), it is equivalent for the contingent claim  $H$  to have a Föllmer-Schweizer decomposition and to possess a locally risk-minimizing strategy. But then with the martingale representation property of  $(\mathcal{F}_t)$  it immediately follows that the Föllmer-Schweizer decomposition is equivalent to the following linear forward backward stochastic differential equation

$$V_t = H - \int_t^T Z_s dW_s - \int_t^T Z_s^1 \frac{\mu_s}{\sigma_s}$$

where  $\mu$  and  $\sigma$  are the drift and volatility of the (discounted) spot process.

**Remark 5.** See also the article by ElKaroui, Peng and Quenez on backward stochastic differential equations in finance (19), Proposition 1.1 for another BSDE description of locally risk-minimizing strategies.

## 4

# Convex Local Risk Minimization

## 4.1 Measuring Risk

So as to extend the quadratic approach to a more meaningful risk-measure while still keeping enough tractability for problems to be considered, we introduce the following set of functions:

**Definition 19.** *The set  $\mathcal{R}$  of admissible functions for measuring risk is made of functions from  $\mathbb{R}$  to  $\mathbb{R}$  which are: strictly convex, positive, null at zero and twice continuously differentiable, and such that  $f \in \mathcal{R}$  is of quadratic growth in the strong sense ( $f'(x) = cx$  for  $|x| > A$ ).*

As a consequence of this definition the first derivative  $f'$  is null at zero for  $f \in \mathcal{R}$ .

**Assumption 3.** *Henceforth whenever we use a function noted  $f$  it will belong to  $\mathcal{R}$ .*

**Remark 6.** *For our derivations in continuous time we will restrict our focus on functions in  $\mathcal{R}$  which are three times continuously differentiable.*

**Remark 7.** *We consider functions of quadratic growth in order to simplify integrability issues and work in  $L^2(P)$ . The strong sense is required to have relatively simple proof of the admissibility of optimal strategies in discrete time. It can be relaxed in continuous time.*

We should also add that the latter assumptions have no impact on the financial meaning of the function given that our approach is to find those strategies which are locally optimal. Locally meaning that the only behaviour that really matters is the one of  $f$  around zero provided we consider prices evolution models with continuous paths.

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

As a matter of fact our approach is “twice” local since we will use the same local perturbations as in the quadratic approach and also a risk-measure which is local, in the sense that it measures the risk over one time step (infinitesimal in continuous time). We already noted in the introduction that using a notion of local risk has the advantage over that of global risk that it can easily encompass the hedging of more exotic products such as American options. In the quadratic approach, and in discrete time, we showed that considering either of the two risks leads to the same characterization of optimal strategies. Yet the global risk is chosen over the local risk because its generalization to continuous time appeared more obvious. Indeed the sum of the one step costs is then the difference of the portfolio values minus the trading gains:

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{t_n} (\Delta V_j - \delta_j \Delta X_j) = V_t - V_0 - \int_0^t \delta_u dX_u.$$

Then with the quadratic criterion the calculus of the limiting risk-process and its quotient can be carried out and the same characterization of martingale orthogonality of the costs process can be recovered. With a general convex risk criterion things are different. Obviously the (cumulative) costs process has the same expression and we could measure the risk of the strategy as the conditional expectation of the chosen function of the costs process:  $R_t^g := \mathbb{E}_t[f(C_T - C_t)] = \mathbb{E}_t\left[f(V_t - V_0 - \int_0^t \delta_u dX_u)\right]$ . But since we cannot separate perturbation on the component  $\delta$  and perturbation on the component  $\beta$  (unless  $f$  happens to be quadratic obviously, see last Remark of Chapter I of Schweizer (47) for an account of this fact) we do not have such explicit characterization through the mean self-financing properties of optimal strategies and therefore no bijection between such a space of  $H$ -admissible mean self-financing strategies and predictable processes. On the other hand, by considering local risk instead, we are able to formulate two very similar conditions (martingale orthogonality) to characterize optimality through the introduction of a process which we chose to name the  $f$ -risk process. This characterization then allows for a nice extension in continuous time under mild technical assumptions on the process  $X$  and the strategies under consideration. If we restrain the processes  $X$  modelling the risky asset while still allowing for most common models to be used, then we can even show that the characterization is equivalent with the concept of local risk-minimization.

This chapter therefore begins with the discrete time setting and then extends the results obtained to the continuous time setting.

## 4.2 Discrete time

### 4.2.1 Definitions and Assumptions

We use the same concepts as in the previous chapter in terms of strategies, contingent claim and small perturbations and thus start by defining the (incremental) *costs process* which we will consider for our definition of the local risk:

**Definition 20.** *The (incremental) costs process of a trading strategy  $\phi = (\beta, \delta)$  is defined as the difference*

$$\Delta C_k(\phi) := \Delta V_k - \delta_k \Delta X_k, \quad k = 0, \dots, T$$

We now introduce the local version of a convex criterion for the hedging error of a trading strategy:

**Definition 21.** *The (local)  $f$ -risk process of a trading strategy  $\phi$  is the process*

$$R_k^f(\phi) := E[f(\Delta C_{k+1}(\phi)) | \mathcal{F}_k], \quad k = 0, \dots, T-1.$$

### 4.2.2 Local $f$ -Risk Minimization

The risk-minimization is then carried out the same way as in the quadratic case, only the risk criterion has changed.

**Definition 22.** *A trading strategy  $\phi$  is called locally  $f$ -risk-minimizing if for any trading date  $k$  and any admissible local variation  $\Delta$  of  $\phi$  at  $k$  we have*

$$R_k^f(\phi + \Delta) - R_k^f(\phi) \geq 0 \quad P - a.s.$$

As announced at the beginning of the chapter we only changed the way the risk is being assessed and not the way we specify optimal strategies. However this “small” change has rather “big” implication in tools which can be used to characterize those optimal strategies, as it turns out. One major property of the quadratic criterion which is lost in the general convex case is the separability between the two components of a strategy. We already insisted on this point in the previous chapter and emphasize here again that we cannot solve the minimization problem embedded in the definition of locally  $f$ -risk-minimizing strategies in two separated steps. In the quadratic case, this was indeed realized by first minimizing a conditional covariance with respect to the  $\delta$



#### 4. CONVEX LOCAL RISK MINIMIZATION

---

component and then deriving the corresponding optimal  $\beta$  from the mean self-financing condition. Here we rely on the convexity of  $f$  and on theorem related to minimization of convex functions to characterize optimal strategies. Indeed given the set of conditions imposed on  $f$ ,  $X_k$  and  $V_k$ , we have the existence and uniqueness of the optimal strategy  $\phi^*$ . It is characterized by the first-order optimality equations

$$\mathbb{E}_k (f'(\Delta C_{k+1}(\phi^*))) = 0 \quad (4.1)$$

$$\mathbb{E}_k (f'(\Delta C_{k+1}(\phi^*))\Delta X_{k+1}) = 0 \quad (4.2)$$

**Theorem 1.** *Assume that there is some  $c > 0$  such that  $\text{Var}_k(\Delta X_{k+1}) := E_k(\Delta X_{k+1}^2) - E_k(\Delta X_{k+1})^2 \geq c P - a.s.$ , then there exists a unique locally  $f$ -risk-minimizing trading strategy  $\phi^*$  whose components  $\delta^*$  and  $\beta^*$  solve equations (4.1) and (4.2).*

To prove the theorem, we first need the following lemma:

**Lemma 4.** *Let  $h(x, y, \omega) := E_k(f(U - y - xV))(\omega)$  with  $U$  and  $V$  in  $L^2(P)$  and such that there exists  $c > 0$  with  $\text{Var}_k(V) > c P - a.s.$ . Then for a fixed  $(x, y) \mapsto \omega$   $h(x, y, \omega)$  is elliptic.*

*Proof.* The proof is straightforward from the characterization of ellipticity for  $C^2$  functions of two variables with the help of the Hessian matrix  $H$  of  $h$  for a fixed  $\omega$

$$\det(H) = \begin{vmatrix} E_k(V^2 f''(U - y - xV)) & E_k(V f''(U - y - xV)) \\ E_k(V f''(U - y - xV)) & E_k(f''(U - y - xV)) \end{vmatrix}$$

where the expression for  $H$  is justified by the quadratic growth assumption on  $f$  together with the fact that  $U$  and  $V$  are both in  $L^2(P)$ .

Then  $h$  is elliptic if the smallest eigenvalue of the symmetric matrix  $H$  has a positive lower bound for all  $x$  and  $y$ . But the characteristic polynomial of  $H$  is

$$\det(H - \lambda I) = (E_k(V^2 f''(U - y - xV)) - \lambda)(E_k(f''(U - y - xV)) - \lambda) - (E_k(V f''(U - y - xV)))^2$$

After some algebra in order to compute the discriminant of the second order polynomial, we obtain that the lowest of the two roots is strictly greater than a positive number if and only if the following expression is itself strictly greater than a positive number

$$E_k(V^2 f''(U - y - xV)) - (E_k(V f''(U - y - xV)))^2$$

But then, using the strict convexity of  $f$  it may be written as

$$(E_k(f''(U - y - xV)))^2 \left( E_k \left( V^2 \frac{f''(U - y - xV)}{E_k(f''(U - y - xV))} \right) - \left( E_k \left( V \frac{f''(U - y - xV)}{E_k(f''(U - y - xV))} \right) \right)^2 \right)$$

Now we define a new equivalent measure  $Q$  by its Radon-Nikodym derivative  $\frac{dQ}{dP} = \frac{f''(U-y-xV)}{E_k(f''(U-y-xV))}$  so that we have

$$(E_k(f''(U-y-xV)))^2 \left( E_k^Q(V^2) - (E_k^Q(V))^2 \right)$$

Since by assumption  $Var_k V > c P - a.s.$ ,  $f$  is strictly convex and of quadratic growth, and since  $P$  and  $Q$  have the same null sets, necessarily there exists  $c'$  such that  $Var_k^Q V > c'$  and the lemma is proved.  $\square$

Now the proof of the theorem follows from the construction of a  $\mathcal{F}_k$  minimizer using dyadic rationals.

*Proof.* We use the same function  $h$  as defined in the previous lemma. Firstly for a fixed  $\omega$ ,  $h$  is thus a strictly convex function of  $x$  and  $y$  so that it has a global minimum  $(x^*, y^*)$  if and only if  $(x^*, y^*)$  is a critical point of  $h$ , i.e.  $\nabla h(x^*, y^*) = 0$ . From the lemma we have that  $h$  has a unique global minimum  $P$ -almost surely. Finally we show that  $(x^*, y^*)$  is  $\mathcal{F}_k$ -measurable: let  $D_n = \{j2^{-n} | j \in \mathbb{Z}\}$  be the set of dyadic rational of order  $n$ , we define

$$(x_n(\omega), y_n(\omega)) = \underset{(x,y) \in D_n \times D_n}{\operatorname{argmin}} \{h(x, y, \omega)\}$$

Since  $\omega \mapsto h(x, y, \omega)$  is  $\mathcal{F}_k$ -measurable,  $(x_n, y_n)$  is also  $\mathcal{F}_k$ -measurable. As  $(x_n, y_n)$  is bounded in  $n$   $P$ -a.e. (since  $\lim_{|(x,y)| \rightarrow \infty} h(x, y, \omega) = +\infty$  from the ellipticity and hence coercivity of  $h$ ) and  $h$  is continuous in  $(x, y)$ ,  $(\tilde{x}, \tilde{y}) = \liminf_{n \rightarrow \infty} (x_n, y_n)$  is a  $\mathcal{F}_k$ -measurable minimizer of  $h$  and by uniqueness it is equal to  $(x^*, y^*)$ .  $\square$

**Remark 8.** We note that like in the quadratic case this result only shows how to construct (implicitly) a candidate for the locally risk-minimizing strategy but assumes that  $\phi$  is an  $H$ -admissible strategy hence in  $L^2$ . Using the assumption of strong quadratic growth of the convex function  $f$  we can show that the candidate strategy defined by relations (4.1) and (4.2) is indeed admissible. For that we would need to further assume that there exists  $c \in \mathbb{R}$  with  $0 \leq c < 1$  so that

$$(\mathbb{E}[\Delta X_k | \mathcal{F}_{k-1}])^2 \leq c \mathbb{E}[\Delta X_k^2 | \mathcal{F}_{k-1}] \quad P - a.s., \quad 1 \leq k \leq T$$

Thus we have the same sufficient condition as in the quadratic case which is equivalent to the fact that  $X$  has a bounded mean-variance trade-off process.

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

The set of equations (4.1) and (4.2) is equivalent to the property that the process  $\left(C_k^f\right)_k$  with  $C_k^f = \sum_{i=1}^k f'(\Delta C_i)$  is a martingale (strongly) orthogonal to (the martingale part of)  $X_k$ . This calls for the two following definitions:

**Definition 23.** *Given a trading strategy  $\phi$ , the  $f$ -costs process is the process  $\left(C_k^f\right)_k$  defined by  $C_k^f = \sum_{i=1}^k f'(\Delta C_i(\phi))$  for  $k \in \{1, \dots, T\}$  and  $C_0^f = 0$ .*

**Definition 24.** *A trading strategy  $\phi$  is called pseudo-optimal (for the  $f$ -risk-minimization) if its  $f$ -costs process  $C^f(\phi)$  is a martingale orthogonal to the martingale part of  $X_k$ .*

This definition will be the main ingredient of the extensions to the continuous time setting in the general semimartingale case. Before introducing another characterization of pseudo-optimality (and therefore an equivalent characterization of optimality in discrete time) through a decomposition theorem, we rewrite the martingale orthogonality property of the  $f$ -costs process

$$\Delta C_k^f(\phi) = f'(\Delta C_k) := \Delta M_k^\perp$$

where  $(M_k^\perp)$  is a martingale orthogonal to  $X$ . Since  $f'$  is bijective from the strict convexity of  $f$  and the quadratic bounds imposed, we have

$$\begin{aligned} \Delta C_k &= (f')^{-1}(\Delta M_k^\perp) \\ \Leftrightarrow C_k &= \sum_{j=1}^k (f')^{-1}(\Delta M_j^\perp) \end{aligned}$$

We therefore introduce the concept of a  $g$ -martingale:

**Definition 25.** *An adapted process  $Y$  is a  $g$ -martingale if there exists a martingale  $M$  such that*

$$Y = \sum_{j=1}^k g(\Delta M_j)$$

Likewise we have the notion of a  $g$ -martingale orthogonal to a martingale  $M$ :

**Definition 26.** *An adapted process  $Y$  is a  $g$ -martingale orthogonal to a martingale  $M$  if there exists a martingale  $M^\perp$  orthogonal to  $M$  such that*

$$Y = \sum_{j=1}^k g(\Delta M_j^\perp)$$

With this last definition in hands we can give our third equivalent characterization of a locally  $f$ -risk-minimizing strategy.

**Proposition 3.** *There exists an  $H$ -admissible locally  $f$ -risk-minimization strategy  $\phi$  if and only if the contingent claim  $H$  admits the following decomposition*

$$H = c + \sum_{j=1}^T \delta_j \Delta X_j + L_T^{f'^{-1}, \perp}$$

where  $c$  is a constant,  $\delta$  a predictable process in  $L^2(X)$  and  $L^{f'^{-1}, \perp}$  a  $f'^{-1}$ -martingale orthogonal to the martingale part of  $X$ , with  $L_0^{f'^{-1}, \perp} = 0$ .

*Proof.* The if part of the theorem is straightforward from the remark following definition (24) since an  $H$ -admissible strategy  $\phi$  verifies  $V_T = H$ . For the reverse it is easily checked that the strategy  $\phi = (\beta, \delta)$  with  $\beta$  defined by

$$\beta_k = c + \sum_{j=1}^k \delta_j \Delta X_j + L_k^{f'^{-1}, \perp} - \delta_k X_k, \quad t = 1, \dots, T$$

is pseudo-optimal and thus locally  $f$ -risk-minimizing. □

**Remark 9.** *To further anticipate on the development in continuous we notice that if we assume that  $f^{(3)}(x) \geq 0, \forall x \in \mathbb{R}$  then  $f'$  is convex too and thus the  $f'^{-1}$ -martingale appearing in the decomposition is a supermartingale. Therefore applying this remark to the  $f$ -costs process of a locally risk-minimizing strategy,  $C^f$  possesses a unique Doob decomposition into a decreasing predictable process  $A^f$  and a martingale  $M^f$ . We may then rewrite the pseudo-optimal condition into a so-called backward difference equation*

$$Y_k + \sum_{j=k}^T Z_j^1 \Delta M_j + \sum_{j=k}^T Z_j^2 \Delta M_j^f - \sum_{j=k}^T F(j, Y_j, Z_j^1, Z_j^2) = H$$

where the unknowns are the adapted processes  $Y, Z^1$  and  $Z^2$ . This result is very close to the one we will see in continuous time with the representation of pseudo-optimal strategies through solutions of a forward backward stochastic differential equation. See Cohen and Elliott (12) for more insights on the theory of backward difference equations.

**Remark 10.** *At this point it is worth checking whether the results obtained in the general convex case agree with the quadratic case of chapter 3. We thus restate our results for  $f(x) = \frac{1}{2}x^2$ . The  $f$ -costs process is then the costs process as defined by*

## 4. CONVEX LOCAL RISK MINIMIZATION

---

equation (3.5) and the characterization obtained in theorem (1) is equivalent to the formulation obtained in Remark (2). Likewise the Follmer Schweizer decomposition of  $H$  is equivalent to the decomposition of theorem (3) since  $f^{t-1} = Id$ .

We next move the continuous time setting and see how the results obtained in the quadratic case and the results obtained in discrete time in the convex case find their counterparts.

### 4.3 Continuous Time Setting

#### 4.3.1 Definitions and Assumptions

We recall that we work with a probability space  $(\Omega, \mathcal{F}, P)$  now equipped with a continuous time filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$  where  $T \in \mathbb{R}_+$  is still a fixed and finite time horizon. As usual we assume the  $(\mathcal{F}_t)$  satisfies the usual conditions of right-continuity and completeness. We also assume that  $\mathcal{F}_0$  is trivial and that  $\mathcal{F}_T = \mathcal{F}$ . The discounted price process  $X$  still satisfies

$$X = (X_t)_{(0 \leq t \leq T)} \text{ is a semimartingale with a decomposition } X = X_0 + M + A$$

such that

$$M = (M_t)_{(0 \leq t \leq T)} \text{ is a square-integrable martingale with } M_0 = 0$$

and

$$A = (A_t)_{(0 \leq t \leq T)} \text{ is a predictable process of finite variation } |A| \text{ with } A_0 = 0.$$

We recall the definition of a trading strategy in continuous time:

**Definition 27.** A trading strategy  $\phi$  is a pair of processes  $\delta = (\delta_t)_{(0 \leq t \leq T)}$ ,  $\beta = (\beta_t)_{(0 \leq t \leq T)}$  satisfying the following conditions

1.  $\delta$  is predictable
2. The process  $\int_0^t \delta_u dX_u$ ,  $(0 \leq t \leq T)$  is a semimartingale of class  $S^2$
3.  $\beta$  is adapted (hence  $V$  is adapted too)
4. The process  $V(\phi)$  defined by  $V_t(\phi) := \delta_t \cdot X_t + \beta_t$ ,  $(0 \leq t \leq T)$  is right-continuous and satisfies  $V_t(\phi) \in L^2(P)$ ,  $(0 \leq t \leq T)$

We recall as well the definition of the costs process in continuous time:

**Definition 28.** *The costs process is the following right-continuous and square-integrable process*

$$C_t(\phi) := V_t(\phi) - \int_0^t \delta_u dX_u, \quad (0 \leq t \leq T)$$

Contrary to the quadratic case we do not associate with the costs process a global risk-measure so we need a partition of the trading interval  $[0, T]$  to measure the local risk accordingly. The (local) risk associated with the costs process in continuous time is then

**Definition 29.** *Given a partition  $\tau$  of  $[0, T]$ , where  $\tau = \{0 = t_0, t_1, \dots, t_k = T\}$  the (local) risk of a trading strategy  $\phi$  at  $t_i \in \tau$  is*

$$\Delta R_{t_i}(\phi) := \mathbb{E} (f (\Delta C_{t_{i+1}}(\phi)) | \mathcal{F}_{t_i})$$

In order to define risk-minimization in a local way, we again need the concept of small perturbations:

**Definition 30.** *A small perturbation is a bounded trading strategy  $\phi = (\beta, \delta)$  such that  $\beta_T = 0$  and  $\delta_T = 0$ .*

Given a contingent claim  $H$  and a trading strategy  $\phi$  generating  $H$ , we want to study the increase of risk at some discrete times when the strategy is perturbed. To do so, given a partition  $\tau$  of  $[0, T]$ , where  $\tau = \{0 = t_0, t_1, \dots, t_k = T\}$ , and a small perturbation  $\Delta$ , we define the process  $r_f^\tau$  the following way:

**Definition 31.** *The  $f$ -risk quotient of a trading strategy  $\phi$  along the partition  $\tau$  is the process*

$$r_f^\tau[\phi, \Delta](t, \omega) = \sum_{t_i, t_{i+1} \in \tau} \frac{\Delta R_{t_i}(\phi + \Delta|_{(t_i, t_{i+1}]})(\omega) - \Delta R_{t_i}(\phi)(\omega)}{t_{i+1} - t_i} \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

The  $f$ -risk quotient is always well-defined since for the case of convex risk-minimization we use the size of the mesh instead of the increase of quadratic variation of the martingale part of  $X$  as time scale for measurement of risk increase due to perturbations.

**Remark 11.** *We emphasize that this definition of the risk quotient of a strategy differs in this case from the quadratic definition (17) since the measurement is made on the*

## 4. CONVEX LOCAL RISK MINIMIZATION

---

incremental (remaining) risk  $\Delta R$  instead of the global (remaining) risk  $\sum_j \Delta R_j$ . On the other hand the definition is equivalent to the one used in section (III.2) of Schweizer (47) for the treatment of American options, apart from the time scale. So as to have the same equivalence between optimal and pseudo-optimality which we obtained in the discrete time setting, we changed our definition from the original article (1). We will see however that in the case of liquidity costs studied in chapter 7, we still need the discrete time perturbations of the risk, though the trading occurs in continuous time. In the cases of interest it does not change the optimality equations though.

### 4.3.2 Local $f$ -Risk-Minimization

Now we can define the local  $f$ -risk-minimization the same way as we did for the discrete time setting

**Definition 32.** For a contingent claim  $H$ , a trading strategy  $\phi$  generating  $H$  is called locally risk-minimizing if for every small perturbation  $\Delta$  and every increasing sequence of partitions  $(\tau_n)_{n \in \mathbb{N}}$  tending to the identity, we have

$$\liminf_{n \rightarrow \infty} r_f^{\tau_n}[\phi, \Delta] \geq 0 \quad P - a.e.$$

As a matter of fact, this definition naturally extends the notion of local minimization of local risk. However this definition might not always be of much practical interest and in the following section, we will introduce the concept of a pseudo-optimal strategy similar to the one introduced in the discrete time setting. When restricting our attention to certain prices processes  $X$  it will be shown that we can concentrate on those strategies indeed.

### 4.3.3 The $f$ -Costs Process

We proceed with defining the  $f$ -costs process which will allow us to characterize pseudo-optimal strategies by analogy with discrete time.

**Definition 33.** For a trading strategy  $\phi$  we define the  $f$ -costs process  $C_t^f(\phi)$  as the following limit, whenever it exists

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{l_n} f' \left( V^{\tau_k^n} - V^{\tau_{k-1}^n} - \int_{\tau_{k-1}^n}^{\tau_k^n} \delta_s^{\tau_{k-1}^n} dX_s \right)$$

where convergence is required in ucp topology, for any increasing 0-convergent sequence  $(\tau_n)$  of Riemann partitions of  $[0, T]$  of length  $l_n$  (i.e.  $\tau_{l_n}^n = T$ ) and where we used the notation  $X^t$  for the process stopped at  $t$ .

To ensure that the  $f$ -costs process is well defined we need to introduce some restrictions on strategies. We shall concentrate on strategies which are  $H$ -admissible according to the following definition:

**Definition 34.** A trading strategy  $\phi = (\beta, \delta)$  is  $H$ -admissible if

$$\begin{cases} V_T = H \text{ } P - a.s. \\ \text{Its costs process } (C_t) \text{ is a semimartingale (and hence } V \text{ itself is one).} \end{cases}$$

We now focus on an  $H$ -admissible strategy  $\phi$  and state a theorem related to the existence of the  $f$ -costs process.

**Theorem 2.** The  $f$ -costs process of an  $H$ -admissible strategy  $\phi$  is well defined and is given according to the following formula

$$\begin{aligned} C_t^f(\phi) &= f''(0) \left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) \\ &\quad + \frac{f^{(3)}(0)}{2} \left( [V, V]_t^c - 2 \int_{0+}^t \delta_{s-} d[V, X]_s^c + \int_{0+}^t \delta_{s-}^2 d[X, X]_s^c \right) \\ &\quad + \sum_{0 < s \leq t} f'(\Delta V_s - \delta_{s-} \Delta X_s) - f''(0)(\Delta V_s - \delta_{s-} \Delta X_s) \end{aligned} \quad (4.3)$$

with notation  $[V, X]^c$  standing for the continuous part of the (càdlàg) quadratic covariation process.

*Proof.* The reasoning is very close to the one used in the proof of Itô formula for general semimartingales in Protter (44). Let  $\mathcal{P}_n$  be an increasing sequence of Riemann partitions of  $[0, T]$ ,  $\mathcal{P}_n = \{0 = t_0^n \leq \dots \leq t_{N_n}^n = T\}$ .

$$C_t^{f, \mathcal{P}_n}(\phi) = \sum_{k=1}^{N_n} f' \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)$$

where we have assumed without loss of generality that  $t$  belongs to the sequence of partitions ( $t = t_{N_n}^n$  with  $\lim_{n \rightarrow \infty} N_n = +\infty$ ). Since  $V$  and  $\int \delta dX$  are càdlàg processes, and  $\sum_s (\Delta V_s)^2$  and  $\sum_s \delta_{s-}^2 (\Delta X_s)^2$  are (absolutely) convergent series, given  $\epsilon > 0$  we can find two sets  $A$  and  $B$  such that  $A$  and  $B$  are disjoint and  $A \cup B$  exhausts the jump times of  $V$  and  $X$  on  $(0, T]$ ,  $A$  being a set of jump times that  $V$  and  $X$  have a.s. a finite number of times and  $B$  being such that  $\sum_{0 < s \leq t} (\Delta V)^2 \leq \epsilon^2$  and  $\sum_{0 < s \leq t} \delta_{s-}^2 (\Delta X)^2 \leq \epsilon^2$ . Thus we have

$$C_t^{f, \mathcal{P}_n}(\phi) = \sum_{k, A} f' \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) + \sum_{k, B} f' \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)$$



#### 4. CONVEX LOCAL RISK MINIMIZATION

---

where  $\sum_{k,A}$  denotes  $\sum_k 1_{\{A \cap (t_{k-1}, t_k] \neq \emptyset\}}$  and  $\sum_{k,B}$  denotes  $\sum_k 1_{\{B \cap (t_{k-1}, t_k] \neq \emptyset\}}$ . The first sum converges to  $\sum_{s \in A} f'(\Delta V_s - \delta_{s-} \Delta X_s)$ . In the second sum we apply Taylor's theorem which says

$$f'(x) = f''(0)x + \frac{1}{2}f^{(3)}(0)x^2 + R(x)$$

where  $|R(x)| \leq r(x)x^2$ , such that  $r : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an increasing function with  $\lim_{u \downarrow 0} r(u) = 0$ . Thus we have

$$\begin{aligned} \sum_{k,B} f' \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) = \\ f''(0) \sum_{k,B} \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) \end{aligned} \quad (4.4)$$

$$+ \frac{1}{2} f^{(3)}(0) \sum_{k,B} \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2 \quad (4.5)$$

$$+ \sum_{k,B} R \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) \quad (4.6)$$

The first sum (4.4) is equal to

$$\sum_k \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) - \sum_{k \in A} \left( V_{t_k} - V_{t_{k-1}} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)$$

which converges to

$$\left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) - \sum_{s \in A} (\Delta V_s - \delta_{s-} \Delta X_s)$$

The second sum (4.5), after developing and switching to obvious and less cumbersome notations, is equal to

$$\sum_{k,B} (V_k - V_{k-1})^2 - 2(V_k - V_{k-1}) \int_{t_{k-1}}^{t_k} \delta_s dX_s + \left( \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2$$

$\sum_{k,B} (V_k - V_{k-1})^2 = \sum_k (V_k - V_{k-1})^2 - \sum_{k,A} (V_k - V_{k-1})^2$  and the first sum converges to  $[V, V]_t$  while  $\sum_{k \in A} (V_k - V_{k-1})^2$  converges to  $\sum_{s \in A} (\Delta V_s)^2$ . Now  $\sum_{k,B} 2(V_k - V_{k-1}) \int_{t_{k-1}}^{t_k} \delta_s dX_s$  converges to  $2 \int_{0+}^t \delta_{s-} d[V, X]_s - 2 \sum_{s \in A} \delta_{s-} \Delta V_s \Delta X_s$ . Finally  $\sum_{k,B} \left( \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2$  is equal to  $\sum_k \left( \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2 - \sum_{k,A} \left( \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2$  and converges to  $\int_{0+}^t \delta_{s-}^2 d[X, X]_s -$

$\sum_{s \in A} \delta_{s-}^2 (\Delta X_s)^2$ . Now we turn to the last term (4.6) of the Taylor's development

$$\begin{aligned} & \left| \sum_{k,B} R \left( V_k - V_{k-1} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right) \right| \\ & \leq \sum_{k,B} r \left( \left| V_k - V_{k-1} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right| \right) \left( V_k - V_{k-1} - \int_{t_{k-1}}^{t_k} \delta_s dX_s \right)^2 = (*) \end{aligned}$$

Assuming that  $\sup_{\omega} \delta \leq K < \infty$  over  $[0, T]$  we have

$$(*) \leq \sup r ((K+1)\epsilon) ([V, V]_t + 2K[V, X]_t + K^2[X, X]_t)$$

We are now ready to take the limit when  $\epsilon$  goes to zero. The last term tends to zero from the property of  $r$  and it remains to prove that the series  $\sum_{s \in A}$  are absolutely convergent. We next proceed by localization, as in Protter (44) by considering first  $U_K = \inf\{t > 0, |\delta| > K\}$ ,  $W_K = \inf\{t > 0, |V| > K\}$  and  $Z_K = \inf\{t > 0, |X| > K\}$  so that  $1_{[0, U_K]} \delta$ ,  $1_{[0, W_K]} V$  and  $1_{[0, Z_K]} X$  are  $[-K, K]$ -valued. Therefore we have that  $|f'(x) - f''(0)x| \leq Cx^2$  for some constant  $C$ . This allows us to write

$$\begin{aligned} & \left| \sum_{s \in A} f'(\Delta V_s - \delta_{s-} \Delta X_s) - f''(0) \left( \sum_{s \in A} \Delta V_s - \delta_{s-} \Delta X_s \right) \right| \\ & \leq C \sum_{s \in A} \Delta V_s^2 - 2\delta_{s-} \Delta V_s \Delta X_s + \delta_{s-}^2 \Delta X_s^2 \\ & \leq C([V, V]_t + 2K|[V, X]_t| + K^2[X, X]_t) < \infty \end{aligned}$$

And the series are absolutely convergent which completes the proof.  $\square$

**Remark 12.** We note from the explicit formula (4.3) that the  $f$ -costs process is itself a semimartingale for an  $H$ -admissible strategy.

**Remark 13.** Conditions imposed on  $\phi$  to be admissible strategies are stronger than what is actually required for the  $f$ -costs process to be well defined. It would be enough to have

$$\begin{cases} V_T = H P - a.s. \\ X \text{ has finite and integrable quadratic variation.} \\ V \text{ has finite and integrable quadratic variation.} \\ V \text{ and } X \text{ have finite and integrable quadratic covariation.} \end{cases}$$

for the  $f$ -costs process to be well defined as seen from the proof. We enforced the condition that  $V$  is a semimartingale so as to have another interesting characterization of pseudo-optimal strategies.

## 4. CONVEX LOCAL RISK MINIMIZATION

---

With the  $f$ -costs process well defined for strategies of interest in continuous time, we can now state the criteria which will characterize pseudo-optimal strategies, by analogy with the discrete time case.

### 4.3.4 Pseudo-Optimal Strategies

**Definition 35.** *An  $H$ -admissible strategy  $\phi$  will be called pseudo-optimal for the local risk-minimization if its  $f$ -costs process is a martingale strongly orthogonal to the martingale part  $M$  of the process  $X$ .*

**Remark 14.** *Since the  $f$ -costs process is well defined for an  $H$ -admissible strategy, the definition always makes sense.*

In the next chapter, we will derive the corresponding set of equations that pseudo-optimal strategies have to solve in two different Markovian frameworks. But before that, we first present the notion of  $g$ -martingale in continuous time, which we will need to study some special cases where we can already have a different representation of the solutions. This is again inspired from the discrete time setting results.

## 4.4 $g$ -Martingales and Orthogonality

We defined pseudo-optimality in continuous time through the  $f$ -costs process, the limit of  $f'$  applied to infinitesimal costs increments, which happened to be semimartingale increments given the assumptions on  $H$ -admissible strategies. In order to have the same representation result which we obtained in discrete time in the form of proposition (3), we need some auxiliary definitions and properties on the limit process.

**Definition 36.** *For a function  $g$  twice continuously differentiable and a general semimartingale  $Y$ , we define the  $g$ -stochastic integral of  $Y$  noted  $\int_0^T g(dY)$  the following limit whenever it exists*

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{l_n} g \left( Y^{\tau_k^n} - Y^{\tau_{k-1}^n} \right)$$

*where convergence is required in ucp topology, for any increasing  $0$ -convergent sequence  $(\tau_n)$  of Riemann partitions of length  $l_n$  (i.e.  $\tau_{l_n}^n = T$ ) of  $[0, T]$ .*

We then have the following theorem:

**Theorem 3.** *Given a semimartingale  $Y$ , the  $g$ -stochastic integral process of  $Y$  is well defined and is given by the semimartingale below*

$$g'(0)(Y_t - Y_0) + \frac{g''(0)}{2}[Y, Y]_t^c + \sum_{0 < s \leq t} g(\Delta Y_s) - g'(0)(\Delta Y_s) \quad \forall t \in [0, T] \quad P - a.s. \quad (4.7)$$

Theorem (2) is thus only a consequence of this more general theorem.

*Proof.* We give our proof for completeness but also refer to the work of Diop (17, 18) for a generalization to time dependent and optional function of semimartingale increments. Essentially the proofs rely on the same ingredients though.

Let  $\mathcal{P}_n$  be a refining sequence of Riemann partitions of  $[0, T]$ ,  $\mathcal{P}_n = \{0 = t_0^n \leq \dots \leq t_n^n = T\}$ . We want to find the limit of the following discretized sum of semimartingale increments

$$\sum_{k=1}^n g(Y_{t_k} - Y_{t_{k-1}}) = (*)$$

Since  $Y$  is a semimartingale, which we assume to be a càdlàg process without loss of generality,  $\sum_s (\Delta Y_s)^2$  is an absolutely convergent serie, given  $\epsilon > 0$  we can find two sets  $A$  and  $B$  such that  $A$  and  $B$  are disjoint and  $A \cup B$  exhausts the jump times of  $Y$  on  $(0, T]$ ,  $A$  being a set of jump times that  $Y$  has almost surely a finite number of times and  $B$  being such that  $\sum_{0 < s \leq t} (\Delta Y)^2 \leq \epsilon^2$ .

Thus we have

$$(*) = \sum_{k,A} g(Y_{t_k} - Y_{t_{k-1}}) + \sum_{k,B} g(Y_{t_k} - Y_{t_{k-1}})$$

where  $\sum_{k,A}$  denotes  $\sum_k 1_{\{A \cap (t_{k-1}, t_k] \neq \emptyset\}}$  and  $\sum_{k,B}$  denotes  $\sum_k 1_{\{B \cap (t_{k-1}, t_k] \neq \emptyset\}}$ . The first sum converges to  $\sum_{s \in A} g(\Delta Y_s)$ . In the second sum we apply Taylor's theorem saying

$$g(x) = g'(0)x + \frac{1}{2}g''(0)x^2 + R(x)$$

where  $|R(x)| \leq r(x)x^2$ , such that  $r : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an increasing function with

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

$\lim_{u \downarrow 0} r(u) = 0$ . Thus we have

$$\sum_{k,B} g(Y_{t_k} - Y_{t_{k-1}}) = g'(0) \sum_{k,B} (Y_{t_k} - Y_{t_{k-1}}) \quad (4.8)$$

$$+ \frac{1}{2} g''(0) \sum_{k,B} (Y_{t_k} - Y_{t_{k-1}})^2 \quad (4.9)$$

$$+ \sum_{k,B} R(Y_{t_k} - Y_{t_{k-1}}) \quad (4.10)$$

The first sum (4.8) is equal to

$$\sum_k (Y_{t_k} - Y_{t_{k-1}}) - \sum_{k \in A} (Y_{t_k} - Y_{t_{k-1}})$$

which converges to

$$Y_t - Y_0 - \sum_{s \in A} \Delta Y_s$$

The second sum (4.9) is equal to

$$\sum_k (Y_k - Y_{k-1})^2 - \sum_{k,A} (Y_k - Y_{k-1})^2$$

where we have used notation  $Y_k$  for  $Y_{t_k}$ . It converges to  $[Y, Y]_t - \sum_{s \in A} (\Delta Y_s)^2$ .

Now we turn to the last term (4.10) of the Taylor's development

$$\left| \sum_{k,B} R(Y_k - Y_{k-1}) \right| \leq \sum_{k,B} r(|Y_k - Y_{k-1}|) (Y_k - Y_{k-1})^2 \quad (4.11)$$

Assuming that  $\sup_{\omega} Y \leq K \leq \infty$  over  $[0, T]$  we have (4.11)  $\leq \sup r(2K\epsilon) [Y, Y]_t$ . We are now ready to take the limit when  $\epsilon$  goes to zero. The last term tends to zero from the property of  $r$  and it remains to prove that the series  $\sum_{s \in A}$  are absolutely convergent. We next proceed by localization by considering first  $U_K = \inf\{t > 0, |Y| > K\}$  so that  $1_{[0, U_K]} Y$  is  $[-K, K]$ -valued. Therefore we have that  $|g(x) - g'(0)x| \leq Cx^2$  for some constant  $C$ . This allows us to write

$$\left| \sum_{s \in A} g(\Delta Y_s) - g'(0) \left( \sum_{s \in A} \Delta Y_s \right) \right| \leq C \sum_{s \in A} \Delta Y_s^2 \leq C([Y, Y]_t < \infty$$

And the series are absolutely convergent which completes the proof.  $\square$

We thus have an application noted  $\int g(\cdot)$  which transforms a semimartingale into a semimartingale. It is interesting to ask whether this application is invertible and in case it is what is its inverse. The next theorem answers positively and is a direct extension of the discrete time case.

**Theorem 4.** *Let  $g$  be a bijective  $C^2$  function from  $\mathbb{R}$  to  $\mathbb{R}$ , with  $g(0) = 0$ , then the  $g$ -stochastic integral seen as an application from the linear space of semimartingales  $\mathcal{S}$  into  $\mathcal{S}$  is itself a bijection and its inverse is given by the  $g^{-1}$ -stochastic integral modulo a constant term.*

*Proof.* It is enough to check that given a semimartingale  $S$  we have

$$\int_0^t g^{-1} d \left( \int_0^s g(dS) \right) = S_t - S_0$$

But from (4.7) we write

$$\int_0^t g(dS) = g'(0) (S_t - S_0) + \frac{g''(0)}{2} [S, S]_t^c + \sum_{0 < s \leq t} g(\Delta S_s) - g'(0)(\Delta S_s)$$

so that

$$\begin{aligned} \int_0^t g^{-1} d \left( \int_0^s g(dS) \right) &= g^{-1'}(0) \left( \int_0^t g(dS) \right) + \frac{g^{-1''}(0)}{2} \left[ \int_0^\cdot g(dS), \int_0^\cdot g(dS) \right]_t^c \\ &\quad + \sum_{0 < s \leq t} g^{-1} \left( \Delta \int_0^s g(dS) \right) - g^{-1'}(0) \left( \Delta \int_0^s g(dS) \right) \end{aligned}$$

Now we have

$$\begin{aligned} g^{-1'}(0) \left( \int_0^t g(dS) \right) &= \frac{1}{g'(0)} \left( g'(0)(S_t - S_0) - \frac{g''(0)}{2} [S, S]_t^c \right. \\ &\quad \left. + \sum_{0 < s \leq t} g(\Delta S_s) - g'(0)(\Delta S_s) \right) \\ \frac{g^{-1''}(0)}{2} \left[ \int_0^\cdot g(dS), \int_0^\cdot g(dS) \right]_t^c &= -\frac{g''(0)}{2g'(0)} [S, S]_t^c \\ \sum_{0 < s \leq t} g^{-1} \left( \Delta \int_0^s g(dS) \right) - g^{-1'}(0) \left( \Delta \int_0^s g(dS) \right) &= \sum_{0 < s \leq t} \Delta S_s - \frac{1}{g'(0)} g(\Delta S) \end{aligned}$$

Summing the last three equations we find

$$\int_0^t g^{-1} \left( d \int_0^s g(dS) \right) = S_t - S_0$$

which is the result expected. □

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

**Remark 15.** *Now this is immediately seen to apply to the  $f$ -costs process since with our standing assumptions on  $f$ ,  $f'$  is strictly positive and maps  $\mathbb{R}$  into  $\mathbb{R}$  and hence is a bijection.*

With the definition of the  $g$ -stochastic integral and the existence theorem, we can introduce the continuous time notions of a  $g$ -martingale

**Definition 37.** *An adapted stochastic process  $Z$  is a  $g$ -martingale if there exists a martingale  $M$  such that  $Z$  is the  $g$ -stochastic integral of  $M$ .*

Likewise we have the definition of a  $g$ -martingale orthogonal to a martingale  $N$

**Definition 38.** *An adapted stochastic process  $Z$  is a  $g$ -martingale orthogonal to  $N$  if there exists a martingale  $M$ , orthogonal to  $N$ , such that  $Z$  is the  $g$ -stochastic integral of  $M$ .*

The last remark together with the last definition find applications to our risk-minimization problem, provided we can have a description of martingales orthogonal to  $M$ . This situation happens typically when we study the special case when the filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$  has the martingale representation property. For instance if  $(\Omega, \mathcal{F})$  is separable, then applying theorem 44 of Protter (44), there is a countable  $L^2$ -martingale basis. In our case for simplicity and we further assume the following

**Definition 39.**  *$M$  and  $N$  form a  $P$ -basis of  $L^2(P)$  if the following conditions are satisfied*

1. *Both  $M$  and  $N$  are square-integrable martingales under  $P$ .*
2.  *$M$  and  $N$  are  $P$ -orthogonal*
3. *Every martingale  $Z$  in  $L^2(P)$  has a unique representation*

$$Z_t = Z_0 + \int_0^t K_u^1 dM_u + \int_0^t K_u^2 dN_u \quad P - a.s. \quad (4.12)$$

*for two predictable processes  $K^1 \in L^2(P_M)$  and  $K^2 \in L^2(P_N)$ .*

As already noted in Schweizer (48), condition (4.12) is equivalent to assuming that the stable subspace generated by  $M$  and  $N$  coincides with the whole space of square-integrable martingales under  $P$ . From the discrete time equivalent characterization of optimal strategies of proposition (3) we are encouraged to look at having the same result in continuous time. Indeed we have the following theorem

**Theorem 5.** *Given a contingent claim  $H$  the two following statements are equivalent*

1. *There exists an  $H$ -admissible pseudo-optimal strategy  $\phi$*
2.  *$H$  admits a decomposition*

$$H = c + \int_0^T \delta_u dX_u + M_T^{f'^{-1}, \perp}$$

where  $c \in \mathbb{R}$  and  $M^{f'^{-1}, \perp}$  is a  $f'^{-1}$ -martingale orthogonal to  $M$ .

*Proof.* That 1.  $\Rightarrow$  2. is immediate from the fact that if there exists an  $H$ -admissible pseudo-optimal strategy  $\phi$ , by definition we have that the  $f$ -costs process  $C_t^f(\phi)$  is a martingale orthogonal to  $M$ . We thus write  $C_t^f(\phi) = M_t^\perp$  and apply theorem (4) to find

$$\int_0^t f'^{-1} \left( dC_u^f(\phi) \right) = V_t - V_0 - \int_0^t \delta_u dX_u = \int_0^t f'^{-1} \left( dM_u^\perp \right)$$

Writing  $\int_0^t f'^{-1} (K_u^2 dN_u) = M_t^{f'^{-1}, \perp}$ , where  $K^2$  is the process arising in the decomposition (4.12) of the  $f$ -costs process, we have for  $t = T$ ,  $H = V_0 + \int_0^T \delta_u dX_u + M_T^{f'^{-1}, \perp}$ . For 2.  $\Rightarrow$  1. let us assume that the contingent claim  $H$  admits the following decomposition

$$H = c + \int_0^T \delta_u dX_u + M_T^{f'^{-1}, \perp}$$

with  $c \in \mathbb{R}$  and  $\delta_u$  a predictable process in  $L^2(X)$  and  $M^{f'^{-1}, \perp}$  a  $f'^{-1}$ -martingale orthogonal to the martingale part  $M$  of  $X$ . We then consider the adapted process  $\beta$  defined by

$$\beta_t = c + \int_0^t \delta_u dX_u + M_t^{f'^{-1}, \perp} - \delta_t X_t$$

We now have to show that the trading strategy  $\phi$  defined by the pair  $(\beta, \delta)$  is indeed pseudo-optimal. But we have  $(V_u = \beta_u + \delta_u X_u)$

$$C_t^f(\phi) = \int_0^t f'(dV_u - \delta_u dX_u) = \int_0^t f'(dM_u^{f'^{-1}, \perp})$$

and by definition of  $M^{f'^{-1}, \perp}$  and again using theorem (4), we see that  $C^f(\phi)$  is a martingale orthogonal to  $M$ .  $\square$



#### 4. CONVEX LOCAL RISK MINIMIZATION

---

We thus have another characterization of pseudo optimality which will prove most useful when considering specific models. One other interesting result we have from using theorem (4) for pseudo-optimal strategies is that it allows for a relatively simple proof of the actual optimality of those strategies with mild technical assumptions on  $X$ . This is important as it means that we can indeed concentrate on these for the purpose of finding locally risk-minimizing strategies. Thus for the last part of this chapter, we add the assumptions that the martingales  $M$  and  $N$  appearing in (39) are continuous and that their quadratic variation processes are absolutely continuous with respect to the Lesbegue measure.

With these assumptions in place we state the last theorem of the chapter

**Theorem 6.** *Let  $\phi$  be a pseudo-optimal strategy for  $H$ , then it is locally risk-minimizing.*

*Proof.* In order to avoid confusion with our notations, we use in the proof notation  $\Delta_{t_i}^{t_{i+1}}U$  for the increment of the process  $U$  between  $t_i$  and  $t_{i+1}$ :  $\Delta_{t_i}^{t_{i+1}}U = U_{t_{i+1}} - U_{t_i}$  and  $\Delta U_t$  for the jump of  $U$  at  $t$ :  $\Delta U = U_t - U_{t-}$ . The proof follows from writing the definition of pseudo-optimal strategies

$$C_t^f(\phi) = M_t^\perp$$

with  $M^\perp$  a martingale orthogonal to  $M$ . With the assumptions on trading strategies we can apply the representation property of our filtration to write  $M_t^\perp = \int_0^t H_u dN_u$ , with  $H$  a predictable process in  $L^2$ . From theorem (4) is equivalent to having

$$C_t(\phi) := V_t - V_0 - \int_0^t \delta_u dX_u = M_t^{f'^{-1}, \perp}$$

where  $M^{f'^{-1}, \perp} = \int f'^{-1} dM^\perp$  is a  $f'^{-1}$ -martingale orthogonal to the martingale part of  $X$ . Therefore we may write the local risk at  $t_i$  as

$$\Delta R_{t_i}(\phi) = \mathbb{E}_{t_i} f \left( \Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp} \right) \quad (4.13)$$

for a given partition  $\tau = \{t_i\}_i$ . Now we write the process  $r_f^\tau$  on  $t = t_i \in \tau$  for a small perturbation  $\Gamma = (\epsilon, \nu)$

$$\begin{aligned} r_f^\tau[\phi, \Gamma](t, \omega) &= \frac{\Delta_{t_i}^{t_{i+1}} R(\phi + \Gamma|_{[t_i, t_{i+1}]}) (\omega) - \Delta_{t_i}^{t_{i+1}} R(\phi) (\omega)}{t_{i+1} - t_i} \\ &= \frac{\mathbb{E}_{t_i} \left( f \left( \Delta_{t_i}^{t_{i+1}} C(\phi + \Gamma|_{[t_i, t_{i+1}]}) \right) \right) (\omega) - \mathbb{E}_{t_i} \left( f \left( \Delta_{t_i}^{t_{i+1}} C(\phi) \right) \right) (\omega)}{t_{i+1} - t_i} \end{aligned}$$

Applying Taylor's formula with remainder term to  $g : (x, y) \mapsto f(x + y)$  in the expectation, we have that

$$\begin{aligned} f\left(\Delta_{t_i}^{t_{i+1}} C(\phi + \Gamma|_{[t_i, t_{i+1}()])}\right) &= f\left(\Delta_{t_i}^{t_{i+1}} C(\phi)\right) - \eta_{t_i} f'\left(\Delta_{t_i}^{t_{i+1}} C(\phi)\right) \\ &\quad - \left(\int_{t_i}^{t_{i+1}} \nu_s dX\right) f'\left(\Delta_{t_i}^{t_{i+1}} C(\phi)\right) + \frac{1}{2} \left(\eta_{t_i} + \int_{t_i}^{t_{i+1}} \nu_s dX\right)^2 g(\tilde{\phi}) \end{aligned}$$

where  $g(\tilde{\phi}) = f''(\Delta C(\tilde{\phi}))$  with  $\tilde{\phi} = (\tilde{\beta}, \tilde{\delta})$  such that  $|\tilde{\beta}| \leq \beta$  and  $|\tilde{\delta}| \leq \delta$ . Rearranging and simplifying we get

$$\begin{aligned} r_f^\top[\phi, \Gamma](t, \omega) &= \eta_{t_i} \frac{\mathbb{E}_{t_i}(f'(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} + \frac{\mathbb{E}_{t_i}\left(\left(\int_{t_i}^{t_{i+1}} \nu_s dX\right) f'(\Delta C_{t_{i+1}}(\phi))\right) (\omega)}{t_{i+1} - t_i} \\ &\quad + \frac{\mathbb{E}_{t_i}\left(\left(\eta_{t_i} + \frac{1}{2} \int_{t_i}^{t_{i+1}} \nu_s dX\right)^2 g(\tilde{\phi})\right) (\omega)}{t_{i+1} - t_i} \end{aligned}$$

Now we replace  $\Delta_{t_i}^{t_{i+1}} R(\phi)$  with its expression (4.13) to find that the first term on the right-hand side is equal to

$$\eta_{t_i} \frac{\mathbb{E}_{t_i}\left(f'\left(\Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp}\right)\right) (\omega)}{t_{i+1} - t_i}$$

We next apply Itô's lemma for a general semimartingale to  $f'\left(\Delta_{t_i}^t M^{f'^{-1}, \perp} 1_{t \geq t_i}\right)$ , between  $t_i$  and  $t_{i+1}$ . This gives

$$\begin{aligned} &f'\left(\Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp}\right) = \\ &\int_{t_i}^{t_{i+1}} f''\left(\Delta_{t_i}^{s-} M^{f'^{-1}, \perp}\right) dM_s^{f'^{-1}, \perp} + \frac{1}{2} \int_{t_i}^{t_{i+1}} f^{(3)}\left(\Delta_{t_i}^{s-} M^{f'^{-1}, \perp}\right) d[M^{f'^{-1}, \perp}, M^{f'^{-1}, \perp}]_s^c \\ &\quad + \sum_{t_i < t \leq t_{i+1}} f'\left(\Delta_{t_i}^t M^{f'^{-1}, \perp}\right) - f'\left(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}\right) - f''\left(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}\right) \Delta M_t^{f'^{-1}, \perp} \end{aligned}$$

We begin by treating the first two terms, leaving the jump term aside. Replacing  $M_t^{f'^{-1}, \perp}$  by its expression (without the jump part  $\sum_{0 < s \leq t} f'^{-1}(\Delta M_s^\perp) - f'^{-1}(0) \Delta M_s^\perp$ ) and computing  $[M^{f'^{-1}, \perp}, M^{f'^{-1}, \perp}]^c$  accordingly, we get after removing the martingale term whose expectation vanishes

$$\begin{aligned} &\mathbb{E}_{t_i}\left(f'\left(\Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp}\right)^c\right) = \\ &\frac{f^{(3)}(0) \mathbb{E}_{t_i}\left(\int_{t_i}^{t_{i+1}} f''\left(\Delta_{t_i}^{s-} M^{f'^{-1}, \perp}\right) d[M^\perp, M^\perp]_s^c\right)}{2f''(0)^3} \\ &\quad + \frac{\mathbb{E}_{t_i}\left(\int_{t_i}^{t_{i+1}} f^{(3)}\left(\Delta_{t_i}^{s-} M^{f'^{-1}, \perp}\right) d[M^\perp, M^\perp]_s^c\right)}{2f''(0)^2} \end{aligned}$$

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

We used the superscript  $c$  on the left hand side to remind that we consider only the continuous part.

Then, by dividing by  $t_{i+1} - t_i$  and taking the limit using the absolute continuity of  $[M^\perp, M^\perp]^c$  with respect to the Lebesgue measure and the left continuity of  $s \mapsto \Delta_{t_i}^{s-} M^{f'^{-1}, \perp}$

$$\lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} \left( f' \left( \Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp} \right)^c \right)}{t_{i+1} - t_i} = 0$$

We now take care of the two terms coming from the jump part of  $f'(\Delta_{t_i}^t M^{f'^{-1}, \perp})$ .

The jump part is

$$\begin{aligned} & \sum_{t_i < t \leq t_{i+1}} f''(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}) \left[ f'^{-1}(\Delta M_s^\perp) - (f'^{-1})'(0) \Delta M_s^\perp \right] \\ & + \sum_{t_i < t \leq t_{i+1}} f'(\Delta_{t_i}^t M^{f'^{-1}, \perp}) - f'(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}) - f''(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}) \Delta M_t^{f'^{-1}, \perp} \end{aligned}$$

With the jump of the process  $M^{f'^{-1}, \perp}$  at  $t$  being

$$\Delta M_t^{f'^{-1}, \perp} = f'^{-1} \left( \Delta M_t^\perp \right)$$

this jump part becomes

$$\begin{aligned} & \sum_{t_i < t \leq t_{i+1}} f' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} + f'^{-1} \left( \Delta M_t^\perp \right) \right) - f' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} \right) \\ & - \sum_{t_i < t \leq t_{i+1}} f'' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} \right) (f'^{-1})'(0) \Delta M_t^\perp \end{aligned} \quad (4.14)$$

Thus if there are jumps, the expectation of the above expression might not vanish, unless  $f$  happens to be quadratic.

The rest of the proof relies on exactly the same argument, except for applying Itô's formula to the product  $\left( \int_{t_i}^{t_{i+1}} \nu_s dX \right) f'(\Delta C_{t_{i+1}}(\phi))$  instead of  $f'(\Delta C_{t_{i+1}}(\phi))$  only to find

$$\begin{aligned} & \left( \int_{t_i}^{t_{i+1}} \nu_s dX \right) f'(\Delta C_{t_{i+1}}(\phi)) = \left( \int_{t_i}^{t_{i+1}} \nu_s dX \right) f' \left( \Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp} \right) \\ & \left( \int_{t_i}^{t_{i+1}} \nu_s dX \right) f' \left( \Delta_{t_i}^{t_{i+1}} M^{f'^{-1}, \perp} \right) = \\ & \int_{t_i}^{t_{i+1}} f'' \left( \Delta_{t_i}^{s-} M^{f'^{-1}, \perp} \right) dM_s^{f'^{-1}, \perp} + \frac{1}{2} \int_{t_i}^{t_{i+1}} f^{(3)} \left( \Delta_{t_i}^{s-} M^{f'^{-1}, \perp} \right) d[M^{f'^{-1}, \perp}, M^{f'^{-1}, \perp}]_s^c \\ & + \sum_{t_i < t \leq t_{i+1}} f'(\Delta_{t_i}^t M^{f'^{-1}, \perp}) - f'(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}) - f''(\Delta_{t_i}^{t-} M^{f'^{-1}, \perp}) \Delta M_t^{f'^{-1}, \perp} \end{aligned}$$

and then using the orthogonality of  $M^\perp$  with  $M$ . Thus both first order terms in the Taylor development vanish, leaving only the positive quadratic term, and therefore meaning that any small perturbation of a pseudo-optimal strategy will lead to an increase of risk. Hence the optimality of pseudo-optimal strategies is proved.  $\square$

A few remarks are in order, following the theorem.

**Remark 16.** *We derived the implication under very mild technical assumptions on the  $L^2$ -basis, basically the only assumption which is needed is that  $M$  and  $N$  be continuous with a quadratic variation process absolutely continuous with respect to Lebesgue measure. The implication still holds provided we have a martingale representation theorem with  $M$  and  $N$  that may have jumps but we then need to impose more requirements on the jump part. For instance it would be sufficient in a market driven by a special semimartingale (in the sense of Jacod and Shiryaev (29)) to have that the compensator of the random measure of jumps  $\nu$  is absolutely continuous with respect to Lebesgue measure, i.e.  $\nu([0, t] \times G) = F(G)t$  for  $G$  a Borel set of  $\mathbb{R}$ , since upon taking expectation of the jump part (4.14) and using Fubini's theorem*

$$\begin{aligned} & \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \left\{ f' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} + f'^{-1}(x) \right) - f' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} \right) \right\} \nu(dx) dt \\ & \quad - \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \left\{ f'' \left( \Delta_{t_i}^{t-} M^{f'^{-1}, \perp} \right) (f'^{-1})'(0)x \right\} \nu(dx) dt \end{aligned}$$

*Thus dividing by  $t_{i+1} - t_i$  and letting  $t_{i+1} \rightarrow t_i$  we get that the jump part vanishes as well and so the pseudo-optimal strategy is indeed optimal.*

**Remark 17.** *In our original article introducing “non-quadratic” local risk-minimization (1) we already relied on pseudo-optimality as the criterion to apply in continuous time by analogy with the situation in discrete time. We also introduced a criterion for actual optimality similar to the one used in chapter 7 where we discuss the impact of liquidity costs on strategies. The criterion is different in that it considers that not only risk is measured at discrete time but also costs. The link between pseudo-optimality and optimality was then achieved in the very special case when the filtration was the natural filtration of strong Markov processes, solutions of stochastic differential equations. We will see in the following chapters 5 and 6 that the new results obtained in the form of theorem 6 make things considerably easier. Also in a general setting considering criterion from definition 32 means that we allow for more general kind of strategies since we can always recover the criterion from our original article by restricting our strategy to “simple” strategies (as in Harrison and Pliska (24) or Cetin, Jarrow and Protter*

#### 4. CONVEX LOCAL RISK MINIMIZATION

---

(8) for instance) which correspond to buy and hold strategies on a predefined and fixed set of times.

**Remark 18.** *Our last remark is directly related to the problem of uniqueness of locally risk-minimizing strategies. Indeed, this is almost straightforward in discrete time from the assumptions made on the risk function  $f$ . In continuous time this is more involved unless we fall in the case described in the remark above, where uniqueness appears as a by-product of the equations explicitly derived.*

*In the general case introduced in this chapter, we would first need to find suitable conditions so that the reverse implication of theorem (6) holds. Here the difficulty comes from the fact that the  $f$ -costs process is defined independently of the actual optimality criterion, contrary to the quadratic case.*

# 5

## Application to Stochastic Volatility Models

This chapter is dedicated to the study of a stochastic volatility model with Markovian solutions which in turn allow to find a characterization of optimal strategies through a non-linear parabolic partial differential equation (PDE). The PDE is naturally obtained from the pseudo-optimal criterion given that the  $f$ -costs can be expressed as a function of the diffusion parameters, assuming smoothness of the strategy components. On the other hand the link between non-linear PDE of quadratic growth in the gradient and quadratic backward stochastic differential equations (as in Kobylanski (33)) arising naturally from an extension of the Feynman-Kac formula is revisited thanks to the equivalence between pseudo-optimality and optimality in this setting, since theorem (6) of chapter 4 applies.

### 5.1 Model Assumptions

Throughout this chapter we model the evolution of  $X$  through an SDE with stochastic volatility, which is given by the following system

$$dX_s = a(s, X_s, Y_s)ds + b(s, X_s, Y_s)dW_s^1 \quad (5.1)$$

$$dY_s = c(s, X_s, Y_s)dt + d(s, X_s, Y_s) \left( \rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2 \right) \quad (5.2)$$

with initial conditions  $X_0 = x$ ,  $Y_0 = y$  and  $(W^1, W^2)$  a standard two-dimensional Wiener process under  $P$ . With this prescription the stochastic factor  $Y$  of the volatility

## 5. APPLICATION TO STOCHASTIC VOLATILITY MODELS

---

has constant instantaneous correlation  $\rho$  with  $X$ , i.e.  $d \langle X, Y \rangle_t = \rho dt$ . We will assume that  $a, b, c$  and  $d$  are Lipschitz continuous functions on  $(0, T) \times \mathbb{R}^2$  taking values in  $\mathbb{R}$  and that there exists a constant  $C$  such that for  $t \in [0, T]$  and  $x, y, x', y' \in \mathbb{R}$

$$|a(t, x, y) - a(t, x', y')| + |b(t, x, y) - b(t, x', y')| + |c(t, x, y) - c(t, x', y')| + |d(t, x, y) - d(t, x', y')| \leq C (|x - x'| + |y - y'|) \quad (5.3)$$

$$|a(t, x, y)|^2 + |b(t, x, y)|^2 + |c(t, x, y)|^2 + |d(t, x, y)|^2 \leq C (1 + |x|^2 + |y|^2) \quad (5.4)$$

These assumptions ensure existence and uniqueness of a strong solution to the system of SDE (5.1), (5.2) and the continuity of the flow  $(t, x, y) \mapsto (X_s^{t,x,y}, Y_s^{t,x,y})$ , where  $X_s^{t,x,y}$  and  $Y_s^{t,x,y}$  are the solutions with initial data  $X_t = x$  and  $Y_t = y$  (see for instance Platen and Kloeden (32)). These properties in turn ensure that the solution is a strong Markov process.

Thus with these diffusion assumptions we will now place ourselves in a Markovian framework and look for the optimal strategy  $\phi$  as a smooth function of the state variables

$$\begin{cases} \delta_t &= \delta(t, X_t, Y_t) \\ V_t &= V(t, X_t, Y_t) \end{cases}$$

### 5.2 Quadratic PDE

We first derive a PDE formulation. For that purpose we use equation (4.3) in order to express the  $f$ -costs process as a function of the diffusion parameters and the strategy

$$\begin{aligned} C_t^f(\phi) &= \int_0^t \left[ f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \right. \\ &\quad + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\ &\quad \left. - f^{(3)}(0) \delta_u \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 \right] du \\ &\quad + \int_0^t f''(0) \left( \left( \frac{\partial V}{\partial X} - \delta_u \right) b_u + \frac{\partial V}{\partial Y} \rho d_u \right) dW_u^1 + \int_0^t f''(0) \frac{\partial V}{\partial Y} \sqrt{1 - \rho^2} d_u dW_u^2 \end{aligned}$$

Now, applying to the strategy  $\phi$  the first pseudo-optimality criterion, i.e. that  $(C_t^f)$  must be martingale under the measure  $P$ , we find the equation satisfied by the portfolio

value  $V$

$$\begin{aligned} f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \\ + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\ - f^{(3)}(0) \delta_u \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 = 0 \end{aligned}$$

with terminal condition  $V_T = H$ .

Applying to the strategy  $\phi$  the second pseudo-optimality criterion, *i.e.* that the martingale  $(C_t^f)$  must be orthogonal to  $X$ , we find the equation satisfied by the optimal hedge  $\delta$

$$\left( \frac{\partial V}{\partial X} - \delta_u \right) b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u = 0$$

We next rewrite the equations in a more standard way

$$\begin{aligned} \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} c_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u = \\ \delta_u a_u + \alpha \left( \left( \frac{\partial V}{\partial X} b_u + \frac{\partial V}{\partial Y} \rho d_u - \delta_u b_u \right)^2 + (1 - \rho^2) \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 \right) \end{aligned} \quad (5.5)$$

$$\frac{\partial V}{\partial X} b_u + \frac{\partial V}{\partial Y} \rho d_u - \delta_u b_u = 0 \quad (5.6)$$

where  $\alpha = -\frac{1}{2} \frac{f^{(3)}(0)}{f''(0)}$ . Inserting equation (5.6) in equation (5.5) and dropping the subscript  $u$  of the time-dependence for ease of reading, we find

$$\frac{\partial V}{\partial u} + \Lambda V = \frac{a}{b} \left( \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho d \right) + \alpha \left( \sqrt{1 - \rho^2} \frac{\partial V}{\partial Y} d \right)^2 \quad (5.7)$$

$$\delta b = \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho d \quad (5.8)$$

where  $\Lambda$  is the infinitesimal generator corresponding to the diffusion equations (5.1) and (5.2) under measure  $P$

$$\Lambda V_u = \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u$$

which corresponds to the second-order elliptic operator defined by  $L = -\Lambda$ .

Equation (5.8) gives the optimal hedge as a function of the portfolio value and its derivatives with respect to the state variables, so as such it may be seen as an extension



## 5. APPLICATION TO STOCHASTIC VOLATILITY MODELS

---

of the Black and Scholes delta. Solving for  $V$  in the quadratic parabolic PDE given by equation (5.7) with boundary condition  $V_T = H$  yields the value of the optimal portfolio.

### 5.2.1 Existence and Uniqueness Results

We first state our last assumptions on the SDE driving price and volatility processes and on the contingent claim  $H$

**Assumption 4.** • *Functions  $a$ ,  $b$ ,  $c$  and  $d$  are bounded, uniformly in  $(t, x, y)$ .*

- *The volatility process of  $X$ , given by function  $a(t, x, y)$  is uniformly lower bounded in  $(t, x, y)$ , with a strictly positive bound (this condition is often referred to in the literature as a non-degeneracy condition for  $X$ ).*
- *The contingent claim  $H$  is bounded.*

We now study the existence and uniqueness of solutions to the quasi-linear parabolic PDE.

We next write equation (5.7) in an Hamiltonian form

$$-\frac{\partial V}{\partial t} + H(t, x, y, V, DV, D^2V) = 0 \text{ in } (0, T) \times \mathbb{R}^2$$

where the Hamiltonian is  $H(t, x, y, u, p, M) = -Tr(aM) - \mu p^1 - \gamma p^2 - F(t, x, y, u, \sigma^t(t, x, y)p)$ , with

$$\sigma = \begin{pmatrix} b & 0 \\ \rho d & (1 - \rho^2)d \end{pmatrix}$$

$p = (p^1, p^2) \in \mathbb{R}^2$ ,  $a = (a_{i,j}) = \sigma \sigma^t$  the variance-covariance matrix, so with  $a_{11} = b^2$ ,  $a_{22} = d^2$  and  $a_{12} = a_{21} = \rho b d$ ,  $M \in \mathcal{S}^2$  the space of  $2 \times 2$  symmetric matrices and finally  $F(t, x, y, u, \sigma^t(t, x, y)p) = \frac{\alpha}{b} \sigma^t(t, x, y)p \cdot e^1 + \alpha \sigma^t(t, x, y)p \cdot e^2$ , with  $e^1 = (1, 0)$  and  $e^2 = (0, 1)$ .

Since solutions of the quadratic PDE (5.7) may not be smooth we introduce the weaker notion of viscosity solutions. We refer to Crandall, Ishii and Lions (14) for more details on this notion.

**Definition 40.** *A lower semicontinuous (resp. upper semicontinuous) function  $u$  is a viscosity subsolution (resp. viscosity supersolution) of (5.5) if for any  $\phi \in C^2([0, T] \times$*

$\mathbb{R}^n$  such that if  $\phi - V$  has a global maximum (resp a global minimum) in  $(t_0, x_0, y_0)$  we have

$$-\frac{\partial \phi}{\partial t}(t_0, x_0, y_0) + H(t_0, x_0, y_0, D\phi(t_0, x_0, y_0), D^2\phi(t_0, x_0, y_0)) \leq 0$$

$$\left[ \text{resp. } -\frac{\partial \phi}{\partial t}(t_0, x_0, y_0) + H(t_0, x_0, y_0, D\phi(t_0, x_0, y_0), D^2\phi(t_0, x_0, y_0)) \geq 0 \right]$$

The function  $u$  is a viscosity solution if it is both a supersolution and a subsolution.

**Theorem 7.** Equation (5.7) has a unique viscosity solution on  $[0, T] \times \mathbb{R}^2$ .

*Proof.* The proof consists in verifying that all hypotheses from theorem 3.2 and theorem 3.8 of Kobylanski (33) hold. Indeed theorem 3.8 and the remark just before it show that there exists a solution of the associated Forward Backward SDE and it is a viscosity solution of the quadratic PDE. Theorem 3.2 which is a comparison theorem for viscosity super and subsolution then provides the uniqueness result.

The conditions to be checked are hypotheses (H4) and (H5) of Kobylanski. Hypothesis (H4) is exactly conditions (5.3) and (5.4). For hypothesis (H5), we have

$$|F(t, x, y, u, \sigma^t(t, x, y)q)| \leq \max(1, \frac{a}{b})|\sigma^t(t, x, y)q|^2 \leq C(1 + |\sigma^t(t, x, y)q|^2)$$

$$\left| \frac{\partial F}{\partial z}(t, x, y, u, \sigma^t(t, x, y)q) \right| = \left| \left( \frac{a}{b}, 2\alpha\sigma^t(t, x, y)q \cdot e^2 \right) \right| \leq C(1 + |\sigma^t(t, x, y)q|)$$

$$\frac{\partial F}{\partial u}(t, x, y, u, \sigma^t(t, x, y)q) = 0 \leq c_\epsilon + \epsilon|\sigma^t(t, x, y)q|^2, \forall \epsilon > 0$$

$$\begin{aligned} & \left| \left( \frac{\partial F}{\partial x}(t, x, y, u, \sigma^t(t, x, y)q), \frac{\partial F}{\partial y}(t, x, y, u, \sigma^t(t, x, y)q) \right) \right| = \\ & \left| \left( \frac{\frac{\partial a}{\partial x}b - a\frac{\partial b}{\partial x}}{b^2} + \alpha\nabla_x\sigma^t(t, x, y)q, \left( \frac{\frac{\partial a}{\partial y}b - a\frac{\partial b}{\partial y}}{b^2} + \alpha\nabla_y\sigma^t(t, x, y)q \right) \right) \right| \leq \\ & C(1 + |\sigma^t(t, x, y)q|^2) \end{aligned}$$

thanks to the non degeneracy condition on  $X$ . □

## 5. APPLICATION TO STOCHASTIC VOLATILITY MODELS

---

### 5.2.2 Complete markets case

The case of complete markets allows us to recover the celebrated Black and Scholes formula ((5), (41)) regardless of the choice we make for the function  $f$ . Indeed, by taking  $d$  the volatility of volatility equal to zero, the optimality equations reduce to

$$\delta_u = \frac{\partial V}{\partial X} \quad (5.9)$$

$$\frac{\partial V}{\partial u} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 = 0 \quad (5.10)$$

Equation (5.9) gives the perfect hedging strategy in that context, since upon suitable boundary conditions it is well known that the PDE (5.10) has a unique solution. Of course one can verify that the  $f$ -costs process is then identically zero, which amounts to having a self-financing strategy that perfectly replicates the contingent claim  $H$ .

### 5.3 Quadratic FBSDE

Given the two equations we found for the optimal portfolio, we may now relate pseudo-optimal strategies for the local risk-minimization with the solution of a FBSDE associated with the diffusion process of the discounted price  $X$ . This is based on the generalisation of the Feynman-Kac formula (see survey paper on BSDE in finance from El Karoui, Peng and Quenez (19) for instance), which links quasi-linear PDE with BSDE.

The next theorem states that if we have a smooth solution to the quadratic PDE (5.7) then it is also a solution to the associated FBSDE.

**Theorem 8.** *Any smooth pseudo-optimal strategy  $\phi = (\beta, \delta)$  for the local risk-minimization yields a solution to the following Forward-Backward stochastic differential equation*

$$\begin{aligned} dX_t &= a_t dt + b_t dW_t^1 \\ dY_t &= c_t dt + d_t (\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2) \\ -dV_s &= g(s, X_s, Y_s, V_s, Z_s^1, Z_s^2) ds - Z_s^1 dW_s^1 - Z_s^2 dW_s^2 \\ V_T &= H \end{aligned}$$

with  $W = (W^1, W^2)$  is a standard two-dimensional Brownian motion and  $g(s, X, Y, Z^1, Z^2) = -\frac{a}{b} Z^1 - \alpha (Z^2)^2$ , with  $V = \beta + \delta X$  and  $Z = (\delta b, \frac{\partial V}{\partial Y} d\sqrt{1 - \rho^2})$ .

*Proof.* The result follows from a straightforward application of the Itô formula to the pseudo-optimal strategy  $\phi^* = (\beta^*, \delta^*)$ , which solves equations (5.7) and (5.8). We get

$$\begin{aligned} dV_t^* &= \frac{\partial V}{\partial t} + \Lambda V_t + \frac{\partial V}{\partial S} \sigma dW_s^1 + \frac{\partial V}{\partial \sigma} \Sigma (\rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2) \\ \Leftrightarrow dV_t^* &= -g(t, S_t, \sigma_t, V_t, \delta_t^*, \sqrt{1 - \rho^2} \frac{\partial V}{\partial \sigma} \Sigma) + \frac{\partial V}{\partial S} \sigma dW_s^1 + \frac{\partial V}{\partial \sigma} \Sigma (\rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2) \\ \Leftrightarrow -dV_t^* &= g(t, S_t, \sigma_t, V_t, \delta_t^*, \sqrt{1 - \rho^2} \frac{\partial V}{\partial \sigma} \Sigma) - \delta_t^* \sigma_s dW_s^1 - \frac{\partial V}{\partial \sigma} \Sigma \sqrt{1 - \rho^2} dW_s^2 \end{aligned}$$

which is the result announced with  $Y = V$  and  $Z = (\delta \sigma, \frac{\partial V}{\partial \sigma} \Sigma \sqrt{1 - \rho^2})$ .  $\square$

The last theorem requires stronger assumptions than what we may need in this context, in that it assumes that we have pseudo-optimal strategies are smooth functions of the state variables. A direct application of theorem (4) allows us to overcome these requirements since we are typically in the case where there the filtration generated by the state variables  $X$  and  $Y$  has the martingale representation property and hence there is a  $P$ -basis with  $M = W^1$  and  $N = W^2$  (see Karatzas and Shreve (31) for instance).

Finally we give the most interesting result which is a direct application of a general result of Kobylanski (33) and the last results of chapter 4

**Theorem 9.** *The unique solution of the quadratic FBSDE (7.1) gives an optimal strategy for the risk-minimization problem with risk function  $f$ .*

*Proof.* With our standing assumptions on the process  $X$  and  $Y$ , we can apply theorem 3.8 of Kobylanski (33) which ensures that there is a unique solution of the quadratic FBSDE (7.1). Now since  $X$  has continuous paths we are in a position to apply theorem (6) to get the desired result upon checking that the BSDE part of FBSDE (7.1) is exactly equivalent to

$$V_t - V_0 - \int_0^t \delta_u dX_u = M_t^{f'^{-1}, \perp}$$

$\square$

## 5. APPLICATION TO STOCHASTIC VOLATILITY MODELS

---

## 6

# Application to Jump Diffusion Models

This chapter is dedicated to the study of a stochastic volatility model with jumps, with Markovian solutions which allow to find a characterization of optimal strategies through a non-linear parabolic partial integro-differential equation (PDE). The PIDE is naturally obtained from the pseudo-optimal criterion given that the  $f$ -costs can be expressed as a function of the diffusion parameters, assuming smoothness of the strategy components. On the other hand the link between non-linear PDE of quadratic growth in the gradient and quadratic backward stochastic differential equations (Kobylanski) arising naturally from an extension of the Feynman-Kac formula is revisited thanks to the equivalence between pseudo-optimality and optimality in this setting, since theorem (6) of chapter 4 applies.

In this section, we want to provide an example of a situation where the non-quadratic risk definitely implies a different hedging strategy, not only through the Taylor expansion around zero of the risk function  $f$ . We therefore model the evolution of  $S$  through an SDE with stochastic volatility and Poisson jumps in the vein of the Bates model (4)

$$\begin{aligned}dX_s &= a(s, X_{s-}, Y_{s-})ds + b(s, X_{s-}, Y_{s-})dW_s^1 + kdN_s \\dY_s &= c(s, X_{s-}, Y_{s-})ds + d(s, X_{s-}, Y_{s-})\left(\rho dW_s^1 + \sqrt{1-\rho^2}dW_s^2\right)\end{aligned}$$

with initial conditions  $X_0 = x$ ,  $Y_0 = y$  and  $(W^1, W^2)$  a standard two-dimensional Wiener process under  $P$ .  $N_t$  is a Poisson process of intensity  $\lambda$  and the amplitude of

## 6. APPLICATION TO JUMP DIFFUSION MODELS

---

the jumps  $k$  has probability distribution  $K$ . We also assume that  $W_t = (W_t^1, W_t^2)$ ,  $N_t$  and  $k$  are independent. With this prescription the stochastic factor  $Y$  of the volatility has constant instantaneous correlation  $\rho$  with  $X$ , i.e.  $d \langle X, Y \rangle_t = \rho dt$ . As in the case of stochastic volatility we shall assume that appropriate conditions hold on the adapted processes  $a, b, c, d, K$  and  $\lambda$  so that the set of SDEs has a unique strong solution. With these assumptions we will again place ourselves in a Markovian framework and look for the optimal strategy  $\phi$  as smooth functions of the state variables

$$\delta_t = \delta(t, X_t, Y_t)$$

$$V_t = V(t, X_t, Y_t)$$

### Quadratic PIDE

We first derive a PIDE formulation. For that purpose we express the costs process as a function of the diffusion parameters and the strategy

$$\begin{aligned} C_t^f(\phi) = & \int_0^t \left( f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_{u-} a_u \right) \right. \\ & + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\ & - f^{(3)}(0) \delta_{u-} \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_{u-}^2 b_u^2 \Big) du \\ & + \int_0^t f''(0) \left( \frac{\partial V}{\partial X} - \delta_{u-} \right) \sigma_u dW_u^1 + \int_0^t f''(0) \frac{\partial V}{\partial Y} d_u dW_u^2 \\ & + \int_0^t \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u) K(k) dk dN_u \end{aligned}$$

which we have again obtained from equation (4.3), and with  $\Delta V_u$  the jump in  $V$  when there is a jump  $\Delta X_u$  of size  $k$  on  $X$  at time  $u$  being equal to  $V(u-, X_{u-} + k, Y_{u-}) - V(u-, X_{u-}, Y_{u-})$ . Now, applying to the strategy  $\phi$  the first pseudo-optimality criterion, i.e. that  $(C_t^f)$  must be martingale under the measure  $P$ , we find the equation satisfied

---

by the portfolio value  $V$

$$\begin{aligned}
f''(0) & \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_{u-} a_u \right) \\
& + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
& - f^{(3)}(0) \delta_{u-} \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_{u-}^2 b_u^2 \\
& + \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u) K(k) dk \lambda_u = 0
\end{aligned}$$

with terminal condition  $V_T = H$ .

Applying to the strategy  $\phi$  the second pseudo-optimality criterion, i.e. that the martingale  $(C_t^f)$  must be orthogonal to  $X$ , we find the equation satisfied by the optimal hedge  $\delta$

$$\left( \frac{\partial V}{\partial X} - \delta_{u-} \right) b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u + \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u) k K(k) dk \lambda_u = 0$$

Contrary to the stochastic volatility case, where only the local behaviour of the risk function  $f$  in 0 mattered, the optimal strategy in the jump-diffusion model requires the knowledge of the risk function  $f$  on its whole support.

**Remark 19.** *A formal link with forward backward stochastic differential equations with jumps can be done, by assuming that a smooth enough solution to the PIDE satisfied by  $V$  exists. There are a few recent papers (Matoussi and Wang (40), Barles, Buckdahn and Pardoux (2), Bouchard and Elie (6) or Lejay, Mordecki and Torres (36) for a numerical scheme) which discuss the probabilistic interpretation of solutions of PIDE through solutions of the corresponding FBSDE with jumps in a general framework, but they all require Lipschitz conditions on the source term of the PIDE or on the driver of the FBSDE with jumps. We however mention two articles from Morlais (42, 43) which tackle the issue of BSDE with jumps and with a quadratic growth in the driver but with a very specific form of the latter.*



## 6. APPLICATION TO JUMP DIFFUSION MODELS

---

# 7

## Liquidity

### 7.1 Motivations

We have seen in chapter 4 how the theory of quadratic hedging could be extended so as to use a general convex function to account for the asymmetric nature of the risk arising from trading costs. However these trading costs were assumed to be evaluated in a perfect market, which is to say without considering transaction costs. In this chapter we extend the approach to the case where there are transaction costs on the stock component. Transaction costs are understood to occur as a dependence of the costs on the volume traded and not from the bid/ask spreads where the change of volume (the “gamma”) would be the main factor. As a matter of fact we are especially concerned with continuous time equations in which case a non-zero bid/ask spread would lead to an infinite costs in most cases (due to the infinite variation of the Brownian motion). The local risk is still a convex function of the local costs process and we derive the corresponding (pseudo-)optimal strategies in both discrete time and continuous time settings. We end the chapter by exemplifying the hedging method with two same models used in the “infinite” liquidity case of chapter 5 and 6: a one dimensional stochastic volatility model and a mixture of stochastic volatility and jump diffusion.

### 7.2 Liquidity costs and risk process

Among the number of market imperfections which can be considered when applying a trading strategy are two equally important facts. Firstly there always exists a difference in the prices at which one can either buy or sell an asset, this is know as the bid/ask

## 7. LIQUIDITY

---

spread or bid/offer spread. Incorporating this effect in our theory is rather involved since the costs would then depend on the sign of the amount of risky asset to buy or to sell, and this feature would certainly make the identification of optimal strategies less straightforward. In the quadratic case, this has been nicely tackled by Lamberton, Pham and Schweizer (35) in the discrete time case but its extension to continuous time does not seem obvious. Secondly, neglecting the bid/ask spread, the price depends on the absolute amount of risky asset one buys or sells. To understand how liquidity costs can modify optimal strategies and what corresponding pseudo-optimal criterion we should look at, we start by investigating the situation in a discrete time setting. We study existence and uniqueness of solutions to the minimization problem and to this end we use the same multi-period model as in chapter 4 section 4.2, where the evolution of the risky asset is driven by a strictly positive process  $X_k$ , ( $k = 0, \dots, T$ ) on some probability space  $(\Omega, \mathcal{F}, P)$ .  $\mathcal{F}_k$  then denotes the  $\sigma$ -field of events which are observable up to and including time  $k$ . We assume that  $X_k$  is adapted and square-integrable.  $X_k$  is actually the discounted stock prices process, which is equivalent to having the money market account grow at a zero interest rate.

In this two-asset market, we recall that we are interested in hedging a contingent claim which is described by a square-integrable random variable  $H \in L^2(P)$ . To do so we introduce a trading strategy  $\phi$  represented by two stochastic processes:  $(\delta_k)$ , ( $k = 1, \dots, T$ ) a predictable process and  $(\beta_k)$ , ( $k = 0, \dots, T$ ) adapted to  $\mathcal{F}_k$  and both in  $L^2(P)$ .  $\delta_k$  is the amount of stock held in period  $k$ , ( $= (t_{k-1}, t_k]$ ) and has to be fixed at the beginning of that period, *i.e.* we assume that  $\delta_k$  is  $\mathcal{F}_{k-1}$ -measurable ( $k = 1, \dots, T$ ).  $\beta_k$ , the amount held in the market account in period  $k$ , is allowed to be fixed at the end. We thus relax the usual predictable assumptions on the strategy components the same way as we did in the previous chapter. Since the adjustment at the terminal date  $T$  will be made only on the cash account, we further assume that there will be no liquidity costs on the stock. This means essentially that physically settled options can be dealt with exactly as cash settled ones.

The theoretical value of the portfolio at time  $k$  is its value right after applying the strategy and is given by

$$V_k = \delta_k X_k + \beta_k, \quad (k = 1, \dots, T)$$

$$V_0 = \beta_0$$

We admit only strategies such that each  $V_k$  is square-integrable and which replicate the contingent claim  $H$ , *i.e.* we require  $V_T = H$ , which for instance can always be done through adjusting  $\beta$  at time  $T$ . This is simply recalling definitions 1 and 5.

Applying strategy  $\phi$  induces costs  $\Delta C_k$  at time  $k > 0$ , which are given in the presence of liquidity costs on the stock by

$$\Delta C_k(\phi) = \mathcal{L}((\delta_{k+1} - \delta_k), X_k, t_k) + (\beta_k - \beta_{k-1}) \quad \forall k \in \{1, \dots, T\}$$

with the convention that  $\delta_{T+1} = \delta_T$  and where the function  $\mathcal{L}$  gives the costs of adjusting the stock part and accounts for the liquidity effect

- If  $(\delta_{k+1} - \delta_k) > 0$ , meaning that we have to buy more stocks, we might not necessarily be able to do so at the theoretical price  $X_k$  but rather at a higher price, so that the bigger the quantity to acquire the greater the marginal costs.
- If on the contrary  $(\delta_{k+1} - \delta_k) < 0$ , meaning that we have to sell more stocks, we might not necessarily be able to do so at the theoretical price  $X_k$  but rather at a lower price, so that the bigger the quantity to sell the greater the marginal costs (costs are negative in this case, so that they are smaller in absolute value).

### Assumptions on Liquidity Costs

As a consequence of the liquidity effect observed on real markets and described above, it is legitimate to assume that  $\mathcal{L} : (\mathbb{R}, \mathbb{R}_+, \mathbb{R}_+) \rightarrow \mathbb{R}$  is a strictly increasing and convex function of its first variable, with  $\mathcal{L}(0, \cdot, \cdot) = 0$  and that it is differentiable with respect to its first variable, with  $\frac{\partial \mathcal{L}}{\partial x}(0, X, \cdot) = X$ . As a matter of fact we do not take into account any bid/ask spread at this level. We also neglect the transaction's impact on the price process meaning that there is no feedback effect no matter the quantity. This amounts to assuming that the period of trading is much greater than the relaxation time of the market impact function. Finally we will assume that the first order derivative of  $\mathcal{L}$  with respect to the quantity  $x$ ,  $\frac{\partial \mathcal{L}}{\partial x}$ , is bounded. This means that above a certain quantity to buy or to sell, there is a fixed and finite price available for trading.

If there exists an adapted function  $g$ , *i.e.*  $g = g(x, t, \omega)$  with  $\omega \in \mathcal{F}_k$ , such that the liquidity costs can be written as  $\mathcal{L}((\delta_{k+1} - \delta_k), X_k, t_k) = (\delta_{k+1} - \delta_k)g((\delta_{k+1} - \delta_k), t_k)$ , then  $g$  is called the supply curve. We refer to Cetin et al (8) for more details on the self financing approach in case there is a supply curve. In our case, we will assume that

## 7. LIQUIDITY

---

there exists an increasing density function  $l : (\mathbb{R}, \mathbb{R}_+, \mathbb{R}_+) \rightarrow \mathbb{R}$ ,  $l \in \mathcal{C}^1$  which represents the price to pay for buying a marginal amount of stock so that  $\mathcal{L}$  takes the following form

$$\mathcal{L}(\Delta\delta_{k+1}, X_k, t_k) = \int_0^{\Delta\delta_{k+1}} (x, X_k, t_k) dx \quad (7.1)$$

with then  $(0, X_k, t_k)$ , the marginal costs for entering a transaction whatever its sign being equal to  $X_k$  in the absence of bid/ask spread.

It corresponds to smoothing the orderbook profile which gives the quantity available for a given price. In order to make calculus in continuous time more tractable whilst not narrowing the scope of the paper we shall assume that the marginal costs can be written as a stationary function times the theoretical spot price  $X$ , *i.e.*  $(x, X, t) = l(x)X_t$ . We now note that with these assumptions on the liquidity costs function together with the convexity of the risk function  $f$  we cannot be sure that  $(x, y) \mapsto f(\mathcal{L}(x)X + y)$  is a convex function, unlike the case of “infinite” liquidity.

The (local)  $f$ -risk is then naturally defined as the conditional expectation given information up to time  $k$  of our chosen functional  $f$  of the costs including liquidity costs incurred at time  $k + 1$ . This reads

$$\Delta R_k^f(\phi) = \mathbb{E}_k(f(\Delta C_{k+1})) \quad (7.2)$$

### 7.3 Optimal and pseudo-optimal strategies

As in the previous chapters, optimal strategies will sequentially minimize the risk process, backward in time since they have to replicate the contingent claim at time  $T$  and indeed solve the following problem

**Problem (\*)** Given a contingent claim  $H$ , find  $\phi^*$ , admissible strategy such that

$$\forall k \in (0, \dots, T - 1), \Delta R_k(\phi) \geq \Delta R_k(\phi^*) \forall \phi \text{ admissible,}$$

$$\text{with } \delta_{k+1} = \delta_{k+1}^* \text{ and } \beta_{k+1} = \beta_{k+1}^*$$

We note that Problem (\*) though formulated differently than local  $f$ -risk minimization in chapter 4 is actually equivalent. It is enough to check that if  $\phi$  and  $\phi^*$  are admissible strategies, then  $(\phi - \phi^*)1_{t_k}$  is an admissible local variation of  $\phi$  at  $k$ . Given

### 7.3 Optimal and pseudo-optimal strategies

---

the conditions imposed on  $f \in \mathcal{R}_f$ ,  $\mathcal{L}$ ,  $X_k$  and  $\beta_k$  we have the existence of the optimal strategy which is a solution to the following first-order optimality equations

$$\begin{aligned} & \begin{cases} \mathbb{E}_k (f' (\Delta C_{k+1}(\phi^*))) & = 0 \\ \mathbb{E}_k (f' (\Delta C_{k+1}(\phi^*)) \mathcal{L}'(\delta_{k+2} - \delta_{k+1}, X_{k+1}, t_{k+1})) & = 0 \end{cases} \\ \Leftrightarrow & \begin{cases} \mathbb{E}_k (f' (\Delta C_{k+1}(\phi^*))) & = 0 \\ \mathbb{E}_k (f' (\Delta C_{k+1}(\phi^*)) l(\delta_{k+2} - \delta_{k+1}) X_{k+1}) & = 0 \end{cases} \end{aligned} \quad (7.3)$$

where we have used the notation  $\mathcal{L}'$  for the partial derivatives of  $\mathcal{L}(x, y, z)$  with respect to its first variable.

We then have a theorem for the existence of a locally risk-minimizing strategy:

**Theorem 10.** *Problem (\*) has at least one solution  $\phi^*$  whose components  $\delta^*$  and  $\beta^*$  solve the set of equations (7.3).*

To prove the theorem, we first need the following lemma:

**Lemma 5.** *Let  $h(x, y, \omega) := E_k (f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) (\omega)$  with  $U, V$  and  $X \in L^2(P)$ .*

*If there exists  $c > 0$  such that  $\text{Var}_k(X) > c$  then we have  $\lim_{\|(x,y)\| \rightarrow \infty} h(x, y, \omega) = +\infty$ .*

*Proof.* We write

$$\begin{aligned} h(x, y) &= E_k (f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) \\ &= E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) > 0} 1_{U-x > 0} f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) > 0} 1_{U-x \leq 0} f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) \leq 0} 1_{U-x > 0} f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) \leq 0} 1_{U-x \leq 0} f (\mathcal{L}((U - x), X, t_{k+1}) + (V - y))) \end{aligned}$$

so that we have the following inequality

$$\begin{aligned} h(x, y) &\geq E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) > 0} 1_{U-x > 0} f ((U - x)X_{k+1} + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) > 0} 1_{U-x \leq 0} f ((U - x)A^- + B^- + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) \leq 0} 1_{U-x > 0} f ((U - x)A^+ + B^+ + (V - y))) \\ &+ E_k (1_{\mathcal{L}((U-x), X, t_{k+1}) + (V-y) \leq 0} 1_{U-x \leq 0} f ((U - x)X_{k+1} + (V - y))) \end{aligned}$$

as  $f$  is decreasing on  $\mathbb{R}_-$  and increasing on  $\mathbb{R}_+$  and as there exist  $A^+$ ,  $A^-$ ,  $B^+$  and  $B^-$  such that  $(U - x)X_{k+1} \leq \mathcal{L}((U - x), X, t_{k+1}) \leq A^+(U - x) + B^+$  for  $U - x > 0$  and  $A^-(U - x) + B^- \leq \mathcal{L}((U - x), X, t_{k+1}) \leq (U - x)X_{k+1}$  for  $U - x \leq 0$ . The latter fact

## 7. LIQUIDITY

---

stems from our assumption that  $\mathcal{L}$  has bounded first order derivative with respect to its first variable.

From this inequality and adapting the proof of lemma (4) to show that the three functions  $(x, y) \mapsto f((U - x)X_{k+1}) + (V - y)$  and  $(x, y) \mapsto f((U - x)X_{k+1}) + B^{+/-} + (V - y)$  are elliptic hence coercive we get the desired result

$$\lim_{\|(x,y)\| \rightarrow \infty} h(x, y) = +\infty \quad (7.4)$$

□

The proof of the theorem is then essentially the same as in the “infinite” liquidity case except that we cannot ensure uniqueness of an optimal strategy.

*Proof.* Let  $h(x, y, \omega)$  be the function defined above with  $U, V$  and  $X \in L^2(P)$ . We first observe that because of our assumptions on the liquidity costs, for a fixed  $\omega$ ,  $h$  is a continuous and differentiable function of  $(x, y)$  so that it reaches a minimum  $(x^*, y^*)$  only if  $(x^*, y^*)$  is a critical point of  $h$ , *i.e.*  $\nabla h(x^*, y^*) = 0$ . Secondly we have  $\lim_{\|(x,y)\| \rightarrow \infty} h(x, y, \omega) = +\infty$  from the above lemma,  $P - a.e.$  so that  $h$  has a global minimum  $P$ -almost surely. Finally we show that  $(x^*, y^*)$  is  $\mathcal{F}_k$ -measurable: let  $D_n = \{j2^{-n} | j \in \mathbb{Z}\}$  be the set of dyadic rational of order  $n$ , we define

$$(x_n(\omega), y_n(\omega)) = \underset{(x,y) \in D_n \times D_n}{\operatorname{argmin}} \{h(x, y, \omega)\}$$

Since  $\omega \mapsto h(x, y, \omega)$  is  $\mathcal{F}_k$ -measurable,  $(x_n, y_n)$  is also  $\mathcal{F}_k$ -measurable. As  $(x_n, y_n)$  is bounded in  $n$   $P$ -a.e. and  $h$  is continuous in  $(x, y)$ ,  $(\tilde{x}, \tilde{y}) = \liminf_{n \rightarrow \infty} (x_n, y_n)$  is a  $\mathcal{F}_k$ -measurable minimizer of  $h$ , satisfying  $\nabla h(\tilde{x}, \tilde{y}) = 0$ . □

For the sake of extending the notion of local risk-minimization in a continuous time setting we reinterpret the set of equations (7.3) as follows: defining the processes  $\left(C_k^f\right)_k$  by  $C_k^f = \sum_{i=1}^k f'(\Delta C_i)$ ,  $C_0^f = 0$ , and  $(X_k^S)_k$  with  $X_k^S = X_0 + \sum_{i=1}^k (l(\Delta \delta_i) X_i - l(0) X_{i-1}) = X_0 + \sum_{i=1}^k (l(\Delta \delta_i) X_i - X_{i-1})$ ,  $X_0^S = X_0$ , equations (7.3) are equivalent to having  $C^f$  be a martingale strongly orthogonal to the martingale part of the process  $X^S$ . The first process will again be referred to as the  $f$ -costs process as in chapter 4, while the new process  $X^S$  will be referred to as the supply price process. We shall name this property pseudo-optimality. We also note that in the original case of “infinite” liquidity, so with  $l(\cdot) = 1$ , the supply price process is just the stock price  $X$ , as is expected.

## 7.4 Continuous time setting

Now let  $(\Omega, \mathcal{F}, P)$  be a probability space with a filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$  satisfying the usual conditions of right-continuity and completeness.  $T \in \mathbb{R}^+$  denotes a fixed and finite time horizon. Furthermore, we assume that  $\mathcal{F}_0$  is trivial and that  $\mathcal{F}_T = \mathcal{F}$ . We model the risky asset  $X = (X_t)_{0 \leq t \leq T}$  as a strictly positive semimartingale and we use a right-continuous version of  $X$ . We use the same trading strategies as in chapters 3 and 4 which therefore satisfy the following definition:

**Definition 41.** *A general trading strategy  $\phi$  is then a pair of processes  $\delta = (\delta_t)_{0 \leq t \leq T}$ ,  $\beta = (\beta_t)_{0 \leq t \leq T}$ ,  $(\delta)_t$  being a predictable process and  $(\beta)_t$  being an adapted process.*

Contingent claims which will be considered are still of European type and are thus described by random variables  $H \in L^2(P)$ .

In order to define the processes which are the basic ingredients of pseudo-optimality in continuous time we need to restrict the set of trading strategies to  $H$ -admissible strategies which satisfy the following requirements

$$\left\{ \begin{array}{l} \delta_T = \delta^H \text{ } P - a.s. \\ \beta_T = \beta^H \text{ } P - a.s. \\ \delta \text{ has finite and integrable quadratic variation} \\ \beta \text{ has finite and integrable quadratic variation} \\ \delta \text{ and } \beta \text{ have finite and integrable quadratic covariation} \end{array} \right.$$

The two following sections are dedicated to the definition and expression of these two processes required to characterize pseudo-optimal risk-minimizing strategies by analogy with discrete time.

### 7.4.1 The $f$ -Costs Process (inclusive of liquidity costs)

For a general trading strategy  $\phi$  we define the  $f$ -costs process  $C_t^f(\phi)$  as the following limit, whenever it exists

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{l_n} f' \left( \mathcal{L}(\delta^{\tau_k^n} - \delta^{\tau_{k-1}^n}, X^{\tau_k^n}) + \beta^{\tau_k^n} - \beta^{\tau_{k-1}^n} \right)$$

where convergence happens in ucp topology, for any sequences  $\mathcal{P}_n$  of Riemann partitions of  $[0, T]$  of length  $l_n$  (i.e.  $\tau_{l_n}^n = T$ ). We used the notation  $X^T$  for the process stopped at  $T$ .



## 7. LIQUIDITY

---

We now restrict our attention to  $H$ -admissible strategies and show the existence of its  $f$ -costs process.

**Theorem 11.** *The  $f$ -costs process of an  $H$ -admissible strategy  $\phi$  is well defined and is given by the following formula*

$$\begin{aligned}
C_t^f(\phi) = & \\
& f''(0) \left( V_t - V_0 - \int_{0+}^t \delta_s - dX_s + \frac{1}{2} l'(0) \int_{0+}^t X_s - d[\delta, \delta]_s^c \right) \\
& + \frac{f^{(3)}(0)}{2} \left( [\beta, \beta]_t^c + 2 \int_{0+}^t X_s - d[\beta, \delta]_s^c + \int_{0+}^t X_s^2 - d[\delta, \delta]_t^c \right) \\
& + \sum_{0 < s \leq t} f'(\Delta\beta_s + \mathcal{L}(\Delta\delta_s, X_s)) - f''(0)(\Delta\beta_s + \Delta\delta_s X_s)
\end{aligned} \tag{7.5}$$

with notation  $[X, Y]^c$  standing for the continuous part of the (càdlàg) quadratic covariation process.

*Proof.* Though very similar to the proof given for the  $f$ -costs process expression in case of infinite liquidity, we give the proof for finite liquidity for completeness. Let  $\mathcal{P}_n$  be a refining sequence of Riemann partitions of  $[0, T]$ ,  $\mathcal{P}_n = \{0 = t_0^n \leq \dots \leq t_{l_n}^n = T\}$ .

$$\begin{aligned}
C_t^{\mathcal{P}_n}(\phi) = & \\
& \sum_{k=1}^{l_n} f'(\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}})
\end{aligned}$$

Since  $\beta$ ,  $\delta$  and  $X$  are càdlàg process, and  $\sum_s (\Delta\beta_s)^2$ ,  $\sum_s (\Delta\delta_s)^2$  and  $\sum_s (\Delta X_s)^2$  are (absolutely) convergent series, given  $\epsilon > 0$  we can find two sets  $A$  and  $B$  such that  $A$  and  $B$  are disjoint and  $A \cup B$  exhausts the jump times of  $\beta$ ,  $\delta$  and  $S$  on  $(0, T]$ ,  $A$  being a set of jump times that  $\beta$ ,  $\delta$  and  $S$  have a.s. a finite number of times and  $B$  being such that  $\sum_{0 < s \leq t} (\Delta\beta)^2 \leq \epsilon^2$ ,  $\sum_{0 < s \leq t} (\Delta\delta)^2 \leq \epsilon^2$  and  $\sum_{0 < s \leq t} (\Delta X)^2 \leq \epsilon^2$ .

Thus we have

$$\begin{aligned}
C_t^{\mathcal{P}_n}(\phi) = & \\
& \sum_{k,A} f'(\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) \\
& + \sum_{k,B} f'(\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}})
\end{aligned}$$

where  $\sum_{k,A}$  denotes  $\sum_k 1_{\{A \cap (t_{k-1}, t_k] \neq \emptyset\}}$  and  $\sum_{k,B}$  denotes  $\sum_k 1_{\{B \cap (t_{k-1}, t_k] \neq \emptyset\}}$ . The first sum converges to  $\sum_{s \in A} f'(\mathcal{L}(\Delta\delta_s, X_s) + \Delta\beta_s)$ .

In the second sum we apply Taylor's theorem to  $f'$  and to  $\mathcal{L}$  seen as a function of its first variable

$$f'(x) = f''(0)x + \frac{1}{2}f^{(3)}(0)x^2 + R(x) \quad (7.6)$$

$$\mathcal{L}(x) = l(0)x + \frac{1}{2}l'(0)x^2 + R_{\mathcal{L}}(x) \quad (7.7)$$

where  $|R(x)| \leq r(x)x^2$ , such that  $r : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an increasing function with  $\lim_{u \downarrow 0} r(u) = 0$  and likewise for  $R_{\mathcal{L}}$ . Thus we have

$$\begin{aligned} & \sum_{k,B} f'(\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) = \\ & f''(0) \sum_{k,B} (\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) \end{aligned} \quad (7.8)$$

$$+ \frac{1}{2}f^{(3)}(0) \sum_{k,B} (\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}})^2 \quad (7.9)$$

$$+ \sum_{k,B} R(\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) \quad (7.10)$$

The first sum (7.8) is equal to

$$\begin{aligned} & \sum_k (\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) \\ & - \sum_{k \in A} (\mathcal{L}(\delta_{t_k} - \delta_{t_{k-1}}, X_{t_k}) + \beta_{t_k} - \beta_{t_{k-1}}) \end{aligned}$$

which converges in ucp topology to

$$\begin{aligned} & \left( V_t - V_0 - \int_{0+}^t \delta_s - dX_s + \frac{1}{2}l'(0) \int_{0+}^t X_s - d[\delta, \delta]_s^c \right) \\ & - \sum_{s \in A} (\mathcal{L}(\Delta \delta_s, X_s) + \Delta \beta_s) \end{aligned}$$

The second sum (7.9), after developing and switching to less cumbersome notations, is equal to

$$\sum_{k,B} (\beta_k - \beta_{k-1})^2 + 2(\beta_k - \beta_{k-1})\mathcal{L}(\delta_k - \delta_{k-1}, X_k) + \mathcal{L}(\delta_k - \delta_{k-1}, X_k)^2$$

$\sum_{k,B} (\beta_k - \beta_{k-1})^2 = \sum_k (\beta_k - \beta_{k-1})^2 - \sum_{k,A} (\beta_k - \beta_{k-1})^2$  and the first sum converges to  $[\beta, \beta]_t$  while  $\sum_{k \in A} (\beta_k - \beta_{k-1})^2$  converges to  $\sum_{s \in A} \Delta \beta_s^2$ . Now  $\sum_{k,B} 2(\beta_k - \beta_{k-1})\mathcal{L}(\delta_k - \delta_{k-1}, X_k) = \sum_{k,B} 2X_{k-1}(\beta_k - \beta_{k-1})(\delta_k - \delta_{k-1}) + \sum_{k,B} (X_k - X_{k-1})(\beta_k - \beta_{k-1})(\delta_k - \delta_{k-1})$ . The first term is equal to  $\sum_k 2X_{k-1}(\beta_k - \beta_{k-1})(\delta_k - \delta_{k-1}) - \sum_{k,A} 2X_{k-1}(\beta_k - \beta_{k-1})(\delta_k -$

## 7. LIQUIDITY

$\delta_{k-1}$ ) and converges to  $2 \int_{0+}^t X_s d[\beta, \delta]_s - 2 \sum_{s \in A} X_s \Delta \beta_s \Delta \delta_s$ . The second term is less than  $\sup_{k,B} |(X_k - X_{k-1})| \sum_{k,B} |\beta_k - \beta_{k-1}| |\delta_k - \delta_{k-1}|$  again less than  $\sup_{k,B} |(X_k - X_{k-1})| (\sum_k (\beta_k - \beta_{k-1})^2 + \sum_k (\delta_k - \delta_{k-1})^2)$ . Taking the limit when  $n \rightarrow \infty$  we find that  $|\sum_{k,B} (X_k - X_{k-1})(\beta_k - \beta_{k-1})(\delta_k - \delta_{k-1})| \leq \epsilon \sqrt{[\delta, \delta]_t} \sqrt{[\beta, \beta]_t}$ . Finally

$$\begin{aligned} & \sum_{k,B} X_k^2 (\delta_k - \delta_{k-1})^2 = \sum_{k,B} X_{k-1}^2 (\delta_k - \delta_{k-1})^2 \\ & + 2 \sum_{k,B} X_{k-1} (X_k - X_{k-1}) (\delta_k - \delta_{k-1})^2 + \sum_{k,B} (X_k - X_{k-1})^2 (\delta_k - \delta_{k-1})^2 \end{aligned}$$

The first term is equal to  $\sum_k X_{k-1}^2 (\delta_k - \delta_{k-1})^2 - \sum_{k,A} X_{k-1}^2 (\delta_k - \delta_{k-1})^2$  and converges to  $\int_{0+}^t X_s^2 d[\delta, \delta]_s - \sum_{s \in A} X_s^2 (\Delta \delta_s)^2$ . The second term is less than  $\sup_{k,B} |X_k| \sup_{k,B} |(X_k - X_{k-1})| (\sum_k (\delta_k - \delta_{k-1})^2)$  and if we assume for now that  $S \leq K < \infty$  uniformly in  $t$  then we have  $|\sum_{k,B} X_{k-1} (X_k - X_{k-1}) (\delta_k - \delta_{k-1})^2| \leq K \epsilon [\delta, \delta]_t$ . The last term is less than  $\epsilon^2 [\delta, \delta]_t$  by following the same reasoning. Now we turn to the last term (7.10) of the Taylor's development

$$\begin{aligned} & \left| \sum_{k,B} R(\beta_{t_k} - \beta_{t_{k-1}} + (\delta_{t_k} - \delta_{t_{k-1}}) X_{t_k}) \right| \\ & \leq \sum_{k,B} r(|\beta_{t_k} - \beta_{t_{k-1}} + (\delta_{t_k} - \delta_{t_{k-1}}) X_{t_k}|) (\beta_{t_k} - \beta_{t_{k-1}} + (\delta_{t_k} - \delta_{t_{k-1}}) X_{t_k})^2 \quad (7.11) \end{aligned}$$

Again assuming that  $\sup X \leq K \leq \infty$  over  $[0, T]$  we have (7.11)  $\leq \sup r((K+1)\epsilon) [\delta, \delta]_t$ . We are now ready to take the limit when  $\epsilon$  goes to zero. The last term tends to zero from the property of  $r$  and it remains to prove that the series  $\sum_{s \in A}$  are absolutely convergent. We next proceed by localization, as in Protter (44) by considering first  $V_K = \inf\{t > 0, |\delta| > K\}$ ,  $W_K = \inf\{t > 0, |\beta| > K\}$  and  $Z_K = \inf\{t > 0, |X| > K\}$  so that  $1_{[0, V_K]} \delta$ ,  $1_{[0, W_K]} \beta$  and  $1_{[0, Z_K]} X$  are  $[-K, K]$ -valued. Therefore we have that  $|f'(x) - f''(0)x| \leq Cx^2$  for some constant  $C$ . This allows us to write

$$\begin{aligned} & \left| \sum_{s \in A} f'(\Delta \beta_s + \Delta \delta X_s) - f''(0) \left( \sum_{s \in A} \Delta \beta_s + \Delta \delta X_s \right) \right| \\ & \leq C \sum_{s \in A} \Delta \beta_s^2 + 2 \Delta \beta_s \Delta \delta X_s + \Delta \delta_s^2 X_s^2 \\ & \leq C([\beta, \beta]_t + 2K[\delta, \beta]_t + K^2[\delta, \delta]_t) < \infty \end{aligned}$$

And the series are absolutely convergent which completes the proof.  $\square$

**Corollary 2.** *The  $f$ -costs process of an  $H$ -admissible strategy  $\phi$  can also be expressed*

in terms of the portfolio value  $V$

$$\begin{aligned}
 C_t^f(\phi) &= f''(0) \left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) + f''(0)l'(0) \left( \frac{1}{2} \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \right) \\
 &\quad + \frac{f^{(3)}(0)}{2} \left( [V, V]_t^c - 2 \int_{0+}^t \delta_{s-} d[V, X]_s^c + \int_{0+}^t \delta_{s-}^2 d[X, X]_t^c \right) \\
 &\quad + \sum_{0 < s \leq t} f'(\Delta V_s - \delta_{s-} \Delta X_s + \mathcal{L}(\Delta \delta_s, X_s) - \Delta \delta_s X_s) \\
 &\quad - \sum_{0 < s \leq t} f''(0)(\Delta V_s - \delta_{s-} \Delta X_s) \tag{7.12}
 \end{aligned}$$

*Proof.* The proof is a straightforward application of quadratic variation properties when expressing  $\beta$  as a function of  $V$  in formula (7.5).  $\square$

**Remark 20.** *The additional term in the expression of  $C_t^f(\phi)$  due to the finite liquidity is  $f''(0)l'(0) \left( \frac{1}{2} \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \right)$  and it is non-decreasing given the convexity of both  $f$  and  $\mathcal{L}$ .*

#### 7.4.2 The supply price process

For an  $H$ -admissible trading strategy  $\phi$  we define the supply price process  $X_t^S(\phi)$  as the following limit, whenever it exists

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{l_n} \left( l(\delta^{\tau_k^n} - \delta^{\tau_{k-1}^n}) X^{\tau_k^n} - X^{\tau_{k-1}^n} \right)$$

where convergence happens in ucp topology, for any sequences  $\mathcal{P}_n$  of Riemann partitions of  $[0, T]$  of length  $l_n$ .

For an  $H$ -admissible strategy  $\phi$  we have a similar theorem relative to the existence of the supply price process as for the  $f$ -costs.

**Theorem 12.** *The supply price process  $X^S$  of an  $H$ -admissible strategy  $\phi$  is well defined and is given by the following formula*

$$\begin{aligned}
 X_t^S(\phi) &= X_t + l'(0) \left( \delta_t X_t - \delta_0 X_0 - \int_{0+}^t \delta_{s-} dX_s \right) + \frac{1}{2} l''(0) \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \\
 &\quad + \sum_{0 < s \leq t} (l((\Delta \delta_s) - 1) X_s - l'(0) \Delta \delta_s X_s) \tag{7.13}
 \end{aligned}$$

*Proof.* The proof follows the same lines as theorem 11 so we do not detail it here.  $\square$

## 7.5 Application to stochastic volatility models

In order to derive an explicit formula for the  $f$ -costs and supply price processes and completely characterize pseudo-optimal strategies for the local risk-minimization, we will need to introduce further assumptions on the evolution of  $X$ .

We start with the same setting as in chapter 5, where we recall that  $(X, Y)$  is described by the following set of SDEs

$$dX_s = a(s, X_s, Y_s)ds + b(s, X_s, Y_s)dW_s^1 \quad (7.14)$$

$$dY_s = c(s, X_s, Y_s)dt + d(s, X_s, Y_s) \left( \rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2 \right) \quad (7.15)$$

with initial conditions  $X_0 = x$ ,  $Y_0 = y$  and  $(W^1, W^2)$  a standard two-dimensional Wiener process under  $P$ . With this prescription the stochastic factor  $Y$  of the volatility has constant instantaneous correlation  $\rho$  with  $X$ , i.e.  $d \langle X, Y \rangle_t = \rho dt$ . We will assume that  $a, b, c$  and  $d$  are Lipschitz continuous functions on  $(0, T) \times \mathbb{R}^2$  taking values in  $\mathbb{R}$  and that there exists a constant  $C$  such that for  $t \in [0, T]$  and  $x, y, x', y' \in \mathbb{R}$

$$\begin{aligned} & |a(t, x, y) - a(t, x', y')| + |b(t, x, y) - b(t, x', y')| \\ & + |c(t, x, y) - c(t, x', y')| + |d(t, x, y) - d(t, x', y')| \leq C (|x - x'| + |y - y'|) \end{aligned} \quad (7.16)$$

$$|a(t, x, y)|^2 + |b(t, x, y)|^2 + |c(t, x, y)|^2 + |d(t, x, y)|^2 \leq C (1 + |x|^2 + |y|^2) \quad (7.17)$$

These assumptions ensure existence and uniqueness of a strong solution to the system of SDE (7.14), (7.15) and the continuity of the flow  $(t, x, y) \mapsto (X_s^{t,x,y}, Y_s^{t,x,y})$ , where  $X_s^{t,x,y}$  and  $Y_s^{t,x,y}$  are the solutions with initial data  $X_t = x$  and  $Y_t = y$  (see Platen and Kloeden (32)). These properties in turn ensure that the solution is a strong Markov process.

With these diffusion assumptions we will now place ourselves in a Markovian framework and look for the optimal strategy  $\phi$  as a smooth function of the state variables

$$\delta_t = \delta(t, X_t, Y_t)$$

$$V_t = V(t, X_t, Y_t)$$

## 7.5.1 PDE formulation

So as to derive a set of PDEs satisfied by pseudo-optimal strategies, we first express the  $f$ -costs process as a function of the diffusion parameters and the strategy

$$\begin{aligned}
 C_t^f(\phi) = & \int_0^t \left[ f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \right. \\
 & + f''(0) l'(0) \frac{X_u}{2} \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial \delta}{\partial Y} \rho b_u d_u \right) \\
 & + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
 & \left. - f^{(3)}(0) \delta_u \left( \frac{\partial V}{\partial X} \sigma_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 \right] du \\
 & + \int_0^t f''(0) \left( \frac{\partial V}{\partial X} - \delta_u \right) b_u dW_u^1 + \int_0^t f''(0) \frac{\partial V}{\partial Y} d_u dW_u^2
 \end{aligned}$$

which follows from equation (7.12).

Likewise we express the supply price process

$$\begin{aligned}
 X_t^S(\phi) = & X_t + l'(0) \left( \delta_t X_t - \delta_0 X_0 - \int_0^t \delta_u a_u du - \int_0^t \delta_u b_u dW_u^1 \right) \\
 & + \frac{1}{2} l''(0) \int_0^t \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial \delta}{\partial Y} \rho b_u d_u \right) du
 \end{aligned}$$

which is derived from equation (7.13).

Now, applying to the strategy  $\phi$  the first pseudo-optimality criterion, *i.e.* that  $C$  must be martingale under the measure  $P$ , we find a first fully non-linear PDE satisfied by the strategy  $(V, \delta)$

$$\begin{aligned}
 f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \\
 + f''(0) l'(0) \frac{X_u}{2} \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial \delta}{\partial Y} \rho b_u d_u \right) \\
 + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
 - f^{(3)}(0) \delta_u \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 = 0
 \end{aligned}$$

with terminal condition  $V_T = H$ .

## 7. LIQUIDITY

---

In order to apply to the strategy  $\phi$  the second pseudo-optimality criterion, *i.e.* that the martingale  $C$  must be orthogonal to the martingale part of the supply price process  $X^S$ , we first identify its martingale part

$$X_t^S(\phi) - \mathbb{E}(X_t^S(\phi)) = \int_0^t \left(1 + l'(0)X \frac{\partial \delta}{\partial X}\right) b_u dW_u^1 + \int_0^t l'(0)X \frac{\partial \delta}{\partial Y} d_u dW_u^2$$

so that the second PDE satisfied by the strategy  $(V, \delta)$  is

$$\begin{aligned} \left(\frac{\partial V}{\partial X} - \delta\right) \left(1 + l'(0)X \frac{\partial \delta}{\partial X}\right) b^2 + \frac{\partial V}{\partial Y} \left(1 + l'(0)X \frac{\partial \delta}{\partial X}\right) \rho b d + \\ \left(\frac{\partial V}{\partial X} - \delta\right) \frac{\partial \delta}{\partial Y} l'(0)X \rho b d + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0)X d^2 = 0 \end{aligned}$$

With some rearrangements, the pseudo-optimal strategy  $\phi$  finally solves the following system of parabolic and hyperbolic PDEs

$$\left\{ \begin{array}{l} \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a + \frac{\partial V}{\partial Y} b + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b d = \\ \delta a + \alpha \left( \left( \frac{\partial V}{\partial X} b + \frac{\partial V}{\partial Y} \rho d - \delta b \right)^2 + (1 - \rho^2) \left( \frac{\partial V}{\partial Y} \right)^2 d^2 \right) \\ + l'(0) \frac{X}{2} \left( \left( \frac{\partial \delta}{\partial X} b + \frac{\partial \delta}{\partial Y} \rho d \right)^2 + (1 - \rho^2) \left( \frac{\partial \delta}{\partial Y} \right)^2 d^2 \right) \\ \left( \frac{\partial V}{\partial X} - \delta \right) \left( 1 + l'(0)X \frac{\partial \delta}{\partial X} \right) b^2 + \frac{\partial V}{\partial Y} \left( 1 + l'(0)X \frac{\partial \delta}{\partial X} \right) \rho b d \\ + \left( \frac{\partial V}{\partial X} - \delta \right) \frac{\partial \delta}{\partial Y} l'(0)X \rho b d + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0)X d^2 = 0 \end{array} \right. \quad (7.18)$$

with  $V_T = H$ .

### Complete markets case

We investigate the case of complete markets by setting the volatility of volatility  $d$  equal to zero. The equation for the hedge ratio  $\delta$  then reduces to

$$\left(\frac{\partial V}{\partial X} - \delta\right) \left(1 + l'(0)X \frac{\partial \delta}{\partial X}\right) = 0$$

so that a sufficient condition is that  $V, \delta$  is a solution to

$$\delta = \frac{\partial V}{\partial X} \quad (7.19)$$

$$\frac{\partial V}{\partial u} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} \sigma^2 \left(1 + l'(0)X \frac{\partial^2 V}{\partial X^2}\right) = 0 \quad (7.20)$$

Upon the generalized Black and Scholes PDE ((5), (41)) (7.20) having a solution, which is expected when the contingent claim has a convex payoff, equation (7.19) gives the perfect hedging strategy in that context. As in the “infinite” liquidity case, the solution does not depend on the choice for the function  $f$ . An easy calculation allows to verify that in this case the  $f$ -costs process is identically zero, which amounts to having a self-financing strategy incorporating liquidity costs that perfectly replicates the contingent claim  $H$ . Another remark is that the PDE (7.20) shows that the value of the portfolio being an increasing function of the volatility for a convex payoff, in the presence of liquidity costs, is increased proportionally to the slope of the marginal costs and to the  $\Gamma = \frac{\partial^2 V}{\partial X^2}$  of the option.

### 7.5.2 The minimization problem

Despite the fact that in discrete time a pseudo-optimal strategy, satisfying the set of equations (7.3), might not be optimal, in continuous time, when working with continuous path processes we have a correspondence between the two concepts. First and foremost we need to redefine the notion of optimality we are concerned with, in particular to take into account the fact we did not define the costs of a strategy in continuous time <sup>1</sup>.

Given a partition  $\tau$  of  $[0, T]$ , where  $\tau = \{0 = t_0, t_1, \dots, t_k = T\}$ , and a small perturbation  $\Delta$ , we define the process  $r_f^\tau$  as:

**Definition 42.** *The  $f$ -risk quotient (inclusive of liquidity costs) of a trading strategy  $\phi$  along the partition  $\tau$  is the process*

$$r_f^\tau[\phi, \Delta](t, \omega) := \sum_{t_i, t_{i+1} \in \tau} \frac{\Delta R_{t_i}(\phi + \Delta|_{(t_i, t_{i+1}]})(\omega) - \Delta R_{t_i}(\phi)(\omega)}{t_{i+1} - t_i} 1_{(t_i, t_{i+1}]}(t) \quad (7.21)$$

with  $\Delta R_{t_i}(\Phi) = \mathbb{E}(f(\Delta C_{t_{i+1}})|\mathcal{F}_{t_i})$ .

And optimality is then defined the usual way:

**Definition 43.** *For a contingent claim  $H$ , a trading strategy  $\phi$  generating  $H$  is called locally risk-minimizing if for every small perturbation  $\Delta$  and every increasing sequence of partitions  $(\tau_n)_{n \in \mathbb{N}}$  tending to the identity, we have*

$$\liminf_{n \rightarrow \infty} r_f^{\tau_n}[\phi, \Delta] \geq 0 \quad \mathcal{P} - a.e. \quad (7.22)$$

---

<sup>1</sup>see remark 21 at the end of the chapter for more comments on this



## 7. LIQUIDITY

---

Given the smoothness of the risk function  $f$  and the liquidity costs function  $\mathcal{L}$  we can rewrite the process  $r_f^\tau$  by using a Taylor development around the non-perturbed strategy  $\phi$ . Let  $\Gamma = (\beta, \delta)$  be a small perturbation and let us fix  $t \in [0, T]$ . Because of the definition of the process  $r_f^\tau[\phi, \Gamma]$  and as we work with increasing sequences of partitions, we may assume that  $t$  is one of the  $t_{i(n)}^n$  (we will thereafter drop the superscript  $n$  and simply write  $t_i$  instead), we have

$$\begin{aligned} r_f^\tau[\phi, \Gamma](t, \omega) &= \frac{\Delta R_{t_i}(\phi + \Gamma|_{[t_i, t_{i+1}]}) (\omega) - \Delta R_{t_i}(\phi) (\omega)}{t_{i+1} - t_i} \\ &= \frac{\mathbb{E}_{t_i} (f(\Delta C_{t_{i+1}}(\phi + \Gamma|_{[t_i, t_{i+1}]}) (\omega)) - \mathbb{E}_{t_i} (f(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} \end{aligned}$$

Applying Taylor's formula with remainder term to  $g : (x, y) \mapsto f(\mathcal{L}(x) + y)$  in the expectation, we have that

$$\begin{aligned} f(\Delta C_{t_{i+1}}(\phi + \Gamma|_{[t_i, t_{i+1}]}) (\omega)) &= f(\Delta C_{t_{i+1}}(\phi)) - \beta_{t_i} f'(\Delta C_{t_{i+1}}(\phi)) - \delta_{t_i} \mathcal{L}'(\phi) f'(\Delta C_{t_{i+1}}(\phi)) \\ &\quad + \frac{1}{2} \delta_{t_i}^2 h(\tilde{\phi}) + \frac{1}{2} (\beta_{t_i} + \delta_{t_i} \mathcal{L}'(\tilde{\phi}))^2 g(\tilde{\phi}) \end{aligned}$$

where  $g(\tilde{\phi}) = f''(\Delta C_{t_{i+1}}(\tilde{\phi}))$  and  $h(\tilde{\phi}) = \mathcal{L}''(\tilde{\phi}) f'(\Delta C_{t_{i+1}}(\tilde{\phi}))$  with  $\tilde{\phi} = (\tilde{\beta}, \tilde{\delta})$  such that  $|\tilde{\beta}| \leq \beta$  and  $|\tilde{\delta}| \leq \delta$ . With the assumptions on  $f \in \mathcal{R}$ , namely  $f'(0) = 0$  and  $f''(0) > 0$ , the remainder term  $\delta_{t_i}^2 h(\tilde{\phi}) + (\beta_{t_i} + \delta_{t_i} \mathcal{L}'(\tilde{\phi}))^2 g(\tilde{\phi})$  will remain strictly positive in a neighborhood of  $t_i$  for  $\delta_{t_i}$  and  $\beta_{t_i}$  small enough.

Rearranging and simplifying we get

$$\begin{aligned} r_f^\tau[\phi, \Gamma](t, \omega) &= \beta_{t_i} \frac{\mathbb{E}_{t_i} (f'(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} + \delta_{t_i} \frac{\mathbb{E}_{t_i} (\mathcal{L}'(\phi) f'(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} \\ &\quad + \frac{1}{2} \frac{\mathbb{E}_{t_i} (\delta_{t_i}^2 h(\tilde{\phi})) (\omega)}{t_{i+1} - t_i} + \frac{1}{2} \frac{\mathbb{E}_{t_i} ((\beta_{t_i} + \delta_{t_i} \mathcal{L}'(\tilde{\phi}))^2 g(\tilde{\phi})) (\omega)}{t_{i+1} - t_i} \end{aligned}$$

Since we work with Itô processes, the following stands

$$\begin{aligned} \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} (f'(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} &= \Lambda (f' \circ \Delta C)_{t_i} \\ \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} (\mathcal{L}'(\phi) f'(\Delta C_{t_{i+1}}(\phi))) (\omega)}{t_{i+1} - t_i} &= \Lambda (\mathcal{L}' \cdot f' \circ \Delta C)_{t_i} \end{aligned}$$

and

$$\begin{aligned} \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} \left( h(\tilde{\phi}) \right) (\omega)}{t_{i+1} - t_i} &= \Lambda h_{t_i} \\ \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} \left( g(\tilde{\phi}) \right) (\omega)}{t_{i+1} - t_i} &= \Lambda g_{t_i} \\ \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} \left( \mathcal{L}' g(\tilde{\phi}) \right) (\omega)}{t_{i+1} - t_i} &= \Lambda (\mathcal{L}' \cdot g)_{t_i} \\ \lim_{t_{i+1} \rightarrow t_i} \frac{\mathbb{E}_{t_i} \left( \mathcal{L}'^2 g(\tilde{\phi}) \right) (\omega)}{t_{i+1} - t_i} &= \Lambda (\mathcal{L}'^2 \cdot g)_{t_i} \end{aligned}$$

where  $\Lambda$  is the infinitesimal generator associated with the diffusion (7.14, 7.15)

$$\Lambda h = \frac{\partial h}{\partial X} a + \frac{\partial h}{\partial Y} c + \frac{1}{2} \frac{\partial^2 h}{\partial X^2} b^2 + \frac{1}{2} \frac{\partial^2 h}{\partial Y^2} d^2 + \frac{\partial^2 h}{\partial X \partial Y} \rho b d \quad (7.23)$$

Finally the process  $r_f^\tau$  evaluated in  $t$  is worth

$$\begin{aligned} r_f^\tau[\phi, \Gamma](t, \omega) &= \beta_t \Lambda (f' \circ \Delta C)_t + \delta_t \Lambda (\mathcal{L}' \cdot f' \circ \Delta C)_t \\ &\quad + \frac{1}{2} (\beta_t^2 \Lambda g_t + 2\beta_t \delta_t \Lambda (\mathcal{L}' \cdot g)_t + \delta_t^2 \Lambda (\mathcal{L}'^2 \cdot g + h)_t) \end{aligned}$$

Now we first take the component  $\delta$  of the perturbation equal to zero, that is we perturb only  $\beta$ , so that we have the following first condition for the strategy  $\phi$  to be locally risk-minimizing

$$\beta_t \Lambda (f' \circ \Delta C)_t + \frac{1}{2} \beta_t^2 \Lambda g_t \geq 0 \quad P - a.e. \quad \forall \beta_t$$

As a consequence we must have  $\Lambda (f' \circ \Delta C)_t = 0$ .

Likewise we take the component  $\beta$  equal to zero and we get the following second condition for the strategy  $\phi$  to be locally risk-minimizing

$$\delta_t \Lambda (\mathcal{L}' \cdot f' \circ \Delta C)_t + \frac{1}{2} \delta_t^2 \Lambda (\mathcal{L}'^2 \cdot g + h)_t \geq 0 \quad P - a.e. \quad \forall \delta_t$$

Therefore we must have  $\Lambda (\mathcal{L}' \cdot f' \circ \Delta C)_t = 0$ . But we observe that

$$\begin{cases} \Lambda (f' \circ \Delta C)_t = 0 \\ \Lambda (\mathcal{L}' \cdot f' \circ \Delta C)_t = 0 \end{cases}$$

$\Leftrightarrow$

## 7. LIQUIDITY

---

$$\left\{ \begin{array}{l} f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \\ \quad + f''(0) l'(0) \frac{X}{2} \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial \delta}{\partial Y} \rho b_u d_u \right) \\ \quad + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\ \quad - f^{(3)}(0) \delta_u \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 = 0 \\ \left( \frac{\partial V}{\partial X} - \delta \right) \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) b_u^2 + \frac{\partial V}{\partial Y} \left( 1 + l'(0) X \frac{\partial \delta}{\partial X} \right) \rho b_u d_u \\ \quad + \left( \frac{\partial V}{\partial X} - \delta \right) \frac{\partial \delta}{\partial Y} l'(0) X \rho b_u d_u + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0) X d_u^2 = 0 \end{array} \right.$$

Finally, just as in the “infinite” liquidity case, we see that in this context of stochastic volatility model, the optimal strategies with respect to the local risk-minimization problem are the same as the pseudo-optimal strategies. Likewise, the only requirement to obtain this result is the existence of the infinitesimal generator and its expression in terms of the parameters of the SDE driving the process so as to identify the sets of two equations, which therefore allows again for a straightforward generalization to more general Itô processes.

### 7.6 Application to stochastic volatility/jump diffusion models

Now we study pseudo-optimal strategies in a situation where the stock process may exhibit jumps so as to demonstrate that the global behaviour of the risk function  $f$  can also have an impact. To this end we model the evolution of  $X$  through an SDE with stochastic volatility and Poisson jumps as in chapter 6

$$dX_s = a(s, X_{s-}, Y_{s-}) ds + b(s, X_{s-}, Y_{s-}) dW_s^1 + k dN_s \quad (7.24)$$

$$dY_s = c(s, X_{s-}, Y_{s-}) ds + d(s, X_{s-}, Y_{s-}) \left( \rho dW_s^1 + \sqrt{1 - \rho^2} dW_s^2 \right) \quad (7.25)$$

with initial conditions  $X_0 = x$ ,  $Y_0 = y$  and  $(W^1, W^2)$  a standard two-dimensional Wiener process under  $P$ .  $N_t$  is a Poisson process of intensity  $\lambda$  and the amplitude of the jumps  $k$  has probability distribution  $K$ . We also assume that  $W_t = (W_t^1, W_t^2)$ ,  $N_t$  and  $k$  are independent. With this prescription the stochastic factor  $Y$  of the volatility has constant instantaneous correlation  $\rho$  with  $X$ , i.e.  $d \langle X, Y \rangle_t = \rho dt$ . As in the

---

## 7.6 Application to stochastic volatility/jump diffusion models

above case of stochastic volatility we shall assume that appropriate conditions hold on the adapted processes  $a, b, c, d, K$  and  $\lambda$  so that the set of SDEs has a unique strong solution.

With these assumptions we will again place ourselves in a Markovian framework and look for the optimal strategy  $\phi$  as a smooth function of the state variables

$$\delta_t = \delta(t, X_t, Y_t)$$

$$V_t = V(t, X_t, Y_t)$$

### 7.6.1 PIDE formulation

So as to derive a set of PIDEs satisfied by pseudo-optimal strategies, we first express the  $f$ -costs process as a function of the diffusion parameters and the strategy

$$\begin{aligned} C_t^f(\phi) &= \int_0^t \left( f''(0) \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u a_u \right) \right. \\ &\quad + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\ &\quad + \frac{l'(0)X}{2} \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial \delta}{\partial Y} \rho b_u d_u \right) \\ &\quad \left. - f^{(3)}(0) \delta_{u-} \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 b_u^2 \right) du \\ &\quad + \int_0^t f''(0) \left( \frac{\partial V}{\partial X} - \delta_{u-} \right) b_u dW_u^1 + \int_0^t f''(0) \frac{\partial V}{\partial Y} d_u dW_u^2 \\ &\quad + \int_0^t \int_{\mathbb{R}} f'(\Delta V_u - \delta_{u-} \Delta X_u + \mathcal{L}(\Delta \delta_u, X_u) - \Delta \delta_u X_u) K(k) dk dN_u \end{aligned}$$

which we have obtained from equation (7.12), and with  $\Delta V_u$  the jump in  $V$  when there is a jump  $\Delta X_u$  of size  $k$  on  $X$  at time  $u$  being equal to  $V(u-, X_{u-} + k, Y_{u-}) - V(u-, X_{u-}, Y_{u-})$  and likewise for  $\Delta \delta_u$ .

Now, applying to the strategy  $\phi$  the first pseudo-optimality criterion, *i.e.* that  $C$  must be martingale under the measure  $P$ , we find the PIDE satisfied by the portfolio

## 7. LIQUIDITY

---

value  $V$

$$\begin{aligned}
f''(0) & \left( \frac{\partial V}{\partial u} + \frac{\partial V}{\partial X} a_u + \frac{\partial V}{\partial Y} c_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} b_u^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} d_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u d_u - \delta_u - a_u \right) \\
& + \frac{f^{(3)}(0)}{2} \left( \left( \frac{\partial V}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial V}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial V}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
& + \frac{l'(0)X}{2} \left( \left( \frac{\partial \delta}{\partial X} \right)^2 b_u^2 + \left( \frac{\partial \delta}{\partial Y} \right)^2 d_u^2 + 2 \frac{\partial \delta}{\partial X} \frac{\partial V}{\partial Y} \rho b_u d_u \right) \\
& - f^{(3)}(0) \delta_u - \left( \frac{\partial V}{\partial X} b_u^2 + \frac{\partial V}{\partial Y} \rho b_u d_u \right) + \frac{f^{(3)}(0)}{2} \delta_u^2 - b_u^2 \\
& + \int_{\mathbb{R}} f'(\Delta V_u - \delta_u - \Delta X_u) K(k) dk \lambda_u = 0
\end{aligned}$$

with terminal condition  $V_T = H$ .

In order to apply to the strategy  $\phi$  the second pseudo-optimality criterion, *i.e.* that the martingale  $C$  must be orthogonal to the martingale part of the supply price process  $X^S$ , we first identify its martingale part

$$\begin{aligned}
X_t^S(\phi) - \mathbb{E}(X_t^S(\phi)) & = \int_0^t \left( 1 + l'(0)X \frac{\partial \delta}{\partial X} \right) b_u dW_u^1 + \int_0^t l'(0)X \frac{\partial \delta}{\partial Y} d_u dW_u^2 \\
& + \int_0^t \int_{\mathbb{R}} ((l(\Delta \delta_u) - 1)X_u + k) K(k) dk d\tilde{N}_u
\end{aligned}$$

with  $\tilde{N}$  the compensated Poisson process of  $N$ .

Therefore the second PIDE satisfied by the strategy  $(V, \delta)$  is

$$\begin{aligned}
& \left( \frac{\partial V}{\partial X} - \delta \right) \left( 1 + l'(0)X \frac{\partial \delta}{\partial X} \right) b^2 + \frac{\partial V}{\partial Y} \left( 1 + l'(0)X \frac{\partial \delta}{\partial X} \right) \rho b d \\
& + \left( \frac{\partial V}{\partial X} - \delta \right) \frac{\partial \delta}{\partial Y} l'(0)X \rho b d + \frac{\partial V}{\partial Y} \frac{\partial \delta}{\partial Y} l'(0)X d^2 \\
& + \int_{\mathbb{R}} f'(\Delta V_u - \delta_u - \Delta X_u + \mathcal{L}(\Delta \delta_u, X_u) - \Delta \delta_u X_u) ((l(\Delta \delta_u) - 1)X_u + k) K(k) dk \lambda_u = 0
\end{aligned}$$

We can therefore check that contrarily to the stochastic volatility case, where only the local behaviour of the risk and liquidity costs functions  $f$  and  $\mathcal{L}$  in 0 mattered, finding the optimal strategy in a jump-diffusion model requires the knowledge of both functions on their whole support.

## 7.6 Application to stochastic volatility/jump diffusion models

---

**Remark 21.** *We emphasize that for the case of liquidity costs we took a different route in order to arrive at equations for (pseudo-)optimal strategies. As a matter of fact we made the strong assumption that optimal strategies were “Markovian” which was justified by the strong Markov property of the processes  $X$  and  $Y$ , and then check that the system of PDEs obtained in both cases (optimality and pseudo optimality) were the same. This methodology was inspired by our initial work on non-quadratic local risk-minimization and reflected in our original paper (1). Yet, further to the result obtained in the “infinite” liquidity case, namely theorem (6), we may obtain the same direct relationship (implication) between pseudo-optimal and optimal strategies. The costs process we consider for that purpose is derived from by taking  $f' = Id$  in equation (7.5) to get*

$$C_t(\phi) = \left( V_t - V_0 - \int_{0+}^t \delta_{s-} dX_s \right) + l'(0) \left( \frac{1}{2} \int_{0+}^t X_{s-} d[\delta, \delta]_s^c \right) + \sum_{0 < s \leq t} \mathcal{L}(\Delta\delta_s, X_s) - \Delta\delta_s X_s \quad (7.26)$$

*Optimality would then be defined as in chapter 4, with the risk along the partition being taken as the expectation of the increase of the (continuous) costs process  $C$  between two consecutive times, weighted by function  $f$ . For this implication to hold we need the same assumptions as for theorem (6), namely the existence of a martingale representation for the filtration  $(\mathcal{F}_t)$ . In the case where it is generated by continuous processes, then the orthogonality condition which in the case of finite liquidity is generally not simply on the martingale part of  $X$ , still reduces to the same condition.*

## 7. LIQUIDITY

---

# 8

## Numerical Results and Comparisons

### 8.1 Motivations

The aim of this last chapter is to illustrate numerically the impact of the choice of different risk functions on the optimal strategies. While the strongest point of our method is that it is economically more justified as it allows to put more weight on losses than on gains, being a local minimization of a local risk, we saw in chapter 4 and 5 how we could derive several characterizations of optimal strategies. Particularly in the context of diffusion models optimal strategies may be given as solutions of a quadratic forward backward stochastic differential equation (FBSDE) or alternatively a quasilinear partial differential equation (PDE). Those two characterizations give rise to two numerical methods to compute optimal hedge ratio and portfolio values.

We therefore focus on these two characterizations and show numerical results in a set of two different stochastic volatility models. The two models chosen are taken from the comparative study of quadratic hedging methods undertaken by Heath, Platen and Schweizer (25).

We also present in this chapter one possible way of addressing the issue of pricing, which as previously mentioned, does not have a straightforward answer due to the fact that optimal strategies are not necessarily neither self-financing nor mean self-financing.

Throughout this chapter we consider only European put option as contingent claim so as to have boundedness of the terminal condition in all the equations considered. This



## 8. NUMERICAL RESULTS AND COMPARISONS

---

means that  $H = h(X) = (X_T - K)$  for a fixed strike price  $K$ .

### 8.2 Benchmark Stochastic Volatility Models

We first introduce two of the four stochastic volatility models already used by Heath, Platen and Schweizer (25, 26) as presented in the motivations section. We use the same notations (except for  $\delta$  and  $\beta$  which we replaced by  $\kappa$  and  $\theta$  in models  $S1$ ) and terminology as this will also enable us to verify our result in the case when we choose a quadratic function for  $f$ .

The SDE driving the stock process and its volatility is of the form

$$\begin{aligned}\frac{dX_t}{X_t} &= \mu(t, Y_t)dt + Y_t dW_t \\ dY_t &= a(t, Y_t)dt + b(t, Y_t)dW'_t\end{aligned}$$

with  $P$ -Brownian motions  $W, W'$  with constant instantaneous  $\rho$ , *i.e.*  $d\langle W, W' \rangle_t = \rho dt$  and the choices for the drifts and volatility of volatility are summarized in the table below

Model	Volatility Dynamics $Y$	Appreciation Rate $\mu$
Stein	$dY_t = \kappa(\theta - Y_t)dt + k dW_t$	$\mu(t, Y_t) = \Delta Y_t$
Heston	$d(Y_t)^2 = \kappa(\theta - (Y_t)^2)dt + \Sigma Y_t dW'_t$	$\mu(t, Y_t) = \Delta Y_t$

(8.1)

The references for these two models can be found in Stein and Stein (52) and Heston (27).

The assumptions are that the constants  $k, \kappa, \theta, \Sigma$  are non-negative, with  $\Delta$  and  $\gamma$  real valued and  $\rho \in [-1, 1]$ . The only model with non-zero correlation though is the Heston model.

For the Heston model we also require that the Feller's test for explosions is satisfied, which guarantees the existence and uniqueness of a strictly positive strong solution  $Y$ . This condition reads  $\kappa\theta \geq \frac{1}{2}\Sigma^2$ .

### 8.2.1 Solving the Quadratic Forward-Backward Stochastic Differential Equation

As was shown in chapters 4 and 5 the natural characterization obtained for pseudo-optimal strategy is through a quadratic FBSDE. This was proven in the case when we have a representation theorem for martingales related to the filtration considered and thus holds for stochastic volatility models with the natural filtration of the two state variables  $X$  and  $Y$ .

Therefore we consider in this section the numerical resolution of the following FBSDE

$$\begin{aligned}
 \frac{dX_t}{X_t} &= \mu(t, Y_t)dt + Y_t dW_t^1 \\
 dY_t &= a(t, Y_t)dt + b(t, Y_t)(\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2) \\
 -dV_s &= g(s, X_s, Y_s, V_s, Z_s^1, Z_s^2)ds - Z_s^1 dW_s^1 - Z_s^2 dW_s^2 \\
 V_T &= h(X_T)
 \end{aligned} \tag{8.2}$$

with  $W = (W^1, W^2)$  is a standard two-dimensional Brownian motion and  $g(s, S, \sigma, Y, Z^1, Z^2) = -\frac{\mu}{\sigma}Z^1 - \alpha(Z^2)^2$ .

The literature on numerical schemes for solving a quadratic BSDE or a decoupled FBSDE is not as abundant as in the case of BSDE with Lipschitz drivers. As a matter of fact quadratic BSDE only appeared recently, namely for the pricing and hedging of derivatives in incomplete markets. In most cases, the existing papers are focussing on utility maximization approaches. Yet it is striking that the equations obtained in those frameworks are very close to the one which we obtained in the previous chapters. So we will show how we can use the proposed numerical schemes for our purposes.

To the best of our knowledge two papers propose a numerical method, Imkeller Dos Reis and Zhang (28), and Richou (45), for some fairly general decoupled FBSDE. They both rely on the theory of *BMO martingales* to provide estimate of approximations and thus prove convergence of their numerical schemes.

The article of Richou is concerned with improving the numerical scheme initially proposed by Imkeller Dos Reis and Zhang, through an adapted mesh for the time discretization. Unfortunately the scheme is then proven to work for a special case of volatility function for the forward part, which is a time dependent volatility. Thus this does not apply to our cases of interest. We note however that from the thesis of Richou, where he provides additional numerical results with an actual implementation,

## 8. NUMERICAL RESULTS AND COMPARISONS

---

the scheme with improved time discretization does not seem to perform better than the one with a uniform mesh. So we detail the truncation procedure put in place by Imkeller, Dos Reis and Zhang. For the truncation of the quadratic part, which would otherwise cause troubles for the convergence of the numerical scheme, we introduce the map  $\tilde{h}_n$ , for  $n \in \mathbb{N}$ , which is assumed continuously differentiable and satisfies

- $\tilde{h}_n \rightarrow id$  locally uniformly,  $|\tilde{h}_n| \leq |id|$  and  $|\tilde{h}_n| \leq n + 1$
- 

$$\tilde{h}_n(x) = \begin{cases} (n+1) & , \quad x > n+2 \\ x & , \quad |x| \leq n \\ -(n+1) & , \quad x < -(n+2) \end{cases}$$

- the derivatives of  $\tilde{h}_n$  is absolutely bounded by 1 and converges to 1 locally uniformly.

The construction of such a sequence of functions is given in their paper (28), section 5.

We then set  $h_n : \mathbb{R}^2 \rightarrow \mathbb{R}$  by  $z \mapsto h_n(z) = (\tilde{h}_n(z_1), \tilde{h}_n(z_2))$ ,  $n \in \mathbb{N}$ .

Next we define the truncated driver, for  $n \in \mathbb{N}$   $f_n(t, x, y, z) := f(t, x, y, h_n(z))$ ,  $(t, x, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ . This gives rise to the following family of so-called *truncated BSDE*

$$V_t^n = H(X_T) + \int_t^T g_n(s, X_s, Y_s, Z_s^{1n}, Z_s^{2n}) ds - \int_t^T Z_s^{1n} dW_s^1 - \int_t^T Z_s^{2n} dW_s^2, \quad (8.3)$$

$t \in [0, T], n \in \mathbb{N}$

For the convergence to happen, two sets of conditions on the coefficients are required to hold

**Assumption 5.** • *There exists a positive constant  $K$  such that  $\mu$ ,  $a$  and  $b$  are uniformly Lipschitz continuous with Lipschitz constant  $K$ , and  $\mu(\cdot, 0)$ ,  $a(\cdot, 0)$  and  $b(\cdot, 0)$  are bounded by  $K$*

- *There exists a constant  $M \in \mathbb{R}_+$  such that  $H$  is absolutely bounded by  $M$ ,  $f$  is measurable and continuous in  $(x, y, z^1, z^2)$  and for  $(t, x) \in [0, T] \times \mathbb{R}$ ,  $y, y' \in \mathbb{R}$  and  $(z^1, z^2), (z'^1, z'^2) \in \mathbb{R}^2$  we have*

$$|g(t, x, y, z^1, z^2)| \leq M(1 + \|(z^1, z^2)\|^2)$$

$$|g(t, x, y, z^1, z^2) - g(t, x, y, z'^1, z'^2)| \leq M(1 + \|(z^1, z^2)\| + \|(z'^1, z'^2)\|) \|(z^1, z^2) - (z'^1, z'^2)\|$$

**Assumption 6.** • *The functions  $\mu$ ,  $a$  and  $b$  are continuously differentiable with bounded derivatives in the spatial variable for all  $t \in [0, T]$*

- *$g$  is continuously partially differentiable in  $(x, y, z^1, z^2)$*
- *There exists a constant  $M \in \mathbb{R}_+$  such that for all  $(t, x, y, z^1, z^2) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$*

$$\begin{aligned} \left| \frac{\partial g}{\partial x} \right| &\leq M(1 + \|(z^1, z^2)\|^2) \\ \left| \frac{\partial g}{\partial y} \right| &\leq M(1 + \|(z^1, z^2)\|^2) \\ \left| \frac{\partial g}{\partial z^1} \right| &\leq M(1 + \|(z^1, z^2)\|) \\ \left| \frac{\partial g}{\partial z^2} \right| &\leq M(1 + \|(z^1, z^2)\|) \end{aligned}$$

- *$H$  is a continuously differentiable function satisfying  $|\nabla H| \leq M$*

It is readily checked that with the specific form of the driver in (8.2) and the choice of coefficients driving the SDEs, these conditions are satisfied. Note that there is also an ellipticity condition on the volatility matrix which trivially holds in our setting. So under these assumptions, we have the following theorem (theorem 6 of (28)):

**Theorem 13.** *Fix  $n \in \mathbb{N}$  and let  $X$  be the solution of (1). Let  $(V, Z)$  and  $(V^n, Z^n)_{n \in \mathbb{N}}$  be the solution pairs of (8.2) and (8.3) respectively. Then for all  $p \geq 2$  there exists a positive constant  $C_p$  such that for all  $n \in \mathbb{N}$*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |V_t^n - V_t|^p \right] + \mathbb{E} \left[ \left( \int_0^T |Z_s^n - Z_s|^2 ds \right)^{\frac{p}{2}} \right] \leq C_p \frac{1}{n^{1/2}}. \quad (8.4)$$

Having established the convergence of the truncated FBSDE we shall concentrate on the numerical methods available when the driver is Lipschitz. We recall that there are mainly three different methods proposed in the literature for addressing this numerical issue which actually differ in the way conditional expectations are approximated. The first one relies on quantization techniques as in Lemor (37). The second one uses Malliavin derivatives as in Bouchard and Touzi (7). And finally the most straightforward approach is based on least-square regressions. The two references for that last method are the paper of Gobet, Lemor and Warin (22) and the PhD thesis of Lemor (37). We chose to implement method based on least-square regressions technique and for

## 8. NUMERICAL RESULTS AND COMPARISONS

---

that purpose we used basis functions taken from Lemor (37): they are multidimensional polynomials, hypercubes indicators and hypercubes indicators with low degree polynomials for the theoretical value process approximation.

### 8.2.2 Solving the Nonlinear Partial Differential Equation

We consider in this section the PDE associated with the FBSDE (8.2) or derived from the martingale and orthogonality conditions on the  $f$ -costs process as in chapter 5. With the notations introduced at the beginning of the chapter, we have

$$\frac{\partial V}{\partial u} + \Lambda V = \frac{\mu}{Y} \left( \frac{\partial V}{\partial X} XY + \frac{\partial V}{\partial Y} \rho b \right) + \alpha \left( \sqrt{1 - \rho^2} \frac{\partial V}{\partial Y} b \right)^2$$

with terminal condition  $V_T = H(X_T) = (X_T - K)_+$  and

$$\Lambda V_u = \frac{\partial V}{\partial X} \mu_u X + \frac{\partial V}{\partial Y} a_u + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} X^2 Y^2 + \frac{1}{2} \frac{\partial^2 V}{\partial Y^2} b_u^2 + \frac{\partial^2 V}{\partial X \partial Y} \rho b_u XY$$

Its numerical approximation will serve as a benchmark for our FBSDE schemes, as the convergence of the former is usually much better as for the latter. In order to solve it numerically, we use a finite elements method with convection terms corresponding to the first order spatial derivatives treated by the *Characteristic Galerkin* method (see Kuzmin (34) for an introduction) so as to obtain an unconditionally stable scheme. In order to deal with the non-linear terms when  $f^{(3)}(0) \neq 0$  we use Gauss-Newton iterations.

The grid is chosen big enough so that we can use Dirichlet boundary conditions in the  $X$  direction:  $V(X_{min}) = K - X_{min}$  and  $V(X_{max}) = 0$  whereas in the  $Y$  direction we chose free boundary conditions.

The optimal hedge is then computed according to the following formula

$$\delta = \frac{\partial V}{\partial X} + \frac{\partial V}{\partial Y} \rho \frac{b}{XY}$$

by approximating the partial derivatives from the grid values.

### 8.2.3 Discrete Time Approximation

Given that the problem of local risk-minimization was firstly stated in a discrete time setting, we also compare the results of the previous sections with the risk-minimization program applied to the discretization of the SDEs. The method used works as follows

## 8.2 Benchmark Stochastic Volatility Models

---

- We compute the state variables  $X$  and  $Y$  at  $n$  hedging dates for a number  $N$  of simulation paths
- We approximate the risk function  $f$  with its *3rd*-order Taylor expansion at 0.
- We rewrite the risk function accordingly.
- We perform a least-square fit onto basis functions (the same as for the FBSDE approximation).
- We carry out the minimization on each path given the fact that the risk function is now an explicit function of the value  $V$  and the hedge  $\delta$ , as a *3rd*-order polynomial.

So we have  $R_k = \mathbb{E}_k (f (V_{k+1} - V - \delta(S_{k+1} - S_k)))$ , with  $f(x) = \frac{1}{2}x^2 + \frac{\alpha}{6}x^3$ , this gives

$$R_k = \mathbb{E}_k \left( \frac{1}{2} (V_{k+1} - V - \delta(S_{k+1} - S_k))^2 + \frac{\alpha}{6} (V_{k+1} - V - \delta(S_{k+1} - S_k))^3 \right)$$

which yields after developing and factorizing terms in powers of  $\delta$  and  $V$

$$\begin{aligned} R_k &= \frac{1}{2} \mathbb{E}_k (V_{k+1}^2) + \frac{\alpha}{6} \mathbb{E}_k (V_{k+1}^3) \\ &+ V \left( -\mathbb{E}_k (V_{k+1}) - \frac{\alpha}{2} \mathbb{E}_k (V_{k+1}^2) \right) + V^2 \left( \frac{1}{2} + \frac{\alpha}{2} \mathbb{E}_k (V_{k+1}) \right) - V^3 \frac{\alpha}{6} \\ &+ \delta \left( -\mathbb{E}_k (V_{k+1} S_{k+1}) + S_k \mathbb{E}_k (V_{k+1}) - \frac{\alpha}{2} \mathbb{E}_k (V_{k+1}^2 S_{k+1}) + \frac{\alpha}{2} S_k \mathbb{E}_k (V_{k+1}^2) \right) \\ &+ \delta^2 \left( \frac{1}{2} \mathbb{E}_k (S_{k+1}^2) - S_k \mathbb{E}_k (S_{k+1}) + \frac{1}{2} S_k^2 + \frac{\alpha}{2} \mathbb{E}_k (V_{k+1} S_{k+1}^2) \right. \\ &\quad \left. - \alpha S_k \mathbb{E}_k (V_{k+1} S_{k+1}) + \frac{\alpha}{2} S_k^2 \mathbb{E}_k (V_{k+1}) \right) \\ &+ \delta^3 \left( -\frac{\alpha}{6} \mathbb{E}_k (S_{k+1}^3) + \frac{\alpha}{2} S_k \mathbb{E}_k (S_{k+1}^2) - \frac{\alpha}{2} S_k^2 \mathbb{E}_k (S_{k+1}) + \frac{\alpha}{6} S_k^3 \right) \\ &+ \delta V \left( \mathbb{E}_k (S_{k+1}) - S_k + \alpha \mathbb{E}_k (V_{k+1} S_{k+1}) - \alpha S_k \mathbb{E}_k (V_{k+1}) \right) + \delta V^2 \left( -\frac{\alpha}{2} \mathbb{E}_k (S_{k+1}) + \frac{\alpha}{2} S_k \right) \\ &+ \delta^2 V \left( -\frac{\alpha}{2} \mathbb{E}_k (S_{k+1}^2) + \alpha S_k \mathbb{E}_k (S_{k+1}) - \frac{\alpha}{2} S_k^2 \right) \end{aligned}$$

Thus in the Monte Carlo implementation we have to compute a total of 9 conditional expectations  $\mathbb{E}_k (V_{k+1})$ ,  $\mathbb{E}_k (V_{k+1}^2)$ ,  $\mathbb{E}_k (V_{k+1}^3)$ ,  $\mathbb{E}_k (S_{k+1})$ ,  $\mathbb{E}_k (S_{k+1}^2)$ ,  $\mathbb{E}_k (S_{k+1}^3)$ ,  $\mathbb{E}_k (V_{k+1} S_{k+1})$ ,  $\mathbb{E}_k (V_{k+1} S_{k+1}^2)$  and  $\mathbb{E}_k (V_{k+1}^2 S_{k+1})$ . We do so using a linear regression algorithm which relies on a singular value decomposition routine. We then solve numerically the minimization problem using Fletcher-Reeves conjugate gradient algorithm.

### 8.2.4 Numerical Results

#### Convergence Tests

We produce convergence graphs for two sets of inputs in the Heston case. The option we consider is a 1Y put option with a strike value of 100. The first set allows to simply recover the Black and Scholes model and corresponds to  $X_0 = 100$  and  $Y_0 = 0.2$  all other parameters being set to zero. The reference values obtained by applying Black Scholes formula is 7.9656 for the portfolio initial value and  $-0.4602$  for the initial delta.

We obtained the following results by solving the FBSDE.

The first two graphs show the convergence with the number of paths, with  $n = 100$  and polynomial basis up to order 3.

Using the minimization in the Monte Carlo we got

The next series of two graphs shows the convergence with the number of time steps, with  $N = 100000$  and polynomial basis up to order 3

Results obtained by solving the FBSDE

Results obtained using the minimization in the Monte Carlo

Now we perform the exact same convergence tests with the values used in the paper of Heath, Platen and Schweizer. These are  $\Delta = 0.5$ ,  $\kappa = 5.0$ ,  $\theta = 0.04$ ,  $\Sigma = 0.5$ . We used  $\rho = -0.3$  and  $\alpha = 0$ .

We show only the results for the FBSDE in the same order as above (convergence with respect to the number of paths and then convergence with respect to the number of time steps)

Finally we give the optimal hedge  $\delta$  and theoretical portfolio value  $V$  for one year put options of different strikes for different levels of correlation

They have been obtained with  $N = 100000$  paths and  $n = 100$  and averaged over 50 runs.

### 8.3 Mean Costs

In the quadratic framework, Föllmer and Schweizer showed that the optimal strategies have zero costs on average. This is lemma (3) of chapter 3. We retrieve this property when considering  $f$  quadratic. On the other hand, for risk functions that have a non-zero third order derivative  $f^{(3)}(0) \neq 0$ , the average costs will generally not vanish. It is

then interesting to compute this value and look at some properties such as its empirical distribution. This will give a hint as to which price should the option be sold. In the quadratic case, the authors suggested to use the initial optimal hedge portfolio value since the strategy is mean self-financing. We use the FBSDE characterization of the optimal strategy to show how we may compute the average costs of an optimal strategy

$$-dV_t = g(t, X_t, Y_t, V_t, \delta_t, \sqrt{1 - \rho^2} \frac{\partial V}{\partial Y} b) - \delta Y_s dW_s^1 - \frac{\partial V}{\partial Y} b \sqrt{1 - \rho^2} dW_s^2$$

Thus we find that the costs are expressed as

$$V_T - V_t - \int_t^T \delta_s dS_s = \alpha \int_t^T \frac{\partial V^2}{\partial Y} b^2 (1 - \rho^2) + \frac{\partial V}{\partial Y} b \sqrt{1 - \rho^2} dW_s^2$$

The average costs are then

$$\mathbb{E}_t \left( V_T - V_t - \int_t^T \delta_s dS_s \right) = \alpha (1 - \rho^2) \mathbb{E}_t \left( \int_t^T \frac{\partial V^2}{\partial Y} b^2 \right) \quad (8.5)$$

From this last expression we see that the average costs for the  $f$ -risk-minimizing strategy are expressed as the expected squared costs for the local risk-minimizing strategy in the quadratic case.

So as to compute them we can either use the Monte Carlo implementation and run a forward calculation and then average the results along the paths, or we can write the PDE for  $C(t, X_t, Y_t) = \mathbb{E}_t \left( V_T - V_t - \int_t^T \delta_s dS_s \right)$

$$\frac{\partial C}{\partial t} + x\mu \frac{\partial C}{\partial x} + a \frac{\partial C}{\partial y} + \frac{1}{2} x^2 y^2 \frac{\partial^2 C}{\partial x^2} + \frac{1}{2} b^2 \frac{\partial^2 C}{\partial y^2} + xyb\rho \frac{\partial^2 C}{\partial x \partial y} + \alpha (1 - \rho^2) \frac{\partial V^2}{\partial Y} b^2 = 0$$

on  $(0, T) \times (0, \infty) \times \mathbb{R}$  with boundary condition

$$C(T, x, y) = 0$$

for  $(x, y) \in (0, \infty)$ . We solve this equation with a finite element method and since the computation of  $\frac{\partial V}{\partial Y}$  is a source term of the PDE, we do this along with the computation of  $V$ .



## 8. NUMERICAL RESULTS AND COMPARISONS

---

### Numerical Results

We finally present the numerical results obtained by solving the two PDEs (8.5) and (5.5) for the Heston model with the second set of parameters as described in the previous section, with correlation  $\rho = 0$ .

We price one year put options, in the money ( $K = 110$ ), at the money ( $K = 100$ ) and out of the money ( $K = 90$ ) and we give the values of theoretical portfolio at initial time  $(V, \delta)$  together with the mean costs  $C$ .

The parameters of the finite elements PDE solver are as follows: the grid in  $X$  variable goes from 40 to 300 and has 100 discretization points. The grid in  $Y$  variable goes from 0.0001 to 1 and has 50 discretization points. The elements are quadratic on each triangle of the mesh (P2 elements). The time discretization is uniform with 50 time steps but we still use the *characteristic Galerkin* method to have an unconditionally stable scheme. We start with model  $H1$ . We use 5 different values of  $\alpha$ : from -0.2 to 0.2 The results are summarized in the following graphs

Note that because the correlation is zero in this set of parameters the initial theoretical portfolio value adjusted with the mean costs is the same whichever function  $f$  is chosen. This is clearly seen in that last graph

# References

- [1] Frédéric Abergel and Nicolas Millot. Nonquadratic local risk-minimization for hedging contingent claims in incomplete markets. *SIAM J. Financial Math.*, 2(1):342–356, 2011. 42, 55, 89
- [2] Guy Barles, Rainer Buckdahn, and Etienne Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics An International Journal of Probability and Stochastic Processes*, 60(1):57–83, 1997. 67
- [3] Pauline Barrieu and Nicole El Karoui. Pricing, hedging and optimally designing derivatives via minimization of risk measures, 2007. 3
- [4] David S Bates. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996. 65
- [5] Fischer Black and Myron S Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973. 1, 2, 62, 83
- [6] Bruno Bouchard and Romuald Elie. Discrete-time approximation of decoupled forward-backward sde with jumps. *Stochastic Processes and their Applications*, 118(1):53–75, January 2008. 67
- [7] Bruno Bouchard and Nizar Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2):175–206, June 2004. 95
- [8] Umut Çetin, Robert Jarrow, and Philip Protter. Liquidity risk and arbitrage pricing theory. *Finance and Stochastics*, 8(3):311–341, 08 2004. 4, 56, 71

## REFERENCES

---

- [9] Ales Cerny and Jan Kallsen. Hedging by Sequential Regressions Revisited. *Social Science Research Network Working Paper Series*, March 2008. 2
- [10] Anirban Chakraborti, Ioane Muni Toke, Marco Patriarca, and Frederic Abergel. Econophysics review: I. empirical facts. *Quantitative Finance*, 11(7):991–1012, 2011. 2
- [11] Anirban Chakraborti, Ioane Muni Toke, Marco Patriarca, and Frederic Abergel. Econophysics review: II. agent-based models. *Quantitative Finance*, 11(7):1013–1041, 2011. 2
- [12] Samuel N. Cohen and Robert J. Elliott. A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications*, 120(4):442–466, 2010. 39
- [13] R. Cont and P. Tankov. *Financial Modelling with Jump Processes, Second Edition*. Chapman & Hall/Crc Financial Mathematics Series. Taylor and Francis, 2009. 2
- [14] Michael G Crandall, Hitoshi Ishii, and Pierre-Louis Lions. users guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):67, 1992. 60
- [15] M. H. A. Davis and J. M. C. Clark. A note on super-replicating strategies. *Philosophical Transactions: Physical Sciences and Engineering*, 347(1684):485–494, 1994. 3
- [16] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520, 1994. 2
- [17] Assane Diop. *Sur la convergence de certaines fonctionnelles de semimartingales discrétisées*. PhD thesis, Université Paris VI, 2009. 47
- [18] Assane Diop. Convergence of some random functionals of discretized semimartingales. *Arxiv preprint arXiv10012182*, page 31, 2010. 47
- [19] N El Karoui, S Peng, and M C Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997. 32, 62

- 
- [20] H. Föllmer and A. Schied. *Stochastic finance: an introduction in discrete time*. De Gruyter studies in mathematics. Walter de Gruyter, 2004. 25
- [21] Hans Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*, 5:389–414, 1991. 19, 31, 32
- [22] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. A regression-based monte carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005. 95
- [23] J. Michael Harrison and David M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, June 1979. 1
- [24] J. Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215–260, August 1981. 1, 55
- [25] D. Heath, E. Platen, and M. Schweizer. *Numerical comparison of local risk-minimisation and mean-variance hedging*. Research report (Australian National University. Centre for Mathematics and its Applications). Australian National University, Centre for Mathematics and its Applications, School of Mathematical Sciences, 1999. 91, 92
- [26] David Heath, Eckhard Platen, and Martin Schweizer. A comparison of two quadratic approaches to hedging in incomplete markets. *Mathematical Finance*, 11(4):385–413, 2001. 31, 92
- [27] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–43, 1993. 2, 92
- [28] Peter Imkeller, Goncalo Dos Reis, and Jianing Zhang. Results on numerics for fbsde with drivers of quadratic growth, 2010. 3, 93, 94, 95
- [29] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1987. 55

## REFERENCES

---

- [30] Jan Kallsen. A utility maximization approach to hedging in incomplete markets. *Mathematical Methods of Operations Research*, 50(2):321–338, 1999. 3
- [31] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Graduate texts in mathematics. Springer, 1991. 63
- [32] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Applications of mathematics. Springer-Verlag, 1992. 58, 80
- [33] Magdalena Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28(2):558–602, 2000. 57, 61, 63
- [34] Dmitri Kuzmin. A guide to numerical methods for transport equations, 2010. 96
- [35] Damien Lamberton, Huyen Pham, and Martin Schweizer. Local risk-minimization under transaction costs. *Math. Oper. Res.*, 23(3):585–612, March 1998. 70
- [36] Antoine Lejay, Ernesto Mordecki, and Soledad Torres. Numerical approximation of Backward Stochastic Differential Equations with Jumps. 2007. 67
- [37] Jean-Philippe Lemor. *Approximation par projections et simulations de Monte-Carlo des équations différentielles stochastiques rétrogrades*. These, Ecole Polytechnique X, June 2005. 95, 96
- [38] Benoit Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36:394, 1963. 2
- [39] Benoit Mandelbrot. The variation of some other speculative prices. *The Journal of Business*, 40:393, 1967. 2
- [40] Anis Matoussi and Hao Wang. Probabilistic interpretation for sobolev solution of semilinear parabolic partial integro- differential equations. 2009. 67
- [41] Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics*, 4(1):141–183, Spring 1973. 1, 2, 62, 83
- [42] Marie Amelie Morlais. An extended existence result for quadratic bsdes with jumps with application to the utility maximization problem, 2008. 67

- 
- [43] Marie-Amélie Morlais. A new existence result for quadratic bsdes with jumps with application to the utility maximization problem. *Stochastic Processes and their Applications*, 120(10):1966–1995, 2010. 67
- [44] P.E. Protter. *Stochastic integration and differential equations*. Applications of mathematics. Springer, 2004. 43, 45, 50, 78
- [45] Adrien Richou. Numerical simulation of bsdes with drivers of quadratic growth. 2010. 93
- [46] M. Schweizer and H. Föllmer. Hedging by Sequential Regression: an Introduction to the Mathematics of Option Trading. *ASTIN Bulletin*, 18(2):147–160, 1988. 2
- [47] Martin Schweizer. *Hedging of Options in a General Semimartingale Model*. PhD thesis, Swiss Federal Institute of Technology Zurich, 1988. 5, 19, 21, 22, 23, 24, 25, 27, 28, 30, 34, 42
- [48] Martin Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2):339–363, April 1991. 5, 19, 30, 50
- [49] Martin Schweizer. Mean variance hedging for general claims. Discussion Paper Serie B 167, University of Bonn, Germany, 1992. 3
- [50] Martin Schweizer. Local risk-minimization for multidimensional assets and payment streams. *Banach Center Publications*, 83:213–229, 2008. 30, 32
- [51] Dieter Sondermann and Hans Föllmer. Hedging of non-redundant contingent claims. Discussion paper serie b, University of Bonn, Germany, 1985. 2, 21
- [52] Elias M Stein and Jeremy C Stein. Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4(4):727–52, 1991. 92